PORTFOLIO OPTIMIZATION USING FINANCIAL CHAOS INDEX AND TIME-HOMOGENEOUS TOP-K RANKING

MASOUD ATAEI

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN MATHEMATICS AND STATISTICS
YORK UNIVERSITY
TORONTO, ONTARIO

SEPTEMBER 2020

Abstract

We develop a new stock market index that captures the chaos existing in the market by measuring the mutual changes of asset prices. This new index relies on a tensor-based embedding of the stock market information, which in turn frees it from the restrictive value- or capitalization-weighting assumptions that commonly underlie other various popular indexes. We show that our index is a robust estimator of the market volatility which enables us to characterize the market by performing its regime analysis and the task of segmentation with a high degree of reliability. In addition, we analyze the dynamics and kinematics of the actual market volatility as compared to the implied volatility by introducing a time-dependent dynamical system model. In turn, we will be able to address several long-standing questions on the nature of a relation between equity and option markets. Apart from this, we establish causal relations which exist between our proposed index and an important category of news-based indexes which measure uncertainty in a wide spectrum of economic and financial policies. Another new feature of this work is to define a time-dependent index to capture the overall behavior of the stock market fundamentals. We then design a set of statistical tests which enable us to assess different types of the efficient market hypothesis, revealing weak and semi-strong efficiency of the U.S. stock market for the time frame January 1990-December 2019 using monthly and quarterly time frequencies. In contrast, the stock market during this time frame turned out to be weakly and semi-strongly inefficient as per daily time frequency.

Besides, using the introduced tensor-based embedding of the stock market information, enables us to tackle the problem of time-dependent top-K ranking. More specifically, given N items, all

having positive latent strengths, top-K ranking problem aims to identify the K items ($K \leq N$) receiving the highest ranks based on partially revealed comparisons among the items. This problem has been widely studied in the case for which comparisons are performed independently. However, identifying the top-K rankings becomes a more intricate task when comparisons per se are performed along some temporal dimension. In this work, we further investigate potential impacts of temporality on sequences of top-K items and propose a time-homogeneous ranking scheme. Our framework relies mainly on tensor decompositions, rank centrality, and an innovative continuous extension of the Bradley-Terry-Luce (BTL) model. The proposed continuous BTL model extends the win/loss nature of the logistic model to a continuous setting, further reflecting preference degrees that may exist among the compared items. Our computations, which pertain to the analysis of S&P500 data during the same time period January 2007-December 2019, confirm that the our developed top-K ranking scheme is an effective approach to optimize cardinality-constrained portfolios which involve large volumes of noisy and incomplete data.

KEYWORDS Pairwise Comparisons; Tensor Decompositions; Transfer Entropy; Granger Causality; CBOE Volatility Index; Economic Policy Uncertainty; Equity Market Volatility Tracker; Efficient Market Hypothesis; Stock Market Segmentation and Regime Analysis; Rank Aggregation; top-K ranking; Bradley-Terry-Luce (BTL) model; Rank Centrality; Non-stationary Spatio-temporal Data Analysis; Cardinality-constrained Portfolio Optimization; S&P500.

Acknowledgments

I would like to express my deep sense of appreciation to my supervisors, Professor Michael Chen and Professor Zijiang Yang. Had it not been for their guidance and tremendous support, I would not be able to surmount all obstacles laid in the path of research. Their flexibility and dedication allowed me to explore different areas of mathematics, statistics and computer science that caught my interest while keeping in sight my road map and final goals. Furthermore, I would like to thank Professor Georges Monette and Professor Mike Zabrocki for taking time, to be part of my supervisory as well as examination committee, and for advising me on my research during my Ph.D. studies. I am also grateful to Professor Huang Kai and Professor Radu Campeanu for agreeing to be on my examination committee.

Throughout my study at York University, I benefited a lot from professors and friends who were always helpful. I especially want to thank Professor Vladimir Vinogradov for his unconditional generosity and support, from whom I acquired a profound knowledge of large deviations theory and its related fields of study. My sincere thanks also to Professor Steven Wang, who kindly helped me to build up a solid background in mathematical statistics and artificial intelligence areas. My sincere thanks also to Professor Reza Peyghami for all his support and introducing me to the realm of conic optimization modeling. I would also like to extend my thanks to my Ph.D. peers, Dr. Affan Shoukat, Dr. Allysa Lumley, Justin Miles, Nathan Gold, Mohammed Althubyani, Farid Gassoumov, Georgios Katsimpas, Morgan Wang, Sirous Homayouni and Richard Le. I am also thankful to Primrose Miranda, Anne-Marie Carless, Anne Marie Ridley, Susan Rainey, Madeline Salzarulo and Steven Chen for their administrative support during the whole period of my study.

Last but not least, I want to thank my parents and my beloved sister, Minoo, who supported me thoroughly during the time I was working on my dissertation. I would like to express my deepest gratitude to them who unfailingly helped and encouraged me during the years of my schooling.

Table of Contents

Abstract **ii**

Acknowledgements **iv**

Table of Contents **vi**

List of Tables **ix**

List of Figures **x**

1 Introduction **1**

1.1 Financial Chaos Index 1

1.2 Time-homogeneous Top-K Ranking 4

1.3 Outline 8

1.4 Notations and Preliminaries 11

2 Financial Chaos Index **14**

2.1 Phenomenology of Comparative judgments 14

2.1.1 Pairwise Comparison Matrices 15

2.1.2 Pairwise Comparison Tensors 27
2.2 An Index of Financial Chaos 36
2.2.1 Formulation of the Index 36
2.2.2 Non-stationarity Analysis 39
2.2.3 Tests for Stationarity 40
2.2.4 Random Walk Tests 44
2.3 Relation to Market Regimes 46
2.3.1 Regime Analysis 46
2.3.2 Market Segmentation 55
2.4 Relation to Market Volatility 59
2.4.1 Robust Estimation of the Market Volatility 60
2.4.2 Cointegration Analysis 62
2.4.3 Information Kinematics 65
2.4.4 Information Dynamics 72
2.5 Relation to Market Forces 81
2.5.1 Extended Granger Causality 83
2.5.2 Analysis of Daily Data 86
2.5.3 Analysis of Monthly Data 89
2.6 Relation to Market Fundamentals 96
2.6.1 Feature Selection 96
2.6.2 Market Fundamentals Index 100
2.6.3 Tests of Statistical Independence 103

2.6.4 Block-wise Granger Causality 106
2.7 The Efficient Market Hypothesis 111
2.7.1 Weak EMH 112
2.7.2 Potent EMH 112
2.7.3 Semi-strong EMH 113

3 Time-homogeneous Top-K Ranking **115**
3.1 Time-independent Top-K Ranking 115
3.1.1 Continuous Extension to BTL Model 116
3.1.2 Rank Centrality 117
3.2 Time-homogeneous Top-K Ranking 118
3.2.1 Shape-constrained Polyadic Decomposition 120
3.3 Computational Results 122
3.3.1 Evaluation Criteria 124
3.3.2 Time-homogeneous vs. Time-independent Top-K Ranking 124
3.3.3 Optimal Investment Strategies 125

4 Conclusion **133**

Bibliography **136**

List of Tables

2.1 Results of ADF unit-root test for daily FCI_t. 41

2.2 Results of ADF unit-root test for monthly FCI_t. 41

2.3 Results of ADF unit-root test for quarterly FCI_t. 41

2.4 Results of KPSS test for daily, monthly and quarterly FCI_t. 44

2.5 Results of variance ratio test for daily, monthly and quarterly FCI_t. 46

2.6 Results of regime analysis for daily FCI_t. 54

2.7 Time periods corresponding to obtained temporal segments of the market. 58

2.8 Summary statistics for obtained temporal segments of the market. 59

2.9 Results of Johonson's cointegration test for daily FCI_t and VIX_t. 65

2.10 Statistics computed for the network of monthly extended Granger causal relations. . 94

2.11 Computed p-values for strictly causal relations among monthly time series. 95

2.12 Table of selected fundamental features for the market. 99

2.13 Results of independence tests between quarterly FCI_t and MFI_t. 106

2.14 Results of Granger causality test for $\mathrm{H}_{0,\mathtt{Q}}^{\mathtt{NEWS}}$. 109

2.15 Results of Granger causality test for $\mathrm{H}_{0,\mathtt{Q}}^{\mathtt{PUB}}$. 110

List of Figures

1.1 The structure of Chapter 2 illustrated using a flowchart. 10

2.1 Schematic of a k-length pairwise comparison matrix. 17

2.2 Schematic of a reciprocal pairwise comparison tensor. 34

2.3 Plot of the relative errors for rank-1 polyadic decomposition. 34

2.4 Plot of the average inconsistency measures. 35

2.5 Plot of the regularity measures. 35

2.6 Plot of the daily FCI_t. 38

2.7 Plot of the annotated monthly FCI_t. 38

2.8 Frequency histogram for daily FCI_t. 48

2.9 Probability density functions of market regimes. 53

2.10 Schematic of the obtained segments for the market. 58

2.11 Plots of the daily FCI_t vs VIX_t. 63

2.12 Plot of the long-run equilibrium for daily FCI_t-VIX_t. 67

2.13 Plot of the sample cross-correlation for daily FCI_t-VIX_t. 68

2.14 Plot of the orthogonalized IRF for $\mathrm{FCI}_t \rightsquigarrow \mathrm{VIX}_t$. 72

2.15 Plot of the orthogonalized IRF for $\mathrm{VIX}_t \rightsquigarrow \mathrm{FCI}_t$. 73

2.16 Computed diagram of information flow between FCI_t and VIX_t. 75

2.17 Symbolic diagram of information flow between FCI_t and VIX_t. 77

2.18 Diagram of extended Granger causal relations among daily time series. 87

2.19 Diagram of extended Granger causal relations among monthly time series. 89

2.20 Heatmap of empirical probability measures for instantaneous causal relations among monthly time series. 90

2.21 Heatmap of extended Granger causal measures among monthly time series. 90

2.22 Schematic of a third-order market fundamentals design tensor. 101

2.23 Plot of the relative error of polyadic decomposition for market fundamentals design tensor. 101

2.24 Plot of the quarterly FCI_t vs MFI_t. 103

2.25 Plot of the sample cross-correlation for quarterly FCI_t-MIX_t. 104

3.1 Time-independent vs. time-homogeneous top-K rankings under discrete BTL. . . . 126

3.2 Time-independent vs. time-homogeneous top-K rankings under continuous BTL. . 126

3.3 Time-homogeneous top-K rankings; discrete vs. continuous BTL models. 127

3.4 Overall consistency measures for time-homogeneous and true top-K rankings. . . . 128

3.5 Scatter plots for consecutive consistency measures evaluated at $K = K^*$. 129

3.6 Monthly average returns with no transaction fee. 130

3.7 Monthly average returns with fixed 0.5% transaction fee. 130

3.8 Monthly average returns with fixed 1.0% transaction fee. 131

3.9 Cumulative monthly average returns for time-homogeneous top-K. 131

3.10 Cumulative monthly average returns for true top-K ranking. 132

3.11 Cumulative absolute errors for monthly average returns. 132

Chapter 1

Introduction

This dissertation is organized in two main chapters as described below.

1.1 Financial Chaos Index

In Chapter 2, we set out to analyze the stock markets by resorting to the theory and applications of tensors. We develop a special class of spatio-temporal tensors and show that our created algebraic object is a suitable structure to embed the collective judgment of agents who are present in the market. It is then the use of such effective embedding of the stock market information that allows us to take advantage of different mathematical and statistical properties of the constructed tensors, in order to define an index which captures the chaos existing in the market by measuring the mutual changes of asset prices. The proposed financial chaos index is shown to be indeed a robust estimator of market volatility, which in turn enables us to reliably utilize it for various purposes including analysis of market regimes, performing the task of market segmentation and testing the efficient market hypothesis.

To begin with, we formulate our tensor-based model of the stock market on the basis of a rich body of literature that has been conducted by numerous scholars during the past decades who have

developed the theory of pairwise comparative judgments to its near maturity stage. This theory thus far has mainly relied on mathematical aspects of the so-called pairwise comparison matrices and their special subclass widely known as the reciprocal pairwise comparison matrices, which finds various application in different areas of science including the analytic hierarchy processes, e.g., see (Brunelli, 2014; Ishizaka and Labib, 2009; Ishizaka and Labib, 2011; Mu and Pereyra-Rojas, 2016; Mu and Pereyra-Rojas, 2017a; Mu and Pereyra-Rojas, 2017b; Saaty, 2001a; Saaty, 2001b; Saaty and Vargas, 2012) for review of the foundations and recent developments. Our primary contribution in this work is to extend the theory of pairwise comparative judgments to a tensor domain, using the invaluable findings available on the pairwise comparison matrices and their various properties.

To achieve the above-mentioned goal, we first axiomitize the pairwise comparative judgments, and in light of the defined axiom, we state and re-interpret the most important results that have already been obtained on mathematical and statistical properties of the pairwise comparison matrices and their special subclass of reciprocal matrices in the literature. By re-examining the various properties of the mentioned comparison matrices, we demonstrate that these mathematical objects are among the most suitable structures, if not the only ones, to model the procedures that underlie agents' thought processes, especially those that relate to their judgment mechanisms. We then extend the comparison matrices to higher dimensions and establish an important result on relation of the polyadic decomposition of our constructed tensors and the consistency of the judgments cast by the agents throughout time. Once having constructed the mentioned tensors based upon the pairwise comparison matrices and investigated their various algebraic properties, we utilize them to lay out the foundations for defining the financial chaos index.

Being a robust estimator of the market's actual volatility, subsequently raises an important question on the relation of our proposed financial chaos index as compared to other indexes that rather measure the market's implied volatility. Our further contribution in this work is to examine the relationship between the financial chaos index and VIX, which is a predominant index for measuring the market's implied volatility. We note that other efforts have also been made to characterize VIX and investigate its potential roles in the financial markets, e.g., see (Ben-David,

Graham, and Harvey, 2013; Drechsler and Yaron, 2011; Forbes and Warnock, 2012; Giglio and Kelly, 2018; Martin, 2017; Nagel, 2012; Nagel, 2016; Rey, 2015) for a number of notable research papers concerning the VIX in the literature.

Investigating the relationship between the financial chaos index and VIX, in turn enables us to characterize the relation between the equity and option markets, and address several long-standing questions in finance. For instance, whether there exist a relationship between an option contract written on an asset and its associated price, or for another example, whether options can be viewed as assets being independent from one another on the option market whose prices depend solely on their underlying equities. We address these questions and some other relevant ones in this chapter by systematically studying the kinematics and dynamics of the financial chaos index and VIX using a variety of tools. Namely, we characterize the relationship between these two indexes by carrying out their cointegration analysis, which in turn leads to identification of their long-run equilibrium relationship. Furthermore, by using the orthogonalized impulse-response functions, we determine the possible causal relations existing between two indexes. Also, we formulate a time-dependent dynamical system to model the flow of various types of information such as self-entropy, transfer entropy and approximate entropy between the considered indexes over time.

Another important line of research to which we further contribute in this chapter, concerns the possible causal relations existing between the financial chaos index and a set of indexes which measure uncertainty in various types of policies. What amplifies the importance of our analysis pertains to the fact that the selected policy uncertainty indexes in this work, are developed based on the text mining methods, mainly using the information reflected in the daily newschapters, which itself is a rapidly growing area of research, e.g., see (Baker, Bloom, and Davis, 2016; Baker, Bloom, Davis, and Kost, 2019; Davis and Seminario, 2019; Gentzkow, Kelly, and Taddy, 2019; Hassan, Hollander, Lent, and Tahoun, 2019; Kelly, Manela, and Moreira, 2019; Manela and Moreira, 2017) for survey of recent developments and applications of the text-based methods in the literature.

In this respect and by resorting to the framework of Granger causality, we characterize the

direction of causal relations existing among various news-based policy uncertainty indexes and the stock market's actual volatility index represented by the financial chaos index. Our goal for carrying out such investigation is to unravel the complex set of relations that potentially exist among the considered time series, whose interpretations could possibly benefit various sectors of economy, finance, government and health care, to mention but a few. For instance, in the sequel we show that based on an analysis conducted using the information given for a time period January 1990-December 2019, the health care policy uncertainty is Granger caused by the trade, general regulations and entitlement programs policy uncertainties. However, what stands out as rather a surprising result is that the health care policy uncertainty itself is shown to Granger cause the security, economic, government spending and monetary policy uncertainties as well as the equity market volatility tracker. These outcomes are very much tangible when one considers the case of the COVID-19 and how sufficiently well the causal relations inferred by our analysis using the data from the time period January 1990-December 2019, extrapolate to the circumstances experienced around the globe, and in particular in the U.S., during 2020.

1.2 Time-homogeneous Top-K Ranking

In Chapter 3 we gear our focus towards the top-K ranking problem and develop an algorithm to solve the cardinality-constrained optimization problems. For a large collection of N items such as documents, images, movies or stocks, the problem of inferring a ranking over the items is frequently encountered in the areas related to machine learning and recommendation systems. This problem which is commonly referred to as *rank aggregation* has received a considerable amount of attention in the literature when the goal is to aggregate rankings from *pairwise comparisons*, which itself is particular case of the rank aggregation from *multi-wise comparisons*.

Rank aggregation from paired comparison data arises in many real-world applications in which comparing the items in pairs for the purpose of establishing a partial ranking (ordering) among them, renders a more practical approach than ranking the full list of them. However, broadly

speaking, the main incentives to analyze the compared data fall under one of the following situations:

- The data provided are already in the form of comparisons;
- The comparisons emerge *by design*.

The former situation is frequently encountered when the quantities ascribed to items do not bear a direct significant interpretation, unless the items are being compared to each other. For instance, the judgment on whether the variance of a random variable is considerably high is only subjective. However, the pairwise (or multi-wise) comparisons of variances of a collection of random variables can potentially reveal more detailed information. In fact, the method of comparison is often used to create a frame of reference to which items can be related (Raivola, 1985).

Another possibility to encounter the former situation mentioned above is when each item possesses a set of qualities that cannot be quantified; e.g., there is no direct way to measure the strength of an athlete. In these cases, comparison serves as a practical tool in providing a framework to make judgments on qualities of the items relative to each other. On the other hand, it is quite possible that the data provided are not already in the form of comparisons, but the methodologies developed for the corresponding analysis rely on compared data. In this case, the pairwise (or multi-wise) comparisons are typically synthesized using the data provided, which leads to the situation pertaining to comparisons by design.

To tackle the pairwise rank aggregation problem, various approaches have been introduced in the literature, which could essentially be categorized as follows:

- **Non-active ranking:** In this setting, the pairs of items for comparison are fixed for querying. This category is mainly studied under the following dichotomy:
  - A true global ranking exists over the items. Hence, some probabilistic model could be considered whose outcomes are the observed pairwise comparisons. Most known works focus on the following parametric families:

    - Bradley-Terry-Luce (BTL) model (Bradley, 1954; Bradley, 1955; Bradley and Terry, 1952; Luce, 1959);
    - Thurstone model (Thurstone, 1927);
    - Plackett-Luce model (Luce, 1959; Plackett, 1975);
    - Mallows model (Mallows, 1957).
  - A true global ranking does not necessarily exist over the items. Most influential works pertain to the *minimum feedback arc set tournaments*; see (Alon, 2006) and references therein for more details.

- **Active ranking:** In this setting, the pairs of items for comparison could be selected in an adaptive fashion. Most dominant streams of research in this direction are:
  - Distance-based models (Ailon, 2012; Jamieson and Nowak, 2011; Wauthier, Jordan, and Jojic, 2013);
  - BTL-based models (Ailon, Charikar, and Newman, 2008; Chen, Bennett, Collins-Thompson, and Horvitz, 2013);
  - Plackett-Luce based models (Szörényi, Busa-Fekete, Paul, and Hüllermeier, 2015);
  - Non-parametric models (Heckel, Shah, Ramchandran, and Wainwright, 2016; Mohajer and Suh, 2016).

There is also a vast literature on the methods developed for solving the multi-wise rank aggregation problems; see (Alvo and Philip, 2014; Chen, Li, and Mao, 2018) and references therein for the survey of literature which pertains to this stream of research. Besides, each of the above-mentioned sub-categories itself can further be studied under the following situations:

- The data provided are high-dimensional;
- The data provided are incomplete;
- The data provided are noisy.

In this chapter, we set out to study the potential impacts of temporality on sequences of top-K items. To the best of our knowledge, this problem has not yet been fully exploited in the literature. More specifically, we will concentrate on the non-active rank aggregation from pairwise comparisons under the BTL model, and propose a time-homogeneous top-K ranking scheme which can effectively address the high-dimensional, noisy and incomplete data. In our proposed scheme, the pairwise compared data are required by design. This means that the raw data provided do not necessarily need to be in the form of comparisons. However, our proposed scheme is still applicable to those problems in which the raw data are presented as pairwise comparisons. In what follows, we briefly discuss the underlying reasons for studying the rank aggregation problem under the above-mentioned restrictions.

Since this is quite a long-standing area of research, there is already a large volume of literature on the BTL model and its applications to pairwise rank aggregation problems. For instance, a comprehensive review of the BTL model and its extensions with the emphasis on dependent data structures is presented in (Cattelan, 2012). The authors of (Negahban, Oh, and Shah, 2017) proposed the *rank centrality algorithm* to obtain a global ranking over the items under the BTL model, and it was proved in (Chen, Gopi, Mao, and Schneider, 2017; Chen, Fan, Ma, and Wang, 2017; Jang, Kim, Suh, and Oh, 2016) that the above rank aggregation algorithm recovers the top-K items perfectly under some optimal sample complexity. The rank centrality algorithm has been built upon a direction of research, which is commonly referred to as *spectral ranking*, in which eigenvectors of certain matrices are employed to find global rankings of items; see (Vigna, 2016) for more details on spectral ranking. Other spectral methods for top-K ranking under the BTL model have also been proposed in the literature; see (Chen and Suh, 2015; Shah and Wainwright, 2015; Suh, Tan, and Zhao, 2017) and references therein for more details.

The rank centrality algorithm is the cornerstone of our proposed time-homogeneous ranking scheme. The underlying reasons for making such a choice can be clarified as follows: In fact, what makes the rank centrality distinct from its predecessors in spectral ranking pertains to its use of a particular form of matrix which embeds the information related to a *Markov chain* corresponding to a *random walk* on a graph. In turn, it is the reliance on Markov chains and random walks which

makes the rank centrality a suitable candidate to be incorporated to the studies of temporal top-K ranking problem.

Moreover, resorting to the rank centrality algorithm enables us to restrict the scope of this research to the study of non-active rank aggregation problems and their temporal behavior. This is mainly due to the results reported in (Negahban, Oh, and Shah, 2017) whose authors showed that the performance of the rank centrality will be comparable to the best model-aware algorithm inasmuch as data are generated as per a reasonable model. As a consequence, we omit the study of active rank aggregation problems in this work, and gear our main focus towards the non-active rank aggregation.

Even though the rank centrality plays a crucial role in our proposed scheme, it is the tensorial formulation of the top-K ranking problem that provides the necessary framework for its applications. Formulating the problem in terms of tensors can potentially provide an opportunity to employ decomposition techniques for different purposes such as smoothing, imputation, clustering, segmentation, etc. In this chapter, we employ the polyadic decomposition within our proposed ranking scheme. It will be shown that the applications of such a tensor decomposition technique can lead to development of representative, interpretable and effective models for time-homogeneous top-K ranking problems. A shape-constrained polyadic decomposition model alongside the rank centrality algorithm will then constitute the main components of the proposed time-homogeneous top-K ranking scheme.

1.3 Outline

Chapter 2 and its structure are organized as follows:

In Section 2.1, we define the axiom of pairwise comparative judgment, under which various algebraic properties of the pairwise comparison matrices, and in particular reciprocal pairwise comparison matrices, are elucidated. In this section, we further define a third-order spatio-temporal tensor object using the so-called constitution criterion which embeds the stock market information,

which is then followed by the definition of the so-called consensus tensor.

In Section 2.2, we employ the defined consensus tensor to define the financial chaos index. Thereafter, we characterize it by resorting to unit-root, stationarity and random walk tests. Using the properties of the financial chaos index, we then perform the tasks of regime analysis and market segmentation which are the focus of Section 2.3. Besides, the relation of the financial chaos index with market volatility is explicated in Section 2.4, where various tools are employed to characterize the relationship between our proposed index and the market's implied volatility index (VIX).

In Section 2.5, we further taken into account policy uncertainty indexes alongside financial chaos index and VIX, and utilize extended Granger causality to analyze them. Furthermore, given the importance of the assets' fundamentals, we devote Section 2.6 to develop an index which can effectively capture the overall behavior of stock market fundamentals, whose relation to financial chaos index and other news-based policy uncertainty indexes are also characterized using 1-step blockwise Granger causality. Finally, an amalgam of all the presented material in the chapter manifest themselves in Section 2.7, where statistical tests are designed to examine efficient market hypothesis in its various forms. A schematic road map of the chapter is shown in Figure 1.1.

Chapter 3 and its structure are also organized as follows:

In Section 3.1, the notion of *pairwise advantage* is introduced, and a continuous extension to the conventional discrete BTL model is proposed which reflects possible degrees of preferences among the compared items. Unlike the discrete BTL model that considers only the partial ordering, the continuous extension takes into account the difference between each partial ordering pair as well. Subsequently, the top-K items are identified from the weighted partial ordering using the rank centrality algorithm.

More specifically, the pairwise advantages of the items will be used to construct a reciprocal pairwise comparison matrix at every temporal step holding the spatial relations among the items. Each of these matrices, will then be considered as a sample input to the rank centrality algorithm when identifying a sequence of time-independent top-K items. However, in order to obtain a set of time-homogeneous top-K ranked items, as opposed to the time-independent variation, the

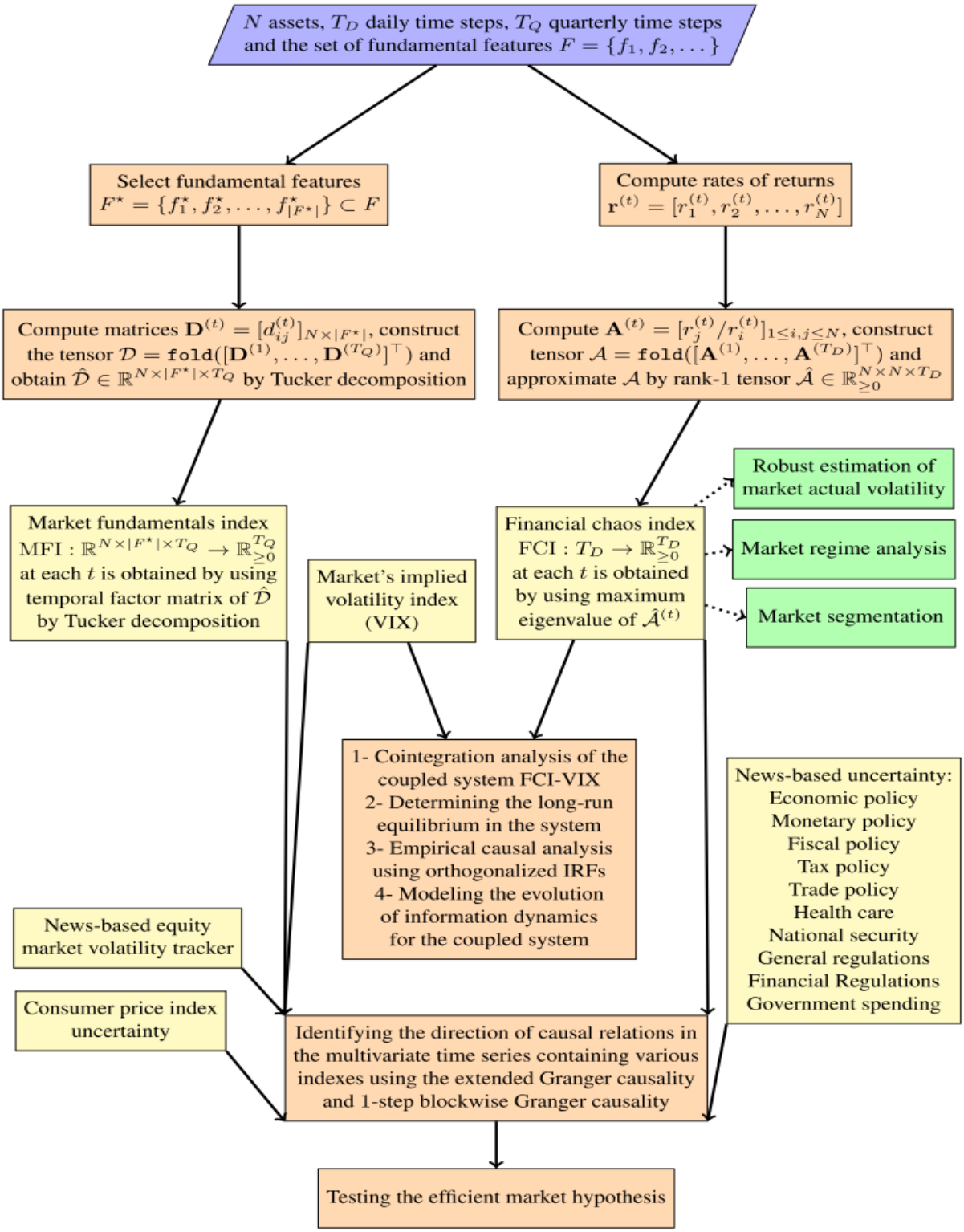

Figure 1.1: The structure of Chapter 2 illustrated using a flowchart, where yellow boxes pertain to various types of indexes, whereas the orange boxes describe the processes involved.

reciprocal pairwise comparison matrix over the time will be treated as a third-order spatio-temporal tensor resulting in a reciprocal pairwise comparison tensor.

A shape-constrained polyadic decomposition model is then employed and solved to smooth out the constructed RPCT while preserving its underlying structure. In turn, the applications of such smoothed RPCT to top-K ranking procedures yield consistent sequences of top-K items. The tensorial formulation of the problem and the related smoothing procedures are discussed in Section 3.2. Finally, computational results for a case study on S&P500 and its related cardinality-constrained portfolio optimization problem are given in Section 3.3, where a set of criteria is developed based on the *Hamming distance* in order to measure consistencies of the top-K ranking sequences.

1.4 Notations and Preliminaries

Throughout the dissertation, scalar quantities, vectors and matrices are denoted using lowercase lightface (e.g., x), lowercase boldface (e.g., $\mathbf{x}$) and uppercase boldface (e.g., $\mathbf{X}$) letters, respectively, where all vectors are presumed to be column ones. Also, boldface calligraphic letters (e.g., $\boldsymbol{\mathcal{X}}$) are used to denote the tensors, whereas their lightface counterparts (e.g., $\mathcal{X}$) would often represent the induced sample spaces. We further use $\mathbb{R}_{\geq 0}$ and $\mathbb{R}_{>0}$ to denote the fields of positive and strictly positive real numbers, respectively, while $\mathbb{C}$ is used to denote the field of complex numbers.

Tensors are algebraic objects which extend the notion of matrices to higher dimensions. The number of dimensions of a tensor is referred to as the *order* or *modes* of the tensor, i.e., $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$ indicates an Nth-order tensor having $K = \prod_{n=1}^{N} I_n$ elements in total. Analogous to matrix rows and columns, *mode-n fibers* of tensors are derived by fixing every index but one, while fixing all but two indexes of the tensors yields hyperplanes known as *slices*. The Frobenius norm of $\boldsymbol{\mathcal{X}}$ is further defined as follows:

$$\|\boldsymbol{\mathcal{X}}\|_F = \sqrt{\langle \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{X}} \rangle} = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_n=1}^{I_N} \sqrt{x_{i_1 i_2 \ldots i_n}^2}\,. \tag{1.1}$$

In addition, the unfolding of tensors is to reshape them into more mathematically tractable lower-dimensional representations. For the tensor $\boldsymbol{\mathcal{X}}$, the *vector unfolding* (also called *vectorization*) yields the vector

$$\mathbf{x} = \texttt{vec}(\boldsymbol{\mathcal{X}}) \in \mathbb{R}^K, \tag{1.2}$$

while reshaping the tensor $\boldsymbol{\mathcal{X}}$ into a matrix having the form

$$\boldsymbol{\mathcal{X}}_{(n)} = [\mathbf{f}_1^{(n)}, \mathbf{f}_2^{(n)}, \ldots, \mathbf{f}_{\frac{K}{I_n}}^{(n)}], \tag{1.3}$$

results in a *mode-n unfolding* (also called *matricization*) of $\boldsymbol{\mathcal{X}}$, where the column vector $\mathbf{f}_i^{(n)}$ represents the i-th *mode-n fiber* such that the matricized tensor has dimensions $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_n \times \frac{K}{I_N}}$.

Different types of products can also be defined for tensors, among which the *mode-n product* finds vast applications. Given a matrix $\mathbf{U} \in \mathbb{R}^{J \times I_n}$, the mode-n product of $\boldsymbol{\mathcal{X}}$ with $\mathbf{U}$ is defined by

$$(\boldsymbol{\mathcal{X}} \times_n \mathbf{U})_{i_1 \ldots i_{n-1} j i_{n+1} \ldots i_n} = \sum_{i_n=1}^{I_n} x_{i_1 i_2 \ldots i_N} u_{j i_n}, \tag{1.4}$$

where the resulting tensor has dimensions $I_1 \times \ldots I_{n-1} \times J_n \times I_{n+1} \ldots I_N$. In other words, the mode-n product of $\boldsymbol{\mathcal{X}}$ with $\mathbf{U}$ is obtained by left multiplication of the matrix $\mathbf{U}$ by each one of the mode-n fibers of $\boldsymbol{\mathcal{X}}$, i.e.,

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} \times_n \mathbf{U} \Longleftrightarrow \boldsymbol{\mathcal{Y}}_{(n)} = \mathbf{U}\boldsymbol{\mathcal{X}}_{(n)}. \tag{1.5}$$

An important approach to envisage tensors, is by considering them as a stack of their frontal slices. This perspective is particularly useful when one deals with third-order spatio-temporal tensors, where the frontal slices correspond to different points of time within the temporal dimension of the tensor. For instance, the frontal slices of the tensor $\boldsymbol{\mathcal{Z}} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ are denoted by $\boldsymbol{\mathcal{Z}}^{(n)} \in \mathbb{R}^{I_1 \times I_2}$ for $n = 1, 2, \ldots, I_3$. Subsequently, we may define an `unfold` operator as follows:

$$\texttt{unfold}(\boldsymbol{\mathcal{Z}}) := [\boldsymbol{\mathcal{Z}}^{(1)}, \boldsymbol{\mathcal{Z}}^{(2)}, \ldots, \boldsymbol{\mathcal{Z}}^{(I_3)}]^\top, \tag{1.6}$$

for which a `fold` operator cabn further be defined satisfying the relation $\texttt{fold}(\texttt{unfold}(\boldsymbol{\mathcal{Z}})) := \boldsymbol{\mathcal{Z}}$.

Furthermore, using the notion of mode-n tensor product, one can employ the Tucker decomposition and approximate the third-order tensor $\boldsymbol{\mathcal{Z}}$ using a dense core tensor $\boldsymbol{\mathcal{H}} \in \mathbb{R}^{R_1 \times R_2 \times R_3}$ with multi-linear rank (R_1, R_2, R_3) and a set of factor matrices $\mathbf{A} \in \mathbb{R}^{I_1 \times R_1}$, $\mathbf{B} \in \mathbb{R}^{I_2 \times R_2}$ and $\mathbf{C} \in \mathbb{R}^{I_3 \times R_3}$, i.e.,

$$\boldsymbol{\mathcal{Z}} \approx \widehat{\boldsymbol{\mathcal{Z}}} = \boldsymbol{\mathcal{H}} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}. \tag{1.7}$$

Another technique for tensor decomposition which finds a wide range of applications in science is the *polyadic decomposition* which approximates $\boldsymbol{\mathcal{Z}}$ by the sum of rank-1 tensors (also called *atoms*), i.e.,

$$\boldsymbol{\mathcal{Z}} \approx \widehat{\boldsymbol{\mathcal{Z}}} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r, \tag{1.8}$$

where the notation "$\circ$" represents *vector outer product*, and R is a given positive integer known as the rank of $\widehat{\boldsymbol{\mathcal{Z}}}$. It is worth mentioning that in the case where R is identified to be a minimal rank, the decomposition given by equation (1.8) is referred to as the *canonical polyadic decomposition* (CPD). Furthermore, the tensor $\boldsymbol{\mathcal{Z}}$ is said to be a rank-1 tensor if it can be written as the outer product of three vectors, i.e.,

$$\boldsymbol{\mathcal{Z}} = \mathbf{a} \circ \mathbf{b} \circ \mathbf{c}, \tag{1.9}$$

where $\mathbf{a} \in \mathbb{R}^{I_1}$, $\mathbf{b} \in \mathbb{R}^{I_2}$ and $\mathbf{c} \in \mathbb{R}^{I_3}$.

Furthermore, the Hadamard product of two matrices with the same size is denoted by "$*$" and is defined as their elementwise product, and the *indicator function* is defined as $\mathbb{1}(x) = 0, 1$ if $x \leq 0$ or $x > 0$, while the notation x^+ is understood as $x \vee 0 = \max(x, 0)$.

Chapter 2

Financial Chaos Index

2.1 Phenomenology of Comparative judgments

We consider a general exchange economy consisting of a universe $S := \{S_1, S_2, \dots, S_N\}$ of N assets traded by $M \gg 1$ agents over a finite time horizon $\mathcal{T} := \{t_1, t_2, \dots\} \subset \{1, 2, \dots, T\}$, where T denotes the maximum number of discrete time periods under consideration. Given the set of alternatives S, we are concerned with the situation where an agent, say m, performs a complete set of pairwise comparisons on alternatives contained in S and assigns a strictly positive quantity $\mu_m^{(t)}$ to her *degree of preference* on every comparison at time $t \in \mathcal{T}$. We will show that under mild conditions of the so-called *pairwise comparative judgment* axiom and the strict positivity of $\mu_m^{(t)}$ for every $t \in \mathcal{T}$, the preference gauge assigned by agent m to every comparison is bound to take on form of a ratio between two strictly positive quantities. Various consequences of such an implication and its phenomenological aspects that underlie the faculty of judgment in human beings will be also discussed throughout this section.

2.1.1 Pairwise Comparison Matrices

Let us assume that for a generic agent m, there exist a directed and complete edge-weighted *comparison multigraph* $\mathcal{G}_m^{(t)} = (S, \mathcal{E}_m^{(t)})$ with N nodes, whose edge set $\mathcal{E}_m^{(t)}$ contains the preference degrees assigned by the agent to every pair of the assets that are compared at time $t \in \mathcal{T}$. A *walk* or *path* on $\mathcal{G}_m^{(t)}$ is defined as a sequence $S_1, S_2, \ldots, S_k$ of nodes where $(S_i, S_{i+1}) \in \mathcal{E}_m^{(t)}$ for $i = 1, 2, \ldots, k-1$, whose length is determined by the number of edges traversed to reach the terminal node starting from the initial node. In this setting, we presume that traversing a self-loop would increase the walk length by an increment. By a cycle we understand a closed walk with identical initial and terminal nodes. The weight matrix associated to $\mathcal{G}_m^{(t)}$ is defined as follows:

Definition 2.1.1 (PCM). *A square matrix $\mathbf{A} \in \mathcal{A}_{\mathrm{PCM}}^{(t)} \subset \mathbb{R}_{>0}^{N \times N}$ is said to be a pairwise comparison matrix (PCM) if its elements elicit the agent's preference degrees on pairs of assets that are compared at a given time $t \in \mathcal{T}$, where $\mathcal{A}_{\mathrm{PCM}}^{(t)}$ represents the family of all pairwise comparison matrices at time t.*

According to this definition, the (i, j)-th element of the k-th power $\mathbf{A}^k$ of the pairwise comparison matrix $\mathbf{A}$ can be interpreted as the preference degree of asset S_j over S_i based on a walk of length k performed on the comparison multigraph $\mathcal{G}_m^{(t)}$. Then completeness of $\mathcal{G}_m^{(t)}$ guarantees that the weight matrix $\mathbf{A}$ is irreducible as every pair of nodes are connected by a path of arbitrary a certain length. Moreover, since every pair of the nodes in $\mathcal{G}_m^{(t)}$ is connected via a certain path of length k, the matrix is primitive, and such a primitivity property strengthens the irreducibility condition by further positing k-connectivity of $\mathcal{G}_m^{(t)}$.

Let us now introduce the relations $\prec_k$, $\sim_k$ and $\preceq_k$ which denote *k-length strict preference*, *indifference* and *preference-indifference* relations on S, respectively. These relations, which extend those proposed in (Fishburn, 1970; Suppes, 1999) for the special case $k = 1$, which are defined as follows:

$$\forall \mu_m^{(t)} \in R_{>0}, \exists \prec_k, \exists \sim_k \colon (S_i \prec_k S_j) \vee (S_j \prec_k S_i) \vee (S_i \sim_k S_j). \tag{2.1}$$

The formula (2.1) can be perceived as the fact that a generic agent m believes that for all her

subjective preference degrees measured at a given time $t \in \mathcal{T}$, there exist relations $\prec_k$ and $\sim_k$ such that the asset S_j would be either strictly better than the asset S_i based on a walk of length k on $\mathcal{G}_m^{(t)}$ (e.g., $S_i \prec_k S_j$) or strictly worse (e.g., $S_j \prec_k S_i$), or S_i and S_j should be judged as being indifferent from one another (e.g., $S_i \sim_k S_j$). We assume that $\prec_k$ is a strict partial order and define the k-length indifference $\sim_k$ as the absence of strict preference such that

$$x \sim_k y \Leftrightarrow (x \not\prec_k y) \wedge (y \not\prec_k x). \tag{2.2}$$

We are then ready to formulate the so-called pairwise comparative judgment axiom as follows:

Axiom (PCJ). *An agent m is said to satisfy a pairwise comparative judgment (PCJ) axiom, if for every $t \in \mathcal{T}$, $k \geq 1$, $i, j, l \in \{1, 2, \ldots, N\}$ and $\mu_m^{(t)} > 0$, there exist relations $\prec_k$, $\preceq_k$ and $\sim_k$ such that*

(i). $S_i \prec_k S_j \iff \mu_m^{(t)}(S_i \prec_k S_j) > \mu_m^{(t)}(S_j \prec_k S_i)$;

(ii). $S_i \sim_k S_j \iff \mu_m^{(t)}(S_i \sim_k S_j) = \mu_m^{(t)}(S_j \sim_k S_i)$;

(iii). $\mu_m^{(t)}(S_i \preceq_k S_j) = \mu_m^{(t)}(S_i \preceq_k S_l)\, \mu_m^{(t)}(S_l \preceq_k S_j)$.

The properties (i) and (ii) stated in the PCJ Axiom imply that for every $x, y, z \in S$, the k-length preference $\prec_k$ and indifference $\sim_k$ relations satisfy the following properties

$$\begin{aligned}
(x \sim_k x), &\qquad \text{reflexivity of indifference;} \\
\neg(x \prec_k x), &\qquad \text{irreflexivity of preference;} \\
(x \sim_k y) \implies (y \sim_k x), &\qquad \text{symmetry of indifference;} \\
(x \prec_k y) \implies (y \not\prec_k x), &\qquad \text{asymmetry of preference;} \\
(x \prec_k y) \wedge (y \prec_k z) \implies (x \prec_k z), &\qquad \text{transitivity of preference;} \\
(x \sim_k y) \wedge (y \sim_k z) \implies (x \sim_k z), &\qquad \text{transitivity of indifference.}
\end{aligned}$$

It is also noted that the k-length preference-indifference relation $\preceq_k$ is defined as the union of

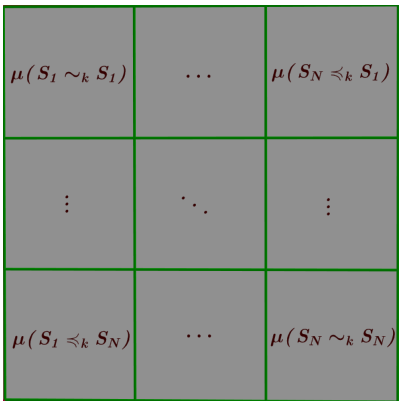

Figure 2.1: Schematic of a k-length PCM.

k-length preference and indifference relations, i.e.,

$$x \preceq_k y \Leftrightarrow (x \sim_k y) \vee (x \prec_k y). \tag{2.3}$$

As a result, transitivity of $\sim_k$ amounts to the fact that $\preceq_k$ is transitive and complete, which in turn implies that $\preceq_k$ is a weak order. A schematic of a k-length PCM constructed using the introduced relational notions is depicted in Figure 2.1. We point out that in the special case $k = 1$, part (iii) of the PCJ Axiom is often referred to as the *consistency property* of PCMs. It is used for quantifying the judgments consistently and ensures interconnectivity among all activities which in turn permits their quantitative assessment. Let us propose the following statistic for measuring the average degree of consistency on $\mathcal{G}_m^{(t)}$:

$$\mathrm{cDeg}(\mathcal{E}_m^{(t)}) = \frac{1}{N^3} \sum_{i,j,l=1}^{N} \mathbb{I}[\epsilon_{il}^{(t)} \epsilon_{lj}^{(t)} = \epsilon_{ij}^{(t)}]. \tag{2.4}$$

Here, one has $\mathrm{cDeg} = 1$ for fully consistent $\mathcal{G}_m^{(t)}$, while for $\mathrm{cDeg}(\mathcal{G}_m^{(t)})$ values reasonably close to 1 we would rather refer to $\mathcal{G}_m^{(t)}$ as a *near-consistent* comparison multigraph. Furthermore, were it to be $\mathrm{cDeg} = 0$, the comparison multigraph $\mathcal{G}_m^{(t)}$ would be *inconsistent*, rendering impossibility of existence of a quantitative scheme which would allow any associative comparison performed among N^3 triplets of assets.

For convenience of the reader and in light of the PCJ Axiom, we present the following series of technical lemmas on properties of PCMs and their special classes, most of which have already been known.

Lemma 2.1.1 (Rank Property)**.** *Under the PCJ* Axiom*, it holds for agent* m *that for every* $t \in \mathcal{T}$*,*

$$\forall \mathbf{A} \in \mathcal{A}_{\text{PCM}}^{(t)} : \text{rank}\,(\mathbf{A}) = 1. \tag{2.5}$$

Proof. First, note that by consistency property (iii) of the PCJ Axiom, the matrix $\mathbf{A}$ can be written as follows:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21}a_{11} & a_{21}a_{12} & \cdots & a_{21}a_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}a_{11} & a_{N1}a_{12} & \cdots & a_{N1}a_{1N} \end{pmatrix}.$$

Subsequently, observe that each row of this matrix is a constant multiplier of first row. Hence, $\text{rank}(\mathbf{A}) = 1$. ■

Now, let us provide some intuition behind Lemma 2.1.1. For a strictly positive matrix (which is the case for any PCM), the rank of the matrix indicates the number of vertices on the convex hull of the data contained in the matrix. Thus, $\text{rank}\,(\mathbf{A}) = 1$ delineates that a single class is sufficient to cluster the contained data, further implying existence of strong dependencies among various entries of the matrix. In other words, the fact that $\text{rank}\,(\mathbf{A}) = 1$, implies that there exist only one piece of information embedded within the matrix $\mathbf{A}$.

In turn, Lemma 2.1.1 yields

Lemma 2.1.2 (Reciprocity Property)**.** *Under the PCJ* Axiom*, for every* $\mathbf{A} \in \mathcal{A}_{\text{PCM}}^{(t)}$ *constructed*

by agent m, we have that for each $t \in \mathcal{T}$,

$$\begin{aligned} \forall i = 1, 2, \ldots, N : a_{ii} &= 1, \\ \forall i, j = 1, 2, \ldots, N : a_{ij} &= \frac{1}{a_{ji}}. \end{aligned} \tag{2.6}$$

Proof. By Lemma 2.1.1, $\mathrm{rank}(\mathbf{A}) = 1$. Thus, the matrix $\mathbf{A}$ can be written as the outer product of two vectors, i.e.,

$$\mathbf{A} = \mathbf{u} \circ \mathbf{v}^{\intercal}, \tag{2.7}$$

where $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{N}_{>0}$. Subsequently, using the consistency property (iii) of the PCJ Axiom, for every $i = 1, 2, \ldots, n$, we get

$$\begin{gathered} a_{ii} = u_i v_i = (u_i v_k)(v_k v_i) = (u_i v_i)(u_k v_k) \implies \\ a_{ii} = a_{ii}(u_k v_k) \implies \\ u_k v_k = 1, \quad \forall k \in \{1, 2, \ldots, N\}. \end{gathered}$$

It is then evident that

$$u_k = \frac{1}{v_k}, \quad \forall k \in \{1, 2, \ldots, N\}. \tag{2.8}$$

In view of equation (2.8), the elements of $\mathbf{A}$ are essentially expressed by element-wise division of vectors $\mathbf{u}$ and $\mathbf{v}$, i.e., $a_{ij} = \frac{u_i}{v_j}$. This then leads to $a_{ij} = \frac{1}{a_{ji}}$ which completes the proof. ■

The assertions of Lemma 2.1.2 suggest us to concentrate on a particular subclass of $\mathcal{A}^{(t)}_{\mathrm{PCM}}$ which satisfy the reciprocity property at every $t \in \mathcal{T}$. It will be shown that equation (2.6) fulfilled for members of such subclass of PCMs make them a preferred choice for modeling the procedures that underlie agents' judgment faculties. In other words, employing the above-mentioned subclass of PCMs enables us to connect mechanisms of the agents' ways of thought to their quantification of the so-called *secondary qualities*. In this respect, it is worth to provide the following definition.

Definition 2.1.2 (RPCM)**.** *A square matrix $\mathbf{A} \in \mathcal{A}^{(t)}_{\mathrm{RPCM}} \subset \mathcal{A}^{(t)}_{\mathrm{PCM}}$ is said to be a reciprocal pairwise comparison matrix (RPCM) if*

(i) $\mathbf{A}$ *has unit diagonal elements, i.e.,* $a_{ii} = 1$ *for all* $i = 1, 2, \ldots, N$,

(ii) $\mathbf{A}$ *has reciprocal off-diagonal elements, i.e.,* $a_{ij} = \frac{1}{a_{ji}}$ *for all* $i, j = 1, 2, \ldots, N$,

where $\mathcal{A}^{(t)}_{\mathrm{RPCM}}$ *represents the family of all reciprocal pairwise comparison matrices at time* t *which is a proper subclass of* $\mathcal{A}^{(t)}_{\mathrm{PCM}}$.

An important, plausibly unique, characterization property pertaining to RPCMs is stated in the following lemma.

Lemma 2.1.3 (Generalized Idempotent Property)**.** *Under the PCJ* Axiom *and for agent* m*, the following generalized idempotent relation*

$$\forall k \geq 1, \forall \mathbf{A} \in \mathcal{A}^{(t)}_{\mathrm{RPCM}} : \mathbf{A}^k = N^{k-1}\mathbf{A} \tag{2.9}$$

holds for every $t \in \mathcal{T}$.

Proof. It is straightforward. ■

We clarify possible implications of Lemma 2.1.3 as follows. For any RPCM associated to agent m's set of beliefs, Lemma 2.1.3 asserts that the diagonal elements of its k-th power are equal to those of the original RPCM (whose diagonal elements are all ones) times N^{k-1}. Thus, we may write

$$\forall k \geq 1, \forall i \in \{1, 2, \ldots, N\}, \mu^{(t)}_m(S_i \sim_k S_i) = N^{k-1}\, \mu^{(t)}_m(S_i \sim_1 S_i). \tag{2.10}$$

To gain more insight on implications of the generalized idempotent property of RPCMs and comprehend the rationale behind the results provided by equation (2.10), it deserves to distinguish between $S_i \sim_k S_i$ and $S_i \sim_1 S_i$ from the information-theoretic point of view. To this end, we first note that $S_i \sim_1 S_i$ corresponds to traversing a self-loop on $\mathcal{G}^{(t)}_m$, i.e., comparing an asset to itself. However, the information gain for such a comparison is clearly zero. Hence, we may refer to the event of performing the comparison $S_i \sim_1 S_i$ as a *redundant action*. On the other hand (and in contrast to traversing self-loops), communicating the cycles of $\mathcal{G}^{(t)}_m$, i.e., $S_i \sim_k S_i$ where $k \geq 2$,

can potentially increase the amount of information gain for the agent. This stems from the fact that the subjective gauge of the agent about dominance relations existing among assets becomes more refined each time she assimilates the information pertaining to the cycles of $\mathcal{G}_m^{(t)}$ whose lengths are strictly greater than one.

Now, since the relation $S_i \sim_k S_i$ corresponds to a cycle of length $k-1$ for every $i \in \{1, 2, \ldots, N\}$, there exist at most $k-1$ intermediary nodes in each cycle. According to the *maximum entropy principle*, we can assume that the agent's assimilation of information on each intermediary node of the cycle follows the discrete uniform probability model with each individual probability equal to $(\frac{1}{N})^{k-1}$. Hence, it is to be expected that agent m would continue exploring the cycles of $\mathcal{G}_m^{(t)}$ rationally until she commits the so-called redundant action, after which she stops the exploration. Define a new random variable Z which denotes the number of trials that takes for agent m to exploit the cycles of $\mathcal{G}_m^{(t)}$ until she commits a redundant action. Evidently, Z is a geometrically distributed random variable with the probability of success (traversing a self-loop) equal to $(\frac{1}{N})^{k-1}$. As a result, the expected number of Bernoulli trials that takes for agent m to commit a redundant action is

$$\mathbb{E}[Z] = \frac{1}{(\frac{1}{N})^{k-1}} = N^{k-1}. \tag{2.11}$$

In turn, this implies that the information gain gets increased N^{k-1}-fold by traversing the cycles of length $k-1$ on $\mathcal{G}_m^{(t)}$, e.g., compare to equation (2.10).

It is worth to mention that the kinds of interpretations like the one provided above which are consistent with common sense, potentially amplify the importance of RPCMs as being tools deemed appropriate for modeling agents' underlying mechanisms of judgment, further providing justifications for the use of comparative judgment framework in a broader sense.

Moreover, the above lemmas yield that eigenvalues of an RPCM are either N or 0 (with algebraic multiplicity $N-1$).

Lemma 2.1.4 (Characteristic Polynomial)**.** *For agent m and every $\mathbf{A} \in \mathcal{A}_{\mathrm{RPCM}}^{(t)}$ where $t \in \mathcal{T}$, the*

characteristic polynomial for $\mathbf{A}$ *takes on the following form:*

$$p_{\mathbf{A}}(\lambda) = \lambda^N - N\lambda^{N-1}. \tag{2.12}$$

Proof. It is well-known that the characteristic polynomial of a square matrix is given by

$$p_{\mathbf{A}}(\lambda) = \lambda^N - [\mathrm{Tr}(\mathbf{A})]\,(\lambda^{N-1}) + \cdots + (-1)^N \det(\mathbf{A}). \tag{2.13}$$

By Lemma 2.1.1, all the terms of $p_{\mathbf{A}}(\lambda)$ except the first two leading terms equal zero. It remains to note that $\mathrm{Tr}(\mathbf{A}) = N$ for any RPCM, which yields

$$p_{\mathbf{A}}(\lambda) = \lambda^N - N(\lambda^{N-1}). \tag{2.14}$$

In turn, this completes the proof. ■

Next, we use equation (2.12) for the characteristic polynomial of $\mathbf{A}$ to determine the eigenvalues of the matrix.

Lemma 2.1.5 (Perron-Frobenius Eigenvalue)**.** *For agent* m*, it holds that*

$$\forall \mathbf{A} \in \mathcal{A}^{(t)}_{\mathrm{RPCM}}, \lambda_{\max} = N \tag{2.15}$$

for every $t \in \mathcal{T}$*, where* $\lambda_{\max}$ *represents the largest eigenvalue of* $\mathbf{A}$*, also called the Perron-Frobenius eigenvalue.*

Proof. Recall that for a matrix $\mathbf{A}$ with strictly positive elements, the Perron-Frobenius theorem states that $\mathbf{A}$ has a positive real eigenvalue $\lambda_{\max}$, which is strictly greater than the absolute values (moduli) of all the other eigenvalues. On the other hand, by Lemma 2.1.4 the roots of the characteristic polynomial of $\mathbf{A}$ are either 0 or N. In addition, Lemma 2.1.1 implies that $\mathrm{rank}(\mathbf{A}) = 1$. Hence, $\mathbf{A}$ has a single nonzero eigenvalue $\lambda_{\max} = N$ with all the remaining eigenvalues equal to zeros. ■

An another interesting property of RPCMs is that they belong to a family of matrices for which there is a relation between their rank, order and trace (see equation (2.16)).

Lemma 2.1.6 (Trace-Order-Rank Property)**.** *Under the PCJ* Axiom *and for agent m, the specific characteristics of a generic matrix $\mathbf{A} \in \mathcal{A}_{\mathrm{PCM}}^{(t)}$ satisfies the following relationship:*

$$\mathrm{Tr}(\mathbf{A}) = N \cdot \mathrm{rank}(\mathbf{A}). \tag{2.16}$$

Proof. By Lemma 2.1.3, we have

$$\mathbf{A}^k = (N^{k-1})\mathbf{A}, \quad \forall k \geq 1. \tag{2.17}$$

For $k = 2$, equation (2.17) yields that

$$\mathbf{A}^2 = N\mathbf{A}. \tag{2.18}$$

The proof then follows from the fact that equation (2.18) implies the validity of equation (2.16). See (Harville, 2018) for more details on this implication. ■

Further on, since for any RPCM we have $\mathrm{Tr}(\mathbf{A}) = \lambda_{\max}$, in such case Lemma 2.1.6 stipulates that

$$\lambda_{\max} = N \cdot \mathrm{rank}(\mathbf{A}). \tag{2.19}$$

Equation (2.19) can be interpreted as follows: It is known that if we take any vector corresponding to the largest eigenvalue $\lambda_{\max}$, then the action of the matrix on that specific vector attains its maximum value. Next, recall that $\mathrm{rank}(\mathbf{A})$ quantifies the amount of information that is preserved by such action. Hence, equation (2.19) means that the maximum action of the matrix equals its complexity (provided that the order N of the matrix signifies its complexity) multiplied by the amount of information preserved by such action. In turn, this provides an additional useful property for RPCMs and substantiates their use to model the thought processes of particular agents. Namely, equation (2.19) guarantees that the possibility for the loss of information is eliminated inasmuch as action of RPCMs on vectors chosen along the eigenvector corresponding to $\lambda_{\max}$ are

taken into account. This is due to the fact since $\mathrm{rank}(\mathbf{A}) = 1$ for RPCMs, then reducing the dimensionality of the data set would require one to project each data point along some unit vector. Moreover, it is of great interest to choose the unit vector in a way that projection of the data points along it would retain as much of the variation of the data points as possible. To achieve this goal, it is well-known that the eigenvector corresponding to $\lambda_{\max}$ can be chosen for the purpose of dimensionality reduction since such an eigenvector is the direction along which the data set would have the maximum amount of variance.

Next, another important aspect of RPCMs that deserves attention pertains to sensitivity of their eigenvalues to perturbations of their elements. A possible approach for conducting such sensitivity analysis can be carried out by studying the condition number of the matrix since the change in the output of a function for a small changes in its input arguments can be quantified by the condition number of the function. For a simple eigenvalue λ_j, it was shown in (Wilkinson, 1988) that $\mathrm{cond}(\lambda_j; \mathbf{A})$, expressed in terms of the left and right eigenvectors of $\mathbf{A}$, measures the sensitivity of λ_j to small perturbations in elements of the matrix. Yet, another useful representation for $\mathrm{cond}(\lambda_j; \mathbf{A})$ is given in (Smith, 1967) whose author puts forward the following relationship

$$\mathrm{cond}(\lambda_j; \mathbf{A}) = \frac{\|\mathrm{adj}(\lambda_j I - \mathbf{A})\|_2}{p'_{\mathbf{A}}(\lambda_j)}, \tag{2.20}$$

where $\|\mathbf{A}\|_2$ denotes the *largest singular value* of $\mathbf{A}$, and $p'_{\mathbf{A}}(\lambda_j)$ is the derivative of the characteristic polynomial of $\mathbf{A}$ w.r.t. λ which is evaluated at $\lambda = \lambda_j$. For the matrix $\mathbf{A}$, being an RPCM with a simple eigenvalue $\lambda_{\max} = N$, then equation (2.20) simplifies as follows:

$$\mathrm{cond}(\lambda = N; \mathbf{A}) = \frac{\|\mathrm{adj}(NI - \mathbf{A})\|_2}{N^{N-1}}. \tag{2.21}$$

Therefore, it becomes clear that for a matrix $\mathbf{A} \in \mathcal{A}^{(t)}_{\mathrm{RPCM}}$ with a sufficiently large order N, whose majority of the elements are from the same order of magnitude and none of its elements is near-zero or near-infinity, the eigenvalue of the matrix would be relatively insensitive to changes occurred in all of its elements, unless the structure of the matrix undergoes a considerable perturbation. We refer to this property of the largest eigenvalue of RPCMs as *robustness*.

Lemma 2.1.7 (Robustness of Perron-Frobenius Eigenvalue)**.** *The largest eigenvalue* $\lambda_{\max}$ *of a generic matrix* $\mathbf{A} \in \mathcal{A}_{\mathrm{RPCM}}^{(t)}$ *is robust w.r.t. perturbations occurred in the matrix, provided that the elements of the matrix are homogeneous, i.e., they are uniformly bounded from above and below.*

Proof. See (Saaty, 1993). ■

An application of the sensitivity analysis for RPCMs is illustrated by the following example:

Example 2.1.1. *Consider the following matrix* $\mathbf{A} \in \mathcal{A}_{\mathrm{RPCM}}^{(t)}$

$$\mathbf{A} = \begin{pmatrix} 1 & \alpha & \beta \\ \frac{1}{\alpha} & 1 & \gamma \\ \frac{1}{\beta} & \frac{1}{\gamma} & 1 \end{pmatrix},$$

where $\alpha, \beta, \gamma > 0$*. Then, it can be shown that the adjugate of* $3I - \mathbf{A}$ *is as follows:*

$$\mathrm{adj}(3I - \mathbf{A}) = \begin{pmatrix} 3 & 2\alpha + \frac{\beta}{\gamma} & 2\beta + \alpha\gamma \\ \frac{2}{\alpha} + \frac{\gamma}{\beta} & 3 & \frac{\beta}{\alpha} + 2\gamma \\ \frac{2}{\beta} + \frac{1}{\alpha\gamma} & \frac{\alpha}{\beta} + \frac{2}{\gamma} & 3 \end{pmatrix}.$$

This enables us to compute $\mathrm{cond}(\lambda = 3; \mathbf{A})$*. We forego to provide an explicit relation for the condition number due to its excessive length. Now, assume that the original matrix* $\mathbf{A}$ *is constructed using the triplet* $(\alpha, \beta, \gamma) = (1.001, 0.995, 1.005)$ *and the perturbed matrices are created using the triplets* $(1.002, 0.980, 1.015)$*,* $(0.950, 0.775, 1.015)$*,* $(1.875, 0.205, 0.580)$ *and* $(5.225, 3.170, 0.001)$*. Then it can be shown that the condition number for the original matrix is* 1.000*, and for the perturbed matrices the condition numbers are equal to* 1.000*,* 1.015*,* 2.030 *and* 417.709*, respectively. It is then observed that perturbations occurred in the neighborhood of the original matrix do not influence the condition number substantially, whereas an structural change in the matrix, as in the latter case, leads to a huge change in the value of the condition number.*

Besides, the derivation of $\|\cdot\|_2$ *is often too involved and complex. Hence, one can potentially obtain*

an overestimate of $\mathrm{cond}(\lambda = 3; \mathbf{A})$ *by resorting to the well-known inequality* $\|\cdot\|_2 \leq \|\cdot\|_F$. *Yet, another conservative upper bound on* $\mathrm{cond}(\lambda = 3; \mathbf{A})$ *is obtained by observing that*

$$\mathrm{adj}(3I - \mathbf{A}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 2 \begin{pmatrix} 1 & \alpha & \beta \\ \frac{1}{\alpha} & 1 & \gamma \\ \frac{1}{\beta} & \frac{1}{\gamma} & 1 \end{pmatrix} + \begin{pmatrix} 1 & \frac{\beta}{\gamma} & \alpha\gamma \\ \frac{\gamma}{\beta} & 1 & \frac{\beta}{\alpha} \\ \frac{1}{\alpha\gamma} & \frac{\alpha}{\beta} & 1 \end{pmatrix}.$$

Then, $\mathrm{cond}(\lambda = 3; \mathbf{A})$ *would be bounded from above by summing the Frobenius norms of each decomposed matrix, i.e., an upper bound on* $\mathrm{cond}(\lambda = 3; \mathbf{A})$ *is*

$$\left(\frac{1}{N^{N-1}}\right) \max\left\{\max(N, \alpha, \beta, \gamma, \alpha\gamma, \frac{\beta}{\gamma}), \frac{1}{\min(\alpha, \beta, \gamma, \alpha\gamma, \frac{\beta}{\gamma})}\right\}.$$

In turn, this implies that insofar as the elements of the matrix $\mathbf{A}$ *are uniformly bounded, perturbations in the neighborhood of* $\mathbf{A}$ *do not affect its maximum eigenvalue substantially.*

The concluding properties of RPCMs we discuss here pertains to the eigenvalues of their perturbed matrices. Namely, the following lemma holds.

Lemma 2.1.8. *The Perron-Frobenius eigenvalue* $\lambda_{\max} = N$ *for the matrix* $\mathbf{A}$ *if and only if the matrix* $\mathbf{A} \in \mathcal{A}^{(t)}_{\mathrm{RPCM}}$. *Otherwise,* $\lambda_{\max} > N$ *for any matrix being a perturbation of* $\mathbf{A}$.

Proof. See (Saaty, 1993). ■

Furthermore, the following lemma delineates the relation between Perron-Frobenius eigenvalues of RPCMs and their consistency properties (property (iii) of the PCJ Axiom).

Lemma 2.1.9. *The matrix* $\mathbf{A} \in \mathcal{A}^{(t)}_{\mathrm{RPCM}}$ *for some* $t \in \mathcal{T}$, *is consistent if and only if the Perron-Frobenius eigenvalue of the matrix is* $\lambda_{\max} = N$.

Proof. See (Saaty, 1993). ■

As per Lemmas 2.1.8 and 2.1.9, any perturbation of the matrix $\mathbf{A} \in \mathcal{A}_{\mathrm{RPCM}}^{(t)}$ can be seen as a departure from the fulfillment of the condition (iii) of the PCJ Axiom. This is due to the fact that $\lambda_{\max} = N$ for any $\mathbf{A} \in \mathcal{A}_{\mathrm{RPCM}}^{(t)}$ at every $t \in \mathcal{T}$. Therefore, at any given point of time t, one may write the definition of consistency as follows:

$$a_{il}^{(t)} a_{lj}^{(t)} = a_{ij}^{(t)}, \quad i, j, l = 1, 2, \ldots, N. \tag{2.22}$$

This implies that the average degree of consistency on $\mathcal{G}_m^{(t)}$ for agent m yields $\mathrm{cDeg}(\mathcal{E}_m^{(t)}) = 1$ using equation (2.4) for all $t \in \mathcal{T}$. In turn, this permits the agent m to quantify her judgments to an extent that she can perform associative comparisons among all N^3 triplets of assets.

2.1.2 Pairwise Comparison Tensors

In this section, we employ tensors as a natural structure to embody the information pertaining to the mutual interactions among price changes in the assets. Let us start by defining the so-called *pairwise comparison tensors* and their *reciprocal* counterparts.

Definition 2.1.3 (PCT)**.** *A third-order spatio-temporal tensor $\boldsymbol{\mathcal{A}} \in \mathcal{A}_{\mathrm{PCT}}^{|\mathcal{T}|} \subset \mathbb{R}_{>0}^{N \times N \times |\mathcal{T}|}$ is said to be a pairwise comparison tensor (PCT) if it is formed by concatenating $|\mathcal{T}|$ number of PCMs as its frontal slices, where $\mathcal{A}_{\mathrm{PCT}}^{|\mathcal{T}|}$ denotes the family of PCTs having temporal dimensions $|\mathcal{T}|$.*

Definition 2.1.4 (RPCT)**.** *A generic tensor $\boldsymbol{\mathcal{A}} \in \mathcal{A}_{\mathrm{RPCT}}^{|\mathcal{T}|}$ is said to be a reciprocal pairwise comparison tensor (RPCT) if each of its frontal slices is an RPCM, where $\mathcal{A}_{\mathrm{RPCT}}^{|\mathcal{T}|}$ denotes the family of RPCTs having temporal dimensions $|\mathcal{T}|$, which is a proper subclass of $\mathcal{A}_{\mathrm{PCT}}^{|\mathcal{T}|}$.*

In what follows, we extend the concept of matrix inconsistency introduced in (Saaty, 1977; Saaty, 1993) to the domain of tensors and define a function to measure the inconsistencies of PCTs. Keeping this in mind, recall Lemma 2.1.8 which states that the largest eigenvalue $\lambda_{\max}$ of any matrix obtained by perturbing its associated $\mathbf{A} \in \mathcal{A}_{\mathrm{RPCM}}^{(t)}$ would have its largest eigenvalue being greater than N. More so, it was discussed that the consistency property (iii) of the PCJ

Axiom is satisfied only if a given matrix belongs to $\mathcal{A}_{\mathrm{RPCM}}^{(t)}$ (e.g., see Lemma 2.1.9). Hence, one may envisage the statistic $\lambda_{\max} - N$ as being a measure of departure of the judgments cast by agent m from the consistency condition of the PCJ Axiom. An average perturbation statistic can then be defined as $\frac{\lambda_{\max}-N}{N-1}$.

Definition 2.1.5 (Inconsistency Function)**.** *For a tensor $\boldsymbol{\mathcal{A}} \in \mathcal{A}_{\mathrm{PCT}}^{|\mathcal{T}|}$, the inconsistency function $\psi : \mathcal{T} \to \mathbb{R}_{\geq 0}^{|\mathcal{T}|}$ is defined as follows:*

$$\psi(t) := \frac{\lambda_{\max}^{(t)} - N}{N-1}. \tag{2.23}$$

Here, $\lambda_{\max}^{(t)}$ represents the largest eigenvalue of the PCM corresponding to the t-th frontal slice of $\boldsymbol{\mathcal{A}}$.

In turn, Definition 2.1.5 suggests one to examine temporal inconsistencies which agent m may confront in the protean of making her judgments. It can be stated that the judgments cast by agent m are temporally consistent if her judgments at every point of time $t \in \mathcal{T}$ satisfy the condition (iii) of the PCJ Axiom. That is, the inconsistency function for a pairwise comparison tensor $\boldsymbol{\mathcal{A}} \in \mathcal{A}_{\mathrm{PCT}}^{|\mathcal{T}|}$ is zero if every frontal slice of $\boldsymbol{\mathcal{A}}$ is an RPCM. In other words, in an ideal case for which the reciprocal pairwise comparison tensors are used for agent m's absorption of the market information, one can ensure that the mechanisms underlying the agent's subjective opinion remain persistent over time, i.e., $\mu_m^{(t)} = \mu_m$ for all $t \in \mathcal{T}$ (one may find this situation similar to the scenario involving stationary frequencies of an ergodic Markov chain).

Such regularity of the inconsistency measures obtained for a PCT can also be defined by using the concept of *approximate entropy* (ApEn). In order to define the ApEn for a sequence of inconsistency measures $\{\psi(1), \psi(2), \dots\}$ of length $|\mathcal{T}|$, firstly two blocks of longitude $l \leq |\mathcal{T}|$ are defined by $\breve{\psi}(i) = \{\psi(i), \dots, \psi(i+l-1)\}$ and $\breve{\psi}(j) = \{\psi(j), \dots, \psi(j+l-1)\}$, and then the distance between them is calculated using $d(\breve{\psi}(i), \breve{\psi}(j)) = \max_{k=1,\dots,l}(|\psi(i+k-1) - \psi(j+k-1)|)$.

Thereafter, the following statistic

$$C_i^l(r) = \frac{1}{|\mathcal{T}| - l + 1} \sum_{j=1}^{|\mathcal{T}|-l+1} \mathbb{I}[d(\breve{\psi}(i), \breve{\psi}(j)) \leq r] \tag{2.24}$$

is defined, where the summation term counts the number of consecutive blocks of longitude l which are similar to the given block $\breve{\psi}(i)$ within a given resolution r. Subsequently, the logarithmic frequency with which longitude-l blocks that are close together stay together for the next increment defines the ApEn for the given sequence of inconsistency measures, i.e.,

$$\mathring{\psi} = \frac{1}{|\mathcal{T}| - l + 1} \sum_{i=1}^{|\mathcal{T}|-l+1} C_i^l(r). \tag{2.25}$$

ApEn defined by equation (2.25) enjoys several mathematical and statistical properties among which the following are noteworthy: 1- ApEn is a statistic which is robust, meaning that it is insensitive to artifacts or outliers; 2- ApEn is not altered by translations or scaling applied uniformly to all elements contained within the considered sequence of inconsistency measures; 3- nonlinearity causes greater values of ApEn, e.g, see (Pincus and Goldberger, 1994); 4- ApEn is model-free, which makes it a promising statistic for the analysis of data series whose generating sources are unknown; and 5- Computation of ApEn requires equally-spaced measurements over time, e.g, see (Pincus, 1991).

The above-mentioned properties of ApEn, among others, make it a suitable candidate to measure the regularity of the inconsistency measures. This has been formally defined as follows:

Definition 2.1.6 (Regularity Measure)**.** *For a tensor $\boldsymbol{\mathcal{A}} \in \mathcal{A}_{\mathrm{PCT}}^{|\mathcal{T}|}$, the regularity measure of inconsistencies $\mathring{\psi} : \mathbb{R}_{\geq 0}^{|\mathcal{T}|} \to \mathbb{R}_{\geq 0}$ is defined by the approximate entropy of the inconsistency function $\psi(t)$ given by equation* (2.23)*.*

By appealing to the applications of $\mathring{\psi}$, one can then quantify the regularity of the inconsistencies, e.g., tackling the question of whether a pairwise comparison tensor $\boldsymbol{\mathcal{A}} \in \mathcal{A}_{\mathrm{PCT}}^{|\mathcal{T}|}$ is regularly inconsistent (consistent), or it is irregularly inconsistent. Basically, the regularity measure $\mathring{\psi}$ of inconsistencies estimates the amount of randomness found in $\psi(t)$ without requiring any prior

knowledge of the source that generates the time-dependent inconsistency values. In other words, lower values of $\mathring{\psi}$ attest that $\psi(t)$ is persistent, repetitive and predictive in the sense that patterns repeat themselves throughout the series. On the other hand, higher values of $\mathring{\psi}$ imply independence between particular temporal values of $\psi(t)$ and hence are indicators towards lower number of repeated patterns and smaller values of randomness, e.g., see (Delgado-Bonal and Marshak, 2019; Pincus, 2008) for more details on approximate entropy.

Remark 2.1.1. *A pairwise comparison tensor $\boldsymbol{\mathcal{A}} \in \mathcal{A}_{\mathrm{PCT}}^{|\mathcal{T}|}$ is said to be regularly consistent if*

$$(\mathring{\psi} = 0) \wedge (\forall t : \psi(t) = 0), \tag{2.26}$$

and regularly inconsistent if

$$(\mathring{\psi} = 0) \wedge (\exists t : \psi(t) > 0). \tag{2.27}$$

Otherwise, $\boldsymbol{\mathcal{A}}$ is said to be irregularly inconsistent if

$$(\mathring{\psi} > 0) \wedge (\exists t : \psi(t) > 0), \tag{2.28}$$

where the magnitude of $\mathring{\psi}$ delineates the amount of irregularity.

The RPCTs constructed for agent m according to Definition 2.1.3, reflect the ideal situations in which the agent satisfies the conditions of the PCJ Axiom throughout time. In reality, however, the values of the comparison tensor could deviate substantially from the ideal case. In order to illustrate this phenomenon, consider tensor $\boldsymbol{\mathcal{A}} \in \mathcal{A}_{\mathrm{RPCT}}^{|\mathcal{T}|}$ whose perturbed tensor is denoted by $\tilde{\boldsymbol{\mathcal{A}}} \in \mathcal{A}_{\mathrm{PCT}}$ and referred to as the *consensus pairwise comparison tensor*. The following theorem shows that $\tilde{\boldsymbol{\mathcal{A}}}$ has to be a rank-1 estimate of $\boldsymbol{\mathcal{A}}$ to guarantee that the inconsistency function (2.23) can be recovered.

Theorem 2.1.1 (Consensus PCT)**.** *Let $\boldsymbol{\mathcal{A}} \in \mathcal{A}_{\mathrm{RPCT}}^{|\mathcal{T}|}$ be a tensor whose associated consensus tensor is denoted by $\tilde{\boldsymbol{\mathcal{A}}} \in \mathcal{A}_{\mathrm{PCT}}$. Also, let $\tilde{\boldsymbol{\mathcal{A}}}$ be a rank-1 tensor, i.e., $\tilde{\boldsymbol{\mathcal{A}}} = \tilde{\mathbf{z}} \circ (\tilde{\mathbf{x}} \circ \tilde{\mathbf{y}}^{\intercal})$, which solves the*

following constrained polyadic decomposition model:

$$\min_{\tilde{\mathbf{x}},\tilde{\mathbf{y}},\tilde{\mathbf{z}}} \ \|\boldsymbol{\mathcal{A}} - \tilde{\mathbf{z}} \circ (\tilde{\mathbf{x}} \circ \tilde{\mathbf{y}}^{\intercal})\|_F \tag{2.29a}$$

$$\text{s.t. } \tilde{z}_t(\tilde{\mathbf{x}} \circ \tilde{\mathbf{y}}^{\intercal}) > \mathbf{0}_N, \quad t = 1, 2, \ldots, |\mathcal{T}|, \tag{2.29b}$$

$$\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in \mathbb{R}^N_{>0}, \tag{2.29c}$$

$$\tilde{\mathbf{z}} \in \mathbb{R}^{|\mathcal{T}|}_{\geq 0}. \tag{2.29d}$$

Here, $\mathbf{0}_N$ denotes an $N \times N$ matrix whose entries are all zeros, and inequality (2.29b) *is evaluated component-wise. Then, the elements of the vector $\tilde{\mathbf{z}}$ depend on the largest eigenvalues of the frontal slices of $\tilde{\boldsymbol{\mathcal{A}}}$ (up to some positive scaling).*

Proof. Consider the matrix $\mathbf{A} \in \mathcal{A}^{(t)}_{\text{RPCM}}$ at a given time $t \in \mathcal{T}$, and let $\mathbf{u}$ and $\mathbf{v}$ represent the left and right eigenvectors of its associated Perron-Frobenius eigenvalue $\lambda_{\max}$, respectively, i.e.,

$$\mathbf{u}^{\intercal}\mathbf{A} = \lambda_{\max}\mathbf{u}^{\intercal} \quad \text{and} \quad \mathbf{A}\mathbf{v} = \lambda_{\max}\mathbf{v},$$

where strict positivity of $\mathbf{A}$ renders $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N_{>0}$. Then, by Perron-Frobenius theorem we have

$$\lim_{k\to\infty} \left(\frac{\mathbf{A}}{\lambda_{\max}}\right)^k = \left(\frac{1}{\mathbf{v}^{\intercal}\mathbf{u}}\right)\mathbf{u}\mathbf{v}^{\intercal}. \tag{2.30}$$

However, Lemma 2.1.3 implies that

$$\mathbf{A}^k = N^{k-1}\mathbf{A}. \tag{2.31}$$

Hence,

$$\mathbf{A} \sim \left(\frac{\lambda^k_{\max}}{N^{k-1}}\right)\left(\frac{1}{\mathbf{v}^{\intercal}\mathbf{u}}\right)\mathbf{u}\mathbf{v}^{\intercal} = \lambda_{\max}\left(\frac{1}{\mathbf{v}^{\intercal}\mathbf{u}}\right)\mathbf{u}\mathbf{v}^{\intercal}, \tag{2.32}$$

which implies that $\mathbf{A}$ is a rank-1 matrix multiplied by a constant which depends on its largest eigenvalue $\lambda_{\max}$.

On the other hand, the rank-1 polyadic decomposition of the tensor $\boldsymbol{\mathcal{A}} \in \mathcal{A}^{|\mathcal{T}|}_{\text{RPCT}}$ can be written as the outer product of a rank-1 matrix by a vector which resides on the temporal dimension of the

tensor, i.e.,

$$\boldsymbol{\mathcal{A}} \approx \tilde{\boldsymbol{\mathcal{A}}} = \tilde{\mathbf{z}} \circ (\tilde{\mathbf{x}} \circ \tilde{\mathbf{y}}^{\intercal}), \tag{2.33}$$

where $\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in \mathbb{R}^{N}_{>0}$ and $\tilde{\mathbf{z}} \in \mathbb{R}^{|\mathcal{T}|}_{\geq 0}$. Therefore, elements of the vector $\tilde{\mathbf{z}}$ can be regarded as the terms which depend on the largest eigenvalues of the corresponding matrices which pertain to frontal slices of the consensus tensor $\tilde{\boldsymbol{\mathcal{A}}}$.

Consequently, in order to recover the inconsistency function $\psi(t)$ at each time instant, one can first obtain $\tilde{\boldsymbol{\mathcal{A}}}$ by solving model (2.29) and then proceed by computing the largest eigenvalue of the approximated tensor at each of its frontal slices, and thereafter evaluate the inconsistency function given by equation (2.23). ■

The following remark elaborates on the relation between the inconsistency function $\psi(t)$ and polyadic decomposition of RPCTs.

Remark 2.1.2. *In order to recover the inconsistency measures, one may first approximate the tensor* $\boldsymbol{\mathcal{A}} \in \mathcal{A}^{|\mathcal{T}|}_{\mathrm{RPCT}}$ *by solving model* (2.29) *and then proceed by computing the largest eigenvalue of the approximated tensor* $\tilde{\boldsymbol{\mathcal{A}}}$ *at each of its frontal slices. Subsequently, the inconsistency measure at each time instant is derived by evaluating the function* $\psi(t)$ *given by* (2.23).

It is further noted that the error of approximation in rank-1 polyadic decomposition of an RPCT is directly linked to its average inconsistency measure, as elucidated in the following example.

Example 2.1.2. *Let* $\mathbf{r}^{(t)} \in R^{N}_{>0}$ *denote the vector of lag-*l *daily rates of return for* N *assets, i.e., the lag-*l *rates of return for assets are computed using the following relation:*

$$r^{(t)} = \frac{C^{(t)}}{C^{(t-l)}}, \tag{2.34}$$

where $C^{(t)}$ *denotes the adjusted closing price for the assets at time* $t \in \mathcal{T}$.

In our simulations, we considered $N = 811$ *assets, selected to be those that have ever appeared in the list of* $S\&P500$ *at a time period from January* 1990 *to December* 2019 *for which their price*

information did not contain any missing values[1]*. Also, in our datasets, the lag-*1 *rates of return were comprised of* $|\mathcal{T}| = 7558$ *and* $|\mathcal{T}| = 360$ *daily and monthly time steps, respectively.*

The RPCTs were constructed for each given lag value (shown schematically in Figure 2.2), which was then followed by computation of their associated inconsistency function $\psi(t)$*, average inconsistency measure* $\bar{\psi}(t)$ *as well as their regularity measure* $\mathring{\psi}(t)$[2]*. The values for the lag parameter* l *we considered in our simulations ranged from* 1 *through* 10*.*

Figures 2.3 to 2.5 depict, respectively, the rank-1 *poliadic decomposition errors, average inconsistency measures and regularity measures for different values of the lag parameter* l*. It is clear from these figures that the error of polyadic decomposition grows nonlinearly by increasing values of the lag parameter, whereas the growth of the average inconsistency measures exhibits a linear trend. These observations were also verified empirically, and led to the formulation of the following relations:*

$$\epsilon(l) \approx \epsilon(1)\sqrt{l}, \tag{2.35}$$

and

$$\bar{\psi}(l) \approx \bar{\psi}(1)l, \tag{2.36}$$

where $\epsilon(l)$ *and* $\bar{\psi}(l)$ *denote the relative error of approximation and average inconsistency measure at lag* l*, respectively. Furthermore, it is evident that* $\mathring{\psi}(t)$ *decreases as the values of the lag parameter increase.*

In the above discussion, we analyzed the mechanisms of judgment casting for a specific agent in the market. However, our ultimate goal is to expand our analysis to the entire participants of the market. For this purpose, we are to make a reasonably realistic assumption based on which the closing prices of assets are the only public source of information available or being of interest to the agents. Then, our analytical framework developed for an individual agent can be extended to encompass the totality of the agents by noting that, for an individual agent, what distinguishes

[1]Data were retrieved from the Center for Research in Security Prices (CRSP) provided by Wharton Research at the University of Pennsylvania (WRDS).

[2]All models were solved on Amazon web services (AWS) using a server with Xen hypervisor, 128 cores and 4TB of RAM memory.

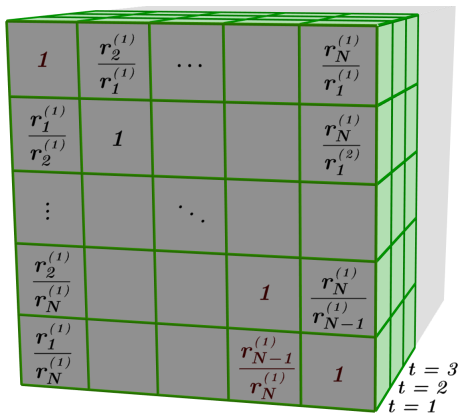

Figure 2.2: Schematic of an RPCT.

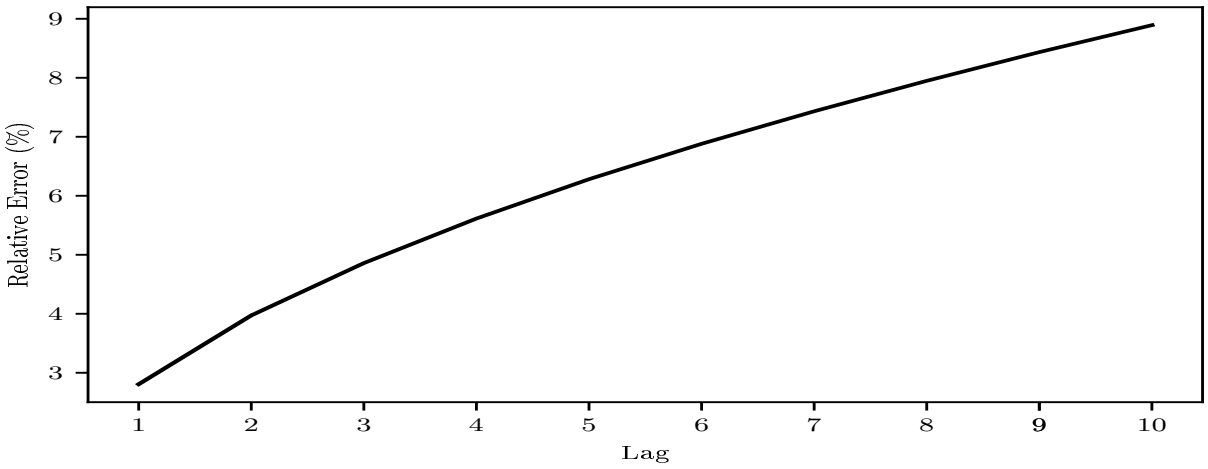

Figure 2.3: Plot of the relative errors for rank-1 polyadic decomposition at different lag values.

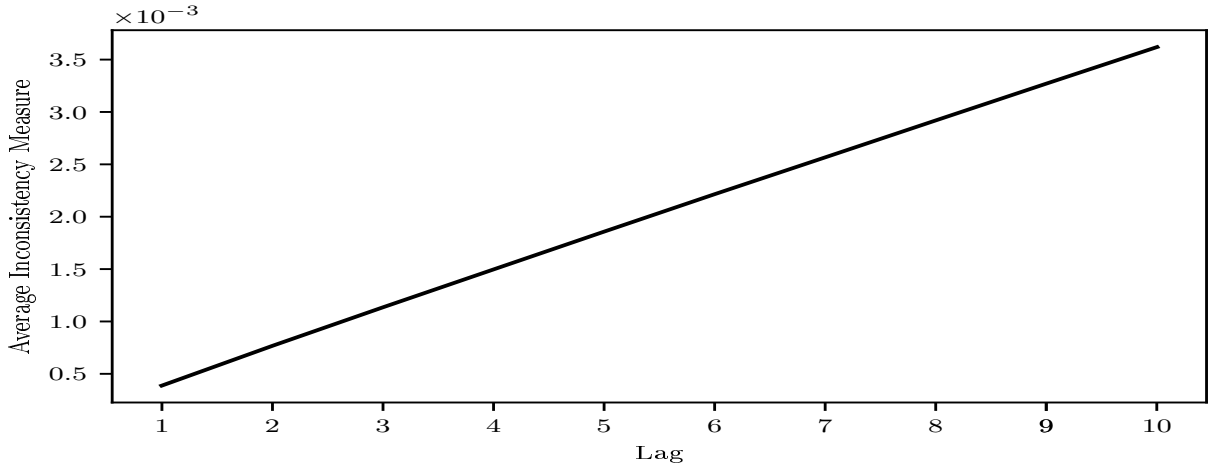

Figure 2.4: Plot of the average inconsistency measures at different lag values.

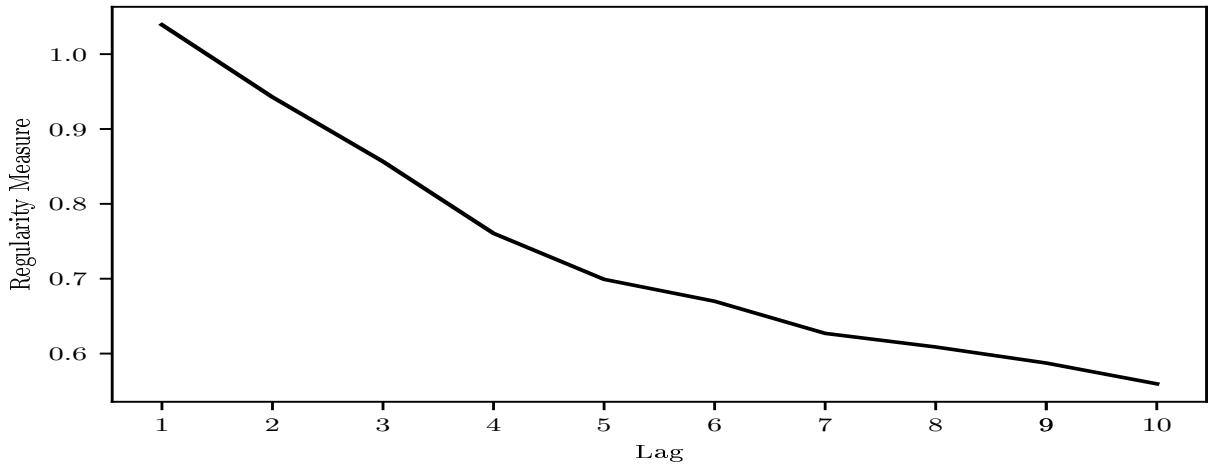

Figure 2.5: Plot of the regularity measures at different lag values.

different RPCTs from each other pertains to the lag criterion used in the computation of the rates of return. Thus, the problem of multiple subjective criteria existing across the market (as there are $M \gg 1$ agents who construct their RPCTs using their particular lag values), can potentially be reduced to that of finding an appropriate value for l which represents the lag value used by majority of the agents. Let us call the criterion constructed above, which is based on such a lag value the *constitution* criterion (denoted by $\mathcal{C}$); we adapt the term constitution from (Gibbard, 2014) where the author addresses Arrow dilemma (Arrow, 2012) and refers to amalgamation of individual preferences as a constitution.

We reckon that the best constitution criterion is the one corresponding to the lag value using which the constructed consensus PCT yields smallest value of the average inconsistency $\bar{\psi}(t)$ and the largest value of the regularity of inconsistencies $\mathring{\psi}(t)$. This is because larger values of $\mathring{\psi}(t)$ indicate higher independence among inconsistencies of PCMs at different points of the time horizon. Thus, opting for the largest value of $\mathring{\psi}(t)$ ensures that $\mathcal{G}_{\mathcal{C}}^{(t)}$ under consideration will more likely to remain independent at different time instants, making it possible to draw more accurate conclusions on relations existing among the assets at intersequent time instants.

2.2 An Index of Financial Chaos

In this section, we employ the concepts and tools presented in Section 2.1 to introduce a stock market index which captures chaotic behavior in asset prices. In the sequel, we then characterize the so-called *financial chaos index* (FCI) by resorting to *unit root*, *stationarity* and *random walk* tests for the daily realizations of the index as well as its monthly and quarterly ones.

2.2.1 Formulation of the Index

Let us recall that the consensus PCT derived by using the constitution criterion could potentially embody the stock market information about mutual asset price changes in an effective way. In turn, this enables us to formally define the financial chaos index as follows:

Definition 2.2.1 (FCI). *Consider a consensus PCT given by* $\tilde{\boldsymbol{\mathcal{A}}} \in \mathcal{A}_{\mathrm{PCT}}^{|\mathcal{T}|}$ *whose associated RPCT (denoted by* $\boldsymbol{\mathcal{A}} \in \mathcal{A}_{\mathrm{RPCT}}^{|\mathcal{T}|}$*) is constructed using lag-*l *rates of return for some value of* l *as its constitution criterion. The financial chaos index* $\mathrm{FCI} : \mathcal{T} \to \mathbb{R}_{\geq 0}^{|\mathcal{T}|}$ *is then defined via the inconsistency function given by equation* (2.23) *as follows:*

$$\mathrm{FCI}(t) := \psi(t) = \frac{\lambda_{\max}^{(t)} - N}{N - 1}, \tag{2.37}$$

where $\lambda_{\max}^{(t)}$ *represents the largest eigenvalue of the matrix corresponding to the* t*-th frontal slice of* $\tilde{\boldsymbol{\mathcal{A}}}$*. Throughout the chapter, we further denote by* FCI_t *(or by* ψ_t*) the time series which takes on the observed FCI values.*

For instance, consider the data stated in Example 2.1.2 and recall that the lag-1 rate of return was deemed suitable for the purpose of constitution criterion. Using the lag-1 rates of return for computing FCI_t, we sketch the plots of daily and annotated monthly FCI_t in Figures 2.6 and 2.7 for a time period from January 1990 to December 2019, respectively. Note that we obtain the monthly FCI_t by taking average of the daily FCI_t values during each month. As it is seen from these figures, the FCI_t spikes whenever the stock market undergoes chaotic fluctuations in response to various anomalous political, geopolitical, economical, financial, fiscal, health and psychological events. For instance, such spikes were observed during the first and second Gulf wars, the explosion of the dot-com bubble, September 11, the 2002 stock market crash, SARS coronovirus pandemic in 2003, failure of Lehman Brothers, the debt-ceiling dispute in 2011, Chinese stock market crash in 2016, OPEC oil cut, the 2018 world-wide stock market down fall, and the recent US-China trade tensions, to mention but a few. Also, for practical reasons addressed in the sequel, we utilize the notion of the quarterly FCI_t, which is defined as the average of daily FCI values during each quarter.

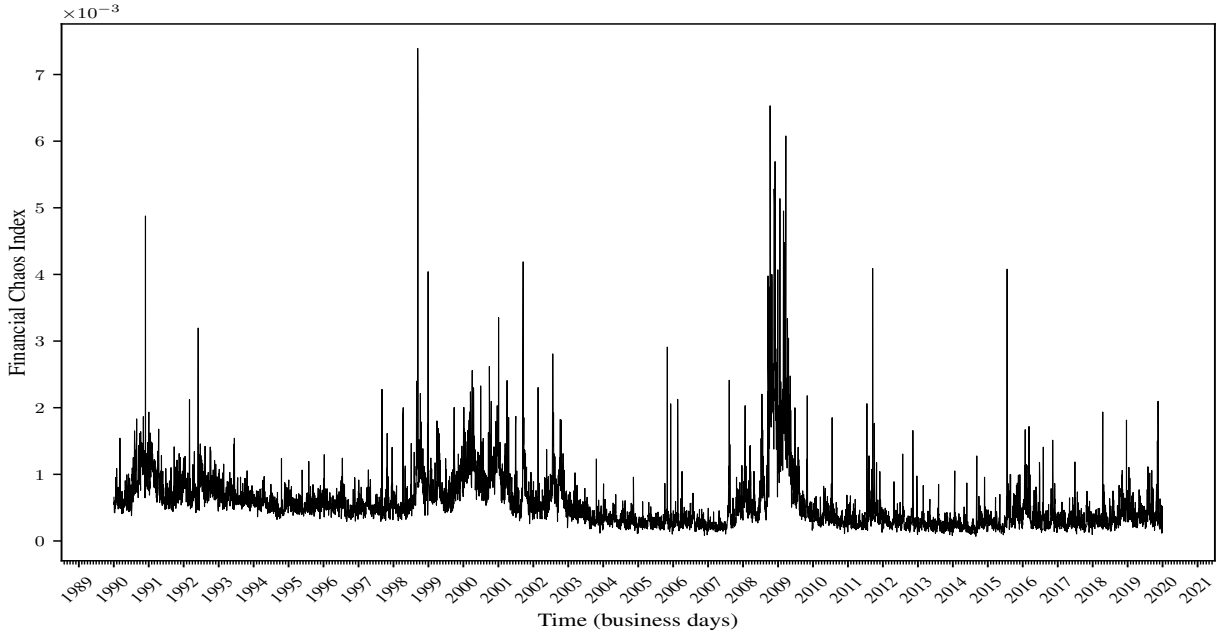

Figure 2.6: Plot of the daily FCI_t for a time period from January 1990 to December 2019.

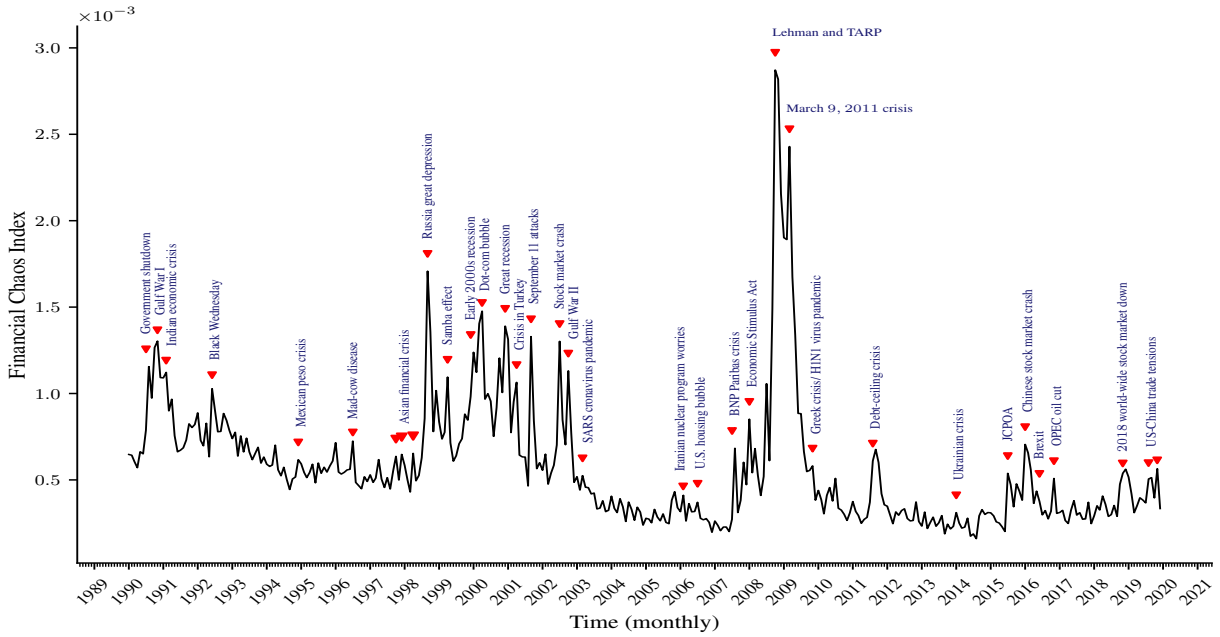

Figure 2.7: Plot of the annotated monthly FCI_t portrayed during January 1990-December 2019.

2.2.2 Non-stationarity Analysis

Given a realization of the process FCI_t, the main goal of this subsection is to characterize daily as well as monthly and quarterly FCI_t by performing necessary unit root tests. To detect whether the underlying stochastic process possesses a root of the unit circle in $\mathbb{C}$, we employ the *Augmented Dickey-Fuller* (ADF) (Said and Dickey, 1984) test which is based on estimating the following regression model:

$$\psi_t = \lambda + \delta t + \phi\psi_{t-1} + \sum_{j=1}^{p} \beta_j \Delta\psi_{t-j} + \epsilon_t, \tag{2.38}$$

where λ and δ are recognized as the *shift* and *trend* coefficients, respectively. Furthermore, the autoregressive–moving-average (ARMA) structure of the errors in model (2.38) are approximated by p-lagged difference terms represented by $\Delta\psi_{t-j}$, where the length of the lag parameter p, is set in a way to ensure that the innovation process $\{\epsilon_t\}$ (having mean zero and homoskedastic) is serially uncorrelated. It makes sense then to introduce the following set of null hypotheses:

$$\begin{cases} \mathrm{H}_{0,1}^{\mathtt{ADF}} : \phi = 1, \lambda = 0, \delta = 0, \\ \mathrm{H}_{0,2}^{\mathtt{ADF}} : \phi = 1, \lambda \neq 0, \delta = 0, \\ \mathrm{H}_{0,3}^{\mathtt{ADF}} : \phi = 1, \lambda \neq 0, \delta \neq 0, \end{cases}$$

where $\mathrm{H}_{0,1}^{\mathtt{ADF}}$, $\mathrm{H}_{0,2}^{\mathtt{ADF}}$ and $\mathrm{H}_{0,3}^{\mathtt{ADF}}$ correspond to the hypotheses on ψ_t to be a simple $I(1)$, an $I(1)$ with shift and finally an $I(1)$ with shift and a time trend, respectively, where $I(1)$ denotes a unit-root process as opposed to $I(0)$ which is a process lacking unit root. It is important to introduce the corresponding alternative hypotheses as follows:

$$\begin{cases} \mathrm{H}_{1,1}^{\mathtt{ADF}} : |\phi| < 1 \Rightarrow \psi_t \sim I(0), \\ \mathrm{H}_{1,2}^{\mathtt{ADF}} : |\phi| < 1 \Rightarrow \psi_t \sim I(0) \text{ with shift}, \\ \mathrm{H}_{1,3}^{\mathtt{ADF}} : |\phi| < 1 \Rightarrow \psi_t \sim I(0) \text{ with shift and linear trend}. \end{cases}$$

The ADF test would then reject any individual null hypothesis in favor of its corresponding

stationary alternative if there is strong evidence against ψ_t to be an $I(1)$. In this respect, for each individual null hypothesis $\mathrm{H}_{0,1}^{\mathtt{ADF}}$, $\mathrm{H}_{0,2}^{\mathtt{ADF}}$ or $\mathrm{H}_{0,3}^{\mathtt{ADF}}$, we first employ the *ordinary least squares* (OLS) method to estimate the model parameters, and subsequently the following *standard t-statistic*

$$\mathrm{ADF}_t = \frac{\hat{\phi} - 1}{\mathrm{SE}(\phi)}, \tag{2.39}$$

and *normalized bias statistic*

$$\mathrm{ADF}_n = \frac{|\mathcal{T}|(\hat{\phi} - 1)}{1 - \hat{\beta}_1 - \cdots - \hat{\beta}_p}, \tag{2.40}$$

are computed in order to evaluate the significance of the obtained values of the shift and trend coefficients. If a computed test statistic is less than the corresponding critical value, then the underlying null hypothesis gets rejected in favor of its alternative.

Tables 2.1 to 2.3 summarize the outcomes obtained for testing the unit-root hypotheses for daily, monthly and quarterly FCI_t, respectively, where the critical values of the tests were computed according to (MacKinnon, 1994; MacKinnon, 2010). Using either of test statistics, the results imply that $\mathrm{H}_{0,2}^{\mathtt{ADF}}$ and $\mathrm{H}_{0,3}^{\mathtt{ADF}}$ get rejected for both daily and monthly time series at 1% levels of significance. Hence, there is a sufficient evidence to reject the hypotheses on daily and monthly FCI_t being any type of $I(1)$ processes with shift or linear trends. Note that the $\mathrm{H}_{0,2}^{\mathtt{ADF}}$ and $\mathrm{H}_{0,3}^{\mathtt{ADF}}$ for quarterly FCI_t using t-statistics get rejected in favor of their alternatives at only 5% level of significance. On the contrary, there is lack of sufficient evidence to reject $\mathrm{H}_{0,1}^{\mathtt{ADF}}$ for all three time series at 1% level of significance, suggesting that these time series are non-stationary processes at the mentioned significance level.

2.2.3 Tests for Stationarity

It is well documented in the literature that the unit root tests may suffer from a number of shortcomings such as lack of ability to discern unit-root processes from weakly stationary ones. In addition, it is enlightening to note that failure to reject $\mathrm{H}_{0,1}^{\mathtt{ADF}}$ for any of time series does not necessarily mean that the time series considered possesses a unit root. To mitigate some drawbacks of the ADF test,

Table 2.1: Results of ADF unit-root test for daily FCI_t.

		t-statistic			normalized bias statistic		
		$\mathrm{H}_{0,1}^{\mathtt{ADF}}$	$\mathrm{H}_{0,2}^{\mathtt{ADF}}$	$\mathrm{H}_{0,3}^{\mathtt{ADF}}$	$\mathrm{H}_{0,1}^{\mathtt{ADF}}$	$\mathrm{H}_{0,2}^{\mathtt{ADF}}$	$\mathrm{H}_{0,3}^{\mathtt{ADF}}$
p-value		0.015	0.001	0.001	0.018	0.001	0.001
Test statistics		-2.431	-6.132	-6.768	-11.709	-82.057	-100.787
Critical values	1% level	-2.567	-3.433	-3.962	-13.685	-20.581	-29.271
	5% level	-1.942	-2.862	-3.413	-8.038	-14.073	-21.663
	10% level	-1.617	-2.567	-3.128	-5.713	-11.236	-18.212

Table 2.2: Results of ADF unit-root test for monthly FCI_t.

		t-statistic			normalized bias statistic		
		$\mathrm{H}_{0,1}^{\mathtt{ADF}}$	$\mathrm{H}_{0,2}^{\mathtt{ADF}}$	$\mathrm{H}_{0,3}^{\mathtt{ADF}}$	$\mathrm{H}_{0,1}^{\mathtt{ADF}}$	$\mathrm{H}_{0,2}^{\mathtt{ADF}}$	$\mathrm{H}_{0,3}^{\mathtt{ADF}}$
p-value		0.042	0.004	0.001	0.049	0.001	0.001
Test statistics		-2.017	-3.789	-4.657	-8.004	-30.460	-49.964
Critical values	1% level	-2.574	-3.452	-3.987	-13.568	-20.302	-28.709
	5% level	-1.942	-2.870	-3.424	-7.983	-13.927	-21.338
	10% level	-1.617	-2.572	-3.135	-5.685	-11.141	-17.976

Table 2.3: Results of ADF unit-root test for quarterly FCI_t.

		t-statistic			normalized bias statistic		
		$\mathrm{H}_{0,1}^{\mathtt{ADF}}$	$\mathrm{H}_{0,2}^{\mathtt{ADF}}$	$\mathrm{H}_{0,3}^{\mathtt{ADF}}$	$\mathrm{H}_{0,1}^{\mathtt{ADF}}$	$\mathrm{H}_{0,2}^{\mathtt{ADF}}$	$\mathrm{H}_{0,3}^{\mathtt{ADF}}$
p-value		0.064	0.010	0.018	0.091	0.002	0.002
Test statistics		-1.828	-3.492	-3.851	-5.953	-27.147	-35.556
Critical values	1% level	-2.586	-3.491	-4.043	-13.272	-19.619	-27.412
	5% level	-1.943	-2.886	-3.450	-7.882	-13.586	-20.613
	10% level	-1.614	-2.579	-3.151	-5.629	-10.914	-17.453

it is beneficial to conduct additional set of statistical tests with the goal of analyzing stationarity of the time series. To test stationarity of ψ_t, we employ *Kwiatkowski, Phillips, Schmidt and Shin* (KPSS) test (Hobijn, Franses, and Ooms, 2004; Kwiatkowski, Phillips, Schmidt, and Shin, 1992), which is based on the following structural model:

$$\psi_t = \lambda + \delta t + v_t, \tag{2.41}$$

$$v_t = r_t + \epsilon_t, \tag{2.42}$$

$$r_t = r_{t-1} + u_t, \quad (\text{with } r_0 = 0), \tag{2.43}$$

where $\{u_t, t \geq 0\}$ are i.i.d. random variables with zero mean and finite variance σ^2, while ϵ_t and r_t are stationary and random walk processes, respectively. The null hypotheses of stationarity could then be stated as follows:

$$\begin{cases} \mathrm{H}_{0,1}^{\mathtt{KPSS}} : \sigma^2 = 0, \lambda \neq 0, \delta = 0, \\ \mathrm{H}_{0,2}^{\mathtt{KPSS}} : \sigma^2 = 0, \lambda \neq 0, \delta \neq 0, \end{cases}$$

where $\mathrm{H}_{0,1}^{\mathtt{KPSS}}$ and $\mathrm{H}_{0,2}^{\mathtt{KPSS}}$ correspond to the hypotheses that ψ_t are *level stationary* as well as *level and trend stationary* processes, respectively. The KPSS test would reject each individual $\mathrm{H}_{0,1}^{\mathtt{KPSS}}$ and $\mathrm{H}_{0,2}^{\mathtt{KPSS}}$ in favor of their corresponding alternatives given by

$$\begin{cases} \mathrm{H}_{1,1}^{\mathtt{KPSS}} : \sigma^2 \neq 0, \lambda \neq 0, \delta = 0 \Rightarrow z_t \sim I(1) \text{ with shift}, \\ \mathrm{H}_{1,2}^{\mathtt{KPSS}} : \sigma^2 \neq 0, \lambda \neq 0, \delta \neq 0 \Rightarrow z_t \sim I(1) \text{ with shift and linear trend}, \end{cases}$$

if the computed KPSS statistic is evaluated to be larger than its corresponding critical value at a given level of significance. The KPSS test would then be carried out by noting that under such null hypotheses, λ and δ can be estimated using OLS. In turn, this enables one to derive the t-th OLS residual e_t as follows:

$$e_t = \psi_t - bt - c, \tag{2.44}$$

for some slope b and intercept c. This leads to the formulation of the KPSS statistic as follows:

$$\text{KPSS} = \frac{\sum_{t=1}^{|\mathcal{T}|}\left(\sum_{i=1}^{t} e_i\right)^2}{|\mathcal{T}|^2\hat{\sigma}^2}, \tag{2.45}$$

where

$$\hat{\sigma}^2 = \frac{1}{|\mathcal{T}|}\left[\sum_{t=1}^{|\mathcal{T}|} e_t^2 + 2\sum_{j=1}^{L}\left(1-\frac{j}{1+L}\right)\left(\sum_{s=j+1}^{|\mathcal{T}|} e_s e_{s-j}\right)\right]. \tag{2.46}$$

Observe that the coefficients $\left(1-\frac{j}{1+L}\right)$ for $j = 1, 2, \ldots, L$, correspond to *Newey-West weights*. It is worth mentioning that, under the normality assumption on $\{\epsilon_t\}$, the test statistic given by equation (2.45) coincides with the *LM statistic*. Hence, the bandwidths involved in data-driven lag selection methods can be computed effectively when determining the *Newey-West estimators* (Newey and West, 1994) and the *Bartlett kernel* (Andrews, 1991).

Table 2.4 summarizes the results obtained for testing the stationarity null hypotheses $\mathrm{H}_{0,1}^{\texttt{KPSS}}$ and $\mathrm{H}_{0,2}^{\texttt{KPSS}}$ for daily, monthly and quarterly FCI_t with the critical values of the tests computed based on (Kwiatkowski, Phillips, Schmidt, and Shin, 1992). According to Table 2.4, $\mathrm{H}_{0,1}^{\texttt{KPSS}}$ is rejected at all three significance levels for daily FCI_t, whereas $\mathrm{H}_{0,2}^{\texttt{KPSS}}$ gets rejected only at 5% level of significance for this time series. Thus, the hypothesis that daily ψ_t is a stationary process should be rejected as there exist a sufficient evidence to support the belief that this time series has a certain time-dependent (non-stationary) structure. Consequently, we may conclude that the time series ψ_t corresponding to the daily realizations of FCI_t during January 1990-December 2019 is $I(1)$. Similarly, the same rationale holds true for the monthly and quarterly FCI_t, save for the fact that these time series behave in a more stationary manner as compared to their daily counterparts. That is, for the monthly FCI_t, both $\mathrm{H}_{0,1}^{\texttt{KPSS}}$ and $\mathrm{H}_{0,2}^{\texttt{KPSS}}$ get rejected at 5% level of significance, while for the quarterly data only the former null hypothesis gets rejected at the two specified significance levels. It is noted that $\mathrm{H}_{0,2}^{\texttt{KPSS}}$ gets rejected at 10% level of significance for the quarterly FCI_t.

Table 2.4: Results of KPSS test for daily, monthly and quarterly FCI_t.

		Lags	$\mathrm{H}_{0,1}^{\mathtt{KPSS}}$	$\mathrm{H}_{0,2}^{\mathtt{KPSS}}$
p-value	Daily	51	0.000	0.025
	Monthly	11	0.013	0.469
	Quarterly	5	0.030	0.638
Test statistics	Daily		2.406	0.177
	Monthly		0.694	0.058
	Quarterly		0.547	0.046
Critical values	1% level		0.743	0.217
	5% level		0.461	0.148
	10% level		0.348	0.119

2.2.4 Random Walk Tests

In order to test whether FCI_t is a random walk, we employ the *variance ratio test* developed in (Lo and MacKinlay, 1988) based on which the null hypothesis model is given by the following recursive relation:

$$\psi_t = \lambda + \psi_{t-1} + \epsilon_t, \tag{2.47}$$

where λ represents a *shift* constant, and $\{\epsilon_t\}$ are zero-mean innovations. By making a remark that the variance of the increments of a random walk is a linear function of the length of the observation period, we put forward the following null hypotheses:

$$\mathrm{H}_0^{\mathtt{HOMO}} : \{\epsilon_t\} \text{ are i.i.d. with } N(0, \sigma_0^2),$$

and

$$\mathrm{H}_0^{\mathtt{HTRO}} : \mathbb{E}[\epsilon_t] = 0 \text{ and } \mathbb{E}[\epsilon_t \epsilon_{t-l}] = 0 \text{ for all } t \text{ and any } l \neq 0.$$

Note that the former null hypothesis is associated to the case of homoscedastic increments and the latter corresponds to heteroscadasticity-consistent increments. In order to examine either of $\mathrm{H}_0^{\mathtt{HOMO}}$

or $\mathrm{H}_0^{\texttt{HTRO}}$, the unknown parameters λ and σ_0^2 are first estimated as follows:

$$\hat{\lambda} = \frac{1}{|\mathcal{T}|q} \sum_{t=1}^{|\mathcal{T}|q} (\psi_t - \psi_{t-1}), \tag{2.48}$$

and

$$\hat{\sigma}_0^2 = \frac{1}{|\mathcal{T}|q - 1} \sum_{t=1}^{|\mathcal{T}|q} (\psi_t - \psi_{t-1} - \hat{\lambda})^2, \tag{2.49}$$

where $q < \frac{|\mathcal{T}|}{2}$ denotes a period parameter used to create the overlapping return horizons. Then, the corresponding test statistic is formulated in terms of the ratio of the variances as follows:

$$\overline{M}_r(q) = \frac{\bar{\sigma}^2(q)}{\hat{\sigma}_0^2} - 1, \tag{2.50}$$

which has an underlying normal distribution with

$$\bar{\sigma}^2(q) = \frac{1}{q(|\mathcal{T}|q - q + 1)(1 - \frac{q}{|\mathcal{T}|q})} \sum_{t=1}^{|\mathcal{T}|q} (\psi_t - \psi_{t-1} - q\hat{\lambda})^2. \tag{2.51}$$

Table 2.5 presents the outcomes of the variance ratio test performed on daily, monthly and quarterly FCI_t. In order to assess whether daily FCI_t satisfies the requirements for being a random walk, recall that it was shown earlier that the daily FCI_t is an $I(1)$ process in a given time period January 1990-December 2019 at 1% level of significance (see Table 2.2). Besides, every random walk is known to be an $I(1)$ process. This suggests that daily FCI_t can potentially be a random walk. However, note that non-stationarity of a time series is not a sufficient condition for it to be a random walk (Tsay, 2005). This fact has also been reflected in p-values of the variance ratio test obtained for the daily FCI_t, where the two null hypotheses are rejected at all three levels of significance and based on any value of the period parameter q. Hence, daily FCI_t is rejected to be a random walk. On the other hand, there is insufficient evidence to reject the random walk null hypotheses on monthly and quarterly FCI_t. Hence, we can conclude that these two time series are random walks at 1% level of significance.

Table 2.5: Results of variance ratio test for daily, monthly and quarterly FCI_t.

		homoscedasticity				heteroscadasticity-consistent			
Period (q)		2	4	8	16	2	4	8	16
p-value	Daily	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Monthly	0.044	0.003	0.011	0.012	0.391	0.189	0.209	0.172
	Quarterly	0.078	0.514	0.191	0.168	0.476	0.789	0.537	0.415
Test statistics	Daily	-23.019	-30.339	-24.011	-17.731	-7.353	-7.297	-6.924	-6.237
	Monthly	-2.011	-2.960	-2.558	-2.533	-0.856	–1.312	-1.253	-1.365
	Quarterly	1.762	-0.652	-1.308	-1.378	0.711	-0.267	-0.617	-0.815
Critical values	1% level	-2.576	-2.576	-2.576	-2.576	-2.576	-2.576	-2.576	-2.576
	5% level	-1.960	-1.960	-1.960	-1.960	-1.960	-1.960	-1.960	-1.960
	10% level	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645

2.3 Relation to Market Regimes

In this section, we take advantage of the properties pertaining to the financial chaos index and conduct the tasks of regime analysis (based on daily realizations) and market segmentation (based on monthly values). It will be shown that the U.S. stock market during the time period from January 1990 to December 2019 had been driven by multiple underlying regimes (states) rather than a single one, and further, the market could be divided into at least 6 time segments within which the corresponding FCI_t values are time-homogeneous.

2.3.1 Regime Analysis

We begin this subsection by visually inspecting the daily and monthly realizations of FCI_t using Figures 2.6 and 2.7, which suggests that there exist apparent multiple underlying regimes for the market rather than a single one. In what follows, we utilize the daily FCI_t realizations to investigate such an observation in a more rigorous manner using the approach introduced and employed in (Everitt, 1996; Powell, Roa, Shi, and Xayavong, 2007) for the statistical inference on underlying regimes of time series. More specifically, we point out that the authors of (Campbell, 2000; Camp-

bell and Cochrane, 1999) have previously established that cyclical changes of economic states give rise to the corresponding changes in the expected returns on stocks. According to these authors, under absence of regime changes the stock returns would have constant mean and volatility (the case of a random walk with constant drift). Herein, we draw a similarity between the notion of the time variation in stock returns mentioned above and our new financial chaos index. More precisely, we show that under market regime switching, the shift of average level of daily FCI_t can be modeled as random switching from one mean level to another one. It is worth mentioning that the stock market switches among various regimes generally occur in response to changes that happen in the underlying forces driving the market.

Further, Figure 2.6 suggests that huge fluctuations in the values of FCI_t are rare, and the set containing such large values has potentially a much smaller cardinality as compared to the sets which correspond to low and intermediate magnitudes. Thus, the *power-law distribution* can be used to explain the rarity phenomenon. On the other hand, Figure 2.8 depicts that *lognormal distribution* can potentially be a suitable candidate to model the low and intermediate values of FCI_t. Hence, we opt for the use of *modified lognormal power-law* (MLP) distribution to analyze the daily FCI_t realizations which was introduced and developed in (Basu, Gil, and Auddy, 2015; Basu and Jones, 2004). More specifically, the MLP distribution behaves like a lognormal distribution at low and intermediate values of FCI_t, having a characteristic peak and turnover features unique to lognormal distributions, whereas it transitions to a power-law distribution (also widely known as Pareto-type distribution) at high values of FCI_t.

Let us define Ψ_t to be a random variable associated to FCI_t (per daily time frequency) whose realization is denoted by $\boldsymbol{\psi} = [\psi_1, \psi_2, \ldots, \psi_{\mathcal{T}}]^{\intercal}$. The MLP distribution then takes advantage of the fact that later time evolution of Ψ_t could skew the lognormal density of initial values of the random variable. More precisely, the subsequent time of growth of the random variable representing quantities of Ψ_t could itself be considered as a random variable. For instance, if we draw a quantity $\boldsymbol{\psi}_{t_0}$ randomly from a lognormal distribution of Ψ_t at time $t = t_0$ (with location and shape parameters μ and σ, respectively), the subsequent growth of Ψ_t following the time $t = t_0$ can then

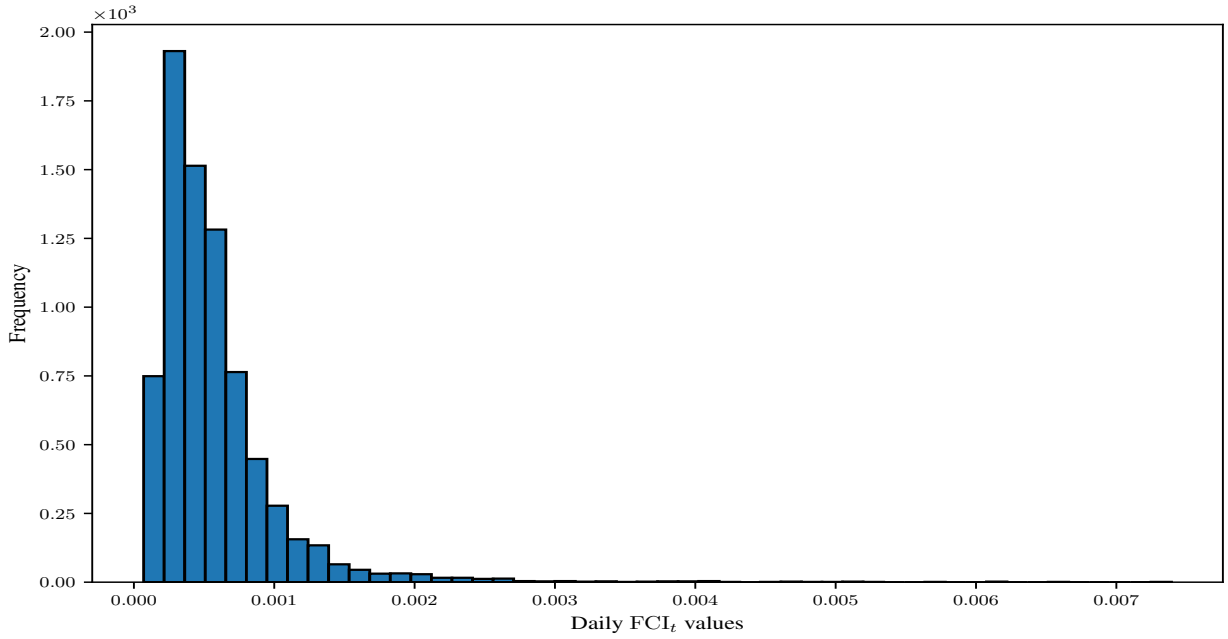

Figure 2.8: Frequency histogram for daily FCI_t during January 1990-December 2019.

be characterized by an exponential distribution, i.e.,

$$\frac{\partial \Psi_t}{\partial t} = \kappa \Psi_t, \tag{2.52}$$

with a fixed value for the *growth rate* $\kappa > 0$. Accordingly, a lognormal distribution is maintained having the mean of its logarithmic values being shifted to $\mu + \kappa(t_0 + dt)$ after, say a fixed time dt, is elapsed, whereas the shape parameter σ remains unaltered. Thereafter, time can be treated as an exponentially distributed random variable with *stopping rate* parameter $\zeta > 0$ and the probability density function

$$f(t) = \zeta \exp\{-\zeta t\}. \tag{2.53}$$

The probability density function for Ψ_t can then be derived as being

$$g(\psi_t; \mu, \sigma, \omega) = \left(\frac{\omega}{2}\right)\left(\frac{1}{\psi_t}\right)^{1+\omega} \exp\left\{\omega\mu + \frac{\omega^2\sigma^2}{2}\right\} \operatorname{erfc}\left\{\frac{1}{\sqrt{2}}\left(\omega\sigma - \frac{\log\psi_t - \mu}{\sigma}\right)\right\}, \tag{2.54}$$

where $\psi_t > 0$, $\mu \in (-\infty, \infty)$, $\sigma > 0$, $\omega = \frac{\zeta}{\kappa}$ is a strictly positive parameter known as *power-law*

tail, and erfc denotes the *complementary Gauss error function*. Also, the corresponding cumulative distribution function of the density function (2.54) takes on the form:

$$G(\psi_t;\mu,\sigma,\omega)=\tfrac{1}{2}\mathrm{erfc}\left\{-\tfrac{\log\psi_t-\mu}{\sqrt{2}\sigma}\right\}-\tfrac{1}{2\psi_t^{\omega}}\exp\left\{\omega\mu+\frac{\omega^2\sigma^2}{2}\right\}\mathrm{erfc}\left\{\tfrac{1}{\sqrt{2}}\left(\omega\sigma-\tfrac{\log\psi_t-\mu}{\sigma}\right)\right\}. \tag{2.55}$$

It is noteworthy to mention that $G(\psi_t;\mu,\sigma,\omega)$ gets dominated by its first term in the case where $\psi_t\to 0$. Specifically, the limit coincides with the cumulative distribution function of a lognormal distribution with parameters μ and σ. On the contrary, the MLP behaves as a pure power-law distribution as $\sigma\to 0$.

The mean and variance of the distribution are also derived as follows:

$$\mathbb{E}[\Psi_t]=\omega\exp\left\{2\mu+\sigma^2\right\}\left(\frac{\exp\left\{\sigma^2\right\}}{\omega-2}-\frac{\omega}{(\omega-1)^2}\right),\quad \omega>1, \tag{2.56}$$

and

$$\mathrm{Var}[\Psi_t]=\frac{\omega}{\omega-1}\exp\left\{\mu+\frac{\sigma^2}{2}\right\},\quad \omega>2, \tag{2.57}$$

respectively.

Let us assume that there are R market regimes for some given value $R\geq 1$. These regimes are generally hidden and the value of R is unknown. To make inference on the underlying regimes, we assume the following functional for the global frequency distribution of Ψ_t:

$$h(\psi_t;\boldsymbol{\theta})=\sum_{r=1}^{R}\pi_r g(\psi_t;\boldsymbol{\theta}_r)=\sum_{r=1}^{R}\pi_r\left(\tfrac{\omega_r}{2}\right)\left(\tfrac{1}{\psi_t}\right)^{1+\omega_r}\exp\left\{\omega_r\mu_r+\frac{\omega_r^2\sigma_r^2}{2}\right\}\mathrm{erfc}\left\{\tfrac{1}{\sqrt{2}}\left(\omega_r\sigma_r-\tfrac{\log\psi_t-\mu_r}{\sigma_r}\right)\right\}, \tag{2.58}$$

where $\boldsymbol{\theta}=[\boldsymbol{\theta}_1,\boldsymbol{\theta}_2,\cdots,\boldsymbol{\theta}_R]^{\intercal}$ denotes the vector of unknown parameters whose components are comprised of $\boldsymbol{\theta}_r=[\mu_r,\sigma_r,\omega_r,\pi_r]$, each of which corresponding to a single regime r for $r=$

$1, 2, \cdots, R$, i.e.,

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1^{\intercal} \\ \boldsymbol{\theta}_2^{\intercal} \\ \vdots \\ \boldsymbol{\theta}_R^{\intercal} \end{pmatrix} = \begin{pmatrix} \mu_1 & \sigma_1 & \omega_1 & \pi_1 \\ \mu_2 & \sigma_2 & \omega_2 & \pi_2 \\ \vdots & \vdots & \vdots & \vdots \\ \mu_R & \sigma_R & \omega_R & \pi_R \end{pmatrix}. \tag{2.59}$$

Here, π_r denotes the proportion of observations that is associated to each regime.

The probability density function given by equation (2.58) corresponds to a finite mixture and can be used to determine the total number R of the regimes that underlie the global frequency distribution and further estimate the parameter vectors associated to each regime. Note that every value of the regime parameter r corresponds to a single hypothesis in the *alternative component specification* of the model, see (Hamilton, 1994) for more details.

Subsequently, assume that $|\mathcal{T}|$ realizations of the financial chaos index are classified into L discrete categories such that:

$$B_l = [b_l - \frac{b_l - b_{l-1}}{2}, b_l + \frac{b_{l+1} - b_l}{2}], \tag{2.60}$$

for $l = 2, \ldots, L-1$. Furthermore, we have the following categories

$$B_1 = (0, b_1 + \frac{b_2 - b_1}{2}], \tag{2.61}$$

and

$$B_L = [b_L - \frac{b_L - b_{L-1}}{2}, \infty), \tag{2.62}$$

in which $b_1 = \min\{\mathrm{FCI}_t\}$ and $b_L = \max\{\mathrm{FCI}_t\}$. Further, we determine the total number of discrete bins using the ***Rice rule***, which yields $L = 2|\mathcal{T}|^{\frac{1}{3}} = 41$. Hence, the error should have a *multinomial distribution*. In order to verify this, let $\hat{p}_l$ and $p_l(\boldsymbol{\theta})$ denote the observed and expected

proportions of Ψ_t in a given category l, respectively. For instance, for $l = 1, 2, \ldots, L$, we have

$$\begin{aligned} p_l(\boldsymbol{\theta}) = \mathbb{P}[\Psi_t \in B_l|\boldsymbol{\theta}] = \int_{\inf(B_l)}^{\sup(B_l)} h(\psi_t;\boldsymbol{\theta})d\psi_t &= \int_{\inf(B_l)}^{\sup(B_l)} \sum_{r=1}^{R} \pi_r g(\psi_t;\boldsymbol{\theta}_r)d\psi_t \\ &= \sum_{r=1}^{R} \pi_r \int_{\inf(B_l)}^{\sup(B_l)} g(\psi_t;\boldsymbol{\theta}_r)d\psi_t \\ &= \sum_{r=1}^{R} \pi_r \left[G(\sup(B_l);\boldsymbol{\theta}_r) - G(\inf(B_l);\boldsymbol{\theta}_r)\right]. \end{aligned} \tag{2.63}$$

Then, the corresponding likelihood function takes on the following form:

$$\mathcal{L}(\boldsymbol{\theta}|\hat{\mathbf{p}}) = \prod_{l=1}^{L} \left(p_l\left(\boldsymbol{\theta}\right)\right)^{|\mathcal{T}|\hat{p}_l}, \tag{2.64}$$

where $\hat{\mathbf{p}} = [\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_L]^\intercal$.

By equation (2.64), the log-likelihood function is derived as follows:

$$\log \mathcal{L}(\boldsymbol{\theta}|\hat{p}) = |\mathcal{T}| \sum_{l=1}^{L} \hat{p}_l \log p_l(\boldsymbol{\theta}). \tag{2.65}$$

Subtracting the constant term $|\mathcal{T}| \sum_{l=1}^{L} \hat{p}_l \log \hat{p}_l$ from this function and multiplying it by 2, would then convert the function given by equation (2.65) into

$$\log \mathcal{L}(\boldsymbol{\theta}|\hat{\mathbf{p}}) = \begin{cases} 2|\mathcal{T}| \sum\limits_{l=1}^{L} \hat{p}_l \log \left[\dfrac{p_l(\boldsymbol{\theta})}{\hat{p}_l}\right], & \text{if} \quad \hat{p}_l > 0, \\ 0, & \text{if} \quad \hat{p}_l = 0, \end{cases} \tag{2.66}$$

which can conveniently be expressed in terms of KL-divergence of two distributions such that

$$\log \mathcal{L}(\boldsymbol{\theta}|\hat{\mathbf{p}}) = -2|\mathcal{T}| \sum_{l=1}^{L} D_{\mathrm{KL}}(\hat{p}_l||p_l(\boldsymbol{\theta})). \tag{2.67}$$

Note that RHS of equation (2.67) has an approximate χ^2 distribution with $L - 4R - 1$ degrees of freedom when all involved parameter estimates are assumed to be random variables.

Finally, in order to estimate the vectors of parameters $\boldsymbol{\theta}$ for a fixed value of R, the following constrained optimization model is to be solved:

$$\max_{\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\omega},\boldsymbol{\pi}} \ \log \mathcal{L}(\boldsymbol{\theta}|\hat{\mathbf{p}}) \tag{2.68a}$$

$$\text{s.t.} \ \sum_{r=1}^{R} \pi_r = 1, \tag{2.68b}$$

$$\mu_r \in (-\infty, \infty), \qquad r = 1, 2, \cdots, R, \tag{2.68c}$$

$$\sigma_r \in (0, \infty), \qquad r = 1, 2, \cdots, R, \tag{2.68d}$$

$$\omega_r \in (0, \infty), \qquad r = 1, 2, \cdots, R, \tag{2.68e}$$

$$\pi_r \in (0, 1], \qquad r = 1, 2, \cdots, R, \tag{2.68f}$$

where $\boldsymbol{\mu} = [\mu_1, \mu_2, \cdots, \mu_R]^\intercal$, $\boldsymbol{\sigma} = [\sigma_1, \sigma_2, \cdots, \sigma_R]^\intercal$, $\boldsymbol{\omega} = [\omega_1, \omega_2, \cdots, \omega_R]^\intercal$ and $\boldsymbol{\pi} = [\pi_1, \pi_2, \cdots, \pi_R]^\intercal$. Then, the null hypothesis that a model is based on R regimes fits the data sufficiently well gets accepted if its corresponding p-value of χ^2 test is greater than a specified level of significance.

Recall that in practice, the actual value of R is unknown, and it is usually identified through conducting some experiments. For this purpose, we perform goodness-of-fit tests for each regime hypothesis by taking advantage of the asymptotic distribution of the function given by equation (2.67). Table 2.6 reports the results for regime analysis of daily FCI_t during January 1990-December 2019 where the constrained optimization models (2.68) were solved for $R = 1, 2, 3, 4$ and 5 using a *trust-region Newton-conjugate-gradient* algorithm.

By referring to Table 2.6, it is seen that both 2- and 3-regime models are accepted to represent the underlying probability distribution of the daily FCI_t values at all three levels of significance $\alpha = 0.01, 0.05, 0.10$, while other models including the random walk model of single regime do not pass the goodness-of-fit tests. Therefore, we opt to choose the 3-regime model as the regime-delegating paradigm of financial chaos index realizations due to its lower χ^2-value. As a consequence, being clustered in more than 1 regime, leads us to conclude that the 1-regime model does not describe the data well, which suggests that the occurrence of patterns along the time horizon

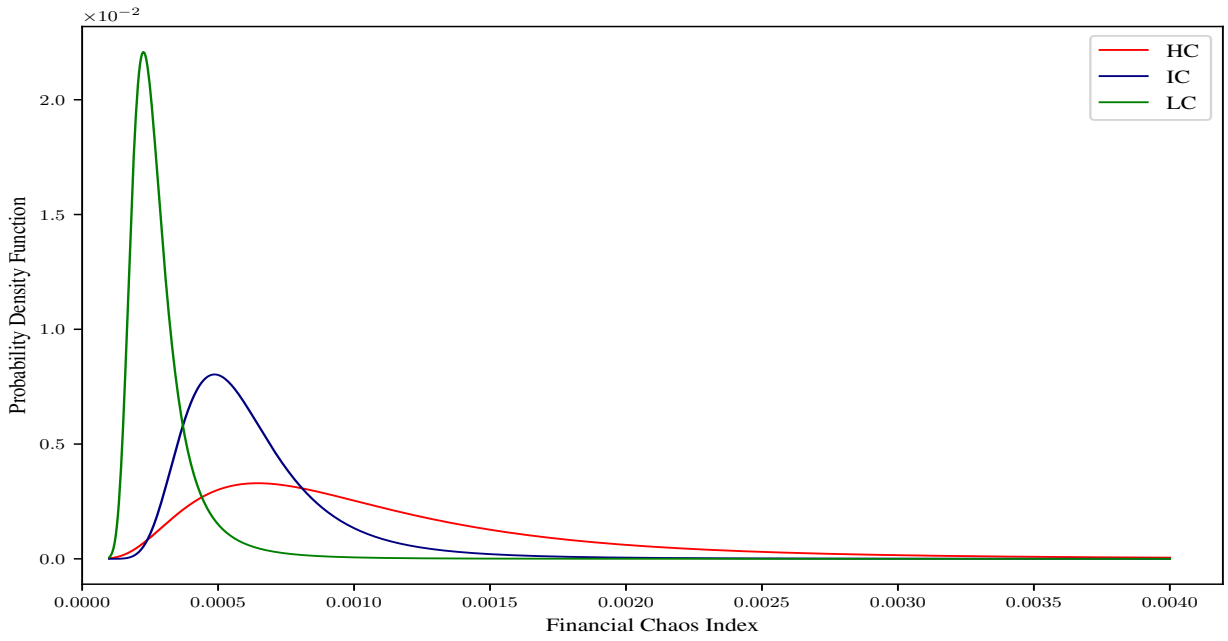

Figure 2.9: Plots of the probability density functions of market regimes.

does not follow a uniform distribution.

Further, our results obtained for the 3-regime case imply that the dominating regime (with a weight of 54.8%) corresponds to an *intermediate-chaos* (IC) regime of the stock market having the mean value of 0.0006 and the standard deviation of 0.0003, whereas less dominant *low-chaos* (LC) regime of the market (with a weight of 35.5%) has the mean and standard deviation values equal to 0.0003 and 0.0001, respectively. It is also interesting to note that both of these regimes feature relatively small values of the standard deviations, which implies that the variability of the chaos pertaining to these states of the market are quite predictable. On the other hand, it is evident that the *high-chaos* (HC) regime of the market (with a weight of 9.7%, the mean value of 0.0012 and the standard deviation of 0.0009) has a relatively negligible contribution to the underlying mixture distribution of FCI_t. Yet, in contrast to the former two regimes mentioned above, the high-chaos regime indicates a substantially higher variability which implies a lower stability of its underlying market state. Plots of probability density functions for various regimes underlying the stock market are further depicted in Figure 2.9.

Table 2.6: Results of regime analysis for daily FCI_t during January 1990-December 2019.

	1-regime	2-regime	3-regime	4-regime	5-regime
μ_1	-7.988	-8.414	-8.502	-8.287	-8.405
μ_2		-7.646	-7.697	-7.949	-7.796
μ_3			-7.204	-7.883	-7.446
μ_4				6.397	-7.057
μ_5					-6.296
σ_1	0.525	0.306	0.218	0.479	0.091
σ_2		0.245	0.308	0.535	0.204
σ_3			0.573	0.519	0.213
σ_4				0.922	0.295
σ_5					0.421
ω_1	3.159	2.662	3.584	2.131	0.091
ω_2		2.627	3.837	4.593	0.204
ω_3			3.869	3.644	0.213
ω_4				2.292	0.295
ω_5					0.421
π_1	1.000	0.522	0.355	0.246	0.039
π_2		0.478	0.548	0.093	0.206
π_3			0.097	0.651	0.225
π_4				0.010	0.245
π_5					0.285
χ^2	96.743	24.249	18.594	61.310	73.663
p-value	0.000	0.983	0.995	0.002	0.000

In addition, in order to determine which regime is dominant at each specific time, we define a random variable S_t which denotes the state of the FCI_t for a given time instant t, whose support is comprised of three states $s_t = \mathrm{HC}, \mathrm{IC}, \mathrm{LC}$. Then, the conditional probability density function that regime $S_t = s_t$ generates a particular observation ψ_t is given by

$$h(S_t = s_t|\psi_t; \boldsymbol{\theta}) = \frac{h_{s_t}(\psi_t; \boldsymbol{\theta}_{s_t})}{h(\psi_t; \boldsymbol{\theta})}. \tag{2.69}$$

2.3.2 Market Segmentation

As it was exhibited in Section 2.3.1, the daily FCI_t data stream is made of consecutive regimes that are separated by abrupt changes, owing to the fact that the underlying model producing the data switches multiple times among various regimes. In such a situation, the realization of the monthly FCI_t could be further characterized by resorting to *retrospective change-point detection* methods. In this framework, the time series which are also referred to as signals, are segmented into several *homogeneous* sub-signals.

The literature on the segmentation of time series by various change point detection methods is vast. A recent and thorough review of these methods is reported in (Truong, Oudre, and Vayatis, 2018). In our case, we perform the change-point detection on a high-dimensional mapping of $\bar{\psi}_t$ which is implicitly defined by a *kernel function*. This method is non-parametric and model-free, and it can be used to segment the time series without having any prior knowledge on the form of the underlying probability distribution that generates the data, see (Arlot, Celisse, and Harchaoui, 2019; Desobry, Davy, and Doncarli, 2005; Harchaoui and Cappé, 2007; Harchaoui, Moulines, and Bach, 2009) for more details. In short, $\bar{\psi}_t$ is first mapped onto a *reproducing Hilbert space* $\mathcal{H}$ (i.e., a Hilbert space in which the point evaluation of functions is a certain continuous linear functional) for which the associated kernel function is denoted by

$$k(\cdot,\cdot) : \mathbb{R}^{\mathcal{T}}_{\geq 0} \times \mathbb{R}^{\mathcal{T}}_{\geq 0} \to \mathbb{R}. \tag{2.70}$$

Further, the related mapping function $f : \mathbb{R}^{\mathcal{T}}_{\geq 0} \to \mathcal{H}$ is defined by

$$f(\bar{\psi}_t) = k(\bar{\psi}_t, \cdot) \in \mathcal{H}, \tag{2.71}$$

leading to the following definitions for the inner-product and norm

$$\langle f(\bar{\psi}_s) | f(\bar{\psi}_t) \rangle_{\mathcal{H}} = k(\bar{\psi}_s, \bar{\psi}_t), \tag{2.72}$$

and

$$\left\| f(\bar{\psi}_t) \right\|_{\mathcal{H}}^2 = k(\bar{\psi}_t, \bar{\psi}_t), \tag{2.73}$$

respectively, for any samples $\bar{\psi}_s, \bar{\psi}_t \in \mathbb{R}_{\geq 0}^{\mathcal{T}}$.

Next, assume that the kernel $k(\cdot, \cdot)$ is translation invariant, i.e.,

$$k(\bar{\psi}_s, \bar{\psi}_t) = \phi(\bar{\psi}_s - \bar{\psi}_t), \quad \forall s, t, \tag{2.74}$$

where ϕ is a bounded continuous positive definite function on $\mathbb{R}$. Then, the mapping $f(\cdot)$ is shown to transform piecewise i.i.d. signals into piecewise constant signals within the feature space $\mathcal{H}$. That is, the signal becomes piecewise constant once it is mapped in the high-dimensional feature space $\mathcal{H}$ if it is comprised of independent random variables with piecewise constant distributions under the assumption that $k(\cdot, \cdot)$ is translation invariant, see (Sriperumbudur et al., 2008) for more details.

Then, our change-point detection problem is aimed to detect mean-shifts in the embedded signal whose cost function is considered to be the *average scatter measure* (Harchaoui and Cappé, 2007) given by

$$\text{cost}(\bar{\psi}_t) := \sum_{t=a+1}^{b} \left\| f(\bar{\psi}_t) - \bar{\mu} \right\|_{\mathcal{H}}^2, \tag{2.75}$$

where $\bar{\mu} \in \mathcal{H}$ denotes the empirical mean of the embedded sub-signal $\{f(\bar{\psi}_t)\}_{t=a+1}^{b}$. Note that explicit computation of the mapped data is not required since it can be shown with some effort that the kernel cost function given by equation (2.75) can be expressed as follows:

$$\text{cost}(\bar{\psi}_t) = \sum_{t=a+1}^{b} k(\bar{\psi}_t, \bar{\psi}_t) - \frac{1}{b-a} \sum_{s,t=a}^{b} k(\bar{\psi}_s, \bar{\psi}_t). \tag{2.76}$$

To compute the kernel cost function given by equation (2.76), we employ the *Gaussian kernel* (also known as the *radial basis function*). This yields that

$$\text{cost}(\bar{\psi}_t) = (b-a) - \frac{1}{b-a} \sum_{s,t=a}^{b} \exp(-\gamma \left\| \bar{\psi}_s - \bar{\psi}_t \right\|^2), \tag{2.77}$$

where γ denotes a *bandwidth* parameter.

For a fixed number of change points, say $K^\star$, the change-point detection problem translates to the following discrete optimization problem:

$$\min \sum_{k=0}^{K^\star} \text{cost}(\{\bar{\psi}_t\}_{t=t_k}^{t_{k+1}}) \tag{2.78a}$$

$$\text{s.t. } |\mathbf{t}| = K^\star, \tag{2.78b}$$

$$0 = t_0 < t_1 < \cdots < t_{K^\star} < t_{K^\star+1} = |\mathcal{T}|, \tag{2.78c}$$

where $\mathbf{t} = [t_0, t_1, \cdots, t_{K^\star}, t_{K^\star+1}]^\intercal$ denotes a vector containing the change points augmented by two dummy indexes $t_0 := 0$ and $t_{K^\star+1} := |\mathcal{T}|$. By noting the additive nature of the objective function (2.78a), this optimization problem can be solved recursively by the dynamic programming that relies on the observation that (2.78) is equivalent to the following optimization problem:

$$\min \text{cost}(\{\bar{\psi}_t\}_{t=0}^{t^\star}) + \sum_{k=0}^{K^\star-1} \text{cost}(\{\bar{\psi}_t\}_{t=t_k}^{t_{k+1}}) \tag{2.79a}$$

$$\text{s.t. } t^\star \leq |\mathcal{T}| - K^\star, \tag{2.79b}$$

$$t^\star = t_0 < t_1 < \cdots < t_{K^\star-1} < t_{K^\star} = |\mathcal{T}|. \tag{2.79c}$$

This model delineates that once the optimal partition having $K^\star - 1$ elements of all sub-signals is known, then the first change point of the optimal segmentation is expected to be relatively easily evaluated.

Figure 2.10 depicts the results of solving model (2.79) for the realization $\bar{\psi}_t$ of a monthly FCI_t during the time period from January 1990 to December 2019 using the Gaussian kernel cost function given by equation (2.77). This figure indicates that the monthly FCI_t switches cyclically among multiple regimes. Further, Tables 2.7 and 2.8 report the time segments for each sub-signal and their summary statistics.

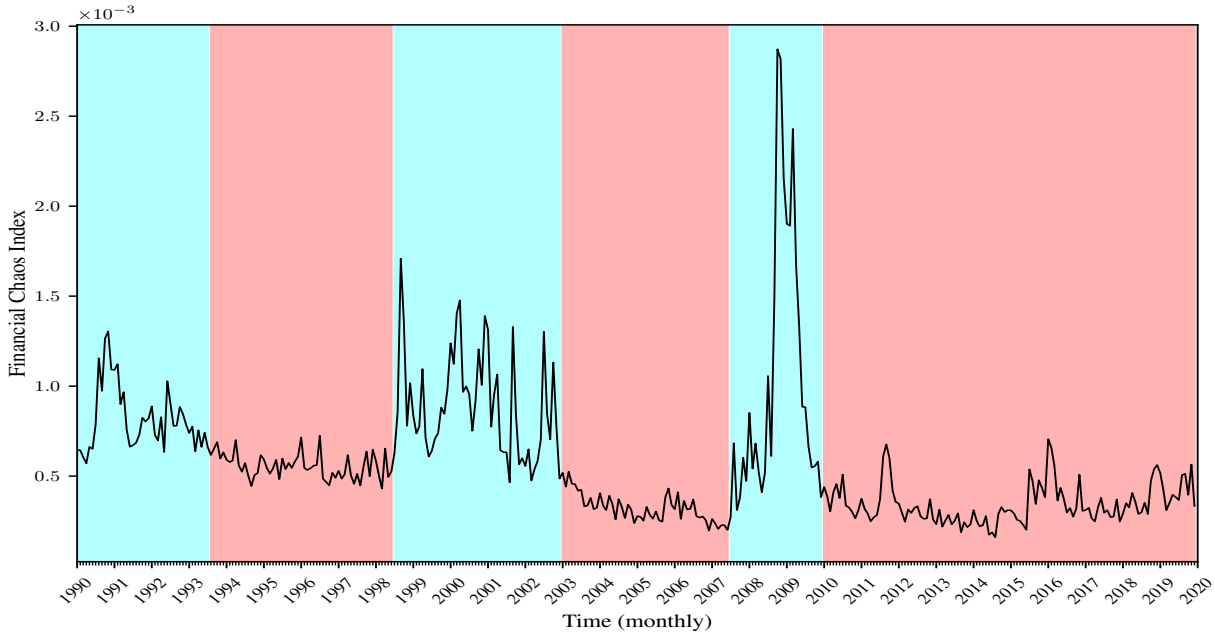

Figure 2.10: Schematic of the obtained segments based on monthly FCI_t during January 1990-December 2019.

Table 2.7: Time periods corresponding to obtained temporal segments of the market.

Segment	Time Period	Segment	Time Period
1	$1990/01 - 1993/06$	4	$2003/01 - 2007/09$
2	$1993/07 - 1998/06$	5	$2007/10 - 2009/12$
3	$1998/07 - 2002/12$	6	$2010/01 - 2019/12$

Table 2.8: Summary statistics for obtained temporal segments of the market.

Segment	1	2	3	4	5	6
Mean	8.211×10^{-4}	5.573×10^{-4}	8.871×10^{-4}	3.287×10^{-4}	1.101×10^{-3}	3.447×10^{-4}
Median	7.796×10^{-4}	5.455×10^{-4}	8.392×10^{-4}	3.163×10^{-4}	6.812×10^{-4}	3.143×10^{-4}
Maximum	1.303×10^{-3}	7.245×10^{-4}	1.708×10^{-3}	6.830×10^{-4}	2.872×10^{-3}	7.053×10^{-4}
Minimum	5.708×10^{-4}	4.301×10^{-4}	4.651×10^{-4}	1.973×10^{-4}	3.819×10^{-4}	1.601×10^{-4}
Standard deviation	1.804×10^{-4}	6.983×10^{-5}	2.917×10^{-4}	8.997×10^{-5}	7.731×10^{-4}	1.088×10^{-4}
Number of samples	42	60	54	57	27	120
Autocorrelation						
1	7.007×10^{-1}	2.314×10^{-1}	4.254×10^{-1}	4.109×10^{-1}	8.224×10^{-1}	6.493×10^{-1}
2	5.985×10^{-1}	1.470×10^{-1}	1.226×10^{-1}	3.758×10^{-1}	6.041×10^{-1}	4.462×10^{-1}
3	4.338×10^{-1}	1.475×10^{-1}	2.910×10^{-1}	4.592×10^{-1}	4.311×10^{-1}	3.448×10^{-1}
4	8.480×10^{-2}	8.691×10^{-2}	1.789×10^{-2}	2.562×10^{-1}	2.968×10^{-1}	1.565×10^{-1}
5	-1.111×10^{-1}	-2.474×10^{-2}	-4.716×10^{-2}	1.007×10^{-1}	7.126×10^{-2}	1.318×10^{-1}
6	-2.718×10^{-1}	3.197×10^{-1}	7.959×10^{-2}	2.501×10^{-1}	-1.681×10^{-1}	1.974×10^{-1}

2.4 Relation to Market Volatility

In this section, we first establish that the financial chaos index is a robust estimator of the stock market's actual volatility and then exploit its relation to *CBOE volatility index* (VIX) which is a prominent measure for the market's expected volatility implied by S&P500 options. Our goal would be to test whether FCI_t and VIX_t, also referred to as the *fear index*, both reflect similar patterns of unusual behavior pertaining to the stock market. Then, we investigate their joint long-run behavior as well as their short-run kinematics, and also their possible causal relations. Thereafter, we borrow a variety of tools developed in information theory to model the dynamics of the actual and implied volatility by formulating a time-dependent dynamical model for the coupled FCI_t-VIX_t system.

2.4.1 Robust Estimation of the Market Volatility

The term volatility, although seemingly naive on the surface, bears an intricate nature and meaning which makes its quantification a daunting task. One widely-conventional definition of the volatility ascribes it to the amount of *dispersion* of returns for a given set of assets. However, we argue here that the dispersion statistic and its variants are not the only available choices for quantification of volatility, and our developed financial chaos index is in fact an alternative measure for the market volatility, which is superior to some dispersion-based measures from a certain point of view.

Let us consider the tensor $\boldsymbol{\mathcal{A}} \in \mathcal{A}_{\mathrm{RPCT}}^{|\mathcal{T}|}$ constructed using some constitution criterion $\mathcal{C}$ as a model for embedding the stock market information whose associated comparison multigraph at each time instant $t \in \mathcal{T}$ is given by $\mathcal{G}_{\mathcal{C}}^{(t)} = (S, \mathcal{E}_{\mathcal{C}}^{(t)})$. Further, recall that the value of FCI_t would become zero for some $t = t_0$ only if the t_0-th frontal slice of $\boldsymbol{\mathcal{A}}$ would satisfy the PCJ Axiom. In other words, the fulfillment of the consistency property of the PCJ Axiom would be equivalent to existence of a set of edge weights $\mathcal{E}_{\mathcal{C}}^{(t_0)}$ based on which all activities on $\mathcal{G}_{\mathcal{C}}^{(t_0)}$ are interconnected.

Besides, it was shown in Section 2.1 that small perturbations in the neighborhood of $\boldsymbol{\mathcal{A}}$ would have negligible impact on the value of FCI_{t_0+dt}, implying presence of a similar set of interconnecting patterns on $\mathcal{G}_{\mathcal{C}}^{(t_0+dt)}$, a feature also known as near-consistency. On the contrary, a relatively large value of FCI_{t_0+dt} as compared to FCI_{t_0} would indicate a major structural change in $\boldsymbol{\mathcal{A}}$ for the time period $[t_0, t_0 + dt]$. Hence, drastically different interconnecting patterns would rather appear at time $t_0 + dt$ as compared to those delineated by $\mathcal{E}_{\mathcal{C}}^{(t_0)}$.

Moreover, the interconnecting patterns pertaining to $\mathcal{E}_{\mathcal{C}}^{(t_0+dt)}$ compared to the ones found on $\mathcal{E}_{\mathcal{C}}^{(t_0)}$ would lack the near-consistency property to a great extent if FCI_{t_0+dt} is relatively large. In view of this, making inference on dominance of relations existing among the assets at $t = t_0 + dt$ would become a near impossible task due to the lack of proper edge weights $\mathcal{E}_{\mathcal{C}}^{(t_0+dt)}$ permitting existence of a quantitative comparison scheme among all the assets involved in $\mathcal{G}_{\mathcal{C}}^{(t_0+dt)}$.

Insofar as the above-mentioned rationale is concerned, one possible definition for the market volatility that we propose is formulated in terms of the financial chaos index, which is presented

below.

Definition 2.4.1 (Market Volatility)**.** *Let the mutual information on the asset prices in the market be modeled using the tensorial formulation based on a given constitution criterion $\mathcal{C}$, and let $\mathcal{E}_{\mathcal{C}}^{(t_0)}$ be a random edge set related to the comparison multigraph $\mathcal{G}_{\mathcal{C}}^{(t_0)}$ at time $t = t_0$. Then, the market is said to be at a high-volatile (or high-chaos) regime at time $t = t_0$ if the random RPCM given by $\mathbf{A}_{\mathcal{C}}^{(t_0)}$ associated to $\mathcal{E}_{\mathcal{C}}^{(t_0)}$ satisfies at least one of the following four properties:*

1. *(sensitivity)* $\exists \delta > 0, \forall \epsilon > 0 : \lim\limits_{N\to\infty} \mathbb{P}\Big[\dfrac{\left\|\mathrm{adj}(NI - \mathbf{A}_{\mathcal{C}}^{(t_0)})\right\|_2}{N^{N-1}} < \delta\Big] < \epsilon,$

2. *(consistency)* $\exists \delta > 0, \forall \epsilon > 0 : \lim\limits_{N\to\infty} \mathbb{P}\left[\dfrac{\sum\limits_{i,j,l=1}^{N} \mathbb{I}[a_{il}^{(t_0)} a_{lj}^{(t_0)} = a_{ij}^{(t_0)}]}{N^3} < \delta\right] < \epsilon,$

3. *(discrepancy)* $\exists \delta > 0, \forall \epsilon > 0 : \lim\limits_{N\to\infty} \mathbb{P}\Big[\dfrac{\left\|\mathbf{A}^{(t_0)} - \mathbf{A}^{(t_0+dt)}\right\|_F}{\max\{\left\|\mathbf{A}^{(t_0)}\right\|_F, \left\|\mathbf{A}^{(t_0+dt)}\right\|_F\}} < \delta\Big] < \epsilon,$

4. *(homogeneity)* $\exists \delta > 0, \exists K > 0, \forall \epsilon > 0 : \lim\limits_{N\to\infty} \mathbb{P}\left[\dfrac{\sum\limits_{i,j=1}^{N} \left(\mathbb{I}[a_{ij}^{(t_0)} > K] + \mathbb{I}[a_{ij}^{(t_0)} < \frac{1}{K}]\right)}{2N^2} < \delta\right] < \epsilon.$

According to the definition of the market volatility provided above, $\mathrm{FCI}_{t=t_0}$ can be regarded as a suitable statistic to quantify the amount of market volatility at time $t_0 \in \mathcal{T}$. In the following, we present several properties of our developed market volatility statistic:

1. The statistic $\mathrm{FCI}_{t=t_0}$ is a robust estimator of the market volatility at time $t_0 \in \mathcal{T}$ as it relies on the largest eigenvalue of the RPCM at $t = t_0$, and hence, it possesses all the properties of the inconsistency measure $\psi(t = t_0)$ defined by equation (2.23).

2. In contrast to various dispersion-based approaches, the financial chaos index does not require sample points (of size greater than one) realized from a particular asset's returns to quantify the market volatility.

3. In contrast to various dispersion-based approaches, which utilize some value- or capitalization-weighting scheme to quantify the market volatility, the financial chaos index captures a different feature which pertains to quantifying the market volatility by measuring the inconsistency of mutual changes in assets' returns.

2.4.2 Cointegration Analysis

In Section 2.4.1 we showed that the financial chaos index is a representative statistic for estimating the market volatility. However, it remains an important task to investigate the relationship between the actual volatility measured by FCI_t and the implied volatility which indicates the forward-looking expectation of the volatility of the market. To this end, we employ the widely used VIX_t index of the market to study the relationships between the actual and implied types of volatility in the market.

To begin with, let us review the result obtained in Section 2.2 on daily FCI_t being an $I(1)$ process. As a consequence, the use of cointegration techniques deems necessary for co-behavior analysis of two time series. By cointegrating these two time series, the possible short-run and long-run behavior existing between the series can potentially reveal themselves. More specifically, by analyzing the cointegrated relationship between FCI_t and VIX_t, a *long-run equilibrium trajectory* can be defined, such that any departure from that path induces *equilibrium correction* that move the coupled system back towards their stable trajectory. Figure 2.11 depicts the plots of standardized daily FCI_t and VIX_t during a time period from January 1990 to December 2019, based on which clear patterns of co-movement are observed between the series.

To test for cointegration rank between FCI_t and VIX_t, we employ Johansen's procedure (Johansen, 1988) whose initial steps are to specify and estimate a *vector autoregressive* (VAR) model

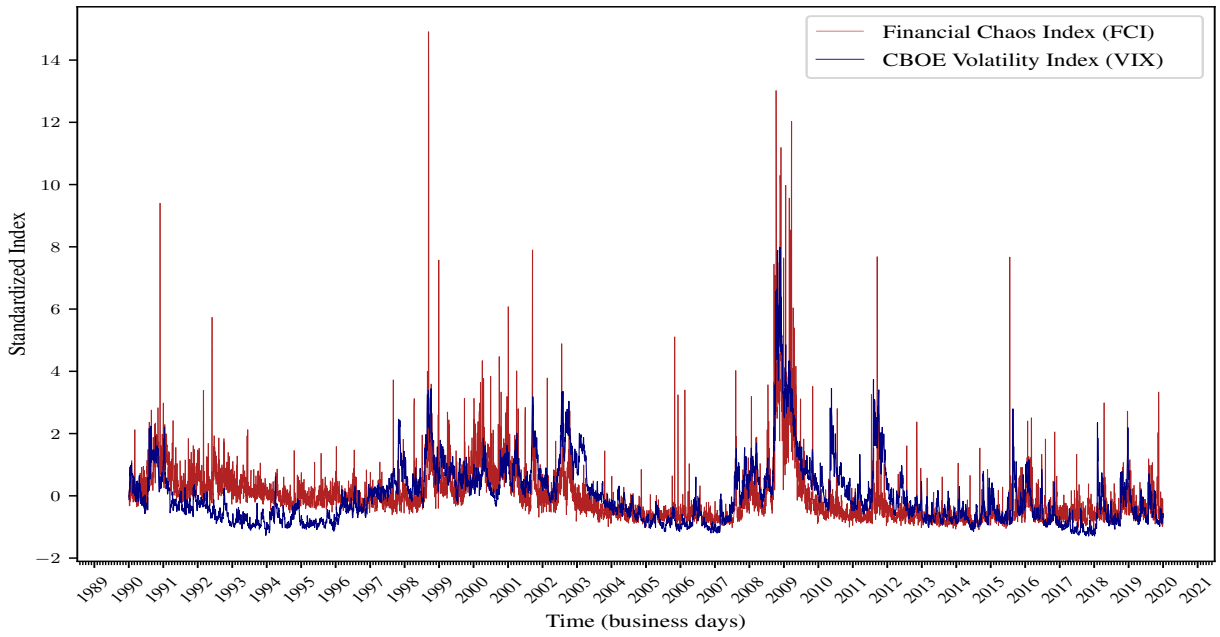

Figure 2.11: Plots of the standardized daily FCI_t vs VIX_t during January 1990-December 2019.

of order p for $\mathbf{X}_t = (\mathrm{FCI}_t, \mathrm{VIX}_t)^\intercal$ and then determine the number of cointegrating vectors by performing a likelihood ratio test for the rank of the matrix $\mathbf{\Pi}$, which is also called the *long-run impact matrix*. This is followed by estimating the resulting *vector error correction model* (VECM) by the method of maximum likelihood. More specifically, the level VAR(p) model

$$\mathbf{X}_t = \mathbf{\Phi}\mathbf{D}_t + \sum_{i=1}^{p} \mathbf{\Pi}_i \mathbf{X}_{t-i} + \boldsymbol{\epsilon}_t, \quad \forall t \in \mathcal{T}, \tag{2.80}$$

is first to be constructed. Here, $\mathbf{D}_t$ contains the deterministic terms of the model. Subsequently, in order to make the cointegrating information more visible, model (2.80) is transformed into the following VECM:

$$\Delta\mathbf{X}_t = \mathbf{\Phi}\mathbf{D}_t + \mathbf{\Pi}\mathbf{X}_{t-1} + \sum_{i=1}^{p-1} \mathbf{\Gamma}_i \Delta\mathbf{X}_{t-i} + \boldsymbol{\epsilon}_t, \tag{2.81}$$

where $\mathbf{\Phi} = \mathbf{\Phi}_1 + \mathbf{\Phi}_2 + \cdots + \mathbf{\Phi}_p - \mathbf{I}_2$. Also, noted that $\mathbf{\Gamma}_i$ are called *short-run impact matrices* which are derived according to $\mathbf{\Gamma}_i = -\sum_{k=i+1}^{p} \mathbf{\Phi}_k$ for $i = 1, 2, \ldots, p-1$. Subsequently, the *trace*

statistic

$$\mathrm{LR}_{\mathrm{trace}}(r_0) = -|\mathcal{T}| \sum_{i=r_0+1}^{2} \log(1-\hat{\xi}_i) \tag{2.82}$$

can be used to test the nested hypotheses

$$\mathrm{H}_0(r_0) : r = r_0,$$

versus one-sided alternative

$$\mathrm{H}_1(r_0) : r > r_0,$$

where $\hat{\xi}_1 \geq \hat{\xi}_2$ are the estimated eigenvalues of the 2×2 matrix $\mathbf{\Pi}$. In addition, an LR statistic called the *maximum eigenvalue statistic* is introduced in (Johansen, 1995), given by

$$\mathrm{LR}_{\max}(r_0) = -|\mathcal{T}| \log(1-\hat{\xi}_{r_0+1}), \tag{2.83}$$

which is known to be applicable for testing the following null hypothesis:

$$\mathrm{H}_0(r_0) : r = r_0,$$

against its alternative

$$\mathrm{H}_1(r_0) : r = r_0 + 1.$$

By noting that $\mathrm{rank}(\mathbf{\Pi}) \leq 1$, we may then formulate the error-correction cointegrating null hypotheses and their alternatives as follows:

$$\begin{cases} \mathrm{H}_{0,1}^{\mathrm{EC}} : r = 0, \\ \mathrm{H}_{0,2}^{\mathrm{EC}} : r = 1, \end{cases} \qquad \begin{cases} \mathrm{H}_{1,1}^{\mathrm{EC}} : r > 0, \\ \mathrm{H}_{1,2}^{\mathrm{EC}} : r > 1, \end{cases}$$

where r denotes the cointegration rank. These hypothesis could subsequently be tested in the following order $\mathrm{H}_{0,1}$ and $\mathrm{H}_{0,2}$ such that $\mathrm{H}_{0,2}$ can only be rejected if $\mathrm{H}_{0,1}$ is rejected, i.e., the testing procedure gets terminated by returning the result $r = 0$ if one fails to reject $\mathrm{H}_{0,1}$, otherwise $\mathrm{H}_{0,2}$ is

Table 2.9: Results of Johonson's cointegration test for daily FCI_t and VIX_t.

		$\mathrm{LR}_{\mathrm{trace}}$		$\mathrm{LR}_{\mathrm{max}}$	
		$\mathrm{H}_{0,1}^{\mathtt{EC}}$	$\mathrm{H}_{0,2}^{\mathtt{EC}}$	$\mathrm{H}_{0,1}^{\mathtt{EC}}$	$\mathrm{H}_{0,2}^{\mathtt{EC}}$
p-value		0.001	0.051	0.001	0.051
Test statistics		119.601	4.118	115.484	4.118
Critical values	1% level	16.359	6.940	15.090	6.940
	5% level	12.321	4.130	11.225	4.130
	10% level	10.474	2.976	9.475	2.976

tested to determine whether $r = 1$ or not.

Table 2.9 reports the results for testing $\mathrm{H}_{0,1}$ and $\mathrm{H}_{0,2}$ using Johansen's procedure. As per this table, $\mathrm{H}_{0,1}$ gets rejected at all the considered significance levels using either $\mathrm{LR}_{\mathrm{trace}}$ or $\mathrm{LR}_{\mathrm{max}}$ test statistics, which implies that $r > 0$. On the other hand, $\mathrm{H}_{0,2}$ gets rejected only at 10% level of significance, whereas there is insufficient evidence to reject $\mathrm{H}_{0,2}$ at 1% significance level by using either of these test statistics. Consequently, we may conclude that $\mathbf{X}_t$ is cointegrated with one cointegrating vector (i.e., $\mathrm{rank}(\mathbf{\Pi}) = 1$), which in turn implies that there exist a common $I(1)$ trend driving both FCI_t and VIX_t. It is worth mentioning that this common trend is more likely to be the so-called *implicit common efficient price* which permanently inflicts movements in both time series by following the new market information. Further, note that the optimal lag value in our experiments was identified to be $p = 16$ by fitting VAR models to the data and evaluating the performance of various model selection criteria such as AIC, BIC and HQC.

2.4.3 Information Kinematics

In order to investigate the causal relations which might (or might not) exist within the coupled economical system $\mathrm{FCI}_t - \mathrm{VIX}_t$, we first note that $\mathrm{rank}(\mathbf{\Pi}) = 1$, ergo the matrix $\mathbf{\Pi}$ can be decomposed as follows:

$$\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}^{\intercal}, \tag{2.84}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $(\mathcal{T} \times 1)$ vectors. It is known that matrix $\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}^{\intercal}$ plays an important role in separating the short-run and long-run relationships between FCI_t and VIX_t. For instance, $\boldsymbol{\alpha}$ determines the corresponding *error-correction speeds* which distributes the influence of FCI_t and VIX_t onto the temporal behavior of $\Delta\mathbf{X}_t$, whereas $\boldsymbol{\beta}$ provides the basis vectors for the space of cointegrating relations. That is, $\boldsymbol{\beta}^{\intercal}\mathbf{X}_{t-1}$ can be considered as the error-correction term, reflecting long-run equilibrium relations between FCI_t and VIX_t.

Therefore, the model given by equation (2.81) can be rewritten as follows:

$$\Delta\mathbf{X}_t = \boldsymbol{\alpha}\boldsymbol{\beta}^{\intercal}\mathbf{X}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta\mathbf{X}_{t-i} + \boldsymbol{\epsilon}_t, \tag{2.85}$$

where the deterministic terms vanish since FCI_t and VIX_t are both $I(1)$ without the shift or trend terms (similar unit root and stationarity tests can be conducted for VIX_t to demonstrate that it is an $I(1)$ without shift or trend terms). Note that such a substitution (i.e., $\boldsymbol{\Phi}\mathbf{D}_t = 0$) can be justified by the work presented in (Johansen, 1995) in which the author considers the following restriction on the form of the deterministic terms in model given by equation (2.81):

$$\boldsymbol{\Phi}\mathbf{D}_t = \boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t. \tag{2.86}$$

Furthermore, it might be promising to consider VECM as a VAR model subject to cointegration constraints. Thus, the long-run behavior of the endogenous variables will be restricted by VEC expressions in the presence of stupendous short-run dynamic fluctuations, where the VEC expressions will converge to their cointegration relation (Zou, 2018). On the basis of the experiments which we have conducted, the impact matrix was estimated to be

$$\widehat{\boldsymbol{\Pi}} = \begin{bmatrix} -0.139 & 0.121 \\ 0.004 & -0.004 \end{bmatrix}. \tag{2.87}$$

In turn, it is possible to use the impact matrix given by equation (2.87) in order to characterize the long-run relationship between FCI_t and VIX_t. We illustrate this relationship schematically in

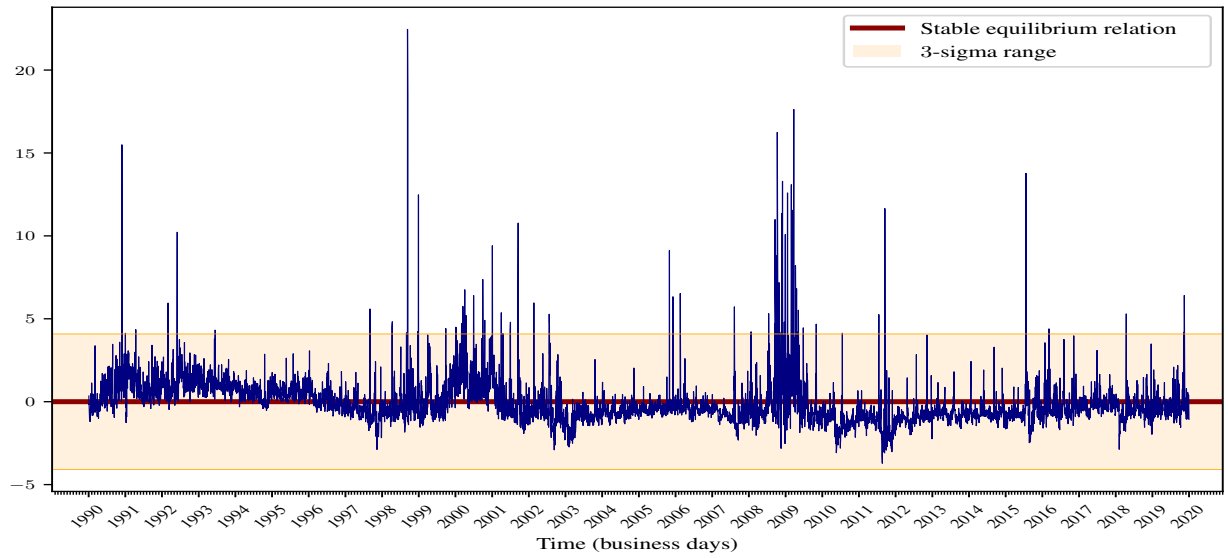

Figure 2.12: Plot of the long-run equilibrium relationship between daily FCI_t and VIX_t during January 1990-December 2019.

Figure 2.12, where the horizontal axis represents a stable relation between variables. This figure indicates to a presence of a long-run common movement between FCI_t and VIX_t along the considered time horizon. Note that FCI_t and VIX_t are tied together through cointegration, and they both respond to disturbances which make adjustments towards the long-run equilibrium relationship. In this respect, Figure 2.12 reveals that there exist several departures of large magnitude from the equilibrium behavior of the two time series. Such fluctuations are observed to occur mainly during the time periods when the market has undergone severe crises, some of which for instance were listed on previous pages to have happened within the time frame $2008 - 2009$.

In contrast to the analysis of the long-run equilibrium, identification of the short-run relations appear to be a more difficult task since there usually are quite a few matrices of coefficients whose interpretations are often not trivial. To surmount this obstacle, we incorporate the notions of the *impulse-response function* (IRF) and the *cross-correlation function* (XCF) along with the notion of approximate entropy within our framework to study the existing short-run relations.

The sample cross-correlation function, which measures the similarities between FCI_t and var-

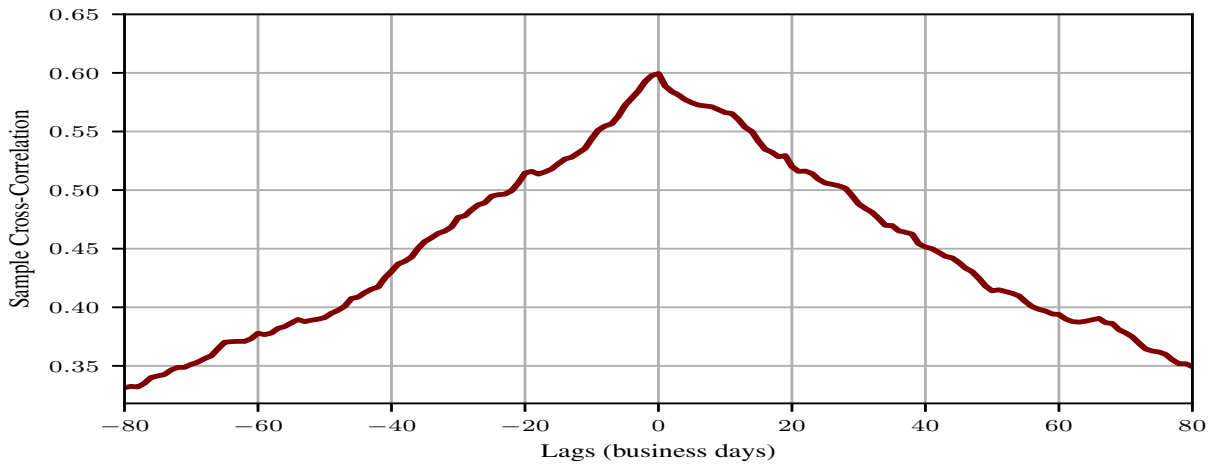

Figure 2.13: Plot of the sample cross-correlation between daily FCI_t and VIX_t during January 1990-December 2019.

ious time-lagged versions of VIX_t as a function of the lag parameter, is introduced as follows:

$$\mathrm{XCF}(p) = \frac{c(p)}{\sqrt{Var(\mathrm{FCI}_t)}\sqrt{Var(\mathrm{VIX}_t)}}, \quad p = 0, \pm 1, \pm 2. \ldots . \tag{2.88}$$

Here, we estimate the lag-p cross-covariance $c(p)$ as follows:

$$c(p) = \begin{cases} \dfrac{1}{|\mathcal{T}|} \sum\limits_{t=1}^{|\mathcal{T}|-p} (\mathrm{FCI}_t - \overline{\mathrm{FCI}_t})(\mathrm{VIX}_{t+p} - \overline{\mathrm{VIX}_t}), & p = 0, 1, 2, \ldots, \\ \dfrac{1}{|\mathcal{T}|} \sum\limits_{t=1}^{|\mathcal{T}|+p} (\mathrm{VIX}_t - \overline{\mathrm{VIX}_t})(\mathrm{FCI}_{t-p} - \overline{\mathrm{FCI}_t}), & p = 0, -1, -2, \ldots . \end{cases} \tag{2.89}$$

As depicted in Figure 2.13, the time delay at which two signals are aligned best corresponds to $p = 0$, where the XCF is roughly 60%. However, it is observed that the amount of XCF remains relatively high even when the time lag parameter is considered to be as large as $p = \pm 80$ business days, implying that the impacts of any change in either of the time series could persist in the other series for several time periods ahead.

In order to track an impact that either time series makes on the other, we confine ourselves to the framework of *empirical causal analysis* which emphasizes the *orthogonalized impulse-response functions* to conduct the causal analysis in between these two time series. Basically, an impulse-

response function describes the evolution of the reaction from one time series along a specified time horizon after a shock is inflicted on the other time series at a given moment.

Let us revisit VAR(p) model given by equation (2.80) and assume that its innovations $\{\boldsymbol{\epsilon}_t\}$ are $N(0, \boldsymbol{\Sigma})$, which are further assumed to be uncorrelated over time. Recall that the deterministic term in equation (2.80) was shown to vanish. Hence, the VAR(p) model can be expressed as follows:

$$\mathbf{X}_t = \sum_{i=1}^{p} \boldsymbol{\Pi}_i \mathbf{X}_{t-i} + \boldsymbol{\epsilon}_t, \quad \forall t \in \mathcal{T}. \tag{2.90}$$

Then, in order to estimate the effect of a one-unit shock in $\boldsymbol{\epsilon}_t$ to endogenous variables of the system, we decompose the variance-covariance matrix into a lower triangular matrix using the *Cholesky decomposition*. Let $\mathbf{P}$ denote the Cholesky decomposition of $\boldsymbol{\Sigma}$ such that $\mathbf{P}\mathbf{P}^\intercal = \boldsymbol{\Sigma}$ and $\mathbf{P}$ is a lower triangular matrix, and assume that $\mathbf{u}_t$ is such that $\mathbf{u}_t = \mathbf{P}^{-1}\boldsymbol{\epsilon}_t$. Our experiments, led to the following estimate for $\mathbf{P}$:

$$\widehat{\mathbf{P}} = \begin{bmatrix} 0.0003 & 0.0000 \\ 0.1336 & 1.4898 \end{bmatrix}. \tag{2.91}$$

We note that the orthogonalized impulse response function would reflect a one standard-deviation impulse to $\mathbf{u}_t$ since

$$\mathbb{E}[\mathbf{u}_t \mathbf{u}_t^\intercal] = \mathbf{P}^{-1}\mathbb{E}[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^\intercal](\mathbf{P}^\intercal)^{-1} = \mathbf{I}_2. \tag{2.92}$$

Furthermore, by converting the VAR(p) to an *infinitive moving average process* MA(∞), one obtains that

$$\mathbf{X}_t = \sum_{p=0}^{\infty} \boldsymbol{\Phi}_p \mathbf{P} \mathbf{u}_{t-p}, \tag{2.93}$$

where $\boldsymbol{\Phi}_p$ is the MA(∞) matrix of coefficient for the p-th lag. As pointed out in (Lütkepohl, 2005), these matrices of coefficients are related to the autoregressive matrices via the following set

of relationships:

$$\mathbf{\Phi}_0 = \mathbf{I}_2, \tag{2.94}$$

$$\mathbf{\Phi}_1 = \mathbf{\Pi}_1, \tag{2.95}$$

$$\mathbf{\Phi}_2 = \mathbf{\Phi}_1\mathbf{\Pi}_1 + \mathbf{\Pi}_2, \tag{2.96}$$

$$\vdots \tag{2.97}$$

$$\mathbf{\Phi}_i = \mathbf{\Phi}_{i-1}\mathbf{\Pi}_1 + \mathbf{\Phi}_{i-2}\mathbf{\Pi}_2 + \cdots + \mathbf{\Phi}_{i-p}\mathbf{\Pi}_p. \tag{2.98}$$

Subsequently, the time-horizon h responses on the i-th time series due to a one-unit shock in the j-th equation at a given time instant t are given by

$$\frac{\partial x_{i,t+h}}{\partial \epsilon_{j,t}} = \{\mathbf{\Phi}_h\}_{ij}. \tag{2.99}$$

Figures 2.14 and 2.15 depict the effect of the impact of a one-unit shock in FCI_t on VIX_t for 80 business days ahead and vice versa, respectively. In order to characterize the regularity and information content of the orthogonalized IRFs, the approximate entropy of the signals are further computed, yielding 0.098 for $\mathrm{FCI}_t \rightsquigarrow \mathrm{VIX}_t$ and 0.160 for $\mathrm{VIX}_t \rightsquigarrow \mathrm{FCI}_t$. These quantities imply that FCI_t is more sensitive to changes in VIX_t rather than the other way around, as the orthogonalized IRF for $\mathrm{VIX}_t \rightsquigarrow \mathrm{FCI}_t$ attains a higher value of approximate entropy, and hence, it has a higher degree of irregularity and unpredictability. It is also worth to mention the synchronicity which is evident like that when both orthogonalized IRF graphs reach their maximuma at day 10 after the shocks were being imposed. Besides, a magnitude of response for the case of $\mathrm{VIX}_t \rightsquigarrow \mathrm{FCI}_t$ is considerably larger than its counterpart $\mathrm{FCI}_t \rightsquigarrow \mathrm{VIX}_t$, which stipulates that a one-unit shock in VIX_t causes a more intense response pertaining to FCI_t as compared to the case where the shock is imposed on FCI_t.

One rationale that partially accounts for the above-mentioned assessments is rooted in the fact that the financial chaos index is a robust estimator of the mutual price changes of the market assets. That is, any substantial marginal increase in the value of FCI_t would demand majority

of the underlying assets (S&P500) to depart significantly from their steady-state price levels. As a consequence, any abrupt change in the value of FCI_t due to a shock could potentially be an important indicator of some structural changes in the underlying assets. In contrast to abrupt changes in FCI_t, imposing a one-unit shock in FCI_t is less likely to be a cause of significant structural changes in the asset prices. Thus, the response of such a shock cannot have a large magnitude since VIX_t would respond sharply to a shock in FCI_t only if the underlying equities (S&P500) utilized in its computation undergo major changes. The foregoing statement accounts for the relatively lower magnitude of VIX_t response to a shock in FCI_t, as depicted in Figure 2.14. This figure further discloses that an increase in the FCI_t values causes positive increments in VIX_t response with the positive effect peak to occur within the first few days, which is then followed by an oscillation until it meets the highest-magnitude response. Thereafter, the response of VIX_t to a shock in FCI_t remains positive for the remainder of the time horizon.

Conversely, the response of FCI_t to a shock in VIX_t has a slightly different behavior than that explained above. As depicted in Figure 2.15, VIX_t has a positive *contemporaneous effect* on FCI_t and this positive response persists during the whole time horizon. However, it is relevant to stress the presence of a steep decline in response of FCI_t which occurs immediately after the mentioned contemporaneous effect at time instant t_0 (the "double dip" observed within $[t_0, t_0+5]$). In order to interpret this empirical observation, we find it useful to recall that the lower values of the financial chaos index are linked to a market in which the *collective judgment* of its participants remain consistent over time (here, we understand consistency in the same sense as that described under the PCJ Axiom). Thus, a positive increase in FCI_t caused by the contemporaneous effect could potentially be understood as a deviation from the *collective consistency condition*. On the other hand, the response from FCI_t decreases immediately after a day is elapsed since the collective judgment of the participants in the market becomes more aware of the inconsistencies that have occurred due to an increased amount of the expected volatility. Thus, the actual market volatility reacts to this phenomenon by correcting its value. However, the response from FCI_t spikes once again after a few days, albeit this time on the reason that the implied volatility becomes realized in the market, which further makes the task of casting a series of consistent judgments less probable

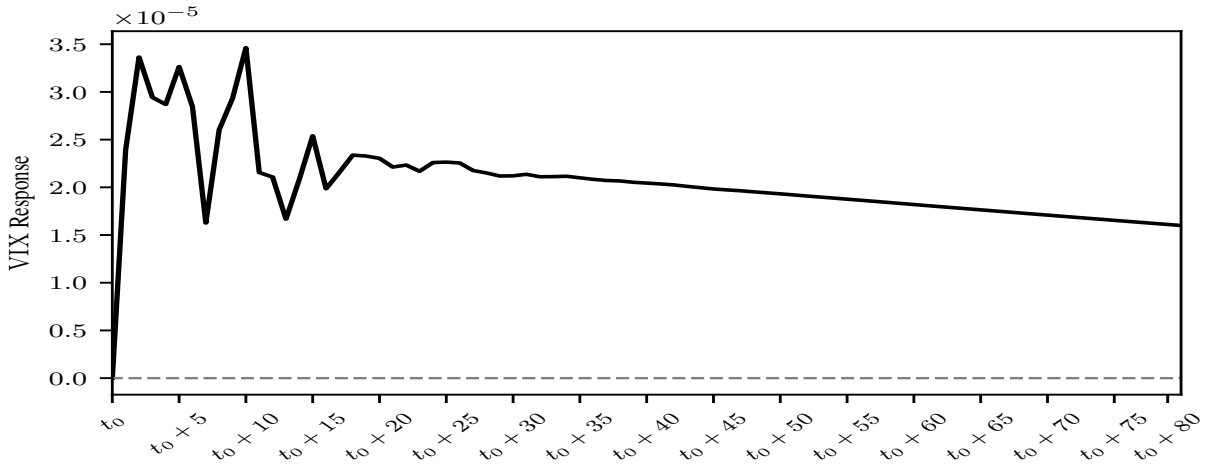

Figure 2.14: Plot of the orthogonalized IRF for $\mathrm{FCI}_t \rightsquigarrow \mathrm{VIX}_t$ during January 1990-December 2019, $\mathrm{ApEn} = 0.098$.

for the agents, which in turn gives rise to increased values of the FCI_t response.

To conclude our discussion, it can be said that VIX_t has a stronger causal effect on FCI_t as compared to the opposite. Hence, the implied volatility of the market plays a key role in characterization of the market's actual volatility, further implying that a set of feedback and feedforward causal relations govern the interactions between FCI_t and VIX_t.

2.4.4 Information Dynamics

In the previous subsection, we investigated the short- as well as long-run co-behavior of FCI_t and VIX_t and showed the existence of a potential bi-directional causal relation between these two time series. In what follows, we consider an additional aspect of the dynamical relation existing between the series which pertains to the scheme by which the information content flows from one time series to another and their possible interactions. For this purpose, we first assume that the coupled system FCI_t-VIX_t is a given isolated system. We will then introduce and utilize the notions of *transfer entropy* and *self-entropy* to quantify the dynamics of the information flow within the coupled system by taking into account the temporal evolution of the existing information dynamics.

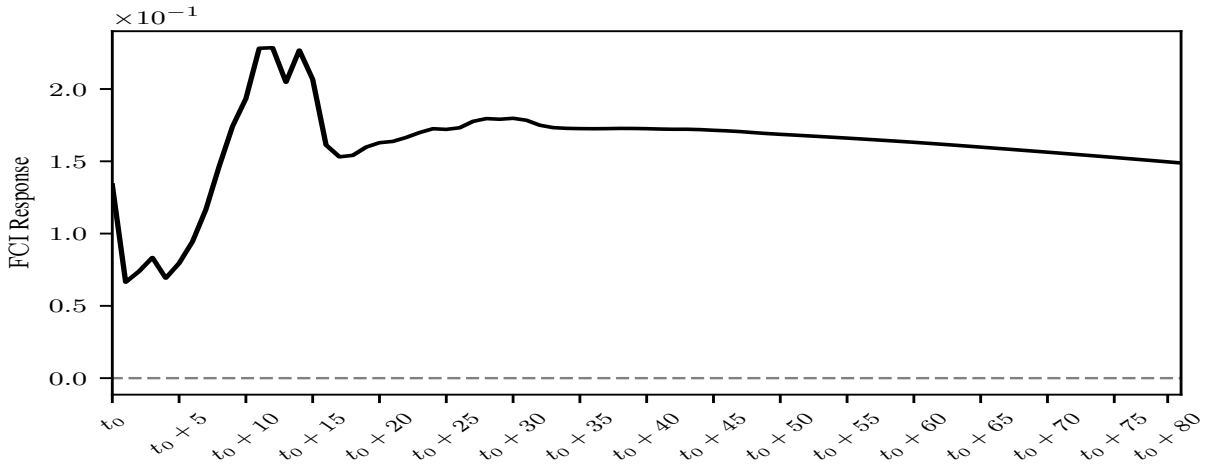

Figure 2.15: Plot of the orthogonalized IRF for $\mathrm{VIX}_t \rightsquigarrow \mathrm{FCI}_t$ during January 1990-December 2019, $\mathrm{ApEn} = 0.160$.

Let V and Z denote two dynamical processes which further represent a time-evolutionary coupled dynamical system, and let us assume that Z is assigned to be the target process. Then, one can employ some standard information-theoretic measures to study the information relevant to Z (Cover and Thomas, 2012; Faes and Porta, 2014). A pivotal measure of such kind is the well-known *Shannon entropy* which quantifies the amount of information carried by Z in terms of its average uncertainty, i.e.,

$$H(Z) = -\sum p(z) \log p(z), \tag{2.100}$$

where $p(z)$ denotes the underlying probability mass function of the discrete random variable Z. Also, another important quantity to consider is the *conditional entropy*. It is defined by

$$H(Z|V) = -\sum p(z, v) \log p(z|v), \tag{2.101}$$

and represents the average uncertainty in Z when a realization of the random variable V is known.

The dynamical properties of our coupled system can then be studied by introducing the notion of the *transition probability* which is the probability of transition of the discrete-time system from its past states to the current one. Denote by Z_t the present state of the process (time series) Z, while $\mathbf{Z}_t^- = [Z_{t-1}, Z_{t-2}, \cdots]^\intercal$ represents the *state variable* of the process Z, which contains the

complete information related to the past of the process Z. The *self-entropy* for Z is then defined as follows:

$$S_Z = \sum p(z_t, \mathbf{z}_t^-) \log \frac{p(z_t|\mathbf{z}_t^-)}{p(z_t)}, \tag{2.102}$$

where z_t and $\mathbf{z}_t^-$ denote the corresponding realizations of Z_t and $\mathbf{Z}_t^-$, respectively. Basically, the self-entropy assesses the influence of the past states of the target process Z onto its present state. In other words, the self-entropy given by equation (2.102) measures the extent to which the uncertainty about Z_t is reduced by the knowledge of $\mathbf{Z}_t^-$. The self-entropy for the process Z ranges from zero to $H(Z_t)$, where the zero corresponds to the situation where the information from the past states ($\mathbf{Z}_t^-$) does not contribute to reduction of uncertainty in Z_t, and the latter quantity is associated to the case where the whole uncertainty is dissipated about Z_t after acquiring the knowledge of $\mathbf{Z}_t^-$.

Besides, in order to assess the influence of imparting the past states of the process V_t on Z_t, i.e. to evaluate the information contained in $\mathbf{V}_t^-$, one can make use of the well-known *transfer entropy* measure (Schreiber, 2000), which is given by the following formula:

$$T_{V\to Z} = \sum p(z_t, \mathbf{v}_t^-, \mathbf{z}_t^-) \log \frac{p(z_t|\mathbf{v}_t^-, \mathbf{z}_t^-)}{p(z_t|\mathbf{z}_t^-)}. \tag{2.103}$$

Similar to self-entropy, the smallest value $T_{V\to Z} = 0$ is attained when $\mathbf{V}_t^-$ does not induce an uncertainty reduction in Z_t beyond what is already contributed by $\mathbf{Z}_t^-$. In contrast, the largest value of transfer entropy $T_{V\to Z} = H(Z_t|\mathbf{Z}_t^-)$ is taken on when the totality of uncertainty about Z_t that was not already dissipated by the knowledge of $\mathbf{Z}_t^-$ is reduced by the information provided through imparting $\mathbf{V}_t^-$. In other words, the direct information measure, signified by the transfer entropy, quantifies the loss of information when assuming V_t does not causally influence Z_t when it actually does, e.g., see to (Amblard and Michel, 2013) for more details.

Note that the self-entropy and transfer entropy measures introduced above, can also be expressed as follows:

$$S_Z = H(Z_t) - H(Z_t|\mathbf{Z}_t^-), \tag{2.104}$$

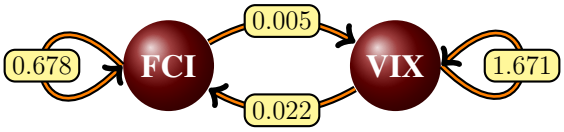

Figure 2.16: Computed diagram of information flow between FCI_t and VIX_t during January 1990-December 2019.

and

$$T_{V\to Z} = H(Z_t|\mathbf{Z}_t^-) - H(Z_t|\mathbf{V}_t^-, \mathbf{Z}_t^-). \tag{2.105}$$

The values of self-entropy and transfer entropy computed[3] using daily realizations of FCI_t and VIX_t during the time period from January 1990 to December 2019 are depicted in Figure 2.16, whereby it can be seen that $T_{\mathrm{VIX}\to\mathrm{FCI}}$ is considerably larger than $T_{\mathrm{FCI}\to\mathrm{VIX}}$. This suggests that the time-dependent information transfer is much stronger in one direction than the other one. This observation supports our previous discussions on orthogonalized impulse-response functions where it was shown that the magnitude of impact on FCI_t responding to a shock in VIX_t was observed to be much greater than that of VIX_t reacting to shock in FCI_t. Therein, it was also discussed that a shock to VIX_t had a contemporaneous effect on FCI_t, but not vice versa.

A subsequent combination of the results obtained using the above information-theoretic measures, namely, the transfer entropy and self-entropy, with those derived by employing the orthogonalized impulse-response functions leads us to draw a conclusion on the direction of causal relations between these two processes. Namely, we conclude that there is a substantial evidence favoring existence of a bi-directional causal relation between FCI_t and VIX_t. Furthermore, this causal relation implies that VIX_t influences FCI_t instantaneously with a significant dispense of information, whereas the opposite casual relation does not share a similar instantaneous nature. Yet, it still contributes to a substantial transfer of information from FCI_t to VIX_t, where the dispensed

[3]Transfer entropy quantities were computed using the RTransferEntropy package in R programming language (Behrendt, Dimpfl, Peter, and Zimmermann, 2019), while the self-entropy quantities were computed using the ITS MATLAB toolbox for practical computation of information dynamics available online at `http://www.lucafaes.net/its.html` which is developed based on the methodologies presented in (Faes, Porta, and Nollo, 2015; Faes, Porta, Nollo, and Javorka, 2017; Xiong, Faes, and Ivanov, 2017).

information are most likely to be transferred in a time-lagged manner.

In fact, the results described above address several long-standing questions in finance on the nature of a relation between equity markets and option markets. Our results suggest that there exist a relationship between an option contract written on an asset and its associated price, since both values $T_{\mathrm{FCI}\to\mathrm{VIX}}$ and $T_{\mathrm{VIX}\to\mathrm{FCI}}$ obtained are strictly positive. In view of the strict positivity of $T_{\mathrm{VIX}\to\mathrm{FCI}}$, we can conclude that options convey new information to *price discovery*[4] of their underlying asset, implying departure of the market from being a *complete*[5] one.

Also, by considering VIX_t to be an amalgam of the information reflected in the prices of all the S&P500 options, our analysis further rejects the hypothesis which is often made by economists who view options as assets being independent from one another on the option market whose prices depend solely on their underlying equities. Our rationale for rejecting the above-mentioned hypothesis is due to the fact that a sufficiently large value of S_{VIX} is an indicator of presence of potentially significant internal relational clusters in the option market.

Next, we extend our analysis to the situation where one concerns with the evolution of information dynamics between FCI_t and VIX_t over time. To this end, we formulate the following dynamical system model to describe the various mechanisms of exchange of information between the considered processes. Namely, we propose the following time-dependent dynamical system model:

$$\begin{cases} \dfrac{\partial F}{\partial t} = (\gamma+\theta)F^2 - \alpha F, \\ \\ \dfrac{\partial V}{\partial t} = \beta V^2 - \delta V, \end{cases} \tag{2.106}$$

where $F := \mathrm{FCI}_t$ and $V := \mathrm{VIX}_t$. Also, parameters α, β, γ, δ and θ represent $T_{\mathrm{FCI}\to\mathrm{VIX}}$, $T_{\mathrm{VIX}\to\mathrm{FCI}}$, S_{FCI}, S_{VIX}, $\mathrm{ApEn(iVrF)}$, respectively, where iVrF denotes the orthogonalized response function of FCI_t to an impulse in VIX_t. These parameters are also illustrated in Figure 2.17. It deserves

[4]The overall process of setting the price of an asset, whether explicitly or in some inferred manner.

[5]In complete market, the transaction costs are negligible and perfect information are conveyed across the market. Meaning that, all participants of the market have access to perfect and instantaneous knowledge of all market prices and their own utility and cost functions. Furthermore, in every possible state of the world, there is a price for every asset in the complete market.

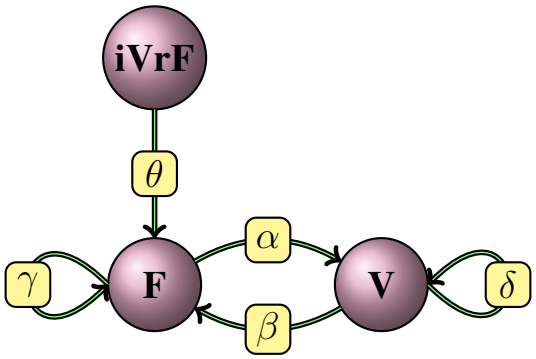

Figure 2.17: Symbolic diagram of information flow between FCI_t and VIX_t.

to mention that, the term θ is involved in model (2.106) as a result of the contemporaneous effect that a one-unit change in VIX_t exerts on FCI_t. Note that the approximate entropy can potentially be viewed as a measure of information content in the orthogonalized IRF for $\mathrm{VIX}_t \rightsquigarrow \mathrm{FCI}_t$ during the specified time horizon.

In order to characterize the model (2.106) in more in-depth, we perform its stability analysis. It can be shown with some effort that the four critical points of this dynamical system are as follows:

$$(F_1, V_1) = (\frac{\alpha}{\gamma+\theta}, \frac{\delta}{\beta}),\ (F_2, V_2) = (0, 0),\ (F_3, V_3) = (\frac{\alpha}{\gamma+\theta}, 0),\ \text{and}\ (F_4, V_4) = (0, \frac{\delta}{\beta}).$$

Note that these critical points are all feasible as FCI_t and VIX_t can take on zero values from the theoretical point of view. Without the loss of generality, we impose a set of non-zero constraints on β and $\gamma + \theta$. These constraints will prevent unboundedness of the critical points which can lead to the cases of unbounded actual as well as implied types of market volatility.

Thereafter, the Jacobian matrix is obtained as

$$J = \begin{pmatrix} 2(\gamma+\theta)F - \alpha & \theta \\ \theta & 2\beta V - \delta \end{pmatrix}. \tag{2.107}$$

Subsequently, we compute the eigenvalues at each of the four critical points, which yield

$$(\lambda_1, \lambda_1^{'}) = (\frac{\alpha}{\delta + \theta}, \frac{\delta}{\beta}),\ (\lambda_2, \lambda_2^{'}) = (-\alpha, -\delta),\ (\lambda_3, \lambda_3^{'}) = (\alpha, -\delta),\ \text{and}\ (\lambda_4, \lambda_4^{'}) = (-\alpha, \frac{\delta}{\beta}).$$

Thereafter, we apply the *center manifold theorem* to study stability of the critical points by deriving their dimensions of stable ($\mathrm{Dim}(E^S)$), unstable ($\mathrm{Dim}(E^U)$) and center ($\mathrm{Dim}(E^C)$) manifolds. Using the previously computed values of $\alpha = 0.005$, $\beta = 0.022$, $\gamma = 0.678$, $\delta = 1.671$ and $\theta = 0.160$, the first critical point is evaluated to be $(F_1, V_1) = (0.006, 75.954)$. This critical point also has the following manifolds' dimensions:

$$\mathrm{Dim}(E^S) = 0,\ \mathrm{Dim}(E^U) = 2,\ \text{and}\ \mathrm{Dim}(E^C) = 0,$$

which indicates that in fact it is a saddle point.

It is interesting to note that the elements of $(F_1, V_1) = (0.006, 75.954)$ reside on roughly all-time highs of their respective time series FCI_t ($\max = 0.0073$ and $\mathrm{mean} = 0.0006$) and VIX_t ($\max = 80.860$ and $\mathrm{mean} = 19.146$). In other words, there could potentially exist hidden acting forces which prevent the market from retaining its volatility state at the neighborhoods of extrema. This assertion can be verified by noting that the average values for FCI_t and VIX_t during the considered time period January 1990-December 2019 are both substantially less than their extrema, an observation which could well be due to the fact that (F_1, V_1) is a source.

It is also promising to analyze various properties of the manifolds for (F_1, V_1) conditioned on different values of the parameters α and δ, i.e.,

$$\begin{cases} \mathrm{Dim}(E^S) = 0,\ \mathrm{Dim}(E^U) = 2,\ \text{and}\ \mathrm{Dim}(E^C) = 0, & \text{if} \quad \alpha > 0 \text{ and } \delta > 0, \\ \mathrm{Dim}(E^S) = 0,\ \mathrm{Dim}(E^U) = 1,\ \text{and}\ \mathrm{Dim}(E^C) = 1, & \text{if} \quad \alpha = 0 \text{ and } \delta > 0, \\ \mathrm{Dim}(E^S) = 0,\ \mathrm{Dim}(E^U) = 1,\ \text{and}\ \mathrm{Dim}(E^C) = 1, & \text{if} \quad \alpha > 0 \text{ and } \delta = 0, \\ \mathrm{Dim}(E^S) = 0,\ \mathrm{Dim}(E^U) = 0,\ \text{and}\ \mathrm{Dim}(E^C) = 2, & \text{if} \quad \alpha = 0 \text{ and } \delta = 0. \end{cases}$$

For the second case where $\alpha = 0$ and $\delta > 0$, which also pertains to the same equilibrium solution of (F_4, V_4), the critical point is estimated to yield $(0, 75.954)$ whose dimension of the stable manifold is one. In this case, typical trajectories of its phase plane travel in closed elliptical orbits in some direction in the neighborhood of the critical point. Such a point is not asymptotically stable as trajectories may never approach $(0, 75.954)$. However, by starting off sufficiently close to this point, the trajectory will remain close to the critical point in some sense. The dilemma for the behavior of nonlinear systems, which is the case here, is to identify whether the direction of the trajectory points inwards or outwards. The fact is that the trajectory can be either pulled towards or pushed away from the critical point in nonlinear systems. Nevertheless, this can be resolved by considering the dimension of the unstable manifold (being equal to one), which in turn implies that the elliptical orbits around (F_1, V_1) possess the property of pushing away any point that is located in its neighborhood.

A financial interpretation of the above-mentioned phenomenon is that if the actual volatility of the market (measured by FCI_t) tends to zero and there is no time-dependent information transfer from FCI_t to VIX_t (i.e., $\alpha = 0$) then the option pricing mechanism becomes completely *speculative* and the only source of information for option pricing would be the contribution of the past information of options (i.e., measured by δ). In other words, for this scenario options would be priced without any relevance to the behavior of their underlying assets. However, as time evolves the implied volatility gradually becomes realized in the market by infusing β and θ amounts of information into FCI_t, which leads to the FCI_t being pushed away from its near-zero orbits. In turn, this increases the value of α from zero to a certain positive number. As a consequence, the dynamics of the information exchange between FCI_t and VIX_t gets restored to a great extent as the time evolves. Let us also stress that in case α takes on a non-zero value, the dynamical system (2.106) becomes governed by the manifolds of (F_4, V_4), and it exhibits a saddle point behavior in the neighbourhood of $(0^+, \frac{\delta}{\beta})$, i.e.,

$$\mathrm{Dim}(E^S) = 1,\ \mathrm{Dim}(E^U) = 1,\ \text{and}\ \mathrm{Dim}(E^C) = 0.$$

Further, the case where $\alpha > 0$ and $\delta = 0$ is also of interest. In this case, which also coincides with (F_3, V_3), the actual volatility of the market is nonzero (e.g., 0.006), whereas the implied volatility does equal zero. Similar to the previous case, the existence of a stable manifold of dimension 1 renders the trajectories of the critical point to travel along the elliptical orbits at a neighborhood of $(0.006, 0)$. However, the presence of an unstable manifold of dimension one, implies that the direction of such trajectories would be outwards and any point located sufficiently close to $(0.006, 0)$ will be pushed away from the critical point.

In context of the information provided above, from financial point of view the critical point $(0.006, 0)$ implies that the securities would be fully priced out based on their underlying assets, which eliminates any possibility for *arbitrage* in the volatility pricing. This means that the opportunity for trading a *delta neutral portfolio* of options and the underlying equities will not exist since the traders would be hindered from taking advantage of possible differences between the forecasted actual volatility of the options' underlying assets and the implied volatility of the options. Also, this scenario is of a transient nature, since the flow of information from FCI_t to VIX_t (i.e., $\alpha > 0$) will cause an increase in the self-entropy of VIX_t (i.e., δ), which in turn increases the magnitude of θ. This results in the growth of α and γ. Hence, the amount of information infused to VIX_t gets magnified as the time evolves, and the dynamical system (2.106) soon retain to a stable state. Further, note that upon increase of the values in δ, the dynamical system gets driven by the conditions of the manifolds governing the critical point (F_3, V_3), which might yield existence of a saddle point-type behavior about $(\frac{\alpha}{(\gamma+\theta)}, 0^+)$, i.e.,

$$\mathrm{Dim}(E^S) = 1,\ \mathrm{Dim}(E^U) = 1,\ \text{and}\ \mathrm{Dim}(E^C) = 0.$$

The final goal of our analysis of stability of the dynamical system (2.106) boils down to investigating the behavior of the dynamical system around the origin. For the latter of case $\alpha = \beta = 0$, the critical points (F_1, V_1) and (F_2, V_2) coincide, and are situated in the center of manifold (since $\mathrm{Dim}(E^C) = 2$). In such case, the direction of elliptical orbits existing at a neighborhood of $(0, 0)$ would become completely unpredictable. That is, the economic stagnation states implied by

$\mathrm{FCI}_t = \mathrm{VIX}_t = 0$ for both equity as well as option markets can either persist for a prolonged period of time, or alternatively the underlying market forces would push the direction of trajectories located in the vicinity of the critical point $(0, 0)$ outwards, so that the markets would regain their competitiveness states. Although the stock market has never undergone yet the above-described situation of the actual and implied types of volatility being both exactly equal to zero, theoretically it seems possible that the former case would govern the stock market at the origin. A substantiating evidence affirming the conclusion drawn above manifests itself when one considers small positive perturbations occurred in the values of α and β, which leads to

$$\mathrm{Dim}(E^S) = 2,\ \mathrm{Dim}(E^U) = 0,\ \text{and}\ \mathrm{Dim}(E^C) = 0.$$

In turn, this would make the origin a sink critical point, and hence it is going to be asymptotically stable. However, at present, such assertion does not appear to have any practical value, but this insight still has some theoretical interest.

2.5 Relation to Market Forces

In Section 2.4, we investigated causal relations that could potentially exist between FCI_t and VIX_t by employing the tools provided by the information theory alongside the orthogonalized impulse-response functions. However, in that section, the coupled system FCI_t-VIX_t was assumed to be an isolated one, i.e., it is not influenced by any exogenous factor. In reality, such an assumption is arguably far from truth, which is due to presence of numerous forces driving the stock market movements. In this section, we relax the constraint of isolation by expanding our considered system to include indexes that measure instability across various economy, government and financial sectors.

In the sequel, we will rely on the well-established *Granger causality* framework to study the direction of possible causal relations in the multivariate system of our time series (e.g., a multivariate time series consisting of FCI_t, VIX_t and other indexes which measure uncertainty in various

policies). In contrast to an approach based on the use of the transfer entropy which is framed in terms of *resolution of uncertainty*, the Granger causality framework is based on the statistical notion of *prediction* via vector autoregressive models. Hence, it provides a more convenient setting for performing statistical tests of inference on detected causal relations. Such a prediction-related property is particularly helpful when one aims at testing the market efficiency whose associated problems are commonly formulated in terms of the concept of *prediction*.

Additional sources of information that we take into account in this section revolve around the publicly available economic and financial information provided by a collection of particular newschapters. Two predominant daily indexes of such kind that we consider herein are *economic policy uncertainty* (EPU) developed in (Baker, Bloom, and Davis, 2016) and *equity market volatility* (EMV) tracker presented in (Baker, Bloom, Davis, and Kost, 2019). More specifically, the former index is constructed in a way that it reflects the frequency of articles in 10 leading U.S. newschapters that include the following trio of terms: *economy* or *economic*; *uncertain* or *uncertainty*; and one or more of the category of terms containing *Congress*, *deficit*, *Federal Reserve*, *legislation*, *regulation* and the *White House*. On the other hand, EMV is constructed by obtaining daily counts of articles that contain at least one term in the categories *economy* or *economic*; *uncertain* or *uncertainty*; and one or more terms from the category of *equity market*, *equity price*, *stock market* and *stock price*. It is noted that the U.S.-related articles used in the construction of EMV exceed 1000 newschapters and are retrieved from the *Access World News' NewsBank service*.

Besides, it is known that reactions to major shifts in various sectors of economy would usually manifest themselves in a slower pace than the daily time frequency. On this reason, we also employ monthly EPU and EMV indexes alongside a number of monthly category-specific uncertainty indexes for the United States which are constructed in a similar manner to EPU. The main difference consists in more restrictive criteria imposed on the terms that are related to economy, policy and uncertainty, e.g., see (Baker, Bloom, and Davis, 2016) for more detail on such list of the employed terms. The monthly uncertainty indexes considered in this work are: *monetary policy uncertainty* (MON), *fiscal policy uncertainty* (FIS), *tax policy uncertainty* (TAX), *government spending policy uncertainty* (GOV), *health care policy uncertainty* (HLTH), *national security pol-*

icy uncertainty (SEC), *entitlement programs policy uncertainty* (ENT), *general regulations policy uncertainty* (GREG), *financial regulations policy uncertainty* (FREG), *trade policy uncertainty* (TRD) as well as the *consumer price index uncertainty* (CPI).

2.5.1 Extended Granger Causality

Let the triple $(\Omega_t, \mathscr{F}_t, \mathbb{P})$ denote a *probability space*, where Ω_t comprises a set containing all market-relevant information *available* up to time t, and $\mathscr{F}_t$ represents a set of *known* information defined on Ω_t, which is further assumed to form a σ-algebra and to be a filtration process, i.e., $\mathscr{F}_1 \subset \mathscr{F}_2 \subset \cdots \subset \mathscr{F}_{|\mathcal{T}|}$. Note that the existing dichotomy used for distinguishing the modes of information is mainly because of the following issue: the set Ω_t is referred to as the available information set since, it contains all the possible events generated as the outcome of interplays among different financial, economic, psychological and political forces underlying the market. On the other hand, the set $\mathscr{F}_t$ is confined to only σ-algebra subsets of Ω_t. Hence, the information generated by $\mathscr{F}_t$ cover smaller amounts of the available information that are known up to time t, rendering its appellation.

Further, let us assume that there is a finite set of random vectors $\{\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_t\}$ which constitute possible multivariate distributions related to the discrete-time stochastic processes $\{\mathbf{Y}(\omega, t) : t = 1, 2, \ldots\}$ such that the K-vector process $\mathbf{Y}_t : \Omega_t^K \to \mathcal{Y}_t \subset \mathbb{R}_{\geq 0}^K$ is a *measurable map* from $(\Omega_t^K, \bigotimes_{k=1}^K \mathscr{F}_{k,t})$ to $(\mathcal{Y}_t, \mathscr{B}_{\mathcal{Y}_t})$ defined by

$$\mathbf{Y}_t^{-1}(B) \equiv \{\omega : \mathbf{Y}(\omega, t) \in B\} \in \bigotimes_{k=1}^{K} \mathscr{F}_{k,t}, \qquad \forall B \in \mathscr{B}_{\mathcal{Y}_t}.$$

Here, K denotes the totality of different time series contained in $\mathbf{Y}_t$, the term $\bigotimes_{k=1}^K \mathscr{F}_{k,t}$ represents the Cartesian product among K involved σ-algebras, and $\mathscr{B}_{\mathcal{Y}_t}$ stands for the smallest σ-algebra containing all open sets in the space $\mathcal{Y}_t \subset \mathbb{R}_{\geq 0}^K$. Note in passing that the Cartesian product of two σ-algebras is not necessarily a σ-algebra. Nevertheless, let $\mathcal{J}_t = \bigotimes_k \mathscr{F}_{k,t}$ denote the Hilbert space spanned by the components $\mathscr{F}_{k,t}$. Then, $\mathcal{J}_t$ could be considered as the set of the known information

reflected in the past and present of $\mathbf{Y}_{k,t}$, $k = 1, 2, \ldots, K$, which are common to all time periods. For two Hilbert spaces $\mathcal{J}_t$ and $\mathcal{J}_t^{'}$, we represent the space spanned by all their components by $\mathcal{J}_t \bigoplus \mathcal{J}_t^{'}$ where $\bigoplus$ denotes an *orthogonal direct sum operator*.

Subsequently, in order to identify the causal relations potentially existing among the time series contained in $\mathbf{Y}_t$, we introduce the notion of *Granger causality* (GC) and consider its applications. Essentially, the Granger causality is based on the following underlying assumptions: 1- cause precedes effect, thus the future values of a time series cannot cause its past values; 2- cause provides *unique* information about the effect which is not available elsewhere in any other variable. The time series $\mathbf{Y}_{i,t}$ is then said *not* to Granger-cause $\mathbf{Y}_{j,t}$ $(i, j = 1, 2, \ldots, K; i \neq j)$ w.r.t. the information set $\mathcal{J}_t$ if and only if the future values of $\mathbf{Y}_{j,t}$ predicted on the basis of the past and current information contained in $\mathcal{J}_t$, do not yield substantial improvement in the accuracy of prediction as compared to the situation in which $\mathbf{Y}_{j,t}$ would have been predicted based on the information in $\mathcal{J}_t \bigominus \mathscr{F}_{i,t}$. Here, $\mathcal{J}_t \bigominus \mathscr{F}_{i,t}$ denotes a shorthand for representing all the information contained in $\mathcal{J}_t$ except for those pertaining to $\mathbf{Y}_{i,t}$.

More formally, the Granger causality can be defined as follows:

Definition 2.5.1 (Granger Causality)**.** *Given a K-vector process $\mathbf{Y}_t$ and its known information set $\mathcal{J}_t$, we define different modes of the Granger causality as follows:*

1. *(GC up to time-horizon $h > 0$):*
   $\mathbf{Y}_{i,t} \overset{(\mathrm{h})}{\nrightarrow} \mathbf{Y}_{j,t} \Leftrightarrow \forall k \in \{1, 2, \ldots, h\} : \mathbb{E}\left[\mathbf{Y}_{j,t+k} | \mathcal{J}_t\right] = \mathbb{E}\left[\mathbf{Y}_{j,t+k} | \mathcal{J}_t \bigominus \mathscr{F}_{i,t}\right]$,

2. *(GC at any time-horizon $h \geq 0$):*
   $\mathbf{Y}_{i,t} \overset{(\infty)}{\nrightarrow} \mathbf{Y}_{j,t} \Leftrightarrow \forall h \geq 0 : \mathbb{E}\left[\mathbf{Y}_{j,t+h} | \mathcal{J}_t\right] = \mathbb{E}\left[\mathbf{Y}_{j,t+h} | \mathcal{J}_t \bigominus \mathscr{F}_{i,t}\right]$,

3. *(GC at time-horizon $h \geq 0$):*
   $\mathbf{Y}_{i,t} \overset{\mathrm{h}}{\nrightarrow} \mathbf{Y}_{j,t} \Leftrightarrow \mathbb{E}\left[\mathbf{Y}_{j,t+h} | \mathcal{J}_t\right] = \mathbb{E}\left[\mathbf{Y}_{j,t+h} | \mathcal{J}_t \bigominus \mathscr{F}_{i,t}\right]$.

It is worth mentioning that in the case where the time-horizon parameter h present in Definition 2.5.1 is strictly positive, it indicates presence of a *time lag* upon conveyance of information

from a particular cause to its effect. On the contrary, the case of $h = 0$ corresponds to *instantaneous* or *zero-lag* effects existing among various time series in $\mathbf{Y}_t$. Instantaneous effect can be seen as a measure for the common information between $\mathbf{Y}_{i,t}$ and $\mathbf{Y}_{j,t}$ that are not shared with their past.

Even though the literature on the analysis of Granger causal patterns under presence of lagged effects only is vast, there are only a very few works which consider impacts of the instantaneous relations. Such a consideration is important, since the lagged causal relations are likely to be misdetermined if the analysis disregards presence of possible instantaneous effects, e.g., see (Faes and Sameshima, 2014) for a detailed discussion. In the sequel, we adhere to the work presented in (Schiatti, Nollo, Rossato, and Faes, 2015) whose authors formulated an extension of the Granger causality and provided a framework to analyze strictly causal ($h > 0$) and instantaneous ($h = 0$) causal relations concurrently; henceforth, the notation eGC will be used to denote the extended Granger causality mentioned above.

On the basis of an approach introduced in (Schiatti, Nollo, Rossato, and Faes, 2015), initially a VAR model of order p is considered which has the following form:

$$\mathbf{Y}_t = \sum_{k=0}^{p} \mathbf{B}_t \mathbf{Y}_{t-k} + \boldsymbol{\mathcal{E}}_t, \tag{2.108}$$

where matrix $\mathbf{B}_t \in \mathbb{R}^{K \times K}$ and vector $\boldsymbol{\mathcal{E}}_t$ represent the regression coefficients and model residuals, respectively. The causal relation from $\mathbf{Y}_{i,t}$ to $\mathbf{Y}_{j,t}$ is then determined by first computing the so-called *unrestricted* model

$$\mathbf{Y}_{j,t} = \sum_{k=0}^{p} \mathbf{B}_{j,t} \mathbf{Y}_{t-k} + \boldsymbol{\mathcal{E}}_{j,t}, \tag{2.109}$$

and then solving the associated *restrictive* regression model

$$\mathbf{Y}_{j,t} = \sum_{k=0}^{p} \mathbf{B}'_{j,t} \mathbf{Y}^{(i)}_{t-k} + \boldsymbol{\mathcal{E}}'_{j,t}. \tag{2.110}$$

Note that, in model given by equation (2.109), $\mathbf{B}_{j,t}$ corresponds to the j-th row of $\mathbf{B}_t$ which represents a part of the model given by equation (2.108) related to the j-th time series in $\mathbf{Y}_t$. Further-

more, the model given by equation (2.110) describes $\mathbf{Y}_{j,t}$ w.r.t. $\mathcal{J}_t \bigominus \mathscr{F}_{i,t}$, where $\mathbf{Y}_t^{(i)}$ is obtained by excluding the i-th time series in $\mathbf{Y}_t$.

An important measure for quantifying the error reduction when predicting $\mathbf{Y}_{j,t}$ w.r.t. $\mathcal{J}_t$ (as compared to the case of predicting it w.r.t. $\mathcal{J}_t \bigominus \mathscr{F}_{i,t}$) is defined by the following *log-likelihood ratio*:

$$\mathrm{eGC}_{i\to j} = \log \frac{|\Sigma_j|}{|\Sigma'_j|}, \tag{2.111}$$

which serves as the test statistic of extended Granger causality, where $\Sigma_j = \mathrm{cov}(\boldsymbol{\mathcal{E}}_{j,t})$ and $\Sigma'_j = \mathrm{cov}(\boldsymbol{\mathcal{E}}'_{j,t})$. The *extended GC measure* defined above is always positive, and it could be further interpreted from information-theoretic perspective as the extent to which the information conveyed from $\mathbf{Y}_{i,t}$ to $\mathbf{Y}_{j,t}$ helps reduce the ambiguity in $\mathbf{Y}_{j,t}$. Moreover, the transfer entropy introduced in Section 2.3 would reduce to Granger causality when the prediction functions are linear and developed based on the VAR models. It has already been shown that the GC measure formulated based on the strictly causal ($h > 0$) counterparts of equation (2.108), equation (2.109) and equation (2.110) for a VAR model with i.i.d. multivariate Gaussian residuals would asymptotically be equivalent to the transfer entropy, e.g., see (Barnett and Bossomaier, 2012).

2.5.2 Analysis of Daily Data

Figure 2.18 presents a diagram of the extended Granger causal relations existing among various daily time series contained in the vector

$$\mathbf{Y}_t = (\mathrm{FCI}_t, \mathrm{VIX}_t, \mathrm{EMV}_t, \mathrm{EPU}_t)^\intercal .$$

In view of this graph, it is interesting to stress existence of the instantaneous circular triad among FCI_t, VIX_t and EMV_t. This observation reconfirms the results we obtained previously in Section 2.4 regarding the direction of causal relations between FCI_t and VIX_t, where it was concluded that VIX_t would exert its influence on FCI_t instantaneously, while it was deemed that FCI_t would cause VIX_t in a time-lagged manner.

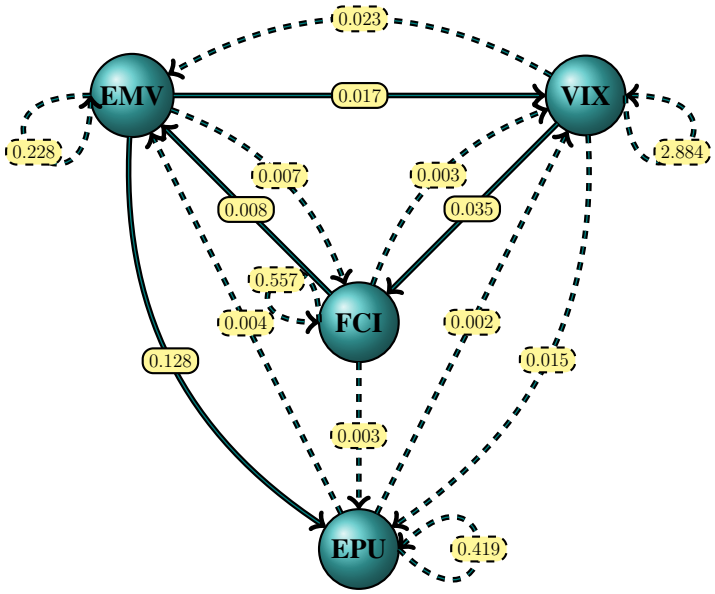

Figure 2.18: Diagram of extended Granger causal relations among daily time series at 1% level of significance during January 1990-December 2019. Instantaneous and lagged relations are indicated using solid and dashed lines, respectively. The amount of extended GC measure (2.111) is also presented for each relation within a box.

In order to investigate whether the daily news-based uncertainty indexes Granger cause FCI_t by strictly (time-lagged) causal relations, we formulate the following multiple null hypothesis:

$$\mathrm{H}_{0,\mathtt{D}}^{\mathtt{NEWS}} = \left(\mathrm{EMV}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t\right) \bigcap \left(\mathrm{EPU}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t\right).$$

In addition, the p-values associated to $\mathrm{EMV}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$ and $\mathrm{EPU}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$ are computed as being $p_1 = 0.000$ and $p_2 = 0.036$ based on the average of 500 replications, respectively. Subsequently, in order to test $\mathrm{H}_{0,\mathtt{D}}^{\mathtt{NEWS}}$ for the daily time frequency, note that from the category of news-based indexes only EMV_t significantly Granger causes the FCI_t in a strict manner. Thus, ***Bonferroni's test***, e.g., see (Bland and Altman, 1995), appears to be an appropriate method for examining the multiple null hypothesis $\mathrm{H}_{0,\mathtt{D}}^{\mathtt{NEWS}}$, as Bonferroni's approach is often considered to be powerful when there are very few large effects in the system. To this end, let $\mathcal{J}_{t,\mathtt{D}}^{\mathtt{PRC}} = \mathscr{F}_{1,t}$ and $\mathcal{J}_{t,\mathtt{D}}^{\mathtt{NEWS}} = \mathscr{F}_{3,t} \bigotimes \mathscr{F}_{4,t}$ denote the set of known information for assets' daily prices and news-based uncertainties, respectively. Note that $\mathcal{J}_{t,\mathtt{D}}^{\mathtt{PRC}}$ is constructed separately from $\mathcal{J}_{t,\mathtt{D}}^{\mathtt{NEWS}}$ since the price-related information does not share a common probability space with news-based information. The Bonferroni's method would then reject $\mathrm{H}_{0,\mathtt{D}}^{\mathtt{NEWS}}$ at level of significance α if

$$\min\{p_1, p_2\} \leq \frac{\alpha}{2}. \tag{2.112}$$

Once the inequality (2.112) is examined for various values of α, it becomes clear that $\mathrm{H}_{0,\mathtt{D}}^{\mathtt{NEWS}}$ gets rejected at significance level 1% w.r.t. the information set $\mathcal{J}_{t,\mathtt{D}}^{\mathtt{PRC}} \bigoplus \mathcal{J}_{t,\mathtt{D}}^{\mathtt{NEWS}}$, implying that the mutual daily price changes of assets are time-lagged Granger caused by news-based uncertainty indexes.

It is worth mentioning that even though VIX_t was not involved in any of the testing procedures described above, its inclusion in $\mathbf{Y}_t$ is a necessary step when conducting Granger causality analysis in order to avoid possible inference of spurious causal relations since the VIX_t was shown to be cointegrated with FCI_t, e.g., see Section 2.4.2 for more details.

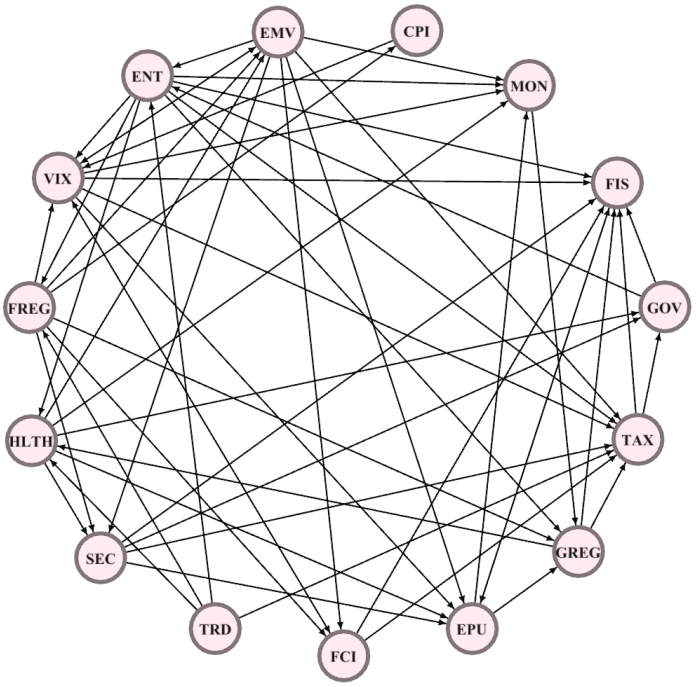

Figure 2.19: Diagram of extended Granger causal relations among monthly time series in at 1% level of significance during January 1990-December 2019. Some Self-loops were omitted for clarity; the remaining self-loops pertain to FCI_t, EMV_t, EPU_t, MON_t, HLTH_t, SEC_t, GREG_t, FREG_t, TRD_t and CPI_t.

2.5.3 Analysis of Monthly Data

Figure 2.19 depicts the extended Granger causal relations for the time series contained in $\mathbf{Z}_t$ comprised of the monthly time series, i.e.,

$$\mathbf{Z}_t = (\mathrm{FCI}_t, \mathrm{VIX}_t, \mathrm{EMV}_t, \mathrm{EPU}_t, \mathrm{MON}_t, \mathrm{FIS}_t, \mathrm{TAX}_t, \mathrm{GOV}_t, \mathrm{HLTH}_t, \mathrm{SEC}_t, \mathrm{ENT}_t, \mathrm{GREG}_t, \mathrm{FREG}_t, \mathrm{TRD}_t, \mathrm{CPI}_t)^{\intercal}.$$

Furthermore, Figures 2.20 and 2.21 portray heatmaps of the empirical probability measures for instantaneous causal relations as well as the extended Granger causal measures among monthly time series contained in $\mathbf{Z}_t$, respectively.

In view of Figures 2.19 to 2.21, one can observe a rather different set of causal relations existing among monthly FCI_t, VIX_t, EMV_t and EPU_t as compared to their daily counterparts. Specifically, the direction of instantaneous and strict causal relations between FCI_t and VIX_t are reversed in the

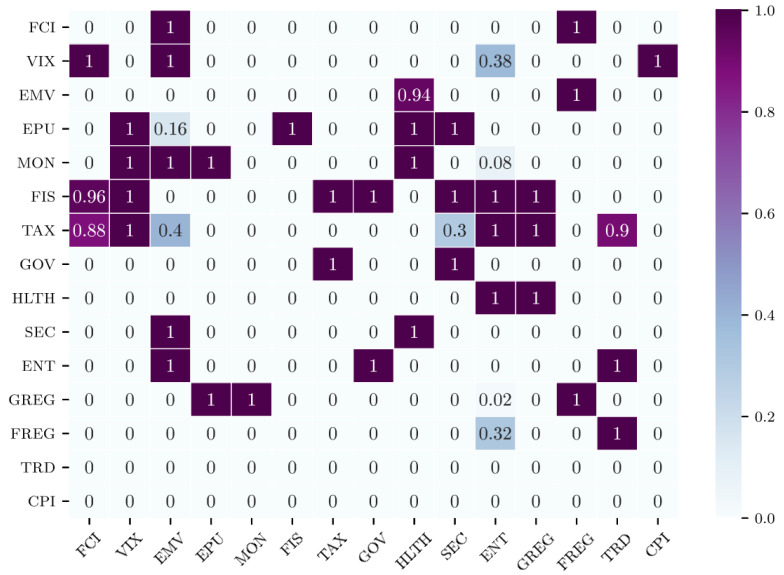

Figure 2.20: Heatmap of empirical probability measures for instantaneous causal relations among monthly time series at 1% level of significance during January 1990-December 2019. The (i, j)-element represents the probability that the j-th time series Granger causes the i-th time series instantaneously.

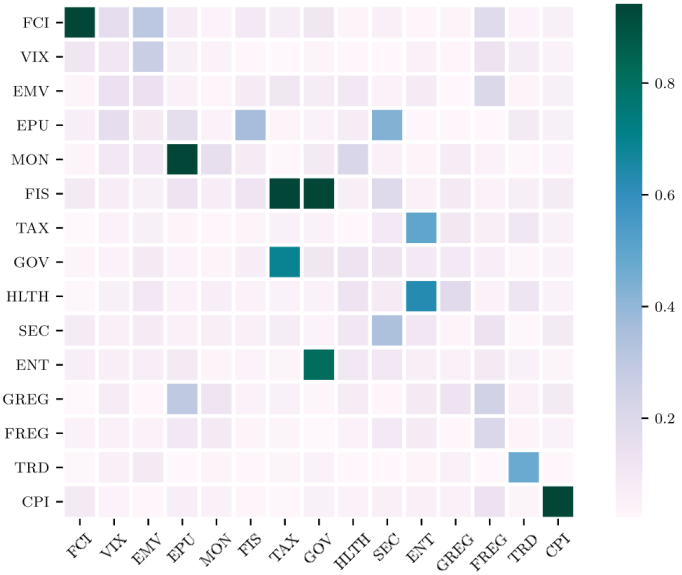

Figure 2.21: Heatmap of extended Granger causal measures among monthly time series during January 1990-December 2019. The (i, j)-element represents the measure from component j to i, i.e., $\mathrm{eGC}_{j \to i}$.

case of monthly time series, where FCI_t can be seen to exert its impact on VIX_t instantaneously, whereas VIX_t would strictly Granger cause FCI_t. Also, on the basis of these figures, it is seen that EMV_t would instantaneously Granger cause all three monthly time series FCI_t, VIX_t and EPU_t.

We also find it remarkable to note other important causal relations revealed by Figures 2.19 to 2.21. For instance, on the one hand, the health care policy uncertainty is seen to have been Granger caused by the trade, general regulations and entitlement programs policy uncertainties. On the other hand, it is rather surprising to find out that the health care policy uncertainty has Granger caused the security, economic, government spending and monetary policy uncertainties as well as the equity market volatility tracker during the considered period of time. These results are especially tangible when one takes into account the case of the COVID-19 and how accurately the causal relations inferred by analyzing the data pertaining to January 1990-December 2019 extrapolated to the circumstances experienced particularly in the U.S. during 2020. These information per se can potentially be valuable for analysts and decision makers in the areas related to public health care policies. Another observation, potentially being of interest for the government decision makers, pertains to the causal relations related to the fiscal policy uncertainty. More specifically, it is seen that the fiscal policy uncertainty has been Granger caused by several important monthly indexes including the security, economic, general regulations, tax, government spending and entitlement programs policy uncertainties as well as the FCI and VIX.

The network of the extended Granger causal relations depicted in Figure 2.19 can further be analyzed more rigorously from graph-theoretical point of view. In particular, the associated *network diameter* is computed to be 7 which is an indicator of the number of steps that are required for the two most distant nodes in the network to reach one another, see (Brandes, 2001) for more details. The *average path length* is further computed to be 2.326 which signifies a relatively efficient network for information flow. We note that the average path length provides a statistic of communication efficiency for the entire graph, which is measured as the shortest possible path between all nodes in the network and is bounded by the value of the network diameter. Also, the *graph density*, being an statistic of the level of connected edges within the network relative to the total possible value, is computed to be 0.314, which implies that the analyzed network has a sparse

structure rather than a dense one, where about 31% of nodes are connected with one another.

Table 2.10 provides further statistics computed for the network depicted in Figure 2.19. As per results reported in this table, the *hyperlink-induced topic search* (HITS) approach is used to compute the *authority* and *hub* statistics for each node. Basically, the former statistic measures how valuable the information stored by a particular node is, while the latter one quantifies the quality of the links to and from that particular node, e.g., see (Kleinberg, 1999) for more details. We note that higher authority scores are indicative of nodes that are linked by many hubs (in-degree connections), while the hub values are obtained by calculating the number of outbound links (out-degree connections). On the basis of this table, it is seen that the highest authority scores are associated with FIS and TAX nodes, whereas VIX, EMV and ENT nodes attain the highest hub values. The most authoritative nodes of the network can alternatively be identified by computing the *page rank proximity* statistic, e.g., see (Brin and Page, 1998) for more details. By resorting to the results provided by Table 2.10, the FIS and TAX are identified as the nodes attaining the highest values for the page rank proximity, which are aligned with the results obtained by using HITS algorithm.

Moreover, Table 2.10 reports the *betweenness centrality* computed for each node of the analyzed network. This statistic plays a pivotal role in understanding how information influence and flow within the network, by measuring how important an individual node is for the other nodes when the network is traversed. In other words, the betweenness centrality indicates how often a specific node lies on the shortest path between two other nodes, e.g., see (Barrat, Barthelemy, Pastor-Satorras, and Vespignani, 2004) for more details. Based on the results obtained, it is evident that ENT node attains the highest value of betweenness centrality among other nodes which are contained in the network depicted in Figure 2.19. In turn, this implies the importance of the entitlement programs policy uncertainty index among the considered monthly indexes. More precisely, we can observe that the entitlement programs policy uncertainty index Granger causes financial and general regulations, health care, tax, fiscal and monetary policy uncertainties as well as the VIX, while it is Granger caused by the trade and government spending policy uncertainties as well as the EMV.

Lastly, the *bridging coefficient* is computed for each node of the network using the method elucidated in (Hwang, Kim, Ramanathan, and Zhang, 2008), whose results are further presented in Table 2.10. In a nutshell, the bridging coefficient is defined as the average probability of leaving the direct neighbor subgraph of a specific node. Specifically, nodes with relatively higher values of bridging coefficient signify potential bottlenecks to information flow within the network. On the basis of Table 2.10, it is seen that CPI and GOV nodes attain the highest values of bridging coefficient among the other nodes. In turn, it is implied that the circulated information within the network leave the direct neighbor subgraphs of these two nodes with high probabilities as compared to the rest of the nodes present within the network.

Besides, in order to test whether news-based uncertainty indexes Granger cause FCI_t by their strictly causal relations, we formulate the following set of null hypotheses:

$$\mathrm{H}_{0,k}^{\texttt{NEWS}} : \mathbf{Z}_{k,t} \overset{h>0}{\nrightarrow} \mathrm{FCI}_t, \quad \forall k \in \{3, 4, \ldots, K\}.$$

Let $\mathcal{J}_{t,\texttt{M}}^{\texttt{PRC}} = \mathscr{F}_{1,t}$ and $\mathcal{J}_{t,\texttt{M}}^{\texttt{NEWS}} = \bigotimes_{k=3}^{K} \mathscr{F}_{k,t}$ $(K = 15)$ denote the set of known assets' monthly mutual prices changes and news-based policy uncertainty information, respectively. The null hypothesis of Granger causality for the monthly news-based information can then be expressed in terms of the following multiple null hypothesis :

$$\mathrm{H}_{0,\texttt{M}}^{\texttt{NEWS}} = \bigcap_{k=3}^{K} \mathrm{H}_{0,k}^{\texttt{NEWS}}.$$

Table 2.11 reports the p-values of different news-based policy uncertainty indexes which are not strictly Granger causing FCI_t. As per the p-values reported in this table, one can observe numerous small effects existing in the news-based indexes not Granger causing FCI_t. Thus, ***Fisher's combination test*** presented in (Fisher, 1992) would deem appropriate to be used in this case for testing $\mathrm{H}_0^{\texttt{NEWS}}$. It is worth noting that Fisher's combination test has almost the full power, i.e., its type-II error is high when the global null hypothesis is false. Further, Fisher's combination test is based on the fact that if p-values are uniformly distributed in $(0, 1]$, then the negative of their

Table 2.10: Statistics computed for the network of monthly extended Granger causal relations.

Node	Authority	Hub	Page Rank Proximity	Betweenness Centrality	Bridging Coefficient
FCI	0.229	0.255	0.368	0.003	0.208
VIX	0.300	0.422	0.424	0.128	0.084
EMV	0.193	0.426	0.412	0.156	0.079
EPU	0.304	0.211	0.583	0.109	0.123
MON	0.368	0.041	0.583	0.017	0.176
FIS	0.439	0.064	0.700	0.029	0.087
TAX	0.474	0.121	0.636	0.071	0.103
GOV	0.138	0.125	0.466	0.147	0.389
HLTH	0.183	0.251	0.424	0.191	0.091
SEC	0.194	0.285	0.412	0.018	0.124
ENT	0.158	0.440	0.400	0.246	0.085
GREG	0.197	0.230	0.519	0.198	0.133
FREG	0.134	0.245	0.304	0.114	0.088
TRD	0.000	0.199	0.000	0.000	0.288
CPI	0.051	0.063	0.250	0.000	0.567

Table 2.11: Computed p-values for strictly causal relations among monthly time series.

Relation	p-value	Relation	p-value	Relation	p-value
$\mathrm{EPU}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$	0.081	$\mathrm{GOV}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$	0.012	$\mathrm{GREG}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$	0.706
$\mathrm{MON}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$	0.564	$\mathrm{HLTH}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$	0.681	$\mathrm{TRD}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$	0.528
$\mathrm{FIS}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$	0.034	$\mathrm{SEC}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$	0.239	$\mathrm{CPI}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$	0.229
$\mathrm{TAX}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$	0.114	$\mathrm{ENT}_t \overset{h>0}{\nrightarrow} \mathrm{FCI}_t$	0.788		

logarithms all follow the standard exponential distribution, i.e.,

$$-\log p_k \sim \mathrm{Exp}(1), \quad k = 3, 4, \ldots, K. \tag{2.113}$$

Subsequently, if we assume that all the p-values are independent, relation (2.113) suggests formulating the following test statistic for the Fisher's combination test:

$$T_F = -2 \sum_{k=3}^{K} p_k, \tag{2.114}$$

which has a χ^2 distribution with $2\left(K - 2 - \sum_{k=3}^{K} \mathbb{I}\left(\mathbb{P}_n\left[\mathbf{Z}_{k,t} \overset{h=0}{\nrightarrow} \mathrm{FCI}_t\right] = 1\right)\right)$ degrees of freedom under the global null hypothesis, where $\mathbb{P}_n[\cdot]$ denotes an *empirical probability measure*. For the computed p-values reported in Table 2.11, the test statistic equals $T_F = 35.214$. Hence, the p-value of the χ^2_{22} distribution for $\mathrm{H}_0^{\mathtt{NEWS}}$ is approximately 0.037. Finally, we conclude that there is insufficient evidence to reject at 1% level of significance the global null hypothesis $\mathrm{H}_0^{\mathtt{NEWS}}$ stating that the news-based policy uncertainty indexes do not strictly Granger cause FCI_t w.r.t. the information set $\mathcal{J}_{t,\mathtt{M}}^{\mathtt{PRC}} \bigoplus \mathcal{J}_{t,\mathtt{M}}^{\mathtt{NEWS}}$. In turn, this implies that the mutual price changes of the assets are not time-lagged Granger caused by news-based uncertainty indexes.

2.6 Relation to Market Fundamentals

This section of the chapter is devoted to study the changes appearing in assets' prices and the relations that could possibly exist between them and their associated fundamental features. Taking into account the information pertaining to assets' fundamentals, is a key step to potentially comprise the set of all known information about assets which is to be publicly accessible. Such set of all known information would be comprised of assets' fundamentals information augmented by the introduced newschapter-based uncertainty information as well as the assets' historical and current prices. Assuming that these three sources of information are the ones accessible to public, in turn enables us to design a set of statistical hypothesis tests to examine the efficient market hypothesis in all its different modes. The importance of our work on efficient market hypothesis is that it provides a guideline for traders and portfolio managers who aim at allocating their investments at optimal positions. Were it to be shown that the market is inefficient, whether on daily, monthly or quarterly frequencies, then this announces the opportunity for the investors to take advantage of the historical and current prices of assets to discover patterns therein which can provide them with a forecast of the future values of the assets. On the contrary, if the market is shown to be efficient, other set of investment strategies should potentially be employed which do not rely on prediction-based approaches of asset prices.

2.6.1 Feature Selection

Let $F = \{f_1, f_2, \dots\}$ denote a space containing all the fundamental features available for the assets contained in S. Our goal is then to select a certain subset $F^\star \subset F$ whose included fundamental features would best distinguish the assets from one another. Our approach towards feature selection is rather particular, since it relates the *quantitative* fundamental features to the *subjective* stock market industry taxonomies that have been developed and proven effective to classify the stocks into different categories. To be more precise, we select the fundamentals $F^\star$ out of all possible subsets of F which provide the so-called *best-match* (in a sense that we will describe in the

proceeding) for the number of classes given by our specified industry taxonomy.

To begin with, industry classification schemes are deemed indispensable elements of the financial and economic analysis which facilitate measuring economic activities and constructing fame sector indexes. There exist a number of different taxonomies, each being designed on the basis of specific sets of goals, scopes and orientations. Recently, a review of the definitions and construction methods for the taxonomies used worldwide was presented in (Phillips and Ormsby, 2016). One of the popular taxonomies used in North America is the *global industry classification standard* (GICS), based on which S&P500 components are classified into 11 *sectors*, 24 *industry groups*, 69 *industries* and 158 *sub-industries*. Among these categories, the former one containing sectors' information plays by far a crucial role in many of the financial and economic decisions made concerning stock markets. For instance, according to GICS, the main sectors of the stock markets are: 1- *energy*, 2- *materials*, 3- *industrials*, 4- *consumer discretionary*, 5- *consumer staples*, 6- *health care*, 7- *financials*, 8- *information technology*, 9- *telecommunication services*, 10- *utilities* and 11- *real estates*. Nevertheless, what concerns our feature selection approach pertains to the number of S&P500 components classified into the sub-industry category. Namely, we implement a clustering algorithm to process our constructed *design matrix*, i.e.,

$$\mathbf{D} = \begin{bmatrix} \mathbf{d}_1^\intercal \\ \mathbf{d}_2^\intercal \\ \vdots \\ \mathbf{d}_N^\intercal \end{bmatrix} = \begin{matrix} \\ s_1 \\ s_2 \\ \vdots \\ s_N \end{matrix} \begin{matrix} \begin{matrix} f_1^\star & f_2^\star & \cdots & f_{|F^\star|}^\star \end{matrix} \\ \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1|F^\star|} \\ d_{21} & d_{22} & \dots & d_{2|F^\star|} \\ \vdots & \vdots & \vdots & \vdots \\ d_{N1} & d_{N2} & \dots & d_{N|F^\star|} \end{bmatrix} \end{matrix}.$$

Then, we compare the number of clusters obtained by executing the clustering algorithm to the number of sub-industries provided by GICS scheme. In the case where these two quantities match closely enough (or alternatively called the best-match), the subset $F^\star$ of the feature space would be proclaimed optimal. However, there is one obstacle lying ahead of this approach, since the number of clusters is commonly given *a priori* in a vast majority of the clustering algorithms, which defeats our goal of approximating the optimal value for the number of classes when we intend to cluster

the design matrix $\mathbf{D}$.

A promising approach to surmount this obstacle is to resort to the use of the *affinity propagation* algorithm. This algorithm which was first introduced in (Frey and Dueck, 2007), eliminates the need for specifying the number of clusters before running the algorithm. Also, unlike other clustering algorithms, affinity propagation deals with the clustering problem through representing the clusters by the so-called *exemplar* data points. More precisely, the goal of executing such algorithm would be to identify a subset of the N sample points from the set of vectors $\mathbf{d}_1^\intercal, \mathbf{d}_2^\intercal, \ldots, \mathbf{d}_N^\intercal$ as exemplars and then assign the remaining sample points to one of those exemplars. One advantage of affinity propagation algorithm, which we would like to highlight is that all the data points are considered as potential exemplars simultaneously, see (Dueck, 2009) for more technical details.

Aside from the success achieved by using the affinity propagation algorithm for clustering the data points contained in the design matrix $\mathbf{D}$, there remains another impediment which concerns the construction mechanism of the matrix $\mathbf{D}$ itself. That is, the temporal dimension of features for every asset is disregarded when the values for features are represented by a set of single quantities. It is needless to say that such an impediment can be easily mitigated if the data are represented using a *design tensor* rather than a matrix. However, our rationale for the use of matrices is that the configuration of stocks classified on the basis of the industry taxonomies such as GICS scheme are usually not prolonged. Hence, the results obtained by running the affinity propagation algorithm would potentially match the number of elements in the sub-industry category only if a relatively short time span is used for the analysis. A common sense concerning stock markets affirms that the configuration of stocks in industry taxonomies continually undergoes major changes roughly every 4-5 years, which is mainly due to ever-changing technological apparatuses that are being invented which influence assets' infrastructures.

On this reason, we construct our design matrix $\mathbf{D}$ by taking average of fundamentals' values over 4 years of the time span, i.e., from the first quarter of 2016 to the fourth quarter of 2019[6]. Then, the affinity propagation algorithm is executed to cluster N assets for different design matrices

[6]Data were retrieved from Compustat server provided by Wharton Research at the University of Pennsylvania (WRDS).

Table 2.12: Table of selected fundamental features for the market.

Feature	Feature
Acquisitions	Inventories - Total
Assets - Other - Total	Investing Activities - Net Cash Flow
Assets and Liabilities	Investing Activities - Other
Capital Expenditures	Long-Term Debt - Total
Cash and Short-Term Investments	Long-Term Debt - Issuance
Cash Dividends	Liabilities Netting Other Adjustments
Common/Ordinary Equity - Total	Noncontrolling Interest
Comprehensive Income - Total	Non-Operating Income (Expense) - Total
Cost of Goods Sold	Operating Activities - Net Cash Flow
Debt in Current Liabilities	Operating Income Before Depreciation
Earnings Per Share (Basic)	Preferred/Preference Stock (Capital) - Total
Earnings Per Share (Diluted)	Pretax Income
Excess Tax Benefit of Stock Options	Receivables - Total
Extraordinary Items	Sale of Common and Preferred Stock
Financing Activities - Net Cash Flow	Sale of Investments
Funds from Operations - Other	Sale of PPE and Investments - Gain/Loss
Quarterly Income Before Extraordinary Items	Sales/Turnover (Net)
Annual Income Before Extraordinary Items	Short-Term Investments- Total
Income Taxes	Special Items
Intangible Assets - Total	Stockholders Equity

constructed using various subsets of the available fundamental features. Thereafter, our results obtained by running the clustering algorithm are compared to the number of elements in GICS sub-industry category, and a subset $F^\star$ is identified as optimal if it has cardinality close to the number of sub-industries, namely 158. Note that the total number of fundamentals considered in our experiments exceeded 600 quarterly and annual items, and the optimal $F^\star$ that we determined by conducting extensive experiments contained 40 fundamentals (see Table 2.12 for the list of the selected features). It is also worth mentioning that running the affinity propagation algorithm for $F^\star$ resulted in 156 categories, which is slightly less than the desired number 158. Hence, we may assert attainment of the so-called best-match and proclaim the determined $F^\star$ to be an optimal set of features w.r.t. the above-described criterion.

2.6.2 Market Fundamentals Index

In the proceeding discussions, we will demonstrate a mechanism to construct an index for the stock market fundamentals which can capture their mutual overall behavior. Next, we will show that our developed *market fundamentals index* (MFI) is related to FCI, VIX and some other policy uncertainty indexes.

Let us denote by $\boldsymbol{\mathcal{D}} \subset \mathbb{R}^{N \times |F^\star| \times |\mathcal{T}|}$ a *design tensor* accommodating the values of the selected fundamental features for each asset throughout time (e.g., throughout $|\mathcal{T}| = 120$ quarters), whose schematic is depicted in Figure 2.22. In order to derive the so-called market fundamentals index $\mathrm{MFI} : \mathbb{R}^{N \times |F^\star| \times |\mathcal{T}|} \to \mathbb{R}_{\geq 0}$, it appears to be necessary to characterize the complexity of the design tensor $\boldsymbol{\mathcal{D}}$, so that an appropriate set of tools could potentially be identified and utilized for formulating the index. As a preliminary step, we perform a polyadic decomposition on $\boldsymbol{\mathcal{D}}$ to estimate its rank R. Namely,

$$\boldsymbol{\mathcal{D}} \approx \widehat{\boldsymbol{\mathcal{D}}} = \sum_{r=1}^{R} \mathbf{s}_r \circ \mathbf{f}_r \circ \mathbf{t}_r, \tag{2.115}$$

where $\mathbf{s} \in \mathbb{R}^N$, $\mathbf{f} \in \mathbb{R}^{|F^\star|}$ and $\mathbf{t} \in \mathbb{R}^{|\mathcal{T}|}$. Figure 2.23 exhibits the relative error of polyadic decomposition for the approximated tensor $\widehat{\boldsymbol{\mathcal{D}}}$, which reveals that the elements (atoms) of polyadic decomposition do not successfully capture the underlying structure of the design tensor $\boldsymbol{\mathcal{D}}$, even when the value for rank is considered to be as large as 100. In turn, this renders the structure of $\boldsymbol{\mathcal{D}}$ to be a more complex nonlinear one.

On this reason, it is more suitable to decompose the tensor $\boldsymbol{\mathcal{D}}$ using a method which captures its embedded nonlinear patterns. One possible choice of the method from those employed in our work is to approximate $\boldsymbol{\mathcal{D}}$ by using the Tucker decomposition. The Tucker decomposition approximates $\boldsymbol{\mathcal{D}}$ with a small-size core tensor $\boldsymbol{\mathcal{H}} \in \mathbb{R}^{R_1 \times R_2 \times R_3}$ and three factor matrices $\mathbf{S} \in \mathbb{R}^{N \times R_1}$, $\mathbf{F} \in \mathbb{R}^{|F^\star| \times R_2}$ and $\mathbf{T} \in \mathbb{R}^{|\mathcal{T}| \times R_3}$ as follows:

$$\boldsymbol{\mathcal{D}} \approx \widehat{\boldsymbol{\mathcal{D}}} = \boldsymbol{\mathcal{H}} \times_1 \mathbf{S} \times_2 \mathbf{F} \times_3 \mathbf{T}, \tag{2.116}$$

where all the columns contained in the factor matrices $\mathbf{S}$, $\mathbf{F}$ and $\mathbf{T}$ are orthonormal, and the vector

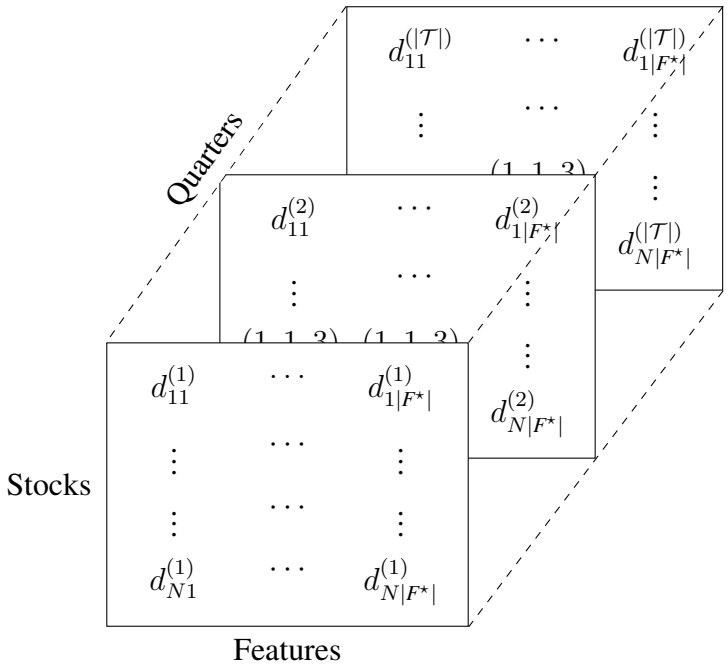

Figure 2.22: Schematic of a third-order market fundamentals design tensor during January 1990-December 2019.

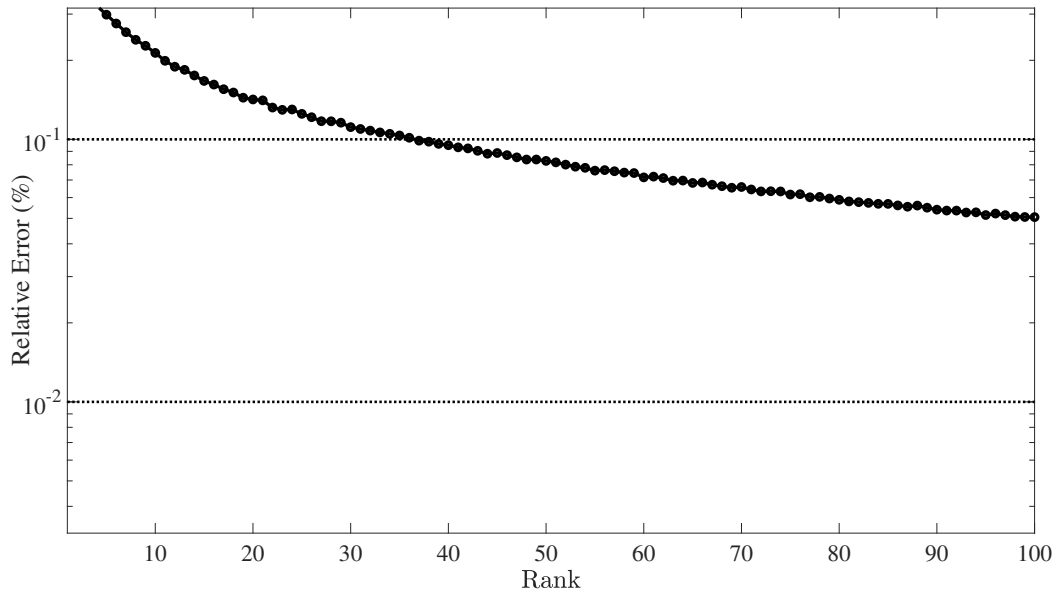

Figure 2.23: Plot of the relative error of polyadic decomposition of $\boldsymbol{\mathcal{D}}$ for different values of R.

$\mathbf{R} = [R_1, R_2, R_3]^\intercal$ is referred to as the *multilinear rank* of $\widehat{\boldsymbol{\mathcal{D}}}$. By having an initial guess of $\boldsymbol{\mathcal{H}}$, the factor matrices are then obtained by solving the following optimization problem:

$$\min_{\mathbf{S},\mathbf{F},\mathbf{T}} \|\boldsymbol{\mathcal{D}} - \boldsymbol{\mathcal{H}} \times_1 \mathbf{S} \times_2 \mathbf{F} \times_3 \mathbf{T}\|_F. \tag{2.117}$$

In order to solve model (2.117), we fix $R_1 = R_2 = 69$ for the core tensor $\boldsymbol{\mathcal{H}}$ which coincides with the number of industries in GICS scheme, as we are inclined to believe that making such a choice could effectively address the correlations existing among different industry categories of the stocks. Subsequently, we solve the optimization problem for a variety of $R_3 > 1$ values and obtain the relative error of its approximation given by

$$rel_{error} = \frac{\|\boldsymbol{\mathcal{D}} - \widehat{\boldsymbol{\mathcal{D}}}\|_F}{\|\boldsymbol{\mathcal{D}}\|_F}. \tag{2.118}$$

An approximate tensor $\widehat{\boldsymbol{\mathcal{D}}}$ would then be deemed ideal if its relative approximation error given by equation (2.118) is minimal. Note that $\widehat{\boldsymbol{\mathcal{D}}}$ becomes a relatively accurate approximation of $\boldsymbol{\mathcal{D}}$ for an appropriately chosen core tensor $\boldsymbol{\mathcal{H}}$. However, for the sake of smoothing out its noise, we opt for choosing a tensor approximation $\widehat{\boldsymbol{\mathcal{D}}}$ whose relative error of approximation is not substantially small, in order to ensure that $\widehat{\boldsymbol{\mathcal{D}}}$ captures only the most significant and salient information embedded in $\boldsymbol{\mathcal{D}}$. As per our conducted experiments, the value $R_3 = 20$ was considered to be a proper choice for the temporal rank, yielding 6.811% of the relative error of approximation.

Once the factor matrices are obtained, we gear our focus to 20 time series contained within the columns of the factor matrix $\mathbf{T}$. Each of these time series of length $|\mathcal{T}|$ could potentially be seen as a random vector which captures some aspects of the temporal information contained in $\boldsymbol{\mathcal{D}}$. As a consequence, one may derive a time-dependent statistic for the information contained in $\boldsymbol{\mathcal{D}}$ by averaging the time series which comprise the factor matrix $\mathbf{T}$. We refer to this statistic as the market fundamental index whose formal definition is given below.

Definition 2.6.1 (MFI)**.** *Let $\boldsymbol{\mathcal{D}}$ be a tensor composed of the information related to the fundamental features of assets over time which is smoothed out by a tensor $\widehat{\boldsymbol{\mathcal{D}}} = \boldsymbol{\mathcal{H}} \times_1 \mathbf{S} \times_2 \mathbf{F} \times_3 \mathbf{T}$, where*

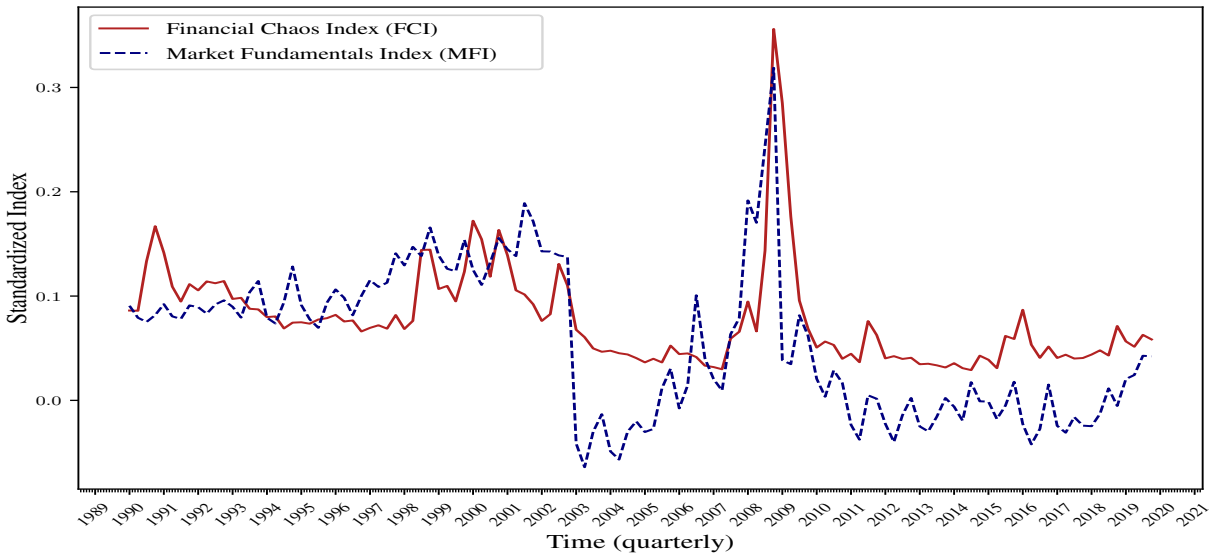

Figure 2.24: Plot of the standardized quarterly FCI_t vs quarterly MFI_t during January 1990-December 2019.

$\boldsymbol{\mathcal{H}} \in \mathbb{R}^{R_1 \times R_2 \times R_3}$, $\mathbf{S} \in \mathbb{R}^{N \times R_1}$, $\mathbf{F} \in \mathbb{R}^{|F^\star| \times R_2}$ *and* $\mathbf{T} \in \mathbb{R}^{|\mathcal{T}| \times R_3}$. *Then, the market fundamentals index (MFI) is defined by the following function:*

$$\mathrm{MFI}(t) := \frac{1}{R_3} \sum_{k=1}^{R_3} |\mathbf{T}_{kt}|. \tag{2.119}$$

By analogy to the case of FCI, we also denote the underlying time series of MFI by MFI_t. Figure 2.24 compares the standardized quarterly realizations of FCI_t and MFI_t during the time period January 1990-December 2019, according to which obvious patterns of co-movements are observed.

2.6.3 Tests of Statistical Independence

In what follows, we investigate whether the above-introduced *market fundamentals index* is dependent of and correlated with the *financial chaos index*. We consider carrying out such an analysis being substantial since its conclusions will suggest a path of possible development for further ideas in the forthcoming subsection.

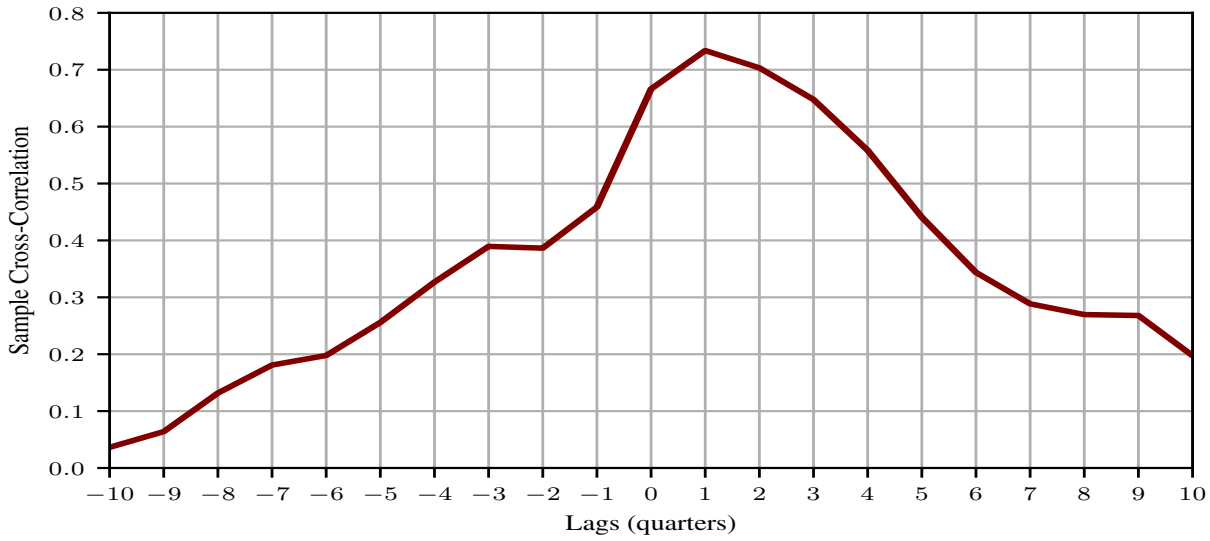

Figure 2.25: Plot of the sample cross-correlation for quarterly FCI_t-MIX_t during January 1990-December 2019.

To begin with, we compute the sample cross-correlation between quarterly FCI_t and MFI_t for various lag values using the relation given by equation (2.88), whose results are displayed in Figure 2.25 implying that high level of correlation exist between these two time series.

However, the fact that two time series are highly correlated, does not necessarily imply that they are dependent, e.g., see (Dean and Dunsmuir, 2016) for more details. More rigorously, we employ statistical hypothesis tests of independence for these two time series in order to characterize them in more in-depth. For this purpose, we adhere to the framework developed in (Gretton and Györfi, 2008; Gretton and Györfi, 2010) for consistent non-parametric tests of independence between quarterly FCI_t and MFI_t. Let $(\mathrm{FCI}_{t=1}, \mathrm{MFI}_{t=1}), \ldots, (\mathrm{FCI}_{t=n}, \mathrm{MFI}_{t=n})$ be an $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$-valued random sample whose corresponding pairs are drawn from certain distributions defined on the same probability space. Also, denote the joint and marginal *cumulative distribution functions* of FCI_t and MFI_t by $\mathbb{G}_{\mathrm{FCI}_t,\mathrm{MFI}_t}(\cdot,\cdot)$, $\mathbb{G}_{\mathrm{FCI}_t}(\cdot)$ and $\mathbb{G}_{\mathrm{MFI}_t}(\cdot)$, while their *empirical distribution functions* are denoted by $\mathbb{G}_{n,\mathrm{FCI}_t,\mathrm{MFI}_t}(\cdot,\cdot)$, $\mathbb{G}_{n,\mathrm{FCI}_t}(\cdot)$ and $\mathbb{G}_{n,\mathrm{MFI}_t}(\cdot)$.

Clearly, the null hypothesis that FCI_t and MFI_t are independent can be formalized as follows:

$$\mathrm{H}_0^{\texttt{INDP}} : \mathbb{G}_{\mathrm{FCI}_t,\mathrm{MFI}_t} = \mathbb{G}_{\mathrm{FCI}_t} \times \mathbb{G}_{\mathrm{MFI}_t}.$$

In order to test such null hypothesis $\mathrm{H}_0^{\texttt{INDP}}$, we introduce an L_1-based test statistic which employs the given finite partitions $\mathcal{P}_n = \{A_{n,1}, \ldots, A_{n,m_n}\}$ and $\mathcal{Q}_n = \{B_{n,1}, \ldots, B_{n,m'_n}\}$ of $\mathbb{R}_{\geq 0}$. Namely, define

$$L_n(\mathbb{G}_{n,\mathrm{FCI}_t,\mathrm{MFI}_t}, \mathbb{G}_{n,\mathrm{FCI}_t} \times \mathbb{G}_{n,\mathrm{MFI}_t}) := \sum_{A\in\mathcal{P}_n} \sum_{B\in\mathcal{Q}_n} |\mathbb{G}_{n,\mathrm{FCI}_t,\mathrm{MFI}_t}(A \times B) - \mathbb{G}_{n,\mathrm{FCI}_t}(A)\mathbb{G}_{n,\mathrm{MFI}_t}(B)|, \tag{2.120}$$

where the empirical measures associated with the samples $(\mathrm{FCI}_{t=1}, \mathrm{MIX}_{t=1}), \ldots, (\mathrm{FCI}_{t=n}, \mathrm{MIX}_{t=n})$, $\mathrm{FCI}_{t=1}, \ldots, \mathrm{FCI}_{t=n}$, and $\mathrm{MFI}_{t=1}, \ldots, \mathrm{MFI}_{t=n}$ are defined for any pair of Borel sets A and B as follows:

$$\mathbb{G}_{n,\mathrm{FCI}_t,\mathrm{MFI}_t}(A \times B) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\Big[(\mathrm{FCI}_{t=i}, \mathrm{MFI}_{t=i}) \in A \times B\Big], \tag{2.121}$$

$$\mathbb{G}_{n,\mathrm{FCI}_t}(A) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\Big[\mathrm{FCI}_{t=i} \in A\Big], \tag{2.122}$$

and

$$\mathbb{G}_{n,\mathrm{MFI}_t}(B) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\Big[\mathrm{MFI}_{t=i} \in B\Big], \tag{2.123}$$

respectively. Furthermore, an alternative test statistic used for assessing independence of two time series is the *log-likelihood statistic* given by

$$I_n(\mathbb{G}_{n,\mathrm{FCI}_t,\mathrm{MFI}_t}, \mathbb{G}_{n,\mathrm{FCI}_t} \times \mathbb{G}_{n,\mathrm{MFI}_t}) := \sum_{A\in\mathcal{P}_n} \sum_{B\in\mathcal{Q}_n} \mathbb{G}_{n,\mathrm{FCI}_t,\mathrm{MFI}_t}(A \times B) \log \frac{\mathbb{G}_{n,\mathrm{FCI}_t,\mathrm{MFI}_t}(A \times B)}{\mathbb{G}_{n,\mathrm{FCI}_t}(A)\mathbb{G}_{n,\mathrm{MFI}_t}(B)}. \tag{2.124}$$

In the case where the values of the computed test statistics given by equation (2.120) or equation (2.124) exceed the corresponding given thresholds at the specified levels of significance, the respective null hypotheses of independence should be rejected.

Table 2.13 presents the results of running independence tests between quarterly FCI_t and MFI_t time series. As per to the results reported therein, the null hypothesis $\mathrm{H}_0^{\texttt{INDP}}$ gets rejected at 1% level

Table 2.13: Results of independence tests between quarterly FCI_t and MFI_t.

		L_1-based	log-likelihood
Test statistics		0.839	1.073
Threshold	1% level	0.514	0.464
	5% level	0.492	0.416
	10% level	0.479	0.391

of significance using either the L_1-based test statistic or the log-likelihood one. Also, note that, for our experiments the sample size was $n = |\mathcal{T}| = 120$, and the number of discretisations per dimension was selected to be $m_n = m_n^{'} = 6$. The latter choice was made based on the results obtained in Section 2.3.2, where the market was shown to be segmented into 6 parts during 30 years time span considered.

2.6.4 Block-wise Granger Causality

Let us first assume that the market fundamentals index is already at our disposal. Hence, we may devote the remainder of this section to discover the existence of causal relations from *policy uncertainty indexes* to FCI_t by further taking into account the assets' fundamentals information. To this end, we first formulate a conditional null hypothesis and then perform a 1-*step block-wise Granger causality* test to identify and characterize possible causal relations. A reason for choosing only 1-step-ahead Granger causality tests, is due to the observation that selecting a minimal possible time-horizon size (i.e., $h = 1$ or equivalently, one single quarter of the time unit) could be sufficient to capture potential changes in macroeconomic factors and assets' fundamentals that may influence the stock market trends. In addition, inspection of Figure 2.25 supports our choice for the size of time-horizon, whereby it is observed that the maximum cross-correlation between FCI_t and MFI_t occurs at 1-lag quarter having its value to be 73.378%.

Denote by $\mathcal{J}_{t,\mathtt{Q}}^{\mathtt{PRC}} = \mathscr{F}_{\mathrm{FCI},t}$ the set of known information for FCI_t up to a quarter of time unit t. Also, we will slightly abuse the notation, by using $\mathbf{F}_t$ to represent FCI_t. Let us now define the

following quarterly multivariate time series

$$\mathbf{U}_t = (\mathrm{VIX}_t, \mathrm{MFI}_t)^\intercal,$$

and

$$\mathbf{W}_t = (\mathrm{EMV}_t, \mathrm{EPU}_t, \mathrm{MON}_t, \mathrm{FIS}_t, \mathrm{TAX}_t, \mathrm{GOV}_t, \mathrm{HLTH}_t, \mathrm{SEC}_t, \mathrm{ENT}_t, \mathrm{GREG}_t, \mathrm{FREG}_t, \mathrm{TRD}_t, \mathrm{CPI}_t)^\intercal,$$

with their respective sets of the known information represented by $\mathscr{F}_{\mathrm{VIX},t} \bigotimes \mathscr{F}_{\mathrm{MFI},t}$, and

$$\mathcal{J}_{t,\mathtt{Q}}^{\mathtt{NEWS}} = \bigotimes_{k=1}^{K} \mathscr{F}_{\mathbf{W}_k,t},$$

where $K = 13$.

Next, we would like to test whether the group of individual univariate components of $\mathbf{W}_t$ are not a 1-step Granger cause of FCI_t, given the information on the respective components of $\mathbf{U}_t$. However, since both components of $\mathbf{U}_t$ are statistically dependent on FCI_t, i.e., VIX_t dependents through cointegration (e.g., see Section 2.4.2) and MFI_t by implications of the independence tests executed in Section 2.6.3, then possibilities of inferring spurious causalities would potentially be eliminated if these common dependencies (VIX_t and FCI_t) are *conditioned out*.

To this end, we express the joint distribution of the multivariate processes $(\mathbf{F}_t, \mathbf{W}_t, \mathbf{U}_t)^\intercal$ in VAR(p) form as follows:

$$\begin{pmatrix} \mathbf{F}_t \\ \mathbf{W}_t \\ \mathbf{U}_t \end{pmatrix} = \sum_{k=1}^{p} \begin{pmatrix} b_{ff,k} & b_{fw,k} & b_{fu,k} \\ b_{wf,k} & b_{ww,k} & b_{wu,k} \\ b_{uf,k} & b_{uw,k} & b_{uu,k} \end{pmatrix} \begin{pmatrix} \mathbf{F}_{t-k} \\ \mathbf{W}_{t-k} \\ \mathbf{U}_{t-k} \end{pmatrix} + \begin{pmatrix} \epsilon_{f,t} \\ \boldsymbol{\epsilon}_{w,t} \\ \boldsymbol{\epsilon}_{u,t} \end{pmatrix}. \tag{2.125}$$

Here, the residual variance-covariance matrix of the error vectors takes on a form as follows:

$$\boldsymbol{\Sigma} = \operatorname{cov} \begin{pmatrix} \epsilon_{f,t} \\ \boldsymbol{\epsilon}_{w,t} \\ \boldsymbol{\epsilon}_{u,t} \end{pmatrix} = \begin{pmatrix} \Sigma_{ff} & \boldsymbol{\Sigma}_{fw} & \boldsymbol{\Sigma}_{fu} \\ \boldsymbol{\Sigma}_{wf} & \boldsymbol{\Sigma}_{ww} & \boldsymbol{\Sigma}_{wu} \\ \boldsymbol{\Sigma}_{uf} & \boldsymbol{\Sigma}_{uw} & \boldsymbol{\Sigma}_{uu} \end{pmatrix}. \tag{2.126}$$

Note that that the relation given by equation (2.125) does not involve neither the shift nor a linear trend components due to the fact that FCI_t is an $I(1)$ process with no shift or linear trend, as per Section 2.2.2. Subsequently, by analogy to Section 2.5, we express the *unrestricted* and *restricted* regression models by

$$\mathbf{F}_t = \sum_{k=1}^{p} b_{ff,k}\mathbf{F}_{t-k} + \sum_{k=1}^{p} b_{fw,k}\mathbf{W}_{t-k} + \sum_{k=1}^{p} b_{fu,k}\mathbf{U}_{t-k} + \epsilon_{f,t}, \tag{2.127}$$

and

$$\mathbf{F}_t = \sum_{k=1}^{p} b^{'}_{ff,k}\mathbf{F}_{t-k} + \sum_{k=1}^{p} b^{'}_{fu,k}\mathbf{U}_{t-k} + \epsilon^{'}_{f,t}, \tag{2.128}$$

respectively, where the variables used for conditioning that are contained in $\mathbf{U}_t$, are present in both of the above models.

In turn, the null hypothesis on no causality is formulated as follows:

$$\mathrm{H}_{0,\mathtt{Q}}^{\mathtt{NEWS}} : \forall k \in \{1, 2 \ldots, p\}, b_{wf,k} = 0,$$

whose corresponding alternative hypothesis is:

$$\mathrm{H}_{1,\mathtt{Q}}^{\mathtt{NEWS}} : \exists k \in \{1, 2 \ldots, p\}, b_{wf,k} \neq 0.$$

Similar to equation (2.111), the test statistic based on the generalized variance of the regression models is utulized, which estimates the *model prediction error* and is given by the following log-likelihood ratio:

$$\mathrm{GC}_{\mathbf{W}_t \overset{h=1}{\nrightarrow} \mathbf{F}_t | \mathbf{U}_t} = \log \frac{|\Sigma^{'}_{ff}|}{|\Sigma_{ff}|}. \tag{2.129}$$

Table 2.14: Results of Granger causality test for $\mathrm{H}_{0,\mathtt{Q}}^{\mathtt{NEWS}}$.

		χ^2-test	F-test
p-value		0.129	0.234
Test statistics		63.663	1.224
Critical values	1% level	78.660	1.923
	5% level	69.832	1.585
	10% level	65.422	1.431

Table 2.14 summarizes the results obtained for testing $\mathrm{H}_{0,\mathtt{Q}}^{\mathtt{NEWS}}$. Both χ^2 and F underlying distributions for the test statistics were considered. As per these results, there is no sufficient evidence to reject the null hypothesis $\mathrm{H}_{0,\mathtt{Q}}^{\mathtt{NEWS}}$ at 1% level of significance using either of these test statistics. In turn, this implies that the mutual quarterly price changes of the assets are 1-step Granger caused by news-based uncertainty indexes, given the known information sets pertaining to VIX_t and MFI_t. In other words, the null hypothesis does not get rejected w.r.t. the known information set $\mathcal{J}_{t,\mathtt{Q}}^{\mathtt{PRC}} \bigoplus \mathcal{J}_{t,\mathtt{Q}}^{\mathtt{NEWS}} | \mathscr{F}_{\mathrm{VIX},t} \bigotimes \mathscr{F}_{\mathrm{MFI},t}$.

Moreover, it is very important to investigate the impacts of news-based information ($\mathcal{J}_{t,\mathtt{Q}}^{\mathtt{NEWS}}$) augmented with the fundamentals' information ($\mathscr{F}_{\mathrm{MFI},t}$), on the price-generating mechanisms of the assets. As we may assume without the loss of generality that such an augmented set of information alongside the price-related sets of known information ($\mathscr{F}_{\mathrm{FCI},t}$ and $\mathscr{F}_{\mathrm{VIX},t}$) contains all the *publicly known information* about the assets. To this end, we reshape the configuration of the elements contained in the multivariate time series $\mathbf{W}_t$ and $\mathbf{U}_t$, and introduce two new time series

$$\widetilde{\mathbf{U}}_t = \mathrm{VIX}_t,$$

and

$$\widetilde{\mathbf{W}}_t = (\mathrm{MFI}_t, \mathrm{EMV}_t, \mathrm{EPU}_t, \mathrm{MON}_t, \mathrm{FIS}_t, \mathrm{TAX}_t, \mathrm{GOV}_t, \mathrm{HLTH}_t, \mathrm{SEC}_t, \mathrm{ENT}_t, \mathrm{GREG}_t, \mathrm{FREG}_t, \mathrm{TRD}_t, \mathrm{CPI}_t)^{\intercal},$$

Table 2.15: Results of Granger causality test for $\mathrm{H}_{0,\mathtt{Q}}^{\mathtt{PUB}}$.

		χ^2-test	F-test
p-value		0.014	0.134
Test statistics		98.603	1.409
Critical values	1% level	100.425	2.072
	5% level	90.531	1.665
	10% level	85.527	1.486

whose respective sets of known information are represented by $\mathscr{F}_{\mathrm{VIX},t}$ and

$$\mathcal{J}_{t,\mathtt{Q}}^{\mathtt{PUB}} = \left(\bigotimes_{k=2}^{K'} \mathscr{F}_{\widetilde{\mathbf{W}}_k,t} \right) \bigoplus \mathscr{F}_{\mathrm{MFI},t},$$

where $K' = 14$. Thereafter, the corresponding *unrestricted* and *restricted* regression models are re-derived, but for this time it is to be done by substituting $\widetilde{\mathbf{U}}_t$ for $\mathbf{U}_t$ and $\widetilde{\mathbf{W}}_t$ for $\mathbf{W}_t$. Subsequently, we reformulate a null hypothesis as follows:

$$\mathrm{H}_{0,\mathtt{Q}}^{\mathtt{PUB}} : \forall k \in \{1, 2 \ldots, p\}, \tilde{b}_{\tilde{w}f,k} = 0,$$

and test it versus the following alternative:

$$\mathrm{H}_{1,\mathtt{Q}}^{\mathtt{PUB}} : \exists k \in \{1, 2 \ldots, p\}, \tilde{b}_{\tilde{w}f,k} \neq 0.$$

The corresponding test statistic is further given by

$$\mathrm{GC}_{\widetilde{\mathbf{W}}_t \overset{h=1}{\nrightarrow} \mathbf{F}_t | \widetilde{\mathbf{U}}_t} = \log \frac{|\widetilde{\Sigma}'_{ff}|}{|\widetilde{\Sigma}_{ff}|}. \tag{2.130}$$

According to the results presented in Table 2.15, there is insufficient evidence to reject the null hypothesis that the market fundamentals index alongside newschapter-based uncertainty indexes

do not 1-step Granger cause FCI_t at 1% level of significance, i.e., the null hypothesis does not get rejected w.r.t. the known information set $\mathcal{J}_{t,\mathtt{Q}}^{\mathtt{PRC}} \bigoplus \mathcal{J}_{t,\mathtt{Q}}^{\mathtt{PUB}} | \mathscr{F}_{\mathrm{VIX},t}$. In turn, this implies that the quarterly price changes exhibited by the mutual behavior of the assets in the stock market are 1-step Granger caused by macroeconomic and/or policy uncertainty news as well as the fundamentals-related information.

2.7 The Efficient Market Hypothesis

In the last section of our chapter, we employ the previously developed tools and concepts and concentrate on the long-standing *efficient market hypothesis* (EMH). The literature on EMH and its definition is vast and extensive; we refer the interested reader to (Degutis and Novickytė, 2014; Gabriela ğiĠan, 2015; Lehmann, 1990; Naseer and Tariq, 2015; Shi and Delacroix, 2018; Stein, 1989), to list a few. However, in some broad sense, EMH stipulates that market prices fully reflect all available information, further implying that the sequences of the asset price changes would behave in a more random manner in the case where the market exhibits larger degrees of efficiency, to the extent that in a completely efficient market the price changes would be completely unpredictable.

One possible approach that we follow in this section is to formulate the EMH using the notion of the financial chaos index. Note that such a possibility stems from the remark made in Section 2.2 where it was discussed that FCI_t could potentially be considered as a consistent and robust estimator of the market volatility. Hence, we present three versions of the EMH formulated in terms of the financial chaos index which are in the same spirit as to the definitions provided in (Jensen, 1978).

We stress the fact that failure to reject EMH, or alternatively accept it, has few implications for efficiency of the market price formations insofar as there is insufficient existence of a thorough and explicit economic model accounting for price-generating mechanisms, e.g., see (Lo and MacKinlay, 1988) for similar arguments. State it differently, the price formation of the assets in the stock

market during the considered time period is deemed efficient only in relation to our underlying tensorial economic model.

Throughout, we will use $\mathcal{J}_{t,\mathtt{D}}$, $\mathcal{J}_{t,\mathtt{M}}$ and $\mathcal{J}_{t,\mathtt{Q}}$ to denote the sets of some known information per daily, monthly and quarterly time frequencies, respectively.

2.7.1 Weak EMH

A *weak* form of the EMH stipulates that given the past and present values of the assets, it is impossible to predict their future values. The weak form of EMH can be defined in terms of our developed financial chaos index as follows:

Definition 2.7.1 (Weak EMH)**.** *Let $\mathcal{J}_t^{\mathtt{PRC}}$ denote a set of known information for financial chaos index at time t. Assume without the loss of generality that $\mathcal{J}_t^{\mathtt{PRC}}$ entails all the known information pertaining to the observed historical prices of N particular assets up to time $t \in \mathcal{T}$. Then, the market is said to be weakly efficient w.r.t. $\mathcal{J}_t^{\mathtt{PRC}}$ if FCI_t is a random walk.*

As per our analysis and discussion given in Section 2.2.4, the daily FCI_t did not satisfy the conditions for being a random walk. Thus, we may conclude that the stock market during the time period January 1990-December 2019 was weakly inefficient w.r.t. $\mathcal{J}_{t,\mathtt{D}}^{\mathtt{PRC}}$. On the contrary, for the monthly and quarterly time frequencies, there exist insufficient evidence to reject the random walk hypotheses on FCI_t. As a consequence, the stock market during the given time period can be regarded as being weakly efficient w.r.t. both $\mathcal{J}_{t,\mathtt{M}}^{\mathtt{PRC}}$ and $\mathcal{J}_{t,\mathtt{Q}}^{\mathtt{PRC}}$ at 1% level of significance.

2.7.2 Potent EMH

Here, we introduce a concept of the *potent* EMH to investigate whether news-based uncertainty indexes Granger cause the past and current values of the assets. To be more precise, if the news-based information sets really influence the prices in an *efficient market*, then there should only exist instantaneous Granger causal relations from news-based information on the prices. In other

words, one would not expect to find any strictly Granger causal relation from such news-based information on the prices in an *efficient market*, for any strictly causal relation would indicate that there already exist news-based information which have not been utilized efficiently, e.g., compare to (Kirchgässner, Wolters, and Hassler, 2012). This can be made rigorous by using the financial chaos index. Specifically, the potent EMH can be defined as follows:

Definition 2.7.2 (Potent EMH)**.** *Let $\mathcal{J}_t^{\texttt{PRC}}$ and $\mathcal{J}_t^{\texttt{NEWS}}$ denote sets of known information for financial chaos index and a collection of news-based uncertainty indexes at time t, respectively. Assume without the loss of generality that the set $\mathcal{J}_t^{\texttt{PRC}} \bigoplus \mathcal{J}_t^{\texttt{NEWS}}$ entails all the known information pertaining to the observed historical prices of specific N assets as well as the macroeconomic and various types of policy uncertainty news up to time $t \in \mathcal{T}$. Then, the market is said to be potently efficient w.r.t. $\mathcal{J}_t^{\texttt{PRC}} \bigoplus \mathcal{J}_t^{\texttt{NEWS}}$ if FCI_t is a random walk, and there is no strictly Granger causal relation from the collection of news-based policy uncertainty indexes on FCI_t.*

As per the above definition, we may state that the U.S. stock market during the time period January 1990-December 2019 was potently inefficient w.r.t. $\mathcal{J}_{t,\texttt{D}}^{\texttt{PRC}} \bigoplus \mathcal{J}_{t,\texttt{D}}^{\texttt{NEWS}}$, since the daily FCI_t during the considered time period did not satisfy the conditions for being a random walk, and further the null hypothesis that daily FCI_t was time-lagged (strictly) Granger caused by news-based uncertainty indexes at all the 1% levels of significance (e.g., see Section 2.5.2 for more details). On the other hand, it was shown that for the monthly and quarterly data, FCI_t was a random walk, and also that the news-based policy uncertainty indexes did not strictly Granger cause FCI_t (e.g., see Sections 2.5.3 and 2.6.4 for more information). Therefore, the stock market during the given time period could be proclaimed to have been potently efficient w.r.t. $\mathcal{J}_{t,\texttt{M}}^{\texttt{PRC}} \bigoplus \mathcal{J}_{t,\texttt{M}}^{\texttt{NEWS}}$ and $\mathcal{J}_{t,\texttt{Q}}^{\texttt{PRC}} \bigoplus \mathcal{J}_{t,\texttt{Q}}^{\texttt{NEWS}}$ at 1% level of significance.

2.7.3 Semi-strong EMH

Our final goal is to investigate whether all the publicly known information Granger cause the past and current values of the assets. By assuming that the set of all publicly known information is comprised of news-based information augmented with the information pertaining to fundamental

features of the assets, a *semi-strong* version of EMH can be made rigorous using the notion of the financial chaos index as follows:

Definition 2.7.3 (Semi-strong EMH)**.** *Let $\mathcal{J}_t^{\texttt{PRC}}$ and $\mathcal{J}_t^{\texttt{PUB}}$ denote sets of known information for financial chaos index and publicly accessible information at time t, respectively. Assume without the loss of generality that the set $\mathcal{J}_t^{\texttt{PRC}} \bigoplus \mathcal{J}_t^{\texttt{PUB}}$ entails all the known information pertaining to the observed historical prices of specific N assets as well as the publicly accessible information provided by the news-based sources and assets' fundamentals up to time $t \in \mathcal{T}$. Then, the market is said to be semi-strongly efficient w.r.t. $\mathcal{J}_t^{\texttt{PRC}} \bigoplus \mathcal{J}_t^{\texttt{PUB}}$ if FCI_t is a random walk, and there is no strictly Granger causal relation from the collection of news-based policy uncertainty indexes and/or market fundamental index on FCI_t.*

For the daily and monthly data, the semi-strong form of EMH basically reduces to the potent version of EMH since the information pertaining to the fundamentals of assets are typically accessible to agents only per quarterly or annual time frequencies, i.e., $\mathcal{J}_{t,\texttt{D}}^{\texttt{PUB}} = \mathcal{J}_{t,\texttt{D}}^{\texttt{NEWS}}$ and $\mathcal{J}_{t,\texttt{M}}^{\texttt{PUB}} = \mathcal{J}_{t,\texttt{M}}^{\texttt{NEWS}}$. Thus, we may conclude that the U.S. stock market was semi-strongly inefficient w.r.t. $\mathcal{J}_{t,\texttt{D}}^{\texttt{PRC}} \bigoplus \mathcal{J}_{t,\texttt{D}}^{\texttt{PUB}}$ at the significance level 1%, while the market was semi-strongly efficient w.r.t. $\mathcal{J}_{t,\texttt{M}}^{\texttt{PRC}} \bigoplus \mathcal{J}_{t,\texttt{M}}^{\texttt{PUB}}$ at 1% level of significance.

However, the set $\mathcal{J}_{t,\texttt{Q}}^{\texttt{PUB}}$ is comprised of the known information on collection of policy uncertainty indexes together with the market fundamentals index. Since the quarterly FCI_t was shown to be a random walk and time-lagged Granger causal relations did not exist from the news-based policy uncertainty indexes and/or market fundamental index on FCI_t (e.g., see Section 2.6.4 for more details), the market during the time period January 1990-December 2019 could be claimed to have been semi-strongly efficient w.r.t. $\mathcal{J}_{t,\texttt{Q}}^{\texttt{PRC}} \bigoplus \mathcal{J}_{t,\texttt{Q}}^{\texttt{PUB}}$ at 1% level of significance.

Chapter 3

Time-homogeneous Top-K Ranking

3.1 Time-independent Top-K Ranking

Assume that a list of N items is given, each of which having a positive latent strength denoted by $w_i^\star$, $i = 1, 2, \ldots, N$. The problem is then to infer these strengths based on their observed pairwise comparisons, so that the top-K items are identified.

In order to estimate the latent strengths $w_i^\star$, we find it practical to introduce the notion of *advantage* for pairs of items. Let $a_{ij}^{(t)}$ denote the advantage of item j over item i at a given time t. In the special case of stock markets and using the introduced constitution criterion lag-1 rate of return (see Chapter 2 for more details), one may formulate the pairwise advantages as

$$a_{ij}^{(t)} = \frac{r_j^{(t)}}{r_i^{(t)}} \,, \tag{3.1}$$

where $r_i^{(t)}$ represents the *lag*-1 *rate of return* of asset i at a given time t, i.e.

$$r_i^{(t)} = \frac{c_i^{(t)}}{c_i^{(t-1)}} \,. \tag{3.2}$$

Here, $c_i^{(t)}$ is the adjusted closing price for the asset i at time t. The quantity $a_{ij}^{(t)}$ in equation (3.1) is

a potential measure for relative profitability of asset j w.r.t. asset i, and it could be a natural statistic to use for those traders who wish to maximize their marginal capital returns. In fact, the larger $a_{ij}^{(t)}$ becomes, the more confidence could a trader put on out-performance of the asset j over asset i. It is noteworthy that if $a_{ij}^{(t)}$ is close enough to one, then a lack of sufficient basis for comparing the pair (i, j) of assets is ascertained.

3.1.1 Continuous Extension to BTL Model

In what follows, we develop a continuous extension to the discrete BTL model. The primary application of our proposed model is to infer the latent strengths of the items when it is deemed necessary to consider the continuous nature of the pairwise comparisons. It is noted that, the data-continuity concept dealt with in this section differs from time-continuity which is discussed later.

Let us first recall the definition and some basic properties of the BTL model; see (Cattelan, 2012) and references therein for more details. Based on this model, outcomes of pairwise comparisons performed among the items would necessarily have a win/loss nature; i.e. the binary variable

$$y_{ij}^{(t)} = \begin{cases} 1\,, & \text{if } a_{ij}^{(t)} > 1\,, \\ 0\,, & \text{otherwise}\,, \end{cases} \tag{3.3}$$

can be employed to indicate an outcome of the pairwise comparison at time t. Accordingly, item j is preferred over item i if it has a larger advantage value. It is noted that the non-measurable sets $\{a_{ij}^{(t)} = 1\}$ are assumed to imply non-preference states among the compared items. Thereafter, (3.3) is considered to be a realization of a particular Bernoulli $(0/1)$ random variable denoted by $Y_{ij}^{(t)}$, such that

$$\mathbb{P}\left(Y_{ij}^{(t)} = 1\right) = \frac{w_j^\star}{w_i^\star + w_j^\star} \tag{3.4}$$

for all $i, j = 1, 2, \ldots, N$, and $t = 1, 2, \ldots, T$, where T stands for the total number of comparisons. In the context of stock markets, T may correspond to the total number of trading days in a specific

week, month, year, etc.

The win/loss nature of the discrete BTL model finds a wide range of applications in real-world problems. However, this model does not provide any information on *how strongly* the compared items are preferred over each other. To alleviate this drawback, we re-define the binary variable in (3.3) using the following non-negative continuous variable

$$y_{ij}^{(t)} = \begin{cases} \log a_{ij}^{(t)} \,, & \text{if } a_{ij}^{(t)} > 1 \,, \\ 0 \,, & \text{otherwise} \,. \end{cases} \tag{3.5}$$

In this setting, the item j would further be assigned some *degree of preference* when it is observed to outperform the item i. One may then associate (3.5) to a continuous random variable represented by $Y_{ij}^{(t)}$, for $i, j = 1, 2, \ldots, N$, such that

$$\mathbb{P}\left(Y_{ij}^{(t)} > \delta\right) = \frac{w_j^\star}{w_i^\star + w_j^\star} \,,$$

where $\delta \geq 0$ is a given constant. Similarly to its discrete counterpart, (3.5) should be 0 in the case where item j is not the preferred one within the pair (i, j) of items. Also, for sufficiently small δ, it could be shown that the continuous BTL model would reduce to the discrete model.

3.1.2 Rank Centrality

Once y_{ij}^t's are observed for all pairwise comparisons throughout the time span $t = 1, 2, \ldots, T$, one may proceed to obtain the top-K items. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote a directed *comparison graph* where the *vertex set* $\mathcal{V} = \{1, 2, \ldots, N\}$ contains the given items, and the *edge set* $\mathcal{E}$ includes weights of all pairs of items which are compared. According to the rank centrality algorithm (Negahban, Oh,

and Shah, 2017), edge weights of $\mathcal{G}$ can be evaluated as follows

$$w_{ij} = \frac{\sum\limits_{t=1}^{T} y_{ij}^{(t)}}{\sum\limits_{t=1}^{T} y_{ij}^{(t)} + \sum\limits_{t=1}^{T} y_{ji}^{(t)}} \,. \tag{3.6}$$

Next, consider a random walk on graph $\mathcal{G}$ which is characterized by the following transition matrix

$$p_{ij} = \begin{cases} \frac{1}{\deg_{\max}^{+}} w_{ij} \,, & i \neq j \,, \\ 1 - \frac{1}{\deg_{\max}^{+}} \sum\limits_{k \neq i} w_{ik} \,, & i = j \,, \end{cases} \tag{3.7}$$

for every pair $(i, j) \in \mathcal{E}$, where $\deg_{\max}^{+}$ represents the maximum out-degree of $\mathcal{G}$. Finally, the principal left eigenvector $\boldsymbol{\pi}$ of $\mathbf{P} = [p_{ij}]_{1 \leq i,j \leq N}$ is computed, and afterwards the items acquiring the K largest entries of $\boldsymbol{\pi}$ would form the top-K ranking list. Here, the reciprocal $1/\deg_{\max}^{+}$ represents the scaling factor which ensures that $\mathbf{P}$ is a legitimate *probability transition matrix*. A more detailed outline of this procedure is presented in Algorithm 1.

It is remarked that, inasmuch as (3.3) is taken into consideration, the numerator of equation (3.6) counts the total number of observations in which the item j has been preferred over item i throughout T comparisons. However, in the case where (3.5) is taken into account instead, one may view the numerator of equation (3.6) as a *scaled average preference degree* of the item j over item i throughout the comparisons performed.

3.2 Time-homogeneous Top-K Ranking

The situation which is often encountered in problems having temporal dimensions is that, the top-K items obtained using Algorithm 1 would vary considerably from one period of time to another. These fluctuations emerge primarily as a consequence of complex interactions occurring among the items. In this section, we propose a time-homogeneous alternative to Algorithm 1. Our goal is to obtain some consistent time-dependent sequence of top-K items, where consistency is

Algorithm 1 Time-independent top-K ranking

Input K: number of items to be included in the top list,
N: total number of items,
T: total number of comparisons, and
$\mathbf{A}^{(t)} = \left[a_{ij}^{(t)}\right]_{1\leq i,j\leq N}$: RPCMs for $t = 1, 2, \ldots, T$.

Define comparison graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ characterized by the vertex and edge sets $\mathcal{V}$ and $\mathcal{E}$, respectively.

Compute edge weights $\mathbf{W} = [w_{ij}]_{1\leq i,j\leq N}$ on $\mathcal{G}$ using

$$w_{ij} = \frac{\sum\limits_{t=1}^{T} \mathbb{1}(\log a_{ij}^{(t)})}{\sum\limits_{t=1}^{T} \mathbb{1}(\log a_{ij}^{(t)}) + \sum\limits_{t=1}^{T} \mathbb{1}(\log a_{ji}^{(t)})}$$

and

$$w_{ij} = \frac{\sum\limits_{t=1}^{T} (\log a_{ij}^{(t)})^{+}}{\sum\limits_{t=1}^{T} (\log a_{ij}^{(t)})^{+} + \sum\limits_{t=1}^{T} (\log a_{ji}^{(t)})^{+}}$$

for the discrete and continuous BTL models, respectively.

Compute $\deg_{\max}^{+}$ of $\mathcal{G}$.

Compute time-independent transition matrix $\mathbf{P} = [p_{ij}]_{1\leq i,j\leq N}$, where

$$p_{ij} = \begin{cases} \dfrac{1}{\deg_{\max}^{+}} \left(w_{ij}\right), & \text{if } (i,j) \in \mathcal{E}, \\ 1 - \dfrac{1}{\deg_{\max}^{+}} \left(\sum\limits_{k:(i,k)\in\mathcal{E}} w_{ik} \right), & \text{if } i = j, \\ 0, & \text{if } (i,j) \notin \mathcal{E}. \end{cases}$$

Compute $\boldsymbol{\pi} = [\pi_1, \pi_2, \ldots, \pi_N]^{\mathsf{T}}$, the principal left eigenvector of $\mathbf{P}$.

Output the list of items acquiring the K largest entries of $\boldsymbol{\pi}$.

defined according to number of changes that may occur in consecutive periods of the sequence. As an application of such an approach to consistency, a trader could be considered who re-balances his/her portfolio comprised of K assets every τ days using the above criterion. For such a trader, it is then often profitable to maintain a portfolio which is subject to the least number of assignment changes through consecutive periods of time. However, the problem of optimizing such a portfolio becomes more challenging when one deals with a relatively short time horizon τ and a substantially small values of K, in which case employment of the above-mentioned criterion can prove useful.

3.2.1 Shape-constrained Polyadic Decomposition

Let us now assume that the time period under study is partitioned into M sub-periods such that each sub-period m contains $T^{(m)}$ temporal steps. Hence, $T = \sum_{m=1}^{M} T^{(m)}$ would represent the total number of the temporal steps. The RPCT is then formed by populating the frontal slices of a third-order tensor $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{N\times N\times T}$ by T number of RPCMs. Although $\boldsymbol{\mathcal{A}}$ can potentially reflect the existing spatio-temporal interactions among the items, this tensor is often too noisy to delineate the most significant relations. However, this impediment can effectively be dispensed by smoothing $\boldsymbol{\mathcal{A}}$ out and encapsulating its most prominent information.

To smooth out the advantage tensor $\boldsymbol{\mathcal{A}}$, we formulate the following constrained polyadic decomposition model

$$\min_{\mathbf{x}_r,\mathbf{y}_r,\mathbf{z}_r} \left\| \boldsymbol{\mathcal{A}} - \sum_{r=1}^{R} \mathbf{x}_r \circ \mathbf{y}_r \circ \mathbf{z}_r \right\|_F \tag{3.8a}$$

$$\text{s.t.} \left(\sum_{r=1}^{R} z_{r,t} \left(\mathbf{x}_r \circ \mathbf{y}_r \right) \right) * \left(\sum_{r=1}^{R} z_{r,t} \left(\mathbf{x}_r \circ \mathbf{y}_r \right)^{\intercal} \right) = \mathbf{1}_N \,, \quad t = 1, \ldots, T \,, \tag{3.8b}$$

$$\sum_{r=1}^{R} z_{r,t} \left(\mathbf{x}_r \circ \mathbf{y}_r \right) > \mathbf{0}_N, \qquad t = 1, \ldots, T \,, \tag{3.8c}$$

$$\mathbf{x}_r, \mathbf{y}_r \in \mathbb{R}^N \,, \qquad r = 1, \ldots, R \,, \tag{3.8d}$$

$$\mathbf{z}_r \in \mathbb{R}^T \,, \qquad r = 1, \ldots, R \,, \tag{3.8e}$$

where $\mathbf{1}_N$ and $\mathbf{0}_N$ denote $N \times N$ all ones and all zeros matrices, respectively, and the inequality in (3.8c) is evaluated component-wise. The objective function (3.8a) minimizes the error of approximation w.r.t. Frobenius norm. Further, the first and second sets of constraints ensure that every frontal slice of $\hat{\boldsymbol{\mathcal{A}}}$ would have the reciprocal structure. The model (3.8) is then solved for different values of R, and a minimal estimate of R (if it exists) is considered to be the rank of $\hat{\boldsymbol{\mathcal{A}}}$. It is noted that $\hat{\boldsymbol{\mathcal{A}}}$ becomes a relatively accurate approximate of $\boldsymbol{\mathcal{A}}$ for sufficiently large values of R. However, for the smoothing purposes, one may only opt for possibly small estimates of R to ensure that $\hat{\boldsymbol{\mathcal{A}}}$ captures the most significant information embedded in $\boldsymbol{\mathcal{A}}$.

It is evident that the proposed model (3.8) entails a set of strictly non-negative constraints which impose additional complexities. Instead, we formulate an alternative model that is equivalent to (3.8) with a relatively lower complexity. This new formulation decomposes the transformed $\log(\boldsymbol{\mathcal{A}})$, and it reads as follows:

$$\min_{\tilde{\mathbf{x}}_r, \tilde{\mathbf{y}}_r, \tilde{\mathbf{z}}_r} \left\| \log(\boldsymbol{\mathcal{A}}) - \sum_{r=1}^{R} \tilde{\mathbf{x}}_r \circ \tilde{\mathbf{y}}_r \circ \tilde{\mathbf{z}}_r \right\|_F \tag{3.9a}$$

$$\text{s.t.} \sum_{r=1}^{R} \tilde{z}_{r,t} \left(\tilde{\mathbf{x}}_r \circ \tilde{\mathbf{y}}_r + (\tilde{\mathbf{x}}_r \circ \tilde{\mathbf{y}}_r)^{\intercal} \right) = \mathbf{0}_N \,, \quad t = 1, \ldots, T \,, \tag{3.9b}$$

$$\tilde{\mathbf{x}}_r, \tilde{\mathbf{y}}_r \in \mathbb{R}^N \,, \qquad r = 1, \ldots, R \,, \tag{3.9c}$$

$$\tilde{\mathbf{z}} \in \mathbb{R}^T \,, \qquad r = 1, \ldots, R \,. \tag{3.9d}$$

In this model, $\log(\boldsymbol{\mathcal{A}})$ is taken element-wise, and the first set of constraints ensures that every frontal slice of the estimate tensor $\tilde{\boldsymbol{\mathcal{A}}}$ defined by

$$\log(\boldsymbol{\mathcal{A}}) \approx \tilde{\boldsymbol{\mathcal{A}}} = \sum_{r=1}^{R} \tilde{\mathbf{x}}_r \circ \tilde{\mathbf{y}}_r \circ \tilde{\mathbf{z}}_r \,, \tag{3.10}$$

is a skew-symmetric matrix. By solving model (3.9) and obtaining $\tilde{\boldsymbol{\mathcal{A}}}$, one can then reconstruct $\hat{\boldsymbol{\mathcal{A}}} = e^{\tilde{\boldsymbol{\mathcal{A}}}}$ and use it within the scope of top-K ranking procedures.

We point out that the constrained optimization model (3.9) could possibly be solved employing *Tensorlab* (Vervliet, Debals, and De Lathauwer, 2016), which facilitates the process of imposing

constraints on the frontal slices of $\tilde{\boldsymbol{\mathcal{A}}}$. Yet, a more practical approach to solve (3.9), perused in this work, is to first relax the set of constraints (3.9b), and then employ the standard techniques implemented in *Tensorlab* to derive $\tilde{\boldsymbol{\mathcal{A}}}_{rel}$. Thereafter, each frontal slice of $\tilde{\boldsymbol{\mathcal{A}}}_{rel}$ is converted into a skew-symmetric matrix in order to re-construct $\tilde{\boldsymbol{\mathcal{A}}}$. For instance, the t-th frontal slice of $\tilde{\boldsymbol{\mathcal{A}}}_{rel}$, say $\tilde{\mathbf{A}}_{rel}^{(t)}$, can be converted into a skew-symmetric matrix by means of the following equation

$$\tilde{\mathbf{A}}^{(t)} = \frac{1}{2}\left(\tilde{\mathbf{A}}_{rel}^{(t)} - \left(\tilde{\mathbf{A}}_{rel}^{(t)}\right)^{\intercal}\right), \tag{3.11}$$

where $\tilde{\mathbf{A}}^{(t)}$ denotes the t-th frontal slice of $\tilde{\boldsymbol{\mathcal{A}}}$.

Applications of the smoothed advantage tensor $\hat{\boldsymbol{\mathcal{A}}}$ to top-K ranking could potentially lead to time-homogeneous results. For instance, in the case where top-K items constitute a portfolio's assets and under the assumption that re-balancing takes place on a monthly basis, the corresponding *smoothed RPCMs* (i.e. frontal slices of $\hat{\boldsymbol{\mathcal{A}}}$) can first be extracted for each month and then the top-K rankings may be determined for any particular month by applying the rank centrality algorithm. Further details of the time-homogeneous top-K ranking are outlined in Algorithm 2.

We also note that the advantage tensor $\boldsymbol{\mathcal{A}}$ is more likely to be incomplete in many practical situations, which in turn imposes some analytical as well as computational challenges. However, such impediments could essentially be surmounted since the problem of decomposing incomplete tensors is well-studied in the literature; see (Cichocki et al., 2016; Cichocki et al., 2017; Huang, Sidiropoulos, and Liavas, 2016; Sidiropoulos et al., 2017; Vervliet and De Lathauwer, 2016; Vervliet, Debals, and De Lathauwer, 2017; Vervliet, Debals, Sorber, and De Lathauwer, 2014) for more details.

3.3 Computational Results

In this section, we present our computations which concern the performance evaluation for the proposed schemes. In order to complement our assessment, we retrieved the S&P500 daily adjusted

Algorithm 2 Time-homogeneous top-K ranking

Input K: number of items to be included in the top list,
N: total number of items,
T: total number of comparisons, and
$\hat{\mathbf{A}}^{(t)} = \left[\hat{a}_{ij}^{(t)}\right]_{1\le i,j\le N}$: smoothed RPCMs for $t = 1, 2, \ldots, T$.

Define comparison graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ characterized by the vertex and edge sets $\mathcal{V}$ and $\mathcal{E}$, respectively.

Compute edge weights $\mathbf{W} = [w_{ij}]_{1\le i,j\le N}$ on $\mathcal{G}$ using

$$w_{ij} = \frac{\sum\limits_{t=1}^{T} \mathbb{1}(\log \hat{a}_{ij}^{(t)})}{\sum\limits_{t=1}^{T} \mathbb{1}(\log \hat{a}_{ij}^{(t)}) + \sum\limits_{t=1}^{T} \mathbb{1}(\log \hat{a}_{ji}^{(t)})}$$

and

$$w_{ij} = \frac{\sum\limits_{t=1}^{T} (\log \hat{a}_{ij}^{(t)})^{+}}{\sum\limits_{t=1}^{T} (\log \hat{a}_{ij}^{(t)})^{+} + \sum\limits_{t=1}^{T} (\log \hat{a}_{ji}^{(t)})^{+}}$$

for the discrete and continuous BTL models, respectively.

Compute $\deg_{\max}^{+}$ of $\mathcal{G}$.

Compute time-homogeneous transition matrix $\mathbf{P} = [p_{ij}]_{1\le i,j\le N}$, where

$$p_{ij} = \begin{cases} \dfrac{1}{\deg_{\max}^{+}}\,(w_{ij})\,, & \text{if } (i,j) \in \mathcal{E}\,, \\ 1 - \dfrac{1}{\deg_{\max}^{+}} \left(\sum\limits_{k:(i,k)\in\mathcal{E}} w_{ik} \right), & \text{if } i = j\,, \\ 0\,, & \text{if } (i,j) \notin \mathcal{E}\,. \end{cases}$$

Compute $\boldsymbol{\pi} = [\pi_1, \pi_2, \ldots, \pi_N]^{\intercal}$, the principal left eigenvector of $\mathbf{P}$.

Output the list of items acquiring the K largest entries of $\boldsymbol{\pi}$.

closing prices [7] for the period of time commenced from January 2007 to December 2019. This dataset contains historical prices for $N = 503$ assets which were traded for a total of $T = 3271$ trading days during $M = 156$ months.

[7] Wharton Research Data Service (WRDS) was used to retrieve the historical prices. The datasets are available at `http://wrds.wharton.upenn.edu`.

3.3.1 Evaluation Criteria

In many applications, it is often desirable to obtain an appropriate value for K and a partitioning configuration for M sub-periods, so that the sequence of top-K items would be highly consistent. We define two variants of measures for consistency of the top-K items by resorting to applications of the Hamming distance. Let us recall that the Hamming distance between two equal-sized binary vectors is defined to be the number of positions for which the entries differ. One may then define a measure on consistency of any particular top-K list w.r.t. its preceding sub-period as follows:

$$\mu(K,m) = 1 - \frac{1}{2K} \left\| \mathbf{L}_K^{(m)} - \mathbf{L}_K^{(m-1)} \right\|_H . \tag{3.12}$$

Here, $\mathbf{L}_K^{(m)} \in \{0,1\}^N$ denotes the top-K list of the m-th sub-period for $m = 1, \ldots, M$, whose i-th element is one if the item i appears in the list. Note that the Hamming norm $\|\cdot\|_H$ is computed as the quantity of non-zero entries contained in the binary vector. For any particular sub-period, say m, we refer to the quantity $\mu(K,m)$ as *consecutive consistency measure*, in that, it measures the existing similarity between top-K lists of two consecutive sub-periods m and $m-1$. In turn, the criterion presented above can be employed to define an *overall consistency measure* for a sequence of top-K rankings as follows:

$$\bar{\mu}(K) = \frac{1}{M-1} \sum_{m=2}^{M} \mu(K,m) . \tag{3.13}$$

3.3.2 Time-homogeneous vs. Time-independent Top-K Ranking

In order to construct the RPCT $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{503\times503\times3271}$, the RPCMs $\mathbf{A}^{(t)}$ were first computed for every trading day $t = 1, 2, \ldots, T^{(m)}$, within each individual month $m = 1, 2, \ldots, M$ by using equation (3.1). Then, they were employed to obtain the time-independent and time-homogeneous sequences of the top-K items.

The rank was estimated by solving (3.9) and examining the corresponding error of approximation for every $R = 1, 2, \ldots, 10$. Our experiments revealed that, $R = 1$ yielding a 2.83% error

of approximation was a suitable candidate for the rank of $\hat{\boldsymbol{\mathcal{A}}}$ since the underlying structure of $\boldsymbol{\mathcal{A}}$ was greatly recovered at the above-mentioned error level, while the noise contaminating $\boldsymbol{\mathcal{A}}$ was reduced substantially. It is worthwhile to mention that our experiments with larger values of R led to errors of approximation which were all less than 2.83%. For instance, the original tensor $\boldsymbol{\mathcal{A}}$ was almost fully recovered by choosing $R = 10$ (the error of approximation was obtained to be 0.00%). However, such a near-perfect recovery would not be desirable for smoothing purposes as it implies a compromise made on the effectiveness of the noise reduction procedure. Accordingly, $R = 1$ was selected to ensure that $\hat{\boldsymbol{\mathcal{A}}}$ preserves informative aspects of $\boldsymbol{\mathcal{A}}$, while smoothing out somewhat less significant information.

Figure 3.1 and Figure 3.2 compare the performances of time-independent and time-homogeneous top-K rankings obtained under both discrete and continuous BTL models, respectively. Observe that the time-homogeneous rankings gain higher overall consistency measure $\bar{\mu}(K)$ under both BTL models for every $K = 1, 2, \ldots, 100$. Furthermore, Figure 3.3 reveals that the time-homogeneous top-K ranking yields higher point-wise overall consistency measure under the continuous BTL model, almost everywhere (with the exception of $K = 8$). This indicates that extending the win/loss nature of the discrete BTL model to a continuous setting leads to more consistent results.

3.3.3 Optimal Investment Strategies

Let us reconsider the portfolio optimization problem with the goal to allocate K assets, selected from the list of S&P500, so as to maximize the overall capital return in a given time period. Assume without the loss of generality that re-balancing occurs at the end of each month $m = 1, 2, \ldots, M$. In other words, this scheme corresponds to a fixed partitioning configuration of M sub-periods.

We discussed before that $\boldsymbol{\mathcal{A}}$ contains the outcomes for pairwise comparisons conducted on advantages of the items at every temporal step, i.e. $t \in \{1, 2, \ldots, T\}$. Hence, the top-$K$ items obtained for a certain month, can potentially characterize a profitable K-subset of the S&P500 assets during that specific month. Consequently, one may be able to deal with the objective of

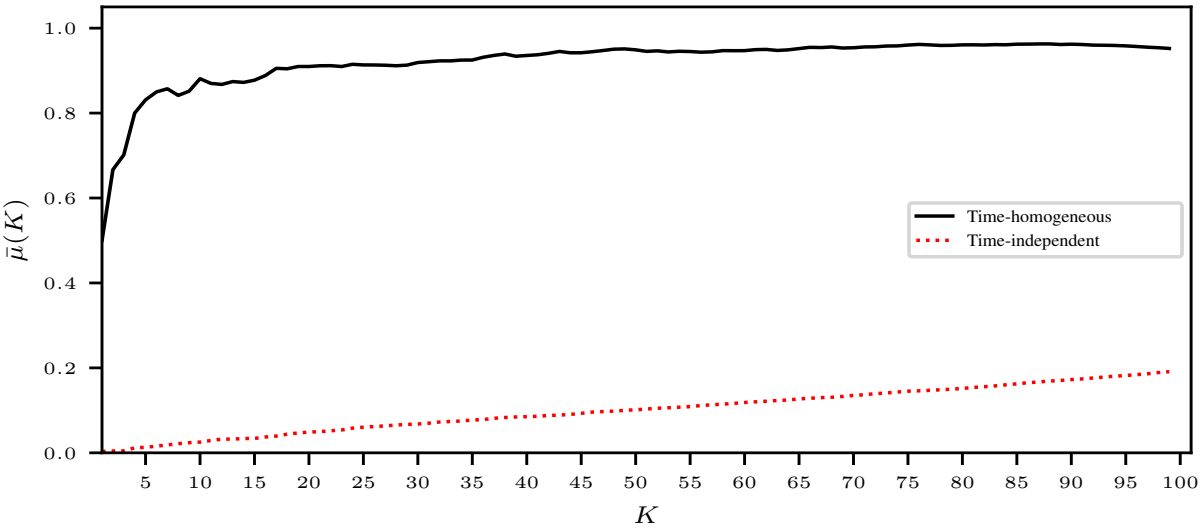

Figure 3.1: Time-independent vs. time-homogeneous top-K rankings under discrete BTL.

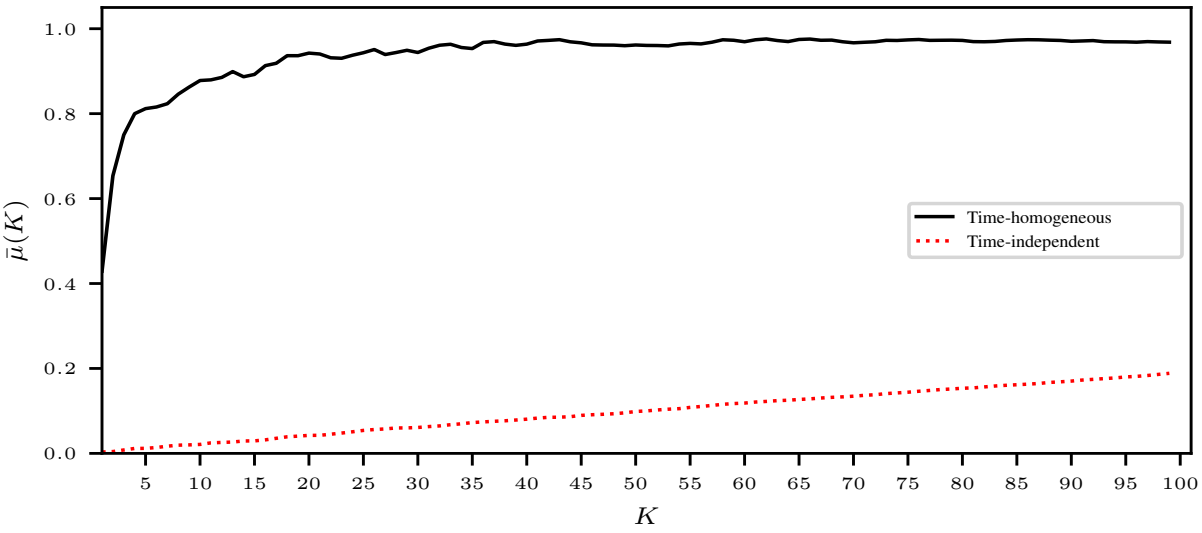

Figure 3.2: Time-independent vs. time-homogeneous top-K rankings under continuous BTL.

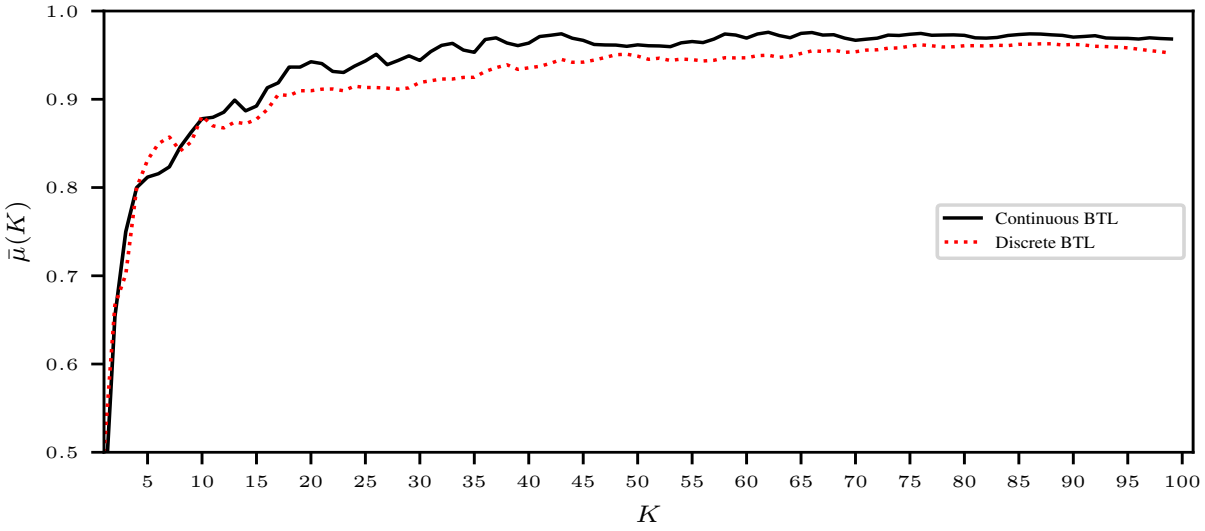

Figure 3.3: Time-homogeneous top-K rankings; discrete vs. continuous BTL models.

maximizing the portfolio's overall return in an implicit way by considering K to be the major contributor to re-balancing costs of the portfolio. Thus, the choice of small values for K can substantially reduce the transaction costs incurred when the portfolio is re-balanced.

We obtain such a desirable K^* by first solving the following optimization problem

$$\kappa^\star = \underset{1\leq\kappa\leq N}{\operatorname{argmax}} \left| \frac{\partial^2}{\partial\kappa^2}\, \bar{\mu}(\kappa) \right|, \tag{3.14}$$

and then rounding $\kappa^\star$ to its nearest integer. In fact, this procedure attempts to find a certain K^* such that for $K > K^*$, $\bar{\mu}(K)$ would typically increase with a less significant rate. It is pointed out that equation (3.14) can be solved by resorting to the use of standard techniques that are available to approximate the Hamming norm; see (Clifford and Starikovskaya, 2016; Cormode, Datar, Indyk, and Muthukrishnan, 2002) for more details.

Figure 3.4 depicts an illustration of $\bar{\mu}(K)$ evaluated at every $K = 1, 2, \ldots, N$, using the time-homogeneous top-K ranking under the continuous BTL model. It is evident that the rates of increase in $\bar{\mu}(K)$ are less significant when $K > 10$. Therefore, $K^* = 10$ can be considered as an estimate for the minimal K that solves (3.14), for which the overall consistency measure evaluates

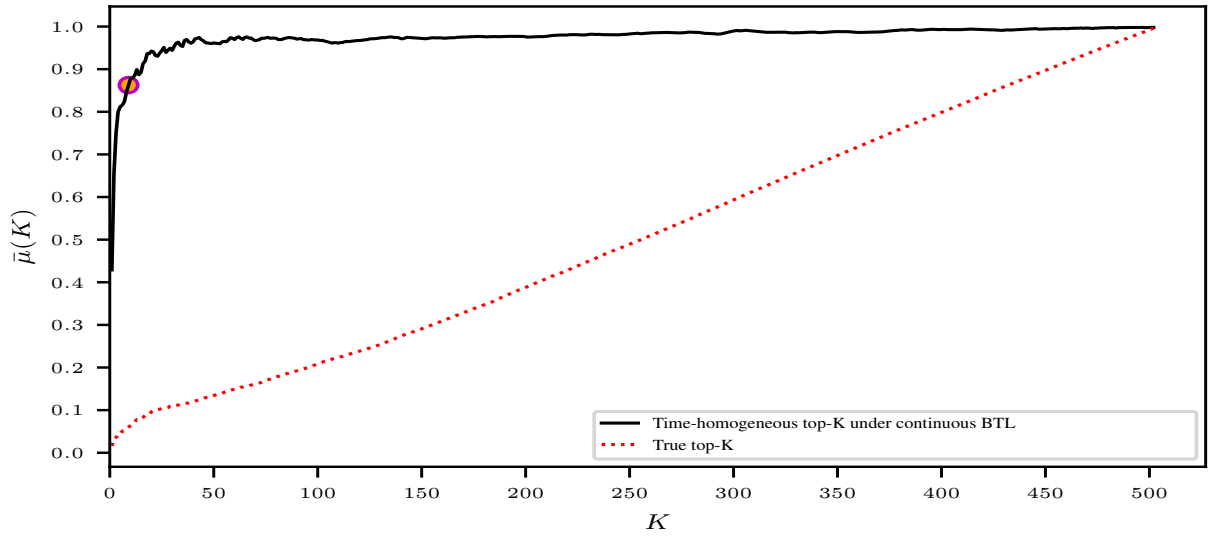

Figure 3.4: Overall consistency measures for time-homogeneous and true top-K rankings.

to $\bar{\mu}(K^*) = 0.87$.

Figure 3.4 further illustrates the overall consistency measures obtained under the *true* top-K rankings. The true top-K ranking at each month is defined as a set containing K assets with the highest expected daily returns. It is seen that $\bar{\mu}(K)$ would approach one under both time-homogeneous and true ranking schemes when K is large enough. In contrast, $\bar{\mu}(K)$ exhibits a linear growth behavior under the true rankings, whereas the time-homogeneous model reaches significantly larger values of $\bar{\mu}(K)$ relatively quickly.

Next, we obtain the consecutive consistency measures for both of these models when K is fixed at $K = K^*$. As shown in Figure 3.5, the majority of the true top-K assets are subject to change when the portfolio is re-balanced at the end of each month. On the contrary, a smaller number of amendments are required when adapting the time-homogeneous top-K ranking scheme.

Also, Figures 3.6 to 3.11 are presented to provide more insight on how the profitability of top-K sequences are influenced by transaction costs. It is evident as compared from these figures that the time-homogeneous model becomes a competitive model when transaction costs are further taken into account. Note that the term $\bar{R}(K, m)$ represents the average return of the top-K assets at

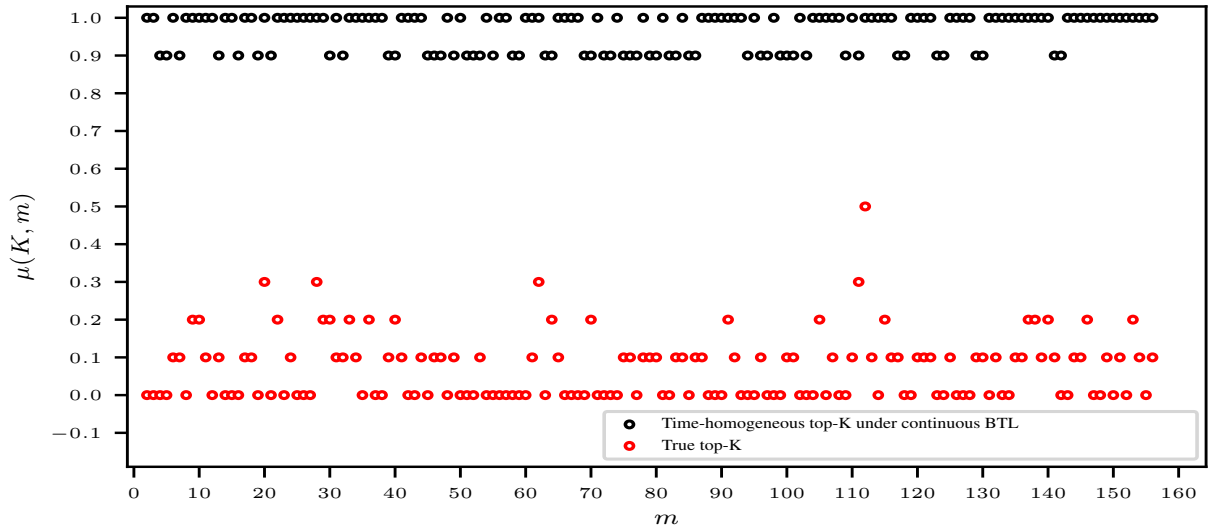

Figure 3.5: Scatter plots for consecutive consistency measures evaluated at $K = K^*$.

a given month m, while $e(K, m)$ denotes the absolute difference between the m-th month average returns of the true and time-homogeneous ranking models.

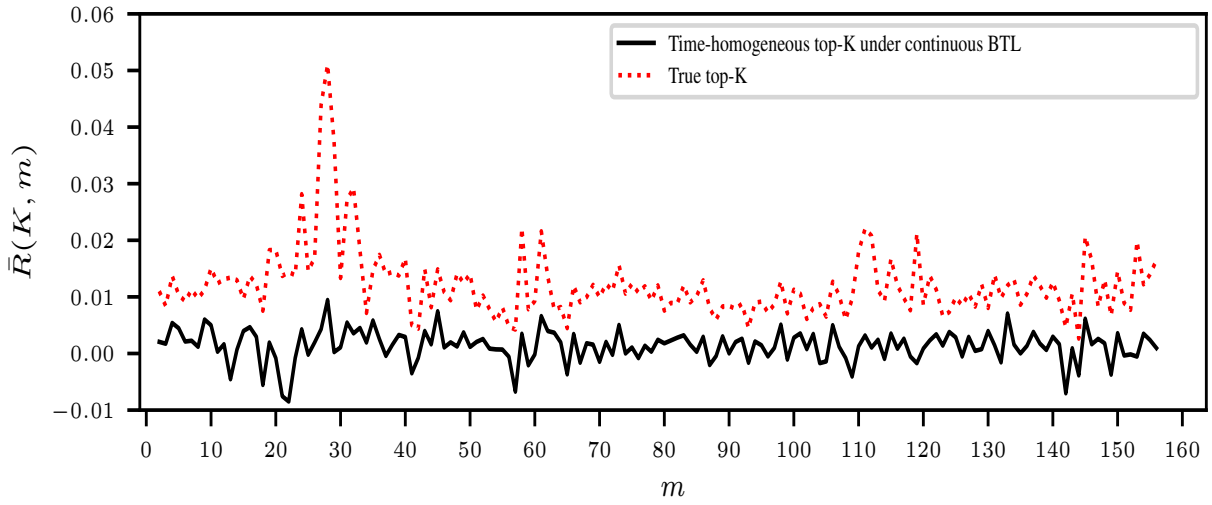

Figure 3.6: Monthly average returns with no transaction fee.

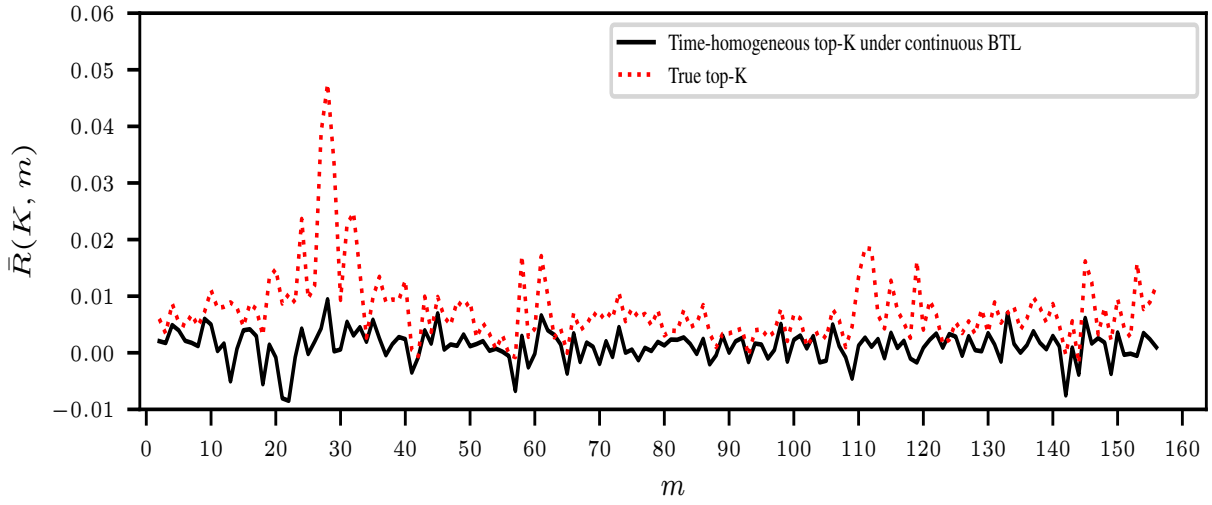

Figure 3.7: Monthly average returns with fixed 0.5% transaction fee.

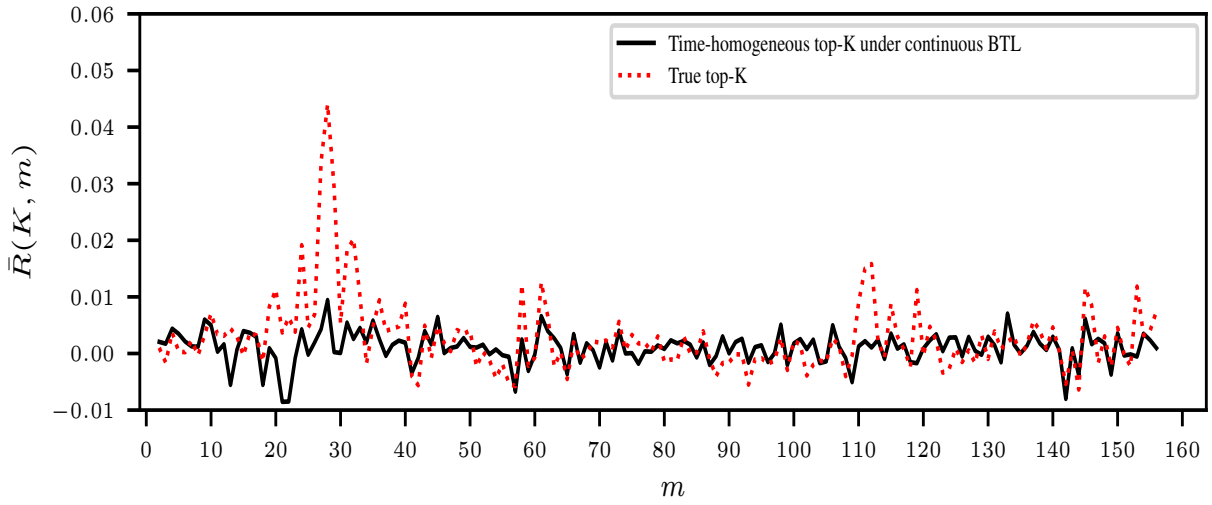

Figure 3.8: Monthly average returns with fixed 1.0% transaction fee.

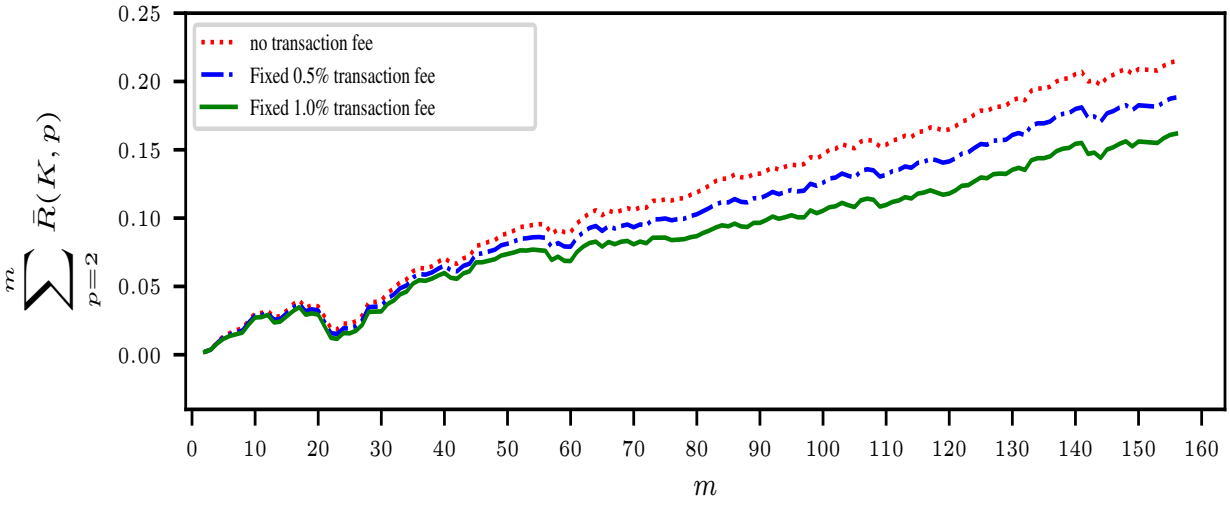

Figure 3.9: Cumulative monthly average returns for time-homogeneous top-K.

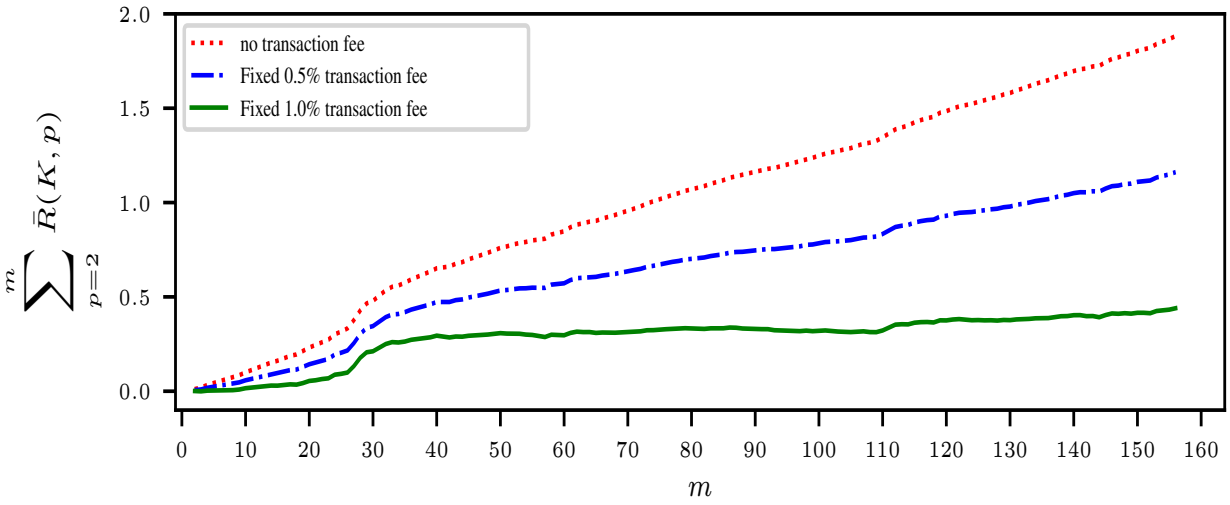

Figure 3.10: Cumulative monthly average returns for true top-K ranking.

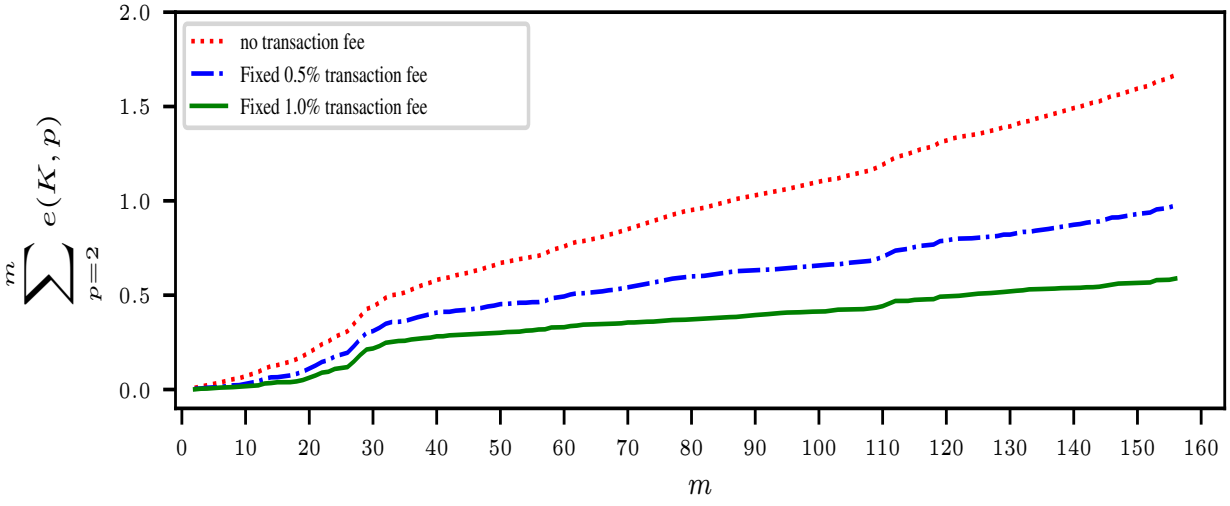

Figure 3.11: Cumulative absolute errors for monthly average returns.

Chapter 4

Conclusion

In this dissertation, we first axiomitized the pairwise comparative judgments and demonstrated that the so-called reciprocal pairwise comparison matrices were the only members of the family of pairwise comparison matrices that would satisfy the formulated axiom. Subsequently, various algebraic properties of the reciprocal pairwise comparison matrices such as their rank, trace, powers, eigenvalues, etc, were discussed, and it was shown that the reciprocal pairwise comparison matrices are suitable structures to model the procedures that underlie agents' thought processes and their mechanisms of casting judgments. Thereafter, the sensitivity of the eigenvalues of the reciprocal pairwise comparison matrices to perturbations in their entries were investigated, and it was established that the largest eigenvalues of the considered matrices would be robust w.r.t. such perturbations.

Next, we defined a special class of spatio-temporal tensors by extending the reciprocal pairwise comparison matrices to a tensor domain, which in turn enabled us to effectively embed the collective judgment of the agents throughout time. An inconsistency function was then developed based on the largest eigenvalues of the frontal slices of the constructed tensors, in order to examine the consistency of the judgments cast by the agents throughout time. A relationship between the mentioned inconsistency function and the rank-1 estimates of the constructed tensors was further established, upon which the financial chaos index was defined. Several properties of the financial

chaos index were also studied, including its robustness for estimating the market volatility, which enabled us to reliably perform the tasks of market regime analysis as well as its segmentation by resorting to the applications of the proposed index.

Furthermore, we clarified the connection between the equity and option markets by studying the relationship between the financial chaos index (as a measure of market's actual volatility) and the VIX (as a measure of market's implied volatility). More so, the causal relations existing among the financial chaos index, VIX, and various types of news-based policy uncertainty indexes were exploited using the applications of the extended Granger causality. By further defining an index to capture the overall behavior of the assets' fundamentals, referred to as the market fundamentals index, the analysis of the causal relations was expanded to encompass the latter index as an additional element alongside the already considered indexes (the financial chaos index, VIX and news-based policy uncertainty indexes), where the causal relations were identified using a 1-step blockwise Granger causality approach.

Subsequently, the results obtained by analyzing the causal relations mentioned above, were employed in testing the efficient market hypothesis. It was shown that the U.S. stock market for the time frame January 1990-December 2019, for either of the monthly or quarterly time frequencies, was weakly, potently and semi-strongly efficient at 1% level of significance. On the contrary, at the same level of significance, the stock market during this time period turned out to be weakly, potently and semi-strongly inefficient as per daily time frequency.

Our further contribution in this research was to introduce the notion of pairwise advantage and developed a continuous extension to the discrete BTL model that empowers one to approach the problem of rank aggregation from pairwise comparisons more effectively by incorporating the degrees of preferences which exist among the items compared. Pairwise advantages of the items were then utilized to construct advantage matrices corresponding to the temporal steps, which were subsequently considered as inputs to the rank centrality algorithm in order to identify sequences of time-independent top-K items. To obtain a set of time-homogeneous top-K ranked items, we then proposed a tensorial formulation of the pairwise rank aggregation problem and constructed

the so-called reciprocal pairwise comparison tensors which were smoothed out by solving a shape-constrained polyadic decomposition model. It was shown that applications of the proposed time-homogeneous ranking scheme yielded consistent sequences of top-K items when applied to the cardinality-constrained portfolio optimization problem of S&P500.

Bibliography

Ailon, Nir (2012). "An active learning algorithm for ranking from pairwise preferences with an almost optimal query complexity". In: *Journal of Machine Learning Research* 13.Jan, pp. 137–164.

Ailon, Nir, Moses Charikar, and Alantha Newman (2008). "Aggregating inconsistent information: Ranking and clustering". In: *Journal of the ACM (JACM)* 55.5, p. 23.

Alon, Noga (2006). "Ranking tournaments". In: *SIAM Journal on Discrete Mathematics* 20.1, pp. 137–142.

Alvo, Mayer and LH Philip (2014). *Statistical methods for ranking data*. Springer.

Amblard, Pierre-Olivier and Olivier JJ Michel (2013). "The relation between Granger causality and directed information theory: A review." In: *Entropy,* 15.1, pp. 113–143.

Andrews, Donald WK (1991). "Heteroskedasticity and autocorrelation consistent covariance matrix estimation." In: *Econometrica: Journal of the Econometric Society,* 59.3, pp. 817–858.

Arlot, Sylvain, Alain Celisse, and Zaid Harchaoui (2019). "A kernel multiple change-point algorithm via model selection." In: *Journal of Machine Learning Research,* 20.162, pp. 1–56.

Arrow, Kenneth J (2012). *Social Choice and Individual Values*. Vol. 12. Yale university press.

Baker, Scott R, Nicholas Bloom, and Steven J Davis (2016). "Measuring economic policy uncertainty." In: *The Quarterly Journal of Economics,* 131.4, pp. 1593–1636.

Baker, Scott R, Nicholas Bloom, Steven J Davis, and Kyle J Kost (2019). *Policy News and Stock Market Volatility.* Working Paper 25720. National Bureau of Economic Research.

Barnett, Lionel and Terry Bossomaier (2012). “Transfer entropy as a log-likelihood ratio.” In: *Physical Review Letters,* 109.13, p. 138105.

Barrat, Alain, Marc Barthelemy, Romualdo Pastor-Satorras, and Alessandro Vespignani (2004). “The architecture of complex weighted networks.” In: *Proceedings of the National Academy of Sciences,* 101.11, pp. 3747–3752.

Basu, Shantanu, M Gil, and Sayantan Auddy (2015). “The MLP distribution: a modified lognormal power-law model for the stellar initial mass function.” In: *Monthly Notices of the Royal Astronomical Society,* 449.3, pp. 2413–2420.

Basu, Shantanu and CE Jones (2004). “On the power-law tail in the mass function of protostellar condensations and stars.” In: *Monthly Notices of the Royal Astronomical Society,* 347.3, pp. L47–L51.

Behrendt, Simon, Thomas Dimpfl, Franziska J Peter, and David J Zimmermann (2019). “RTransferEntropy—Quantifying information flow between different time series using effective transfer entropy.” In: *SoftwareX,* 10, p. 100265.

Ben-David, Itzhak, John R Graham, and Campbell R Harvey (2013). “Managerial miscalibration.” In: *The Quarterly Journal of Economics,* 128.4, pp. 1547–1584.

Bland, J Martin and Douglas G Altman (1995). “Multiple significance tests: the Bonferroni method.” In: *The British Medical Journal,* 310.6973, p. 170.

Bradley, Ralph Allan (1954). “Rank analysis of incomplete block designs: II. Additional tables for the method of paired comparisons”. In: *Biometrika* 41.3-4, pp. 502–537.

— (1955). “Rank analysis of incomplete block designs: III. Some large-sample results on estimation and power for a method of paired comparisons”. In: *Biometrika* 42.3/4, pp. 450–470.

Bradley, Ralph Allan and Milton E Terry (1952). “Rank analysis of incomplete block designs: I. The method of paired comparisons”. In: *Biometrika* 39.3/4, pp. 324–345.

Brandes, Ulrik (2001). “A faster algorithm for betweenness centrality.” In: *Journal of Mathematical Sociology,* 25.2, pp. 163–177.

Brin, Sergey and Lawrence Page (1998). “The anatomy of a large-scale hypertextual web search engine.” In: *Computer Networks and ISDN Systems,* 30, pp. 1–30.

Brunelli, Matteo (2014). *Introduction to the Analytic Hierarchy Process.* Springer.

Campbell, John Y (2000). "Asset pricing at the millennium." In: *The Journal of Finance,* 55.4, pp. 1515–1567.

Campbell, John Y and John H Cochrane (1999). "By force of habit: A consumption-based explanation of aggregate stock market behavior." In: *Journal of Political Economy,* 107.2, pp. 205–251.

Cattelan, Manuela (2012). "Models for paired comparison data: A review with emphasis on dependent data". In: *Statistical Science*, pp. 412–433.

Chen, Xi, Paul N Bennett, Kevyn Collins-Thompson, and Eric Horvitz (2013). "Pairwise ranking aggregation in a crowdsourced setting". In: *Proceedings of the sixth ACM international conference on Web search and data mining*. ACM, pp. 193–202.

Chen, Xi, Sivakanth Gopi, Jieming Mao, and Jon Schneider (2017). "Competitive analysis of the top-K ranking problem". In: *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, pp. 1245–1264.

Chen, Xi, Yuanzhi Li, and Jieming Mao (2018). "A nearly instance optimal algorithm for top-K ranking under the multinomial logit model". In: *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, pp. 2504–2522.

Chen, Yuxin, Jianqing Fan, Cong Ma, and Kaizheng Wang (2017). "Spectral Method and Regularized MLE Are Both Optimal for Top-K Ranking". In: *arXiv preprint arXiv:1707.09971*.

Chen, Yuxin and Changho Suh (2015). "Spectral MLE: Top-K rank aggregation from pairwise comparisons". In: *International Conference on Machine Learning*, pp. 371–380.

Cichocki, Andrzej, Namgil Lee, Ivan Oseledets, Anh-Huy Phan, Qibin Zhao, and Danilo P Mandic (2016). "Tensor networks for dimensionality reduction and large-scale optimization: Part 1. Low-rank tensor decompositions". In: *Foundations and Trends® in Machine Learning* 9.4-5, pp. 249–429.

Cichocki, Andrzej, Anh-Huy Phan, Qibin Zhao, Namgil Lee, Ivan Oseledets, Masashi Sugiyama, and Danilo P Mandic (2017). "Tensor networks for dimensionality reduction and large-scale

optimization: Part 2. Applications and future perspectives". In: *Foundations and Trends® in Machine Learning* 9.6, pp. 431–673.

Clifford, Raphaël and Tatiana Starikovskaya (2016). "Approximate Hamming distance in a stream". In: *arXiv preprint arXiv:1602.07241*.

Cormode, Graham, Mayur Datar, Piotr Indyk, and S Muthukrishnan (2002). "Comparing data streams using hamming norms (how to zero in)". In: *VLDB'02: Proceedings of the 28th International Conference on Very Large Databases*. Elsevier, pp. 335–345.

Cover, Thomas M and Joy A Thomas (2012). *Elements of Information Theory*. John Wiley & Sons.

Davis, Steven J and Cristhian Seminario (2019). "Diagnosing the stock market reaction to Trump's surprise election victory." In:

Dean, Roger T and William TM Dunsmuir (2016). "Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models." In: *Behavior Research Methods,* 48.2, pp. 783–802.

Degutis, Augustas and Lina Novickytė (2014). "The efficient market hypothesis: A critical review of literature and methodology." In: *Ekonomika,* 93, pp. 7–23.

Delgado-Bonal, Alfonso and Alexander Marshak (2019). "Approximate entropy and sample entropy: A comprehensive tutorial." In: *Entropy,* 21.6, p. 541.

Desobry, Frédéric, Manuel Davy, and Christian Doncarli (2005). "An online kernel change detection algorithm." In: *IEEE Transactions on Signal Processing,* 53.8, pp. 2961–2974.

Drechsler, Itamar and Amir Yaron (2011). "What's vol got to do with it." In: *The Review of Financial Studies,* 24.1, pp. 1–45.

Dueck, Delbert (2009). "Affinity propagation: Clustering data by passing messages". PhD thesis. University of Toronto.

Everitt, Brian S (1996). "An introduction to finite mixture distributions." In: *Statistical Methods in Medical Research* 5.2, pp. 107–127.

Faes, Luca and Alberto Porta (2014). "Conditional entropy-based evaluation of information dynamics in physiological systems." In: *Directed Information Measures in Neuroscience*. Springer, pp. 61–86.

Faes, Luca, Alberto Porta, and Giandomenico Nollo (2015). "Information decomposition in bivariate systems: theory and application to cardiorespiratory dynamics." In: *Entropy,* 17.1, pp. 277–303.

Faes, Luca, Alberto Porta, Giandomenico Nollo, and Michal Javorka (2017). "Information decomposition in multivariate systems: definitions, implementation and application to cardiovascular networks." In: *Entropy,* 19.1, p. 5.

Faes, Luca and K Sameshima (2014). "Assessing connectivity in the presence of instantaneous causality." In: *Methods in Brain Connectivity Inference Through Multivariate Time Series Analysis*. Taylor & Francis, pp. 87–112.

Fishburn, Peter C (1970). "Intransitive individual indifference and transitive majorities." In: *Econometrica: Journal of the Econometric Society,* 38.3, pp. 482–489.

Fisher, Ronald Aylmer (1992). "Statistical methods for research workers." In: *Breakthroughs in Statistics*. Springer, pp. 66–70.

Forbes, Kristin J and Francis E Warnock (2012). "Capital flow waves: Surges, stops, flight, and retrenchment." In: *Journal of International Economics,* 88.2, pp. 235–251.

Frey, Brendan J and Delbert Dueck (2007). "Clustering by passing messages between data points." In: *Science,* 315.5814, pp. 972–976.

Gabriela ğiĠan, Alexandra (2015). "The efficient market hypothesis: Review of specialized literature and empirical research." In: *Procedia Economics and Finance,* 32, pp. 442–449.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy (2019). "Text as data." In: *Journal of Economic Literature,* 57.3, pp. 535–74.

Gibbard, Allan F (2014). "Intransitive social indifference and the Arrow dilemma." In: *Review of Economic Design,* 18.1, pp. 3–10.

Giglio, Stefano and Bryan Kelly (2018). "Excess volatility: Beyond discount rates." In: *The Quarterly Journal of Economics,* 133.1, pp. 71–127.

Gretton, Arthur and László Györfi (2008). "Nonparametric independence tests: Space partitioning and kernel approaches." In: *International Conference on Algorithmic Learning Theory*. Springer, pp. 183–198.

Gretton, Arthur and László Györfi (2010). "Consistent nonparametric tests of independence." In: *Journal of Machine Learning Research,* 11.46, pp. 1391–1423.

Hamilton, James Douglas (1994). *Time Series Analysis*. Vol. 2. Princeton University Press, Princeton, NJ.

Harchaoui, Zaid and Olivier Cappé (2007). "Retrospective mutiple change-point estimation with kernels." In: *2007 IEEE/SP 14th Workshop on Statistical Signal Processing*. IEEE, pp. 768–772.

Harchaoui, Zaid, Eric Moulines, and Francis R Bach (2009). "Kernel change-point analysis." In: *Advances in Neural Information Processing Systems*, pp. 609–616.

Harville, David A (2018). *Linear Models and the Relevant Distributions and Matrix Algebra*. CRC Press.

Hassan, Tarek A, Stephan Hollander, Laurence van Lent, and Ahmed Tahoun (2019). "Firm-level political risk: Measurement and effects." In: *The Quarterly Journal of Economics,* 134.4, pp. 2135–2202.

Heckel, Reinhard, Nihar B Shah, Kannan Ramchandran, and Martin J Wainwright (2016). "Active Ranking from Pairwise Comparisons and when Parametric Assumptions Don't Help". In: *arXiv preprint arXiv:1606.08842*.

Hobijn, Bart, Philip Hans Franses, and Marius Ooms (2004). "Generalizations of the KPSS-test for stationarity." In: *Statistica Neerlandica,* 58.4, pp. 483–502.

Huang, Kejun, Nicholas D Sidiropoulos, and Athanasios P Liavas (2016). "A flexible and efficient algorithmic framework for constrained matrix and tensor factorization". In: *IEEE Transactions on Signal Processing* 64.19, pp. 5052–5065.

Hwang, Woochang, Taehyong Kim, Murali Ramanathan, and Aidong Zhang (2008). "Bridging centrality: graph mining from element level to group level." In: *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 336–344.

Ishizaka, Alessio and Ashraf Labib (2009). "Analytic hierarchy process and expert choice: Benefits and limitations." In: *OR Insight,* 22.4, pp. 201–220.

— (2011). "Review of the main developments in the analytic hierarchy process." In: *Expert Systems With Applications.* 38.11, pp. 14336–14345.

Jamieson, Kevin G and Robert Nowak (2011). "Active ranking using pairwise comparisons". In: *Advances in Neural Information Processing Systems*, pp. 2240–2248.

Jang, Minje, Sunghyun Kim, Changho Suh, and Sewoong Oh (2016). "Top-K ranking from pairwise comparisons: When spectral ranking is optimal". In: *arXiv preprint arXiv:1603.04153*.

Jensen, Michael C (1978). "Some anomalous evidence regarding market efficiency." In: *Journal of Financial Economics,* 6.2/3, pp. 95–101.

Johansen, Søren (1988). "Statistical analysis of cointegration vectors." In: *Journal of Economic Dynamics and Control,* 12.2-3, pp. 231–254.

— (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Kelly, Bryan T, Asaf Manela, and Alan Moreira (2019). "Text selection." In:

Kirchgässner, Gebhard, Jürgen Wolters, and Uwe Hassler (2012). *Introduction to Modern Time Series Analysis.* Springer Science & Business Media.

Kleinberg, Jon M (1999). "Authoritative sources in a hyperlinked environment." In: *Journal of Association for Computing Machinery,* 46.5, pp. 604–632.

Kwiatkowski, Denis, Peter CB Phillips, Peter Schmidt, and Yongcheol Shin (1992). "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?." In: *Journal of Econometrics,* 54.1-3, pp. 159–178.

Lehmann, Bruce N (1990). "Fads, martingales, and market efficiency." In: *The Quarterly Journal of Economics,* 105.1, pp. 1–28.

Lo, Andrew W and A Craig MacKinlay (1988). "Stock market prices do not follow random walks: Evidence from a simple specification test." In: *The Review of Financial Studies,* 1.1, pp. 41–66.

Luce, R Duncan (1959). *Individual choice behavior*. John Wiley & Sons, New York.

Lütkepohl, Helmut (2005). *New Introduction to Multiple Time Series Analysis*. Springer Science & Business Media.

MacKinnon, James G (1994). "Approximate asymptotic distribution functions for unit-root and cointegration tests." In: *Journal of Business & Economic Statistics,* 12.2, pp. 167–176.

— (2010). *Critical Values for Cointegration Tests.* Working Paper. Queen's Economics Department.

Mallows, Colin L (1957). "Non-null ranking models. I". In: *Biometrika* 44.1/2, pp. 114–130.

Manela, Asaf and Alan Moreira (2017). "News implied volatility and disaster concerns," in: *Journal of Financial Economics* 123.1, pp. 137–162.

Martin, Ian (2017). "What is the Expected Return on the Market?." In: *The Quarterly Journal of Economics,* 132.1, pp. 367–433.

Mohajer, Soheil and Changho Suh (2016). "Active top-K ranking from noisy comparisons". In: *Communication, Control, and Computing (Allerton), 2016 54th Annual Allerton Conference on*. IEEE, pp. 875–882.

Mu, Enrique and Milagros Pereyra-Rojas (2016). *Practical Decision Making: An Introduction to the Analytic Hierarchy Process (AHP) Using Super Decisions V2.* Springer.

— (2017a). *Practical Decision Making Using Super Decisions V3: An Introduction to the Analytic Hierarchy Process.* Springer.

— (2017b). "Understanding the analytic hierarchy process." In: *Practical Decision Making*. Springer, pp. 7–22.

Nagel, Stefan (2012). "Evaporating liquidity." In: *The Review of Financial Studies,* 25.7, pp. 2005–2039.

— (2016). "The liquidity premium of near-money assets." In: *The Quarterly Journal of Economics,* 131.4, pp. 1927–1971.

Naseer, Mehwish and Bin D Tariq (2015). "The efficient market hypothesis: A critical review of the literature." In: *IUP Journal of Financial Risk Management,* 12.4, pp. 48–63.

Negahban, Sahand, Sewoong Oh, and Devavrat Shah (2017). "Rank centrality: Ranking from pairwise comparisons". In: *Operations Research* 65.1, pp. 266–287.

Newey, Whitney K and Kenneth D West (1994). "Automatic lag selection in covariance matrix estimation." In: *The Review of Economic Studies,* 61.4, pp. 631–653.

Phillips, Ryan L and Rita Ormsby (2016). "Industry classification schemes: An analysis and review." In: *Journal of Business & Finance Librarianship,* 21.1, pp. 1–25.

Pincus, Steve (2008). "Approximate entropy as an irregularity measure for financial data." In: *Econometric Reviews,* 27.4-6, pp. 329–362.

Pincus, Steven M (1991). "Approximate entropy as a measure of system complexity." In: *Proceedings of the National Academy of Sciences,* 88.6, pp. 2297–2301.

Pincus, Steven M and Ary L Goldberger (1994). "Physiological time-series analysis: what does regularity quantify?." In: *American Journal of Physiology-Heart and Circulatory Physiology,* 266.4, H1643–H1656.

Plackett, Robin L (1975). "The analysis of permutations". In: *Applied Statistics*, pp. 193–202.

Powell, John, Rubén Roa, Jing Shi, and Viliphonh Xayavong (2007). "A test for long-term cyclical clustering of stock market regimes." In: *Australian Journal of Management,* 32.2, pp. 205–221.

Raivola, Reijo (1985). "What is comparison? Methodological and philosophical considerations". In: *Comparative Education Review* 29.3, pp. 362–374.

Rey, Hélène (2015). "Dilemma not trilemma: the global financial cycle and monetary policy independence." In:

Saaty, Thomas L (1977). "A scaling method for priorities in hierarchical structures." In: *Journal of Mathematical Psychology,* 15.3, pp. 234–281.

— (1993). "What is relative measurement? The ratio scale phantom." In: *Mathematical and Computer Modelling,* 17.4-5, pp. 1–12.

— (2001a). "Fundamentals of the analytic hierarchy process." In: *The Analytic Hierarchy Process in Natural Resource and Environmental Decision Making*. Springer, pp. 15–35.

— (2001b). "The seven pillars of the analytic hierarchy process." In: *Multiple Criteria Decision Making in the New Millennium*. Springer, pp. 15–37.

Saaty, Thomas L and Luis G Vargas (2012). *Models, Methods, Concepts & Applications of the Analytic Hierarchy Process.* Springer Science & Business Media.

Said, Said E and David A Dickey (1984). “Testing for unit roots in autoregressive-moving average models of unknown order.” In: *Biometrika,* 71.3, pp. 599–607.

Schiatti, L, Giandomenico Nollo, G Rossato, and Luca Faes (2015). “Extended Granger causality: a new tool to identify the structure of physiological networks.” In: *Physiological Measurement,* 36.4, p. 827.

Schreiber, Thomas (2000). “Measuring information transfer.” In: *Physical Review Letters,* 85.2, p. 461.

Shah, Nihar B and Martin J Wainwright (2015). “Simple, robust and optimal ranking from pairwise comparisons”. In: *arXiv preprint arXiv:1512.08949*.

Shi, Shouyong and Alain Delacroix (2018). “Should buyers or sellers organize trade in a frictional market?.” In: *The Quarterly Journal of Economics,* 133.4, pp. 2171–2214.

Sidiropoulos, Nicholas D, Lieven De Lathauwer, Xiao Fu, Kejun Huang, Evangelos E Papalexakis, and Christos Faloutsos (2017). “Tensor decomposition for signal processing and machine learning”. In: *IEEE Transactions on Signal Processing* 65.13, pp. 3551–3582.

Smith, Russell A (1967). “The condition numbers of the matrix eigenvalue problem.” In: *Numerische Mathematik,* 10.3, pp. 232–240.

Sriperumbudur, Bharath K, Arthur Gretton, Kenji Fukumizu, Gert Lanckriet, and Bernhard Schölkopf (2008). “Injective Hilbert space embeddings of probability measures.” In: *21st Annual Conference on Learning Theory (COLT 2008)*. Omnipress, pp. 111–122.

Stein, Jeremy C (1989). “Efficient capital markets, inefficient firms: A model of myopic corporate behavior.” In: *The Quarterly Journal of Economics,* 104.4, pp. 655–669.

Suh, Changho, Vincent YF Tan, and Renbo Zhao (2017). “Adversarial Top-K Ranking”. In: *IEEE Transactions on Information Theory* 63.4, pp. 2201–2225.

Suppes, Patrick (1999). *Introduction to Logic*. Dover Publications.

Szörényi, Balázs, Róbert Busa-Fekete, Adil Paul, and Eyke Hüllermeier (2015). “Online rank elicitation for plackett-luce: A dueling bandits approach”. In: *Advances in Neural Information Processing Systems*, pp. 604–612.

Thurstone, Louis L (1927). “A law of comparative judgment”. In: *Psychological Review* 34.4, p. 273.

Truong, Charles, Laurent Oudre, and Nicolas Vayatis (2018). “A review of change point detection methods.” In: *ArXiv:1801.00718,* pp. 1–31.

Tsay, Ruey S (2005). *Analysis of Financial Time Series*. Second Edition. Wiley Series in Probability and Statistics. Wiley-Interscience.

Vervliet, Nico and Lieven De Lathauwer (2016). “A randomized block sampling approach to canonical polyadic decomposition of large-scale tensors”. In: *IEEE Journal of Selected Topics in Signal Processing* 10.2, pp. 284–295.

Vervliet, Nico, Otto Debals, and Lieven De Lathauwer (2016). “Tensorlab 3.0-Numerical optimization strategies for large-scale constrained and coupled matrix/tensor factorization”. In: *Signals, Systems and Computers, 2016 50th Asilomar Conference on*. IEEE, pp. 1733–1738.

Vervliet, NICO, OTTO Debals, and LIEVEN De Lathauwer (2017). “Canonical polyadic decomposition of incomplete tensors with linearly constrained factors”. In: *Technical Report 16–172, ESAT–STADIUS, KU Leuven, Belgium*.

Vervliet, Nico, Otto Debals, Laurent Sorber, and Lieven De Lathauwer (2014). “Breaking the curse of dimensionality using decompositions of incomplete tensors: Tensor-based scientific computing in big data analysis”. In: *IEEE Signal Processing Magazine* 31.5, pp. 71–79.

Vigna, Sebastiano (2016). “Spectral ranking”. In: *Network Science* 4.4, pp. 433–445.

Wauthier, Fabian, Michael Jordan, and Nebojsa Jojic (2013). “Efficient ranking from pairwise comparisons”. In: *International Conference on Machine Learning*, pp. 109–117.

Wilkinson, James H (1988). *The Algebraic Eigenvalue Problem*. USA: Oxford University Press, Inc.

Xiong, Wanting, Luca Faes, and Plamen C Ivanov (2017). “Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations.” In: *Physical Review E,* 95.6, p. 062114.

Zou, Xiaohua (2018). “VECM model analysis of carbon emissions, GDP, and international crude oil prices.” In: *Discrete Dynamics in Nature and Society,* 2018, pp. 1–11.