

# Individual Sentiments and Collective Decisions: Cryptocurrencies as a Prism for the Wider Market

Ngoc-An Nguyen-Pham, B.Sc.

Supervised by Prof Martin Crane and Dr Marija Bezbradica



A thesis presented for the degree of Doctor of Philosophy

SCHOOL OF COMPUTING  
DUBLIN CITY UNIVERSITY

December 2024

# Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed (Typed Name): An Nguyen

ID No.: 20216202

Date: 18/12/2024

# Acknowledgements

*This has been one of the most wonderful journeys of my life. Four years may not seem long, but for me, it is more than enough to fill a significant part of my memories.*

This research is financially supported by the Science Foundation Ireland Centre for Research Training in Artificial Intelligence under grant number 18/CRT/6223.

First of all, I would like to express my deepest gratitude to my supervisors, Martin and Marija, for their guidance and support over the past 4 years. Throughout this journey, I have experienced various emotions: happiness, pride, disappointment, nervousness, and also moments of self-doubt, but you both have always stood by me. None of my work would have been done without your valuable suggestions, insights and feedback. I am especially grateful to Martin for everything. Thank you for always being there to help, even when you were not well. I feel very lucky to be your student and I have learned so much from you, not just in terms of knowledge but also in your working style. *I still keep the noen yellow winter jacket you bought for me on my first day in Ireland!*

I would like to thank my parents, My and Trinh, and my brother, An, for being my unwavering rock having always been my constant source of motivation. Knowing that you are always there for me, ready to support me whenever I need it, makes me feel safe and warm. *I am not young anymore but I know that you guys are always concerned about me.*

I cannot forget to thank my Vietnamese brothers living here in Ireland, especially Cuong, Phuc, and Le. My life here would be very boring if you guys were not my friends! Also, to Zing and Chau - my best friends of 8 years, it is amazing that after all this time, we are still as close as ever, *cảm ơn hai đứa này vì đã luôn chịu khó lắng nghe và tâm sự với tao qua điện thoại nhé!*

Finally, I want to dedicate this last paragraph to my beardie bear, Alan. Thank you for being such an important part of my life, for taking care of me and for being my excellent listener. I used to be very timid, but since you came into my life, I have become braver and more confident. *Bhuel, go hionraic, tá do bhia beagán gan blas, ach tá sé sin ceart go leor - is cócaire maith mé!*

By the way, I wish I had the magical pocket of Doraemon... :)

# List of Publications

**1) Cryptocurrency Volatility Index: an Efficient Way to Predict the Future CVI**

An Nguyen, Martin Crane, Marija Bezbradica

30th Irish Conference on Artificial Intelligence and Cognitive Science

**2) The Cryptocurrency Market in Transition before and after Covid-19: An Opportunity for Investors?**

An Nguyen, Tai Tan Mai, Marija Bezbradica, Martin Crane

Journal Entropy

**3) Volatility and Returns Connectedness in Cryptocurrency Markets: Insights from Graph-based Methods**

An Nguyen, Tai Tan Mai, Marija Bezbradica, Martin Crane

Journal Physica A: Statistical Mechanics and its Applications

**4) Herding Unmasked: Insights into Cryptocurrencies, Stocks and US ETFs**

An Nguyen, Martin Crane, Thomas Conlon, Marija Bezbradica

Journal PLOS One

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Context of Study . . . . .	3
1.2	Research Objectives . . . . .	7
1.3	Research Contributions . . . . .	9
<b>2</b>	<b>Literature Review</b>	<b>11</b>
2.1	Introduction . . . . .	11
2.2	Modern Financial Markets in Uncertain Times . . . . .	12
2.3	An Overview of Correlations and Graph-based Methods in Financial Markets . . . . .	16
2.3.1	Investigating the Correlations in Financial Assets . . . . .	16
2.3.2	Graph-based Methods and its Applications . . . . .	19
2.4	Herding Behavior Discovery . . . . .	21
2.4.1	Herding Behavior in Traditional Markets . . . . .	21
2.4.2	Herding Behavior in the Cryptocurrency Market . . . . .	25
<b>3</b>	<b>Dataset Description</b>	<b>28</b>
3.1	Introduction . . . . .	28
3.1.1	First Dataset . . . . .	29
3.1.2	Second Dataset . . . . .	30
3.2	A Note on the Granularity of Examined Time Series . . . . .	33
3.3	Statistical Property Comparison . . . . .	35
3.4	Remarks . . . . .	39

---

<b>4</b>	<b>Methodology</b>	<b>40</b>
4.1	Introduction . . . . .	40
4.2	Return and Volatility Calculations from Price Time Series . . . . .	41
4.3	Metrics for Analysis and Evaluation . . . . .	43
4.3.1	Correlation Metrics . . . . .	43
4.3.2	Graph-based Metrics . . . . .	45
4.3.3	Symbolic Metrics . . . . .	49
4.4	Correlation Matrix Formation and Random Matrix Theory . . . . .	50
4.5	Notion of Graphs under Financial Markets Context . . . . .	51
4.6	Noise and Trend in the Financial Markets: A Removal Scheme . . . . .	52
4.6.1	Introduction of Noise and Trend . . . . .	52
4.6.2	Denoising in Finance . . . . .	54
4.6.3	Detrending in Finance . . . . .	57
4.7	Louvain Community Detection . . . . .	60
4.8	Herding detection . . . . .	63
4.9	Remarks . . . . .	64
<b>5</b>	<b>Experimental Results and Discussion</b>	<b>65</b>
5.1	A Note for this Chapter . . . . .	65
5.2	Price Returns-Based Co-Movement in the Cryptocurrency Market . . . . .	66
5.2.1	Introduction . . . . .	66
5.2.2	Timeline Division . . . . .	67
5.2.3	Cryptocurrency Graph Structures under Noise and Trend Effects . . . . .	69
5.2.4	The Evolution of graph Structures with Levels of Granularity . . . . .	73
5.2.5	Normal Times Versus Turbulent Times: A Graph Structure Perspective . . . . .	77
5.2.6	Understanding Investment Decisions of Cryptocurrency Traders Based on Ranking Distribution . . . . .	80
5.2.7	Remarks . . . . .	85
5.3	Volatility-Based Co-Movement in the Cryptocurrency Market . . . . .	87

---

5.3.1	Introduction . . . . .	87
5.3.2	Choosing an Optimal Time Window Length for graph Construction . . . . .	88
5.3.3	Volatility vs Returns: Unveiling the Differences in Cryptocurrency Graph Structures . . . . .	90
5.3.4	Can Changes in Volatility-based Correlation graph be Explained by Investors' Sentiments? . . . . .	98
5.3.5	Remarks . . . . .	100
5.4	Measuring the Influence of Individual Cryptocurrencies on the Cryptocurrency Market . . . . .	102
5.4.1	Introduction . . . . .	102
5.4.2	Determining Which Cryptocurrencies are Dominant . . . . .	103
5.4.3	Remarks . . . . .	111
5.5	Herding Behavior among Different Financial Assets . . . . .	112
5.5.1	Introduction . . . . .	112
5.5.2	Timeline Division . . . . .	113
5.5.3	Herding Behavior in the Cryptocurrency Market . . . . .	115
5.5.4	An Extension to Traditional Markets . . . . .	118
5.5.5	Herding Behavior Across Cryptocurrencies, Stocks and US ETFs: A Look at Community Structures . . . . .	122
5.5.6	Herding Behavior Across Cryptocurrencies, Stocks and US ETFs: Unveiling Results . . . . .	125
5.5.7	Remarks . . . . .	133
<b>6</b>	<b>Conclusion</b>	<b>136</b>
6.1	Research Questions and Key Findings . . . . .	136
6.2	Limitations . . . . .	139
6.3	Future Works . . . . .	140
	<b>Bibliography</b>	<b>143</b>

**A Details of Assets in each Community during each Sub-period 175**

# List of Figures

3.1	Distribution moments of all assets across three types of investment vehicle from the second dataset . . . . .	37
3.2	All assets (e.g. stocks, US ETFs and cryptocurrencies) in the second dataset visualized by pairs of statistical indicators	38
4.1	Density distribution of eigenvalues (blue) obtained from our empirical correlation matrix and the corresponding . . . . .	56
4.2	Market component from our correlation matrix. . . . .	58
4.3	An example of cleaning noise and trend effects in a correlation matrix . . . . .	59
4.4	Louvain community detection. a) MST (Original graph), b) Community detection applied to the MST. . . . .	62
5.1	The reaction of the general public and global economy to the Covid-19 pandemic . . . . .	69
5.2	Cryptocurrency graph structures before and after denoising and detrending, using daily data . . . . .	72
5.3	Graph structure for the first time window, community detection is applied using Louvain method . . . . .	74
5.4	Graph structure for the second time window, community detection is applied using Louvain method . . . . .	75
5.5	Graph structure for the third time window, community detection is applied using Louvain method . . . . .	76

5.6	The ranking distributions of cryptocurrencies. Each community is represented by a circular shape . . . . .	83
5.7	The similarity of each pair of graph structures using different window sizes, observed by <i>betweenness centrality</i> . . . . .	89
5.8	The similarity between two graph structures derived from 2 corresponding volatility formulas, observed by <i>betweenness centrality</i> . . . . .	90
5.9	Changes in graph structures every 30 minutes before denoising and detrending . . . . .	91
5.10	Changes in graph structures every 30 minutes after denoising and detrending . . . . .	93
5.11	The movements of <i>betweenness centrality</i> and <i>Fear and Greed</i> index from 09/03/2019 until 06/04/2021 . . . . .	99
5.12	The influence of each cryptocurrency on the correlation graph over time (before denoising and detrending) . . . . .	104
5.13	The influence of each cryptocurrency on the correlation graph over time (denoised and detrended) . . . . .	104
5.14	The correlation between 34 cryptocurrencies in terms of the number of transactions from 13/02/2019 until 06/04/2021 (on a 30-minute basis) . . . . .	108
5.15	Time-varying community structures of cryptocurrencies, stocks and US ETFs during: a) Pre-Covid-19, b) Covid-19 pandemic, c) Post-Covid-19. . . . .	124
5.16	Time-varying community structures of cryptocurrencies, stocks and US ETFs during: a) Bull market period, b) Ukraine-Russia conflict. . . . .	125
5.17	Herding detection results (a) and sector distribution (b) of each community during the Pre-Covid-19 period. . . . .	127

5.18 Herding detection results (a) and sector distribution (b) of each community during the Covid-19 pandemic period. . . . 128

5.19 Herding detection results (a) and sector distribution (b) of each community during the Post-Covid-19 period. . . . . 128

5.20 Herding detection results (a) and sector distribution (b) of each community during the Bull market period. . . . . 129

5.21 Herding detection results (a) and sector distribution (b) of each community during the Ukraine-Russia conflict period. . 129

# List of Tables

3.1	A list of 34 cryptocurrencies used in our study. Abbreviations are put in parentheses. . . . .	29
3.2	Average rankings of cryptocurrencies during the period between February 2019 and April 2021 . . . . .	30
3.3	Three types of investment vehicle in the second dataset and their corresponding assets . . . . .	32
3.4	Distribution of Business Sectors in the Stock Dataset . . . . .	34
3.5	The average distance between stocks and US ETFs, stocks and cryptocurrencies, and US ETFs and cryptocurrencies, evaluated for each statistical indicator . . . . .	38
4.1	Testing the market efficiency for each cryptocurrency using 48 data points, over the period from February 2019 to April 2021. . . . .	43
4.2	Correlation Metrics Used in this Study . . . . .	44
4.3	Graph-based Metrics Used in this Study . . . . .	46
4.4	Graph-based Metrics Used in this Study (Continue) . . . . .	47
4.5	Metrics for Graph Comparison Used in this Study . . . . .	48
4.6	PageRank-based Metrics for Node Assessment . . . . .	49
5.1	Timeline partition for the HitBTC cryptocurrency dataset (time windows split to take into consideration the Covid-19 pandemic). . . . .	69

5.2	Connection strength measured by <i>Residuality Coefficient</i> and <i>Mean Weight</i> . . . . .	71
5.3	The growth of graph structures over time for each level of granularity, measured by <i>Betweenness Centrality</i> and <i>Degree Assortativity</i> . . . . .	77
5.4	Number of communities at each level of granularity across three time windows . . . . .	77
5.5	The growth of graph structures over time measured by <i>Betweenness Centrality</i> and <i>Degree Assortativity</i> . . . . .	78
5.6	Similarity in graph structures between different time windows	79
5.7	Distribution of rankings and the percentage of high-ranking cryptocurrencies in each community . . . . .	82
5.8	Average distance between a data point and the mean of a graph-based metrics (using returns) . . . . .	94
5.9	The percentage of 3 SAX segments in each period . . . . .	95
5.10	Average distance between a data point and the mean of a graph-based metric for the entire period, the spike in March 2020 and the spike in February 2021 (using volatility) . . . . .	96
5.11	Total Accumulation (TA) for each cryptocurrency with the effect of noise and trend . . . . .	105
5.12	Total Accumulation (TA) for each cryptocurrency after removing noise and trend . . . . .	105
5.13	Dominance Score (DS) for each cryptocurrency with the effect of noise and trend . . . . .	106
5.14	Dominance Score (DS) for each cryptocurrency after removing noise and trend . . . . .	106
5.15	Five sub-periods declaration. . . . .	115
5.16	Herding detection results for cryptocurrencies . . . . .	116
5.17	Herding detection results for stocks and US ETFs . . . . .	119

5.18 Average path length (number of edges) between two assets in six cases, namely between stock and stock, between US ETF and US ETF, between cryptocurrency and cryptocurrency, between stock and US ETF, between stock and cryptocurrency, and between US ETF and cryptocurrency. . . . . 125

A.1 Assets in each community during the Pre-Covid-19. The red color indicates cryptocurrencies, the blue color indicates US ETFs and the black color indicates stocks . . . . . 176

A.2 Assets in each community during the Covid-19 pandemic. The red color indicates cryptocurrencies, the blue color indicates US ETFs and the black color indicates stocks . . . . . 177

A.3 Assets in each community during the Post-Covid-19. The red color indicates cryptocurrencies, the blue color indicates US ETFs and the black color indicates stocks . . . . . 178

A.4 Assets in each community during the Bull market time. The red color indicates cryptocurrencies, the blue color indicates US ETFs and the black color indicates stocks . . . . . 179

A.5 Assets in each community during the Ukraine-Russia Conflict. The red color indicates cryptocurrencies, the blue color indicates US ETFs and the black color indicates stocks . . . . . 180

# List of Abbreviations

MST	Minimum Spanning Tree
PCA	Principal Component Analysis
VIX	Volatility Index
S&P500	Standard and Poor's 500
CSAD	Cross-Sectional Absolute Deviation
PR	PageRank
SAX	Symbolic Aggregate Approximation
RMT	Random Matrix Theory
TA	Total Accumulation
DS	Dominance Score
BTC	Bitcoin
ETH	Ethereum
BAT	Basic Attention Token
MKR	Maker
FUN	FunToken
DAI	Dai

# Abstract

## Individual Sentiments and Collective Decisions: Cryptocurrencies as a Prism for the Wider Market

**Ngoc-An Nguyen-Pham**

Cryptocurrencies have become a vital asset class with rising investment across demographics. However, their high risk is a well-known stylized fact. Therefore, understanding the interconnected movements between cryptocurrencies and their connections with other asset classes is crucial for reducing portfolio risk, especially during market crashes. Building on this context, we first examine how cryptocurrency correlations change over time, especially between stable and turbulent periods. Next, we explore how investor sentiments (e.g. fear, neutral and greed) and investment behaviors (e.g. herding and contrarian) influence these correlations. Lastly, we extend the study to traditional assets, including stocks and US ETFs, to compare correlations and investment behaviors between cryptocurrency and traditional markets across various market conditions.

Using graph-based methods such as Minimum Spanning Tree, Louvain and Girvan-Newman algorithms, we discover consistent changes in cryptocurrency correlations over time. However, within turbulent and stable periods, the correlations exhibit distinct patterns. Specifically, they increase during turbulent times like market crashes, global economic crises, political turmoil and bullish markets. Conversely, they remain low during stable times. Understanding these patterns helps investors make strategic investment decisions and optimize their portfolios. Furthermore, by adopting a popular herding detection model Cross-sectional Absolute Deviation, we find that the correlation in the cryptocurrency market is strongly influenced by investor sentiments and investment behaviors. In other words, the reactions of investors towards the market signal its upcoming correlation pattern as well as possible movements. Therefore, we suggest that investor sentiment is a good indicator that helps investors anticipate future changes and movements in the market, preparing

them to adjust their portfolios if necessary. Notably, we discover that although the cryptocurrency market is gradually becoming similar to traditional markets, there remains a significant gap, not only in the correlation between them but also in statistical characteristics and in how investors engage with them. The distinction of cryptocurrencies from traditional markets suggests a great potential for them to be used as hedges in a portfolio.

In addition to new findings, we introduce two novel techniques. First, a noise and trend removal scheme is applied to the assets' correlations using the Random Matrix Theory and Market Component concept, which has never been considered in the quantitative finance literature. Secondly, this is the first study observing herding behavior at a community level, using the Louvain algorithm.

# Chapter 1

## Introduction

### 1.1 Context of Study

The cryptocurrency market has evolved rapidly since the introduction of Bitcoin in 2009, attracting a vast community of investors. Out of them, well-known names worth mentioning include Barry Silbert, Michael Saylor and Elon Musk [1]. The popularity of these digital assets is validated by impressive numbers. For instance, 70,000 USD is the value of one Bitcoin in April 2024 [2]. Also this year, more than 2.4 million cryptocurrencies are in circulation, accounting for over 2 trillion USD in total [2]. Notably, there has been an increase from 5 million in 2016 to 560 million cryptocurrency users worldwide to date [3].

Unlike traditional asset classes such as stocks, bonds, or fiat currencies, cryptocurrencies are decentralized, meaning that they are not regulated by any intermediaries such as government and banks [4]. Instead, investors who participate in trading activities in the cryptocurrency market are the main factor driving its worth [5]. Consequently, this emerging market is famous for its large volatility in prices, which radically exceeds the fluctuations seen in traditional markets [6]. This unique characteristic, while offering opportunities for high profits, also exposes investors to substantial risk.

Given the opportunities and challenges in the cryptocurrency market, the primary goal of this PhD study is to help investors better understand the characteristics

of cryptocurrencies to build up more effective investment strategies. To achieve this, we focus on examining the time-varying correlation between cryptocurrencies and how these changes relate to the investment behavior of investors (for instance, the herding phenomenon among investors towards a market consensus under different market conditions) over time. Specifically, our hypothesis is that correlations among cryptocurrencies are strongly influenced by how investors behave. By analyzing and extracting characteristics in these correlations, we can gain insights into how investors make decisions and react to different market conditions. We expect that investor behavior will shift across different market phases, whether stable, bullish, bearish, or during a crisis. It is our hope that an understanding of correlation patterns between different assets in a market along with the tendencies of investors in making decisions across various scenarios, especially during turbulent times, can assist in creating more suitable asset allocation strategies, depending on a particular market circumstance. This knowledge, we believe, can also help in avoiding risks such as economic bubbles and market crashes, caused by investors themselves.

In financial markets, the correlation between two assets refers to their similarity in terms of a financial indicator's movements over a period of time, such as price, return, and volatility. When two assets are completely similar, any upward or downward movement in one asset is mirrored simultaneously in the other [7]. Several metrics have been introduced to assess the correlation between cryptocurrencies themselves, and between cryptocurrencies and traditional assets (e.g. stocks, bonds, oil and fiat currencies). One of the first studies on this topic was conducted by Yermack [8], who tested the daily dependency between Bitcoin and national currencies (e.g., EUR, JPY, GBP) using *Pearson* correlation measurement. Since then, a variety of statistics-based correlation measures such as Wavelet Coherence, Vector Autoregressive (VAR), Granger causality and q-Dependent Detrended Correlation have been widely adopted on larger datasets, comprising different cryptocurrencies and asset classes. This diversification paints a more comprehensive picture of the dependencies present in the financial markets [9–12]. Notably, with the fast growth

of Deep Learning, several methods such as Multilayer Perceptron (MLP) and Long Short Term Memory (LSTM) have been successfully applied to cryptocurrency-related correlation studies [13, 14].

In this PhD study, we adopt the Pearson method, while acknowledging that using the Pearson correlation measurement has several limitations, such as being sensitive to outliers and not being able to capture non-linear relationships [15]. However, we have chosen to use it in our study for the following reasons:

- Pearson correlation is a widely accepted method in the literature for both cryptocurrency [16–18] and traditional asset markets [19–21]. This strongly reinforces our belief in the applicability of this method of correlation calculation for our problem.
- This correlation measurement is suitable for time series with repeated observations, as is the case in financial time series, unlike other methods that require independent observations [22].
- Existing literature has proposed various methods, such as Spearman [23] and Kendall [24], to address the non-linearity problem. However, these methods convert rational numbers into integer rankings, which can lead to a loss of important information from financial time series [22]. Since our analysis relies heavily on price time series, preserving rational numbers is crucial. Therefore, using these non-linear metrics may present challenges for our research.

With this way of measuring, the correlation between two assets ranges from -1 to 1, where -1 represents a complete difference between two assets (i.e. negative linear correlation), 0 indicates no linear correlation, and 1 reveals a complete similarity (i.e. positive linear correlation). In other words, the higher the Pearson correlation value, the more similar the two assets are.

Thanks to graph-based methods, correlations between multiple assets can be explored and observed visually [25]. Its core idea is to construct a graph where the distance between two objects in the graph reflects their correlation, i.e. the closer

they are, the stronger they are correlated, i.e. the more similar they are. In the financial context, a graph comprises nodes and edges, where each node represents an asset and an edge between two nodes depicts the relationship between them, which tends to be their similarity in terms of a financial indicator's movements (i.e. price, return or volatility). Various graph construction methods have been proposed in the literature, including Minimum Spanning Tree (MST) [26], k-Nearest Neighbors (kNN) [19], Planar Maximally Filtered Graph (PMFG) [21], and Threshold Weighted-Minimum Dominating Set (TW-MDS) [27], to name but a few. The popularity of graph-based methods can be attributed to several factors: 1) their ability to observe the time-varying connectedness among multiple objects; 2) the diversity of metrics available in Graph Theory to analyze graph structures, such as betweenness centrality, degree assortativity, and closeness centrality. These metrics can be used to learn and understand the underlying characteristics of a graph; 3) their ease of implementation in comparison with statistical and Neural Networks methods. To this end, we adopt the Minimum Spanning Tree (MST) approach for graph construction in this study. Furthermore, from a graph, community detection algorithms can be applied to extract deeper potential hidden information from the graph, which is an important part of our experiments.

Along with the evolution of financial markets, various events tend to occur bringing either negative or positive effects, i.e. disturbing and devaluing financial assets during a market crash while propelling their worth during a market boom. In recent years, several events have significantly impacted these markets, such as the US-China trade war 2018-2019 [28], the Covid-19 pandemic followed by the global economic crisis in 2020 [29], and the Ukraine-Russia conflict in 2022 [30]. By focusing on these events, our research aims to reveal how the graph structures of cryptocurrencies evolve under these events and how investor behavior adapts in response to these shocks. Additionally, we aim to compare these event-driven periods with stable times to highlight differences in such correlations as well as in investor behavior.

Notably, we extend our study to traditional assets, including stocks and US ETFs, to facilitate a comparison between the cryptocurrency market and traditional markets. This allows us to identify both common patterns and divergences between the two. Moreover, this extension seeks to address an intriguing question: Is the cryptocurrency market becoming more similar to traditional markets, or does it continue to evolve uniquely?

The remainder of the PhD thesis is structured as follows: Chapter 2 covers an overview of the relevant studies. Chapter 3 describes two datasets used in our experiments. Chapter 4 presents terminologies, methodologies and pre-processing procedures. Chapter 5 discusses experimental results followed by their hypotheses and implications. Last but not least, a conclusion along with limitations from our study and potential future works are provided in Chapter 6.

## 1.2 Research Objectives

As stated previously, we center our focus on the time-varying correlations in the cryptocurrency market and the relationship between these correlations with the investment behavior of investors under different market conditions. From this, we structure this research into 4 key research questions.

Given the extensive use of return time series in the existing financial market literature, across diversified topics such as forecasting, portfolio optimization, algorithmic trading [31] and also correlation analysis, we determine this financial indicator as the key parameter in our experiments. Starting from this, our first objective is to discover the correlations between cryptocurrencies over time using return time series. In particular, our first research question includes:

- RQ1-1: How does the structure of the return-based correlation graph evolve over time? In particular, does it differ between stable and turbulent periods?
- RQ1-2: If so, could these changes be linked to shifts in investors' strategies? In other words, does the investment behavior of investors influence the corre-

lations between cryptocurrencies?

We note that this research question and its corresponding experiments form the foundation of our long-term study, helping to establish key parameters and directions for subsequent research questions and experiments, such as the appropriate level of granularity for time series to construct correlation graphs and the necessity to apply denoise and detrend techniques to the correlations before delving into thorough analysis. These will be discussed in later chapters.

Building on the returns-based correlation graphs, we notice the limited use of volatility time series in not only cryptocurrencies-related studies but also in other asset classes. In [32], the authors confirm that return information alone cannot fully capture all phenomena in a market. Volatility, on the other hand, has the potential to offer new insights that enrich our understanding of the market. With this in mind, we seek to explore the time-varying correlation graphs constructed from the volatility information of cryptocurrencies, which is our second research question:

- RQ2-1: Do volatility-based graph structures exhibit different behaviors from returns-based graph structures, across different market conditions?
- RQ2-2: If so, what additional insights can we extract from these differences?

The findings from the first two research questions might suggest a link between the changes in correlation graph structures and the collective behavior of investors - i.e. their tendency to herd, a phenomenon where investors follow the majority's actions or the so-called market consensus. In other words, this behavior could be a key factor driving the shifts in graph structures. This leads us to the next research question:

- RQ3-1: Can we detect herding behavior in the cryptocurrency market during boom/bust times?
- RQ3-2: If so, is it responsible for changes in the graph structures?

Lastly, we expand the third research question into traditional assets, including stocks and US ETFs, which contributes to the literature in two ways. Firstly, to compare the differences and similarities between cryptocurrencies and traditional assets. Secondly, to provide a more comprehensive understanding for those diversifying their portfolio by investing in different types of investment vehicles.

- RQ4-1: What are the similarities and differences between herding behavior in cryptocurrencies and traditional assets (i.e. stocks and US ETFs) in times of market turbulence?
- RQ4-2: How does herding behavior manifest in a diversified market that comprises both cryptocurrencies and traditional assets?

### 1.3 Research Contributions

We divide our contributions into two categories, namely data-driven contributions and technical contributions.

Regarding data-driven contributions, firstly, we use high-frequency data of 30 minutes granularity for all time series examined in this PhD study. Although a few papers work on higher frequency data, such as 1 minute and 5 minutes [33], they only account for a small portion of the existing literature. The common practice is using daily or monthly data, potentially causing a loss of important information on different types of investors and their behavior from the time series. This appears to be a major concern in the cryptocurrency market, since it is well-known for its high price fluctuations, even within a day [34–36]. Thus, using a low frequency means that we are ignoring valuable information (e.g., the intraday fluctuations of a time series) on purpose. As a result, this can adversely affect the correlation extracted from the dataset, potentially leading to inaccurate correlation-using experiments. Secondly, we use a more recent dataset covering several recent events, namely the US-China trade war in 2019, the Covid-19 pandemic followed by the global economic crisis in 2020, the Bull market period in 2021-2022 and the Ukraine-Russia conflict started

in 2022. Therefore, even though the topics of correlations in financial markets as well as investor behavior are not new, experimental results from recent years provide fresh insights into how these dynamics have evolved, especially in response to recent significant global events. Thirdly, our study stands out from the existing literature by utilizing volatility time series, whereas most previous research has primarily focused on returns. This usage is expected to discover characteristics that are hidden in returns-based results. Lastly, to our knowledge, this is the first study examining herding behavior across three different types of investment vehicle (e.g. cryptocurrencies, stocks and US ETFs). Existing studies tend to focus on one type of investment vehicle only.

When it comes to technical contributions, one of our novel techniques contributing to the existing literature is noise and trend removal in the correlation graphs, which will be introduced in Section 4.6. In brief, we remove the noise and trend (if any) from the correlations among the assets before delving into further analysis. These effects, while existing in the financial markets [6, 37, 38], are not widely taken into consideration by previous studies. To this end, the removal of noise and trend effects opens an opportunity to observe and explore underlying interactions between different assets. Hence, we can obtain novel findings that are hidden behind the noise and trend. Secondly, we integrate a community detection algorithm into herding behavior experiments. Consequently, the herding behavior is detected at a small scale level, instead of the entire market as in existing studies, which eventually provides a more detailed result.

Overall, this study provides readers with an understanding of how correlations in the cryptocurrency market change, especially during major events impacting the financial markets. Moreover, the investment behavior of investors behind a particular graph structure is also explored. Last but not least, readers also gain knowledge about the differences and similarities between herding behavior in the cryptocurrency market and traditional markets. With these insights, investors can establish more effective investment strategies in specific market conditions.

# Chapter 2

## Literature Review

### 2.1 Introduction

Our research takes the correlation between financial assets and herding behavior among investors as its central focus, especially during uncertain times when the financial markets experience a crash or a boom. Therefore, this chapter discusses the existing studies relevant to these topics to prepare readers with enough context and essential knowledge for the following chapters. We first introduce several major events over the 4-year period from 2019 to 2023 that have had a great influence on the global economy and financial markets, both cryptocurrencies and traditional ones. These events will also be of interest to our study and will be further explored through our experiments. The next section covers existing studies working on correlation-related topics, where we introduce various approaches adopted in the literature to calculate and analyze the correlation in financial markets, followed by common findings revealed by previous researchers in this domain. Likewise, the last section presents an up-to-date literature on herding behavior in the cryptocurrency and traditional markets.

## 2.2 Modern Financial Markets in Uncertain Times

Over the four years from 2019 to 2023, global financial markets have experienced several major events, including but not limited to the US-China trade war [28], the Covid-19 pandemic [29], and the more recent Russia-Ukraine conflict [30]. Although these events stemmed from different causes, they all have had a negative impact on financial markets, influencing both traditional and cryptocurrency assets to varying extents. Additionally, this period also covered other less well-known events like the oil price reduction caused by the shale revolution in the US [39], the price war between Saudi Arabia and Russia [40], and the layoff wave in tech companies [41], but their impact was mainly on a particular industry, rather than at the worldwide level.

### US-China Trade War

The US-China trade war took place in January 2018, following the imposition of initial tariffs on Chinese imports by the new US administration at that time [28]. The primary goals behind this decision were to reduce the US trade deficits at the time, stop the requirements of American companies to transfer technology and protect US technology companies from intellectual property theft [42]. As a countermeasure, China also levied different tariffs on US products [28] as well as imposing legislation preventing their citizens from using technology products made in the US [43, 44]. Although the trade war started early in 2018, it escalated during 2019 [45]. Consequently, financial markets in both the US and China were affected, especially in the commodity, energy and technology sectors [45–47]. Major stocks in both countries as well as the ETFs that track them fell sharply [48, 49]. Trade war headlines attracted huge attention throughout the year, and announcements relating to the tariff imposition caused the stock and index values to fluctuate [50]. Indeed, it has been shown that there was a wave of selling pressure in 2019 because the stock prices fluctuated due to the panic of investors caused by the trade war [51]. In China, the economy also exhibited a decline in GDP growth rate in 2019 [42], accompanied by

decreased returns across various sectors in financial markets [52]. The uncertainty in financial markets due to this trade war was also seen in other countries across the world [47, 53]. The event ended in January 2020 after both countries signed off on the so-called *Phase One agreement* trade deal [54].

## **Covid-19 Pandemic and Global Economic Crisis**

Right after the trade war, the world encountered another challenge in the form of the Covid-19 pandemic [29]. This event disrupted economies across the world. In particular, due to the fast spread speed and serious consequence of this health crisis, different regulations were applied to stop the virus, such as closing national borders, constraining mobility as well as stopping business activities worldwide [55, 56]. Subsequently, these measures profoundly influenced the economies of numerous countries, eventually leading to a global economic downturn in March 2020 as a response to various governments' efforts to control the disease spreading [57]. This is considered one of the biggest financial crises ever, even worse than previous ones such as the Black Monday 1987 crash, the 2007-2008 Global Financial Crisis [58] and others. Various studies highlighted that this pandemic disturbed and devalued financial markets, from traditional assets such as stocks and bonds [59–61] to safe-haven assets such as gold [62, 63] to cryptocurrencies [16, 59, 64]. The pandemic's strongest effects were observed during its first wave (from March until June 2020), with most assets experiencing a big drop in prices [65]. Fortunately, although the world continued facing different Covid-19 waves, its impact on financial markets lessened significantly, with their prices gradually recovering to pre-pandemic levels [66].

## **The Russia-Ukraine Conflict**

As the impacts of Covid-19 were subsiding through intervention measures such as vaccines and social distancing, another major event disrupted world economies. On 24 February 2022, Russia officially declared war on Ukraine. Numerous studies

have dug into the impact of this event on the financial markets [67–72]. Russia, a significant player in oil and gas production, has long been a crucial fuel supplier to many countries, particularly in the EU. Furthermore, Ukraine has served as a critical transit route for Russian petroleum exports to Europe. Consequently, this conflict had a severe effect on the energy sector [73]. In particular, when tensions rose and military actions escalated, oil prices tended to increase due to fears of potential supply disruptions. Conversely, when diplomatic efforts appeared to make progress or when there was a decrease in hostilities, oil prices tended to ease [74, 75]. Moreover, increased sanctions on Russia and efforts by countries, notably in Europe, to reduce dependence on Russian energy further triggered oil and gas shortages, the energy prices in Europe thereby surged during the first half of 2022 until alternative suppliers were found later on [76, 77]. Additionally, the EU’s embargo on seaborne crude oil and the G7’s price cap mechanism, which both took effect on 05/12/2022 also acted as a stimulus for the global oil prices already high for seasonal reasons [76].

Regarding other sectors, most studies have focused on the EU financial markets, since they were directly linked to Russia and Ukraine in commercial and political activities. Generally, it has been shown that the war had heterogeneous impacts on the EU financial markets. In particular, while most countries revealed a drop in their stock market prices after the war took place [70, 71, 78], those closest to Russia were found to have been impacted greater [79]. Interestingly, stocks and indices in several countries experienced an upward trend during the war, such as Belgium, Denmark, Luxembourg, Netherlands and Spain. This contradiction is attributed to both the commercial and political activities of EU countries with Russia. In particular, countries with higher exports to Russia experienced a negative impact due to the economic sanctions and the fear among investors of reduced exports. In contrast, countries that experienced a positive impact had lower commercial ties with Russia, and more importantly, their investors’ sentiments were positive because NATO would not engage in armed conflict over this matter [70]. On the other hand, this war seemed not to have distorted the financial markets in Asian countries and

other developed ones such as the US [30, 72, 80]. Nevertheless, studies [71, 80, 81] emphasized that, aside from the energy sector, the effect of this event on financial markets globally is negligible compared to previous events such as the Covid-19 pandemic, the trade war and past financial crises.

## **Other Pressing Events**

Apart from the aforementioned major events, several other events have also occurred that were mainly influential within a particular industry. For instance, the shale revolution in the US led to an increase in oil and natural gas production but also caused an oversupply issue because the global consumption of oil and gas decreased [82] and green energy sources began to emerge [83]. The world oil market thereby fluctuated with downward trends at some points in 2019 [39]. Additionally, although the energy market generally earned positive returns in the year 2019 compared to the previous year, this sector placed dead last in the Standard & Poor's 500 index in 2019 [84]. After this year, another oil-related event occurred was the price war between Russia and Saudi Arabia, as a consequence of the Covid-19 pandemic [40]. Specifically, the Organization of the Petroleum Exporting Countries (OPEC) suggested reductions in oil production amidst the COVID-19 pandemic due to the drop in consumption demand. While Saudi Arabia agreed with this proposal in early March 2020 to stabilize the oil markets, Russia withdrew from the agreement and continued increasing production as well as supply. This led to a further drop by 10% in oil prices which had already been shrunk dramatically since the onset of the pandemic. The price of oil further dropped later on after Saudi Arabia announced a decline in their oil prices as a consequence of the agreement breakdown [40].

During the year 2022, tech companies laid off their staff, which was caused by several factors, i.e. high inflation rate, high interest rate, supply chain crisis and geopolitical turmoil. Moreover, concerns about a financial recession in 2022/2023 led to preventative measures such as freezing recruitments as early as 2022 [85]. Additionally, people were bombarded with bad news about layoff announcements

and downturns in many companies, especially the giant tech ones such as Amazon, Alibaba, Microsoft, Tesla and Meta [41]. Findings reveal that layoff announcements significantly and negatively impact firm stock returns in both the short and medium term [41, 86].

## **2.3 An Overview of Correlations and Graph-based Methods in Financial Markets**

### **2.3.1 Investigating the Correlations in Financial Assets**

The correlation between financial assets has been a widely researched topic over a considerable period of time, particularly within traditional markets like stocks, bonds, and fiat currencies [7, 8, 19, 87]. This research has been beneficial to market participants in helping them understand the underlying interactions among various assets, which offer several advantages. One key benefit of understanding these correlations is to empower investors to construct well-diversified portfolios. Specifically, by allocating assets with low correlations, investors can potentially reduce portfolio volatility and improve risk-adjusted returns [88, 89]. Moreover, knowledge of correlations helps identify opportunities for hedging against losses. In this regard, investors can use assets that have a negative or low correlation with existing assets in their portfolio as hedges to ease the losses, especially during turbulent times such as market crashes [90, 91]. Apart from that, expanding correlation analysis to microeconomic variables such as interest rate, inflation rate and GDP growth rate can potentially inform investors about future movements of their investments [92].

One of the best-known methods for measuring the correlation is the Pearson correlation coefficient, which is adopted widely in the literature. For example, the authors in [7] calculated the monthly correlations between various market indices from different countries across all continents between 1980 and 2010, using the Pearson method. This 30-year period covers several major events negatively impacting

the global economy and financial markets like Black Monday in 1987, Dot-com Bubble in 2001 and Subprime Mortgage Crisis in 2008. Their findings indicate that the markets in different countries tend to move in sync during such market crashes, as reflected by a surge in their correlations. Similarly, Nie in [19] applied the same method to analyze the time-varying correlation between 83 stock assets from the S&P 100 index over a 16-year period from January 2005 to July 2020. In addition, to assess the correlations more accurately, they used the so-called *Influence Strength* (IS) index. Their results reinforced earlier conclusions, showing that the IS index surged during periods of crisis, particularly in 2008 and 2020. In addition to the Pearson method, various alternative ones have been proposed. One of them stems from Liu et al. [93] who introduced *Mutual Information* as a novel approach to take account for non-linear effects in intra-day S&P stock data. Other methods, such as Wavelet coherence and Fast Fourier Transform analysis, have also been utilized in various studies [87, 94]. We recall that the correlations in our study are measured by the Pearson metric.

Different approaches used to study the correlations in traditional markets have been applied to cryptocurrencies, such as Pearson correlation [26], BEKK-GARCH model [95] and Wavelet coherence [9]. A common conclusion in the existing literature is that the correlation among cryptocurrencies is transient. However, in contravention to this, Ethereum tends to be the central correlation transmitter to other coins throughout different periods. In other words, it tends to be highly correlated with most of the examined cryptocurrencies, while this characteristic does not exist in other coins [26, 96, 97]. Apart from that, similar to traditional markets, this emerging market also displays a different behavior during turbulent times, since the correlations between cryptocurrencies are also strengthened when the market is influenced by a major event, such as the Covid-19 pandemic that occurred in 2020 [64, 96].

In addition, some studies have addressed the issue of data granularity that has been a concern in traditional markets, where low-frequency data such as daily

and weekly-recorded time series are typically used. However, studies using high-frequency data only account for a small portion of the available literature. One example of this is the work conducted by Briola et al. who utilized high-frequency resolutions, including one hour and four hours as well as daily data, from 25 large-cap cryptocurrencies traded on the FTX exchange to discover the structural evolution of cryptocurrency correlation graphs at different data frequencies. With the main techniques being Pearson correlation and Minimum Spanning Tree, they found an increase in the complexity of the graphs' shapes for coarser time resolutions, meaning that cryptocurrencies tend to converge into a bigger group as resolution increases [98]. On the other hand, Kwapien et al. in [99] used multiple timescales starting at 10 to 360 minutes and proposed an opposite statement. Specifically, they revealed that low timescales lead to a more centralized graph, while higher timescales result in a more distributed and correlated one. In this regard, the authors explained this difference by pointing to liquidity and capitalization disparities, noting that cryptocurrencies with lower capitalization are traded less frequently than those with large capitalization, causing market information to spread over such cryptocurrencies more slowly. Notably, this study was among the few that removed the trend effect from the original dataset.

The correlation between cryptocurrencies and traditional assets has also attracted massive attention. One of the earliest studies on this subject was conducted by Yermack [8], who examined the daily co-movements between Bitcoin and national currencies (e.g., EUR, JPY, GBP) using Pearson correlation measurement. Since then, a variety of statistical measures such as Wavelet Coherence, Vector Autoregressive (VAR), Granger causality and q-Dependent Detrended Correlation have been widely employed on larger datasets that include a wide range of cryptocurrencies and other asset classes. These diverse approaches and datasets provide a broader view of the dependencies within the financial markets [10–12]. Overall, researchers demonstrated that the relationships between cryptocurrencies and traditional assets vary over time, with correlations frequently increasing and decreasing. However, one

consistent finding is that different asset classes become more closely related to each other during periods of economic turmoil, such as the recent event Covid-19 pandemic. Additionally, the cryptocurrency market is becoming more mature over time, and its correlation with other asset classes gradually increases [12, 100].

### 2.3.2 Graph-based Methods and its Applications

Regarding graph-based methods, several approaches to constructing a graph have been proposed. In [16], the authors constructed a Minimum Spanning Tree (MST) from 128 cryptocurrencies before and during the peak of the Covid-19 pandemic. Their goal was to discover changes in this graph structure under the pandemic's influence. They argued that the graph was more dispersed during the year 2019 (i.e. before the pandemic). By contrast, a centralized graph emerged during the first half of 2020, where cryptocurrencies formed a big community with the central node being USDT. A more complex alternative to MST is the Planar Maximally Filtered Graph (PMFG), which retains more edges along with a more complicated construction procedure. Using this method, Zhao et al. in [21] explored the changes in the correlation graph structure constructed from 322 stocks in the S&P500 index between 1994 and 2014. Similar to the majority of the literature, they captured the distinctive graph structure of the stock market during the market crashes. Besides, they also found that changes in the PMFG structure Granger-cause movements in the VIX index<sup>1</sup>. Therefore, the market volatility can potentially be predicted by analyzing the PMFG structure. Another increasingly popular approach is the Triangulated Maximally Filtered Graph (TMFG). This is one of the techniques employed in [101] to investigate the collapse of the Terra project. In particular, by constructing the TMFG of 61 cryptocurrencies from the beginning until mid-May 2022, the authors were able to identify key moments of market stress during the examined period, such as the intense selling pressure on Bitcoin on 05/05/2022, which they identified as the potential fuse for the process that led to the Terra

---

<sup>1</sup>The index measuring the market volatility. For more details, see [https://www.cboe.com/tradable\\_products/vix/](https://www.cboe.com/tradable_products/vix/)

project's failure. The authors also observed the marginalization of LUNA from the rest of the cryptocurrencies after 07/05/2022. Other methods like Frobenius distance-based kNN, multiplex graph and Graphical LASSO (GLASSO) are not widely adopted compared to the previous ones. However, they have shown their effectiveness in specific cases [19, 102, 103]. To date, Minimum Spanning Tree remains the most common method as it appears in the majority of current research, its popularity reinforces our confidence to utilize Minimum Spanning Tree in our study.

Thanks to graph-based methods, community detection is also employed which enriches the understanding of the correlation. This process aims to partition nodes of a graph into separate groups (called *communities*) by maximizing the correlation between nodes within a community while maximizing the distinction between different communities. In other words, this detection results in communities where nodes belonging to a community are most similar to each other in terms of their movements (e.g. returns and volatility) while nodes from different communities exhibit different patterns in movements. In [25], the authors applied the Louvain algorithm on the MST of 119 cryptocurrencies to cluster potential communities. They selected a 10-month period from April 2017 until February 2018 to study and divided it into two sub-periods to observe the evolution of the community structure. Their results suggest that the community structure does not persist over time, meaning that cryptocurrencies from different communities might form a new group as time evolves. With the communities found by the same method, Gavin and Crane in [104] went one step further by using Principal Component Analysis (PCA) to find an optimal portfolio out of 200 cryptocurrencies in circulation. On the other hand, another application of the Louvain method is discussed in [105]. Specifically, these authors used the daily prices of 465 stocks listed in the Tehran exchange to construct the Threshold-Based Correlation Graph, then applied the community detection algorithm to this graph for further analysis. By observing the stocks in each community and their constituent industries, the authors noticed that stocks

within the same industry tend to cluster together. This finding suggests that stock price correlations are heavily influenced by the industrial sector. A number of other methods are also being used to cluster similar entities but are less popular such as Girvan-Newman, Clauset algorithm, Stochastic block model (SBM), Latent Dirichlet Allocation (LDA) and Markov Random Field (MRF) [18, 106]. One obstacle from existing studies is that some used a specific community detection algorithm only, raising doubt about the robustness of the community structure algorithm. To this end, we first use the Louvain method to detect communities in our dataset and then adopt the Girvan–Newman method to examine the robustness of the communities found earlier.

## 2.4 Herding Behavior Discovery

### 2.4.1 Herding Behavior in Traditional Markets

The discussion of herding behavior among investors dates back over a century, paralleling the long historical development of traditional markets [107]. One of the earliest topics discovered in this area deals with herding behavior under different market conditions. In [108], Aharon applied 2 popular herding detection methods, CSSD and CSAD, to a dataset comprising historical daily data spanning 1990 to 2019. This dataset consists of 10 size-ranked portfolios constructed from stocks traded on the NYSE, AMEX and NASDAQ, covering different major events like the Asian crisis, the Dot-Com bubble, and the 2007-2008 Global Financial Crisis. He found evidence of herding within these stock assets throughout the examined period. Moreover, by integrating the volatility index (VIX index) into the herding models, he noticed that herding signals intensified when market volatility - as measured by the VIX index - increased. The same conclusion was also reached by Ahn et al [109] who used different datasets and herding detection methods. Specifically, based on the daily US stock return time series of 137 firms listed in the Standard & Poor 500 (S&P500) from January 1992 to December 2021, they utilized the Power law

exponent-based model to calculate the time-varying herding magnitude and then compared the results with an economic uncertainty index adopted from [110]. With access to longer and more recent data, the authors confirmed that, apart from the major events covered in the above study, herding was also more pronounced during the Covid-19 pandemic.

However, one downside of these studies is that they detected the herding over an entire period, leading to a lack of details. Thankfully, several studies filled this gap by using shorter periods at different points in time, creating a more complete picture of this phenomenon. For example, Chiang and Zheng in [111] observed the herding of 63 stocks originally from Thailand between January 1997 and December 1998 - a period that covers the Asian Financial Crisis. Using the CSAD model and daily return time series, they suggested the presence of herding in these stocks during the crisis. On the other hand, in [112], the authors collected monthly prices from the 20 most traded stocks in the Egyptian market over 12 years from July 2002 until May 2014. As distinct from most studies, the authors used a state-space model to detect the herding while dividing the data sample into pre-crisis, during the 2007-2008 crisis and the post-crisis period. Their findings revealed an absence of herding behavior before the crisis, transitioning to evidence of herding during the financial crisis and persisting afterward.

Notably, with the huge effect of the Covid-19 pandemic, many studies have explored herding behavior during this unprecedented time [113–118]. In [113], Bouri et al. selected 49 daily stock market indices representing 49 different countries sourced from Morgan Stanley Capital International (MSCI) indices<sup>2</sup>. Using the well-known CSAD model and a rolling window with a size of 102 days, they examined the time-varying global herding from January 2019 until August 2020. Their idea was to assess whether financial markets around the world exhibited the herding phenomenon during the onset of the pandemic. Indeed, it was shown that herding

---

<sup>2</sup>MSCI Inc., based in New York City, is an American financial company that offers a wide range of equity, fixed income, real estate indices, multi-asset portfolio analysis tools, as well as ESG and climate-related products. For more details, see <https://www.msci.com/indexes#featured-indexes>

was apparent from March to May 2020, coinciding with the worst period of the global economy caused by the pandemic. Other studies focusing on different types of investment vehicle, timeframes and methodologies also obtained the same results as investors tended to mimic each other's investment decisions during the first wave of the pandemic (first and second quarter of 2020) [119, 120].

Another aspect explored in the study of herding behavior is the comparison between different countries. A work by [121] conducted a literature review about herding behavior in different countries from 1995 to 2021, its authors found differences in external factors driving herding within different regions. For instance, global risk factors were identified as primary culprits influencing the stock market and giving rise to herding behavior among investors in the Turkish market [122]. On the other hand, the Chinese market tended to be affected by factors such as analyst recommendations, short-term investors' horizons as well as systemic risk [123]. Moreover, the authors also noticed that the majority of markets in developing countries were exposed to herding among retail investors (i.e. non-professional traders relying on personal judgment and subjectivity to make investment decisions). The purchasing power of these investors tends to be relatively small as limited by their personal earning ability. On the other hand, retail investors in developed countries tended to display less herding behavior due to the greater availability of information and knowledge of stock prices.

Another study conducted by Chen and Zheng [124] employed the Simulated Method of Moments estimator to investigate the herding behavior in the Chinese and US stock markets. The experiment was based on the Shanghai Composite Index representing the China market and the S&P500 Index representing the US market over the period between 01/01/1992 and 10/30/2017, including 6739 daily observations. The authors revealed both similar and different patterns in herding behavior between the two countries. Furthermore, they stated that the Chinese stock market was mainly driven by behavioral sentiment dynamics due to the switching behaviors of investors while the US market was mainly driven by fundamental factors.

The authors explained that this contradiction was attributed to the lack of a sufficiently transparent environment as well as comparatively few sophisticated investors in China, as compared to well-developed markets like the US. Additionally, over-speculative trading and the limited influence of fundamental factors in China further contributed to this phenomenon. More recently, Yang and Chuang in [115] investigated herding behavior across three different countries, including China, Taiwan and the US. They collected stock returns for non-financial firms listed on the Taiwan Stock Exchange (TWSE), the Shenzhen Stock Exchange (SZSE), the Shanghai Stock Exchange (SSE), and the New York Stock Exchange (NYSE) from January 2001 until June 2021. To observe the herding behavior, they combined the CSAD model with the Kalman filter and GARCH models. The authors discovered a relatively similar result to the above study. Specifically, Taiwan and China exhibited identical herding patterns with significant herding observed in two sub-periods: 2001-2004, coinciding with two local events, namely the Taiwan Strait crisis and the SARS pandemic and 2007-2009, corresponding to the Global Financial Crisis. By contrast, although herding was also pronounced during 2007-2009 in the US, no herding was apparent before this sub-period. Instead, this developed market revealed another instance of herding during the Covid-19 pandemic in 2020.

In general, most studies in traditional markets found herding displays during events that disrupt financial and economic stability. As explained in [117, 118, 125], during such events, investors tend to abandon their personal information and are influenced by financial news as well as the actions of other investors. Moreover, investors in traditional markets tend to be easily influenced by word of mouth, given the increase in uncertainty and negative news surging, with this being especially true among retail investors. Additionally, during uncertain times, both retail and institutional investors behave similarly and noise trading<sup>3</sup> is shown to increase during periods of high volatility. On the other hand, institutional investors are dominant

---

<sup>3</sup>Noise trading refers to buying and selling of financial assets based on market rumors, trends, or investor emotions, rather than on information such as earning reports, growth potential, or economic indicators (e.g. interest rates, inflation rates, etc). Noise traders make decisions based on the former sources rather than sound financial analysis.

when the markets experience an upward trend, resulting in contrarian herding during the positive market condition [109].

## 2.4.2 Herding Behavior in the Cryptocurrency Market

Cryptocurrencies are a relatively new asset class compared to traditional assets. Although the concept of digital coins was introduced in 2009 with the emergence of Bitcoin as the pioneering cryptocurrency [126], trading activities in this market only really commenced in 2013 [14]. As a result, studies on herding behavior in this domain mainly focus on the period from 2013 onward. In [127], Poyser identified herding in the top 100 cryptocurrencies by market capitalization from late April 2013 until mid-October 2019. He first used a Markov-switching model to segment the sample data into different regimes, then relied on a CSAD model to detect the time-varying herding in each regime. He found that herding behavior changed over time, with evidence of both herding and its absence throughout the examined period. Notably, herders were more active from mid-2017 until the end of the period since the signal of herding intensified during this interval. Additionally, in the second experiment, the author distinguished between 2 market conditions, namely positive and negative returns, to observe the herding behavior in each situation. The results revealed that cryptocurrency investors were more likely to follow the consensus when the market exhibited positive returns, while this tendency was diminished during periods of negative returns. Similarly, a study conducted by Junior et al. [128] collected more recent data, including 80 cryptocurrencies with the highest market capitalization from coinmarketcap.com between July 2015 and March 2020. In this work, they applied a state-space model adapting the standardized-beta method to detect the herding and used a regression model to test the relationship between the herding behavior and the market volatility. Like the previous study, the authors indicated a mix of the presence and absence of herding throughout the considered period. However, the absence of herding was revealed during most of 2019. Additionally, they found that the herding in cryptocurrencies became more pronounced

when the market volatility increased. This finding tallies with various studies, using different datasets and methodologies [129, 130].

Similar to traditional markets, the Covid-19 pandemic has also attracted significant attention in the cryptocurrency market. In [129], the authors gathered intraday trading data for 9 of the largest cryptocurrencies on the basis of market cap such as Bitcoin, Ethereum and Litecoin, covering a 2-year period from January 2019 to January 2021. They applied Patterson and Sharma’s herding intensity measure [131] to each cryptocurrency in two distinct sub-periods: the Pre-Covid-19 period (January 2019 - December 2019) and the Covid-19 pandemic period (January 2020 - January 2021). Unlike other aforementioned approaches, this method can detect the herding activities in each asset. Results suggested a dramatically greater herding intensity during the latter sub-period across all examined cryptocurrencies. Despite the fact that all cryptocurrencies showed the herding phenomenon during this time, no evidence of herding was found in Bitcoin, Ethereum, and Ripple during the Pre-Covid-19 sub-period. These findings align with those of Youssef and Waked who executed a hybrid model combining CSAD and Markov-switching methods on 43 cryptocurrency price time series with daily frequency [132]. In particular, their dataset spans from 28/04/2013 to 11/11/2020, divided into two sub-samples based on the Covid-19 pandemic, viz the Pre-Covid-19 from the beginning until the end of 2019 and the Covid-19 outbreak thereafter. As expected, only the second sample (i.e. the Covid-19 outbreak) revealed the herding evidence. Furthermore, for the second part of the work, the authors used the coronavirus Media Coverage Index (MCI)<sup>4</sup> which calculates the percentage of all news sources covering the novel coronavirus topic to test the relationship between news and the herding magnitude. Interestingly, they found that news about the pandemic significantly drove the herding behavior among investors in the cryptocurrency market, indicating the fact that cryptocurrency traders are mainly influenced by news and people’s sentiments. The

---

<sup>4</sup>The Coronavirus MCI is calculated using all news sources that report on Covid-19. The index scale is from 0 to 100, where a value of 50 on a given day indicates that 50% of all news outlets are covering Covid-19-related stories. This index is published by Avenpack. For more details, see <https://coronavirus.ravenpack.com/>.

result is in line with most of the existing literature [133–136].

More recently, a few studies have utilized a newer dataset that covers up to the year 2023 [137, 138]. However, instead of segmenting the sample into smaller intervals like the previous studies did, they mainly retained the whole period which caused ambiguity for herding detection models. Consequently, their findings tended to suggest an absence of herding throughout the examined period, even during highly uncertain times such as the outbreak of the Covid-19 pandemic.

From the existing literature, it is evident that most studies investigating the financial markets (both traditional and cryptocurrency) focus on periods before the year 2022, while only a few works extend into more recent times. To this end, our work contributes to the existing literature by using more up-to-date data, which spans from April 2019 until May 2023. Moreover, although several studies divided their data into smaller sub-periods, these divisions seem not to be sufficiently refined. To address this, we segment our data into finer sub-periods corresponding to major events that occurred in the financial markets throughout the examined period. Further details of this timeline division will be described in Section 5.5.2.

# Chapter 3

## Dataset Description

### 3.1 Introduction

Our study consists of various experiments aiming to answer different research questions whose requirements of data vary, such as entities in a dataset, timeframe, external variables, etc. Therefore, different datasets are collected in our work to cope with different demands. In particular, our first set of data focuses particularly on the cryptocurrency market, comprising 34 historical high-frequency price time series representing 34 corresponding cryptocurrencies sourced from the HitBTC exchange<sup>1</sup>, a platform for digital asset and currency exchange to quickly and securely trade cryptocurrencies. On the other hand, the second part of our work expands to traditional assets, namely stocks and ETFs, which requires us to use a broader dataset. To this end, our second dataset consists of 222 historical price time series in three types of investment vehicle, including 27 cryptocurrencies, 49 ETFs and 146 stocks, which are collected from the First Rate platform<sup>2</sup>, a leading provider offering high-frequency intraday prices of financial assets on a variety of markets. In both cases, we use tick-by-tick data for all time series, meaning that the price value is collected whenever a pair of bid-ask is matched. Thus, an asset might have more than one price value within a second.

---

<sup>1</sup>Website address: <https://hitbtc.com/>

<sup>2</sup>Website address: <https://firstratedata.com/>

### 3.1.1 First Dataset

The first dataset comprises historical prices of 34 cryptocurrencies observed solely on the HitBTC exchange from 13 February 2019 until 06 April 2021, with a hybrid of high and low market capitalizations. Specifically, the highest rank is 1 (Bitcoin) while the lowest rank is 260 (FunToken), according to the price-checking website Coinmarketcap<sup>3</sup> in April 2021, the full list of coins can be seen in Table 3.1. The choice of these cryptocurrencies is to ensure the availability of data and minimize the percentage of missing values. In other words, these were the most frequently traded coins on the HitBTC exchange during the examined period.

Table 3.1: **A list of 34 cryptocurrencies used in our study. Abbreviations are put in parentheses.**

Cryptocurrencies					
Argur (REP)	Bitcoin SV (BSV)	Ethereum Classic (ETC)	MaidSafeCoin (MAID)	Ontology (ONT)	Tron (TRX)
Bancor (BNT)	Cardano (ADA)	FunToken (FUN)	Maker (MKR)	Ox (ZRX)	Verge (XVG)
Basic Attention Token (BAT)	Decentraland (MANA)	ICON (ICX)	Monero (XMR)	QTUM	Zcash (ZEC)
Bitcoin (BTC)	Dogecoin (DOGE)	IOST	Nem (XEM)	Ripple (XRP)	Zilliqa (ZIL)
Bitcoin Cash (BCH)	EOS	Lisk (LSK)	NEO	Stellar (XLM)	
Bitcoin Gold (BTG)	Ethereum (ETH)	Litecoin (LTC)	OMG Network (OMG)	Tezos (XTZ)	

Regarding the HitBTC exchange, it is chosen for our study for three main reasons. Firstly, data accessibility and availability are critical considerations. Among the limited available sources, HitBTC provides the highest number of cryptocurrencies while maintaining a low rate of missing data. Since our objective is to use as many cryptocurrencies as possible, this exchange is our top priority. Secondly,

<sup>3</sup>Website address: <https://coinmarketcap.com>

while HitBTC is not the largest cryptocurrency exchange, it is recognized as one of the major global exchanges [139]. Thirdly, this exchange is frequently used in the literature [139–141].

Regarding the ranking<sup>4</sup> of each cryptocurrency, we collect monthly rankings from the Historical Snapshots page on Coinmarketcap from February 2019 to April 2021 and average them to obtain an overall ranking for each coin. This practice enables us to have a general assessment in terms of the size of a cryptocurrency during the considered period. A major coin tends to catch more attention of the general public and is utilized more often in cryptocurrency-related activities. This feature is displayed in Table 3.2

**Table 3.2: Average rankings of cryptocurrencies during the period between February 2019 and April 2021**

BTC	ETH	XRP	BCH	LTC	ADA	BSV	EOS	XLM
1	2	4	6	7	10	11	11	12

TRX	XMR	XTZ	NEO	XEM	ETC	MKR	DOGE	ZEC
15	15	18	22	25	28	30	32	32

ONT	BAT	OMG	QTUM	ZRX	BTG	ICX	LSK	REP
37	40	51	51	51	59	60	64	66

ZIL	IOST	XVG	MANA	MAID	BNT	FUN
70	90	91	110	117	144	202

### 3.1.2 Second Dataset

The second dataset comprises 222 assets categorized into 3 types of investment vehicle, namely cryptocurrency (27 assets), stock (146 assets) and the US ETF (49 assets). Each time series starts on 01/04/2019 until 03/05/2023. The selection of these assets is based on four criteria:

- **Data availability:** all the assets are available on the FirstRate platform.

<sup>4</sup>The ranking of a cryptocurrency is determined by its market capitalization (current price  $\times$  number of tokens in circulation). The higher the market capitalization, the higher the corresponding rank assigned to the cryptocurrency.

- Market capitalization: since we are limited to little more than 50 assets for cryptocurrencies and US ETFs, we try to use all of them. On the other hand, there are thousands of stocks available to use so we narrow these down to the top 200 stocks whose companies are in the top 200 in terms of market capitalization, according to the web page Companies Market Cap<sup>5</sup>.
- High trading frequency: To avoid missing values, we only use assets with high trading frequency. Specifically, we remove any stocks, cryptocurrencies, and US ETFs that have more than 1%, 10%, and 12% missing values, respectively. These thresholds are the lowest we can set while still maintaining a reasonable number of assets in each type of investment vehicle. Using stricter thresholds would result in a shortage of data. The fact is that after filtering the assets, the majority of the remaining have less than 1 percent of missing values.
- In the case of US ETFs: the ability to act as a proxy for a corresponding US stock index.

We note that US ETFs are considered a separate type of investment vehicle in this study. This decision is made due to the unique characteristics of ETFs. In particular, they act as a proxy for stock indices, offering higher liquidity and lower trading costs than the stock market [142–144]. In this way, an ETF can be compared to a tracker fund in that it will seek to replicate a particular index, representing the overall performance of a set of assets [144].

Since the official trading time of traditional markets is between 09:30 a.m. and 4:00 p.m., we select this timeframe as the boundary for our time series, meaning that the cryptocurrency data is assumed to only be available within this trading time, even though these coins are traded nonstop. Other price values outside of this timeframe are eliminated. This practice facilitates the comparison between assets from different classes and also preserves their equality. A list of assets in each type of investment vehicle is given in Table 3.3

---

<sup>5</sup><https://companiesmarketcap.com/>

Table 3.3: Three types of investment vehicle in the second dataset and their corresponding assets

Classes	Assets
Cryptocurrencies	ADA (Cardano), BAT (Basic Attention Token), BCH (Bitcoin Cash), BNT (Bancor), BSV (Bitcoin SV), BTC (Bitcoin), DAI (Dai), DASH (Dash), DOGE (Dogecoin), EOS (EOS), ETH (Ethereum), ETC (Ethereum Classic), LTC (Litecoin), NEO (Neo), OMG (OMG Network), QTUM (Qtum), REP (Augur), TRX (TRON), USDT (Tether), UST (TerraUSD), XLM (Stellar), XMR (Monero), XRP (Ripple), XTZ (Tezos), XVG (Verge), ZEC (Zcash), ZRX (Ox)
US ETFs	CIBR (First Trust Nasdaq Cybersecurity ETF), DIA (SPDR Dow Jones Industrial Average ETF), DJD (Invesco Dow Jones Industrial Average Dividend ETF), DJP (iPath Bloomberg Commodity Index Total Return ETN), DVY (iShares Select Dividend ETF), FDLO (Fidelity Low Volatility Factor ETF), FDN (First Trust Dow Jones Internet Index Fund), FNGU (MicroSectors FANG+ Index 3X Leveraged ETN), FTXO (First Trust Nasdaq Bank ETF), IBB (iShares Nasdaq Biotechnology ETF), IDU (iShares U.S. Utilities ETF), IEO (iShares U.S. Oil & Gas Exploration & Production ETF), IJH (iShares Core S&P Mid-Cap ETF), IJJ (iShares S&P Mid-Cap 400 Value ETF), IJK (iShares S&P Mid-Cap 400 Growth ETF), IJR (iShares Core S&P Small-Cap ETF), IJS (iShares S&P Small-Cap 600 Value ETF), IJT (iShares S&P Small-Cap 600 Growth ETF), ITB (iShares U.S. Home Construction ETF), ITOT (iShares Core S&P Total U.S. Stock Market ETF), IUSG (iShares Core S&P U.S. Growth ETF), IVE (iShares S&P 500 Value ETF), IWB (iShares Russell 1000 ETF), IWM (iShares Russell 2000 ETF), IWV (iShares Russell 3000 ETF), IYE (iShares U.S. Energy ETF), IYF (iShares U.S. Financials ETF), IYJ (iShares U.S. Industrials ETF), IYR (iShares U.S. Real Estate ETF), IYT (iShares Transportation Average ETF), IYW (iShares U.S. Technology ETF), IYY (iShares Dow Jones U.S. ETF), KBWB (Invesco KBW Bank ETF), OEF (iShares S&P 100 ETF), ONEQ (Fidelity Nasdaq Composite ETF), PNQI (Invesco Nasdaq Internet ETF), QQQ (Invesco QQQ Trust), QTEC (First Trust NASDAQ-100 Technology Sector Index Fund), QYLD (Global X NASDAQ 100 Covered Call ETF), ROBT (First Trust Nasdaq Artificial Intelligence and Robotics ETF), RPG (Invesco S&P 500 Pure Growth ETF), RPV (Invesco S&P 500 Pure Value ETF), SLYG (SPDR S&P 600 Small Cap Growth ETF), SLYV (SPDR S&P 600 Small Cap Value ETF), SLY (SPDR S&P 600 Small Cap ETF), SOXX (iShares PHLX Semiconductor ETF), SPY (SPDR S&P 500 ETF), SVXY (ProShares Short VIX Short-Term Futures ETF), XTN (SPDR S&P Transportation ETF).
Stocks	AAPL (Apple Inc.), ABBV (AbbVie Inc.), ABB (ABB Ltd.), ABT (Abbott Laboratories), ACN (Accenture plc), ADBE (Adobe Inc.), ADI (Analog Devices, Inc.), ADP (Automatic Data Processing, Inc.), AIR (AAR Corp.), ALV (Autoliv Inc.), AMAT (Applied Materials, Inc.), AMD (Advanced Micro Devices, Inc.), AMGN (Amgen Inc.), AMT (American Tower Corporation), AMZN (Amazon.com Inc.), ASML (ASML Holding NV), ASX (ASE Technology Holding Co., Ltd.), AVGO (Broadcom Inc.), AXP (American Express Company), AZN (AstraZeneca PLC), BABA (Alibaba Group Holding Limited), BAC (Bank of America Corp.), BA (The Boeing Company), BDX (Becton, Dickinson and Company), BHP (BHP Group Limited), BKNG (Booking Holdings Inc.), BKR (Baker Hughes Company), BLK (BlackRock Inc.), BMY (Bristol Myers Squibb Company), BP (BP plc), BSX (Boston Scientific Corporation), BTI (British American Tobacco plc), BUD (Anheuser-Busch InBev SA/NV), BX (The Blackstone Group Inc.), CAT (Caterpillar Inc.), CB (Chubb Limited), CFR (Cullen/Frost Bankers, Inc.), CI (Cigna Corporation), CMCSA (Comcast Corporation), CNI (Canadian National Railway Company), COP (ConocoPhillips), COST (Costco Wholesale Corporation), CP (Canadian Pacific Railway Limited), CRM (Salesforce.com Inc.), CSCO (Cisco Systems, Inc.), CSL (Carlisle Companies Incorporated), CVS (CVS Health Corporation), CVX (Chevron Corporation), C (Citigroup Inc.), DEO (Diageo plc), DE (Deere & Company), DHR (Danaher Corporation), DIS (The Walt Disney Company), DTE (DTE Energy Company), ELV (Eaton Vance Corporation), EL (The Estée Lauder Companies Inc.), EOG (EOG Resources Inc.), EQNR (Equinor ASA), ETN (Eaton Corporation plc), GE (General Electric Company), GILD (Gilead Sciences, Inc.), GOOG (Alphabet Inc.), GS (The Goldman Sachs Group, Inc.), HCA (HCA Healthcare, Inc.), HDB (HDFC Bank Limited), HD (The Home Depot, Inc.), HON (Honeywell International Inc.), HSBC (HSBC Holdings plc), IBM (International Business Machines Corporation), IBN (ICICI Bank Limited), IDEX (Ideanomics, Inc.), INTC (Intel Corporation), INTU (Intuit Inc.), ISRG (Intuitive Surgical, Inc.), ITW (Illinois Tool Works Inc.), JNJ (Johnson & Johnson), JPM (JPMorgan Chase & Co.), KO (The Coca-Cola Company), LIN (Linde plc), LLY (Eli Lilly and Company), LMT (Lockheed Martin Corporation), LOW (Lowe's Companies, Inc.), LRCX (Lam Research Corporation), MA (Mastercard Incorporated), MCD (McDonald's Corporation), MC (Moelis & Company), MDLZ (Mondelez International, Inc.), MDT (Medtronic plc), META (Meta Platforms, Inc.), MMC (Marsh & McLennan Companies, Inc.), MO (Altria Group, Inc.), MRK (Merck & Co., Inc.), MSFT (Microsoft Corporation), MS (Morgan Stanley), MUFG (Mitsubishi UFJ Financial Group, Inc.), MU (Micron Technology, Inc.), NEE (NextEra Energy, Inc.), NFLX (Netflix, Inc.), NKE (NIKE, Inc.), NOW (ServiceNow, Inc.), NVDA (NVIDIA Corporation), NVO (Novo Nordisk A/S), NVR (NVR, Inc.), NVS (Novartis AG), ORCL (Oracle Corporation), PANW (Palo Alto Networks, Inc.), PBR (Petroleo Brasileiro S.A. - Petrobras), PDD (Pinduoduo Inc.), PEP (PepsiCo, Inc.), PFE (Pfizer Inc.), PG (Procter & Gamble Co.), PLD (Prologis, Inc.), PM (Philip Morris International Inc.), PYPL (PayPal Holdings, Inc.), QCOM (QUALCOMM Incorporated), SAP (SAP SE), SBUX (Starbucks Corporation), SCHW (The Charles Schwab Corporation), SHEL (Sealed Air Corporation), SHOP (Shopify Inc.), SLB (Schlumberger Limited), SNY (Sanofi), SONY (Sony Group Corporation), SO (The Southern Company), SYK (Stryker Corporation), TD (The Toronto-Dominion Bank), TJX (The TJX Companies, Inc.), TMO (Thermo Fisher Scientific Inc.), TMUS (T-Mobile US, Inc.), TM (Toyota Motor Corporation), TSLA (Tesla, Inc.), TSM (Taiwan Semiconductor Manufacturing Company Limited), TTE (TotalEnergies SE), TXN (Texas Instruments Incorporated), T (AT&T Inc.), UL (Unilever PLC), UNH (UnitedHealth Group Incorporated), UNP (Union Pacific Corporation), UPS (United Parcel Service, Inc.), VRTX (Vertex Pharmaceuticals Incorporated), VZ (Verizon Communications Inc.), V (Visa Inc.), WFC (Wells Fargo & Co.), WMT (Walmart Inc.), XOM (Exxon Mobil Corporation), ZTS (Zoetis Inc.)

Moreover, unlike cryptocurrencies and US ETFs, the stock market is distinctly classified into 11 sectors based on the specific business activities of companies. These include Utilities, Real Estate, Materials, Information Technology, Industrials, Healthcare, Financials, Energy, Consumer Staples, Consumer Discretionary and Communication Services [145]. This classification system is a fundamental feature of the stock market, providing a structured way to analyze and compare companies that have similar business models, which is absent in cryptocurrencies and US ETFs. However, while there is no official classification for ETFs and most ETFs contain assets from different stock sectors, we emphasize the existence of ETFs designed to track the performance of a specific stock sector. For instance, IYW represents the iShares U.S. Technology ETF which tracks technology stocks. Thus, although we do not classify ETFs, we keep in mind this characteristic. As a result, each asset is assigned to a specific sector, either cryptocurrency, US ETF, or one of the aforementioned stock sectors. We use this information in our experiments for a deeper investigation and analysis.

We note that the number of assets in each stock sector is different and depends on its asset's market capitalization. For instance, Technology, Finance and Healthcare stocks tend to have very high market capitalization while Communication Services and Real Estate tend to have lower market capitalization [146]. Consequently, sectors that tend to have high market capitalization assets have a higher number of entities than the others. Further details for this classification are provided in Table 3.4

## **3.2 A Note on the Granularity of Examined Time Series**

Our first dataset comprises only cryptocurrencies, traded throughout the day and on all seven days per week. Therefore, there is no official trading hour for cryptocurrencies as they are traded nonstop. In other words, the price values of each

Table 3.4: Distribution of Business Sectors in the Stock Dataset

Information Technology	Healthcare	Financials	Consumer discretionary	Consumer staples	Industrials	Energy	Utilities	Materials	Real estate	Communication services
ASX	ABBV	AXP	ALV	BTI	ABB	BKR	DTE	BHP	AMT	
AAPL	ABT	BAC	AZMN	BUD	AIR	BP	NEE	LIN	PLD	
ACN	AMGN	BLK	BKNG	DEO	BA	COP	SO			
ADBE	AZN	BX	COST	EL	CAT	CVX				
ADI	BDX	CB	CSL	GE	CNI	EQNR				
ADP	BMJ	CFR	DIS	HON	CP	PBR				
AMAT	BSX	C	HD	ITW	DE	SHEL				
AMD	CI	GS	IDEX	KO	ETN	SLB				
ASML	CVS	HDB	LOW	MDLZ	LMT	TTE				
AVGO	DHR	HSBC	MCD	MO	UNP	EOG				
BABA	ELV	IBN	NKE	PEP	UPS	XOM				
CMCSA	GILD	JPM	NVR	PG	PYPL					
CRM	HCA	MC	SBUX	PM						
CSCO	ISRG	MMC	TJX	UL						
GOOG	JNJ	MS	TM	WMT						
IBM	LLY	MUFG	TSLA							
INTC	MDT	SCHW								
INTU	MRK	TD								
LRCX	NVO	WFC								
MA	NVS									
META	PFE									
MSFT	SNY									
MU	SYK									
NFLX	TMO									
NOW	UNH									
NVDA	VRTX									
ORCL	ZTS									
PANW										
PDD										
QCOM										
SAP										
SHOP										
SONY										
TMUS										
TSM										
TXN										
T										
VZ										
V										

cryptocurrency are available all day, i.e. from 00:00 a.m. until 23:59 p.m. On the other hand, besides cryptocurrencies, our second dataset contains a large portion of traditional assets, including stocks and ETFs. As stated previously, traditional assets strictly comply with the official trading hours between 09:30 a.m. and 16:00 p.m. from Monday to Friday as well as outside of public holidays. Thus, the price values of each time series in this dataset are only available from 09:30 a.m. until 16:00 p.m. The difference in timeframes leads to the difference in possible granularities that can be utilized between the two datasets. In particular, to guarantee data availability and avoid missing values, we consider the finest level of granularity that meets our requirements. To this end, we select four granularities for the first dataset, including 30-min, 6-hour, 12-hour, and 24-hour. By contrast, due to the limitation in trading hours for stock and ETF assets in the second dataset, we only select 30-minute and 24-hour frequencies.

From these granularities, a common option is selected for both datasets. Specifically, out of 3 types of investment vehicle, i.e. stocks, cryptocurrencies and US

ETFs, cryptocurrencies show higher fluctuations compared to traditional assets. In [147], the authors used 5 min data of Bitcoin prices traded on three different exchanges, Kraken, Bitstamp and Btcbbox, during the period between 2017 and 2021 to calculate the realized volatility (the assessment of variation in returns for an asset by analyzing its historical returns within a defined time period) of this most stable and popular cryptocurrency. Results showed that although Bitcoin is the most valuable and trustworthy cryptocurrency, its volatility fluctuates from 4.8 to 7.5. By contrast, with the same level of granularity, the stock market seems to be more stable, as the realized volatility stood at roughly 0.58 during normal times [148] and increased to just around 1.0 during turbulent ones [149]. These facts suggest that the cryptocurrency price fluctuations are dramatic even within a 5-minute period, while traditional assets, although also volatile, are much more stable at this fine-grained level. Consequently, using a low-frequency dataset such as 12h or 24h appears to cause a loss of important information that influences experimental results, especially in cryptocurrency-related problems. Therefore, our priority is to select the finest level possible in order to retain the highest percentage of information in each asset. To this end, the 30-minute frequency was selected as the best option to use due to the limitations of cryptocurrency data availability. The choice of 30-min frequency will be discussed comprehensively in Section 5.2.4.

### 3.3 Statistical Property Comparison

Statistical characteristics of an asset are captured via its distribution moments, i.e. mean ( $\mu$ ), standard deviation ( $\sigma$ ), kurtosis ( $\kappa$ ) and skewness ( $\tilde{\mu}_3$ ). We note that these moments of each asset are calculated using the asset's log-returns as our experiments throughout this study primarily rely on log-return time series - a transformation derived from price time series (as detailed in Section 4.2). In this section, we examine the differences and similarities in these statistical properties across cryptocurrencies, stocks and US ETFs from the second dataset introduced earlier.

Since there are four statistical indicators, a 4-dimensional visualization is undoubtedly the best choice. However, it is not feasible to visualize the values in such way as it is difficult to observe. Thus, we utilize the so-called Principal Component Analysis (PCA)[150] method to map the 4-dimensional space into a 3-dimensional one. Briefly, PCA is a statistical approach that simplifies a dataset by converting its original dimensions (features) into a reduced set of independent features, referred to as *principal components* [151]. The transformed space has fewer dimensions (e.g. fewer features) but still retains the essential information from the original space. This technique is widely used to simplify complex datasets, aiding visualization, analysis, and feature extraction for machine learning.

Figure 3.1 displays the allocation of all assets from the second dataset, in the PCA-based 3-dimensional embedding space. The closer the two assets are, the more similar they are in terms of distribution moments. We notice an intertwinement between stocks and US ETFs as they are close to each other, forming an area with a high density of assets, indicating shared traits in such statistical properties between these traditional assets. On the other hand, cryptocurrencies display a separation from traditional assets, meaning that the statistical properties of cryptocurrencies differ more from traditional assets than the differences between stocks and US ETFs. To reinforce this result, we calculate the *average pairwise distance* between these three types of investment vehicle for each statistical property. Specifically, for each statistical property, we calculate the absolute difference between every asset from one investment vehicle and every asset from another investment vehicle, then average the differences to obtain the *average pairwise distance* between the two investment vehicles. The formula for the *average pairwise distance* between two investment vehicles is described in Equation 3.1.

$$\text{average pairwise distance}_{A,B} = \frac{\sum_{i \in A, j \in B} |\text{stat}_i - \text{stat}_j|}{\|A\| + \|B\|} \quad (3.1)$$

Where  $A, B$  represent two types of investment vehicle;  $i, j$  are the assets in  $A, B$ , respectively;  $\text{stat}_i$  represents a distribution moment of asset  $i$ ;  $\|\cdot\|$  is the number

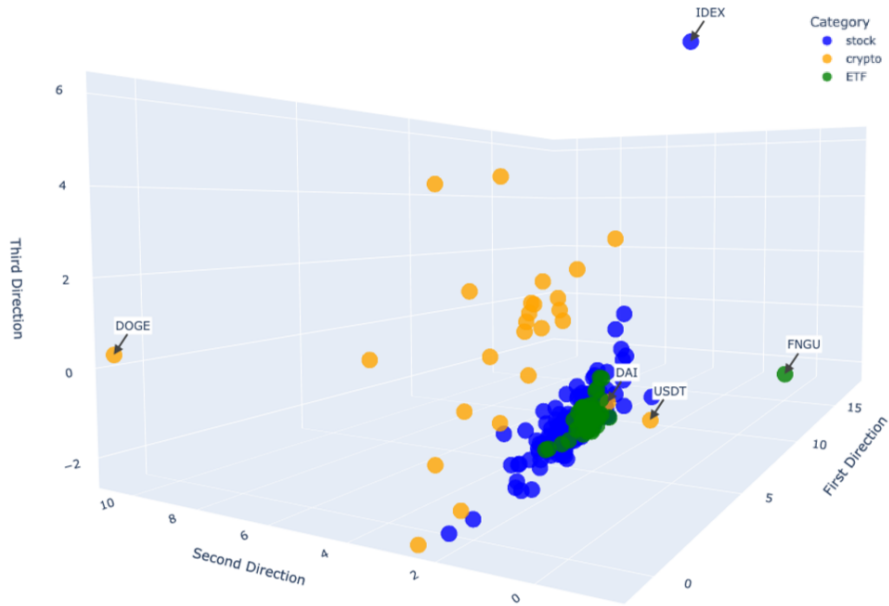


Figure 3.1: **Distribution moments of all assets across three types of investment vehicle from the second dataset obtained by using 3-dimensional PCA. Those further from the rest are marked by their names.**

of assets in a type of investment vehicle.

Table 3.5 displays the *average pairwise distance* between stocks, US ETFs and cryptocurrencies in each statistical property, including mean, standard deviation, kurtosis and skewness. In all cases, the distance between stocks and US ETFs is the lowest, compared to the distance between these traditional assets and cryptocurrencies, meaning that the statistical properties of stocks and US ETFs are more similar to each other while cryptocurrencies show a broader gap from the others. Notably, the great similarity in statistical properties between stocks and US ETFs and the separation of cryptocurrencies from these traditional assets is strongly evidenced through the mean and standard deviation. Overall, this result emphasizes the gap between cryptocurrencies and traditional assets such as stocks and US ETFs in terms of statistical properties, which explains the separation between cryptocurrencies and other assets in Figure 3.1.

Figure 3.2 provides details about these fundamental statistical properties for all assets in the second dataset, categorized into three types of investment vehicle, namely cryptocurrencies (orange color), stocks (blue color) and US ETFs (green

Table 3.5: The average distance between stocks and US ETFs, stocks and cryptocurrencies, and US ETFs and cryptocurrencies, evaluated for each statistical indicator

Comparison	Mean	Std	Kurtosis	Skewness
Stocks vs US ETFs	0.000031	0.002223	43.669915	1.452585
Stocks vs Cryptocurrencies	0.000077	0.018741	57.014938	1.678618
US ETFs vs Cryptocurrencies	0.000070	0.019731	47.683847	1.509051

color).

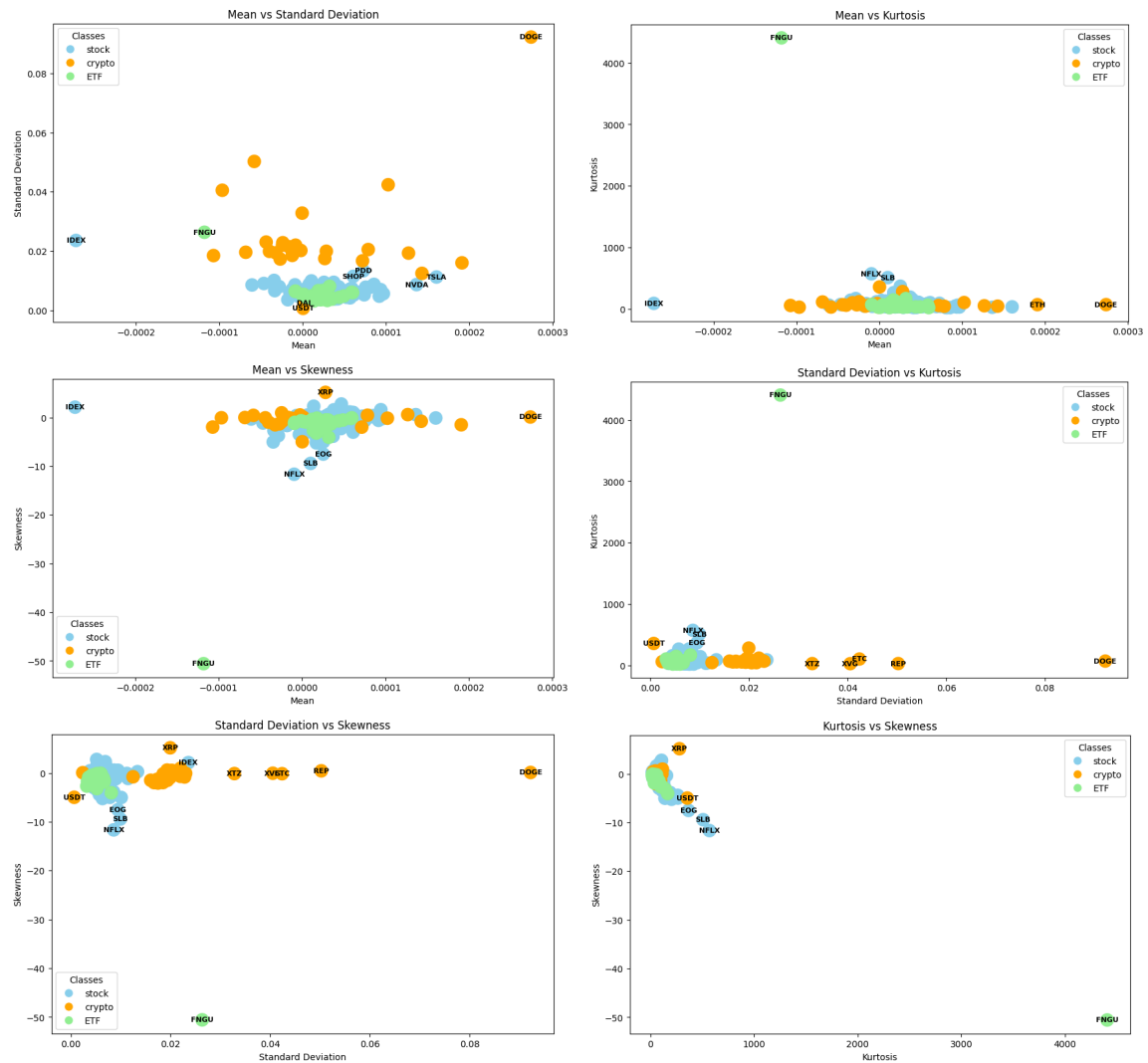


Figure 3.2: All assets (e.g. stocks, US ETFs and cryptocurrencies) in the second dataset visualized by pairs of statistical indicators, including  $\mu - \sigma$ ,  $\mu - \kappa$ ,  $\mu - \tilde{\mu}_3$ ,  $\sigma - \kappa$ ,  $\sigma - \tilde{\mu}_3$ , and  $\kappa - \tilde{\mu}_3$ . Where  $\mu = \text{Mean}$ ,  $\sigma = \text{Standard Deviation}$ ,  $\kappa = \text{Kurtosis}$ , and  $\tilde{\mu}_3 = \text{Skewness}$ .

### 3.4 Remarks

This chapter introduces two datasets used in our research, each of them serves a specific purpose. In particular, the first dataset comprising only cryptocurrencies is used for the first 3 experiments, which are equivalent to Sections 5.2-5.4. These experiments answer cryptocurrencies-related questions only. On the other hand, the second dataset comprising not only cryptocurrencies but also stocks and US ETFs is used for the remaining experiments, where an expansion to traditional markets is conducted. These experiments are discussed in Section 5.5. Additionally, we emphasize the use of 30-min frequency for our datasets in this chapter. The reason for this choice is interpreted comprehensively in Section 5.2.4.

Lastly, we compare the differences and similarities in statistical characteristics between cryptocurrencies, stocks and US ETFs. Our results indicate that stocks and US ETFs share similar distribution moments (e.g. mean, standard deviation, kurtosis and skewness) to some extent. By contrast, these statistical indicators in cryptocurrencies exhibit unique characteristics, separating them from the aforementioned traditional assets. Given the difference between cryptocurrencies and traditional assets in statistical characteristics, we seek to examine whether this distinction remains unchanged when looking at other aspects, such as the way investors participate in these markets and their price movements under different market conditions. These promising questions will be answered in the following Chapters.

# Chapter 4

## Methodology

### 4.1 Introduction

We recall that our research revolves around the issue of correlation between different financial assets, i.e. how do they relate to each other at a certain time and how does this relationship vary over time? Subsequently, we utilize the investors' behavior, such as the decision-making tendency for investment, their emotions and preferences, under a market condition (i.e. stable, crisis, bull/bear), to explain our experimental results. With this regard, our main techniques are graph-based methods, namely graph construction, minimum spanning tree and community detection, along with graph-based metrics for evaluation of this. Moreover, herding detection techniques used to detect the herding behavior among investors also play a crucial role in our research for a further investigation of the relationship between investors' behavior and assets' correlation. This chapter introduces step-by-step all techniques, metrics and terminologies that are necessary for this study.

## 4.2 Return and Volatility Calculations from Price Time Series

Given a price time series  $x_i = (x_1^i, x_2^i, x_3^i, \dots, x_T^i)$  of an asset  $i$  with length  $T$ , the log-return time series  $r_i = (r_1^i, r_2^i, r_3^i, \dots, r_T^i)$  is obtained using the formula  $r_t^i = \log(x_t^i/x_{t-1}^i)$ , where  $x_t^i$  and  $r_t^i$  represent the price and return value at time  $t$ , respectively [152].

The volatility  $v_t^i$  of asset  $i$  at time  $t$  is calculated using the last  $l$  data points from the corresponding return time series. The value of  $l$  depends on the experiments' purposes as well as the datasets used. In this study, it is chosen to be 48 (justification below). We calculate the volatility in 2 ways to ensure the stability and transparency of experimental results.

1. Using the standard deviation of log-returns [153]:

$$v_t^i = \sqrt{\frac{\sum_{k=0}^{l-1} (r_{t-k}^i - \mu)^2}{l}} \quad (4.1)$$

where  $\mu$  is the mean of these log-returns.

2. Using the moving average of squared log-returns [154]:

$$v_t^i = \sqrt{\frac{\sum_{k=0}^{l-1} r_{t-k}^i{}^2}{l}} \quad (4.2)$$

The choice of  $l = 48$  is explained by two reasons: Firstly, we use a 30-minute dataset, which results in 48 data points within a day (24 hours). In other words, we capture the intraday movements of a cryptocurrency by employing a sliding window of size 48 with a sliding step of 30 minutes. Secondly, by examining the market efficiency [155] in cryptocurrencies, a window of 48 data points is sufficiently long for a volatility measurement to capture genuine market movements rather than arbitrary fluctuations. In brief, market efficiency refers to the extent to which asset prices

fully reflect all available information such as earnings announcements, macroeconomic indicators and news, as proposed by the Efficient Market Hypothesis [155]. Therefore, if an asset shows a high degree of market efficiency, its prices as well as price movements reflect what is happening in the market in real-time. In this study, we test the weak-form efficiency (because we only have the price time series)<sup>1</sup> for each cryptocurrency using a time window of 48 data points. This procedure follows the approach proposed by [157], where two key metrics are employed: *permutation entropy*<sup>2</sup> and *statistical complexity*<sup>3</sup>. Specifically, given an empirical time series in a corresponding time window, the authors first calculated permutation entropy and statistical complexity for the time series, then estimated the 95% random confidence interval for each measure by shuffling the time series randomly and calculating permutation entropy and statistical complexity for 30 independent realizations. The idea is based on the theory that if two measures (e.g. permutation entropy and statistical complexity) derived from the empirical time series are within the 95% confidence intervals of randomly shuffled time series, then the empirical time series share common characteristics with random time series. In this case, no dependence patterns exist in the empirical time series, and the future values are independent of past values, indicating that the time series exhibits market efficiency. We found that the majority of assets in our study retain their efficiency for roughly 70 percent of the observed period. This percentage is relatively high for digital currencies, which suggests the equivalence between a coin's prices and the relevant information circulating on that coin at that time, according to [157]. Therefore, the window size of

---

<sup>1</sup>Weak-form market efficiency theory: prices of securities at every instant fully reflect all available information of past prices. This means future price movements cannot be predicted by using past prices [156].

<sup>2</sup>*Permutation entropy* is a non-parametric complexity measure based on ordinal patterns, used to analyze the temporal structure of both linear and non-linear time series. A low value of permutation entropy (near 0) suggests that future values are predictable based on past data. Conversely, a high permutation entropy (near 1) indicates that the values in the time series are independent and follow a uniform distribution. This measure is derived by examining the frequency of various ordinal patterns (permutations) within the dataset [158].

<sup>3</sup>*Statistical complexity* measures the structural complexity of dependence patterns in a time series. A higher level of statistical complexity indicates that these patterns occur in a more intricate manner, meaning the time series exhibits greater fluctuations and the relationship between future and past values is more complex [159].

48 data points can capture and explain the genuine movements of a cryptocurrency.

The results of this experiment are shown in Table 4.1.

Table 4.1: **Testing the market efficiency for each cryptocurrency using 48 data points, over the period from February 2019 to April 2021.**

Crypto	Percentage		
ADA	70,00%	MANA	54,32%
BAT	55,00%	MKR	58,09%
BCH	70,37%	NEO	70,10%
BNT	63,80%	OMG	69,17%
BSV	70,32%	ONT	69,34%
BTC	69,37%	QTUM	69,77%
BTG	71,17%	REP	68,38%
DOGE	70,00%	TRX	70,21%
EOS	69,53%	XEM	70,80%
ETC	68,72%	XLM	70,63%
ETH	70,08%	XMR	69,34%
FUN	43,44%	XRP	70,89%
ICX	70,90%	XTZ	69,81%
IOST	67,61%	XVG	70,08%
LSK	70,08%	ZEC	71,43%
LTC	70,70%	ZIL	69,64%
MAID	69,81%	ZRX	69,47%

### 4.3 Metrics for Analysis and Evaluation

This section introduces the metrics used in this study, categorized into 3 species, namely correlation metrics, graph-based metrics and symbolic metrics.

#### 4.3.1 Correlation Metrics

To evaluate the correlation between two assets, we employ 3 different metrics, including Pearson [160], Spearman [161] and Kendall’s Tau [162]. The three metrics share the same outcome in which the values fall within  $[-1, 1]$ , where  $-1$  indicates the complete difference between the two time series while  $1$  indicates the complete similarity, with respect to a financial indicator’s movements such as price, return and volatility. In other words, the higher the value is, the more similar the two time series are. However, the Pearson metric isolates itself from the other two by its perspective. Specifically, the Spearman and Kendall’s Tau metrics measure the

monotonic relationship between two time series by comparing the relative ranks of their elements. Two time series are said to have a perfect monotonic relationship if the ranks of corresponding elements in the two series are aligned, even if the actual values of the two series differ. In this regard, these two metrics assess the non-linear similarity between two time series. Pearson correlation, however, assesses the similarity based on the linearity between two time series. Thus, they are completely similar only when each corresponding pair of values between the two time series are completely similar. In this study, Pearson is set to be the main metric for measuring the correlation since it also covers the characteristics of non-linear metrics such as Spearman and Kendall’s Tau, while the latter is used for verification.

Given two time series  $x_i$  and  $x_j$  with length  $T$  of two assets  $i$  and  $j$ , respectively. The correlation between them as defined by each metric is provided in Table 4.2.

Table 4.2: **Correlation Metrics Used in this Study**

Metric	Formula
Pearson [160]	$\frac{\sum_{k=0}^{T-1} (x_k^i - \bar{x}_i)(x_k^j - \bar{x}_j)}{\sqrt{\sum_{k=0}^{T-1} (x_k^i - \bar{x}_i)^2 \sum_{k=0}^{T-1} (x_k^j - \bar{x}_j)^2}}$ <p style="text-align: center;">Where <math>\bar{x}_i, \bar{x}_j</math> are the average values of <math>x_i, x_j</math>, respectively.</p>
Spearman [161]	$1 - \frac{6 \sum_{k=0}^{T-1} (\text{rank}(x_k^i) - \text{rank}(x_k^j))^2}{T(T^2 - 1)}$ <p style="text-align: center;">Where <math>\text{rank}(x_k^i)</math> is the magnitude rank of <math>x_k^i</math> out of all observations of <math>x_i</math>.</p>
Kendall’s Tau [162]	$\frac{C - D}{T(T - 1)/2}$ <p style="text-align: center;">Where <math>C</math> is the number of concordant pairs and <math>D</math> is the number of discordant pairs.</p>

### 4.3.2 Graph-based Metrics

As mentioned before, a correlation graph informs the correlations among a group of assets during a certain period of time where each node represents an asset and each edge connecting two nodes represents how correlated the two corresponding assets are. To extract information from such a graph, we adopt several graph-based metrics, grouped into three categories:

1. Metrics for graph evolution assessment: These metrics allow us to observe the structural change of a graph over time.
2. Metrics for graph comparison: These metrics are used to compare structures of different graphs in order to highlight similarities and differences between their structures.
3. Metrics for node assessment: These metrics are used to observe the influence of each node on the whole graph to identify which nodes have the weakest/strongest influence on the graph.

#### Metrics for Graph Evolution Assessment

To assess the structural change of a graph over time, we use 5 different metrics, as shown in Tables 4.3-4.4.

#### Metrics for Graph Comparison

To make a comparison between two graphs' structures, we utilize 2 metrics, as shown in Table 4.5.

---

<sup>4</sup>The term "compressed graph" is used to describe a graph where nodes are close to each other, forming large, interconnected communities (usually 1 or 2 communities), rather than being distributed into multiple smaller communities.

<sup>5</sup>A scattered graph is the opposite of a compressed graph. In a scattered graph, the nodes are more distant from each other, forming several smaller, distinct communities.

Table 4.3: **Graph-based Metrics Used in this Study**

Metric	Formula and Description
Degree Assortativity [163]	$\frac{\frac{1}{m} \sum_{(i,j) \in E} k_i k_j - \left\{ \frac{1}{m} \sum_{(i,j) \in E} \frac{1}{2} [k_i + k_j] \right\}^2}{\frac{1}{m} \sum_{(i,j) \in E} \frac{1}{2} [k_i^2 + k_j^2] - \left\{ \frac{1}{m} \sum_{(i,j) \in E} \frac{1}{2} [k_i + k_j] \right\}^2}$ <p>Where <math>E</math> represents a set of edges in the graph, <math>k_i</math> and <math>k_j</math> are the degrees of node <math>i</math> and <math>j</math> of the edge <math>(i, j) \in E</math>, and <math>m</math> is the number of edges in the graph.</p> <p>Degree assortativity measures the similarity between node degrees in a graph. Specifically, it measures the tendency of nodes with a similar number of connections (degree) to link to one another. A graph with high assortativity coefficient has high-degree nodes connecting to other high-degree nodes, and low-degree nodes tend to link with other low-degree nodes. Conversely, a graph with low assortativity coefficient has high-degree nodes connecting to low-degree nodes. The assortativity coefficient ranges from -1 (disassortative) to 1 (assortative). A small coefficient indicates a more compressed graph<sup>4</sup>, while a large coefficient suggests a more scattered graph<sup>5</sup>.</p>
Betweenness Centrality [164]	$\sum_{i \neq v \neq j} \frac{\sigma_{ij}(v)}{\sigma_{ij}}$ <p>Where <math>\sigma_{ij}</math> is total the number of shortest paths from node <math>i</math> to node <math>j</math> and <math>\sigma_{ij}(v)</math> is the number of those paths that pass through <math>v</math> (not where <math>v</math> is an end point).</p> <p>Betweenness centrality measures a node's importance in connecting others. It counts how many shortest paths pass through a node. A node with high betweenness centrality means that it lies on many shortest paths, indicating influence in information or resource flow. A high average betweenness centrality implies a more scattered graph with small communities, while a low average suggests a more compressed graph forming large communities.</p>

### Metrics for Node Assessment

To calculate the impact of each node on the whole graph, we use the so-called metric *PageRank* - the same metric used by Google to rank web pages in search results [170]. The impact magnitude of a node is assessed based on the number of

Table 4.4: Graph-based Metrics Used in this Study (Continue)

Metric	Formula and Description
Closeness Centrality [165]	$\frac{N-1}{\sum_{j \in V, j \neq i} d(j, i)}$ <p>Where <math>N</math> is the number of nodes, <math>V</math> is a set of vertices in the graph and <math>d(j, i)</math> is the length of the shortest path from node <math>j</math> to node <math>i</math>.</p> <p>Closeness centrality measures the distance from a node to all other nodes in the graph. A node with a high closeness centrality is one that is close to many other nodes. Thus, it has a higher influence on the graph. This metric is used to identify the central node (the node that directly connects to many other nodes) in the graph. To use this metric, we average the closeness centrality of all nodes. Eventually, a high value suggests a more compressed graph, while a low value suggests a more scattered graph.</p>
Residuality Coefficient [166]	$\frac{\sum_{[i \neq j, e(i, j) > L]} e(i, j)^{-1}}{\sum_{[i \neq j, e(i, j) \leq L]} e(i, j)^{-1}}$ <p>Where <math>e(i, j)</math> is the edge weight between node <math>i</math> and node <math>j</math>, <math>L</math> is a threshold chosen manually. In this study, <math>L</math> is the highest edge weight ensuring the connectivity of the so-called Minimum Spanning Tree, which will be introduced in Section 5.5.5.</p> <p>Residuality coefficient compares the relative strength of the connections above and below a threshold value. An increase in this metric means that nodes are further from each other. By contrast, nodes are closer to each other if this metric decreases.</p>
Mean Weight [167]	$\frac{\sum_{(i, j) \in E} e(i, j)}{N-1}$ <p>Where <math>N</math> is the number of nodes, <math>E</math> is a set of edges in the graph and <math>e(i, j)</math> is the edge weight node <math>i</math> and node <math>j</math>.</p> <p>Mean weight calculates the average edge weight of a graph. An increase in this metric means that nodes are further from each other. By contrast, nodes are closer to each other if this metric decreases.</p>

Table 4.5: Metrics for Graph Comparison Used in this Study

Metric	Formula
V-measure [168]	$\frac{2 \times h \times c}{h + c}$ <p>Where:</p> $h \text{ (homogeneity)} = 1 - \frac{H(A,B)}{H(A)}$ $c \text{ (completeness)} = 1 - \frac{H(B,A)}{H(B)}$ $H(A, B) = - \sum_{k=1}^{ A } \sum_{c=1}^{ B } \frac{a_{ck}}{N} \log \frac{a_{ck}}{\sum_{c=1}^{ B } a_{ck}}$ $H(B) = - \sum_{c=1}^{ B } \frac{\sum_{k=1}^{ A } a_{ck}}{n} \log \frac{\sum_{k=1}^{ A } a_{ck}}{n}$ $H(B, A) = - \sum_{c=1}^{ B } \sum_{k=1}^{ A } \frac{a_{ck}}{N} \log \frac{a_{ck}}{\sum_{k=1}^{ A } a_{ck}}$ $H(A) = - \sum_{k=1}^{ A } \frac{\sum_{c=1}^{ B } a_{ck}}{n} \log \frac{\sum_{c=1}^{ B } a_{ck}}{n}$ <p><math>A</math> and <math>B</math> contain the community index of each asset in the first and second graph, respectively. <math>a_{ck}</math> is the number of nodes that are an element of community <math>c</math> and also an element of community <math>k</math>. <math>N</math> is the number of nodes (same for two graphs). <math>n</math> is the number of communities in the first graph. The communities in each graph are found using the community detection algorithm which is represented in Section 4.7</p> <p>The values of <i>v-measure</i> range from 0 to 1, the higher the value is, the more similar two graphs are.</p>
Eigenvalue method [169]	$\sum_{i=1}^k (\lambda_{1i} - \lambda_{2i})^2$ <p>Where <math>\lambda_{1i}, \lambda_{2i}</math> are eigenvalues of the laplacians (<math>L_1, L_2</math>) of the first and second graphs, respectively. Here, <math>L_1 = D_1 - A_1</math>, <math>L_2 = D_2 - A_2</math>, with <math>A_1, A_2</math> are adjacency matrices<sup>6</sup>, <math>D_1, D_2</math> are diagonal matrices of degrees of the first and second graphs, respectively. <math>k</math> is chosen such that the top <math>k</math> eigenvalues of both laplacians contain at least 90% of the information in the graphs (e.g. explain at least 90% of total variance in the graphs).</p> <p>The range of values given by this metric is unbounded on the positive side <math>[0, \infty)</math> and values close to 0 indicate a strong similarity between two graphs while high values show dissimilarity.</p>

links pointing to it. Specifically, the PageRank value  $PR(i)$  of a node  $i$  is equal to

$$(1 - d) + d \times \sum_{j \in C_i} \frac{PR(j)}{|m_j|}.$$

Where  $d$  is a damping factor with its values between 0 and 1. In this study, it is set to be 0.85, which is a common choice in the existing literature [171].  $C_i$  represents the set of nodes that are linked to node  $i$ .  $|m_i|$  refers to the number of outbound links from node  $j$ .

PageRank values range from 0 to 1. Nodes with higher values are considered more influential, and thus transfer their influence to a wider range of nodes in the graph.

Since this metric is only used for a certain period, to summarise the overall PageRank result for several continual periods, we define two PageRank-based metrics, as shown in Table 4.6.

Table 4.6: **PageRank-based Metrics for Node Assessment**

Metric	Formula
Total Accumulation	$\sum_{k=1}^T pr_k$ <p>Where <math>T</math> is the number of periods and <math>pr_k</math> is the PageRank value at the <math>k^{th}</math> period of a node.</p> <p>The value given from this metric is always equal to or greater than 0. The higher the value is, the more influential the corresponding node is.</p>
Dominance Score	$\sum_{k=1}^T hpr_k$ <p><math>hpr_k = 1</math> if its corresponding <math>pr_k</math> is the highest out of all nodes' PageRank values at the <math>k^{th}</math> period, otherwise 0.</p> <p>Where <math>T</math> is the number of periods. The value given from this metric is always equal to or greater than 0. The higher the value is, the more influential the corresponding node is.</p>

### 4.3.3 Symbolic Metrics

Instead of using real values, another type of evaluation relies on symbols, such as the one used in our study called *Symbolic Aggregate Approximation* (SAX). In particular,

this technique is designed to find similar patterns of behavior within a single time series even when the exact magnitudes of values differ due to factors like trend and seasonality [172]. For this, its key aim is to transform the original real values of the time series into a set of symbols before starting an assessment or evaluation. In addition, the effect of SAX is to reduce the dimensionality of the time series.

To implement SAX, we begin by splitting the time series's original values into different segments depending on our problem. Subsequently, each segment of values is assigned to a symbol (i.e. a, b, c, etc). As a result, from the original time series of real values, we obtain a time series of symbols, where different real values belonging to the same segment are assigned to the same symbol.

For this study, SAX is used to compare similarities and differences in movements between different time series. Thus, we calculate the percentage of occurrence for each symbol in each time series and then make a comparison between different time series based on these percentages.

## 4.4 Correlation Matrix Formation and Random Matrix Theory

Given a set of  $N$  time series  $z_k$ ,  $k = 1, 2, \dots, N$ , the correlation between each pair of time series  $C_{i,j}$  is obtained using the Pearson correlation metric, as described in Section 4.3.1. Subsequently, the correlation matrix  $C$  is formed from correlation coefficients obtained above such that  $C = (C_{ij})_{1 \leq i \leq N, 1 \leq j \leq N}$ .

One issue raised from this type of matrix is the question of how reliable these correlations are, in other words, whether the correlation matrix shows genuine and authentic relationships between the considered time series or it simply contains random information which is usually referred to as noise [173]. To this end, our experimental results would be invalid if the correlation matrices only contained random information. Therefore, it is essential to examine the reliability of the correlation matrices before delving into further experiments. Thanks to the findings

associated with Random Matrix Theory (RMT) [174], this hypothesis can be examined. Specifically, given a  $m \times n$  random matrix  $\mathbf{N}$  whose elements are distributed randomly with zero mean and unit variance, the eigenvalue distribution of the correlation matrix  $\mathbf{C}_\mathbf{N} = \frac{1}{m}\mathbf{N}\mathbf{N}^\top$  follows the Marchenko–Pastur probability density function [175] if the Quality Factor  $Q = \frac{m}{n} \geq 1$  holds when the number of timestamps  $m \rightarrow \infty$  and the number of features  $n \rightarrow \infty$ :  $P(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}$ , where  $P$  is the Marchenko–Pastur probability density function,  $\lambda$  is an eigenvalue of  $\mathbf{C}_\mathbf{N}$ ,  $\lambda_\pm = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}$  are upper and lower limits, respectively.

From RMT, eigenvalues falling outside of  $[\lambda_-, \lambda_+]$  are assumed to deviate from its expected predictions and potentially contain essential information [176, 177]. Hence, we can use this theory to test the reliability of the relationships in our empirical data [173]. That is, if an empirical correlation matrix actually has real valuable information, it must have eigenvalues that are outside the bounds of  $[\lambda_-, \lambda_+]$ . Otherwise, the empirical correlation matrix can be taken to contain mainly random noise. In this study, RMT has been used to test our correlation matrices in all experiments. The results show that all correlation matrices are not random and contain valuable information.

## 4.5 Notion of Graphs under Financial Markets Context

Given a correlation matrix  $C$ , the corresponding distance matrix  $D$  is obtained where each element  $d_{ij} = \sqrt{2(1 - c_{ij})}$  is the distance between node  $i$  and  $j$ . This transformation is necessary since the correlation coefficient lacks characteristics of a metric [178], therefore connections learned from a correlation matrix lack topological characteristics that prevent the comparison of the distance between nodes. Here,  $d_{ij}$  ranges from 0 to 2 with 0 meaning that two nodes are in the same position, which indicates the complete similarity in movements (e.g. price, returns and volatility) between them, while 2 is the furthest distance between two nodes, which indicates

the complete difference in movements between them. Following this, a graph in this study is constructed from the aforementioned distance matrix  $D$  where each node represents a financial asset and an edge connecting two nodes appears for their correlation represented by their distance. Therefore, assets with similar behavior (e.g. similar volatility/prices/returns' movements) have small distance values between them while assets that are more different from each other have greater distance values [179].

One downside of the graph is that it is dense, making it memory-intensive and time-consuming for processing because all nodes are linked together while there are edges with negligible information [180]. As a result, a considerable amount of computing time is required to process the graph and redundant information might confuse experimental results. Thus, we extract a subset of the graph called *Minimum Spanning Tree* (MST)[181] to reduce the graph size and remove the redundant information. This subset keeps edges with the highest correlation degree, guarantees the characteristics of a tree and removes other edges. Hence, only the important part of a graph is kept. As a result, this technique reduces a graph size of  $N$  nodes and  $\frac{N*(N-1)}{2}$  edges to  $N$  nodes and  $N - 1$  edges. This MST structure will be used for all experiments in this study. Therefore, from this section onwards, a graph refers to the MST structure, not the original dense version.

## 4.6 Noise and Trend in the Financial Markets: A Removal Scheme

### 4.6.1 Introduction of Noise and Trend

Noise has been shown to exist in both traditional and cryptocurrency markets with the latter being undoubtedly more apparent, according to [6]. In particular, the authors reveal that the average daily signal-to-noise ratio for the cryptocurrency market is only 36%, which is significantly lower than the 90% average daily signal-

to-noise ratio for established US stock exchanges like NYSE and NASDAQ during their examined period from March 2017 to November 2017.

One common factor for such effect in financial markets comes from noise traders, who tend to be unprofessional with restricted financial knowledge, mimicking decisions of professional and popular traders while ignoring their own perspectives as well as being easily seduced by apparently relevant news [182, 183]. Consequently, these noise traders cause fluctuations in the markets. Additionally, systemic risks such as newly enacted regulations and changes in financial policies are also potential influencers for markets' fluctuations [184, 185]. It is shown that both cryptocurrencies and traditional assets react to such news after only a few minutes of the announcements. Furthermore, this effect can last from several hours to even several weeks [184, 186]. When it comes to cryptocurrencies, the noise effect is also attributed to specific sources such as *Pump and Dump* schemes<sup>7</sup>, which have become a typical characteristic that distinguishes digital assets from traditional ones [187]. Besides, this young market is also famous for the dominance of noise (naïve) investors, compared to other, more traditional ones like stocks or currencies [111, 188], which explains its high percentage of noise as mentioned earlier.

Regarding the trend effect, this phenomenon refers to the co-movements in prices across different assets. Consequently, the underlying characteristics of each asset are masked by the existence of trend. One of the causes is the manipulation of social media and breaking news, leading to a so-called action *herding behavior* among investors [189] where the majority of investors invest in the same assets, causing similar fluctuations between them. Another cause lies in the fundamental characteristics of the assets that inadvertently cause a trend among them. For instance, the co-movements in cryptocurrencies can be driven by the mutual blockchain protocols that they share. That is, as of 2024, there are nearly 10,000 active cryptocurrencies

---

<sup>7</sup>A *Pump and Dump* scheme is a form of investment fraud in which the price of an asset is artificially inflated (pumped) by an individual/institution on purpose or false/misleading/exaggerated information. Once the price rises significantly due to increased demand, people start selling off their holdings (dump), causing the price to plummet. As a result, unsuspecting investors who bought in during the pump phase suffer heavy losses, while the frauders profit.

as reported by a famous digital coins-tracking website Coinmarketcap. Whereas, only around 10 blockchain protocols are used to govern those coins [190]. Consequently, the same blockchain protocol is used by many cryptocurrencies. Regarding traditional assets such as stocks and ETFs, these are heavily subject to macro-economic forces since investors in traditional markets have a tendency to allocate their investment portfolios based on the changes in macro-economic factors, which would move assets in the same way, especially during the economic uncertainty [109].

While using original values with the effect of noise and trend is common in the existing literature and proves to obtain meaningful findings, we expect that observing the markets after removing such noise and trend effects is worthwhile and might bring up different angles that are hidden behind these effects. Therefore, apart from doing experiments with the original dataset, we implement a deeper investigation by also running those experiments with the exclusion of noise and trend effects.

### 4.6.2 Denoising in Finance

The basic idea is that the noise in assets leads to the noise in the corresponding correlation matrix between the assets, which is the foundation element of all experiments in this study. Hence, to overcome the noisy issue in the assets, we directly remove the noise from the correlation matrix. For this, we first re-use Random Matrix Theory (RMT) discussed in Section 4.4 to identify the noise in the correlation matrix. We recall that from this theory, an eigenvalue  $\lambda$  obtained from our empirical correlation matrix  $\mathbf{C}$  is considered to exhibit random behavior if it belongs to the noise regime, i.e.  $\lambda \in [\lambda_-, \lambda_+]$ . This eigenvalue typically represents noise and does not carry meaningful structural information about the underlying data. Conversely, those eigenvalues that fall outside of  $[\lambda_-, \lambda_+]$  are assumed to contain informative signals. However, in practice, existing literature only focuses on eigenvalues that are greater than  $\lambda_+$  because they are associated with principal components that explain most variance in the data [191, 192]. Eigenvalues below  $\lambda_-$ , although outside the

noise regime, are typically small and thereby capture negligible variance (e.g. they are not noise but they contain negligible information) [193]. Therefore, we follow the common practice in the literature by also removing these eigenvalues [191, 192]. To this end, our target in the denoising task is to remove the eigenvalues less than or equal to  $\lambda_+$  from the empirical correlation matrix  $\mathbf{C}$ .

*Eigenvector Clipping* [192] is our choice of method for this noise removal. One advantage of this approach is that it does not require training parameters, which guarantees the reliability and robustness of its outcomes. This feature distinguishes Eigenvector Clipping from other methods such as Linear Shrinkage [194], Rotationally Invariant, Optimal Shrinkage [195] and Non-linear Shrinkage [196] which require users to specify multiple parameters, leading to a concern of choosing an appropriate parameter value. Additionally, Eigenvector Clipping is straightforward to implement, with guaranteed efficiency as it keeps the trace of the correlation matrix remains unchanged after the cleaning process [197]. Preserving the trace of a matrix is crucial since it maintains consistency, interpretability, and the fundamental properties of the matrix. This method has shown good performance in different studies and has been applied widely to different topics such as education, portfolio optimization and signal processing [198–201].

Given eigenvalues  $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_n$  and corresponding eigenvectors  $v_1, v_2, \dots, v_n$  of our empirical correlation matrix  $\mathbf{C}$ , we can identify  $k \leq n$  such that  $\lambda_k > \lambda_+$  and  $\lambda_{k+1} \leq \lambda_+$ . As a result,  $\lambda_i, \forall i \leq k$  are noisy eigenvalues, according to Random Matrix Theory. The Eigenvector Clipping replaces these noisy eigenvectors with their average value in which the trace of the correlation matrix  $\mathbf{C}$  after denoising is similar to its origin. The denoised correlation matrix  $\mathbf{C}_{denoised}$  is defined by [202] to be:

$$\mathbf{C}_{denoised} = \sum_{i=1}^n \lambda_i^* v_i v_i^T, \lambda_i^* = \begin{cases} \frac{\lambda_{k+1} + \lambda_{k+2} + \dots + \lambda_n}{n-k}, \forall i \geq k+1 \\ \lambda_i, \forall i \leq k \end{cases} \quad (4.3)$$

Figure 4.1 displays an example of the eigenvalues' density distribution (shown

in blue) obtained from our empirical correlation matrix, which is constructed by 34 cryptocurrencies on the HitBTC exchange during the duration from 2019-07-02 to 2019-07-24 and the corresponding Marchenko-Pastur distribution (shown in red). As can be seen, most eigenvalues fall below  $\lambda_+$  (green vertical line). On the other hand, a few eigenvalues are greater than the threshold, carrying important information. We would like to retain these informative eigenvalues while filtering the impact of the noisy eigenvalues (within the Marchenko-Pastur distribution) as well as the ones carrying negligible information (below  $\lambda_-$ ) from the correlation matrix. The sub-figure in the top right corner is obtained by removing the largest eigenvalue in order to help readers compare the Marchenko-Pastur distribution with empirical eigenvalues more easily.

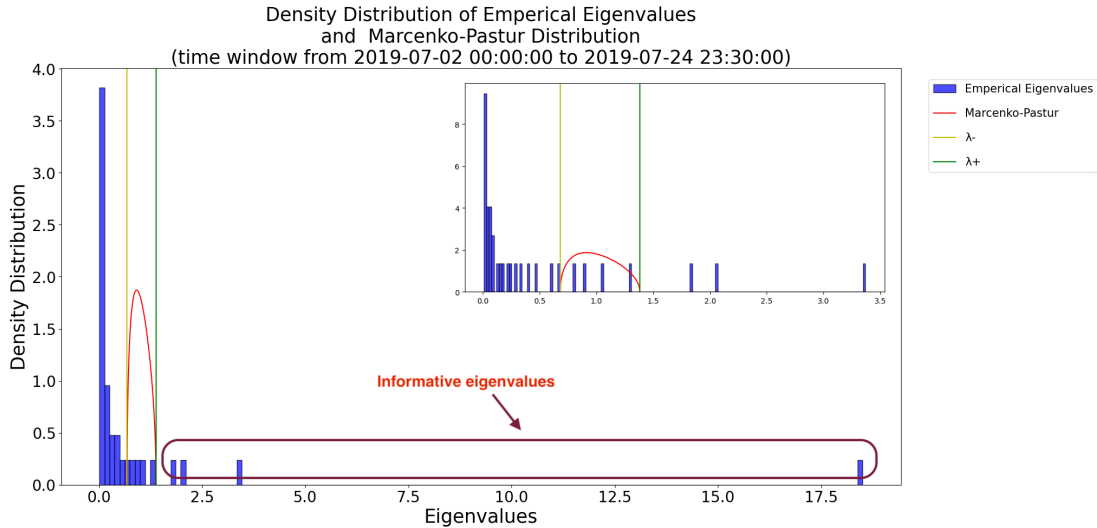


Figure 4.1: **Density distribution of eigenvalues (blue) obtained from our empirical correlation matrix and the corresponding Marchenko-Pastur distribution (red). The yellow and green vertical lines represent the minimum ( $\lambda_-$ ) and maximum ( $\lambda_+$ ) expected eigenvalues of the Marchenko-Pastur distribution, respectively. The top right sub-figure is obtained by removing the largest eigenvalue.**

Regarding the noise removal, we note that there are two main types of noise, including inherent noise and measurement noise [192]. While the inherent noise refers to the natural randomness within the system, the measurement noise tends to be arise from inaccuracies in the data collection or recording process such as price rounding, sampling errors, and missing data points. Our denoising approach,

which relies on RMT, addresses both types of noise but does not explicitly separate them. Moreover, one downside of RMT is that it is sensitive to the length of time series. Specifically, increasing the length of the time series for the same dataset causes  $\lambda_-$  and  $\lambda_+$  to become closer to each other. Eventually, when the length is finite and significantly large, the possibility for the eigenvalues to fall within the Marcenko-Pastur distribution is zero, RMT thereby fails to identify the noisy eigenvalues. On the other hand, if the length of time series is not very large compared to the number of assets, the empirical correlation matrix might be dominated by measurement noise. Consequently, there might be collective effects in the system which are hidden and masked by the measurement noise [203]. These are limitations that will be considered in future work.

### 4.6.3 Detrending in Finance

Implementing the detrending process on a Financial-correlation matrix between assets is based on the concept of *market component* [177], which refers to the first eigenvalue  $\lambda_1$  (the largest eigenvalue) and the corresponding eigenvector  $v_1$  of a correlation matrix. Such a component represents the common systematic risks or common investor responses to them that impact most assets in a market and also provides insights into the overall relationship between individual assets and the market as a whole. Removing the "market component" from a correlation matrix highlights the unique relationships between individual assets, facilitating the understanding of their specific dependencies and individual movements that might be hidden by the trend effect.

Given this concept, we detrend our empirical correlation matrix by removing the first eigenvalue and eigenvector. To this end, the denoised and detrended correlation matrix  $\mathbf{C}_{\text{cleaned}}$  is obtained by subtracting the first eigenvalue and eigenvector from the denoised correlation matrix  $\mathbf{C}_{\text{denoised}}$ .

$$\mathbf{C}_{\text{cleaned}} = \mathbf{C}_{\text{denoised}} - \lambda_1 v_1 v_1^T \quad (4.4)$$

Figure 4.2 displays an example of the market component in our empirical correlation matrix constructed by 34 cryptocurrencies on the HitBTC exchange between 2019-07-02 and 2019-07-24. Observably, there is a blue bar standing significantly apart from the others which represents the market component. We note that the market component is always the largest eigenvalue of a correlation matrix.

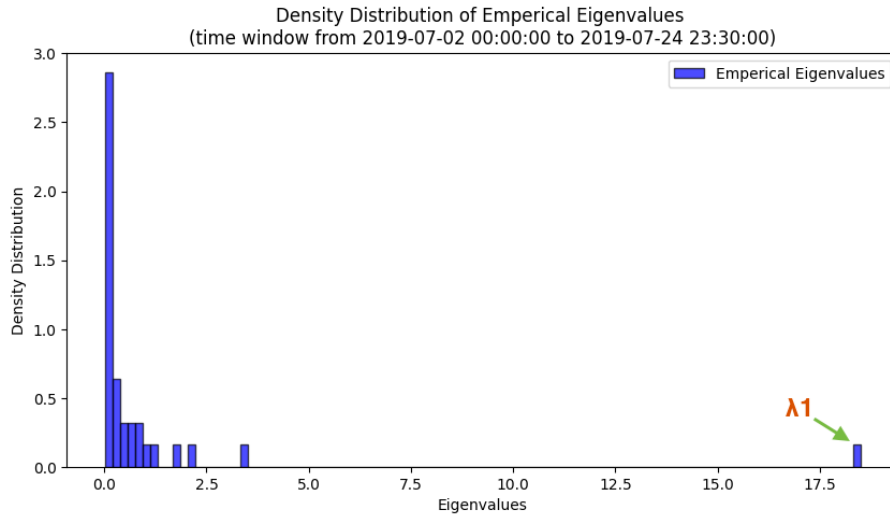
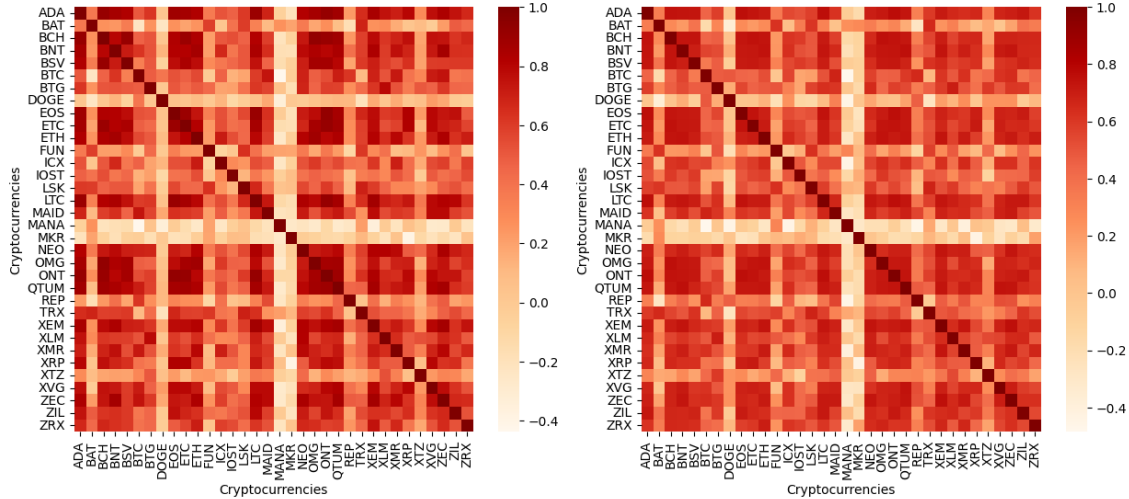


Figure 4.2: **Market component from our correlation matrix.**

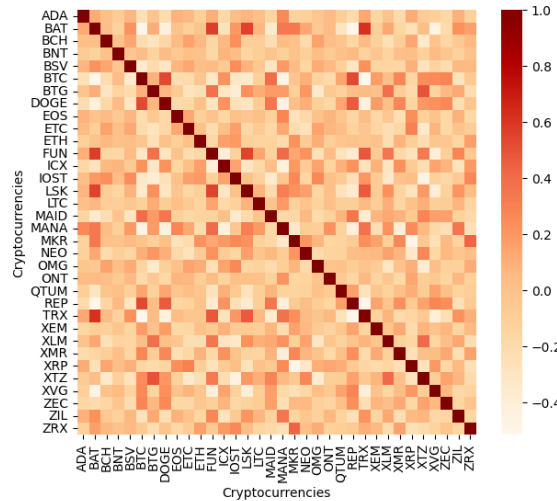
Figure 4.3 displays examples of our empirical correlation matrix mentioned above (a) before denoising and detrending, (b) after denoising and (c) after denoising and detrending, respectively. The intensity of the red color (from white to dark red) represents the correlation between two assets, with darker red indicating a higher correlation. The maximum value of the correlation is 1 which represents a complete similarity between two assets. By contrast, the minimum value of -1 represents a complete difference between two assets. The correlation values in the diagonal line of each correlation matrix are all 1 because these values are derived from calculating the correlation between 2 identical assets. Looking at the original (Figure 4.3a) and denoised correlation matrix (Figure 4.3b), we can infer that the noise actually exists in the assets since the correlation coefficients from the denoised correlation matrix in Figure 4.3b tend to be lower than those in the original correlation matrix, as shown by a brighter red color. In other words, the existence of noise results in the correlation between assets being higher than it should be. Regarding Figure 4.3c,

this correlation matrix looks completely different from the first two cases since it is mostly occupied by low correlation values. It is clear that the correlation between assets weakens significantly as a result of removing the trend effect.



(a) Original correlation matrix

(b) Denoised correlation matrix



(c) Denoised-Detrended correlation matrix

Figure 4.3: An example of cleaning noise and trend effects in a correlation matrix. a) Original correlation matrix with noise and trend, b) correlation matrix after denoising, c) correlation matrix after denoising and detrending.

To support our claims above, we perform a Wilcoxon Signed-Rank test [204] to examine whether the correlation coefficients between cryptocurrencies statistically tend to decrease after removing noise and trend effects. This non-parametric statistical test is specifically designed for paired-sample data, allowing us to draw conclusions about the median of the differences between two samples. For this test,

our null hypothesis is that the differences in correlation coefficients between two correlation matrices are symmetric about a median of zero. By contrast, our alternative hypothesis is that the differences in correlation coefficients between two correlation matrices are symmetric about a median greater than zero. To this end, if the test accepts the alternative hypothesis, it indicates that over 50% of the correlation coefficients in the second correlation matrix are lower than those in the first. Consequently, this would suggest that correlations between cryptocurrencies tend to decrease after removing noise and trend effects. We conduct two Wilcoxon Signed-Rank tests: 1) between the original correlation matrix and the denoised correlation matrix, and 2) between the denoised correlation matrix and the cleaned correlation matrix. Both tests accept the alternative hypothesis, confirming that the correlation coefficients indeed decrease after removing the noise and decrease even further after removing both noise and trend (the median of differences between the original and denoised correlation matrix is 0.05, the median of differences between the denoised and cleaned correlation matrix is 0.52. In both tests, p value is 0.0).

## 4.7 Louvain Community Detection

Community detection aims to partition nodes of a graph into separate groups (called communities) by maximizing the correlation between nodes within a community while maximizing the distinction between different communities. In other words, this detection results in communities where nodes belonging to a community are most similar to each other while nodes from different communities exhibit different characteristics, in terms of the movements of a financial indicator such as prices, returns and volatility. For this technique, the Louvain community detection algorithm [205] is quite straightforward to implement and is also adopted widely in various research topics [200, 206, 207].

The Louvain algorithm is based on modularity optimization whose fundamental principle is searching for a community partition such that it maximizes the density of edges inside communities in comparison with edges between different communities.

This optimization is executed via a measure called *modularity*  $Q$ :

$$Q = \frac{1}{2m} \sum_{i,j} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \gamma(g_i, g_j) \quad (4.5)$$

Where  $A$  represents the adjacency matrix obtained from a graph,  $A_{ij} = 1$  if there is an edge between node  $i$  and node  $j$  and 0 otherwise,  $m = \frac{1}{2} \sum_{i,j} A_{ij}$  is the total number of edges in the graph,  $k_i = \sum_j A_{ij}$  is the node degree of node  $i$ ,  $g_i$  represents the community which node  $i$  belongs to,  $\gamma(g_i, g_j) = 1$  if  $g_i = g_j$  (nodes  $i$  and  $j$  belong to the same community) and 0 otherwise. To this end, the best community partition is obtained when  $Q$  is maximum.

From this idea, the process of Louvain community detection comprises two phases. For the first phase, each node in the graph is initially assigned to a community. As a result, the number of communities at the beginning is as many as the number of nodes. Subsequently, for each node  $i$ , a set of its neighbors can be identified, we take turns calculating the *modularity gain*  $\Delta Q$  when removing  $i$  from its community and placing it in the community of its neighbor. After all neighbors are considered,  $i$  is placed in the community for which this  $\Delta Q$  is positive and maximum. If the gain is not positive, then  $i$  stays in its current community. This process is completed sequentially for all nodes and the whole process is repeated until no further improvement can be achieved, i.e. the  $\Delta Q$  is not positive for all cases and the modularity value  $Q$  reaches its maximum. The formula for the *modularity gain* by placing a node  $i$  into a community of its neighbor  $g_j$  is described as follows:

$$\Delta Q = \left[ \frac{\sum_{in} + 2k_{i,in}}{2m} - \left( \frac{\sum_{out} + k_i}{2m} \right)^2 \right] - \left[ \frac{\sum_{in}}{2m} - \left( \frac{\sum_{out}}{2m} \right)^2 - \left( \frac{k_i}{2m} \right)^2 \right] \quad (4.6)$$

Where  $\sum_{in}$  is the number of edges between nodes inside the community  $g_j$ ,  $\sum_{out}$  is the number of edges linked to nodes inside  $g_j$  (including edges from nodes within other communities),  $k_{i,in}$  is the number of edges from node  $i$  to nodes in  $g_j$ .

The second phase has a similar process to the first phase. However, instead of

working on each node, we consider each community found in the first phase as a node. As a result, the weight of an edge between two new nodes is the number of edges between nodes in the two corresponding communities. After completing this second phase, we obtain the temporary optimal partition. The resulting communities from this second phase will then become the original communities and the two-phase process is reapplied to them. This two-phase algorithm stops when no more changes exist and the *modularity*  $Q$  reaches its global maximum.

Figure 4.4 illustrates an example of a graph and its community detection result derived from the Louvain algorithm. As we mentioned, this graph is a Minimum Spanning Tree (MST) - the form we use for our experiments in this study. Each node in the graph represents an asset. After applying the community detection algorithm, separate communities are identified as shown in Figure 4.4b where each color represents a community and nodes with the same color belong to one community.

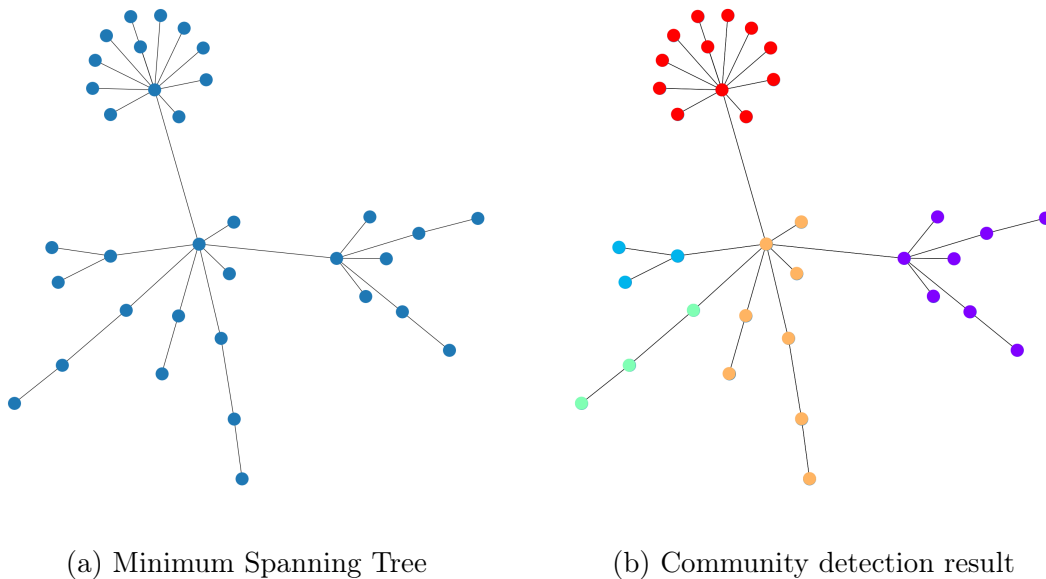


Figure 4.4: **Louvain community detection.** a) MST (Original graph), b) Community detection applied to the MST.

## 4.8 Herding detection

Cross-Sectional Absolute Deviation (CSAD) model is one of the most well-known herding behavior detection approaches that has been used widely in many studies [113, 125, 130, 132, 137]. The fundamental idea of this technique is to examine the relationship between the average return (market return) of a set of assets and the dispersion across each asset return (CSAD). A herding signal is pronounced when the market return changes significantly (i.e. the market experiences huge fluctuations, either going *up* or *down*) while the dispersion between asset returns decreases (i.e. the return movements of different assets become more similar). In this study, we use two forms of the CSAD model for our experiments.

$$CSAD_t = \beta_0 + \beta_1|r_{m,t}| + \beta_2r_{m,t}^2 + \epsilon_t \quad (4.7)$$

$$CSAD_t = \gamma_0 + \gamma_1D^{up}|r_{m,t}| + \gamma_2D^{up}r_{m,t}^2 + \gamma_3D^{down}|r_{m,t}| + \gamma_4D^{down}r_{m,t}^2 + \zeta_t \quad (4.8)$$

Where  $CSAD_t = \sum_{k=1}^T |r_{i,t} - r_{m,t}|/T$ ;  $r_{i,t}$  is the return of asset  $i$  at time  $t$ ;  $r_{m,t} = \sum_{k=1}^N r_{k,t}/N$  is the average of asset returns (also known as the market return) at time  $t$ ;  $\beta_0, \beta_1, \beta_2, \gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4$  are regression coefficients,  $D^{up} = 1$  if  $r_{m,t}$  is positive, otherwise 0;  $D^{down} = 1$  if  $r_{m,t}$  is negative, otherwise 0;  $\epsilon_t, \zeta_t$  are error terms at time  $t$ .

Regarding Equation 4.7, a herding exists when the parameter  $\beta_2$  is significantly negative (a negative coefficient with a confidence level of at least 90%). We expect this coefficient to be negative because it follows the idea of the herding behavior, i.e. a dramatic increase or decrease in the market return ( $r_{m,t}^2$ ) coincides with a decline (negative  $\beta_2$ ) in the dispersion between corresponding assets' returns ( $CSAD_t$ ). Equation 4.8 refers to a more detailed version of Equation 4.7 in which we separate the market into *up* and *down* so that the herding can be detected during the *up/down* market conditions only, instead of considering all market conditions as declared in the first form. The use of both forms for herding detection comes from an expectation that there might be periods when herding does not exist in the general markets but

it might be found in the *up* or *down* markets only and vice versa. To this end, the same logic of Equation 4.7 is applied to Equation 4.8 wherein a significantly negative  $\gamma_2$  refers to a herding signal during the *up* market condition while a significantly negative  $\gamma_3$  refers to a herding signal during the *down* market condition.

## 4.9 Remarks

This chapter describes all the methodologies as well as metrics for analysis and evaluation that are used in our research. In overall, these can be classified into two main categories, namely *graph-based methods* and *herding behavior detection methods*. Regarding the graph-based methods, our main techniques include:

- 1) *Pearson correlation measurement* for calculating the correlation between financial assets.
- 2) *Minimum Spanning Tree (MST)* for graph construction.
- 3) *Eigenvector Clipping* for removing noise from a graph.
- 4) *Market component removal* for removing trend from a graph.
- 5) *Louvain algorithm* for community detection.

Regarding herding behavior detection methods, we adopt two common models to detect the herding signal in a market, namely the Cross-Sectional Absolute Deviation (CSAD) model and its variant, which separates the market into *up* and *down* conditions so that the herding can be detected during the *up/down* market conditions only, instead of considering all market conditions as declared in the former.

The next chapter discusses our experiments and results where these techniques and metrics are mentioned frequently. Therefore, instead of reintroducing these techniques and metrics when using them in the following chapters, we will rather reference this chapter for details.

# Chapter 5

## Experimental Results and Discussion

### 5.1 A Note for this Chapter

This chapter discusses our experiments and corresponding results starting with the cryptocurrency market and then making comparisons between it and traditional markets, including stocks and US ETFs. Our study is constructed under three main bases:

- Firstly, focusing on analyzing the time-varying correlation between different cryptocurrencies and exploring how this correlation relates to investor behavior, especially during turbulent times.
- Since we find that there is a strong link between the changes in the correlation and the investment decisions of investors across different market conditions (i.e. crisis, bull, bear, stable, etc), the second basis is to detect and understand the herding behavior among investors in the cryptocurrency market during these and other conditions.
- Thirdly, we expand this study by investigating the herding behavior in the aforementioned traditional markets and comparing differences as well as similarities between the cryptocurrency market and traditional markets.

The structure for this chapter is as follows: Section 5.2 discusses the time-varying co-movement with respect to price returns of cryptocurrencies (RQ1). Section 5.3

also provides the co-movement analysis between these cryptocurrencies but this is based on their volatility, followed by a comparison between returns and volatility-based results (RQ2). Building on these co-movement experiments, we conduct another experiment that uses correlation information between these cryptocurrencies to assess how much influence each individual cryptocurrency has within the overall correlation graph, which is explained in Section 5.4 (RQ2). Lastly, detection outcomes of herding behavior in each of the cryptocurrencies, stocks and US ETFs are presented in Section 5.5. Furthermore, this section also proposes an expansion of herding behavior detection and analysis by carrying out experiments on a combination of all available types of investment vehicle (i.e. cryptocurrencies, stocks and US ETFs) (RQ3 and RQ4).

It is worth mentioning that we utilize 2 different datasets in this study: the first dataset comprising only cryptocurrencies is used for the first 3 experiments, which are equivalent to Sections 5.2-5.4. On the other hand, the second dataset comprising not only cryptocurrencies but also stocks and US ETFs is used for the remaining experiments, which are equivalent to Section 5.5.

## 5.2 Price Returns-Based Co-Movement in the Cryptocurrency Market

### 5.2.1 Introduction

We begin this study by examining the correlation between cryptocurrencies throughout different market conditions in terms of price returns. Subsequently, explaining experimental outcomes is done by referring to investors' behavior in investment decision-making.

Since cryptocurrencies' prices are collected tick-by-tick, there is no one fixed timescale for all cryptocurrencies leading to an inconsistency between the time series. For this reason, we re-sample the dataset by using data points at a specific

timescale. In particular, we aim to minimize the percentage of missing values as much as possible when upsampling the data. Eventually, 30-min is the lowest frequency that keeps all time series' missing values under 1%. Hence, we choose four different levels of granularity, namely 30 min, 6 h, 12 h and 24 h, constructing four corresponding datasets. Each data point of a dataset is taken to be the price of the last transaction of 34 cryptocurrencies within the considered granularity, such a method of upsampling is commonly adopted in the literature [208]. The selection of a suitable granularity level out of the aforementioned options for cryptocurrency time series in our experiments is also a focus in this section.

Additionally, the influence of noise and trend effects on the correlation between cryptocurrencies followed by a removal scheme are also discussed in this section. These pre-processing steps are useful for highlighting hidden characteristics in financial markets that might not be seen when using original and random-frequency data.

### **5.2.2 Timeline Division**

Given the cryptocurrency time series on the HitBTC exchange as described in 3.1.1, we partition the dataset into separate time intervals based on financial market conditions (i.e. stable, bullish, bearish, etc) in order to construct a correlation graph structure of cryptocurrencies for each financial market condition. We expect that a graph structure during a time interval can explain what has happened to the cryptocurrencies throughout that time. Consequently, each financial market condition should be associated with individual characteristics of the correlation graph structure. On the other hand, constructing the correlation graph by using the entire period or random time intervals might lead to ambiguous results. Indeed, before partitioning the dataset into several sub-periods, we have examined the correlation graph of cryptocurrencies, stocks, US ETFs and their combination over the entire period, the results reveal evidence of herding behavior in all cases. Therefore, no unique pattern is extracted from the results due to the lack of details.

Notably, our dataset covers the Covid-19 pandemic period followed by the global economic downturn in 2020 which adversely influenced the financial markets. As a result, we select this pandemic as a milestone to partition our dataset.

As can be seen from the public attention score towards the Covid-19 pandemic<sup>1</sup> in Figure 5.1a, people started to worry about this disease in January 2020. However, it was not until March 2020 that the Covid-19 pandemic actually caught the attention of people worldwide, since the number of searches quickly peaked. This remained a topic of interest until July 2020. Furthermore, March 2020 was the month in which a pandemic-induced economic recession first occurred, seriously affecting the economy of nations worldwide. This effect is shown in Figures 5.1b-5.1d. In particular, the world GDP's growth rate<sup>2</sup> decreased by 3.3% in 2020 (Figure 5.1b), which is the highest decrease ever, even worse than the Global Financial Crisis and its aftermath in 2007-2009 [209]. Simultaneously, the world's stock markets fluctuated dramatically, which can be seen via the VIX index<sup>3</sup> (Figure 5.1c) and the *S&P500* index<sup>4</sup> (Figure 5.1d), both of which experienced a significant fall during March 2020. However, the economy started to recover afterward, stock markets worldwide became less volatile and the *S&P500* index regained its original pre-pandemic value in July 2020.

Consequently, we split our data into three time windows which correspond to three different stages, including normal time, downturn time and recovery time.

The details for these time windows are shown in Table 5.1.

---

<sup>1</sup>The volume of searches for Covid-19-related terms (i.e. *Covid-19* and *coronavirus disease 19*) on Google, summarised by Google Trends: <https://trends.google.com/trends/>.

<sup>2</sup>The growth rate of the world's GDP is used as a proxy for the development of the global economy in general. This rate is collected from <https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG>.

<sup>3</sup>The VIX index represents the magnitude of fluctuations in the stock market, this index starts at 0 for no upper bound and a higher value implies that the stock market has stronger fluctuations. The daily values of this index are collected from <https://finance.yahoo.com/>.

<sup>4</sup>The *S&P500* index tracks the stock performance of the 500 largest companies listed on stock exchanges in the United States, which is often referred to as the performance of the US economy. The daily values of this index are collected from <https://finance.yahoo.com/>.

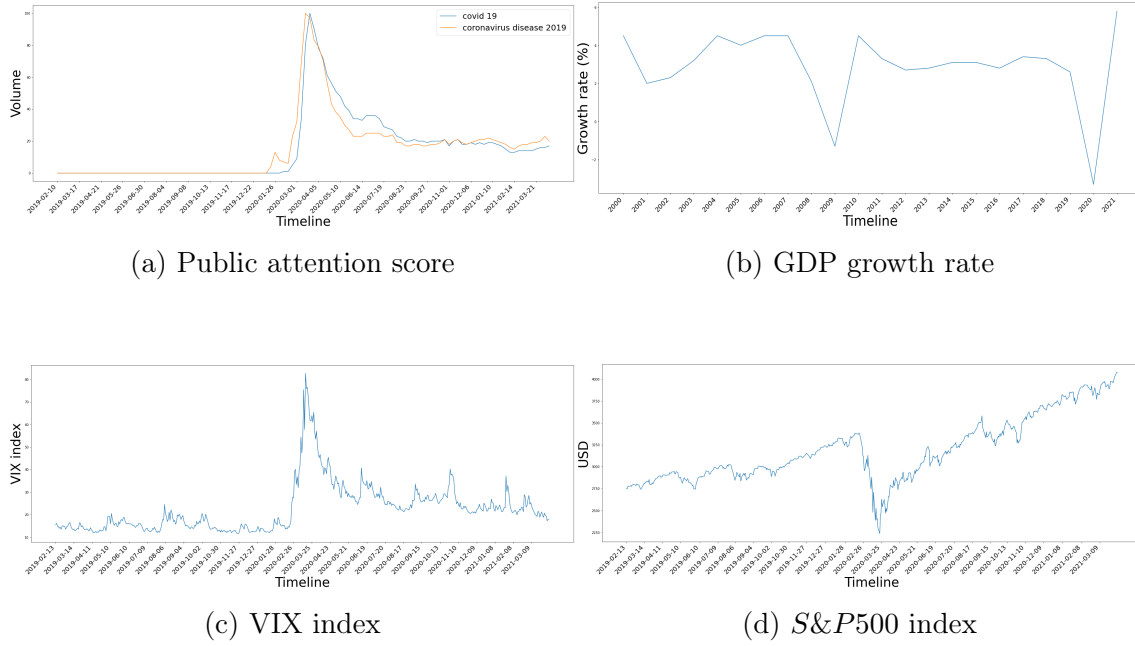


Figure 5.1: The reaction of the general public and global economy to the Covid-19 pandemic. Four factors are considered: (a) Worldwide attention to the pandemic, (b) Global GDP growth, (c) VIX index, (d) S&P500 index.

Table 5.1: Timeline partition for the HitBTC cryptocurrency dataset (time windows split to take into consideration the Covid-19 pandemic).

Time Window	Stage	Time Span	# Days
1	Normal time	13 February 2019 - 31 December 2019	322 days
2	Downturn time	1 January 2020 - 30 June 2020	182 days
3	Recovery time	1 July 2020 - 6 April 2021	280 days

### 5.2.3 Cryptocurrency Graph Structures under Noise and Trend Effects

Given the fact that there are noise and trend effects in the cryptocurrency market, we examine whether these factors affect the cryptocurrency graph structure. Since we have four datasets corresponding to four levels of granularity (e.g. 30 min, 6 h, 12 h and 24 h), we investigate the discrepancy between original and cleaned (after removing noise and trends) graph structures at each granularity by using two graph-based metrics, including *Residually Coefficient* and *Mean Weight* (see 4.4). There is no doubt that the original dataset is informative and valuable, which is proven by numerous findings proposed by previous studies [17, 25, 32], our expectation is that

looking at the market with the absence of noise and trends can enrich the current literature in this area.

We note that although these two metrics are used to examine the connection strength of cryptocurrencies, the *Residuality Coefficient* is known (see [210] for example) to be more vulnerable to the links between cryptocurrencies in different communities, i.e. if the connection strength between cryptocurrencies in different communities increases, the *Residuality Coefficient* will decrease dramatically, and vice versa; the connections between cryptocurrencies within one community do not affect the *Residuality Coefficient* to a great extent [210]. On the other hand, *Mean Weight* is known to be more vulnerable to the links between cryptocurrencies belonging to one community, as it mainly uses the connections within a community to find the average value and ignores the connections between different communities [167]. Thus, the purpose of each of these metrics is to compensate for one/another's shortcomings.

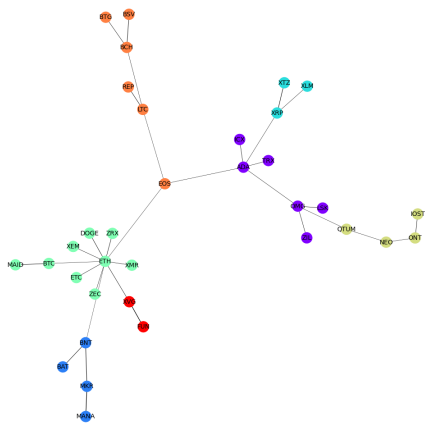
Table 5.2 displays the results of the two metrics before and after removing noise and trend effects from graph structures at four levels of granularity, during three time windows. It is clear that both *Residuality Coefficient* and *Mean Weight* values increase significantly when the effects of noise and trend are dismissed, such phenomenon being true for all granularities. Furthermore, an example of graph structures before and after cleaning is shown in Figure 5.2 to reinforce our finding. As can be seen, the correlation graph structure obtained from daily data changes after noise and trends removal. Moreover, generally, the number of communities in a graph tends to decrease after removing these effects. From these results, we can conclude that the connections between cryptocurrencies are caused mainly by the noise and trend effects. That is, these effects result in different cryptocurrencies becoming closer to each other and forming a group. This phenomenon can be explained by low values of *Residuality Coefficient* and *Mean Weight* in the original graph structures compared to the cleaned ones. Specifically, a value less than 1 of the prior metric means that there are few connections greater than the threshold  $L$ .

Moreover, a small value of the latter metric means that cryptocurrencies within a community are closer to each other. In summary, each community of the original graph structure is compact with strong links inside, which helps the community detection algorithm to easily cluster them. In other words, the difference between different communities is clear because the links between them are weak, i.e. the ones greater than  $L$ . However, after cleaning the correlation matrix, cryptocurrencies that are closely related to each other through noise and trend become further away, meaning that the strong links between some cryptocurrencies are broken. This causes our metrics to increase dramatically, which means that the graph structure starts to expand, forming a sparse graph. For example, the *Residuality Coefficient* of the 30 min original graph structure in the second time window is 0.28, while it is 20 times higher after cleaning the effects of noise and trends. This fact is also true for the rest of our dataset. The result is in line with [99]; these authors did not consider the noise effect but, with the removal of trends, they found that the correlation between the 80 most liquid cryptocurrencies from 1 January 2020 to 1 October 2021 decreased.

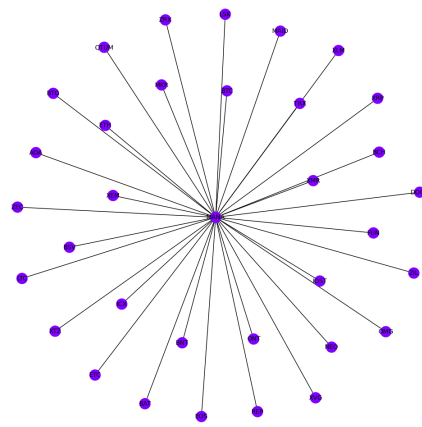
Table 5.2: **Connection strength measured by *Residuality Coefficient* and *Mean Weight*. Four different granularity levels are considered, each with corresponding original and cleaned graph structures.**

Metric	Data Type	Time Window	Granularity			
			30 min	6 h	12 h	24 h
<b>Residuality Coefficient</b>	Original Data	1	0.41	0.11	0.16	0.08
		2	0.28	0.111	0.06	0.05
		3	0.14	0.05	0.07	0.34
	Cleaned data	1	1.69	6.66	14.82	14.40
		2	5.98	8.90	14.41	15.34
		3	2.32	2.99	1.88	1.05
<b>Mean distance</b>	Original Data	1	1.08	0.82	0.80	0.76
		2	0.99	0.71	0.65	0.56
		3	0.98	0.57	0.46	0.45
	Cleaned data	1	1.29	1.38	1.42	1.42
		2	1.40	1.42	1.42	1.42
		3	1.29	1.12	1.01	1.22

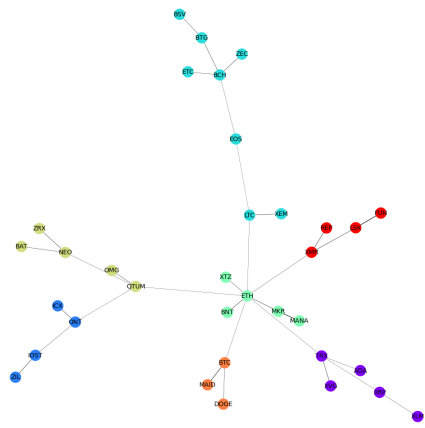
Through constructing graph structures with and without the existence of noise and trend effects, we recognize that the cleaned graph structures provide clear results



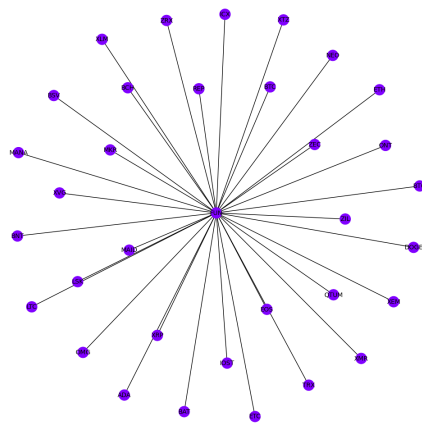
(a) Time window 1, original



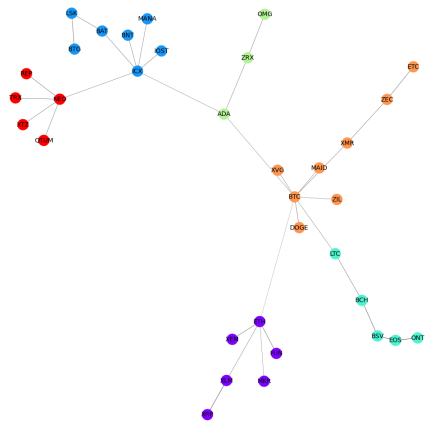
(b) Time window 1, cleaned



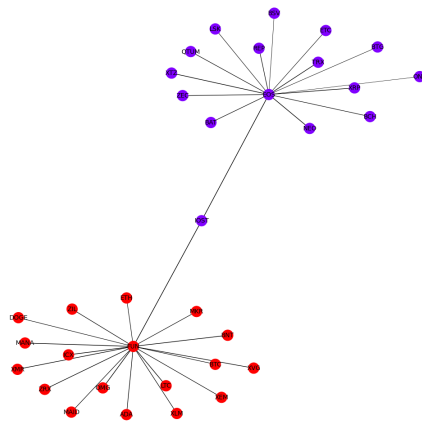
(c) Time window 2, original



(d) Time window 2, cleaned



(e) Time window 3, original



(f) Time window 3, cleaned

Figure 5.2: Cryptocurrency graph structures before and after denoising and detrending, using daily data. For each time window, Louvain algorithm is applied to both original and cleaned graph structures to detect existing communities.

that can be used for more in-depth analysis, while the original ones cause ambiguity as no firm patterns in the cryptocurrency market are detected. Consequently, only results after denoising and detrending are of use in this experiment for further analysis and discussion.

#### 5.2.4 The Evolution of graph Structures with Levels of Granularity

Our next step is to seek an appropriate granularity level that best suits our experiments. Figures 5.3–5.5 show the results of graph structures along with detected communities using the Louvain algorithm with each figure representing a different time window. For each window, four graph structures corresponding to four different levels of granularity are displayed. We found that the graph structures at each level of granularity tend to change over time, as shown by different values of *Betweenness centrality* and *Degree assortativity* across different time windows in Table 5.3. Additionally, if we consider different levels of granularity at the same time, the number of detected communities tends to decrease when the granularity becomes more coarse-grained, this phenomenon is displayed in Table 5.4. For low-grained granularity, such as 24 h, cryptocurrencies build up big communities with few cryptocurrencies acting as central nodes that link directly to the remainder. For example, in Figure 5.3d, MANA acts as a central node that links all other cryptocurrencies together. This explains why community detection techniques cannot distinguish several subsets as the graph in this case is naturally one community. Figure 5.4d shows a similar pattern, while in Figure 5.5d, there are two central nodes that create two big communities with relatively similar sizes.

The phenomenon of most cryptocurrencies belonging to one community at a high level of granularity such as daily can be explained by the nature of the cryptocurrency market. In particular, the cryptocurrency market is well-known for its high volatility compared to other traditional asset classes such as stocks, bonds and commodities [95, 211–213]. In [147], the authors used 5-minute data of Bitcoin prices

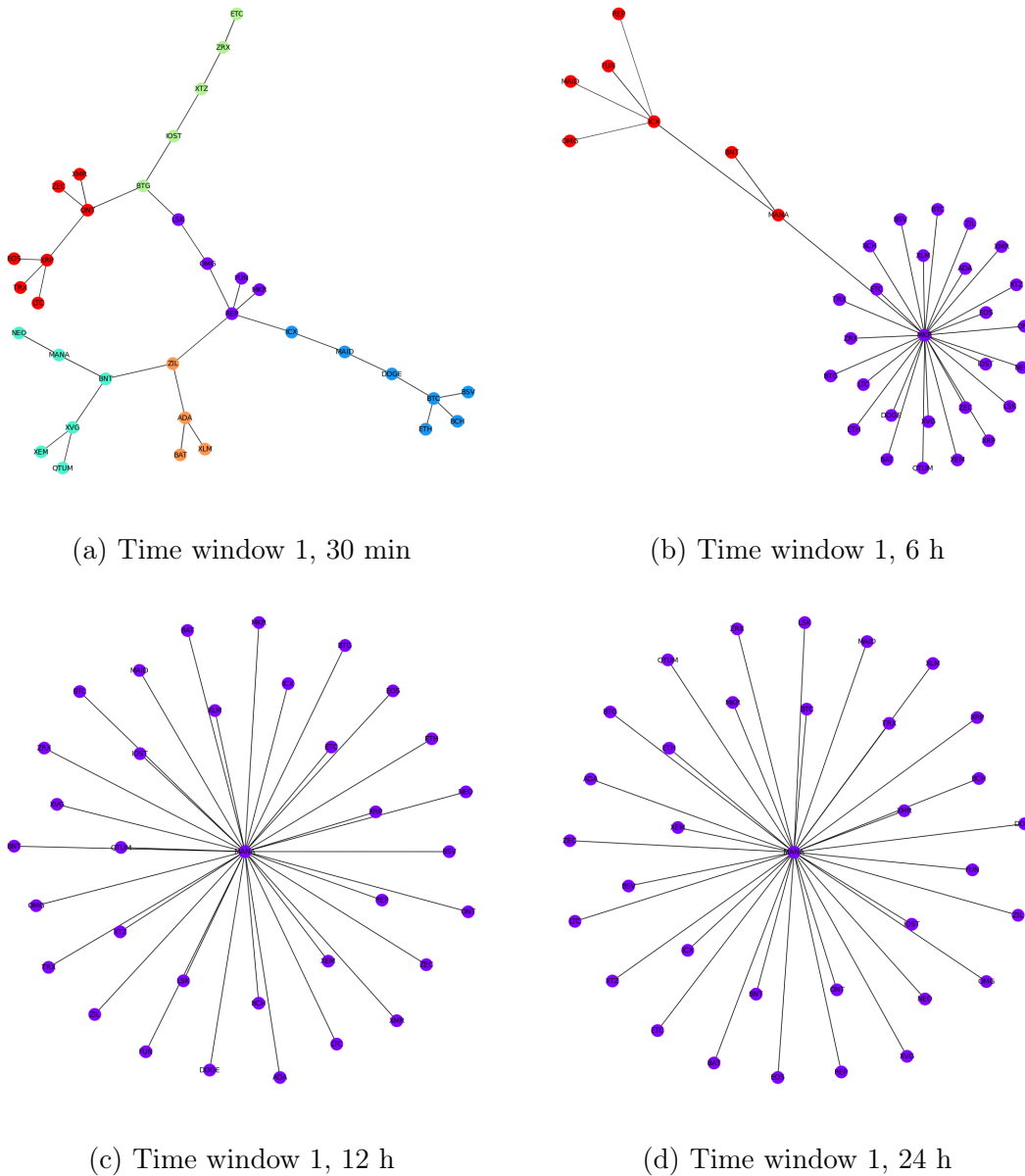


Figure 5.3: **Graph structure for the first time window, community detection is applied using Louvain method. Four different timescales are used, e.g., (a) 30 min, (b) 6 h, (c) 12 h, (d) 24 h.**

traded on three different exchanges, i.e. Kraken, Bitstamp and Btcbox, during the period between 2017 and 2021 to calculate the realized volatility of this most stable and popular cryptocurrency. They found that although Bitcoin is the most valuable and trustworthy cryptocurrency, its volatility fluctuates from 4.8 to 7.5. By contrast, with the same level of granularity, the stock market seems to be more stable, as the realized volatility stood at roughly 0.58 during normal times [148] and increased to just around 1.0 during the Covid-19 pandemic [149]. These facts suggest that

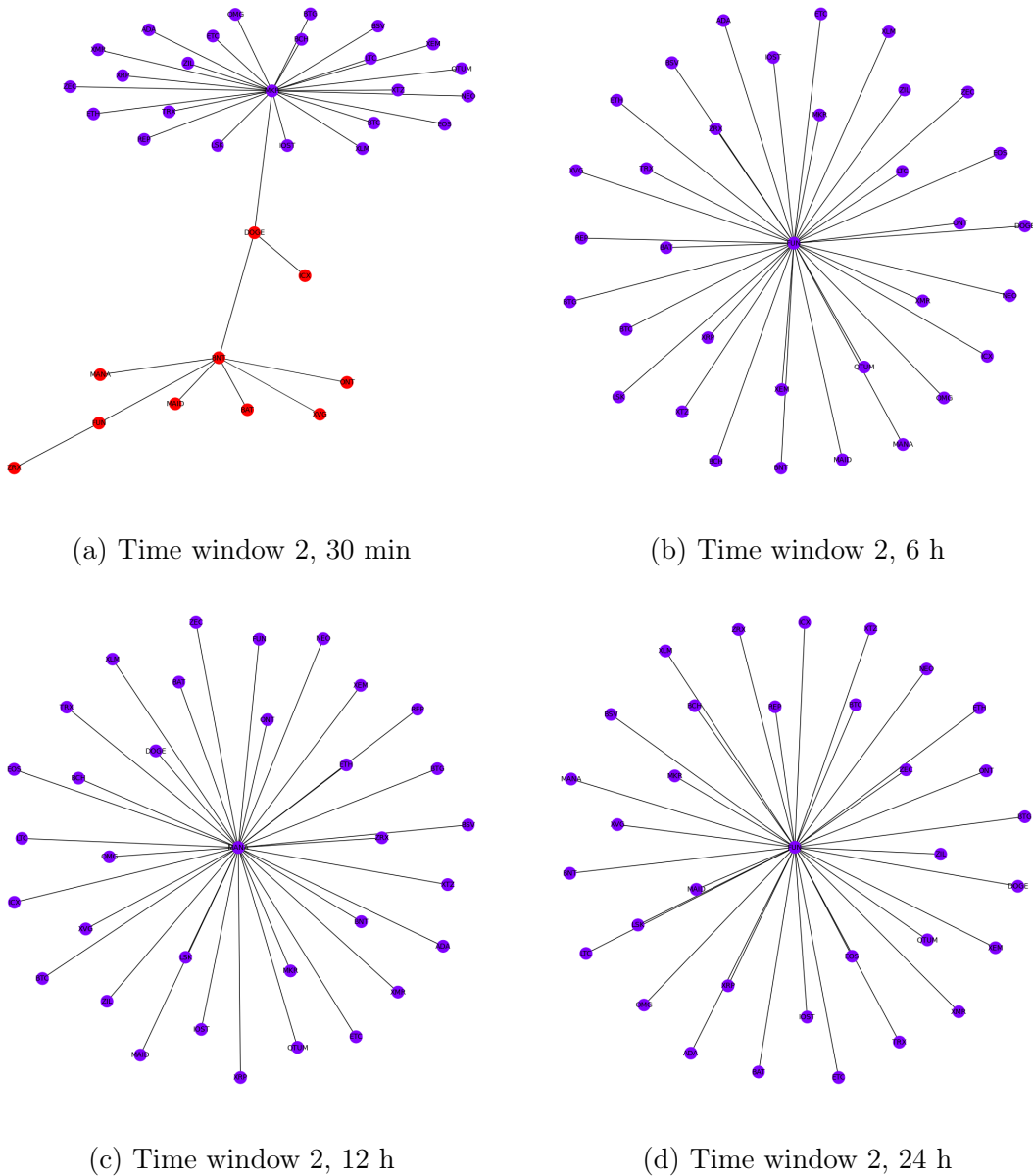


Figure 5.4: **Graph structure for the second time window, community detection is applied using Louvain method. Four different timescales are used, e.g. (a) 30 min, (b) 6 h, (c) 12 h, (d) 24 h.**

the cryptocurrency price fluctuations are dramatic even within a 5-minute period. Consequently, using a low-frequency dataset such as 12 h or 24 h appears to cause a loss of important information in cryptocurrencies, leading to an unclear correlation graph structure, as displayed in Figures 5.3–5.5. This problem has also been described in earlier studies such as [214]. However, existing studies mainly focused on daily data to detect communities in the cryptocurrency market. This is possibly explained by the difficulty in obtaining intraday cryptocurrency data.

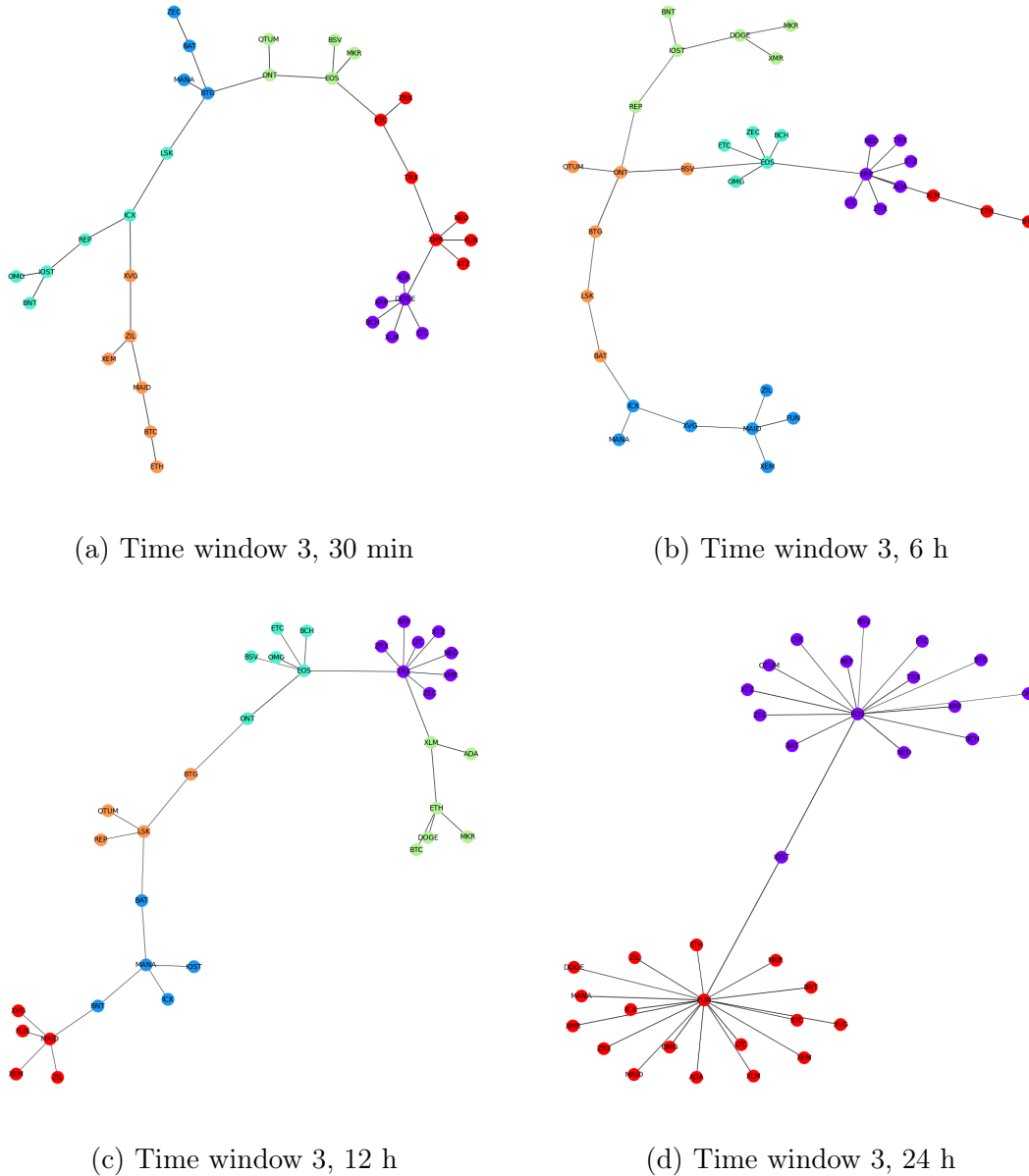


Figure 5.5: **Graph structure for the third time window, community detection is applied using Louvain method. Four different timescales are used, e.g., (a) 30 min, (b) 6 h, (c) 12 h, (d) 24 h.**

Ideally, we would like to use a dataset that is as fine-grained as possible to maintain all information in each cryptocurrency, which increases the informativeness of their graph structure. Unfortunately, our experiments show that for frequencies lower than 30 min, there are a huge amount of missing values as some cryptocurrencies are not traded frequently, thus requiring their removal or imputing a value. This adversely affects the correlation between time series and impacts our analysis. In the end, we choose a 30 min dataset for our experiments. This is also the fixed

Table 5.3: The growth of graph structures over time for each level of granularity, measured by Betweenness Centrality and Degree Assortativity

Metric	Time Window	Granularity			
		30 min	6 h	12 h	24 h
Betweenness Centrality	1	0.15	0.04	0.03	0.03
	2	0.05	0.03	0.03	0.03
	3	0.16	0.14	0.13	0.06
Degree Assortativity	1	- 0.49	- 0.77	-1.00	- 1.00
	2	- 0.72	- 1.00	-1.00	- 1.00
	3	- 0.51	- 0.48	- 0.58	- 0.98

Table 5.4: Number of communities at each level of granularity across three time windows

Granularity	Time Window 1	Time Window 2	Time Window 3
30 min	6	2	6
6 h	2	1	6
12 h	1	1	6
24 h	1	1	2

level of granularity used in other chapters.

### 5.2.5 Normal Times Versus Turbulent Times: A Graph Structure Perspective

To observe the growth of a graph structure over time, we use *Degree Assortativity Coefficient* [215] and *Average Betweenness Centrality* [10]. However, these metrics fail to tell us qualitatively and quantitatively the similarity between the two graphs' structures. Thus, to statistically compare the topological change between two graphs, we use three more metrics, including *v-measure*, *Degree centrality* [94] and *Eigenvalue method* [169, 216].

Table 5.5 shows results of *Betweenness Centrality* and *Degree Assortativity*. Immediately, we can see that there is a huge change occurring in time window 2, which corresponds to the turbulent time caused by the pandemic on both metrics.

Regarding the *Betweenness Centrality*, this metric decreases from 0.15 in time window 1 to 0.05 in the next period before going back to its original value prior to the pandemic outbreak (time window 1). A reasonable explanation for this movement is

Table 5.5: **The growth of graph structures over time measured by Betweenness Centrality and Degree Assortativity.**

Metrics	Time Window 1	Time Window 2	Time Window 3
Betweenness centrality	0.15	0.05	0.16
Degree Assortativity	-0.49	-0.72	-0.51

that the graph structure of the cryptocurrency market during normal times appears to have a dispersive tendency with the whole graph divided into multiple small-size groups such that each group share common characteristics. However, during Covid-19, these synchronize to form a big group. Thus, the number of groups decreases while the size of each group increases. In the recovery time, however, the graph starts to divide into smaller parts again, perhaps because the cryptocurrency market overcame the most connected period and started to go back to its normal behavior.

The *Degree Assortativity* results strongly support those of the *Betweenness Centrality*. In particular, a negative value shows that high-degree nodes are more likely to link to low-degree nodes, which means that each group in the graph has one node acting as a central node connecting to the rest. While the values in time windows 1 and 3 are approximately the same, time window 2 shows a decline of nearly 50 percent. This indicates that the number of connections between high-degree nodes and low-degree nodes increases, i.e. the graph forms big groups with a large number of leaf nodes in each group.

We notice that this time-varying structure is similar to what has been shown in the works of Drozd et al. [16, 99], who stated that the market had a distributed-graph topology or a hierarchical-graph topology in which no node dominates the graph during normal times. However, it became more centralized during the pandemic and started to distribute as this turbulent time was gone. More recently, another work by Nie also confirmed the same result [17].

Table 5.6 shows results of the three similarity metrics for graph structure comparison across different time periods: *normal time* (time window 1), *turbulent time* (time window 2) and *recovery time* (time window 3). Each value shows the similarity between two graph structures observed in two time windows. For *v-measure*, the

higher the value is, the more alike the two graphs are. On the other hand, for the remaining values, a lower value indicates that the two graphs are more similar.

Table 5.6: **Similarity in graph structures between different time windows.** A higher value of metric *v-measure* indicates a greater similarity between two structures, whereas, higher values of metrics *Degree centrality* and *Eigenvalue method* indicate more dissimilarity between two structures.

Metrics	Time Window	1 vs. 2	1 vs. 3	2 vs. 3
	Degree centrality		0.5	0.09
Eigenvalue method		844.45	4.59	759.16
<i>v-measure</i>		0.04	0.32	0.02

The differences between time window 2 and the other two time windows are very clear. In particular, the *v-measure* between time window 1 and 3 is 0.32, meaning that communities found in time window 3 hold roughly one-third of characteristics from time window 1's communities. By contrast, *v-measure* values between time windows 1 and 2 as well as between time windows 2 and 3 are negligible, standing at 0.04 and 0.02, respectively. Additionally, for the topological structures of MSTs, the other two metrics also show the same principle since time windows 1 and 3 share common characteristics and the similarity degree of other cases is nearly zero. Notably, the *Eigenvalue method* shows a significant divergence of time window 2 with others, as shown in Table 5.6.

The severe pandemic and the global downturn of March 2020 together seem to have actually changed the way cryptocurrencies interact with each other. The changes of these interactions have created new communities and broken down old ones, i.e. some cryptocurrencies become closer to each other while others moved further away from each other due to the Covid-19 pandemic and the economic recession. The topological structure during this turbulent time clearly shows completely different patterns compared to the periods when the global market are stable. Furthermore, we noticed that the community structure started to recover back to its pre-Covid-19 levels after June 2020, which coincides with the time the global economy recovered and the Covid-19 pandemic had less impact. During this time, some

characteristics of the graph structure reappeared which resemble those during the normal time (it is obvious that these structures are not fully similar because they change over time, as proven in previous sections and, in addition, after June 2020, the global economy started to recover, but not as well as in the past and the pandemic still had an impact on the economy worldwide to some extent). This is why the *v-measure* between time windows 1 and 3 increased significantly and the corresponding differences measured by *Degree centrality* and *Eigenvalue method* are very small. The community structures for the three time windows are shown in Figures 5.3a, 5.4a and 5.5a.

## 5.2.6 Understanding Investment Decisions of Cryptocurrency Traders Based on Ranking Distribution

Following the clear discrepancies in graph structures between the second time window and the others, in this sub-section, we dig into cryptocurrencies in each community at each time window to understand how these digital currencies form communities over time, especially when experiencing turbulent periods. To this end, the average rankings over the examined period<sup>5</sup> of cryptocurrencies are selected to be representatives in each community. In particular, we use the rankings to examine how they are distributed in each community and compare the distribution between different time windows.

Additionally, to highlight the characteristics of the ranking distribution in each time window, we categorize the cryptocurrencies into two tiers. Specifically, cryptocurrencies ranked 30 or higher are classified as Tier 1, while the rest are classified as Tier 2. Next, we count the number of Tier 1 cryptocurrencies in each time window. Subsequently, we calculate the percentage of Tier 1 cryptocurrencies assigned to each community within that time window.

Table 5.7 displays communities detected by the Louvain algorithm in each time

---

<sup>5</sup>The ranking of a cryptocurrency is measured by its market capitalization (a multiplication between the number of coins in circulation and the current market price of a single coin). We obtain cryptocurrencies' rankings on the <https://coinmarketcap.com> website.

window along with corresponding cryptocurrencies, their rankings and the percentage of Tier 1 cryptocurrencies (Pct(%)) for each community. We found that during the normal time (i.e. time window 1), there is a mix of high-ranking and low-ranking cryptocurrencies in each community. This is reflected in the fact that the percentage of Tier 1 cryptocurrencies is greater than zero across all communities. For example, community 6 has a size of 7 including top-ranking cryptocurrencies such as BTC, ETH and BCH, while also having very low-ranking ones such as MAID and ICX. On the contrary, during the downturn time (i.e. time window 2), we recognize that the community formation of these cryptocurrencies seems to be dramatically different from the previous time window. Specifically, there are only two communities found during this period, compared to 6 in the first one. More importantly, all the highest-ranking cryptocurrencies belong to one community while the majority of low-ranking cryptocurrencies are in the other. This pattern is further emphasized by the stark contrast in the percentages of Tier 1 cryptocurrencies between the two communities, with 100% Tier 1 cryptocurrencies belonging to the second community. By comparing these results with the recovery time (i.e. time window 3), we notice that this latter period shares common characteristics with both normal time and downturn time. In particular, cryptocurrencies started to separate from each other after the downturn time, as evidenced by an increase from 2 communities during the downturn time to 6 communities during the recovery time, which is equal to the normal time. Furthermore, while some communities show a mix between high- and low-ranking cryptocurrencies, there are three communities that are similar to the downturn time, including community 4 with all high-ranking cryptocurrencies (accounting for 54.5% of Tier 1 cryptocurrencies within this time window) and communities 1 and 5 with all low-ranking cryptocurrencies (0% of Tier 1 cryptocurrencies.)

To facilitate a more transparent understanding of our findings, the changes in ranking distributions between different time windows are visualized in Figure 5.6. Each community is represented by a circular shape and the ranking of cryptocur-

Table 5.7: **Distribution of rankings and the percentage of high-ranking cryptocurrencies (Per1)**

in each community during different phases of the financial market: normal time, downturn time and recovery time. The rankings are sorted in ascending order.

Bold values are minimum and maximum ranks in each period.

	Community	Cryptocurrencies	Rankings	Pct(%)
Normal time	1	ADA, XLM, BAT, ZIL	10, 13, 32, 99	12.50
	2	BTG, IOST, XTZ, ZRX, ETC	12, 21, 45, 57, 83	12.50
	3	LSK, OMG, REP, FUN, MKR	26,54, 58, 70, 168	6.25
	4	NEO, MANA, BNT, XVG, XEM, QTUM	19, 31, 41, 86, 117, <b>184</b>	6.25
	5	ONT, ZEC, XMR, XRP, EOS, TRX, LTC	3, 6, 7, 11, 16, 29, 35	37.50
	6	ICX, MAID, DOGE, BTC, BSV, ETH, BCH	<b>1, 2, 5, 9, 34, 84, 130</b>	25.0
Downturn time	1	DOGE, ICX, BNT, MANA, ZRX, FUN, MAID, BAT, XVG, ONT	32, 33, 40, 45, 60, 81, 105, 124, 139, <b>196</b>	0.00
	2	ADA, BCH, BSV, BTC, BTG, EOS, ETH, ETC, IOST, LSK, LTC, MKR, NEO, OMG, QTUM, REP, TRX, XEM, XLM, XMR, XRP, XTZ, ZEC, ZIL	<b>1, 2, 4, 5, 6, 7, 9,</b> 11, 12, 15, 17, 18, 21, 22, 27, 30, 34, 48, 51, 53, 54, 62, 65, 91	100.00
Recovery time	1	BTG, MANA, BAT, ZEC	56, 62, 67, 107	0.00
	2	ONT, QTUM, EOS, BSV, MKR	24, 31, 53, 75, 88	9.10
	3	XVG, ZIL, XEM, MAID, BTC, ETH	<b>1, 2, 38, 48, 109, 136</b>	18.20
	4	ADA, DOGE, XRP, BCH, XLM, LTC	6, 7, 9, 15, 16, 20	54.50
	5	OMG, BNT, IOST, REP, ICX, LSK	68, 78, 85, 100, 101, <b>140</b>	0.00
	6	ETC, ZRX, TRX, NEO, XMR, FUN, XTZ	17, 27, 33, 35, 64, 76, 129	18.20

rencies is represented by the intensity of the blue color, i.e. the darker the blue, the lower the cryptocurrency's rank. At first glance, there is a clear difference between the two communities in Figure 5.6b, with community 1 being much darker than community 2. On the other hand, there is a mix of both bright and dark blue in the majority of communities from the two remaining Figures 5.6a and 5.6c, showing a high level of similarity between them. The exception comes from communities 1, 3 and 5 in Figure 5.6c since community 3 contains all bright colors while the others only have dark colors.

When it comes to these results, the behavior of investors in making investment decisions can be considered a potential explanation for the time-varying graph structures. During normal times, i.e. when the financial market is stable and there is no major event occurring that impacts wider society, investors show contrarian behavior. That is, their decisions to invest in a cryptocurrency are based on their own market analysis and are not influenced by other investors' choices. This might push up the dynamism of the market where a large number of cryptocurrencies with both high and low rankings are invested. As a result, there is a diversification in terms

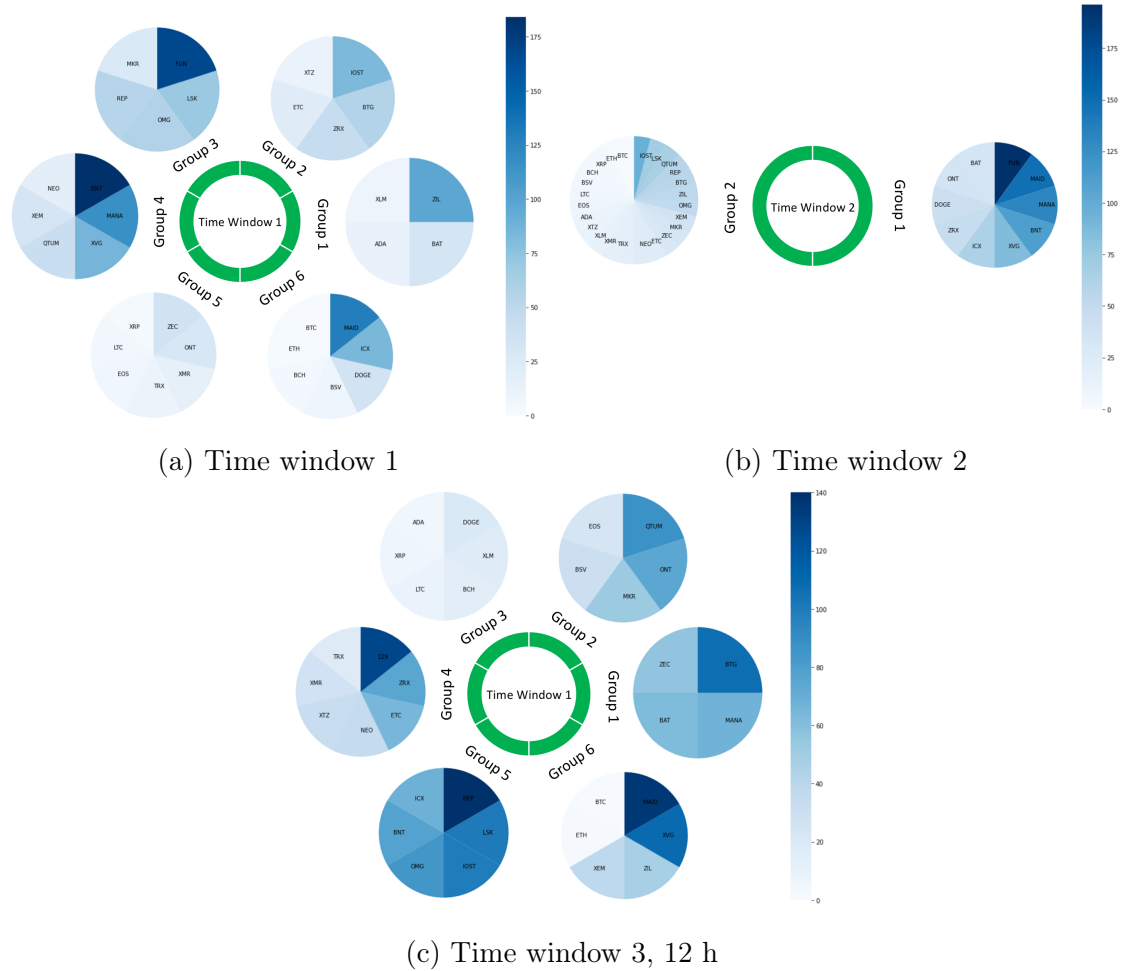


Figure 5.6: **The ranking distributions of cryptocurrencies.** Each community is represented by a circular shape while the rankings of cryptocurrencies in this community are given by the blue color intensity, i.e. the darker the blue, the lower the cryptocurrency’s rank. Three time windows correspond to three market conditions, i.e. a) Time window 1 (normal time), b) Time window 2 (turbulent time), c) Time window 3 (recovery time).

of rankings in each community. Empirically, it is found that there was no herding behavior before the pandemic. For instance, Yarovaya et al. in [217] used hourly price time series of multiple exchanges such as Binance, Bitbay, BitFinex, Coinbase and major cryptocurrencies including BTC, LTC and ETH to find the existence of herding before the start of Covid-19. Based on the *Cross Sectional Absolute Deviation* model, they found that the herding behavior did not exist during this time. By contrast, during turbulent times, investors are panicked by the fluctuations of cryptocurrencies’ price as well as being bombarded by bad news that strongly affects their investment decisions.

Different studies have been carried out to investigate investors' behavior since the onset of the Covid-19 outbreak. Generally, these reached the same conclusion as the pandemic actually increased herding behavior in the cryptocurrency market. In [132], the authors used 43 cryptocurrencies with large market capitalization between 2013 and 2020, they found that investors in the cryptocurrency market tend to follow the consensus and the impact of coronavirus media coverage is significant on the herding behavior. In particular, news related to the coronavirus increases fear and affects the behavior of investors reducing appetite for risk. Consequently, investors disregard their private information and follow others' investment decisions. However, the impact of media is reduced when the market returns to normal. This is in line with different studies that use different datasets and time periods [217–219]. More importantly, the way investors perform herding behavior is that they tend to invest in the major and most frequently traded cryptocurrencies [96]. This can be explained by the fact that high-ranking cryptocurrencies are more mature so they are more stable than the rest and are more likely to retain value under the uncertainty of the global financial market, causing a bias on the part of investors [220]. Consequently, major cryptocurrencies are seen to increase in terms of trading volume and act as a store of value during turbulent times [221]. In other words, there is a risk aversion occurring after the pandemic outburst as described in [189]. Eventually, all high-ranking cryptocurrencies belong to one community, as shown in Figure 5.6b.

Regarding low-ranking cryptocurrencies, we notice that cryptocurrencies with the lowest rankings in our dataset belong to one community during the turbulent time (i.e. time window 2), as shown in Figure 5.6b. One possible reason for this is that low-ranking cryptocurrencies are less likely to be the choice of investment during downturn times because they have negligible value and bring more risk to investors. Instead, they are mainly used for other purposes, such as paying transaction fees, as currency for a smart contract or simply a token on a cryptocurrency platform used to access applications [96]. Being used for the same purpose propels the correlation

between these cryptocurrencies.

All findings that we have shown earlier help us to explain the graph structure in time window 3, which corresponds to the recovery time. During this time, concerns about the pandemic start to decrease, meaning that not only cryptocurrency but also other traditional assets recover with investors' newfound positive attitude bringing them back to normal trading. Cryptocurrency traders start to diversify their portfolios by investing in different low and high-market-capitalized assets as well as making their own decisions [132]. However, one remarkable phenomenon that is worth taking into consideration is risk-taking behavior. A piece of research implemented by Huber et al. [222] used 100 return time series of risky stocks to conduct a survey related to the investment behavior of professional market traders. The responses of more than 800 participants reveal that a number of investors underestimate risk after prolonged exposure to high risk, as they become accustomed to the uncertainty of the economy. Thus, they begin investing in high-risk assets again to gain more profits. This tendency explains the similarity in the graph structures between time windows 1 and 3. However, as we can see in Figure 5.6c, there exists one community with all high-ranking cryptocurrencies (i.e. community 3) and one community with all low-ranking ones (i.e. community 5) as a result of the risk aversion of a portion of investors after the great shock caused by the pandemic, which mean that they still maintain the behavior of avoiding risk during the turbulent time by investing in the most reliable and well-known cryptocurrencies while leaving low-ranking cryptocurrencies for other purposes. Consequently, these two communities behave like the communities in time window 2.

### 5.2.7 Remarks

This section discusses one of our very first experiments which plays an important role in sharpening our studies later on. In other words, this is the onset of our long-term research where future studies rely on these first-ever experiments' results. Specifically, this section aims to examine graph structures of cryptocurrencies across

different periods, representing different market conditions, i.e. stable, turbulent, bearish and bullish. We then identify unique patterns in the graph structures at each market condition and explain the changes in graph structures under different market conditions via investors' behavior in investment decision-making.

To this end, we find that the changes in graph structures across different market conditions reveal corresponding changes in investment decisions of investors. In particular, investors tend to make their own investment decisions based on their personal market analysis and experience during normal times. Therefore, this causes a diversification in the cryptocurrencies chosen to invest in, since not only high- but also low-ranking cryptocurrencies are added to the portfolios of investors. As a result, the graph structure of cryptocurrencies during normal times is scattered with many small communities, composed of both low- and high-ranking assets. On the other hand, investors tend only to trade cryptocurrencies with high market capitalization during turbulent times, while smaller cryptocurrencies are mainly used for other more utilitarian purposes, such as transaction fees, smart contracts tokens or simply to run a digital platform. Consequently, the graph structure during this circumstance is compressed and divided into only two communities, i.e. high-ranking assets and low-ranking assets.

Furthermore, from this section, several aspects and pre-processing steps are determined and fixed for future studies. Firstly, the granularity of 30-min for time series data is proven to be the most suitable frequency for studies. Secondly, we have also examined the influence of noise and trend effects from cryptocurrencies on the correlation among those assets as well as their corresponding graph structure as a consequence. For this, we have raised the awareness of analyzing the financial markets from both perspectives, i.e. with and without noise and trend effects since the results after denoising and detrending might enrich our understanding of the markets.

## 5.3 Volatility-Based Co-Movement in the Cryptocurrency Market

### 5.3.1 Introduction

Following the examination of returns-based graph structures as set out in the previous Section 5.2, we notice that there is limited use of volatility time series as the main data in cryptocurrencies-related experiments since very few works have mentioned this economic indicator to date. As proved in [32], return values cannot explain all phenomena in the cryptocurrency market. Instead, the volatility of cryptocurrencies has the potential, we believe, to provide new insights that enrich our understanding of the market as a whole. From this point of view, we continue our research on cryptocurrency graph structures by using the volatility indicator, which will be discussed in this section.

Specifically, we first examine whether the volatility-based graph structures provide different results compared to the returns-based graph structures. For this, a time-varying comparison between returns and volatility-based graph structures is executed to extract their similarities and differences. To facilitate the comparison, we use a different approach of partitioning the entire period of data into sub-periods from the previous section. That is, instead of identifying fixed sub-periods, we use a sliding window of 30 minutes in this section, which means that a graph structure is constructed every 30 minutes. Hence, the graph structures based on returns and volatilities are compared every 30 minutes, with details about this experiment described in Section 5.3.3. Subsequently, a deeper investigation of volatility-based graph structures is conducted to extract new insights from the previous section.

### 5.3.2 Choosing an Optimal Time Window Length for graph Construction

Unlike the previous section where we construct a cryptocurrency correlation graph for each pre-defined sub-period and the time window size for each of these graphs is subsequently pre-defined to be equal to the length of the corresponding sub-period. In this section, we adopt a sliding window with a sliding step of 30 minutes to construct such graphs. Hence, it is necessary to determine the optimal size for this sliding window. In this regard, we propose two criteria for choosing a suitable window size for graph construction. Firstly, we choose a window size such that the corresponding graph is not sensitive to small changes in window sizes. In other words, the characteristics of the graph are not influenced by small changes in window sizes. Secondly, as declared in Section 4.2, there are two measures that are commonly used for calculating volatility. Since there appears to be no basis for favoring one measure over the other, we target selecting a window size that minimizes the discrepancy in two graphs obtained by two different volatility measures to ensure the consistency of results. We adopt three graph-based metrics, including *degree assortativity*, *betweenness centrality* and *closeness centrality* to assess a graph structure. Thus, the comparison between two graphs is based on these metrics.

For this experiment, we select a set of 28 potential window sizes, spanning from 48 to 1344 data points, which represent 1 day to 28 days (4 weeks) of trading. The rationale behind this selection is that the cryptocurrency market varies very quickly, thus a longer duration might fail to capture significant events happening in the market. Apart from that, we use the linear correlation metric *Pearson* and the non-linear correlation metric *Spearman* to calculate the similarity between two graph structures.

Regarding the first criterion, Figure 5.7 displays the similarity between each pair of two graph structures obtained from two corresponding window sizes (as measured by *Pearson* and *Spearman*). Since the results produced by three graph-based metrics are consistent, we use *betweenness centrality* as a representative for those. We find

that characteristics of graph structures do not change greatly when the window size is large. This suggests that the graph structures remain relatively stable if we use relatively large window sizes. Remarkably, we observed a plateau point starting at a window size of 816 (corresponding to 17 days) with the similarity just slowly increasing from here until the last window size (corresponding to 28 days). As a result, we exclude window sizes that are less than 816.

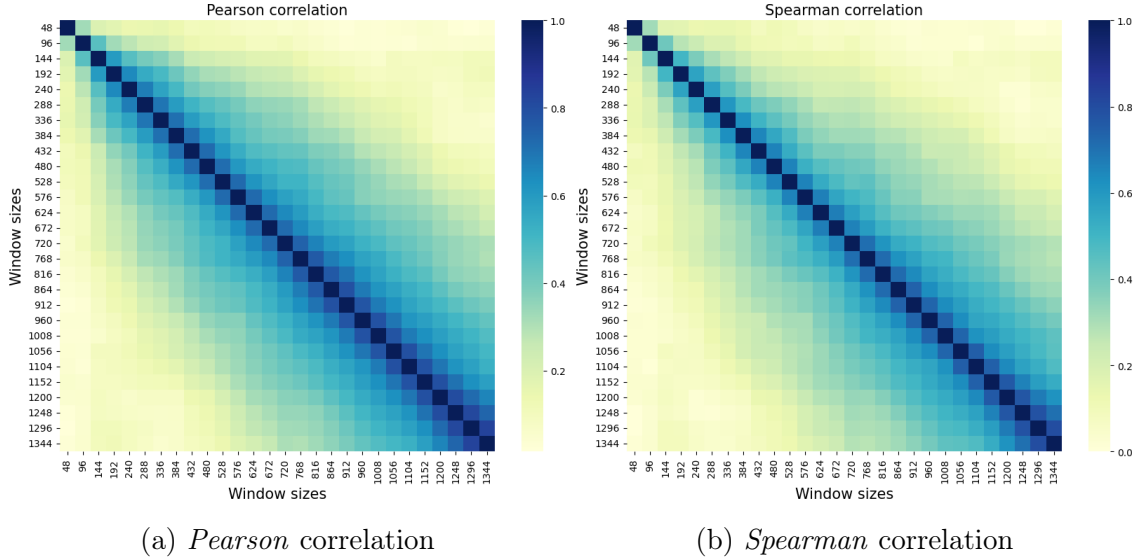


Figure 5.7: **The similarity of each pair of graph structures using different window sizes, observed by *betweenness centrality*. a) Measured by *Pearson*, b) Measured by *Spearman*.**

Regarding the second criterion, we gain a rather similar conclusion as two graph structures obtained from two different volatility measures show a high degree of similarity when using a window size of 816 or higher, indicating a negligible difference in graph structures given the large window size used. The only difference is that *Pearson* and *Spearman* values drop slightly after the window size of 1104 (corresponding to 23 days), before resuming a steady increase from the window size of 1248 (corresponding to 26 days). Details for this result are displayed in Figure 5.8.

In summary, we choose the window size of 1104 data points (23 days) for graph construction. This window size guarantees consistency in graph structures irrespective of the volatility formulas we use and also reduces greatly the changes in graph

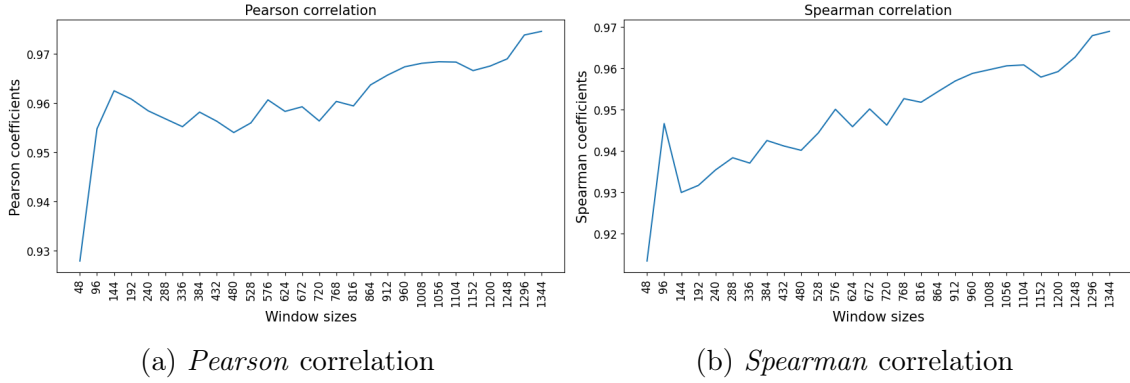


Figure 5.8: **The similarity between two graph structures derived from 2 corresponding volatility formulas, observed by *betweenness centrality*. a) Measured by *Pearson*, b) Measured by *Spearman*.**

structures when changing to another window size. Moreover, this window size is large enough to satisfy the Quality Factor requirement ( $Q > 1$ ) of Random Matrix Theory (as described in Section 4.4) as well as minimizing the time and space complexity.

Additionally, since two volatility formulas show negligible differences in terms of graph-based experimental results when using this window size, without loss of generality, we only use the more well-known formula called *standard deviation-based volatility* (Eqn. 4.1) to conduct the remaining experiments.

### 5.3.3 Volatility vs Returns: Unveiling the Differences in Cryptocurrency Graph Structures

We first determine whether volatility information provides unique traits about the cryptocurrency market. For this, we compare time-varying graph structures constructed using return time series as relative to volatility time series. A graph structure is represented via three graph-assessing metrics, including *degree assortativity*, *betweenness centrality* and *closeness centrality*. Moreover, as emphasized earlier about the importance of considering results with noise and trend effects excluded, we implement this experiment in two scenarios, before and after removing these effects to gain a broader view of cryptocurrencies.

Figure 5.9 illustrates the changes in graph structures every 30 minutes with the

existence of noise and trend, derived from return and volatility time series, respectively. For each figure, red lines indicate the 30-min values of *degree assortativity* (top sub-figure), *betweenness centrality* (middle sub-figure) and *closeness centrality* (bottom sub-figure), calculated on the corresponding graph structure. Blue lines indicate the centered moving average over a 14-day period of each metric while green lines represent the mean value obtained from each metric.

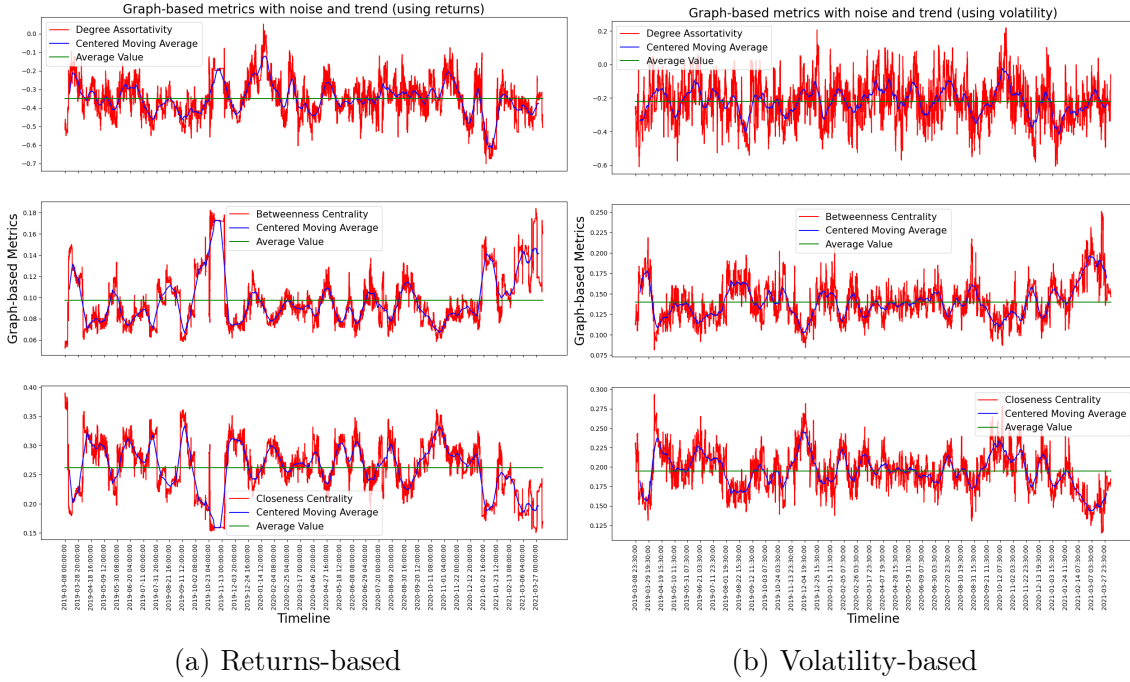


Figure 5.9: Changes in graph structures every 30 minutes before denoising and detrending, constructed by a) return time series, b) volatility time series. Graph structures are assessed by *degree assortativity*, *betweenness centrality* and *closeness centrality*, respectively.

These graph-based metrics show continual fluctuations over the considered period, for both returns and volatility-based cases. This emphasizes the continual changes in the interaction among different cryptocurrencies over time, with the effects of noise and trend present. Furthermore, we do not see any patterns from such fluctuations in the graph structures. In this situation, it moves on to removing noise and trend effects to extract possible hidden patterns from the time-varying graph structures.

Figure 5.10 displays the *degree assortativity*, *betweenness centrality* and *closeness centrality* of returns and volatility-based cryptocurrency graphs every 30 minutes,

respectively, after removing noise and trend effects. The fluctuations of these metrics are also dramatic as those affected by noise and trends previously. Hence, whether noise and trend effects are present or not, the interactions between cryptocurrencies consistently change over time. However, the graph-based metrics after denoising and detrending exhibit different characteristics from those with the existence of noise and trend, as proved by the Wilcoxon signed-rank test [204]. In particular, we conduct the tests to compare the metrics before and after removing these effects. For each metric, namely *Degree assortativity*, *Betweenness centrality* and *Closeness centrality*, our alternative hypothesis is that the median of the differences between the two conditions (before and after removing noise and trend) is not symmetric around zero, which indicates a difference in the result before and after removing noise and trend. The results show that all the tests accept the alternative hypothesis (the medians of differences before and after denoising and detrending for returns-based Degree assortativity, Betweenness centrality, Closeness centrality are 0.51, 0.04 and -0.18, respectively. The medians of differences before and after denoising and detrending for volatility-based Degree assortativity, Betweenness centrality, Closeness centrality are 0.31, -0.01 and 0.01, respectively. For all tests, p value is 0.0). Notably, there are indeed several patterns that can be recognized in both returns and volatility-based results after excluding noise and trends, as discussed below.

Regarding the returns-based results shown in Figure 5.10a, the time-varying values of all three metrics (red lines) tend to cluster at either very low or very high compared to the mean value (green lines). This contrasts with Figure 5.9a where these metrics' values fluctuate around the mean value most of the time. To support our claim, we calculate the average distance of each graph-based metric's values from its corresponding mean value in two scenarios: before and after removing noise and trend. A smaller average distance indicates that the metric's values tend to be closer to its mean throughout the considered period. On the other hand, a large average distance indicates that the metric's values tend to be far away from its mean throughout the considered period. Table 5.8 presents the results, showing

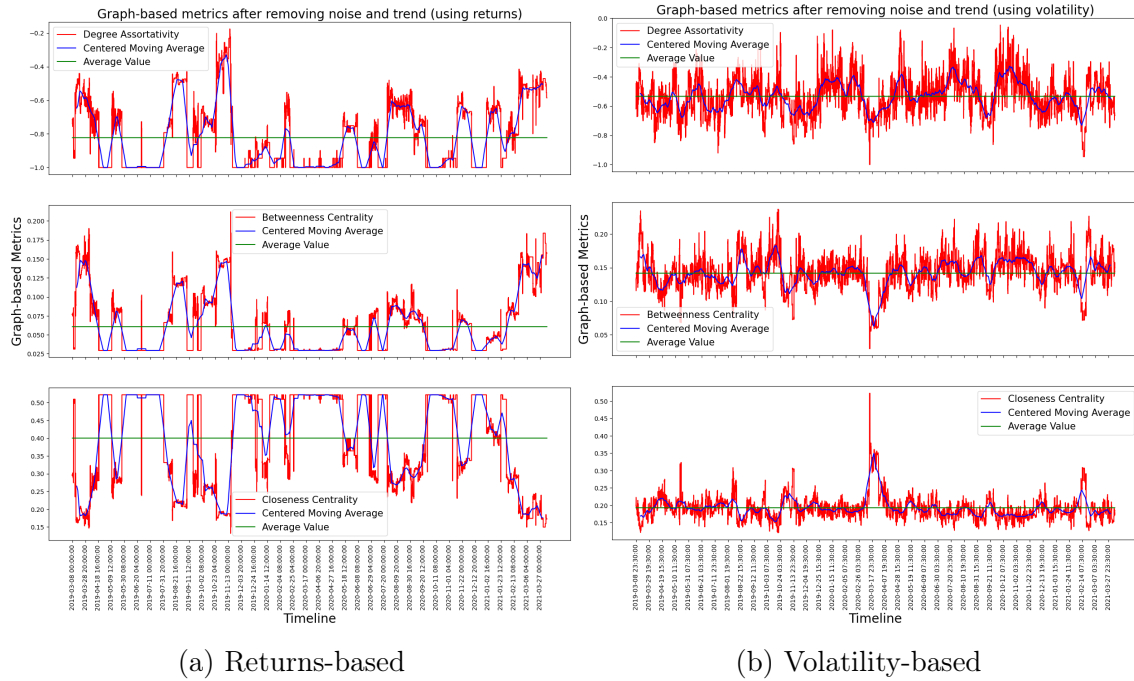


Figure 5.10: Changes in graph structures every 30 minutes after denoising and detrending, constructed by a) return time series, b) volatility time series. Graph structures are assessed by *degree assortativity*, *betweenness centrality* and *closeness centrality*, respectively. Black lines represent the borders of the timeline partition (only for returns-based graphs)

that with noise and trend, the average distance is significantly smaller than after their removal. This difference is especially evident in *Closeness centrality*, where the average distance is 0.03 before and 0.12 after removing noise and trend. Moreover, we also see noticeable patterns in the denoised and detrended graph structures in terms of such values. In particular, they are relatively small during the period between January and July 2020, which means that the graph tends to be compressed during this time, i.e. nodes are close to each other and form one large homogeneous group. On the other hand, large values are dominant during the first period from February 2019 until the end of the same year, implying a scattered graph where nodes are more distant from each other, forming multiple smaller groups. The last period, after July 2020, reveals a mixture of characteristics observed in the first two periods.

For the purpose of verifying this phenomenon, we use the *Symbolic Aggregate Approximation* (SAX) [223] to compare the differences between the three mentioned

Table 5.8: Average distance between a data point and the mean of a graph-based metrics (using returns)

With noise and trend?	Degree Assortativity	Betweenness Centrality	Closeness Centrality
Yes	0.07	0.01	0.03
No	0.18	0.04	0.12

periods relative to each other. To this end, we begin by splitting the metric time series (red lines) into different segments based on the structure of the graph. Specifically, our experiment has revealed that metrics' values within the first 50 percentiles represent a compressed graph<sup>6</sup> (i.e. nodes are close to each other and form a big group). Values surpassing the 80th percentile represent a scattered graph<sup>7</sup> (i.e. nodes are far from each other and spread out). On the other hand, values falling between the 51st and 80th percentile show a mixed graph, displaying characteristics of both compression and scattering. Consequently, three segments are used in this experiment, comprising 1st to 50th percentile (a), 51th to 80th percentile (b) and 81th to 100th percentile (c). Next, we partition the time series into three periods, including 08/03/2019 - 15/01/2020, 16/01/2020 - 23/07/2020 and 24/07/2020 - 06/04/2021. This division is based on the observation that metrics' values between 16/01/2020 and 23/07/2020 indicate a distinct pattern compared to the other periods. Lastly, we calculate the percentage of occurrence for each segment in each of the three periods.

Table 5.9 shows the percentage of each segment (e.g. a,b and c) in each period from the returns-based metrics (Figure 5.10a). Intuitively, the second period clearly displays a different behavior compared to the others. Specifically, the cryptocurrency graph structure tends to be compressed during the second period (16/01/2020 - 23/07/2020) with more than 75 percent of the time displaying an "a" SAX type and

<sup>6</sup>The term "compressed graph" is used to describe a graph where nodes are close to each other, forming large, interconnected communities (usually 1 or 2 communities), rather than being distributed into multiple smaller communities. The structure of a graph is assessed via graph-based metrics such as *Degree Assortativity*, *Betweenness centrality* and *Closeness Centrality* as discussed in Section 4.3.

<sup>7</sup>A scattered graph is the opposite of a compressed graph. In a scattered graph, the nodes are more distant from each other, forming several smaller, distinct communities.

nearly 100 percent of the time showing "a" and "b" SAX types. Conversely, the first and third periods are similar to some extent with more than 50 and 60 percent of the time showing "b" and "c" SAX types, respectively, while the percentage of "a" SAX type is much lower than the second period, indicating that the graph structure tends to be scattered during these periods. The notable difference is that there is an increase of "b" SAX type and a decrease of "a" SAX type between the first and third periods. This result is in line with our findings in Section 5.2.5 where significant changes in the graph structure of cryptocurrencies during the outbreak of the Covid-19 pandemic (first half of 2020) were discovered. On the other hand, the post-outbreak period demonstrates a combination of graph structure characteristics observed in both pre-outbreak and outbreak of the pandemic.

**Table 5.9: The percentage of 3 SAX segments in each period. "a" stands for the first segment, corresponding to the first 50 percentiles of values of graph-based metrics, "b" stands for the second segment, ranging from the 51st percentile to the 80th percentile, "c" stands for the third segment, covering values that are greater than the 80th percentile.**

	SAX type	Betweenness	Degree Assortativity	Closeness
08/03/2019 - 15/01/2020	a	47.6	47.7	47.6
	b	22.3	30.8	22.4
	c	30.1	21.5	30.0
16/01/2020 - 23/07/2020	a	75.3	76.8	75.3
	b	24.2	21.2	24.4
	c	0.50	2.00	0.30
24/07/2020 - 06/04/2021	a	34.3	33.4	33.9
	b	43.8	35.1	43.9
	c	21.9	31.5	22.2

Regarding the volatility-based results displayed in Figure 5.10b, we notice several abnormal spikes where their values are greatly at variance from the rest, especially two spikes including March 2020 and February 2021 which mutually appear in all three metrics. These spikes indicate a compressed graph where its nodes are close to each other and form a big community, rather than broken down into multiple small communities. In other words, the time corresponding to these spikes is the time when most cryptocurrencies share similar movements in terms of volatility, while fewer cryptocurrencies share the same movements during other times. To strengthen

this finding, we compare the average distance from a graph-based metric’s values within each spike to the mean value of the graph-based metric against the average distance from the values over the entire period to the same mean value. Table 5.10 display the results of this experiment. As expected, the values within the spikes are much farther from the mean value compared to values outside of the spikes.

Table 5.10: **Average distance between a data point and the mean of a graph-based metric for the entire period, the spike in March 2020 and the spike in February 2021 (using volatility)**

Period	Degree Assortativity	Betweenness Centrality	Closeness Centrality
Entire	0.10	0.02	0.02
13/03/2020 - 27/03/2020	0.18	0.08	0.15
12/02/2021- 20/02/2021	0.26	0.06	0.08

The first common spike was between 13/03/2020 and 27/03/2020, which was equivalent to the worst time in the first wave of the Covid-19 pandemic, in the sense that the global economy experienced a crash due to the regulations and health protection measures enacted by almost all countries across the globe to defend against the pandemic [224]. This is proven by looking at the huge drop in the *S&P500* index by 30% in Figure 5.1d and also a dramatic increase by more than 500% for the case of the volatility stock index VIX in a matter of 2 months, as displayed in Figure 5.1c). This is also the period when public attention towards the pandemic increased significantly (Figure 5.1a) and the worldwide GDP growth rate dropped to the lowest point of all time (see 5.1b). Notably, The cryptocurrency market also reacted to this global crisis since the majority of digital coins started to lose their values in mid-February 2020 and reached the lowest values on 13/03/2020 as shown on Coinmarketcap, which was also the date graph-based metrics had the highest peak.

The second spike occurred between 12/02/2021 and 20/02/2021, coinciding with a bullish period in the cryptocurrency market as shown on Coinmarketcap. This was the first time the cryptocurrency market experienced a significant surge in prices, not only with major coins such as BTC and ETH but also with minor coins like

BAT and MKR. However, the prices of most cryptocurrencies started to decline again after 20/02/2021. Interestingly, it can be seen that the graph metrics fell back to their values prior to the spike period after 20/02/2021.

We acknowledge that there are other spikes in graph-based metrics. However, they are only consistent in one or two graph-based metrics and the duration for these spikes is much shorter, within one or two days. For instance, three spikes are in *betweenness centrality* and *closeness centrality* but not in *degree assortativity* including May, August and November 2019. One possible driver for this phenomenon might be trading activities associated with illicit practices, such as money laundering, terrorist financing, and corruption [225], which we cannot directly track or collect information on. In addition, another potential reason stems from social media-related activities such as market manipulation or pump-and-dump schemes, where a group of crypto traders tries to create herding to take advantage of the associated upward price movement [226]. Since these spikes are not consistent in all graph-based metrics, we exclude these from our study.

In summary, it is clear that the volatility information provides new results compared to the return one. In particular, the returns-based correlation graph reacts to the change in market condition between January and July 2020 (e.g. during the outbreak of the Covid-19 pandemic and the global economic crisis) by changing its structure and this change lasts until the market recovers to its previous condition. In the meantime, the volatility-based graph can point out the most changing period within a critical event, such as the worst period of the market crash in March 2020 and the most bullish period in 2021. To this end, both economic indicators are useful for understanding the correlation between cryptocurrencies in different market conditions and also the influence of relevant critical events on the cryptocurrency market. However, return information seems to be more suitable for general analysis that does not require a high level of detail to understand fundamental traits while volatility seems to be a favorite indicator for detailed analysis.

### 5.3.4 Can Changes in Volatility-based Correlation graph be Explained by Investors' Sentiments?

We suspect that the spikes from graph-based metrics on the volatility-based graph (Figure 5.10b) can be attributed to herding behaviour [188], which is mainly driven by collective trading decisions [227, 228]. Specifically, investors in the cryptocurrency market tend to follow the market consensus, mimic common investment decisions of the general public and are strongly influenced by the market sentiment, especially during market crashes and bullish periods [188]. That is, they tend to sell off their shares when cryptocurrencies experience a downturn or bad news circulates on the internet, in order to avoid the loss. By contrast, they are more likely to purchase new shares during bullish phases or when positive news about cryptocurrencies spreads among the general public due to the fear of missing out on increasing the cryptocurrency price. As a result, similar trading actions occur across different cryptocurrencies, causing a consistent impact on the volatility of different coins, which in turn, leads to an increase in the correlation within the cryptocurrency market with respect to the volatility. This collective correlation has been discussed in a study by James and Menzies [32], where the authors calculated the proportional contribution of one cryptocurrency to the total volatility of 52 major cryptocurrencies throughout the period between April 2019 and June 2021. They found that the proportion of the market's total volatility is more evenly distributed among all the cryptocurrencies during the Covid-19 market crisis and the bull market of 2021. There was less deviation between different volatilities of individual cryptocurrencies in these two periods, i.e. everything was similarly volatile together.

To reinforce our assumption, we examine the correlation between volatility-based graph's metrics and the Fear and Greed index<sup>8</sup> of investors in the cryptocurrency market. Without loss of generalization, we use the *betweenness centrality* as a repre-

---

<sup>8</sup>Assessing the emotions and sentiments of people towards the cryptocurrency market. The index ranges from 0 to 100, which is classified into 5 levels in total, including extreme fearful (0 - 25), fearful (26-46), neutral (47-54), greedy (55-75) and extreme greedy (76 - 100), website: <https://alternative.me/crypto/fear-and-greed-index/>

sentative metric for all metrics used in our experiments to visualize such correlation, which is displayed in Figure 5.11. In accordance with our expectations, the spike in March 2020 corresponds to the time when people are extremely pessimistic about the market due to the significant decrease in prices. The spike in February 2021, by contrast, corresponds to the time when people are most optimistic about cryptocurrencies, boosted by the continuous growth of this market. This result clearly points out the relationship between the volatility graph structure and people’s sentiments. That is, the cryptocurrency graph constructed by volatility time series is strongly compressed, i.e. the correlation between different cryptocurrencies increases, and the graph forms a big community, when investors overreact to movements in the market, which leads to a herding phenomenon, either buying (positive sentiment) or selling (negative sentiment) behavior. On the other hand, when investors’ sentiments are not extreme, the structure tends to be scattered, forming several communities of small sizes due to distinct movements between different cryptocurrencies. In other words, they exhibit a diversification in investments by not following each other’s investment decisions.

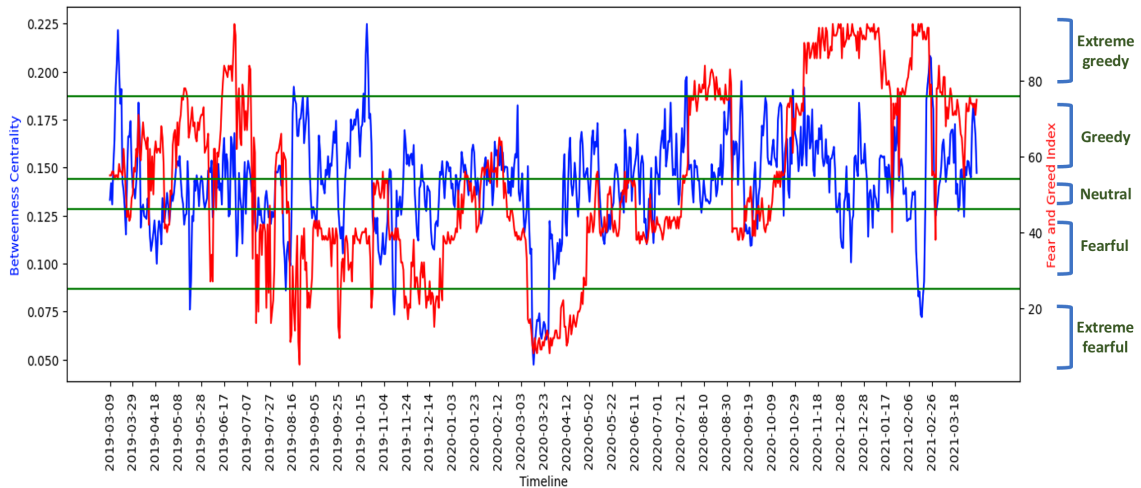


Figure 5.11: The movements of *betweenness centrality* and *Fear and Greed* index from 09/03/2019 until 06/04/2021. The red line refers to the *betweenness centrality*. The blue line refers to the sentiment index. Each green line corresponds to a boundary for a type of sentiment, including extreme fearful, fearful, neutral, greedy and extreme greedy.

We also notice that the magnitude of the first spike in March 2020 is much larger

than that of the second spike in February 2021. This indicates that the Covid-19 pandemic followed by the global economic crisis put more pressure and influence on the cryptocurrency market compared to the bullish time in 2021. In other words, investors reacted to the downturn time more strongly than the bullish time, causing cryptocurrencies to become more correlated. Perhaps, negative news during the outbreak of the pandemic triggered investors to trade irrationally by imitating others' actions in an attempt to avoid substantial losses, which happened to the majority of cryptocurrencies. On the other hand, investors demonstrate greater wisdom and make more rational investment decisions during the positive-return period since they are not under pressure of losing money, thereby reducing herding behavior. These findings contradict several studies conducted before 2020 such as [229, 230], which argued that positive conditions exerted a stronger influence on volatility than negative conditions. However, our result is in line with a recent study utilizing a dataset covering 3 continuous years from 2019 until 2021 [231]. Notably, the stronger influence of negative conditions compared to positive conditions on investors has been well-established in traditional markets and has stood as an enduring fact [232]. This implies that the cryptocurrency market has started sharing similar characteristics with traditional markets, which means that it is becoming progressively more mature. This is aligned with [100], where the authors showed that major and famous cryptocurrencies are becoming more and more similar to the regular stock market.

### 5.3.5 Remarks

This section examines the usefulness of volatility-based correlation graphs in exploring and understanding the time-varying correlation among different cryptocurrencies.

Firstly, we clarify whether this type of graph produces different results compared to those constructed from returns. This comparison is executed via three graph-based metrics, namely *betweenness centrality*, *degree assortativity* and *closeness centrality*, which are used to assess the structure of a graph. We find that

volatility information delivers new insights distinct from return information. Thus, this economic indicator is worth utilizing in experiments to enrich the literature on the cryptocurrency market. Furthermore, from this section, we also notice that return information seems to be more suitable for general analysis to understand fundamental characteristics, where a high level of detail is not necessary. By contrast, volatility information seems to be more suitable for highly detailed analysis. For instance, returns-based correlation graphs react to the change in market condition between January and July 2020 (e.g. during the outbreak of the Covid-19 pandemic and the global economic crisis) by changing its structure and this change lasts until the market recovers to its previous condition. On the other hand, volatility-based graphs manage to identify the most changing period within a critical event, such as the worst period of the market crash in March 2020 and the most bullish period in 2021. Additionally, the outcome of volatility-based graphs suggests their ability to anticipate major events that impact the cryptocurrency market such as economic crises, bull and bear periods.

Secondly, given the changes in volatility-based graph structure over time (represented by three graph-based metrics), especially the most changing periods (i.e. March 2020 and February 2021), we seek to answer the question *what causes these changes?* For this, we aim to explain these changes through investors' sentiments, represented by the Fear and Greed index. Indeed, our experiment confirms a strong link between investors' sentiments and changes in the graph structure. In particular, the volatility correlation between cryptocurrencies increases when investors overreact to the market, either becoming extremely fearful or extremely greedy. Consequently, the corresponding graph tends to be compressed, forming a big community. On the other hand, such a strong correlation between cryptocurrencies mostly disappears when investors' sentiments are not extreme, the corresponding graph in this case is then more likely to be scattered, forming several communities of small sizes. To this end, we suspect that an abnormal increase in the volatility correlation is attributed to the herding behavior among investors.

Additionally, this section also highlights the usefulness of the noise and trend removal scheme, since the experimental results become clearer and exhibit visible patterns after removing these effects. By contrast, correlation graphs constructed by the original time series with the noise and trend present propose confusing results, with unclear patterns. Thus, this section reinforces the importance of running experiments in both cases, with and without noise and trends to extract as much insight as possible.

## 5.4 Measuring the Influence of Individual Cryptocurrencies on the Cryptocurrency Market

### 5.4.1 Introduction

Following the above thorough analysis of the time-varying correlation between cryptocurrencies, we go a little further by looking into the influence of each cryptocurrency on the correlation graph over time. In this experiment, we assess the influence of each cryptocurrency by analyzing its position and connectivity in a correlation graph. In graph theory, influence is often associated with the centrality of a node, where highly connected nodes play a more significant role in the graph's structure and dynamics [233]. Applying this concept to the cryptocurrency correlation graph in this study, cryptocurrencies with strong and numerous connections are considered influential as their price movements tend to be mirrored by a great number of assets in the correlation graph. On the other hand, less influential cryptocurrencies are those that have fewer and weaker connections to other assets in the correlation graph. The measurement of such influence on a correlation graph using node centrality has been adopted widely in the literature [234–236].

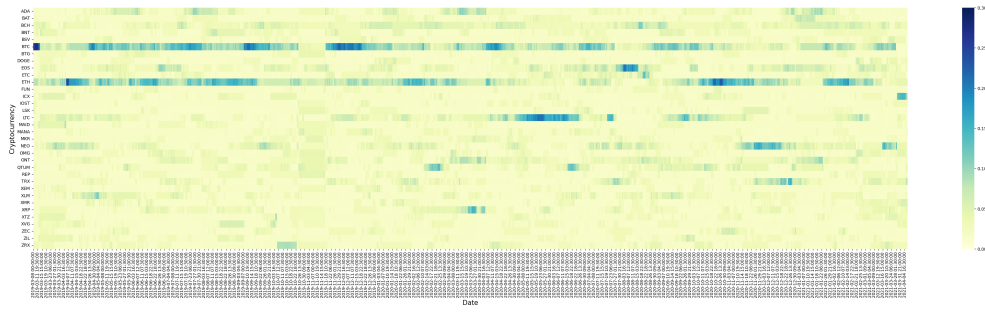
For this purpose, we use a centrality-based algorithm launched by Google which is originally used to assess the importance of web pages based on their links, called *PageRank* [170]. It assigns a numerical score between 0 and 1 to each node, con-

sidering factors such as the number and quality of incoming links. Nodes with higher scores are considered more influential, and thus transfer their information to a wider range of coins in the graph. We calculate the *PageRank* value for each cryptocurrency at each time window, based on the corresponding correlation graph, constructed by returns and volatility time series, respectively. We recall that each time window has 1104 timestamps (23 days), with a sliding step of 30 minutes. To generalize the influence of each cryptocurrency in the graph over the considered period from 2019 to 2021, we use two metrics inherited from *PageRank* that quantify the magnitude of influence for each cryptocurrency, namely *Total Accumulation (TA)* and *Dominance Score (DS)*. These two metrics are described in Table 4.6). Additionally, in this section, we also run experiments under two scenarios, with and without the effect of noise and trends.

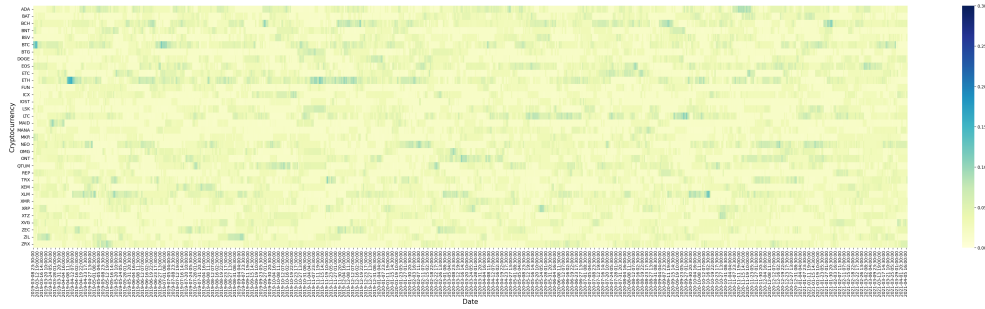
#### 5.4.2 Determining Which Cryptocurrencies are Dominant

Figures 5.12 and 5.13 illustrate the PageRank value over time of each cryptocurrency, calculated from the returns and volatility-based correlation graph, respectively. For each type of graph, we display the results before and after removing noise and trend effects. PageRank values are represented by colors from pale yellow to dark blue, with pale yellow being equivalent to 0, increasing to dark blue, which is equivalent to 1. To this end, these figures are useful for observing the changes in the influence of each cryptocurrency over time.

To generalize the influence of each cryptocurrency on the correlation graph during the 3-year period from 2019 until 2021, we utilize two metrics, viz. *Total Accumulation (TA)* and *Dominance Score (DS)*. The results are shown in Tables 5.11-5.14, each table comprises 4 columns: **Crypto** represents cryptocurrency symbols; **TA/DS** represents Total Accumulation/Dominance Score values; **Rank** represents the ranking of cryptocurrencies; **Percentage** represents the ratio of one cryptocurrency's Total Accumulation/Dominance Score value to the sum of all Total Accumulation/Dominance Score values

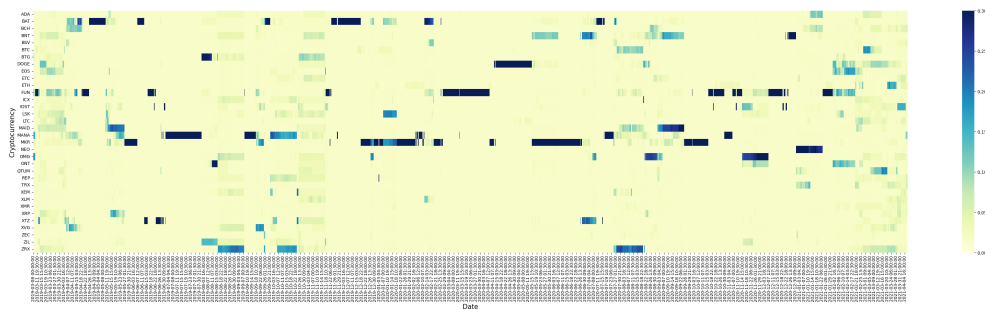


(a) Returns-based results

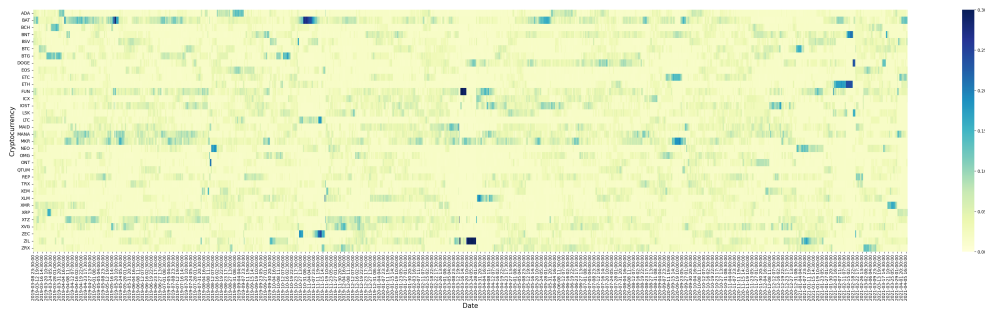


(b) Volatility-based results

Figure 5.12: The influence of each cryptocurrency on the correlation graph over time (before denoising and detrending), constructed from a) return time series, b) volatility time series.



(a) Returns-based results



(b) Volatility-based results

Figure 5.13: The influence of each cryptocurrency on the correlation graph over time (denoised and detrended), constructed from a) return time series, b) volatility time series.

Our experiment demonstrates that returns and volatility yield similar outcomes. In particular, with the influence of noise and trend effects, the prominent cryptocurrencies such as BTC and ETH appear to have a greater influence on the market than the rest, as evidenced by their higher *Total Accumulation* and *Dominance Score* throughout the observed period (see Tables 5.11 and 5.13). This is a common result obtained by various studies whose experiments are based on original data with the presence of noise and trend [189, 237, 238]. By contrast, this characteristic disappears after removing the noise and trend effects. Specifically, smaller and less well-known cryptocurrencies tend to have a larger impact on the market (see Tables 5.12 and 5.14). In contrast, BTC and ETH show negligible impact on the market.

Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage
ETH	1564	2	4.29	BAT	1008	40	2.76	BTC	3414	1	9.35	BAT	822	40	2.25
BCH	1387	6	3.8	DOGE	998	32	2.74	ETH	3090	2	8.46	MKR	797	30	2.18
BTC	1375	1	3.77	OMG	996	51	2.73	LTC	1692	7	4.63	ZIL	791	70	2.17
XLM	1361	12	3.73	XMR	991	15	2.72	NEO	1495	22	4.09	REP	785	66	2.15
NEO	1335	22	3.66	BSV	982	11	2.69	BCH	1473	6	4.03	XTZ	783	18	2.14
LTC	1324	7	3.63	ICX	976	60	2.68	EOS	1443	11	3.95	XVG	783	91	2.14
EOS	1304	11	3.58	BNT	971	144	2.66	ONT	1206	37	3.30	MANA	780	110	2.13
ADA	1265	10	3.47	XVG	965	91	2.64	ADA	1161	10	3.18	ZEC	774	32	2.12
ONT	1211	37	3.32	XEM	964	25	2.64	QTUM	1095	51	3.00	XMR	772	15	2.11
QTUM	1137	51	3.12	XTZ	952	18	2.61	TRX	1054	15	2.89	LSK	756	64	2.07
TRX	1070	15	2.93	MANA	925	110	2.53	XRP	1010	4	2.77	BSV	754	11	2.06
ZEC	1058	32	2.90	REP	908	66	2.49	XLM	985	12	2.70	MAID	754	117	2.06
ETC	1050	28	2.88	BTG	901	59	2.47	BNT	912	144	2.50	DOGE	753	32	2.06
ZRX	1044	51	2.86	MKR	894	30	2.45	ZRX	895	51	2.45	BTG	747	59	2.05
XRP	1037	4	2.84	FUN	860	202	2.36	ETC	879	28	2.41	XEM	733	25	2.01
LSK	1032	64	2.83	MAID	823	117	2.26	ICX	856	60	2.34	FUN	732	202	2.00
ZIL	1010	70	2.77	IOST	803	90	2.20	OMG	848	51	2.32	IOST	703	90	1.93

(a) Volatility

(b) Returns

Table 5.11: **Total Accumulation (TA)** for each cryptocurrency with the effect of noise and trend, observed via volatility (a) and returns (b), respectively. The results are ordered from high to low.

Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage
BAT	1818	40	4.98	BTC	981	1	2.69	MKR	3354	30	9.18	BTC	787	1	2.15
MKR	1627	30	4.46	BSV	975	11	2.67	FUN	3152	202	8.63	BCH	771	6	2.11
FUN	1460	202	4.00	OMG	952	51	2.61	MANA	2318	110	6.35	ZIL	768	70	2.10
MANA	1405	110	3.85	ADA	940	10	2.58	BAT	2130	40	5.83	QTUM	766	51	2.10
XTZ	1309	18	3.59	XEM	937	25	2.57	DOGE	1550	32	4.24	REP	740	66	2.03
ZIL	1288	70	3.53	ETC	931	28	2.55	ZRX	1309	51	3.58	ICX	717	60	1.96
IOST	1260	90	3.45	ZEC	919	32	2.52	BNT	1243	144	3.40	ADA	714	10	1.95
XVG	1206	91	3.30	NEO	907	22	2.49	OMG	1207	51	3.31	XRP	711	4	1.95
MAID	1190	117	3.26	ETH	904	2	2.48	MAID	1177	117	3.22	LTC	709	7	1.94
BNT	1182	144	3.24	LTC	903	7	2.47	XTZ	1054	18	2.89	ETH	707	2	1.93
BTG	1079	59	2.96	EOS	880	11	2.41	NEO	923	22	2.53	XEM	692	25	1.89
REP	1077	66	2.95	TRX	872	15	2.39	IOST	896	90	2.45	TRX	684	15	1.87
DOGE	1074	32	1.94	XMR	863	15	2.37	ONT	877	37	2.40	ETC	676	28	1.85
ZRX	1073	51	2.94	ONT	850	37	2.33	LSK	864	64	2.36	XLM	670	12	1.83
XLM	1069	12	2.93	XRP	837	4	2.30	BTG	856	59	2.34	BSV	653	11	1.79
LSK	1046	64	2.87	QTUM	813	51	2.23	XVG	798	91	2.19	XMR	641	15	1.75
ICX	1042	60	2.86	BCH	811	6	2.22	EOS	794	11	2.17	ZEC	619	32	1.70

(a) Volatility

(b) Returns

Table 5.12: **Total Accumulation (TA)** for each cryptocurrency after removing noise and trend, observed via volatility (a) and returns (b), respectively. The values are ordered from high to low.

Crypto	DS	Rank	Percentage	Crypto	DS	Rank	Percentage	Crypto	DS	Rank	Percentage	Crypto	DS	Rank	Percentage
ETH	4244	2	11.63	ZRX	579	51	1.59	BTC	14204	1	39.89	XLM	5	12	0.01
LTC	3544	7	9.71	DOGE	511	32	1.40	ETH	11481	2	31.43	BNT	4	144	0.01
BCH	3347	6	9.17	MAID	490	117	1.34	LTC	3618	7	9.90	MAID	2	117	0.01
NEO	3094	22	8.48	ICX	475	60	1.30	NEO	1697	22	4.65	BSV	0	11	0.00
BTC	2913	1	7.98	LSK	442	64	1.21	EOS	1356	11	3.71	BTG	0	59	0.00
XLM	2563	12	7.03	BAT	319	40	0.87	QTUM	601	51	1.65	DOGE	0	32	0.00
EOS	1588	11	4.35	OMG	313	51	0.86	TRX	600	15	1.64	FUN	0	202	0.00
QTUM	1499	51	4.11	MANA	248	110	0.68	ADA	521	10	1.43	IOST	0	90	0.00
ONT	1372	37	3.76	XTZ	238	18	0.65	XRP	517	4	1.42	LSK	0	64	0.00
ZIL	1327	70	3.64	XMR	208	15	0.57	ONT	474	37	1.30	MKR	0	30	0.00
ADA	1326	10	3.63	BTG	190	59	0.57	ICX	434	60	1.19	OMG	0	51	0.00
ETC	977	28	2.68	BSV	190	11	0.52	ZRX	278	51	0.76	REP	0	66	0.00
TRX	963	15	2.64	BNT	177	144	0.49	ETC	259	28	0.71	XEM	0	25	0.00
XEM	947	25	2.60	REP	154	66	0.42	BCH	214	6	0.59	XMR	0	15	0.00
XRP	916	4	2.51	MKR	48	30	0.13	BAT	188	40	0.51	XTZ	0	18	0.00
ZEC	675	32	1.85	IOST	6	90	0.02	MANA	66	110	0.18	XVG	0	91	0.00
XVG	598	91	1.64	FUN	0	202	0.00	ZEC	9	32	0.02	ZIL	0	70	0.00

(a) Volatility

(b) Returns

Table 5.13: Dominance Score (DS) for each cryptocurrency with the effect of noise and trend, observed via volatility(a) and returns (b), respectively. The values are ordered from high to low.

Crypto	DS	Rank	Percentage	Crypto	DS	Rank	Percentage	Crypto	DS	Rank	Percentage	Crypto	DS	Rank	Percentage
BAT	4842	40	13.27	BNT	777	144	2.13	MKR	6328	30	17.32	ETH	429	2	1.17
MKR	3358	30	9.20	ETC	716	28	1.96	FUN	5876	202	16.09	LSK	388	64	1.06
IOST	2026	90	5.55	BTC	691	1	1.89	MANA	3782	110	10.35	BTC	346	1	0.95
ADA	1839	10	5.04	XEM	648	25	1.78	BAT	3489	40	9.55	LTC	187	7	0.51
XTZ	1696	18	4.65	OMG	647	51	1.77	ZRX	2747	51	7.52	XEM	64	25	0.18
ZIL	1612	70	4.42	LSK	612	64	1.68	DOGE	2053	32	5.62	XLM	52	12	0.14
DOGE	1601	32	4.39	ONT	557	37	1.53	OMG	1924	51	5.27	ICX	48	60	0.13
MANA	1573	110	4.31	BCH	552	6	1.51	MAID	1864	117	5.10	REP	23	66	0.06
XVG	1471	91	4.03	NEO	541	22	1.48	BTG	1479	59	4.05	TRX	13	15	0.04
REP	1265	66	3.47	ICX	515	60	1.41	NEO	1105	22	3.03	ETC	4	28	0.01
FUN	1231	202	3.37	XMR	488	15	1.34	XTZ	993	18	2.72	ZIL	3	70	0.01
MAID	1165	117	3.19	BSV	458	11	1.26	EOS	683	11	1.87	XRP	2	4	0.01
ETH	1080	2	2.96	TRX	299	15	0.82	QTUM	664	51	1.82	ZEC	2	32	0.01
BTG	1002	59	2.75	EOS	168	11	0.46	IOST	619	90	1.69	BCH	1	6	0.00
XLM	993	12	2.72	XRP	143	4	0.39	BNT	479	144	1.31	ADA	0	10	0.00
ZEC	879	32	2.41	QTUM	111	51	0.30	ONT	448	37	1.23	BSV	0	11	0.00
ZRX	835	51	2.29	LTC	90	7	0.25	XVG	433	91	1.19	XMR	0	15	0.00

(a) Volatility

(b) Returns

Table 5.14: Dominance Score (DS) for each cryptocurrency after removing noise and trend, observed via volatility (a) and returns (b), respectively. The values are ordered from high to low.

The significant influence of minor cryptocurrencies on the market can be attributed to 3 key factors: the prevalence of Pump & Dump schemes, the transaction activity of crypto investors and what investors use minor cryptocurrencies for.

In relation to Pump & Dump, it has become a familiar activity in the cryptocurrency market and is organized regularly by a wide range of investors. A study [226] observed 355 such events on 2 major exchanges, Binance and Yobit. The researchers collected data from 2 different sources, namely Telegram and PumpAnalysis.com during a 4-year period from 2018 until 2021, to analyze the characteristics of coins that are most likely to be targeted by manipulators for a Pump & Dump scheme. Their findings revealed that manipulators target relatively illiquid coins (low-ranking coins). This means that individuals will be more attracted to pumps of low-ranking

coins since these tend to be illiquid [239]. These schemes were found to impact not only the pumped cryptocurrency but also other cryptocurrencies. In particular, a study conducted by Balcilar and Ozdemir [240] investigated the risk spillover effect caused by the Pump & Dump schemes among various cryptocurrencies, using the *Frequency Connectedness*<sup>9</sup> approach. They found evidence of the risk spillover among different cryptocurrencies during pump events. Therefore, a Pump & Dump scheme often originates from a low-ranked cryptocurrency [242, 243], resulting in the transmission of synchronized fluctuations between that coin and other coins.

In order to investigate the connection between investor transaction activity and the significant influence of small cryptocurrencies, we calculate the correlation between our examined cryptocurrencies with respect to the number of transactions recorded every 30 minutes throughout the period, the result is illustrated in Figure 5.14. We collect the number of transactions within every 30 minutes for each cryptocurrency throughout our data period, between 13/02/2019 and 06/04/2021, we then use *Spearman* and *Pearson* correlation metrics to calculate the correlation between each pair of cryptocurrencies (the results are consistent between two correlation metrics). Additionally, we also include the ranking information in this figure to facilitate a comparison between high-ranking and low-ranking cryptocurrencies. Specifically, we classify cryptocurrencies ranked 21<sup>10</sup> or lower as low-ranking, which are colored green in the figure, while the rest are considered high-ranking, which are colored red in the figure.

Our analysis has revealed that low-ranking cryptocurrencies tend to have similar transaction activity patterns. That is, an increase or decrease in the number of transactions of one coin is reflected in trading activity in other coins, resulting in

---

<sup>9</sup>*Frequency Connectedness* measures risk spillover effects across financial markets or assets over different time horizons. Based on spectral analysis of variance decompositions from vector autoregressive (VAR) models, it quantifies how shocks in one variable or market affect others across various frequencies (short-, medium-, and long-term). This method is introduced by Diebold and Yilmaz [241].

<sup>10</sup>Based on our analysis, the number of circulating cryptocurrencies exceeded 5000 as of April 2021. However, only 20 cryptocurrencies had an average market capitalization (throughout the duration of our dataset) of more than 1.5 billion USD. This is the reason for classifying the top 20 cryptocurrencies with the highest market capitalization as high-ranking cryptocurrencies.

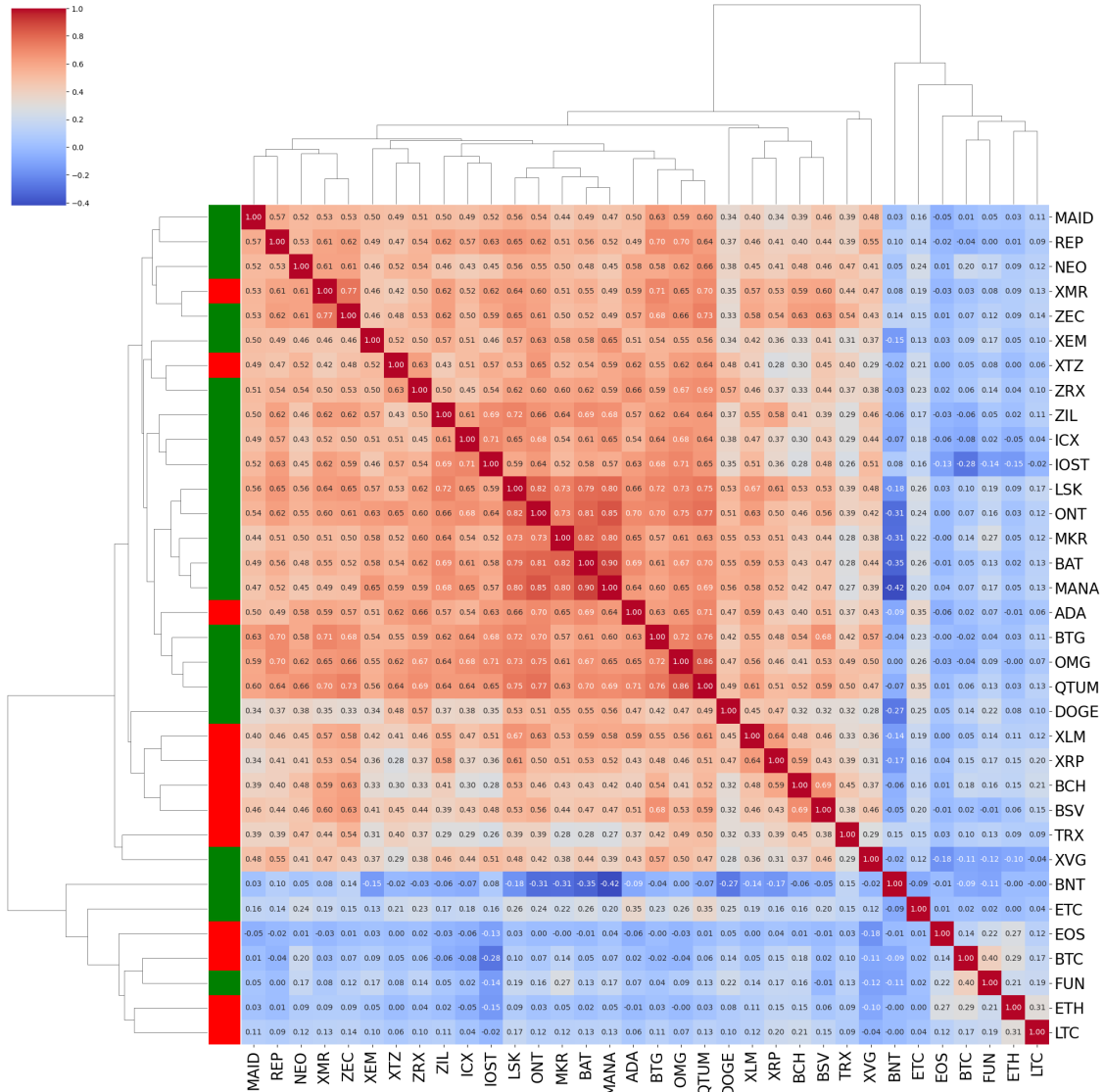


Figure 5.14: The correlation between 34 cryptocurrencies in terms of the number of transactions from 13/02/2019 until 06/04/2021 (on a 30-minute basis). Cryptocurrencies that are ranked 21 or lower are considered low-ranking (marked as green), while the rest are considered high-ranking (marked as red).

a similarity among their volatility movements. However, the biggest coins such as Bitcoin, Ethereum and Litecoin tend to behave independently from the rest since they show nearly zero correlation with other coins, showing that they have their own unique trading patterns. In the cryptocurrency market, only a few assets are well-known and have a large market capitalization, while the majority of coins are unpopular and do not attract much public attention. Therefore, in the absence of noise and trend effects, we should see real connections between different cryptocur-

rencies. That is, when the volatility of a small coin changes, it tends to affect the majority of cryptocurrencies. Conversely, changes in the volatility of a major coin are usually isolated due to its distinct trading pattern.

The significant impact of the largest cryptocurrencies in the correlation graph when considering the noise and trend effects highlights the crucial role of people's sentiment in the cryptocurrency market. As mentioned above, individuals primarily focus on the most prominent cryptocurrencies like BTC, ETH, and LTC, using them as a basis to evaluate the entire market. Consequently, they react when the prices of these coins change and anticipate similar changes in other assets [244, 245]. This results in similar trading decisions, ultimately leading to volatility movements of similar size and direction between different cryptocurrencies. However, without the effects of noise and trend, the dominance of large cryptocurrencies tends to diminish, giving way to the rising influence of minor and less famous cryptocurrencies. This observation highlights the true nature of the cryptocurrency market. That is, small cryptocurrencies are frequently targeted for a Pump & Dump scheme, leading to an increase in their volatility and a spillover effect that impacts a broad spectrum of coins [226]. Additionally, the transaction activity appears to exhibit similarities among the majority of small cryptocurrencies, while major cryptocurrencies demonstrate their own distinctive trading patterns (see Figure 5.14). The similarity in transaction activity among small cryptocurrencies might arise from the lead-lag phenomenon<sup>11</sup> between different cryptocurrencies. That is, small cryptocurrencies often follow the price movements of major cryptocurrencies like Bitcoin (BTC) and Ethereum (ETH) with a slight delay, usually within a few minutes. As a result, it is feasible to generate a profit by purchasing small cryptocurrencies whenever the price of Bitcoin (or ETH) increases, waiting a few minutes until the price of the small coins increases as well, and then selling them [246]. Another possible reason for the similarity in transaction activity is that cryptocurrencies with a low market capitalization are often used as safe havens in a portfolio, especially during turbulent

---

<sup>11</sup>A lead-lag effect describes the situation where one (leading) variable is cross-correlated with the values of another (lagging) variable at later times.

economic and market conditions [247].

Furthermore, we notice another common result from Tables 5.11-5.14. That is, the cryptocurrencies with the highest influence are BAT, MKR and FUN. This seems to stem from their increasing popularity and use cases. Specifically, these cryptocurrencies seem to be gaining progressively higher usage demand and are associated with various cryptocurrencies in the market. In particular:

- **BAT**<sup>12</sup>: Since the beginning, BAT has been shown to have a robust business model and real-world use cases [248]. This cryptocurrency has experienced a continuous surge in usage, partly thanks to the increase in usage of the Brave browser. As of 2021, Brave has 55 million monthly active users, 16 million daily active users and millions of verified creators accepting BAT<sup>13</sup>. As such, BAT is considered one of the most successful alt-coins to date. Moreover, it bridges across Ethereum and Solana blockchains and offers utility to both ecosystems [249, 250].
- **MKR**<sup>14</sup>: MKR is created to govern the MakerDAO platform. Specifically, MakerDAO offers its community of MKR holders the right to vote on risk management and business logic. Thus, this coin plays an important role in the development of the platform [251]. Moreover, MKR is also utilized for paying various fees associated with generating DAI - one of the most famous stablecoins at the time of writing [251]. Initially, DAI served primarily as a means of lending and borrowing crypto assets on MakerDAO. Over time, its utility has expanded to a wide range of applications on the Ethereum blockchain such as the creation of smart contracts. Thanks to the success of the Ethereum blockchain, DAI maintains its popularity and is relevant to various

---

<sup>12</sup>A cryptocurrency invented for revolutionizing the advertising industry on the internet. This coin is used on the Brave web browser as a reward for Brave users when they read ads. This browser has a unique mechanism to block ads and only recommends relevant ads based on users' preferences (<https://basicattentiontoken.org/>).

<sup>13</sup>This information is updated regularly on <https://basicattentiontoken.org/>.

<sup>14</sup>A main token on a peer-to-peer decentralized protocol called MakerDAO. This platform serves as an online bank for cryptocurrencies, facilitating borrowing, lending and savings of various cryptocurrencies (<https://makerdao.com/en/>).

cryptocurrencies [251, 252]. Consequently, MKR has experienced increased usage and is associated with numerous cryptocurrencies, as a result of the DAI's flourishing.

- **FUN**<sup>15</sup>: FUN is the only token accepted for making as well as receiving payment from players and game creators, respectively, on FunFair. Thus, players have to exchange other cryptocurrencies or fiat currencies into FUN to be able to play on FunFair. This online gaming platform has experienced a significant increase in different aspects, including business scale, daily active users, and net worth [253]. Especially, since the outbreak of the pandemic, online gaming platforms have attracted a record amount of users as a result of their being forced to stay at home due to the lockdown policy [254]. Consequently, the demand for using FUN has maintained its increase over time.

### 5.4.3 Remarks

In this section, we adopt the PageRank algorithm to examine the influence of each individual cryptocurrency on the collection of 34 cryptocurrencies from the HitBTC exchange. To this end, we found contradictory results between pre- and post-removal of the noise and trend effects. While the most influential cryptocurrencies in the pre-denoised and detrended collection are high-ranking cryptocurrencies, such as BTC, ETH and BCH, this phenomenon does not hold after the removal of noise and trend effects. Instead, certain minor cryptocurrencies unexpectedly exert a disproportionate influence on the market, such as BAT, MKR and FUN. From these results, investors should be aware of low-ranking cryptocurrencies and not underestimate them.

---

<sup>15</sup>A token on the decentralized, cryptocurrency-based casino gaming platform called FunFair (<https://funfair.games/>), which is the number one online gaming platform in the world.

## 5.5 Herding Behavior among Different Financial Assets

### 5.5.1 Introduction

Herding behavior has become a familiar phenomenon to investors, carrying the potential danger of both undervaluing and overvaluing assets, while also threatening market stability. From the previous sections, it is shown that the correlation among cryptocurrencies changes according to market conditions, especially during turbulent times when this change tends to be the most dramatic, with the correlation reaching the highest points. In this regard, herding behavior was our explanation for such significant changes. Therefore, we are eager to verify this assumption made in the previous experimental results, which is our main reason to work on this topic.

Specifically, this section applies herding detection algorithms followed by statistical analysis to explore characteristics of such phenomenon in financial markets. We begin with the cryptocurrency market. The study is subsequently expanded to traditional markets, including stocks and US ETFs. This experiment is done on each type of investment vehicle independently. To this end, we answer various questions from this experiment, i.e. Which events cause the herding behavior in each type of investment vehicle? How strong is this phenomenon? How are investors' behavior linked to the herding? How resilient is the herding? Subsequently, a comparison between the emerging market like cryptocurrencies and traditional markets is executed.

Notably, to reduce the portfolio risk, investors tend to invest in different types of investment vehicles at the same time [255]. Therefore, after addressing each separate type of investment vehicle independently, we combine all of them and run the herding experiments on the combination, to explore how the herding behavior looks like when different types of investment vehicles are considered at once. Our hypothesis is that the view on each type of investment vehicle differs from a composition of multiple different ones. In particular, a knowledge of the herding patterns of a specific market

is important for understanding its unique features, major events that the market has experienced, and which kind of information/news impacts the market, etc. On the other hand, observing different types of investment vehicle at the same time is critical for understanding the connections between them, highlighting the way investors diversify their portfolios or the impact of each separate type of investment vehicle on their combination. Such insights cannot be obtained by focusing only on a single market.

Additionally, unlike previous sections, here we use a different dataset comprising three types of investment vehicle, namely cryptocurrencies, stocks and US ETFs. For each asset, its historical price at a time step is aggregated from different exchanges. The dataset starts from 01/04/2019 until 03/05/2023, with a granularity of 30 minutes. Details about this dataset are described in Section 3.1.2. This dataset provides recent data covering the most impactful events in financial markets between April 2019 and May 2023. Also, we note that herding detection is performed on different sub-periods partitioned from the dataset based on the corresponding major events occurring in the financial markets. On the other hand, we exclude the experiments on the entire period from this section since we found no remarkable results from these experiments.

### **5.5.2 Timeline Division**

Before coming up with the idea of partitioning the dataset into sub-periods, we executed the herding detection and analysis on the entire period for each type of investment vehicle and also for the combination. We found that performing herding detection on the whole period tends to finish with a 'Yes' result. Furthermore, all types of investment vehicle exhibit herding behavior so no unique patterns can be found. Therefore, it is impossible to conclude anything from this result. This ambiguity encourages us to split the original period into sub-periods so that the herding behavior can be examined on a finer time interval, which enhances the detail level of the experiment.

Throughout the 4-year period from April 2019 until May 2023, the financial markets experienced several pressing events such as the US-China trade war [46, 47, 53], the Covid-19 pandemic [16, 58, 61], and the Ukraine-Russia conflict [67–69, 256]. Since different events might be characterized by different behavioral trends/reactions of investors, these events are our fundamental criteria for splitting the given dataset into smaller intervals. Following this, herding detection-related experiments in this study are run on each sub-period. This practice helps us track the changes in herding behavior over time as well as recognize common features appearing during a particular event.

Based on the aforementioned historical events and the performance of the *S&P500* index (see Sp500 historical data), we partition the original period into 5 sub-periods. In this context, we aim to identify four breakpoints along the timeline, as detailed below:

- The first breakpoint is set at 01/01/2020, which is justified by two key events. First, the US-China trade war officially ended in January 2020. Secondly, the Covid-19 pandemic began to spread to different countries in January 2020 and influenced the global economy accordingly. Consequently, the first sub-period spans from 01/04/2019 to 31/12/2019, which is referred to as the *Pre-Covid-19*.
- As shown in Section 5.2, the period from 01/01/2020 to 30/06/2020 corresponds to the worst Covid-19 pandemic time (first wave) when the financial markets and the worldwide economy experienced a significant downturn. Consequently, 01/07/2020 is selected as the second breakpoint. Therefore, the second sub-period spans from 01/01/2020 to 30/06/2020, which is referred to as the *Covid-19 pandemic*.
- From July 2020, financial markets recovered to their performance observed before the onset of the pandemic. Moreover, the markets have also entered a bullish phase since then that lasted until April 2022. Since this period is too

long, we seek to find a breakpoint for it. We have shown in Section 5.2.6 that although financial markets recovered following the first wave of the Covid-19 pandemic, investor behavior still exhibited characteristics observed both before and during the first wave. With this regard, we treat the remainder of 2020 as the Covid-19 aftermath, reflecting the lingering effects of the initial market crash. Therefore, we select 01/01/2021 as the third breakpoint, resulting in two sub-periods, namely *Post-Covid-19* (between 01/07/2020 and 31/12/2020) and *Bull market* (between 01/01/2021 and 30/04/2022).

- Following the previous partition, the fourth breakpoint is set at 01/05/2022, marking the start of the final sub-period, which spans from 01/05/2022 to 03/05/2023 (the end of our dataset). resulting in the last sub-period from 01/05/2022 to the end of our dataset (03/05/2023). This sub-period is characterized by the ongoing conflict between Ukraine and Russia, so we refer to this sub-period as the *Ukraine-Russia conflict*.

Table 5.15 displays an explicit division for the original period.

Table 5.15: **Five sub-periods declaration.**

Name	Pre-Covid-19	Covid-19 pandemic	Post-Covid-19	Bull market	Ukraine-Russia conflict
<b>Time interval</b>	01/04/2019 to 31/12/2019	01/01/2020 to 30/06/2020	01/07/2020 to 31/12/2020	01/01/2021 to 30/04/2022	01/05/2022 to 03/05/2023

### 5.5.3 Herding Behavior in the Cryptocurrency Market

We begin the herding behavior analysis with the cryptocurrency market. To detect the herding behavior, we adopt the well-known Cross-Sectional Absolute Deviation (CSAD) model. For a more comprehensive analysis, two versions of this model are of use, one distinguishes *up* and *down* markets (positive return and negative return periods) (see Equation 4.8), the other does not distinguish between those two market conditions (see Equation 4.7). Therefore, we provide the results of two versions in this section.

Table 5.16 displays the estimation of  $\beta_2$  from the CSAD model (without distinguishing *up* and *down* market conditions) and  $\gamma_2, \gamma_4$  from the CSAD model for the *up* and *down* market condition, respectively. Only values with the star symbol are statistically significant. Herd behavior is revealed at a period only when its corresponding coefficient (either  $\beta_2, \gamma_2$  or  $\gamma_4$ ) is statistically significantly negative. The more negative the value, the more pronounced the herding.

Table 5.16: **Herding detection results for cryptocurrencies. Herding is pronounced in a period if the corresponding coefficient is statistically significantly negative. We consider three levels of confidence (1% (\*), 5% (\*\*)) and 10% (\*\*\*) to be significant**

Periods	$\beta_2$	$\gamma_2$	$\gamma_4$
Pre-Covid-19	8.01*	11.33*	5.85*
Covid-19 pandemic	-4.16*	-3.79*	-4.20*
Post-Covid-19	-5.24*	-4.71*	-4.74*
Bull market	-2.71*	2.29*	-2.76*
U-R conflict	5.61*	5.11*	6.04*

The parameter  $\beta_2$  is estimated from the equation:

$$CSAD_t = \alpha + \beta_1|r_{m,t}| + \beta_2r_{m,t}^2 + \epsilon_t.$$

The parameters  $\gamma_2$  and  $\gamma_4$  are estimated from the equation:

$$CSAD_t = \alpha + \gamma_1D^{up}|r_{m,t}| + \gamma_2D^{up}r_{m,t}^2 + \gamma_3D^{down}|r_{m,t}| + \gamma_4D^{down}r_{m,t}^2 + \zeta_t.$$

The coefficient  $\beta_2$  reveals herding signals in three sub-periods: during the Covid-19 pandemic, the Post-Covid-19 and the Bull market time. Surprisingly, the strongest herding is found in the Post-Covid-19 while the peak of the Covid-19 pandemic, which we expected to show the strongest herding behavior, only comes in second place. On the other hand, the Pre-Covid-19 and Ukraine-Russia conflict sub-periods reveal an absence of herding behavior. The results remain mostly unchanged after separating the market into *up* and *down* conditions. Specifically, the coefficients  $\gamma_2$  and  $\gamma_4$  are only statistically significantly negative during the sub-periods when  $\beta_2$  is negative, indicating that either the *up* or *down* market condition or both conditions reveal the evidence of herding in these stages. However, while herding is pronounced in both *up* and *down* market conditions during the Covid-19 pandemic and Post-Covid-19 sub-periods, only the downturn times in the Bull market time have the appearance of herding. These results sufficiently capture the prevailing sentiment among cryptocurrency investors. Specifically, they seem to be influenced by mar-

ket fluctuations and breaking news[189], leading to similar investment decisions, i.e. selling off their investments when the market experiences a downturn or bad news circulates on the internet to avoid a loss while purchasing new investments during bullish phases or when positive news about cryptocurrencies spreads among the general public due to the Fear of Missing Out (FOMO) on an increase in the price of the coins[130]. The high degree of sensitivity of investors in this digital market may be attributed to the dominance of irrational (naïve) participants, as confirmed in different studies [188]. These investors participate in the cryptocurrency market primarily because of its perceived great potential for huge profit despite their lack of investment experience. Therefore, they tend to ignore their own analysis and information by following the decisions made by the public due to their lack of financial knowledge [257]. A greater herding magnitude during the Post-Covid-19 compared to the peak of the pandemic can be attributed to the cryptocurrency market's tendency to be more influenced by positive circumstances than negative ones [108, 258]. During this time, the market started to recover from the decline caused by the pandemic, experiencing a price surge since the fourth quarter of 2020. An upward trend in prices along with positive word of mouth, as well as financial news, triggered the positive sentiment of investors together with the FOMO phenomenon that eventually caused a purchase herding in the market [132, 138, 257]. In the Bull market time, although herding is evident, it appears to be weaker than other sub-periods. This might be because many investors had already purchased the coins at the earlier stage and were more likely to hold onto them during this time because of the positive expectations of an upward trend in the future. Thus, they barely purchased at this time, instead, the herding might come from them executing a sell action. Indeed, we notice that there was a downturn phase in this sub-period when the market significantly shrank. Hence, it is possible that a portion of investors began selling off their investments as they either felt panicked by the price drop or were motivated to take profits, resulting in a period of herding [130, 132, 259]. Our result when examining the cryptocurrency market in *up* and *down* conditions also

shows that only  $\gamma_4$  is negative during this phase, meaning that only the downturn time in the Bull market time reveals the herding behavior, which reinforces our hypothesis.

Our results of herding detection in the cryptocurrency market are in line with existing literature which also demonstrates evidence of herding during the first outbreak of the pandemic/ global economic crisis in 2020 and the periods of bullish markets, while there was no herding signal found before the pandemic and after the bullish time. [129, 132, 138, 260]. Remarkably, the herding behavior observed during these three sub-periods validates our inference in Section 5.3.4, where this phenomenon was used to explain the abnormal spikes in the volatility-based correlation graph during the peak of the Covid-19 pandemic and the Bull market time in 2021.

#### **5.5.4 An Extension to Traditional Markets**

Stocks are fundamental components of the global financial system, representing a substantial portion of market capitalization and serving as a key instrument for both individual and institutional investors. On the other hand, (US) ETFs act as proxies for their corresponding (US) indices, which represent the performance of a set of various stocks. This type of investment vehicle has become increasingly important<sup>16</sup>, making it a versatile and influential tool in modern investment strategies. The widespread use in investment portfolios of both stocks and US ETFs emphasizes their worth for inclusion in analysis and experiments. Therefore, we expand our study to these two types of investment vehicle. This experiment serves two purposes: Firstly, to broaden our herding analysis, not just be constrained to the cryptocurrency market, but also expand to traditional markets. Therefore, this study contributes to the literature with a more fulfilling picture of herding behavior in financial markets; Secondly, this experiment facilitates the comparison between young assets like cryptocurrencies and mature assets like stocks and US ETFs, help-

---

<sup>16</sup>As of September 2024, there are 2149 index-tracking ETFs, as reported by <https://www.morningstar.com/funds/will-active-etfs-outnumber-passive-etfs>.

ing us understand better the differences and similarities between them.

Table 5.17 shows the estimation of  $\beta_2$ ,  $\gamma_2$  and  $\gamma_4$  in the case of stocks and US ETFs. As before, only values with the star symbol are statistically significant. Herd behavior is revealed at a period only when its corresponding coefficient (either  $\beta_2$ ,  $\gamma_2$  or  $\gamma_4$ ) is statistically significantly negative. The more negative the value, the more pronounced the herding.

**Table 5.17: Herding detection results for stocks and US ETFs. Herding is pronounced in a period if the corresponding coefficient is statistically significantly negative. We consider three levels of confidence (1% (\*), 5% (\*\*)) and 10% (\*\*\*) to be significant**

Periods	Stocks			US ETFs		
	$\beta_2$	$\gamma_2$	$\gamma_4$	$\beta_2$	$\gamma_2$	$\gamma_4$
Pre-Covid-19	-8.94*	0.01	-9.51*	-4.550*	-2.064	-5.085*
Covid-19 pandemic	-2.81*	-3.05*	-2.70*	-0.995*	-1.297*	-0.998*
Post-Covid-19	10.32*	7.74*	9.88*	8.684*	9.528*	-0.330
Bull market	2.65	6.03**	2.62	31.398*	6.725*	33.658*
U-R conflict	1.44	1.98	1.44	0.155	0.464	-0.417

The parameter  $\beta_2$  is estimated from the equation:

$$CSAD_t = \alpha + \beta_1|r_{m,t}| + \beta_2r_{m,t}^2 + \epsilon_t.$$

The parameters  $\gamma_2$  and  $\gamma_4$  are estimated from the equation:

$$CSAD_t = \alpha + \gamma_1 D^{up}|r_{m,t}| + \gamma_2 D^{up}r_{m,t}^2 + \gamma_3 D^{down}|r_{m,t}| + \gamma_4 D^{down}r_{m,t}^2 + \zeta_t.$$

The herding behavior observed in stocks and US ETFs displays similar patterns, with only the first two sub-periods, namely Pre-Covid-19 and Covid-19 pandemic, having signs of herding behavior. These time intervals cover critical events that significantly impacted traditional markets, i.e. the US-China trade war during the year 2019 and the severe hit of the contagious health disease Covid-19 on the global economy and financial markets in the first half of 2020. Notably, the US-China trade war took place throughout 2019 and continually affected the economies of not only the US and China but also other countries and continents [47, 53]. Moreover, investors were also bombarded by both negative and positive news about this event. Consequently, such news dragged the markets up and down frequently [50]. Indeed, it has been shown that there was a wave of selling assets during this year [51]. This explains the more pronounced herding within the negative market return compared to the positive market return during the Pre-Covid-19, in both stock and US ETF

markets.

This experiment delivers two interesting results. Firstly, herding is more pronounced during the Pre-Covid-19 compared to the Covid-19 pandemic sub-period. Therefore, although the pandemic led to greater losses and fluctuations in traditional financial markets, it is noteworthy that investors exhibited more intense herding behavior before the pandemic, particularly during the turbulent time caused by the US-China trade war. One possible reason for this is the constant fluctuations of markets throughout 2019 when the trade war became more and more intense [48]. In particular, different tariffs and quotas were imposed between these two countries throughout this year, directly impacting financial markets worldwide. As a result, the constant release of negative and positive news gave rise to frequent market movements [50, 51]. Thus, investors actively reacted to market situations, causing herding behavior on a regular basis. The impact of the Covid-19 pandemic was severe and damaging within the second quarter of 2020 but this was followed by a gradual recovery. Hence, investors panicked during the peak of the pandemic but things quickly got back to normal, since they started to make objective investment decisions again, herding then lessened. Another possible reason is that investors in traditional markets are more influenced by macroeconomic factors as well as policy uncertainty. Specifically, US interest rates increased in response to rising inflation. Furthermore, the unemployment rate increased due to the trade war and GDP in 2019 also fell to its lowest since the 2008 crisis [109, 261, 262]. These macroeconomic factors had the potential to scare traditional market traders.

Secondly, the magnitude of herding behavior in the market composed of US ETFs is generally weaker than in the stock and cryptocurrency markets, as indicated by the lower herding coefficients  $(\beta_2, \gamma_2, \gamma_4)$  for US ETFs compared to those for stocks and cryptocurrencies across most sub-periods. This distinction is possibly driven by the stylized facts of US ETFs: the US market itself has a well-established economic environment and a high proportion of rational investors [111, 115]. As a result, since investors are more inclined to rely on their own analysis and expectations

to make investment decisions rather than mimicking other investors, the herding behavior is therefore reduced. This large number of rational investors may be a crucial factor in limiting herding behavior, since the herding frequency in stocks and US ETFs is lower than in cryptocurrencies, as evidenced by Tables 5.16 and 5.17. On the other hand, another stylized fact about ETFs is their low volatility (i.e. low risk) compared to other types of investment vehicle, which limits the potential for short-term profits. Therefore, investors tend to herd in other - more risky - types of investment vehicle to gain high profits while ETFs are preferable for long-term investments [263]. Remarkably, also in US ETFs, Tables 5.16 and 5.17 display an intense dispersion in price movements between these assets during the Bull market sub-period, with  $\beta_2$  exceeding 31.0 and  $\gamma_4$  nearing 34.0. This indicates a substantial divergence in investment decisions among investors in US ETFs during this time.

In summary, we found that investors in all three types of investment vehicle showed evidence of herding during the Covid-19 pandemic sub-period. This is reasonable since the serious pandemic directly affected the global economy and all financial markets. However, other sub-periods reveal different herding patterns between the cryptocurrency market and traditional markets. In particular, herding behavior was found only in stocks and US ETFs throughout 2019, while this phenomenon disappeared in the cryptocurrency market. Indeed, herding during this time appeared to be caused by the trade war between China and the US which was relevant to macroeconomic factors and directly influenced various companies in different sectors. Thus, traditional markets were undoubtedly affected the most. On the other hand, only the cryptocurrency market displayed a herding signal during the Bull market time. This can be explained by the fact that cryptocurrency investors are often swayed more by positive news and positive market returns than negative ones [127]. Whereas, traditional markets are typically more responsive to distressed market conditions than booms [108, 264]. In particular, when traditional markets experience an upward trend, professional investors are found to account for the majority portion [109]. Thus, they tend to make investment decisions by

themselves which prevents them from performing as a herd. For the cryptocurrency market, however, its investors are more likely to be naïve and make irrational decisions, based on word of mouth as well as existing news on social media [64, 188]. Thus, positive news surrounding cryptocurrencies can lure inexperienced investors with the promise of substantial profits.

### **5.5.5 Herding Behavior Across Cryptocurrencies, Stocks and US ETFs: A Look at Community Structures**

After detecting and analyzing herding behavior in each type of investment vehicle independently, we now conduct an experiment on a combination of cryptocurrencies, stocks and US ETFs. This idea stems from the fact that investors tend to select a variety of holdings over different asset classes instead of solely relying on one source to reduce their portfolio risk [255]. By running this experiment, we hope to contribute to the existing literature with a more complete and realistic picture of herding behavior in financial markets. However, making definitive conclusions about herding behavior across this broad combination is ambiguous. Moreover, there is currently no standardized approach to partition assets into distinct subsets for herding detection. To address this, we use the Louvain community detection algorithm to split the combination into smaller subsets of assets (communities) based on their correlation. Subsequently, we apply herding detection techniques to each community.

This practice is grounded in the belief that assets with similar price movements are more likely to reveal evidence of herding behavior among them, while assets with divergent price movements tend to serve as safe havens rather than being herded [265]. Thus, by dividing the assets in this manner, we can preserve the characteristics of portfolio diversification, improve the clarity and granularity of herding signals while mitigating the loss of real herding signals.

Figures 5.15-5.16 illustrate the community detection results on the combination of cryptocurrencies, stocks and US ETFs in each sub-period. The assets in each

community are listed in Appendix A. As can be seen, the community structures evolve over time, reflecting changes in correlations among assets. However, we found several consistent patterns across these five structures. First, cryptocurrencies form several communities that are next to each other, indicating strong correlations within the cryptocurrency market itself. However, there remains a clear separation between the cryptocurrency market and traditional markets, as most cryptocurrencies do not form communities with stocks and US ETFs, an example of this characteristic is highlighted within a purple-shaped oval in each figure. Second, stocks and US ETFs tend to belong to the same community, with a US ETF often acting as the central node, connected by stock assets, suggesting that it is the most influential node within the community, an example of this characteristic is highlighted by a black-shaped oval in each figure. Third, there are some stock sectors in which stocks within the same sector and ETFs that contain these stocks tend to be in the same communities, namely Energy, Healthcare, Information Technology and Financials. On the other hand, this feature does not appear in the remaining sectors. Lastly, we also notice that S&P-related ETFs consistently form a community throughout the sub-periods analyzed. These observations open up potential aspects that our herding detection experiments on a combination of cryptocurrencies, stocks and US ETFs can focus on.

Table 5.18 provides a quantitative analysis of the allocation of stocks, US ETFs and cryptocurrencies in the correlation graphs shown in Figures 5.15 and 5.16. In particular, we calculate the average path length (i.e. the average number of edges) between two assets across six scenarios: 1) between two stocks, 2) between two US ETFs assets, 3) between two cryptocurrencies, 4) between one stock and one US ETF, 5) between one stock and one cryptocurrency, and 6) between one US ETF and one cryptocurrency. As can be seen in the table, the average path length between two cryptocurrencies is consistently the shortest in all sub-periods, suggesting that cryptocurrencies tend to be close to each other. In contrast, the average path length between a cryptocurrency and a traditional asset (stock or US ETF) is much longer

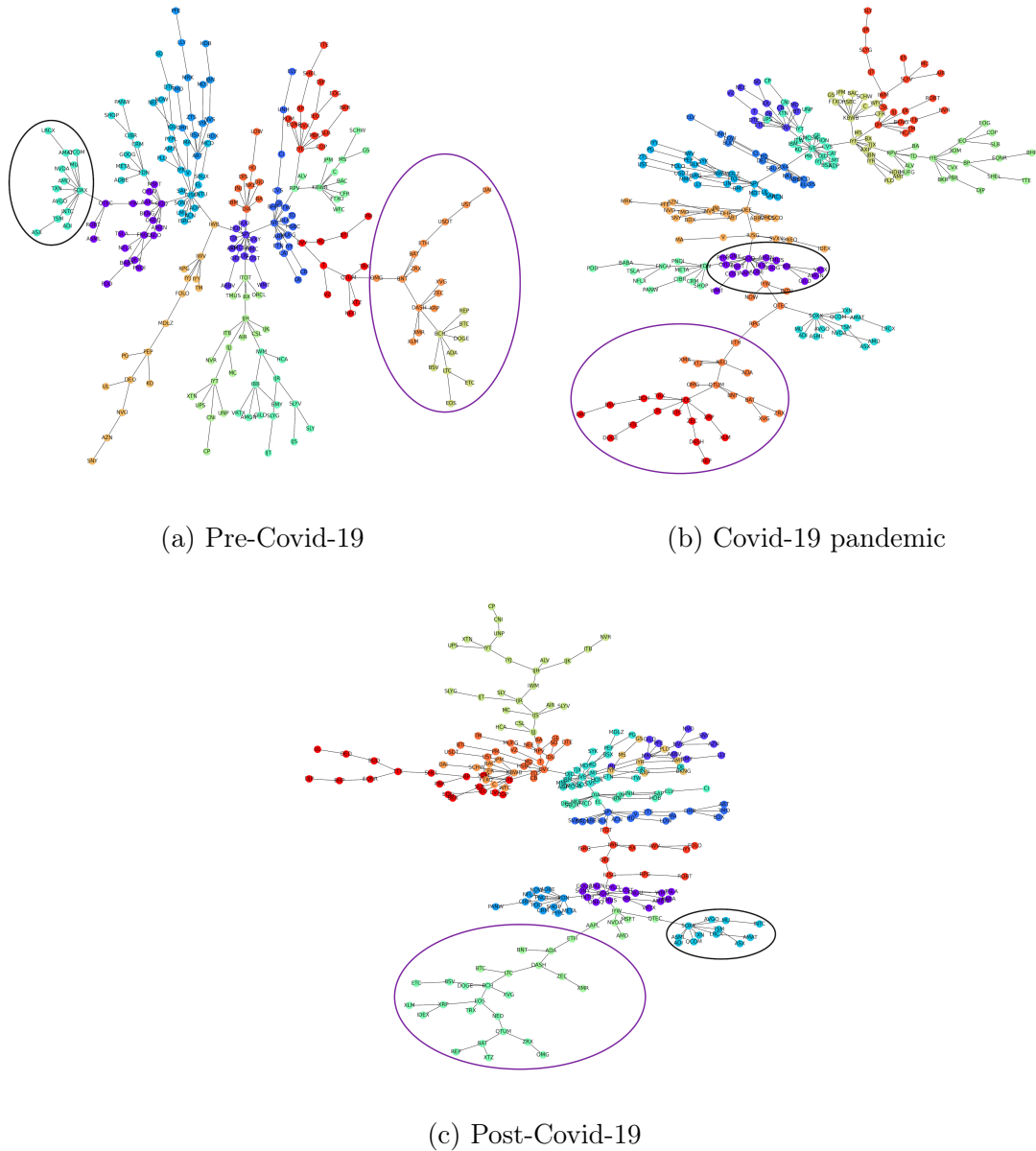


Figure 5.15: **Time-varying community structures of cryptocurrencies, stocks and US ETFs during: a) Pre-Covid-19, b) Covid-19 pandemic, c) Post-Covid-19.**

than that between a stock and a US ETF. This reveals that cryptocurrencies remain separated from most stocks and US ETFs, while stocks and US ETFs tend to be close to each other due to their small average path length. In general, this table supports and reinforces the observations discussed in the previous paragraph about the correlation graphs.

From the communities detected in each sub-period, we then apply the herding detection algorithms to each community. The results of this experiment are discussed

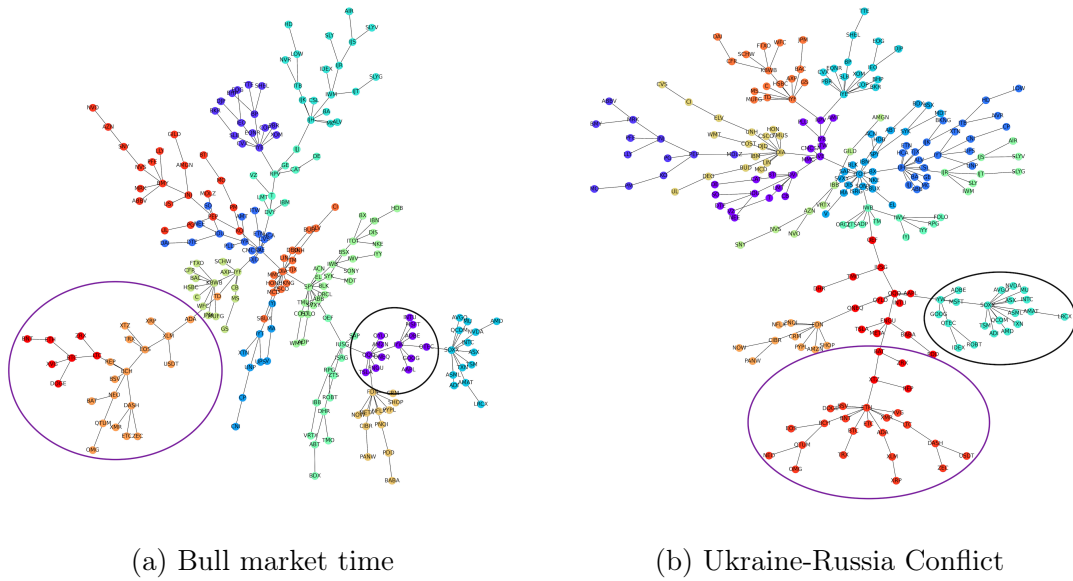


Figure 5.16: **Time-varying community structures of cryptocurrencies, stocks and US ETFs during: a) Bull market period, b) Ukraine-Russia conflict.**

Table 5.18: **Average path length (number of edges) between two assets in six cases, namely between stock and stock, between US ETF and US ETF, between cryptocurrency and cryptocurrency, between stock and US ETF, between stock and cryptocurrency, and between US ETF and cryptocurrency.**

Period	Stock	ETF	Crypto	Stock	Stock	ETF
	vs	vs	vs	vs	vs	vs
	Stock	ETF	Crypto	ETF	Crypto	Crypto
Pre-Covid-19	7.00	5.60	3.70	6.43	11.15	10.63
Covid-19	6.73	6.07	4.83	6.45	12.31	12.37
Post-Covid-19	8.41	8.02	7.27	8.42	14.91	14.65
Bull Time	7.76	6.56	4.68	7.27	11.26	11.02
U-R Conflict	7.07	5.40	4.46	6.32	11.28	10.09

in the next section.

### 5.5.6 Herding Behavior Across Cryptocurrencies, Stocks and US ETFs: Unveiling Results

This section analyzes the herding detection results on a combination of cryptocurrencies, stocks and US ETFs, at a community level. To enhance the detail of herding analysis, we categorize the assets by their business sector, either cryptocurrency, or

US ETF, or one of the 11 stock sectors, as discussed in Section 3.1.2. Therefore, each asset in a community identified by the community detection algorithm is assigned to one of these sectors and the sector distribution within the community is then calculated, indicating the proportion of each sector present. This allows us, for each community in each sub-period, to determine whether herding behavior exists and to analyze the sector distribution associated with that behavior.

We note that due to a large number of herding results, it is challenging to compare general CSAD (Equation 4.7), CSAD for *up* and CSAD for *down* market conditions (Equation 4.8) in each community, between different communities and between different sub-periods. Therefore, we do not differentiate the results of general herding, herding during the *up* and herding during the *down* market condition. Instead, we focus solely on determining whether the communities reveal herding behavior or not. To this end, we consider a community to exhibit herding when either CSAD or CSAD for *up* market condition or CSAD for *down* market condition reveals herding. In addition, the number of assets in each stock sector is different and depends on its assets' market capitalizations. For instance, Technology, Finance and Healthcare stocks tend to have very high market capitalization while Communication Services and Real Estate tend to have lower market capitalization [146]. Consequently, among the 11 stock sectors, Utilities, Materials, Real Estate and Communication Services exhibited a limited number of entities. Therefore, this study concentrates only on the remaining 7 sectors. Details for this classification are provided in Section 3.1.2.

Figures 5.17-5.21 display the herding detection results and sector distribution of each community in each sub-period. Assets in each community are listed in Appendix A. Regarding herding detection figures, the grey bar means the herding based on Equation 4.7 (does not distinguish *up* and *down* markets) presents, the green bar means the herding based on Equation 4.8 for the *up* market condition presents, the red bar means the herding based on Equation 4.8 for *down* market condition presents. The higher the bar, the more pronounced the herding behavior. Regarding the sector distribution figures, sectors are listed on the *y*-axis, each represented by a

unique color. In each community, the longer the bar of a corresponding sector, the greater the sector's percentage within that community.

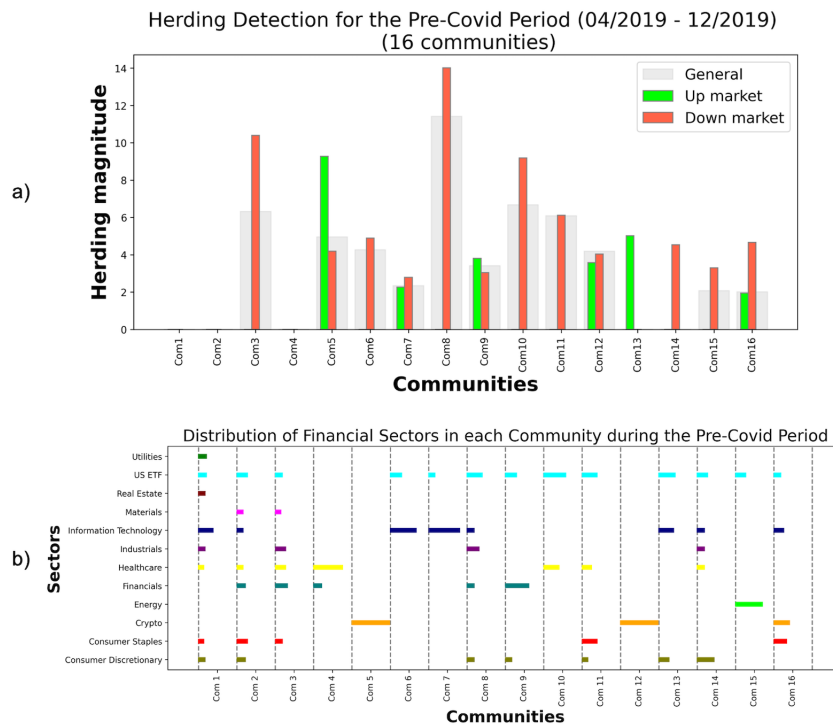


Figure 5.17: Herding detection results (a) and sector distribution (b) of each community during the Pre-Covid-19 period.

The first two sub-periods appear quite similar concerning the density of herding signals, with most communities exhibiting herding behavior (Figures 5.17-5.18). This is contrary to the last three sub-periods since these reveal a great reduction in herding (Figures 5.19-5.21). Perhaps, the high level of herding during the pre- and peak of the pandemic was attributed to two major events, as mentioned previously, the US-China trade war throughout 2019 and the outbreak of Covid-19 in early 2020, respectively. However, as the pandemic came under control and the global economy began to recover, financial markets resumed their normal behavior. Therefore, investors moved away from mimicking each other and made more independent investment decisions, the herding was lessened accordingly.

We observe that cryptocurrencies remain somewhat isolated from traditional markets, as they cluster together, typically forming one to two large communities next to each other, while stocks and US ETFs stay separate from these communi-

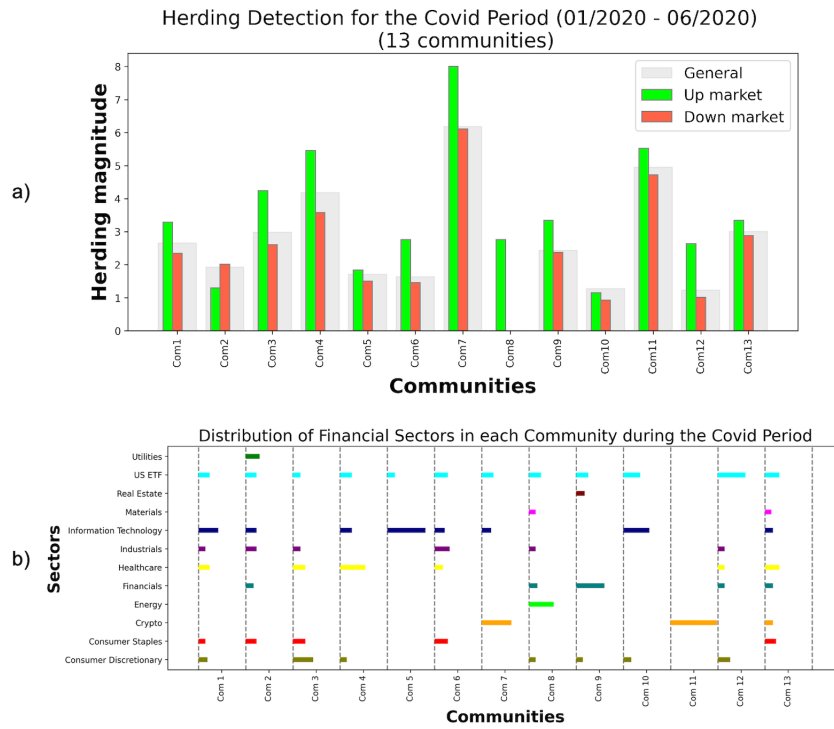


Figure 5.18: Herding detection results (a) and sector distribution (b) of each community during the Covid-19 pandemic period.

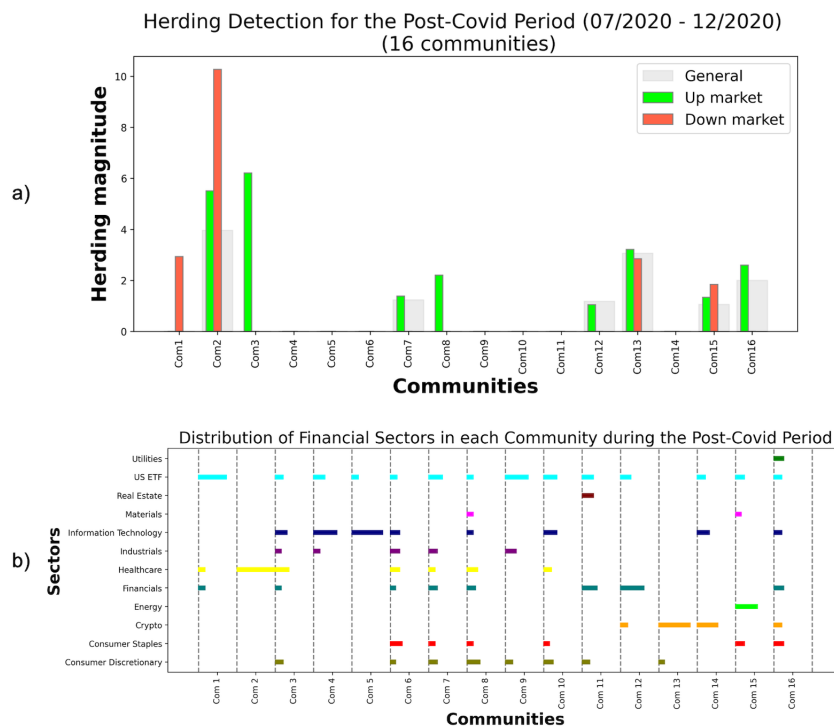


Figure 5.19: Herding detection results (a) and sector distribution (b) of each community during the Post-Covid-19 period.

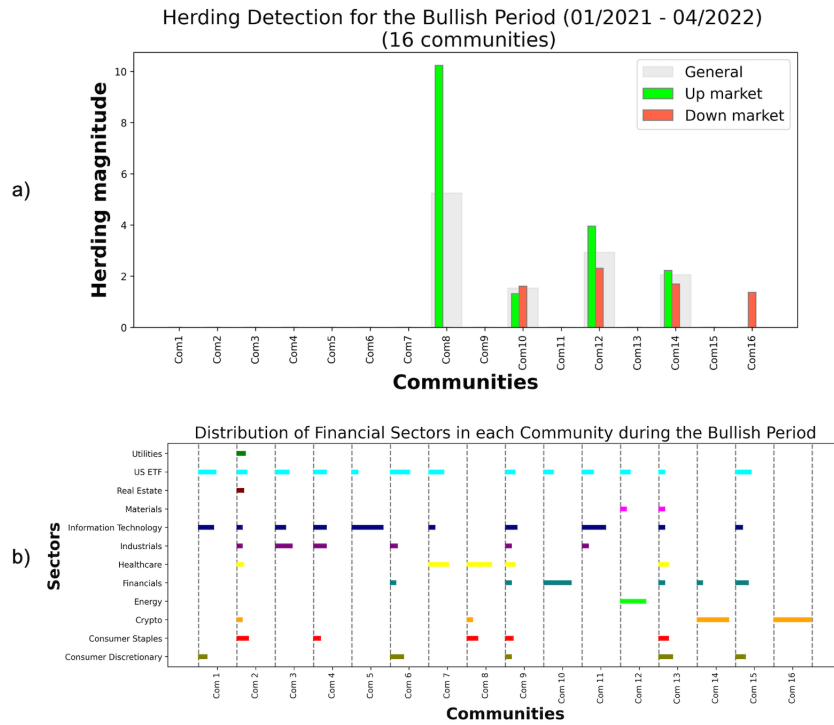


Figure 5.20: Herding detection results (a) and sector distribution (b) of each community during the Bull market period.

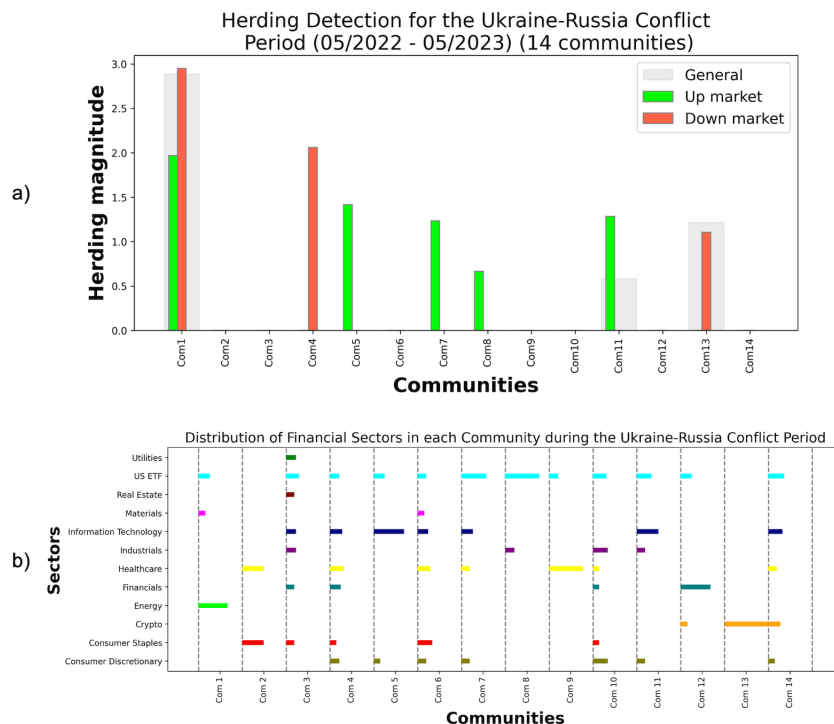


Figure 5.21: Herding detection results (a) and sector distribution (b) of each community during the Ukraine-Russia conflict period.

ties. Looking at different sub-periods, cryptocurrency communities exhibit evidence of herding most of the time. Compared to the herding detection results in the cryptocurrency market as discussed in Section 5.5.3, it is evident that without critical events impacting the entire cryptocurrency market, herding still exists in most cryptocurrencies, but at a community level, i.e herding behavior is pronounced in different subsets of cryptocurrencies, while this phenomenon is absent when looking at all cryptocurrencies at once. Identifying the causes behind this herding is challenging, given the lack of major events influencing the market during those periods [128, 129, 266]. However, it appears that this market is largely driven by investor sentiment [136] and news circulating on online platforms [132]. Additionally, cryptocurrency investors often seem to follow short-term strategies, focusing on price movements rather than the fundamental factors that typically drive traditional markets [130]. Consequently, herding within cryptocurrency communities could be triggered by either the prospect of quick profits or the fear of losses, prompting investors to buy or sell. Moreover, investors might be aware of co-movements between different cryptocurrencies. Thus, when the opportunity to gain profits (or the selling time) arrives, they tend to buy (or sell) the coins that have the same price movements to get higher profits (or mitigate the loss). This could explain why cryptocurrencies formed two distinct communities during the Pre-Covid-19 and bull market time, both showing herding behavior, while no herding was detected across the entire cryptocurrency market in the earlier analysis.

In the other communities, each of them has a mixture of US ETFs and stocks. Herding behavior in these communities follows a similar rule to that seen in cryptocurrency communities since herding signals are present in several communities in each sub-period. Notably, we have uncovered valuable insights from the herding results in stocks and US ETFs. That is, there are four stock sectors, namely Energy, Healthcare, Information Technology and Financials, where the assets and their corresponding ETFs tend to form distinct communities within each sector, throughout the examined sub-periods. Furthermore, the herding signals detected in each of

these sectors align with the timing of economic and political events that specifically impact that sector.

- **Energy sector:** In our case, the energy stocks we examine primarily consist of petroleum treatment and petroleum exploitation firms. Therefore, the herding behavior detected in this sector tends to be mainly driven by oil-related events. We have found that all energy stocks and corresponding ETFs are consistently exposed to herding across all sub-periods. Over the past four years, from 2019 to date, energy companies have faced significant challenges, enduring a series of events that caused substantial fluctuations in their stock prices. These include the US-China trade war, the Covid-19 pandemic, the petroleum oversupply caused by the shale revolution [39, 82], the Russia-Saudi Arabia price war [40, 267] and the Ukraine-Russia war [73].
- **Technology sector:** Technology stocks are grouped into several communities and reveal herding behavior in three sub-periods: Pre-Covid-19 (coinciding with the US-China trade war), Covid-19 pandemic (coinciding with the downturn time caused by the pandemic followed by the global economic recession) and the Ukraine-Russia conflict (aligning with the time when technology companies were laying off workers [41, 85, 268] followed by the AI boom that impelled the Technology sector [269, 270]).
- **Healthcare sector:** Herding behavior was detected in most healthcare stocks and ETFs across three sub-periods, including the Covid-19 pandemic, the Post-Covid-19 and the Bull market time. This herding is likely attributed to heightened investor interest in pharmaceutical companies, due to their strong performance during the pandemic. In particular, many pharmaceutical companies initiated vaccine research as early as 2020 to address global vaccination demand [271–273]. Additionally, other pharmaceutical companies, though not involved in vaccine production, played a vital role by producing necessary components for vaccine production such as nanoparticle reagents or supplying

medical devices for hospitals and daily usage such as the Covid test kits. These contributions pushed their performance during this time [273, 274].

- **Financials sector:** In this experiment, the Financials sector primarily comprises banking companies. We found that this industry was exposed to herding behavior in the first four sub-periods, with the exception of the Ukraine-Russia conflict time. This sector is closely tied to economic conditions and mirrors the performance of the economy, especially banking firms since they play a key role in preserving economic stability and preventing financial crises, through various measures like lowering interest rates [275], imposing new monetary policy comes from the US Federal Reserve [276] and cutting companies' liquidity risks [277]. As a result, economic uncertainty tends to place the Financials sector under distress. Indeed, the global economy has been experiencing ups and downs continually during the periods of herding observed in this stock sector.

For the remaining sectors (i.e. Industry, Consumer Discretionary and Consumer Staples), we noticed that assets within a sector do not group into a large community in any sub-periods. Instead, they are distributed almost evenly across multiple communities. This result indicates that investors did not herd on a sectoral basis in these sectors, even during major events. Interestingly, however, we have observed simultaneous herding behavior between the Technology sector and Standard & Poor (S&P)-related ETFs, as shown in Figure 5.17 (Communities 6 and 7 for Technology stocks, Community 10 for S&P-related ETFs), Figure 5.18 (Community 5 for Technology stocks, Community 12 for S&P-related ETFs) and Figure 5.21 (Community 5 for Technology stocks, Community 8 for S&P-related ETFs). Specifically, when the herding behavior is pronounced in the Technology sector, a set of ETFs tracking small-cap stock indices established by the S&P Global Ratings also show this phenomenon, including IJR, IJS, IJT, IWM, SLYS, SLYV and SLY. This finding is surprising since all the technology stocks studied here are large-cap but their behavior is unexpectedly associated with ETFs tracking the performance of *small-cap*

stocks. This relationship between the Technology sector and the aforementioned ETFs will be explored in future research.

Furthermore, we acknowledge that some communities are formed from multiple sectors and reveal herding behavior, but the reasons for this are unclear as there does not seem to be a clear pattern or convention in these cases. However, most of such communities appear during economic uncertainty times, i.e. the trade war (Figure 5.17) and the Covid-19 pandemic (Figure 5.18). By contrast, this phenomenon was not pronounced from the Bull market time onwards (Figures 5.20-5.21). This suggests that negative market conditions may have contributed to the herding behavior. However, since we do not have a confident justification for this phenomenon, we will leave this result for future exploration.

### **5.5.7 Remarks**

This section detects and analyzes investor herding behavior in different types of investment vehicle separately and also in their combination, including stocks, US ETFs and cryptocurrencies. We enrich the existing literature with three contributions: Firstly, the use of a more up-to-date and granular historical price time series covering several recent events, i.e. the US-China trade war in 2019, the Covid-19 pandemic followed by the global economic crisis in 2020, the Bull market period in 2021-2022 and the Ukraine-Russia conflict which started in 2022. Secondly, the use of three types of investment vehicle, which is rarely conducted by previous studies since they tend to focus on one asset class only. Thirdly, the focus on herding behavior at a community level (small scale) enhances the clarity and granularity of this herding phenomenon.

When examining the herding behavior in each type of investment vehicle, we found similar herding patterns between stocks and US ETFs, with evidence of herding during the Pre-Covid-19 period, coinciding with the US-China trade war, and also during the Covid-19 pandemic. On the other hand, the cryptocurrency market displays quite distinct behavior, with herding occurring during the Covid-19 pan-

demic, the Post-Covid-19 when the market was recovering and the Bull market time when prices surged. In addition, the magnitude of herding in US ETFs tends to be weaker than in stocks and cryptocurrencies.

At the community level, while herding might not be present in a whole market composed of all assets, it might appear within specific subsets of assets. Indeed, each type of investment vehicle studied here revealed herding behavior in certain communities during each sub-period. This is typically driven by specific events impacting those communities, rather than the entire market. Moreover, we notice 4 stock sectors, namely Energy, Healthcare, Information Technology and Financials, in which, for each sector, its assets and relevant ETFs tend to form communities among themselves, with herding behavior driven by sector-specific events.

This work reinforces our hypothesis and interpretation of the results obtained in previous sections, where we used herding behavior to explain the changes in the correlations among cryptocurrencies. Furthermore, we also raise investors' attention towards 4 stock sectors, i.e. Energy, Healthcare, Information Technology and Finance as these seem more sensitive to herding behavior than other sectors. Additionally, from this work, herding behavior should be examined from two perspectives: the entire market and individual subsets (communities) within the market. This is because different perspectives reveal different insights into the herding behavior.

Last but not least, the distinction of the cryptocurrency market from traditional markets shown in this section suggests a possibility to use cryptocurrencies as hedges in a portfolio against other asset classes. However, investors need to keep in mind the potential risks from using cryptocurrencies as hedges due to their fundamental characteristics. In particular, cryptocurrencies are famous for their high volatility, especially during periods of market uncertainty [278]. This high volatility undermines their reliability as stable hedges in a portfolio. For example, the rapid and unpredictable price changes in the cryptocurrency market can lead to significant losses [279]. Consequently, using cryptocurrencies as hedges turns out to add more risks to the portfolio [280]. Moreover, it is shown that cryptocurrency markets ex-

hibit periods of potential arbitrage opportunities across exchanges [281], indicating inefficiencies that can affect their hedging capabilities. Another risk comes from the speculative nature of cryptocurrencies where they are often seen as a bubble without fundamental value [282], leading to concerns about their long-term viability as hedges. The lack of regulations in the cryptocurrency market further exacerbates these risks, as highlighted by Kouam in [283], who argues that cryptocurrencies expose economies to grave economic and financial stability risks. To this end, while cryptocurrencies may offer some hedging benefits, their speculative nature, extreme volatility, and lack of regulation represent substantial challenges that need to be carefully considered.

# Chapter 6

## Conclusion

This study examines the time-varying correlation between cryptocurrencies of different market caps followed by the exploration of the sentiments (e.g. fear, neutral, greed, etc) and investment behavior (e.g. herding behavior, contrarian behavior, etc) of investors behind this correlation, across different market conditions such as stable, recessive, bull and bear markets. Furthermore, an expansion to traditional markets (i.e. stocks and US ETFs) is carried out to gain a better understanding of this emerging digital market in particular and the financial markets in general. Throughout this study, four research questions are covered with several interesting findings. This chapter will revisit our aforementioned research questions and corresponding findings. Moreover, we will also identify the limitations of this study and potential works that can be developed in the future.

### 6.1 Research Questions and Key Findings

We first examine the returns-based graph structure of cryptocurrencies in different market conditions, from stable time (i.e. before the Covid-19 pandemic), to during the peak of the Covid-19 pandemic, to the recovery time and explored the investment behavior of investors behind each graph structure. We find that the significant shift in graph structures between stable and turbulent times highlights the changes in the investment decisions of investors. During stable periods, investors rely on their

own market analysis and experience to make independent investment choices. This leads to a diversified selection of cryptocurrencies in their portfolios, including both high- and low-ranking assets. As a result, the graph structure of cryptocurrencies during stable times has been found to be scattered with many small communities, composed of both low- and high-ranking assets. However, during periods of high market volatility like the peak of the pandemic followed by the global economic crisis, investors typically focus on trading cryptocurrencies with large market capitalizations, while smaller cryptocurrencies are primarily used for other functions, such as paying transaction fees, acting as tokens for smart contracts, or supporting digital platforms. Consequently, the graph structure in this scenario becomes condensed and splits into just two distinct communities: high-ranking assets and low-ranking assets.

Our second research question is developed upon the first research question by examining the volatility-based graph structure of cryptocurrencies. We hypothesize that the volatility indicator can enrich the understanding of the cryptocurrency market, providing new insights that can not be discovered from the return indicator. We find that the volatility-based graph structures deliver unique insights from the return ones. Moreover, return information appears more suitable for general analyses aimed at understanding fundamental characteristics where a high level of detail is unnecessary. In contrast, volatility information might be better suited for highly detailed examinations. Additionally, volatility-based graphs show the potential to predict major events in the cryptocurrency market, such as economic crises, bull and bear trends. In this research question, we also examine the connection between investors' sentiments and the changes in volatility-based graph structures. Our findings confirm a strong relationship between them. In particular, the volatility correlation between cryptocurrencies rises when investors overreact with extreme fear or greed, leading to a compressed graph with one large community. On the other hand, when investor sentiment is more balanced, the correlation weakens, and the graph becomes scattered, forming several communities of small size.

Regarding the third research question, we detect herding behavior in the cryptocurrency market, across 5 sub-periods from 2019 to 2023, including Pre-Covid-19, Covid-19 pandemic, Post-Covid-19, Bull market time, and the Ukraine-Russia conflict time. Out of them, the cryptocurrency market exhibits herding behavior during three sub-periods, including the Covid-19 pandemic, Post-Covid-19 and Bull market time. From these results, typical characteristics of cryptocurrency investors are captured. That is, they are easily influenced by market fluctuations and breaking news, leading to similar investment decisions, eventually causing the herding behavior among themselves. Moreover, this market seems to be more sensitive to positive events such as rising prices and positive news, than negative ones. Notably, the herding behavior seen across these three sub-periods supports our inference from the second research question, where this phenomenon was used to explain the spikes in the volatility-based correlation graph during the peak of the Covid-19 pandemic and the 2021 bull market, which are equivalent to the periods of strongest correlation among cryptocurrencies as well as the most significant changes in the graph structure.

Lastly, our fourth research question extends to traditional assets by comparing herding behavior in cryptocurrencies with stocks and US ETFs. This serves two purposes. Firstly, to highlight the similarities and differences in herding behavior between cryptocurrencies and traditional assets. Secondly, to examine herding behavior in a combination of different types of investment vehicle, including cryptocurrencies, stocks and US ETFs, which is useful for investors valuing portfolio diversification. We find similar herding patterns in stocks and US ETFs across all sub-periods. However, only the Covid-19 pandemic period reveals the herding signals in these traditional assets and also cryptocurrencies. Apart from that, stocks and US ETFs only exhibit herding signals during the Pre-Covid-19 period, aligning with turbulent times caused by the US-China trade war. These findings suggest that herding behavior in cryptocurrencies differs from that in traditional assets. Such differences can be attributed to the differences in underlying characteristics

between them. Specifically, traditional markets are more sensitive to distressed market conditions and macroeconomic factors like interest rates, inflation rates and GDP. By contrast, cryptocurrencies are more sensitive to positive market conditions and strongly influenced by investors' sentiments. The distinction of cryptocurrencies from stocks and US ETFs is reinforced by executing the herding detection on the combination of these assets. To this end, even though the cryptocurrency market is steadily becoming more similar to traditional markets, it still has a long way to go before reaching the same level of maturity.

## 6.2 Limitations

Although the datasets used in this research are larger than most of the existing studies, the number of assets in both cryptocurrency and traditional markets should be ideally higher so that we can make firmer conclusions from our experiments. Secondly, while 30 min granularity has been found to suffice for our calculations, it would be better if we could use a lower level, say 15 min or even finer, so that the information of the time series can be retained as much as possible. Unfortunately, we cannot work at these finer granularities since some assets, mostly cryptocurrencies are not traded sufficiently regularly, causing a large number of missing values at these shorter timescales.

Furthermore, given the popularity and wide use of Pearson correlation measurement, there is also a concern about this metric due to its drawbacks. In particular, it is sensitive to outliers and cannot capture non-linear relationships that might cause misleading results. However, none of the current correlation metrics are comprehensive. Instead, each of them has its own weakness. Therefore, the creation of new approaches for calculating correlation coefficients that overcome the current limitations needs to receive more attention.

Regarding Minimum Spanning Tree (MST), it carries a few limitations. Firstly, this graph is quite sensitive to new data, meaning that its structure might change to some extent when a new asset is added to the graph. Thus, to ensure the consis-

tency of the experimental results, researchers can utilize different graph construction methods and focus on common results between them. Moreover, another notable limitation of the approach is its susceptibility to instability and unreliability during periods of high market correlation, such as global market declines. In such periods, most correlation coefficients among assets have comparable magnitude, with differences often at the noise level. This poses a risk for the MST algorithm as it may promote a wrong candidate to the role of a central hub, thereby creating a misleading representation of the system's structure.

Regarding the noise removal, we recall that there are two main types of noise, including inherent noise and measurement noise. Our denoising approach, which relies on Random Matrix Theory, addresses both types of noise but does not explicitly separate them. Moreover, one downside of Random Matrix Theory is that it is sensitive to the length of time series. Specifically, increasing the length of the time series for the same dataset results in an issue where the possibility for the eigenvalues to fall within the Marcenko-Pastur distribution is zero, Random Matrix Theory thereby fails to identify the noisy eigenvalues. On the other hand, if the length of time series is not very large compared to the number of assets, the empirical correlation matrix might be dominated by measurement noise. Consequently, there might be collective effects in the system which are hidden and masked by the measurement noise.

### 6.3 Future Works

Understanding how assets across different types of investment vehicle are correlated with each other has the potential to shed light on portfolio optimization. Based on the outcome of this study, we can take one step further by constructing and comparing the portfolio optimizations at different market conditions. Therefore, unique characteristics of an optimized portfolio at each market condition, in theory, be learned and analyzed. Moreover, we have noticed that the graph structure of low-frequency data behaves differently from that of high-frequency data. Hence, it

is possible to learn the long-term characteristics of cryptocurrencies based on this structure, which could be potentially beneficial for investors who choose to make a long-term investment decision. Apart from that, future work can expand this study's concepts to other types of investment vehicle, such as commodities, regular currencies and bonds.

As highlighted in the Limitations section (Section 6.2), the use of Pearson correlation as the primary correlation metric has several drawbacks. Therefore, we plan to explore several alternative metrics such as Spearman, Kendall's tau and Wavelet coherence. Unlike Pearson, Spearman and Kendall's tau are non-parametric rank-based measures. By ranking the data points, they reduce the influence of outliers, making it more robust in the presence of extreme values. This is particularly useful in the cryptocurrency market, which exhibits heavy tails and extreme price movements. On the other hand, Wavelet coherence provides a different approach by analyzing time-frequency correlations. Specifically, it captures how correlations between two time series evolve over time and across different frequencies. This metric shows its uniqueness from the previous ones by examining a system at multiple time scales at the same time, allowing for the analysis of both short-term and long-term correlations within a system. Ideally, we would like to run our experiments with these alternative metrics and compare the results to each other.

Regarding herding behavior detection, several findings in this experiment warrant further investigation: Firstly, the differences in herding magnitude between different communities. Secondly, the comparison between CSAD, CSAD for up and CSAD for down markets within each community, between different communities and between different sub-periods. Thirdly, the connections between ETFs tracking the S&P-related indices and technology stocks.

While this study focuses on several factors influencing the investment decision process of an investor such as news, market conditions and personal analysis, future works could include other constituents that might influence investors' decision such as their personal financial conditions as well as cognitive bias (e.g. tendencies,

inclinations, etc).

Lastly, regulatory actions such as bans and taxation and technological risks like security breaches, hacking events, and blockchain outages could significantly distort the development of cryptocurrencies, affecting investment strategies of investors. Therefore, future works could conduct simulation models such as multi-agent systems to simulate the influence of regulatory and technological shocks on the cryptocurrencies as well as the correlation between cryptocurrencies and other asset classes.

# Bibliography

- [1] John Hyatt. *The Top 11 Crypto Investors in 2024*. <https://www.valuewalk.com/top-11-crypto-investors/>. ValueWalk, 2024 (cit. on p. 3).
- [2] Nathan Reiff. *How Much of All Money Is in Bitcoin?* <https://www.investopedia.com/tech/how-much-worlds-money-bitcoin>. Investopedia, 2024 (cit. on p. 3).
- [3] Josh Howarth. *How Many People Use Cryptocurrency?* <https://explodingtopics.com/blog/cryptocurrency-stats>. Exploding Topics, 2024 (cit. on p. 3).
- [4] Chad Albrecht et al. “The use of cryptocurrencies in the money laundering process”. In: *Journal of Money Laundering Control* 22.2 (2019), pp. 210–216 (cit. on p. 3).
- [5] Arturas Sabalionis, Wenbo Wang, and Hail Park. “What affects the price movements in Bitcoin and Ethereum?” In: *The Manchester School* 89.1 (2021), pp. 102–127 (cit. on p. 3).
- [6] Thomas Dimpfl and Franziska J Peter. “Nothing but noise? Price discovery across cryptocurrency exchanges”. In: *Journal of Financial Markets* 54 (2021), p. 100584 (cit. on pp. 3, 10, 52).
- [7] Leonidas Sandoval Junior and Italo De Paula Franca. “Correlation of financial markets in times of crisis”. In: *Physica A: Statistical Mechanics and its Applications* 391.1-2 (2012), pp. 187–208 (cit. on pp. 4, 16).
- [8] David Yermack. “Is Bitcoin a real currency? An economic appraisal”. In: *Handbook of digital currency*. Elsevier, 2015, pp. 31–43 (cit. on pp. 4, 16, 18).

- [9] Azza Bejaoui et al. “Connectedness between emerging stock markets, gold, cryptocurrencies, DeFi and NFT: Some new evidence from wavelet analysis”. In: *Physica A: Statistical Mechanics and its Applications* 619 (2023), p. 128720 (cit. on pp. 4, 17).
- [10] Mehmet Balcilar, Huseyin Ozdemir, and Busra Agan. “Effects of COVID-19 on cryptocurrency and emerging market connectedness: Empirical evidence from quantile, frequency, and lasso networks”. In: *Physica A: Statistical Mechanics and its Applications* 604 (2022), p. 127885 (cit. on pp. 4, 18, 77).
- [11] Imen Omri. “Directional predictability and volatility spillover effect from stock market indexes to Bitcoin: evidence from developed and emerging markets”. In: *The Journal of Risk Finance* ahead-of-print (2023) (cit. on pp. 4, 18).
- [12] Marcin Wątarek, Jarosław Kwapien, and Stanisław Drożdż. “Cryptocurrencies are becoming part of the world global financial market”. In: *Entropy* 25.2 (2023), p. 377 (cit. on pp. 4, 18, 19).
- [13] Steve Hyun et al. “What coins lead in the cryptocurrency market: using Copula and neural networks models”. In: *Journal of Risk and Financial Management* 12.3 (2019), p. 132 (cit. on p. 5).
- [14] Moinak Maiti, Yaroslav Vykylyuk, and Darko Vuković. “Cryptocurrencies chaotic co-movement forecasting with neural networks”. In: *Internet Technology Letters* 3.3 (2020), e157 (cit. on pp. 5, 25).
- [15] Dag Tjøstheim, Håkon Otneim, and Bård Støve. “Statistical dependence: Beyond Pearson’s  $\rho$ ”. In: *Statistical science* 37.1 (2022), pp. 90–109 (cit. on p. 5).
- [16] Stanisław Drożdż et al. “Complexity in economic and social systems: Cryptocurrency market at around COVID-19”. In: *Entropy* 22.9 (2020), p. 1043 (cit. on pp. 5, 13, 19, 78, 114).

- [17] Chun-Xiao Nie. “Analysis of critical events in the correlation dynamics of cryptocurrency market”. In: *Physica A: Statistical Mechanics and its Applications* 586 (2022), p. 126462 (cit. on pp. 5, 69, 78).
- [18] Hiba Rashid Atiya and Huda Naji Nawaf. “Prediction of Link Weight of bitcoin Network by Leveraging the Community Structure”. In: *IOP Conference Series: Materials Science and Engineering*. Vol. 928. 3. IOP Publishing, 2020, p. 032080 (cit. on pp. 5, 21).
- [19] Chun-Xiao Nie. “A network-based method for detecting critical events of correlation dynamics in financial markets”. In: *Europhysics Letters* 131.5 (2020), p. 50001 (cit. on pp. 5, 6, 16, 17, 20).
- [20] J-P Onnela et al. “Dynamic asset trees and Black Monday”. In: *Physica A: Statistical Mechanics and its Applications* 324.1-2 (2003), pp. 247–252 (cit. on p. 5).
- [21] Longfeng Zhao, Wei Li, and Xu Cai. “Structure and dynamics of stock market in times of crisis”. In: *Physics Letters A* 380.5-6 (2016), pp. 654–666 (cit. on pp. 5, 6, 19).
- [22] Stanley Lieberman. “Limitations in the application of non-parametric coefficients of correlation”. In: *American Sociological Review* 29.5 (1964), pp. 744–746 (cit. on p. 5).
- [23] Clark Wissler. “The Spearman correlation formula”. In: *Science* 22.558 (1905), pp. 309–311 (cit. on p. 5).
- [24] Herve Abdi and Paul Molin. “Lilliefors/Van Soest’s test of normality”. In: *Encyclopedia of measurement and statistics* (2007), pp. 540–544 (cit. on p. 5).
- [25] Harshal Chaudhari and Martin Crane. “Cross-correlation dynamics and community structures of cryptocurrencies”. In: *Journal of Computational Science* 44 (2020), p. 101130 (cit. on pp. 5, 20, 69).
- [26] Marcin Wałtorek et al. “Multiscale characteristics of the emerging global cryptocurrency market”. In: *Phys. Rep* 901 (2021), pp. 1–82 (cit. on pp. 6, 17).

- [27] Theophilos Papadimitriou, Periklis Gogas, and Fotios Gkatzoglou. “The evolution of the cryptocurrencies market: A complex networks approach”. In: *Journal of Computational and Applied Mathematics* 376 (2020), p. 112831 (cit. on p. 6).
- [28] Albert Molnar and Patrik Viktor. “The US-China Trade War: A Constructive Analysis of the Causes and Methods of De-Escalation”. In: *The Eurasia Proceedings of Educational and Social Sciences* 32 (2023), pp. 1–7 (cit. on pp. 6, 12).
- [29] Marco Ciotti et al. “The COVID-19 pandemic”. In: *Critical reviews in clinical laboratory sciences* 57.6 (2020), pp. 365–388 (cit. on pp. 6, 12, 13).
- [30] Rima Assaf, Deeksha Gupta, and Rahul Kumar. “The price of war: Effect of the Russia-Ukraine war on the global financial market”. In: *The Journal of Economic Asymmetries* 28 (2023), e00328 (cit. on pp. 6, 12, 15).
- [31] Yuling Huang et al. “A novel deep reinforcement learning framework with BiLSTM-Attention networks for algorithmic trading”. In: *Expert Systems with Applications* 240 (2024), p. 122581 (cit. on p. 7).
- [32] Nick James and Max Menzies. “Collective correlations, dynamics, and behavioural inconsistencies of the cryptocurrency market over time”. In: *Non-linear Dynamics* 107.4 (2022), pp. 4001–4017 (cit. on pp. 8, 69, 87, 98).
- [33] Jarosław Kwapien et al. “Genuine multifractality in time series is due to temporal correlations”. In: *Physical Review E* 107.3 (2023), p. 034139 (cit. on p. 9).
- [34] Alexander Brauneis and Roland Mestel. “Price discovery of cryptocurrencies: Bitcoin and beyond”. In: *Econ. Lett* 165 (2018), pp. 58–61 (cit. on p. 9).
- [35] Rocco Caferra and David Vidal-Tomás. “Who raised from the abyss? A comparison between cryptocurrency and stock market dynamics during the COVID-19 pandemic”. In: *Finance Res. Lett* 43 (2021), p. 101954 (cit. on p. 9).

- [36] Dirk G Baur and Thomas Dimpfl. “The volatility of Bitcoin and its role as a medium of exchange and a store of value”. In: *Empir. Econ* 61.5 (2021), pp. 2663–2683 (cit. on p. 9).
- [37] J Bradford De Long et al. “Noise trader risk in financial markets”. In: *Journal of political Economy* 98.4 (1990), pp. 703–738 (cit. on p. 10).
- [38] Marcos M López de Prado. *Machine learning for asset managers*. Cambridge University Press, 2020 (cit. on p. 10).
- [39] Nathan S Balke, Xin Jin, and Mine Yücel. “The shale revolution and the dynamics of the oil market”. In: *Available at SSRN 3645597* (2020) (cit. on pp. 12, 15, 131).
- [40] Richie Ruchuan Ma, Tao Xiong, and Yukun Bao. “The Russia-Saudi Arabia oil price war during the COVID-19 pandemic”. In: *Energy Economics* 102 (2021), p. 105517 (cit. on pp. 12, 15, 131).
- [41] Rahul Kumar, Dharen Kumar Pandey, and John W Goodell. “Market reactions to layoff announcements during crises: Examining impacts and conditioners”. In: *Finance Research Letters* 58 (2023), p. 104423 (cit. on pp. 12, 16, 131).
- [42] Qi Su. “To what extent did US-China trade war affect the global economy”. In: *SHS Web of Conferences*. Vol. 181. EDP Sciences. 2024, p. 03017 (cit. on p. 12).
- [43] Chad P Bown. “How the United States marched the semiconductor industry into its trade war with China”. In: *East Asian Economic Review* 24.4 (2020), pp. 349–388 (cit. on p. 12).
- [44] Kimberly A Houser. “The Innovation Winter is Coming: How the US-China trade war endangers the world”. In: *San Diego L. Rev.* 57 (2020), p. 549 (cit. on p. 12).

- [45] Lisset-Vanesa Apcho-Ccencho et al. “The impact of international price on the technological industry in the United States and China during times of crisis: commercial war and COVID-19”. In: *Advances in business and management forecasting*. Emerald Publishing Limited, 2021, pp. 149–160 (cit. on p. 12).
- [46] Natalie Fang Ling Cheng et al. “The US-China trade war and the volatility linkages between energy and agricultural commodities”. In: *Energy Economics* 120 (2023), p. 106605 (cit. on pp. 12, 114).
- [47] Muhammad Ali Baig and Hassan Tariq. “US-China Trade War: Regional Implications”. In: *Annals of Social Sciences & Management Studies* 8.3 (2023), pp. 002–4 (cit. on pp. 12, 13, 114, 119).
- [48] Larisa Kapustina et al. “US-China trade war: Causes and outcomes”. In: *SHS Web of Conferences*. Vol. 73. EDP Sciences. 2020, p. 01012 (cit. on pp. 12, 120).
- [49] Drew Richard. *Markets Right Now: Markets shudder as trade tensions flare*. AP News. 2019 August 23 [Cited on 2024 April 04]. AP News. URL: <https://apnews.com/article/6ddfc739ba9a43bba13e7cafd69a5f0d> (cit. on p. 12).
- [50] Lewis Al. *The stock market boomed in 2019. Here’s how it happened*. CNBC. 2019 December 31 [Cited on 2024 April 04]. CNBC. URL: <https://www.cnbc.com/2019/12/31/the-stock-market-boomed-in-2019-heres-how-it-happened.html> (cit. on pp. 12, 119, 120).
- [51] Partington Richard. *Global markets take fright as Trump ramps up US-China trade war*. The Guardian. 2019 August 02 [Cited on 2024 April 04]. The Guardian. URL: <https://www.theguardian.com/business/2019/aug/02/global-markets-take-fright-as-trump-ramps-up-us-china-trade-war> (cit. on pp. 12, 119, 120).
- [52] Chengying He, Rui Chen, and Ying Liu. “US-China trade war and China’s stock market: an event-driven analysis”. In: *Economic research-Ekonomska istraživanja* 35.1 (2022), pp. 3277–3290 (cit. on p. 13).

- [53] Felipe Benguria. “The global impact of the US–China trade war: firm-level evidence”. In: *Review of World Economics* 159.4 (2023), pp. 827–851 (cit. on pp. 13, 114, 119).
- [54] Chad P Bown. “The US–China trade war and Phase One agreement”. In: *Journal of Policy Modeling* 43.4 (2021), pp. 805–843 (cit. on p. 13).
- [55] Susan Martin and Jonas Bergmann. “(Im) mobility in the age of COVID-19”. In: *International Migration Review* 55.3 (2021), pp. 660–687 (cit. on p. 13).
- [56] Suborna Barua. “Understanding Coronanomics: The economic implications of the coronavirus (COVID-19) pandemic”. In: *Available at SSRN 3566477* (2020) (cit. on p. 13).
- [57] Rakesh Padhan and KP Prabheesh. “The economics of COVID-19 pandemic: A survey”. In: *Economic analysis and policy* 70 (2021), pp. 220–237 (cit. on p. 13).
- [58] Zongyun Li et al. “A comparative analysis of COVID-19 and global financial crises: evidence from US economy”. In: *Economic Research-Ekonomska Istraživanja* 35.1 (2022), pp. 2427–2441 (cit. on pp. 13, 114).
- [59] Salim Lahmiri and Stelios Bekiros. “The impact of COVID-19 pandemic upon stability and sequential irregularity of equity and cryptocurrency markets”. In: *Chaos Solit. Fractals* 138 (2020), p. 109936 (cit. on p. 13).
- [60] Mohamed Yousfi et al. “Effects of the COVID-19 pandemic on the US stock market and uncertainty: A comparative assessment between the first and second waves”. In: *Technol. Forecast. Soc. Change* 167 (2021), p. 120710 (cit. on p. 13).
- [61] V Acharya, Robert Engle, and Sascha Steffen. “What explains the crash of bank stock prices during COVID-19?” In: *Asian Bureau of Finance and Economic Research* (2021) (cit. on pp. 13, 114).

- [62] Mohsin Ali, Nafis Alam, and Syed Aun R Rizvi. “Coronavirus (COVID-19)—An epidemic or pandemic for financial markets”. In: *J. Behav. Exp. Finance* 27 (2020), p. 100341 (cit. on p. 13).
- [63] Rahma Chemkha et al. “Hedge and safe haven properties during COVID-19: Evidence from Bitcoin and gold”. In: *The Quarterly Review of Economics and Finance* 82 (2021), pp. 71–85 (cit. on p. 13).
- [64] An Pham Ngoc Nguyen et al. “Volatility and returns connectedness in cryptocurrency markets: Insights from graph-based methods”. In: *Physica A: Statistical Mechanics and its Applications* 632 (2023), p. 129349 (cit. on pp. 13, 17, 122).
- [65] An Pham Ngoc Nguyen et al. “The cryptocurrency market in transition before and after covid-19: An opportunity for investors?” In: *Entropy* 24.9 (2022), p. 1317 (cit. on p. 13).
- [66] Kang Hua Cao et al. “Covid-19’s adverse effects on a stock market index”. In: *Appl. Econ. Lett* 28.14 (2021), pp. 1157–1161 (cit. on p. 13).
- [67] Sinem Derindere Köseoğlu et al. “The impact of the Russian-Ukraine war on the stock market: A causal analysis”. In: *Applied Economics* (2023), pp. 1–11 (cit. on pp. 14, 114).
- [68] Bijoy Chandra Das et al. “Ukraine–Russia conflict and stock markets reactions in Europe”. In: *Global Journal of Flexible Systems Management* 24.3 (2023), pp. 395–407 (cit. on pp. 14, 114).
- [69] Thomas Conlon et al. “Financial market information flows when counteracting rogue states: The indirect effects of targeted sanction packages”. In: *Journal of Economic Behavior & Organization* 217 (2024), pp. 32–62 (cit. on pp. 14, 114).
- [70] Vineeta Kumari, Gaurav Kumar, and Dharen Kumar Pandey. “Are the European Union stock markets vulnerable to the Russia–Ukraine war?” In: *Journal of Behavioral and Experimental Finance* 37 (2023), p. 100793 (cit. on p. 14).

- [71] Luiz Eduardo Gaio et al. “The impact of the Russia-Ukraine conflict on market efficiency: Evidence for the developed stock market”. In: *Finance Research Letters* 50 (2022), p. 103302 (cit. on pp. 14, 15).
- [72] Sabri Boubaker et al. “Heterogeneous impacts of wars on global equity markets: Evidence from the invasion of Ukraine”. In: *Finance Research Letters* 48 (2022), p. 102934 (cit. on pp. 14, 15).
- [73] Ruben Teel. “The effect of the war in Ukraine on herding behaviour in European stock markets”. B.S. thesis. University of Twente, 2023 (cit. on pp. 14, 131).
- [74] Cosimo Magazzino, Marco Mele, and Angelo Quarto. “Does the Oil Price Slump Cause a Cleaner Energy Transition during Coronavirus Disease 2019 in Italy? Fresh Evidence from Quantile and Wavelet Coherency Analyses”. In: *Energy Technology* 12.2 (2024), p. 2301044 (cit. on p. 14).
- [75] Chourouk Haouel. “Assessment of the Impact of Russia’s war on Ukraine on EU and UK oil and gas imports and their Energy Supply Security”. In: *Proceedings of the Central and Eastern European eDem and eGov Days 2023*. 2023, pp. 166–177 (cit. on p. 14).
- [76] Tania Babina et al. “Assessing the impact of international sanctions on Russian oil exports”. In: *Available at SSRN 4366337* (2023) (cit. on p. 14).
- [77] Tomasz Rokicki, Piotr Bórawski, and András Szeberényi. “The impact of the 2020–2022 crises on eu countries’ independence from energy imports, particularly from russia”. In: *Energies* 16.18 (2023), p. 6629 (cit. on p. 14).
- [78] Shaker Ahmed, Mostafa M Hasan, and Md Rajib Kamal. “Russia–Ukraine crisis: The effects on the European stock market”. In: *European Financial Management* 29.4 (2023), pp. 1078–1118 (cit. on p. 14).
- [79] Whelsy Bounou and Alhonita Yatié. “The impact of the Ukraine–Russia war on world stock market returns”. In: *Economics letters* 215 (2022), p. 110516 (cit. on p. 14).

- [80] Imran Yousaf, Ritesh Patel, and Larisa Yarovaya. “The reaction of G20+ stock markets to the Russia–Ukraine conflict “black-swan” event: Evidence from event study approach”. In: *Journal of Behavioral and Experimental Finance* 35 (2022), p. 100723 (cit. on p. 15).
- [81] Feng-lin Wu et al. “Stock market volatility and Russia–Ukraine conflict”. In: *Finance Research Letters* 55 (2023), p. 103919 (cit. on p. 15).
- [82] p.l.c. BP. *Statistical Review of World Energy 2020*. BP p.l.c., 2020 (cit. on pp. 15, 131).
- [83] BT Olanrewaju and OE Olubusoye. *Reduction of Petroleum Consumption*. 2020 (cit. on p. 15).
- [84] Sanzillo Tom. *IEEFA update: Oil and gas stocks place dead last in 2019, again, despite 30% price rise*. Institute for Energy Economics and Financial Analysis. 2020 January 09 [Cited on 2024 April 04]. Institute for Energy Economics and Financial Analysis. URL: <https://ieefa.org/resources/ieefa-update-oil-and-gas-stocks-place-dead-last-2019-again-despite-30-price-rise> (cit. on p. 15).
- [85] Richard Vale. “Forecasting the 2022-23 tech layoffs using epidemiological models”. In: *arXiv preprint arXiv:2305.05210* (2023) (cit. on pp. 15, 131).
- [86] Kamran Eshghi and Vivek Astvansh. “Stock investors’ reaction to layoff announcements: A meta-analysis”. In: *Human Resource Management Journal* (2023) (cit. on p. 16).
- [87] John W Goodell and Stéphane Goutte. “Diversifying equity with cryptocurrencies during COVID-19”. In: *Int. Rev. Financ. Anal* 76 (2021), p. 101781 (cit. on pp. 16, 17).
- [88] Gilles Boevi Koumou. “Diversification and portfolio theory: a review”. In: *Financial Markets and Portfolio Management* 34.3 (2020), pp. 267–312 (cit. on p. 16).

- [89] Duc Khuong Nguyen, Nikolas Topaloglou, and Thomas Walther. “Asset Classes and Portfolio Diversification: Evidence from a Stochastic Spanning Approach”. In: *Available at SSRN 3721907* (2020) (cit. on p. 16).
- [90] Bana Abuzayed and Nedal Al-Fayoumi. “Diversification and hedging strategies of green bonds in financial asset portfolios during the COVID-19 pandemic”. In: *Applied Economics* (2022), pp. 1–11 (cit. on p. 16).
- [91] Yosra Ghabri, Khaled Guesmi, and Ahlem Zantour. “Bitcoin and liquidity risk diversification”. In: *Finance Research Letters* 40 (2021), p. 101679 (cit. on p. 16).
- [92] Hossain Mohammad Reyad et al. “The Impact of Macroeconomic Risk Factors, the Adoption of Financial Derivatives on Working Capital Management, and Firm Performance”. In: *Sustainability* 14.21 (2022), p. 14447 (cit. on p. 16).
- [93] Haitao Liu, Jian Zou, and Nalini Ravishanker. “Clustering high-frequency financial time series based on information theory”. In: *Applied Stochastic Models in Business and Industry* 38.1 (2022), pp. 4–26 (cit. on p. 17).
- [94] Mariana Durcheva and Pavel Tsankov. “Analysis of similarities between stock and cryptocurrency series by using graphs and spanning trees”. In: *AIP Conference Proceedings*. Vol. 2172. 1. AIP Publishing LLC. 2019, p. 090004 (cit. on pp. 17, 77).
- [95] Tony Klein, Hien Pham Thu, and Thomas Walther. “Bitcoin is not the New Gold—A comparison of volatility, correlation, and portfolio performance”. In: *Int. Rev. Financ. Anal* 59 (2018), pp. 105–116 (cit. on pp. 17, 73).
- [96] Paraskevi Katsiampa, Larisa Yarovaya, and Damian Zięba. “High-frequency connectedness between Bitcoin and other top-traded crypto assets during the COVID-19 crisis”. In: *Journal of International Financial Markets, Institutions and Money* 79 (2022), p. 101578 (cit. on pp. 17, 84).

- [97] Carlos Jaureguizar Francés, PG Carles, and Diego Jaureguizar Arellano. “The cryptocurrency market: A network analysis”. In: *Esic Market Economics and Business Journal* 49.3 (2018), pp. 569–583 (cit. on p. 17).
- [98] Antonio Briola and Tomaso Aste. “Dependency structures in cryptocurrency market from high to low frequency”. In: *arXiv preprint arXiv:2206.03386* (2022) (cit. on p. 18).
- [99] Jarosław Kwapień, Marcin Wątopek, and Stanisław Drożdż. “Cryptocurrency Market Consolidation in 2020&2021”. In: *Entropy* 23.12 (2021). ISSN: 1099-4300. DOI: 10.3390/e23121674 (cit. on pp. 18, 71, 78).
- [100] Stanisław Drożdż, Jarosław Kwapień, and Marcin Wątopek. *Signatures of Maturity in Cryptocurrency Market*. MDPI-Multidisciplinary Digital Publishing Institute, 2023 (cit. on pp. 19, 100).
- [101] Antonio Briola et al. “Anatomy of a Stablecoin’s failure: The Terra-Luna case”. In: *Finance Research Letters* 51 (2023), p. 103358 (cit. on p. 19).
- [102] Jue Gong et al. “Spreading of cross-market volatility information: Evidence from multiplex network analysis of volatility spillovers”. In: *Journal of International Financial Markets, Institutions and Money* 83 (2023), p. 101733 (cit. on p. 20).
- [103] Koji Maruhashi et al. “Meta graphical lasso: uncovering hidden interactions among latent mechanisms”. In: *Scientific Reports* 14.1 (2024), p. 18105 (cit. on p. 20).
- [104] Jenna Gavin and Martin Crane. “Community detection in cryptocurrencies with potential applications to portfolio diversification”. In: *arXiv preprint arXiv:2108.09763* (2021) (cit. on p. 20).
- [105] Farideh Rahimnezhad et al. “Application of community detection in stock market analysis”. In: *International Journal of Industrial Engineering and Management Science* 7.1 (2020), pp. 81–91 (cit. on p. 20).

- [106] Di Jin et al. “A survey of community detection approaches: From statistical modeling to deep learning”. In: *IEEE Trans. Knowl. Data. Eng* (2021) (cit. on p. 21).
- [107] Joseph A Schumpeter. *The General Theory of Employment, Interest and Money*. 1936 (cit. on p. 21).
- [108] David Y Aharon. “Uncertainty, fear and herding behavior: Evidence from size-ranked portfolios”. In: *Journal of Behavioral Finance* 22.3 (2021), pp. 320–337 (cit. on pp. 21, 117, 121).
- [109] Kwangwon Ahn et al. “Business cycle and herding behavior in stock returns: theory and evidence”. In: *Financial Innovation* 10.1 (2024), p. 6 (cit. on pp. 21, 25, 54, 120, 121).
- [110] Nicholas Bloom. “Fluctuations in uncertainty”. In: *Journal of economic Perspectives* 28.2 (2014), pp. 153–176 (cit. on p. 22).
- [111] Thomas C Chiang and Dazhi Zheng. “An empirical analysis of herd behavior in global stock markets”. In: *Journal of Banking & Finance* 34.8 (2010), pp. 1911–1921 (cit. on pp. 22, 53, 120).
- [112] Ali Güvercin. “Sentimental herding: The role of regional and global shocks in Egyptian and Saudi stock markets”. In: *Sosyoekonomi* 24.27 (2016), pp. 2–9 (cit. on p. 22).
- [113] Elie Bouri et al. “COVID-19 pandemic and investor herding in international stock markets”. In: *Risks* 9.9 (2021), p. 168 (cit. on pp. 22, 63).
- [114] Iftekhar Hasan, Radu Tunaru, and Davide Vioto. “Herding behavior and systemic risk in global stock markets”. In: *Journal of Empirical Finance* 73 (2023), pp. 107–133 (cit. on p. 22).
- [115] Wan-Ru Yang and Ming-Che Chuang. “Do investors herd in a volatile market? Evidence of dynamic herding in Taiwan, China, and US stock markets”. In: *Finance Research Letters* 52 (2023), p. 103364 (cit. on pp. 22, 24, 120).

- [116] Faheem Aslam et al. “Herding behavior during the COVID-19 pandemic: A comparison between Asian and European stock markets based on intraday multifractality”. In: *Eurasian Economic Review* 12.2 (2022), pp. 333–359 (cit. on p. 22).
- [117] Manish Talwar et al. “Has financial attitude impacted the trading activity of retail investors during the COVID-19 pandemic?” In: *Journal of Retailing and Consumer Services* 58 (2021), p. 102341 (cit. on pp. 22, 24).
- [118] Xiaorui Zhu et al. “Impact of the COVID-19 pandemic on the stock market and investor online word of mouth”. In: *Decision Support Systems* 176 (2024), p. 114074 (cit. on pp. 22, 24).
- [119] Wenyang Huang et al. “Complex network analysis of global stock market co-movement during the COVID-19 pandemic based on intraday open-high-low-close data”. In: *Financial Innovation* 10.1 (2024), p. 7 (cit. on p. 23).
- [120] Richard T Ampofo et al. “An empirical investigation of COVID-19 effects on herding behaviour in USA and UK stock markets using a quantile regression approach”. In: *Journal of Economics and Finance* 47.2 (2023), pp. 517–540 (cit. on p. 23).
- [121] Khalid MousaALHarbi et al. “Herding and retail investors’ decisions: A review literature”. In: *Environment and Social Psychology* 9.4 (2024) (cit. on p. 23).
- [122] Mehmet Balcilar and Riza Demirer. “Effect of global shocks and volatility on herd behavior in an emerging market: Evidence from Borsa Istanbul”. In: *Emerging Markets Finance and Trade* 51.1 (2015), pp. 140–159 (cit. on p. 23).
- [123] Terence Tai-Leung Chong, Xiaojin Liu, and Chenqi Zhu. “What explains herd behavior in the Chinese stock market?” In: *Journal of Behavioral Finance* 18.4 (2017), pp. 448–456 (cit. on p. 23).

- [124] Zhenxi Chen and Huanhuan Zheng. “Herding in the Chinese and US stock markets: Evidence from a micro-founded approach”. In: *International Review of Economics & Finance* 78 (2022), pp. 597–604 (cit. on p. 23).
- [125] Achraf Ghorbel, Yasmine Snene, and Wajdi Frikha. “Does herding behavior explain the contagion of the COVID-19 crisis?” In: *Review of Behavioral Finance* 15.6 (2023), pp. 889–915 (cit. on pp. 24, 63).
- [126] William J Luther. “Bitcoin and the future of digital payments”. In: *The Independent Review* 20.3 (2016), pp. 397–404 (cit. on p. 25).
- [127] Obryan Poyser. “Herding behavior in cryptocurrency markets”. In: *arXiv preprint arXiv:1806.11348* (2018) (cit. on pp. 25, 121).
- [128] Gerson de Souza Raimundo Júnior et al. “Market stress and herding: a new approach to the cryptocurrency market”. In: *Journal of Behavioral Finance* 23.1 (2022), pp. 43–57 (cit. on pp. 25, 130).
- [129] Pinar Evrim Mandaci and Efe Caglar Cagli. “Herding intensity and volatility in cryptocurrency markets during the COVID-19”. In: *Finance Research Letters* 46 (2022), p. 102382 (cit. on pp. 26, 118, 130).
- [130] RN-U-D Jalal et al. “Herding behavior and cryptocurrency: Market asymmetries, inter-dependency and intra-dependency”. In: *The Journal of Asian Finance, Economics and Business* 7.7 (2020), pp. 27–34 (cit. on pp. 26, 63, 117, 130).
- [131] DM Patterson and V Sharma. “Do traders follow each other at the NYSE?” In: *University of Michigan-Dearborn working paper* (2006) (cit. on p. 26).
- [132] Mouna Youssef and Sami Sobhi Waked. “Herding behavior in the cryptocurrency market during COVID-19 pandemic: The role of media coverage”. In: *The North American Journal of Economics and Finance* 62 (2022), p. 101752 (cit. on pp. 26, 63, 84, 85, 117, 118, 130).
- [133] David Vidal-Tomás and Ana Ibañez. “Semi-strong efficiency of Bitcoin”. In: *Finance Research Letters* 27 (2018), pp. 259–265 (cit. on p. 27).

- [134] Omair Haroon and Syed Aun R Rizvi. “COVID-19: Media coverage and financial markets behavior—A sectoral inquiry”. In: *Journal of behavioral and experimental finance* 27 (2020), p. 100343 (cit. on p. 27).
- [135] Tobias Burggraf et al. “Do FEARS drive bitcoin?” In: *Review of Behavioral Finance* 13.3 (2021), pp. 229–258 (cit. on p. 27).
- [136] Muhammad Abubakr Naeem, Imen Mbarki, and Syed Jawad Hussain Shahzad. “Predictive role of online investor sentiment for cryptocurrency market: Evidence from happiness and fears”. In: *International Review of Economics & Finance* 73 (2021), pp. 496–514 (cit. on pp. 27, 130).
- [137] Patrice Racine Diallo et al. “Dynamics of Herding Behavior in Cryptocurrency Markets Amid Market Crashes”. In: *Journal of Economy Culture and Society* 68 (2023), pp. 130–140 (cit. on pp. 27, 63).
- [138] Phasin Wanidwaranan and Santi Termprasertsakul. “Herd behavior in cryptocurrency market: evidence of network effect”. In: *Review of Behavioral Finance* (2023) (cit. on pp. 27, 117, 118).
- [139] Thomas Conlon, Shaen Corbet, and Richard J McGee. “The Bitcoin volume-volatility relationship: A high frequency analysis of futures and spot exchanges”. In: *International Review of Financial Analysis* 91 (2024), p. 103013 (cit. on p. 30).
- [140] Adrian Jobst et al. “Examining liquidity of exchanges and assets and the impact of external events in centralized crypto markets: a 2022 study”. In: *2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE. 2023, pp. 1–6 (cit. on p. 30).
- [141] Paolo Giudici and Gloria Polinesi. “Crypto price discovery through correlation networks”. In: *Annals of Operations Research* 299.1 (2021), pp. 443–457 (cit. on p. 30).

- [142] Ben R Marshall, Nhut H Nguyen, and Nuttawat Visaltanachoti. “Do liquidity proxies measure liquidity accurately in ETFs?” In: *Journal of International Financial Markets, Institutions and Money* 55 (2018), pp. 94–111 (cit. on p. 31).
- [143] Markus S Broman and Pauline Shum. “Relative liquidity, fund flows and short-term demand: evidence from exchange-traded funds”. In: *Financial Review* 53.1 (2018), pp. 87–115 (cit. on p. 31).
- [144] Itzhak Ben-David, Francesco Franzoni, and Rabih Moussawi. “Do ETFs increase volatility?” In: *The Journal of Finance* 73.6 (2018), pp. 2471–2535 (cit. on p. 31).
- [145] Joe Samuel. *The 11 Sectors Of The Stock Market & Their Biggest ETFs*. <https://www.nasdaq.com/articles/the-11-sectors-of-the-stock-market-their-biggest-etfs-2021-05-30>. Nasdaq, 2021 (cit. on p. 33).
- [146] Daly Lyle. *The Largest Companies by Market Cap in 2024*. The Motley Fool. 2024 April 01 [Cited on 2024 April 04]. The Motley Fool. URL: <https://www.fool.com/research/largest-companies-by-market-cap/> (cit. on pp. 33, 126).
- [147] Fei Su, Xinyi Wang, and Yulin Yuan. “The intraday dynamics and intraday price discovery of bitcoin”. In: *Res. Int. Bus. Finance* 60 (2022), p. 101625 (cit. on pp. 35, 73).
- [148] Anastasios Megaritis, Nikolaos Vlastakis, and Athanasios Triantafyllou. “Stock market volatility and jumps in times of uncertainty”. In: *J. Int. Money Finance* 113 (2021), p. 102355 (cit. on pp. 35, 74).
- [149] Faheem Aslam et al. “Intraday volatility spillovers among European financial markets during COVID-19”. In: *Int. J. Financ. Stud.* 9.1 (2021), p. 5 (cit. on pp. 35, 74).
- [150] Michael Greenacre et al. “Principal component analysis”. In: *Nature Reviews Methods Primers* 2.1 (2022), p. 100 (cit. on p. 36).

- [151] MA Girshick. “Principal components”. In: *Journal of the American Statistical Association* 31.195 (1936), pp. 519–528 (cit. on p. 36).
- [152] Ruey S Tsay. *Analysis of financial time series*. John wiley & sons, 2005 (cit. on p. 41).
- [153] G William Schwert. “Stock market volatility”. In: *Financial analysts journal* 46.3 (1990), pp. 23–34 (cit. on p. 41).
- [154] Fernando Pérez-Cruz, Julio A Afonso-Rodriguez, and Javier Giner. “Estimating GARCH models using support vector machines”. In: *Quantitative Finance* 3.3 (2003), p. 163 (cit. on p. 41).
- [155] Elroy Dimson and Massoud Mussavian. “A brief history of market efficiency”. In: *European financial management* 4.1 (1998), pp. 91–103 (cit. on pp. 41, 42).
- [156] PK Mishra. “Weak form market efficiency: Evidence from emerging and developed world”. In: *The Journal of Commerce* 3.2 (2011), p. 26 (cit. on p. 42).
- [157] Higor YD Sigaki, Matjaž Perc, and Haroldo V Ribeiro. “Clustering patterns in efficiency and the coming-of-age of the cryptocurrency market”. In: *Scientific reports* 9.1 (2019), p. 1440 (cit. on p. 42).
- [158] Christoph Bandt and Bernd Pompe. “Permutation entropy: a natural complexity measure for time series”. In: *Physical review letters* 88.17 (2002), p. 174102 (cit. on p. 42).
- [159] Ricardo Lopez-Ruiz, Hector L Mancini, and Xavier Calbet. “A statistical measure of complexity”. In: *Physics letters A* 209.5-6 (1995), pp. 321–326 (cit. on p. 42).
- [160] Israel Cohen et al. “Pearson correlation coefficient”. In: *Noise reduction in speech processing* (2009), pp. 1–4 (cit. on pp. 43, 44).
- [161] Jerrold H Zar. “Spearman rank correlation”. In: *Encyclopedia of biostatistics* 7 (2005) (cit. on pp. 43, 44).

- [162] Maurice George Kendall. “Rank correlation methods.” In: (1948) (cit. on pp. 43, 44).
- [163] Rajat Khanna and Isin Guler. “Degree assortativity in collaboration networks and invention performance”. In: *Strategic Management Journal* 43.7 (2022), pp. 1402–1430 (cit. on p. 46).
- [164] LC Freeman. “A set of measures of centrality based on betweenness”. In: *Sociometry* (1977) (cit. on p. 46).
- [165] Tingting Lan et al. “The impact of pandemic on dynamic volatility spillover network of international stock markets”. In: *Empirical Economics* (2023), pp. 1–30 (cit. on p. 47).
- [166] Paolo Giudici, Paolo Pagnottoni, and Gloria Polinesi. “Network models to enhance automated cryptocurrency portfolio management”. In: *Front. Artif. Intell.* 3 (2020), p. 22 (cit. on p. 47).
- [167] Jiaqi Liang et al. “Correlation-based dynamics and systemic risk measures in the cryptocurrency market”. In: *IEEE 2018 ISI*. IEEE. 2018, pp. 43–48 (cit. on pp. 47, 70).
- [168] Andrew Rosenberg and Julia Hirschberg. “V-measure: A conditional entropy-based external cluster evaluation measure”. In: *2007 EMNLP-CoNLL*. 2007, pp. 410–420 (cit. on p. 48).
- [169] Danai Koutra et al. “Algorithms for graph similarity and subgraph matching”. In: *Proc. Ecol. inference conf.* Vol. 17. Citeseer. 2011 (cit. on pp. 48, 77).
- [170] Sadegh Ilbeigipour and Babak Teimourpour. “A Social Network Analysis Approach to Evaluate the Relationship Between the Mobility Network Metrics and the COVID-19 Outbreak”. In: *Health Services Insights* 16 (2023), p. 11786329231173816 (cit. on pp. 46, 102).
- [171] Monica Bianchini, Marco Gori, and Franco Scarselli. “Inside pagerank”. In: *ACM Transactions on Internet Technology (TOIT)* 5.1 (2005), pp. 92–128 (cit. on p. 49).

- [172] Wei Song et al. “Empirical studies on symbolic aggregation approximation under statistical perspectives for knowledge discovery in time series”. In: *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*. IEEE, 2015, pp. 1040–1046 (cit. on p. 50).
- [173] Vasiliki Plerou et al. “Random matrix approach to cross correlations in financial data”. In: *Phys. Rev. E* 65.6 (2002), p. 066126 (cit. on pp. 50, 51).
- [174] Vasiliki Plerou et al. “A random matrix theory approach to financial cross-correlations”. In: *Physica A* 287.3-4 (2000), pp. 374–382 (cit. on p. 51).
- [175] Claudio Cacciapuoti, Anna Maltsev, and Benjamin Schlein. “Local Marchenko-Pastur law at the hard edge of sample covariance matrices”. In: *J. Math. Phys.* 54.4 (2013), p. 043302 (cit. on p. 51).
- [176] Laurent Laloux et al. “Random matrix theory and financial correlations”. In: *International Journal of Theoretical and Applied Finance* 3.03 (2000), pp. 391–397 (cit. on p. 51).
- [177] Marcos M López de Prado. *Machine learning for asset managers*. Cambridge University Press, 2020 (cit. on pp. 51, 57).
- [178] Mícheál O’Searcoid. *Metric spaces*. Springer Science & Business Media, 2006 (cit. on p. 51).
- [179] Alethea Rea and William Rea. “Visualization of a stock market correlation matrix”. In: *Physica A: Statistical Mechanics and its Applications* 400 (2014), pp. 109–123 (cit. on p. 52).
- [180] Michele Tumminello et al. “Spanning trees and bootstrap reliability estimation in correlation-based networks”. In: *International Journal of Bifurcation and Chaos* 17.07 (2007), pp. 2319–2329 (cit. on p. 52).
- [181] Marcin Wałtorek et al. “Multiscale characteristics of the emerging global cryptocurrency market”. In: *Physics Reports* 901 (2021), pp. 1–82 (cit. on p. 52).

- [182] Per Bak, Maya Paczuski, and Martin Shubik. “Price variations in a stock market with many agents”. In: *Physica A: Statistical Mechanics and its Applications* 246.3-4 (1997), pp. 430–453 (cit. on p. 53).
- [183] Vikash Ramiah, Xiaoming Xu, and Imad A Moosa. “Neoclassical finance, behavioral finance and noise traders: A review and assessment of the literature”. In: *International review of financial analysis* 41 (2015), pp. 89–100 (cit. on p. 53).
- [184] Raphael Auer and Stijn Claessens. “Regulating cryptocurrencies: assessing market reactions”. In: *BIS Quarterly Review September* (2018) (cit. on p. 53).
- [185] Tianle Yang et al. “Fluctuation in the global oil market, stock market volatility, and economic policy uncertainty: a study of the US and China”. In: *The quarterly review of economics and finance* 87 (2023), pp. 377–387 (cit. on p. 53).
- [186] Oguz Ersan et al. “The speed of stock price adjustment to corporate announcements: Insights from Turkey”. In: *Emerging Markets Review* 47 (2021), p. 100778 (cit. on p. 53).
- [187] Tao Li, Donghwa Shin, and Baolian Wang. “Cryptocurrency pump-and-dump schemes”. In: *Available at SSRN 3267041* (2021) (cit. on p. 53).
- [188] José Almeida and Tiago Cruz Gonçalves. “A systematic literature review of investor behavior in the cryptocurrency markets”. In: *Journal of Behavioral and Experimental Finance* (2023), p. 100785 (cit. on pp. 53, 98, 117, 122).
- [189] Paulo Vitor Jordão da Gama Silva et al. “Herding behavior and contagion in the cryptocurrency market”. In: *J. Behav. Exp. Finance* 22 (2019), pp. 41–50 (cit. on pp. 53, 84, 105, 117).
- [190] Blockchain Council. *10 Blockchain Protocols to Know in 2023*. <https://www.blockchain-council.org/blockchain/10-blockchain-protocols/>. Blockchain Council, 2023 (cit. on p. 54).

- [191] Jean-Philippe Bouchaud and Marc Potters. *Theory of financial risk and derivative pricing: from statistical physics to risk management*. Cambridge university press, 2003 (cit. on pp. 54, 55).
- [192] Jean-Philippe Bouchaud and Marc Potters. “Financial applications of random matrix theory: a short review”. In: *arXiv preprint arXiv:0910.1205* (2009) (cit. on pp. 54–56).
- [193] Gordana Ivosev, Lyle Burton, and Ron Bonner. “Dimensionality reduction and visualization in principal component analysis”. In: *Analytical chemistry* 80.13 (2008), pp. 4933–4944 (cit. on p. 55).
- [194] Zdzislaw Burda and Andrzej Jarosz. “Cleaning large-dimensional covariance matrices for correlated samples”. In: *Phys. Rev. E* 105.3 (2022), p. 034136 (cit. on p. 55).
- [195] Joël Bun and A Knowles. “An optimal rotational invariant estimator for general covariance matrices: The outliers”. In: *Preprint* (2018) (cit. on p. 55).
- [196] Olivier Ledoit and Michael Wolf. “Nonlinear shrinkage estimation of large-dimensional covariance matrices”. In: *Ann. Stat.* 40.2 (2012), pp. 1024–1060 (cit. on p. 55).
- [197] Refining Markowitz. *Cleaning correlation matrices*. 2016 (cit. on p. 55).
- [198] Thomas Conlon, Heather J Ruskin, and Martin Crane. “Random matrix theory and fund of funds portfolio optimisation”. In: *Physica A* 382.2 (2007), pp. 565–576 (cit. on p. 55).
- [199] Liusha Yang, Matthew R McKay, and Romain Couillet. “High-dimensional MVDR beamforming: Optimized solutions based on spiked random matrix models”. In: *IEEE Trans. Signal Process* 66.7 (2018), pp. 1933–1947 (cit. on p. 55).
- [200] Tai Tan Mai, Marija Bezbradica, and Martin Crane. “Learning behaviours data in programming education: Community analysis and outcome prediction

- with cleaned data”. In: *Future Generation Computer Systems* 127 (2022), pp. 42–55 (cit. on pp. 55, 60).
- [201] Tai Tan Mai, Martin Crane, and Marija Bezbradica. “Students’ Learning Behaviour in Programming Education Analysis: Insights from Entropy and Community Detection”. In: *Entropy* 25.8 (2023), p. 1225 (cit. on p. 55).
- [202] Tapio Heimo, Kimmo Kaski, and Jari Saramäki. “Maximal spanning trees, asset graphs and random matrix denoising in the analysis of dynamics of financial networks”. In: *Physica A: Statistical Mechanics and its Applications* 388.2-3 (2009), pp. 145–156 (cit. on p. 55).
- [203] Laurent Laloux et al. “Noise dressing of financial correlation matrices”. In: *Physical review letters* 83.7 (1999), p. 1467 (cit. on p. 57).
- [204] SM Taheri and Gholamreza Hesamian. “A generalization of the Wilcoxon signed-rank test and its applications”. In: *Statistical Papers* 54 (2013), pp. 457–470 (cit. on pp. 59, 92).
- [205] Vincent D Blondel et al. “Fast unfolding of communities in large networks”. In: *J. Stat. Mech.: Theory Exp.* 2008.10 (2008), P10008 (cit. on p. 60).
- [206] Abolfazl Dibaji and Sadegh Sulaimany. “Community Detection to Improve Machine Learning Based Heart Disease Prediction”. In: *2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP)*. IEEE, 2024, pp. 1–6 (cit. on p. 60).
- [207] Yuhuan Li et al. “Developing a novel approach in estimating urban commute traffic by integrating community detection and hypergraph representation learning”. In: *Expert Systems with Applications* (2024), p. 123790 (cit. on p. 60).
- [208] Christian T Brownlees and Giampiero M Gallo. “Financial econometric analysis at ultra-high frequency: Data handling concerns”. In: *Computational statistics & data analysis* 51.4 (2006), pp. 2232–2245 (cit. on p. 67).

- [209] Michael Boyle and Amanda Bellucco-Chatham. *The Great Recession*. 2022 (cit. on p. 68).
- [210] Alessandro Spelta and Tanya Araújo. “The topology of cross-border exposures: beyond the minimal spanning tree approach”. In: *Physica A* 391.22 (2012), pp. 5572–5583 (cit. on p. 70).
- [211] Wenjun Feng, Yiming Wang, and Zhengjun Zhang. “Can cryptocurrencies be a safe haven: a tail risk perspective analysis”. In: *Appl. Econ.* 50.44 (2018), pp. 4745–4762 (cit. on p. 73).
- [212] Shaen Corbet et al. “Cryptocurrencies as a financial asset: A systematic analysis”. In: *Int. Rev. Financ. Anal* 62 (2019), pp. 182–199 (cit. on p. 73).
- [213] Saifedean Ammous. “Can cryptocurrencies fulfil the functions of money?” In: *Q. Rev. Econ. Finance* 70 (2018), pp. 38–51 (cit. on p. 73).
- [214] Mickaël Mallinger-Dogan and Mark C Szigety. “Higher-Frequency Analysis of Low-Frequency Data”. In: *J. Portf. Manag* 41.1 (2014), pp. 121–138 (cit. on p. 75).
- [215] Jiaqi Liang, Linjing Li, and Daniel Zeng. “Evolutionary dynamics of cryptocurrency transaction networks: An empirical study”. In: *PloS one* 13.8 (2018), e0202202 (cit. on p. 77).
- [216] Raluca Gera et al. “Identifying network structure similarity using spectral graph theory”. In: *Appl. Netw. Sci.* 3.1 (2018), pp. 1–15 (cit. on p. 77).
- [217] Larisa Yarovaya, Roman Matkovskyy, and Akanksha Jalan. “The effects of a ‘Black Swan’ event (COVID-19) on herding behavior in cryptocurrency markets: Evidence from cryptocurrency USD, EUR, JPY and KRW Markets”. In: *EUR, JPY and KRW Markets* (2020) (cit. on pp. 83, 84).
- [218] D Susana, JK Kavisamathi, and S Sreejith. “Does herding behaviour among traders increase during COVID 19 pandemic? Evidence from the cryptocurrency market”. In: *International Working Conference on Transfer and Diffusion of IT*. Springer. 2020, pp. 178–189 (cit. on p. 84).

- [219] Lars Kaiser and Sebastian Stöckl. “Cryptocurrencies: Herding and the transfer currency”. In: *Finance Res. Lett* 33 (2020), p. 101214 (cit. on p. 84).
- [220] Emna Mnif et al. “Investor behavior and cryptocurrency market bubbles during the COVID-19 pandemic”. In: *Rev. Behav. Finance* (2022) (cit. on p. 84).
- [221] Shaen Corbet et al. “KODAKCoin: a blockchain revolution or exploiting a potential cryptocurrency bubble?” In: *Appl. Econ. Lett* 27.7 (2020), pp. 518–524 (cit. on p. 84).
- [222] Christoph Huber, Jürgen Huber, and Michael Kirchler. “Market shocks and professionals’ investment behavior—evidence from the covid-19 crash”. In: *J. Bank. Financ.* 133 (2021), p. 106247 (cit. on p. 85).
- [223] Vidyapati Jha and Priyanka Tripathi. “Semantic Modelling of Multivariate Time-Series Data in Cognitive IoT”. In: *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*. Vol. 1. IEEE. 2023, pp. 943–948 (cit. on p. 93).
- [224] Winifred Huang et al. “The new challenges of global banking and finance”. In: *The European Journal of Finance* (2023), pp. 1–7 (cit. on p. 96).
- [225] John Collins. “Crypto, crime and control”. In: (2022) (cit. on p. 97).
- [226] Anirudh Dhawan and Tālis J Putniņš. “A new wolf in town? Pump-and-dump manipulation in cryptocurrency markets”. In: *Review of Finance* 27.3 (2023), pp. 935–975 (cit. on pp. 97, 106, 109).
- [227] Ghulame Rubbaniy et al. “COVID-19, Lockdowns and herding towards a cryptocurrency market-specific implied volatility index”. In: *Economics Letters* 207 (2021), p. 110017 (cit. on p. 98).
- [228] Samuel Asante Gyamerah. “Covid-19 Pandemic and Herding Behavior in Cryptocurrency Market”. In: *Applied Finance Letters* 10 (2021), pp. 58–66 (cit. on p. 98).

- [229] Elie Bouri, Georges Azzi, and Anne Haubo Dyhrberg. “On the return-volatility relationship in the Bitcoin market around the price crash of 2013”. In: *Economics* 11.1 (2017) (cit. on p. 100).
- [230] Nidhaledine Ben Cheikh, Younes Ben Zaied, and Julien Chevallier. “Asymmetric volatility in cryptocurrency markets: New evidence from smooth transition GARCH models”. In: *Finance Research Letters* 35 (2020), p. 101293 (cit. on p. 100).
- [231] Shinji Kakinaka and Ken Umeno. “Asymmetric volatility dynamics in cryptocurrency markets on multi-time scales”. In: *Research in International Business and Finance* 62 (2022), p. 101754 (cit. on p. 100).
- [232] David Easley et al. “Liquidity, information, and infrequently traded stocks”. In: *The Journal of Finance* 51.4 (1996), pp. 1405–1436 (cit. on p. 100).
- [233] Sergio Gómez. “Centrality in networks: finding the most important nodes”. In: *Business and consumer analytics: New ideas* (2019), pp. 401–433 (cit. on p. 102).
- [234] Duanbing Chen et al. “Identifying influential nodes in complex networks”. In: *Physica a: Statistical mechanics and its applications* 391.4 (2012), pp. 1777–1787 (cit. on p. 102).
- [235] Shen Zhong, Haotian Zhang, and Yong Deng. “Identification of influential nodes in complex networks: A local degree dimension approach”. In: *Information Sciences* 610 (2022), pp. 994–1009 (cit. on p. 102).
- [236] Tolga Umut Kuzubaş, Inci Ömercikoğlu, and Burak Saltoğlu. “Network centrality measures and systemic risk: An application to the Turkish financial crisis”. In: *Physica A: Statistical Mechanics and its Applications* 405 (2014), pp. 203–215 (cit. on p. 102).
- [237] Jung Yoon Song, Woojin Chang, and Jae Wook Song. “Cluster analysis on the structure of the cryptocurrency market via Bitcoin–Ethereum filtering”. In:

- Physica A: Statistical Mechanics and its Applications* 527 (2019), p. 121339 (cit. on p. 105).
- [238] Jarosław Kwapień, Marcin Wątopek, and Stanisław Drożdż. “Cryptocurrency market consolidation in 2020–2021”. In: *Entropy* 23.12 (2021), p. 1674 (cit. on p. 105).
- [239] Adam Zaremba et al. “Up or down? Short-term reversal, momentum, and liquidity effects in cryptocurrency markets”. In: *International Review of Financial Analysis* 78 (2021), p. 101908 (cit. on p. 107).
- [240] Mehmet Balcilar and Huseyin Ozdemir. “On the Risk Spillover from Bitcoin to Altcoins: The Fear of Missing Out and Pump-and-Dump Scheme Effects”. In: *Journal of Risk and Financial Management* 16.1 (2023), p. 41 (cit. on p. 107).
- [241] Francis X Diebold and Kamil Yilmaz. “On the network topology of variance decompositions: Measuring the connectedness of financial firms”. In: *Journal of econometrics* 182.1 (2014), pp. 119–134 (cit. on p. 107).
- [242] Lanouar Charfeddine and Ahmed Mahrous. “What drives cryptocurrency pump and dump schemes: Coin versus market factors?” In: *Finance Research Letters* 67 (2024), p. 105861 (cit. on p. 107).
- [243] JT Hamrick et al. “The economics of cryptocurrency pump and dump schemes”. In: *Workshop on the Economics of Information Security*. 2019 (cit. on p. 107).
- [244] Peter D DeVries. “An analysis of cryptocurrency, bitcoin, and the future”. In: *International Journal of Business Management and Commerce* 1.2 (2016), pp. 1–9 (cit. on p. 109).
- [245] Husam Bukhary, Kyohei Shibano, and Gento Mogi. “High-Frequency Spillover Analysis of Cryptocurrency in the Exchange Listings”. In: *Edited by Sergey Y. Yurish* (2022), p. 61 (cit. on p. 109).

- [246] Bing Anderson. “A tick-by-tick level measurement of the lead-lag duration between cryptocurrencies: The case of Bitcoin versus Cardano”. In: *Investment Management and Financial Innovations* 20.1 (2023), pp. 174–183 (cit. on p. 109).
- [247] Nikolaos Kyriazis et al. “The differential influence of social media sentiment on cryptocurrency returns and volatility during COVID-19”. In: *The Quarterly Review of Economics and Finance* 89 (2023), pp. 307–317 (cit. on p. 110).
- [248] Anders Bylund. *3 Cryptocurrencies to Buy in Any Bear Market*. <https://www.fool.com/investing/2023/01/30/3-cryptocurrencies-to-buy-in-any-bear-market/>. The Motley Fool, 2023 (cit. on p. 110).
- [249] John Caroline. *Brave Browser Surpasses 25 Million Monthly Active Users*. <https://zycrypto.com/brave-browser-surpasses-25-million-monthly-active-users/>. ZyCrypto, 2021 (cit. on p. 110).
- [250] Yogita Khatri. *Privacy-focused Brave browser added 1M new users in March*. <https://finance.yahoo.com/news/privacy-focused-brave-browser-added-190400408.html>. YahooFinance, 2020 (cit. on p. 110).
- [251] Kriptomat. *What is cryptocurrency Maker (MKR) and how does it work?* <https://kriptomat.io/cryptocurrencies/maker/what-is-maker/>. Kriptomat, 2023 (cit. on pp. 110, 111).
- [252] Martin Young. *How Dai is Dominating the World of NFTs and Digital Art*. <https://beincrypto.com/how-dai-is-dominating-the-world-of-nfts-and-digital-art/>. Beincrypto, 2020 (cit. on p. 111).
- [253] Marcel Pechman. *FunFair (FUN) gains 183% as the blockchain-based gaming industry grows*. <https://cointelegraph.com/news/funfair-fun-gains-183-as-the-blockchain-based-gaming-industry-grows>. Cointelegraph, 2021 (cit. on p. 111).

- [254] Elizabeth Howcroft. *Virtual real estate plot sells for record \$2.4 million*. <https://www.reuters.com/markets/currencies/virtual-real-estate-plot-sells-record-24-million-2021-11-23/>. Reuters, 2021 (cit. on p. 111).
- [255] Thomas Conlon, John Cotter, and Ioannis Ropotos. “Diversification with globally integrated US stocks”. In: *Journal of International Financial Markets, Institutions and Money* 90 (2024), p. 101889 (cit. on pp. 112, 122).
- [256] Mohit Kumar et al. “Exploring Patterns and Correlations Between Cryptocurrencies and Forecasting Crypto Prices Using Influential Tweets”. In: *Machine Learning, Image Processing, Network Security and Data Sciences: 4th International Conference, MIND 2022, Virtual Event, January 19–20, 2023, Proceedings, Part II*. Springer. 2023, pp. 308–321 (cit. on p. 114).
- [257] Manpreet Kaur, Jinesh Jain, and Kirti Sood. ““All are investing in Crypto, I fear of being missed out”: examining the influence of herding, loss aversion, and overconfidence in the cryptocurrency market with the mediating effect of FOMO”. In: *Quality & Quantity* (2023), pp. 1–27 (cit. on p. 117).
- [258] Mansooreh Kazemilari et al. “Network topology of renewable energy companies: minimal spanning tree and sub-dominant ultrametric for the American stock”. In: *Technol. Econ. Dev. Econ* 25.2 (2019), pp. 168–187 (cit. on p. 117).
- [259] Antonis Ballis and Konstantinos Drakos. “Testing for herding in the cryptocurrency market”. In: *Finance Research Letters* 33 (2020), p. 101210 (cit. on p. 117).
- [260] Stephanos Papadamou et al. “Herding behaviour and price convergence clubs in cryptocurrencies during bull and bear markets”. In: *Journal of Behavioral and Experimental Finance* 30 (2021), p. 100469 (cit. on p. 118).
- [261] Maryla Maliszewska, Aaditya Mattoo, and Dominique Van Der Mensbrughe. “The potential impact of COVID-19 on GDP and trade: A preliminary

- assessment”. In: *World Bank policy research working paper* 9211 (2020) (cit. on p. 120).
- [262] Terence Tai Leung Chong and Xiaoyang Li. “Understanding the China–US trade war: causes, economic impact, and the worst-case scenario”. In: *Economic and Political Studies* 7.2 (2019), pp. 185–202 (cit. on p. 120).
- [263] Marco Meier and Christian Maier. “From stocks to ETFs: explaining retail investors’ migration behavior”. In: *Internet Research* 33.4 (2023), pp. 1249–1275 (cit. on p. 121).
- [264] Meryem Duygun, Radu Tunaru, and Davide Vioto. “Herding by corporates in the US and the Eurozone through different market conditions”. In: *Journal of International Money and Finance* 110 (2021), p. 102311 (cit. on p. 121).
- [265] Syed Jawad Hussain Shahzad et al. “Is Bitcoin a better safe-haven investment than gold and commodities?” In: *International Review of Financial Analysis* 63 (2019), pp. 322–330 (cit. on p. 122).
- [266] Daeyoung Jeon et al. “Herding in Cryptocurrency Market and Covid-19: Spillover Effects and Fundamental Information”. In: *Available at SSRN 4759813* (2024) (cit. on p. 130).
- [267] Bassam Fattouh and Adi Imsirovic. “Oil benchmarks under stress”. In: *OIES Energy Comment. April* (2020), pp. 1–8 (cit. on p. 131).
- [268] Q.ai. *Why Were Tech Stocks Down In 2022—And How Long Will The Slump Last?* Forbes. 2023 Jan 19 [Cited on 2024 April 04]. Forbes. URL: <https://www.forbes.com/sites/qai/2023/01/19/why-were-tech-stocks-down-in-2022-and-how-long-will-the-slump-last/?sh=1e907e747f16> (cit. on p. 131).
- [269] Wearden Graeme. *Tech stocks surge as wave of interest in AI drives \$4tn rally*. The Guardian. 2023 May 26 [Cited on 2024 April 04]. The Guardian. URL: <https://www.theguardian.com/technology/2023/may/26/tech->

stock-surge-interest-artificial-intelligence-technology-nvidia-double-value (cit. on p. 131).

- [270] Sebastian Guillen. “The Impact of AI on Company Stock Returns”. In: *Available at SSRN 4595946* (2023) (cit. on p. 131).
- [271] Zacharias Andreadakis et al. “The COVID-19 vaccine development landscape”. In: *Nat Rev Drug Discov* 19.5 (2020), pp. 305–306 (cit. on p. 131).
- [272] Amin Gasmi et al. “A global survey in the developmental landscape of possible vaccination strategies for COVID-19”. In: *Clinical Immunology* 237 (2022), p. 108958 (cit. on p. 131).
- [273] Chad P Bown and Thomas J Bollyky. “How COVID-19 vaccine supply chains emerged in the midst of a pandemic”. In: *The World Economy* 45.2 (2022), pp. 468–522 (cit. on pp. 131, 132).
- [274] Mohammad Arridho Nur Amin. “Policy Analysis of Booster Vaccination against State-Owned Pharmaceutical Stocks”. In: *Jurnal Ekonomi Lembaga Layanan Pendidikan Tinggi Wilayah I* 3.1 (2023), pp. 8–12 (cit. on p. 132).
- [275] Miguel Ampudia and Skander J Van den Heuvel. “Monetary policy and bank equity values in a time of low and negative interest rates”. In: *Journal of monetary economics* 130 (2022), pp. 49–67 (cit. on p. 132).
- [276] Joost V Bats, Massimo Giuliadori, and Aerdt CFJ Houben. “Monetary policy effects in times of negative interest rates: What do bank stock prices tell us?” In: *Journal of financial intermediation* 53 (2023), p. 101003 (cit. on p. 132).
- [277] Christos V Gortsos. “The response of the European Central Bank to the current pandemic crisis: monetary policy and prudential banking supervision decisions”. In: *European Company and Financial Law Review* 17.3-4 (2020), pp. 231–256 (cit. on p. 132).
- [278] Dirk G Baur and Thomas Dimpfl. “Asymmetric volatility in cryptocurrencies”. In: *Economics Letters* 173 (2018), pp. 148–151 (cit. on p. 134).

- [279] Yukun Liu and Aleh Tsyvinski. “Risks and returns of cryptocurrency”. In: *The Review of Financial Studies* 34.6 (2021), pp. 2689–2727 (cit. on p. 134).
- [280] Thomas Conlon, Shaen Corbet, and Richard McGee. “Enduring relief or fleeting respite? Bitcoin as a hedge and safe haven for the US dollar”. In: *Annals of Operations Research* (2024), pp. 1–29 (cit. on p. 134).
- [281] Igor Makarov and Antoinette Schoar. “Trading and arbitrage in cryptocurrency markets”. In: *Journal of Financial Economics* 135.2 (2020), pp. 293–319 (cit. on p. 135).
- [282] Giancarlo Giudici, Alistair Milne, and Dmitri Vinogradov. “Cryptocurrencies: market analysis and perspectives”. In: *Journal of Industrial and Business Economics* 47 (2020), pp. 1–18 (cit. on p. 135).
- [283] Henri Kouam. “Challenges and Implications of Cryptocurrencies, Central Bank Digital Currencies, and Electronic Money”. In: *Corporate Management Ecosystem in Emerging Economies: Global Perspectives*. Springer, 2023, pp. 147–163 (cit. on p. 135).

# Appendix A

## Details of Assets in each Community during each Sub-period

Tables A.1 - A.5 provide the results of community detection on the combination of stocks, cryptocurrencies and US ETFs during different sub-periods, including Pre-Covid-19 (04/2019 - 12/2019), Covid-19 pandemic (01/2020 - 06/2020), Post-Covid-19 (07/2020 - 12/2020), Bull market time (01/2021 - 04/2022), Ukraine-Russia conflict (05/2022 - 05/2023).

Table A.1: Assets in each community during the Pre-Covid-19. The red color indicates cryptocurrencies, the blue color indicates US ETFs and the black color indicates stocks

Community	Assets
1	ACN, ADP, AMT, DTE, EL, INTU, ISRG, LMT, MA, MCD, NEE, NOW, PLD, PYPL, SAP, SBUX, SONY, SO, V, IDU, IUSG, IYR
2	ABBV, AXP, BUD, CMCSA, COST, HON, LIN, MMC, TJX, WMT, OEF, SPY, SVXY
3	ABB, BHP, BLK, CAT, CB, CI, CVS, DE, ELV, ETN, GE, HSBC, ITW, MUFG, TD, UNH, IVE, IYF
4	ABT, BDX, BSX, DHR, HDB, IBN, LLY, MDT, MRK, NVS, PFE, SYK, TMO, ZTS
5	BAT, BNT, DAI, DASH, ETH, OMG, USDT, UST, XLM, XMR, XRP, XVG, ZEC, ZRX
6	ADBE, CRM, GOOG, META, PANW, SHOP, CIBR, FDN
7	ADI, AMAT, AMD, ASX, AVGO, INTC, LRCX, MU, NVDA, QCOM, TSM, TXN, SOXX
8	AIR, BX, CNI, CP, CSL, MC, NVR, ORCL, TMUS, UNP, UPS, IJH, IJJ, IJK, ITB, ITOT, IYT, XTN
9	ALV, BAC, CFR, C, GS, JPM, MS, SCHW, WFC, FTXO, KBWB, RPV
10	AMGN, BMY, GILD, HCA, VRTX, IBB, IJR, IJS, IJT, IWM, SLYG, SLYV, SLY
11	AZN, DEO, KO, MDLZ, NVO, PEP, PG, SNY, TM, UL, FDLO, IWB, IWV, IYJ, IYY, RPG
12	ADA, BCH, BSV, BTC, DOGE, EOS, ETC, LTC, REP
13	AAPL, AMZN, ASML, BABA, BKNG, CSCO, IDEX, MSFT, NFLX, PDD, TSLA, FNGU, IYW, ONEQ, PNQI, QQQ, QTEC, QYLD, ROBT
14	BA, DIS, HD, IBM, JNJ, LOW, NKE, DIA, DJD
15	BKR, BP, COP, CVX, EOG, EQNR, PBR, SHEL, SLB, TTE, XOM, DJP, IEO, IYE
16	BTI, MO, PM, T, VZ, NEO, QTUM, TRX, XTZ, DVY

Table A.2: Assets in each community during the Covid-19 pandemic. The red color indicates cryptocurrencies, the blue color indicates US ETFs and the black color indicates stocks

Community	Assets
1	AAPL, ADBE, AMGN, AMZN, COST, GILD, GOOG, INTC, MSFT, PYPL, SONY, TMUS, VRTX, WMT, <b>IBB</b> , <b>QQQ</b> , <b>QYLD</b>
2	BTI, CB, DE, DTE, ETN, MO, NEE, SO, T, VZ, <b>DVY</b> , <b>IDU</b>
3	ABB, BUD, CI, DEO, DIS, ELV, EL, HD, LOW, MCD, NKE, SBUX, UNH, <b>DIA</b>
4	ABBV, ABT, AZN, BDX, CSCO, DHR, IDEX, JNJ, MA, MRK, NVO, NVS, ORCL, PFE, SNY, TMO, V, <b>IUSG</b> , <b>OEF</b> , <b>ONEQ</b> , <b>SVXY</b>
5	ADI, AMAT, AMD, ASML, ASX, AVGO, LRCX, MU, NVDA, QCOM, TSM, TXN, <b>SOXX</b>
6	ADP, BSX, CAT, CMCSA, CNI, CP, CVS, GE, HON, IBM, ITW, KO, LMT, PM, UNP, UPS, <b>DJD</b> , <b>IVE</b> , <b>IYJ</b> , <b>IYT</b> , <b>XTN</b>
7	INTU, NOW, <b>ADA</b> , <b>BAT</b> , <b>BNT</b> , <b>ETH</b> , <b>NEO</b> , <b>OMG</b> , <b>QTUM</b> , <b>XMR</b> , <b>XTZ</b> , <b>XVG</b> , <b>ZRX</b> , <b>IYW</b> , <b>QTEC</b> , <b>RPG</b>
8	ALV, BA, BHP, BKR, BP, COP, CVX, EOG, EQNR, MUFG, PBR, SHEL, SLB, TD, TTE, XOM, <b>DJP</b> , <b>IEO</b> , <b>IYE</b> , <b>RPV</b>
9	AMT, AXP, BAC, BX, CFR, C, GS, HDB, HSBC, IBN, JPM, MS, PLD, SCHW, TJX, WFC, <b>FTXO</b> , <b>IYF</b> , <b>IYR</b> , <b>KBWB</b>
10	BABA, CRM, META, NFLX, PANW, PDD, SHOP, TSLA, <b>CIBR</b> , <b>FDN</b> , <b>FNGU</b> , <b>PNQI</b>
11	<b>BCH</b> , <b>BSV</b> , <b>BTC</b> , <b>DAI</b> , <b>DASH</b> , <b>DOGE</b> , <b>EOS</b> , <b>ETC</b> , <b>LTC</b> , <b>REP</b> , <b>TRX</b> , <b>XLM</b> , <b>XRP</b> , <b>ZEC</b>
12	AIR, BKNG, CSL, HCA, MC, NVR, TM, <b>IJH</b> , <b>IJJ</b> , <b>IJK</b> , <b>IJR</b> , <b>IJS</b> , <b>IJT</b> , <b>ITB</b> , <b>IWM</b> , <b>ROBT</b> , <b>SLYG</b> , <b>SLYV</b> , <b>SLY</b>
13	ACN, BLK, BMY, ISRG, LIN, LLY, MDLZ, MDT, MMC, PEP, PG, SAP, SYK, UL, ZTS, <b>USDT</b> , <b>UST</b> , <b>FDLO</b> , <b>ITOT</b> , <b>IWB</b> , <b>IWV</b> , <b>IYY</b> , <b>SPY</b>

Table A.3: Assets in each community during the Post-Covid-19. The red color indicates cryptocurrencies, the blue color indicates US ETFs and the black color indicates stocks

Community	Assets
1	BX, ISRG, <b>FDLO</b> , <b>ITOT</b> , <b>IUSG</b> , <b>IWB</b> , <b>IWV</b> , <b>IYY</b> , <b>OEF</b> , <b>ROBT</b> , <b>RPG</b>
2	ABBV, AZN, BMY, GILD, JNJ, LLY, MRK, NVO, NVS, PFE, SNY
3	ABB, ABT, ACN, BDX, BLK, DHR, HD, LOW, MA, ORCL, TMO, V, ZTS, <b>SPY</b> , <b>SVXY</b>
4	ADBE, CRM, META, NFLX, NOW, PANW, PDD, PYPL, SHOP, <b>CIBR</b> , <b>FDN</b> , <b>PNQI</b>
5	ADI, AMAT, ASML, ASX, AVGO, INTC, LRCX, MU, QCOM, TSM, TXN, <b>SOXX</b>
6	ADP, BSX, CAT, CMCSA, CSCO, CVS, DE, ETN, HON, IBM, ITW, KO, LMT, MDLZ, MDT, MMC, PEP, PG, SYK, TJX, <b>DJD</b> , <b>IVE</b>
7	AIR, BA, CSL, GE, HCA, MC, MUFG, TM, <b>IJJ</b> , <b>IJS</b> , <b>RPV</b> , <b>SLYV</b>
8	CI, DIS, ELV, EL, HDB, IBN, LIN, MCD, NKE, SAP, SBUX, UNH, <b>DIA</b>
9	ALV, CNI, CP, NVR, UNP, UPS, <b>IJH</b> , <b>IJK</b> , <b>IJR</b> , <b>IJT</b> , <b>ITB</b> , <b>IWM</b> , <b>IYJ</b> , <b>IYT</b> , <b>SLYG</b> , <b>SLY</b> , <b>XTN</b>
10	AMGN, AMZN, BABA, COST, GOOG, INTU, SONY, TMUS, TSLA, VRTX, WMT, <b>FNGU</b> , <b>IBB</b> , <b>ONEQ</b> , <b>QQQ</b> , <b>QYLD</b>
11	AMT, AXP, BKNG, GS, MS, PLD, <b>IYF</b> , <b>IYR</b>
12	BAC, CFR, C, JPM, SCHW, WFC, <b>DAI</b> , <b>FTXO</b> , <b>KBWB</b>
13	IDEX, <b>BAT</b> , <b>BCH</b> , <b>BSV</b> , <b>DOGE</b> , <b>EOS</b> , <b>ETC</b> , <b>NEO</b> , <b>OMG</b> , <b>QTUM</b> , <b>REP</b> , <b>TRX</b> , <b>XLM</b> , <b>XRP</b> , <b>XTZ</b> , <b>XVG</b> , <b>ZRX</b>
14	AAPL, AMD, MSFT, NVDA, <b>ADA</b> , <b>BNT</b> , <b>BTC</b> , <b>DASH</b> , <b>ETH</b> , <b>LTC</b> , <b>XMR</b> , <b>ZEC</b> , <b>IYW</b> , <b>QTEC</b>
15	BHP, BKR, BP, BUD, COP, CVX, DEO, EOG, EQNR, PBR, SHEL, SLB, TTE, UL, XOM, <b>DJP</b> , <b>IEO</b> , <b>IYE</b>
16	BTI, CB, DTE, HSBC, MO, NEE, PM, SO, TD, T, VZ, <b>USDT</b> , <b>UST</b> , <b>DVY</b> , <b>IDU</b>

Table A.4: Assets in each community during the Bull market time. The red color indicates cryptocurrencies, the blue color indicates US ETFs and the black color indicates stocks

Community	Assets
1	AAPL, ADBE, AMZN, GOOG, INTU, MSFT, TSLA, FNGU, IYW, ONEQ, QQQ, QTEC, QYLD
2	AMT, BTI, CMCSA, CVS, DTE, ETN, HCA, ITW, KO, MO, NEE, PLD, PM, SO, DAI, DJD, IDU, IVE, IYR
3	CNI, CP, MA, UNP, UPS, V, IYJ, IYT, XTN
4	CAT, DE, GE, IBM, LMT, T, VZ, DVY, IJJ, RPV
5	ADI, AMAT, AMD, ASML, ASX, AVGO, INTC, LRCX, MU, NVDA, QCOM, TSM, TXN, SOXX
6	AIR, ALV, BA, CSL, HD, IDEX, LOW, MC, NVR, IJH, IJK, IJR, IJS, IJT, ITB, IWM, SLYG, SLYV, SLY
7	ABT, BDX, DHR, ISRG, SAP, TMO, VRTX, ZTS, IBB, IUSG, OEF, ROBT, RPG
8	ABBV, AMGN, AZN, BMY, GILD, JNJ, LLY, MDLZ, MRK, NVO, NVS, PEP, PFE, PG, SNY, UL, UST
9	ABB, ACN, ADP, BLK, BSX, COST, EL, MDT, ORCL, SYK, TMUS, WMT, FDLO, SPY, SVXY
10	AXP, BAC, CB, CFR, C, GS, HSBC, JPM, MS, MUFG, SCHW, WFC, FTXO, IYF, KBWB
11	BABA, CRM, META, NFLX, NOW, PANW, PDD, PYPL, SHOP, CIBR, FDN, PNQI
12	BHP, BKR, BP, COP, CVX, EOG, EQNR, PBR, SHEL, SLB, TTE, XOM, DJP, IEO, IYE
13	BKNG, BUD, CI, CSCO, DEO, ELV, HON, LIN, MCD, MMC, SBUX, TJX, TM, UNH, DIA
14	TD, ADA, BAT, BCH, BSV, DASH, EOS, ETC, NEO, OMG, QTUM, REP, TRX, USDT, XLM, XMR, XRP, XTZ, ZEC
15	BX, DIS, HDB, IBN, NKE, SONY, ITOT, IWB, IWV, IYY
16	BNT, BTC, DOGE, ETH, LTC, XVG, ZRX

Table A.5: Assets in each community during the Ukraine-Russia Conflict. The red color indicates cryptocurrencies, the blue color indicates US ETFs and the black color indicates stocks

Community	Assets
1	BHP, BKR, BP, COP, CVX, EOG, EQNR, PBR, SHEL, SLB, TTE, XOM, <b>DJP</b> , <b>IEO</b> , <b>IYE</b>
2	ABBV, BMY, JNJ, KO, LLY, MDLZ, MO, MRK, PEP, PFE, PG, PM
3	AMT, BTI, CAT, CB, CMCSA, DE, DTE, ITW, LMT, MMC, NEE, PLD, SO, T, VZ, <b>DVY</b> , <b>IDU</b> , <b>IVE</b> , <b>IYR</b> , <b>RPV</b>
4	ABT, ACN, BDX, BLK, BSX, BX, DIS, EL, HDB, IBN, ISRG, MA, MDT, NKE, SAP, SBUX, SONY, SYK, V, <b>ITOT</b> , <b>SPY</b> , <b>SVXY</b>
5	ADBE, ADI, AMAT, AMD, ASML, ASX, AVGO, GOOG, IDEX, INTC, LRCX, MSFT, MU, NVDA, QCOM, TSM, TXN, <b>IYW</b> , <b>QTEC</b> , <b>ROBT</b> , <b>SOXX</b>
6	BUD, CI, COST, CSCO, CVS, DEO, ELV, HON, IBM, LIN, MCD, TMUS, UL, UNH, WMT, <b>DIA</b> , <b>DJD</b>
7	ADP, ORCL, TM, ZTS, <b>FDLO</b> , <b>IWB</b> , <b>IWV</b> , <b>IYJ</b> , <b>IYY</b> , <b>RPG</b>
8	AIR, <b>IJR</b> , <b>IJS</b> , <b>IJT</b> , <b>IWM</b> , <b>SLYG</b> , <b>SLYV</b> , <b>SLY</b>
9	AMGN, AZN, GILD, NVO, NVS, SNY, VRTX, <b>IBB</b>
10	ABB, ALV, BA, BKNG, CNI, CP, CSL, ETN, GE, HCA, HD, LOW, MC, NVR, TJX, UNP, UPS, <b>IJH</b> , <b>IJJ</b> , <b>IJK</b> , <b>ITB</b> , <b>IYT</b> , <b>XTN</b>
11	AMZN, CRM, NFLX, NOW, PANW, PYPL, SHOP, <b>CIBR</b> , <b>FDN</b> , <b>PNQI</b>
12	AXP, BAC, CFR, C, GS, HSBC, JPM, MS, MUFG, SCHW, TD, WFC, <b>DAI</b> , <b>FTXO</b> , <b>IYF</b> , <b>KBWB</b>
13	<b>ADA</b> , <b>BCH</b> , <b>BNT</b> , <b>BSV</b> , <b>BTC</b> , <b>DASH</b> , <b>DOGE</b> , <b>EOS</b> , <b>ETC</b> , <b>ETH</b> , <b>LTC</b> , <b>NEO</b> , <b>OMG</b> , <b>QTUM</b> , <b>TRX</b> , <b>USDT</b> , <b>XLM</b> , <b>XMR</b> , <b>XRP</b> , <b>XVG</b> , <b>ZEC</b>
14	AAPL, BABA, DHR, INTU, META, PDD, TMO, TSLA, <b>BAT</b> , <b>REP</b> , <b>XTZ</b> , <b>ZRX</b> , <b>FNGU</b> , <b>IUSG</b> , <b>OEF</b> , <b>ONEQ</b> , <b>QQQ</b> , <b>QYLD</b>