

# **Essays on Wavelet-Based Approaches for Analyzing Stock Price Dynamics**

DISSERTATION  
of the University of St.Gallen,  
School of Management,  
Economics, Law, Social Sciences  
and International Affairs  
to obtain the title of  
Doctor of Philosophy in Finance

submitted by

**Christian Vial**

from

Lucerne

Approved on the application of

**Prof. Dr. Karl Frauendorfer**

and

**Prof. Dr. Roland Füss**

Dissertation no. 5028

Difo-Druck GmbH, Untersiema 2020

The University of St.Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, May 25, 2020

The President:

Prof. Dr. Bernhard Ehrenzeller

# Acknowledgements

This dissertation was written as part of my doctoral studies at the University of St. Gallen. A dissertation is an endeavor that is not carried out by any single person, even if it ultimately only bears the author's name. Without the incredible people around me, I would not have managed to complete this dissertation. They have all contributed to it, and I wish to express my deepest thanks to those individuals for their immense and unconditional support.

My deepest gratitude goes to my supervisor Prof. Dr. Karl Frauendorfer for the opportunity to write this thesis under his supervision. He always gave me the freedom to develop my own ideas, but was there for me with his invaluable advice and deep experience whenever needed. I would also like to thank Prof. Dr. Roland Füss, who served as my co-supervisor and constantly supported me in all matters.

My special thanks go to Dr. Mark Kyburz, who edited my thesis so thoroughly and always showed incredible flexibility. Sometimes plans cannot be adhered to. And so it is even more invaluable when someone willingly offers help in such times.

Finally, I would like to thank my family for their unconditional love. They never stopped believing in me and supporting me. My parents, who have always been there for me and inspired me to strive for greater goals. They enabled me to pursue my doctorate. My brother Raphael, who always gave me hope and encouraged me to continue. He has sacrificed incredibly much of his own time for improving this thesis and has contributed to its quality with his invaluable advice. I could always count on him. I do not know how I can thank him enough.

Last but not least, my love Stefanie, who has always been there for me and has shown an incredibly great amount of understanding. She always brightened me up when I needed it and showed me new paths on which I could move forward. I am so indebted to her.

Without those people, I would not have completed this endeavor. I dedicate this thesis to them.

Küssnacht, October 2019

Christian Vial



# Contents

<b>List of Tables</b> . . . . .	<b>vi</b>
<b>List of Figures</b> . . . . .	<b>ix</b>
<b>Abstract</b> . . . . .	<b>xi</b>
<b>Zusammenfassung</b> . . . . .	<b>xiii</b>
<b>Synopsis</b> . . . . .	<b>1</b>
<b>I Analysis of Frequency Dynamics in US Stock Correlations</b> . . . . .	<b>7</b>
1 Introduction . . . . .	7
2 Wavelet Transformation . . . . .	11
2.1 Wavelet and Scaling Filter . . . . .	11
2.2 Maximal Overlap Discrete Wavelet Transformation . . . . .	14
2.3 Wavelet Correlation . . . . .	16
3 Data . . . . .	18
4 Empirical Results . . . . .	19
4.1 Statistics of Multiscale Correlation . . . . .	19
4.2 Multiscale Wavelet Correlation Pairs . . . . .	24
4.2.1 Subsample Analysis . . . . .	24
4.2.2 Full Sample Analysis . . . . .	30
4.2.3 Non-Crisis and Crisis Period Analysis . . . . .	33
4.3 Overall Multiscale Correlation . . . . .	40
4.4 Robustness . . . . .	45
5 Conclusion . . . . .	49
Appendix . . . . .	51

<b>II Multiscale Analysis of the Underlying Structures in US Stock</b>	
<b>Correlations: A Wavelet-Based Random Matrix Theory Approach . . .</b>	<b>55</b>
1 Introduction . . . . .	55
2 Literature Review . . . . .	59
3 Methodology . . . . .	62
3.1 Wavelet Theory . . . . .	63
3.2 Random Matrix Theory . . . . .	66
3.2.1 Eigendecomposition . . . . .	67
3.2.2 Eigenvalue Distribution of Random Matrices . . . . .	68
3.2.3 Distribution of the Bulk of the Eigenvalue Spectrum . . . . .	69
3.2.4 Eigenvector Distribution . . . . .	73
4 Data . . . . .	74
5 Empirical Results . . . . .	75
5.1 Cross-Correlation Statistics . . . . .	75
5.2 Theoretical Eigenvalue Distribution of Wavelet Correlation Matrix . . . . .	78
5.3 Analysis of the Empirical Eigenvalue and Eigenvector Distribution . . . . .	83
5.4 Consistency of the Bulk of the Eigenvalue Spectrum with RMT . . . . .	91
5.5 Interpretation of Correlation Statistics Across Timescales . . . . .	93
5.6 Timescale Characteristics of Correlation . . . . .	103
5.7 Robustness . . . . .	116
6 Conclusion . . . . .	118
<b>III Portfolio Optimization Under Heterogeneous Investment Horizons . . .</b>	<b>121</b>
1 Introduction . . . . .	121
2 Literature Review . . . . .	124
3 Wavelet Methodology . . . . .	125
3.1 Continuous Wavelet Transform . . . . .	126
3.2 Maximal Overlap Discrete Wavelet Transform . . . . .	129
3.3 Wavelet Variance and Covariance . . . . .	131

3.4	Standard Errors and Test Statistics of Wavelet Variance . . . . .	132
4	Wavelet Portfolio Strategy . . . . .	134
4.1	Optimization Method . . . . .	135
4.2	Weight Restrictions . . . . .	136
4.3	Portfolio Formation . . . . .	137
4.4	Performance & Risk Assessment . . . . .	138
5	Data & Data Analysis . . . . .	139
6	Empirical Analysis . . . . .	143
6.1	Coherence Analysis . . . . .	143
6.2	Analysis of Variance . . . . .	147
6.3	Wavelet-Based Portfolio Optimization . . . . .	151
6.4	Robustness . . . . .	171
7	Multiscale Portfolio Optimization . . . . .	178
8	Conclusion . . . . .	180
	Appendix A . . . . .	182
	Appendix B . . . . .	184

<b>Bibliography . . . . .</b>	<b>189</b>
-------------------------------	------------



# List of Tables

## I Analysis of Frequency Dynamics in US Stock Correlations

I.1	Descriptive statistics of the sample . . . . .	21
I.2	Correlation equality across timescales: Subsample analysis . . . . .	29
I.3	Correlation equality across timescales: Full sample analysis . . . . .	32
I.4	Analysis of correlation equality in crisis and non-crisis periods . . . . .	37
I.5	Correlation equality across timescales: Extended sample analysis . . . . .	47
I.6	Analysis of correlation equality using an LA(16) wavelet filter . . . . .	48
I.7	Correlation equality across timescales: Longer time period analysis . . . . .	51
I.8	Analysis of correlation equality using a DB(4) wavelet filter . . . . .	53

## II Multiscale Analysis of the Underlying Structures in US Stock Correlations: A Wavelet-Based Random Matrix Theory Approach

II.1	Descriptive statistics of timescale correlation . . . . .	77
II.2	Eigenvalues and eigenvectors disagreeing with RMT assumptions . . . . .	88
II.3	Variance contribution of eigenvalues at different timescales . . . . .	106
II.4	Analysis of the three largest eigenvalues across different timescales . . . . .	109

## III Portfolio Optimization Under Heterogeneous Investment Horizons

III.1	Descriptive statistics of daily log-returns . . . . .	140
III.2	Analysis of unconditional correlations . . . . .	142
III.3	MODWT variance decomposition of equally-weighted, daily portfolio returns . . . . .	148
III.4	MODWT-decomposed variance of minimum variance portfolios . . . . .	155
III.5	Descriptive statistics of time domain and scale-based minimum-variance portfolios . . . . .	162

III.6 Annualized volatility of time domain and scale-based minimum-variance portfolios . . . . .	168
III.7 Analysis of conventional time domain and scale-based minimum-variance portfolios using different weighting restrictions . . . . .	172
III.8 MODWT-decomposed variance of adaptive-estimation-window scale portfolio approach . . . . .	177
III.9 MODWT-decomposed variance of scale-based minimum variance portfolios: Crisis and non-crisis period analysis . . . . .	184
III.10 Annualized volatility of minimum variance portfolios: Different base date . . . . .	185
III.11 MODWT-decomposed variance of scale-based minimum variance portfolios: Different rebalancing intervals . . . . .	186

# List of Figures

## I Analysis of Frequency Dynamics in US Stock Correlations

I.1	Analysis of correlation probability density functions . . . . .	20
I.2	Wavelet correlation of stock subsample . . . . .	25
I.3	Wavelet correlation in non-crisis and crisis periods . . . . .	35
I.4	Wavelet multiple correlation . . . . .	41
I.5	Wavelet multiple correlation: Extended data sample . . . . .	46

## II Multiscale Analysis of the Underlying Structures in US Stock Correlations: A Wavelet-Based Random Matrix Theory Approach

II.1	Distribution of correlation coefficients . . . . .	76
II.2	Simulated, reshuffled, and theoretical eigenvalue distributions at different timescales . . . . .	82
II.3	Empirical, reshuffled, and theoretical eigenvalue distributions at different timescales . . . . .	84
II.4	Number of significant components (eigenvalues) as a function of the timescale . . . . .	89
II.6	Sector classification using absolute eigenvector components . . . . .	95
II.7	Contribution of the eigenvector to sectors over different timescales . . . . .	96
II.8	Regression of eigenvector portfolios on the market portfolio . . . . .	99
II.9	Log-log plot of inverse participation ratio (IPR) as a function of the eigenvalue . . . . .	101
II.10	Consistency of the eigenvector components across different timescales . . . . .	104
II.11	Evolution of eigenvalues across different timescales . . . . .	110
II.12	Distribution of filtered correlation matrix at different timescales . . . . .	114
II.13	Evolution of eigenvalues across different timescales . . . . .	117

### **III Portfolio Optimization Under Heterogeneous Investment Horizons**

III.1	Wavelet squared coherence between four selected stocks . . . . .	145
III.2	MODWT-decomposed variance contribution . . . . .	150
III.3	Portfolio weights of scale-based minimum variance portfolio strategies . . . . .	165

# Abstract

Financial markets — and stock price movements in particular — are often described using simplified assumptions. However, financial markets are *complex systems* involving various interacting components. Agents in these markets have heterogeneous traits and differ in many respects. Among other characteristics, they have unique preferences, interpret information differently, pursue disparate investment goals, and focus on different investment horizons. These heterogeneities impact agents' buying and selling decisions and ultimately stock prices. As a result, those heterogeneities directly influence interdependencies between stocks and their price dynamics. Existing methods have not been sufficiently able to capture and explain these *complexities*.

For this reason, the present thesis examines stock market mechanisms and interaction patterns using alternative mathematical filtration methods. The focus lies on investigating the price fluctuations of and the interdependencies between stocks across different timescales (time horizons). This analysis is directly linked to the assumption that market agents operate on different investment horizons.

Chapter 1 studies changes in US stock correlations for different time horizons using wavelet decomposition. Wavelet decomposition is a method that allows filtering the dynamics of a time series within certain frequency ranges (time horizons). The empirical observations in this study indicate that stock market correlations do not remain constant across different time horizons. A major deficiency of the analysis in Chapter 1 is the significant degree of randomness hidden in correlation matrices. Chapter 2 therefore examines correlation structures using random matrix theory (RMT). RMT analysis reveals that stock markets are governed by collective market behavior and sectoral factors across different timescales. Based on these insights, Chapter 3 studies portfolio strategies for minimizing risk at specific time horizons (scale-based portfolio strategies). The study demonstrates that (portfolio) variances can be minimized within a targeted frequency range using these scale-based portfolio strategies. Based on these findings, an optimization-method for simultaneous variance-minimization across different frequency bands is proposed.



# Zusammenfassung

Finanzmärkte – und Aktienpreisbewegungen im spezifischen – werden in der Ökonomie und im Finanzwesen oftmals auf Basis vereinfachter Annahmen beschrieben. Aktienmärkte sind jedoch *komplexe Systeme* mit einer Vielzahl interagierender Komponenten. Die unterschiedlichen Teilnehmer an diesen Märkten weisen dabei heterogene Eigenschaften auf und unterscheiden sich in zahlreichen Belangen. Unter anderem zeigen sie abweichende Präferenzen, interpretieren Information unterschiedlich, verfolgen verschiedene Investitionsziele und fokussieren sich auf unterschiedliche Investitionszeiträume. Diese Heterogenitäten wirken sich direkt auf das Kaufverhalten dieser Marktteilnehmer aus. Damit beeinflussen sie letztlich die Preisbewegungen von und die Interaktionsmuster zwischen verschiedenen Aktien im Markt. Bestehende Methoden können diese *komplexen* Zusammenhänge nicht immer erfassen.

Die vorliegende Arbeit untersucht deshalb die Mechanismen und Interaktionsmuster im Aktienmarkt mit Hilfe von alternativen mathematischen Filtrationsmethoden. Dabei steht insbesondere die Erforschung der Veränderung dieser Preisbewegungen und Interaktionsmuster für verschiedene Zeithorizonte im Vordergrund. Diese Analyse steht in direktem Zusammenhang zur Annahme, dass Marktteilnehmer auf verschiedenen Investitionshorizonten operieren.

Kapitel 1 analysiert die Korrelationsveränderungen des US Aktienmarktes für verschiedene Zeithorizonte (Timescales). Hierfür wird die Wavelet Dekompositionsmethode verwendet, welche es erlaubt, Dynamiken für verschiedene Frequenzbereiche (Zeithorizonte) aus der Zeitreihe zu filtrieren. Die Aktienmarktkorrelationen zeigen sich dabei als nicht konstant über die verschiedenen Zeithorizonte. Ein Problem der Analyse in Kapitel 1 besteht im erwiesenermassen grossen Anteil an Zufallskomponenten, welche in der Korrelationsmatrix vorliegen. Kapitel 2 untersucht deshalb die Korrelationsstrukturen anhand der Random Matrix Theory (RMT). RMT erlaubt es, innerhalb der Korrelationsmatrix zwischen informativen Bestandteilen und diesen Zufallskomponenten zu unterscheiden. Dabei zeigt sich, dass der Aktienmarkt durch gesamtmarktbezogene und sektorale Strukturen geprägt ist.

Kapitel 3 analysiert auf Basis dieser Erkenntnisse Portfoliostrategien, welche ein Portfolio über spezifische Zeithorizonte optimieren (Wavelet-basierte Optimierung). Dabei zeigt sich, dass die (Portfolio-)Varianzen in den avisierten Frequenzbereichen tatsächlich mit den Wavelet-basierten Portfoliostrategien minimiert werden können. Abschliessend wird eine Optimierungsmethode für die simultane Varianzminimierung über verschiedene Frequenzbereiche vorgestellt.

# Synopsis

Stock markets are *complex dynamic systems*. They are composed of a myriad of stocks that form *complex* networks with distinct interaction mechanisms and dependency structures. A vast number of agents participate in those markets and thus create system *complexity*. These agents are characterized by many heterogeneities: among others, they have various investment needs, they differ in their interpretation of information and perception of risk, and they vary in their investment perspectives. Interactions between these heterogeneous market agents are ultimately reflected in the *complex* evolution and interplay of stock prices.

Stock markets are further considered *dynamic* as they are subject to constant change. A stock market system continuously adjusts to different market states and is generally non-stationary. This change may result from intrinsic shifts in the behavior of market agents or from external influencing factors.

Economics and finance employ various simplifying assumptions to describe these *complex* and *dynamic* stock market systems. The most central assumption is perhaps the description of stock price fluctuations by means of a Brownian motion. This theory was first proposed by Bachelier (1900) in his groundbreaking and (at least at the time) underappreciated thesis on "Théorie de la Spéculation."<sup>1</sup> It laid the foundation for the Efficient Market Hypothesis (EMH) and now forms the bedrock of modern financial theory. However, other simplifying assumptions have also been used to describe stock prices and stock markets. For example, economists long relied on the assumption of completely rational investors and propagated concepts such as the assumption of representative agents. Heterogeneities had little or no place in these theories or were simply assumed to dissolve in the holistic system. Clearly, these ideas have not remained undisputed in economics since their formulation. As Buiter (2003) highlighted, James Tobin regarded the concept of the representative market agent as one of the "unfortunate theoretical developments in macroeconomic theory." Similarly, Mandelbrot (1963b, 1997) was an early critic of the assumption of

---

<sup>1</sup> Several decades later, the concept of Brownian motion was developed further by Osborne (1959) — an astrophysicist.

Gaussian-distributed price fluctuations and of the EMH. However, especially in the past two decades, researchers have progressively moved away from and questioned the irrefutability of these simplifying assumptions.<sup>2</sup> Not least, this metamorphosis was further accelerated by the global financial crisis of 2008 and by the failure of existing economic models to explain the frequency with which such collapses occur. Stock markets are increasingly regarded as what they are: *complex dynamic systems*. New theories that take into account these *complexities* have evolved over time: for example, adaptive expectations (Arthur, Holland, LeBaron, Palmer & Tayler, 1997; Routledge, 1999, 2001),<sup>3</sup> bounded rationality (Anufriev, Bottazzi & Pancotto, 2006; Brock & Hommes, 1998), Lévy motion (Mandelbrot, 1962, 1963b, 1967), fractional Brownian Motion (Mandelbrot, 1963a, 1965, 1997), adaptive market hypothesis (Lo, Andrew, 2005; Lo, 2004), fractal market theory (Peters, 1994, 1996; Weron & Weron, 2000), and heterogeneous market theories (Müller et al., 1993, 1997). Similarly, new methods for analyzing these systems have been introduced: for example, detrended fluctuation analysis, graph theory, multifractality analysis, random matrix theory, and wavelet theory.

This thesis contributes to the understanding of stock markets as *complex* and *dynamic systems* by examining three facets in more detail: time horizon effects, structural dependency, and changing market conditions. Its main focus lies on investigating how correlations in the stock market change under consideration of different time horizons. Is the stock market structure different in the short term compared to the long term? How do dependencies between stocks adjust as we consider monthly rather than daily time horizons? How do they behave in crisis periods? These and other questions are at the heart of this thesis.

To clarify these questions, I employ methods from *econophysics*<sup>4</sup> and specifically focus on filtration methods: like wavelet theory and random matrix theory. RMT respectively its derivatives have already been broadly applied in economics and finance

---

<sup>2</sup>There is ample empirical evidence that returns of financial assets are not Gaussian-distributed (Bouchaud & Potters, 2003; Dacorogna, Gençay, Müller, Olsen & Pictet, 2001; Lo & MacKinlay, 1999; Mantegna & Stanley, 2004).

<sup>3</sup>See LeBaron (2006) and Hommes (2006) for a detailed review of adaptive expectations models.

<sup>4</sup>I interpret the field of econophysics more broadly than other literature. I understand it not only as the application of mathematical methods from physics — but from many different disciplines — to problems of economics and finance.

while wavelet theory has only recently been introduced to the discipline. Below, I illustrate these methods in more detail and present the various studies comprising this thesis.

The first study uses wavelet analysis to investigate correlations and their changes over different time horizons (bi-daily up to yearly). Stock correlations are usually only considered for a specific time horizon (timescale), i.e., short- or long-term. Wavelet analysis is a method for examining the dynamics of a time series over several frequencies (timescales) and offers many advantages over other methods for timescale analysis.<sup>5</sup> These advantages include transformation without loss of information and simultaneous time and frequency resolution.

Using data of 268 US stocks for the period between June 30, 1980 to June 30, 2018, I find significant evidence that the relationship between stocks changes with time horizon (timescale). I demonstrate that correlations in non-crisis market periods rise for longer time horizons. In crisis periods, on the other hand, the level of correlation is shown to be high irrespective of the time horizon considered. As a result, only minor differences exist between short- and medium-term correlations. Only for longer time horizons is an increase in correlations once again detected. These findings have important implications for portfolio and risk management. They indicate that different investment decisions are necessary for investors with different investment horizons. A limitation of this study is the existence of comparatively high measurement inaccuracies in correlations of longer time horizons.

The second study introduces a more efficient way to examining correlation structures across different time horizons: random matrix theory (RMT).<sup>6</sup> This method is applied here in response to the results of the first study. RMT allows filtering informative and random components from a system based on eigendecomposition. In

---

<sup>5</sup>Alternative timescale methods, for example, include Fourier transformation, windowed Fourier transformation, or temporal aggregation. Temporal aggregation refers to adjusting the sampling rate in the time domain (calculating, e.g., daily, weekly, or monthly returns).

<sup>6</sup>The same sample of 268 US stocks and the same time period as in the first study are used. Thus, the second study effectively extends the investigation of wavelet correlations performed in the first study. However, it focuses on changes in the general correlation structures and substructures for different time horizons rather than on the overall correlation dynamics.

earlier research, RMT has already proven to be an effective instrument for investigating stock correlation structures. However, my study demonstrates the applicability of this method for analyzing wavelet correlations and thus for examining correlations across different time horizons.

Regardless of the time horizon, the largest eigenvalues are found to be closely linked to macroeconomic factors. The dominant (largest) eigenvalue reflects a collective effect of the whole market and thus is associated with a general market factor. In turn, the subdominant (next lower) eigenvalues relate to sectoral factors. While this interpretation remains consistent across different time horizons, the influence of these factors on the correlation structure changes. Almost all eigenvalues associated with sectoral factors increase with the time horizon. This increase is observable in non-crisis and crisis periods. This implies that sectoral structures in the correlation matrix become relatively more influential with increasing time horizon and irrespective of the market state.

For the largest eigenvalue — i.e., the eigenvalue associated with the market factor — more complex structures emerge. In non-crisis periods, the largest eigenvalue similarly increases with the time horizon. In times of crisis, the eigenvalue initially shows no difference at low timescales. Only at higher timescales, it also starts to rise. However, the eigenvalue is significantly higher in times of crisis than in times of non-crisis. This increase is more significant in comparison to the increase of the subdominant eigenvalues, so that — in crisis periods — the market factor represents the dominant component defining correlation at all time horizons.

This discovery of eigenvalue dynamics helps to explain why some capital market models provide meaningful results for monthly data while not exhibiting similar consistency for daily data. At the same time, the results provide important insights into the time-variant structures of stock markets in non-crisis and in crisis times. These observations offer crucial information for risk and portfolio management.

Based on the previous findings, the third study examines the application of portfolio strategies optimized for specific time horizons. This study again uses wavelet decomposition to derive stock covariances of different timescales. These wavelet-

based covariance matrices are then used to compute minimum variance portfolios optimized for particular timescales.

Portfolios are constructed for a sample of 23 US stocks for the period between March 29, 1969 and December 30, 2016. The results show that these timescale-based portfolio optimization strategies can indeed minimize variance at a specific time horizon. Consequently, investors should construct portfolios under consideration of their investment horizon and in accordance with the horizon over which they measure performance.

However, variance at low timescales (short time horizons) is found to contribute significantly more to total variance than variance at high timescales (long time horizons). It is unlikely that investors will fully ignore this high energy (variance) in stock fluctuations at low timescales. Hence, investors will focus on more than one investment horizon. Portfolio strategies that are optimized for one particular timescale might therefore not be pragmatic. In accordance with the multi-horizon preferences of investors, I propose a portfolio optimization method that allows for jointly minimizing variance over multiple time horizons.



# Chapter I

## Analysis of Frequency Dynamics in US Stock Correlations

*Christian Vial*

### 1 Introduction

Correlation measures the interplay of assets in the complex financial market system. Quantifying the degree of correlation is thereby crucial for gaining a general understanding of stock market interdependencies and the management of investment risk. A factor often ignored in this quantification of correlation is the measurement period and a possible variation of correlation over time. Many financial market models assume that correlations remain constant over different time horizons and rarely conceive interrelations between stocks as changing with sampling frequency (daily, weekly, monthly). This intuition is usually justified by the assumption of rational investors, who share an identical investment horizon.

The present study aims to demonstrate that correlations vary with the time horizon (timescale) and thus have a multi-horizon (multiscale) nature. This hypothesis follows from heterogeneous market theories and is in line with the understanding of financial markets as *complex systems*:<sup>1</sup> Heterogeneous market theories assume that stock markets consist of many heterogeneous agents with differing beliefs and investment needs. These heterogeneous agents operate at different investment horizons and process information by considering their respective investment timescale. The impact of new data on stock prices thereby depends on how agents perceive and interpret this information.<sup>2</sup> While certain information is relevant for short-term

---

<sup>1</sup>Prominent representatives of heterogeneous market theories include the fractal market hypothesis (Peters, 1994), the heterogeneous market hypothesis (Dacorogna, Müller, Pictet & Olsen, 1998; Müller et al., 1993), and the adaptive market hypothesis (Lo, 2004).

<sup>2</sup>The concept of heterogeneous investment horizons is closely related to the concept of heterogeneous expectations, where the former is likely a central component of the latter (Chen & Li, 2016; Kirman,

agents, the same information might not equally affect the views and decision-making of long-term agents. In particular, agents with a short-term investment horizon are more interested in the relationship between stocks at higher frequencies (short-term fluctuations), whereas long-term agents focus on comovements of stock returns at lower frequencies (long-term fluctuations) (Candelon, Piplack & Straetmans, 2008; Jammazi & Aloui, 2012; Madaleno & Pinho, 2012, 2014).

Even though agents show perceptual differences and dissimilar investment horizons, this does not imply that they are necessarily driven by different factors. Short- and long-term agents may react to the same market news and economic announcements, yet may weigh and process this information differently. How they incorporate information into their investment decision ultimately defines stock prices and the correlations between stocks: Given the multitude of heterogeneous agents with dissimilar investment horizons and disparate interpretation of information, stock correlations are likely to exhibit time-horizon-variant characteristics.<sup>3</sup> Therefore, heterogeneous market theories equally imply multiscale dependency structures.

This study intends to explore the possible multi-horizon nature of the relationship between stocks. A lack of analytical tools has prevented previous research from studying the properties of correlations for diverse time horizons: In general, conventional time domain methods use different sampling intervals (employing, among others, daily, weekly, or monthly data) to deduce the short- and long-term characteristics of time series. However, an increase in the sampling interval and the associated removal of data points are always accompanied by information loss. So far, correlation analysis with conventional time domain methods has thus mostly been limited to short- and long-term time perspectives (In & Kim, 2006; Masih, Alzahrani & Al-Titi, 2010; Shah, Deo & King, 2016).

This study employs wavelet transformation — a non-parametric filtering technique — to overcome these temporal limitations. In contrast to conventional short- and long-term time domain analysis, wavelet transformation allows for a more granu-

---

2006). Similarly, behavioral finance theories and market anomalies such as bounded rationality, herding, and momentum trading can be integrated into heterogeneous market models and often appear as leitmotifs for the rationale of non-homogeneous market agents.

<sup>3</sup>Information can affect both short- and long-term stock price movements, although to different degrees.

lar filtration of the frequency spectrum while preserving full time series information (Shah et al., 2016). The transformation decomposes stock returns into different timescale (time horizon) components. Thereby, each timescale represents a certain frequency band in the spectrum of possible time horizons. These timescale components (wavelet coefficients) enable deriving multiscale correlations.

Wavelet functions are localized in time and in timescale. Compared to other frequency filtering methods such as the Fourier transformation, wavelet transformation thus allows simultaneously representing correlation in both the time and the timescale (frequency) domain. Additionally, the underlying time series need to be neither periodic nor stationary. Consequently, wavelet transformation is ideally suited to investigating the timescale properties of stock correlations (Gençay, Selçuk & Whitcher, 2003).

This study contributes to existing research by analyzing the timescale cross-correlation structure of a large universe of US stocks using wavelet transformation. Wavelet transformation has previously been used to study timescale cross-correlations between assets (see, e.g., Gençay, Selçuk & Whitcher, 2001a, 2005; In & Kim, 2006; Xu & Gençay, 2003). However, the investigation of timescale correlations in stock markets has mostly been limited to a few stocks. Considering correlation structures for a larger sample provides new insights into the multiscale relationship between stocks and the general functioning of the overall stock market.

Further, this study contributes to existing research by analyzing the timescale properties of correlations during different market states. Longin and Solnik (2001), Ang and Chen (2002), Fenn et al. (2011), and others have shown that correlations vary considerably between normal and distressed market periods. However, these studies are mostly limited to a specific time horizon. In turn, studies that consider multiscale correlation structures usually do not investigate different market states. The present study provides new insights into the structural timescale relationships in stock markets by jointly examining changes in the correlation structure across different market phases and time horizons.

Finally, this study contributes to existing research by examining both individual wavelet correlations and the multivariate wavelet correlation structure. Most pre-

vious studies have been limited to only one of these analyses. However, to gain a deeper understanding of the timescale-varying dependency structures in stock markets, it is crucial to jointly examine individual correlations and the overall correlation structure.

The results for the full observation period show that stock correlations are indeed time-horizon-inconsistent and that the dependency structure between stocks varies with the time horizon. The analysis of different market phases reveals even more complex correlation structures. Distinctive temporal mechanisms define the interaction between stocks in different market states:

In *non-crisis periods*, correlations are generally lower at short-term horizons and increase with the time horizon. A non-negligible number of stocks demonstrate these timescale-variant correlation patterns. The heterogeneous market hypothesis provides a possible explanation for this trend. Lower correlations at short-term horizons might result from a greater influence of firm-specific (idiosyncratic) factors on stock price movements. These firm-specific factors induce lower correlations between stocks. On the other hand, increasing comovement between stocks at long-term horizons might result from a growing influence of long-term common factors. Macroeconomic trends, which affect all stocks, might gain in influence with increasing time horizons.

In contrast, in *crisis periods*, the correlations across different timescales vary only slightly. This homogeneity of correlations in times of crisis might be explained by the fact that stock prices are determined by the same cross-market news during distressed market periods. Information might affect short-term and long-term investors equally during such market states. As a result, asset correlations are high at all timescales.

The remainder of this study is organized as follows. Section 2 presents wavelet transformation to derive stock correlations for different timescales. Section 3 describes the dataset. Section 4 presents the empirical analysis of the stock correlations for different timescales. First, a reduced sample of representative stocks is used to gain an understanding of the mechanisms of timescale correlations. The analysis is then extended to the complete sample of 268 US stocks. In a next step, the timescale sensitivity of correlations is studied with regard to various market states. Then, the over-

all correlation structure is investigated using multivariate analysis based on wavelet multiple correlation. Finally, I test the empirical results for robustness. Section 5 discusses the limitations of this study and suggests avenues for future research.

## 2 Wavelet Transformation

Wavelet transformation decomposes a time series into sets of coefficients that are associated with distinct time horizons (timescales). The elements of each set relate to a particular time location (Conlon, Cotter & Gençay, 2018). Hence, this method partitions the original time series into multiresolution components and represents the signal in the time-timescale-domain (time-frequency-domain).

Wavelet transformation is achieved by filtering the time series with a collection of high-pass and low-pass filters. These high- and low-pass filters are called wavelet and scaling filters, respectively. A high-pass filter attenuates low- while preserving high-frequency characteristics of a time series. Contrastingly, a low-pass filter discards high- and preserves low-frequency features of a signal.

Cascading these filters allows constructing a succession of frequency intervals. The bandwidths of these frequency intervals are halved with every application of the filters, in the descent from high to low frequencies. Thus, the filtering procedure enables decomposing a time series into certain frequency bands (timescales).<sup>4</sup> Correlation matrices for every timescale can then be derived from the collection of these transformed time series.

### 2.1 Wavelet and Scaling Filter

Let  $\{\tilde{h}_l; l = 0, \dots, L - 1\}$  in  $\mathbb{R}^L$  be the wavelet filter, where  $L$  is the width of the filter and is required to be an even number. For the wavelet filter to have width  $L$ , it

---

<sup>4</sup>A band-pass filter (i.e., a filter that passes frequencies within a certain range and attenuates frequencies outside that range) can be constructed by recursively applying a combination of low-pass and high-pass filters.

must hold that  $\tilde{h}_0 \neq 0$  and  $\tilde{h}_{L-1} \neq 0$ . However,  $\tilde{h}_l = 0$  applies for  $l < 0$  and  $l \geq L$  such that  $\{\tilde{h}_l\}$  is an infinite sequence. Further, a wavelet filter must satisfy the basic properties

$$\sum_{l=0}^{L-1} \tilde{h}_l = 0, \quad \sum_{l=0}^{L-1} \tilde{h}_l^2 = \frac{1}{2}, \quad \text{and} \quad \sum_{l=-\infty}^{\infty} \tilde{h}_l \tilde{h}_{l+2n} = 0, \quad (1)$$

for all non-zero integers  $n$ . These properties ensure that the wavelet filter i) sums to zero and thus identifies changes in the data, ii) has half-unit energy, which guarantees variance preservation, and iii) is orthogonal to its even shifts, which facilitates multiresolution analysis. While the second property requires some of the coefficients to deviate from zero, the first condition guarantees that these deviations cancel each other out in sum.

The scaling filter  $\{\tilde{g}_l; l = 0, \dots, L-1\}$  in  $\mathbb{R}^L$  complements the wavelet filter  $\tilde{h}_l$ . It is a quadrature mirror filter. Thus, the frequency response of  $\tilde{g}_l$  (Fourier transform of  $\tilde{g}_l$ ) is the mirror image of the frequency response of  $\tilde{h}_l$  around  $\pi/2$  and the following relation must hold:

$$\tilde{g}_l = (-1)^{l+1} \tilde{h}_{L-1-l}. \quad (2)$$

The basic properties of the scaling filter are thus given by

$$\sum_{l=0}^{L-1} \tilde{g}_l = 1, \quad \sum_{l=0}^{L-1} \tilde{g}_l^2 = \frac{1}{2}, \quad \text{and} \quad \sum_{l=-\infty}^{\infty} \tilde{g}_l \tilde{g}_{l+2n} = 0, \quad (3)$$

for all non-zero integers  $n$ . In contrast to the wavelet filter, the coefficients of the scaling filter sum to 1. The remaining properties of the scaling filter accord with the wavelet filter (half-unit energy and orthogonal to even shifts).

Several different filters fulfill the conditions in formulas 1 and 3. Each of these filters can be suitable for a specific analysis and may best match the underlying dataset. The

choice of a wavelet filter and its respective length is a non-trivial task. A wavelet filter should match the characteristic features of the underlying data series.<sup>5</sup>

The choice of filter length is also characterized by a trade-off between frequency and time localization. A short wavelet filter (time domain) may not capture the complexities in the spectral density of certain time series. It may also introduce undesirable artifacts into the wavelet coefficients. Contrastingly, a filter that is too long decreases the degree of time localization of the decomposed series. Additionally, it renders the influence of boundary conditions more severe (see the following section 2.2) (Masset, 2008; Percival & Walden, 2000).

Comparing different wavelet filters has shown that the Daubechies Least-Asymmetric wavelet filter (symmlet) of length eight LA(8) is an appropriate choice for analyzing the sample in this study.<sup>6</sup> It provides a good balance between time and frequency localization. This conclusion is consistent with the observations of Percival and Walden (2000). Furthermore, the LA(8) wavelet filter is widely used in financial research. According to Gençay, Gradojevic, Selçuk and Whitcher (2010), Gençay et al. (2002), and Gençay et al. (2005), it is well-suited for analyzing financial time series.

---

<sup>5</sup>For example, a smooth wavelet filter may be chosen if the original data series is smooth. Similarly, the filter length should be chosen so that it accurately reflects the frequency information of a time series.

<sup>6</sup>The Daubechies class of filters are best defined in terms of the squared gain function for the Daubechies scaling filter:

$$\mathcal{G}(f) = 2\cos^L(\pi f) \sum_{l=0}^{\frac{L}{2}-1} \binom{\frac{L}{2}-1+l}{l} \sin^{2l}(\pi f),$$

where the length  $L$  is a positive even integer and  $f$  is the frequency. Using the relation between the squared gain function of the scaling function and the squared gain function of the wavelet function  $\mathcal{H}(f) = \mathcal{G}\left(f + \frac{1}{2}\right)$ , the corresponding squared gain function of the wavelet filter is obtained by

$$\mathcal{H}(f) = 2\sin^L(\pi f) \sum_{l=0}^{\frac{L}{2}-1} \binom{\frac{L}{2}-1+l}{l} \cos^{2l}(\pi f).$$

However, these gain functions do not define unique sequences of Daubechies wavelet filters. Procedures known as factorization can be used to obtain specific filters (for further details, see Oppenheim and Schaffer (2009); Bruce and Gao (1996); Härdle, Kerkyacharian, Picard and Tsybakov (1998); Daubechies (1992)). One of these factorizations produces the least asymmetric (LA) class of wavelet filters used in this study Gençay, Selçuk and Whitcher (2002).

## 2.2 Maximal Overlap Discrete Wavelet Transformation

The Maximal Overlap Discrete Wavelet Transform (MODWT) is a special form of wavelet transformation.<sup>7</sup> It is usually implemented on the basis of Mallat's (1989) pyramid algorithm (Gençay et al., 2002). With this algorithm, wavelet transforms are obtained by successively filtering the vector of observations with the wavelet and the scaling filter.

Let  $r$  be a time series of returns  $\{r_t; t = 0, \dots, N - 1\}$  in  $\mathbb{R}^N$  of length  $N$ . Then, the wavelet coefficients  $\tilde{W}_{j,t}$  and scaling coefficients  $\tilde{V}_{j,t}$  for iteration  $j > 0$  and time  $t$  are obtained by:

$$\tilde{W}_{j,t} = \sum_{l=0}^{L-1} \tilde{h}_l \tilde{V}_{j-1,t-2^j-1l \bmod N}, \quad \tilde{V}_{j,t} = \sum_{l=0}^{L-1} \tilde{g}_l \tilde{V}_{j-1,t-2^j-1l \bmod N}, \quad (4)$$

for  $t = 0, 1, \dots, N - 1$  and  $\tilde{V}_{0,t} \equiv r_t$ . Hence, the wavelet and scaling coefficients in the first step of the pyramid algorithm are derived by convolution of the wavelet and scaling filter with the original time series. For every next step, the filtering operation is repeated with the output of the prior scaling coefficients.

The wavelet coefficients  $\tilde{W}_{j,t}$  cover detailed fluctuations of the original time series. They are related to changes for timescale (time horizon)  $\lambda_j$  of length  $2^j - 1$  at decomposition (scale) level  $j$  and time location  $t$ . The scaling coefficients  $\tilde{V}_{j,t}$  represent the overall trend in the time series for scale  $\lambda_j$  and time location  $t$ . Therefore, the wavelet transforms (i.e., the collection of wavelet coefficients of a certain timescale) are filtered time series that are associated with distinct time horizons. For a data series with daily sampling rate, the first scale level coefficients capture fluctuations at frequencies of 2–4 days. The next higher scale level covers frequencies of 4–8 days. This decomposition can be continued up to a maximum level of decomposition, which is given by  $J = \lfloor \log_2(N) \rfloor$ .

This study uses seven levels of decomposition  $J = 7$ . The coefficients for these

---

<sup>7</sup>Several names have been used in the literature to refer to the MODWT: the undecimated DWT, non-decimated DWT, translation-invariant DWT, shift-invariant DWT, time-invariant DWT, stationary DWT, or wavelet frames.

seven levels of decomposition are:  $\tilde{W}_{1,t}$  (2–4 days),  $\tilde{W}_{2,t}$  (4–8 days),  $\tilde{W}_{3,t}$  (8–16 days),  $\tilde{W}_{4,t}$  (16–32 days),  $\tilde{W}_{5,t}$  (32–64 days),  $\tilde{W}_{6,t}$  (64–128 days),  $\tilde{W}_{7,t}$  (128–256 days), and  $\tilde{V}_{7,t}$  (256– days). The collection of wavelet transforms  $\{\tilde{W}_{1,t}, \tilde{W}_{2,t}, \dots, \tilde{W}_{7,t}, \tilde{V}_{7,t}\}$  fully reflects the information of the original time series and provides a seamless time-timescale representation.

Note that the modulo operator in formula 4 results in a circular filtering over the finite time series. This operation is necessary to derive the wavelet coefficients both at the beginning and at the end of the time series. Evidently, this modification leads to a bias in these boundary coefficients. To derive an unbiased statistical measure (e.g., the unbiased wavelet correlation), those wavelet coefficients affected by the boundary condition need to be accounted for (see section 2.3).

The circularity operation also makes an implicit assumption about the continuation of the time series for values outside the observable spectrum. Directly applying this circular filtering operation to the original time series implies the signal to be periodic. This is a rather unrealistic and inadequate assumption, unless strong seasonality effects exist in the time series. An alternative and effective technique borrowed from Fourier analysis is to mirror the time series at its last observation. This mirroring produces a new time series of length  $2N$ . The reflection of the time series mitigates the assumption of periodicity. At the same time, the procedure prevents sudden shifts (discontinuities) at the boundaries of the return series.<sup>8</sup> Therefore, I use this reflection of the time series to derive the wavelet transforms in this study.

The MODWT is related to another wavelet transformation: the Discrete Wavelet Transform (DWT). The MODWT is an undecimated wavelet transformation whereas the DWT is not. Hence, the MODWT does not remove redundant wavelet coefficients through decimation<sup>9</sup>. Due to this redundancy, the MODWT filters in formulas

<sup>8</sup>This reflection has no influence on the sample mean and variance. Both remain identical to those of the original time series.

<sup>9</sup>Decimation is derived from the Latin term "removal of the tenth." Originally, this referred to a practice of military discipline in the Roman army to punish units found guilty of capital offenses. A cohort of renegade soldiers would be divided into groups of ten men. One soldier of each group would be chosen by lots and executed by his comrades. It was a pragmatic approach to punishing a large group of offenders. While bearing a sinister connotation due to its original context, the term decimation has a less violent meaning in signal processing. It refers to the procedure of reducing the sampling rate of a signal. It is also known as downsampling or subsampling.

1 and 3 are rescaled versions of the filters typically used in the DWT. This rescaling is necessary to ensure energy preservation (aggregate variance). It guarantees that the MODWT accurately reflects the variance and covariance contribution at particular timescales.

The MODWT offers some distinct advantages over the DWT (among others): shift-invariance, a non-dyadic time series length requirement, and an asymptotically more efficient (co-)variance estimator. In turn, the DWT also shows some useful properties, like DWT transformed time series being typically decorrelated (in contrast to MODWT).<sup>10</sup> While this study uses the MODWT transformation, the inherent relationship between the two transformations enables benefiting from some of the advantageous properties of the DWT. For example, the decorrelating properties of the DWT are used to derive the confidence intervals of the MODWT wavelet correlations (see next section).

## 2.3 Wavelet Correlation

The MODWT wavelet coefficients represent the changes in the time series at a particular timescale. Percival and Mofjeld (1997) showed that the MODWT is energy-preserving. Hence, the variance of the wavelet coefficients also reflects the energy of the original time series at a certain timescale. The transformation thus allows decomposing the variance of the time series  $r$  on a scale-by-scale basis:

$$\begin{aligned} \|r\|^2 &= \sum_{t=0}^{N-1} r_t^2 = \sum_{j=1}^J \sum_{t=0}^{N-1} \tilde{W}_{j,t}^2 + \sum_{t=0}^{N-1} \tilde{V}_{J,t}^2 \\ &= \sum_{j=1}^J \|\tilde{\mathbf{W}}_j\|^2 + \|\tilde{\mathbf{V}}_J\|^2, \end{aligned} \tag{5}$$

---

<sup>10</sup>For a detailed coverage of the DWT, see Percival and Walden (2000).

where  $\tilde{\mathbf{W}}_j = \{\tilde{W}_{j,t}; t = 0, \dots, N-1\}$ ,  $\tilde{\mathbf{V}}_j = \{\tilde{V}_{j,t}; t = 0, \dots, N-1\}$ , and  $\|\dots\|^2$  is the squared Euclidean norm specifying the energy of the time series. These results can be transferred analogously to covariance decomposition (Whitcher, 1998). Given two return series  $\{r_{p,t}\}$  and  $\{r_{q,t}\}$  of stocks  $p$  and  $q$ , the unbiased wavelet cross-covariance derived from MODWT wavelet coefficients for scale  $\lambda_j$  is defined as

$$\tilde{v}_{p,q}(\lambda_j) = \frac{1}{\tilde{N}_j} \sum_{t=L_j-1}^{N-1} \tilde{W}_{p,j,t} \tilde{W}_{q,j,t}, \quad (6)$$

where  $\tilde{N}_j \equiv N - L_j + 1$  specifies the coefficients unaffected by the boundary conditions,  $L_j \equiv (2^j - 1)(L - 1) + 1$  refers to the length of a filter at scale level  $j$ , and  $\tilde{W}_{p,j,t}$  and  $\tilde{W}_{q,j,t}$  describe the wavelet coefficients of the time series  $\{r_{p,t}\}$  and  $\{r_{q,t}\}$ , respectively.

The wavelet cross-correlation  $\tilde{\rho}_{p,q}(\lambda_j)$  is now specified as

$$\tilde{\rho}_{p,q}(\lambda_j) = \frac{\tilde{v}_{p,q}(\lambda_j)}{\tilde{v}_{p,p}^2(\lambda_j) \tilde{v}_{q,q}^2(\lambda_j)}, \quad (7)$$

where  $\tilde{v}_{p,p}^2(\lambda_j)$  and  $\tilde{v}_{q,q}^2(\lambda_j)$  are equivalent to the variances of the two processes  $p$  and  $q$  at timescale  $\lambda_j$ , respectively. This wavelet cross-correlation measures the relationship between two processes on a scale-by-scale basis.

Confidence intervals for wavelet correlation coefficients are obtained based on large sample properties and by applying the Fisher  $z$ -transformation,  $F(\cdot) \equiv \tanh^{-1}(\cdot)$ . For the correlation estimate  $\hat{p}$  based on  $N$  independent samples,  $\sqrt{N-3}[F(\hat{p}) - F(p)]$  is approximately normally distributed, where  $p$  is the population correlation. Therefore, the approximate  $1 - \alpha$  confidence interval for a wavelet correlation coefficient is given by

$$\left[ \tanh \left\{ F [\tilde{p}_{p,q}(\lambda_j)] - \frac{\Phi^{-1} \left( 1 - \frac{\alpha}{2} \right)}{\sqrt{\hat{N}_j - 3}} \right\}, \right. \\ \left. \tanh \left\{ F [\tilde{p}_{p,q}(\lambda_j)] + \frac{\Phi^{-1} \left( 1 - \frac{\alpha}{2} \right)}{\sqrt{\hat{N}_j - 3}} \right\} \right], \quad (8)$$

where  $\hat{N}_j = 2^{-j}N - \lceil (L-2)(1-2^{-j}) \rceil$  and  $\Phi^{-1}(\cdot)$  specifies the probit function. The operation  $\tanh(\cdot)$  transforms the confidence intervals back to the interval  $[-1, 1]$ . The quantity  $\hat{N}_j$  corresponds to the number of coefficients used in a (decimated) DWT at scale level  $j$ . The distributional assumption of the Fisher  $z$ -transformation is only valid if the observations of the wavelet processes  $\tilde{W}_{p,j,t}$  and  $\tilde{W}_{q,j,t}$  are uncorrelated (Gençay et al., 2002). As mentioned, the DWT is an approximately decorrelating transformation. Consequently, the quantity  $\hat{N}_j$  is a reasonable measure for the sample size in formula 8.

### 3 Data

The dataset for the empirical analysis consists of daily closing prices for all stocks of the S&P 500 index between June 30, 1980 and June 30, 2018 with a full price history (survival period). This results in a sample of 268 stocks and a time series length of 9,935 data points (roughly 38 years). Price history data were obtained from the Center of Research in Security Prices (CRSP). Data for index affiliation were gathered from Compustat.

The restriction to index components with full history is imperative for analyzing long-term dynamics in time series. It ensures that sufficiently long observation periods are available to adequately describe long horizon comovements between stocks. Yet, the restriction might be detrimental with regard to the representative portrayal of overall market behavior. However, a regression of the market capitalization-weighted portfolio of the sample on the return series of the S&P 500 index showed high coher-

ence between the two time series (not shown). The underlying stocks can therefore be regarded as representative of the general market.

The requirement for the full time series history might introduce an imminent survivorship bias to the analysis. Therefore, section 4.4 also tests the results using a larger sample of stocks over a shorter observation period. However, these robustness tests are limited due to the increasing measurement errors at longer time horizons. Hence, the presence of an imminent survivorship bias cannot be rejected completely.

## 4 Empirical Results

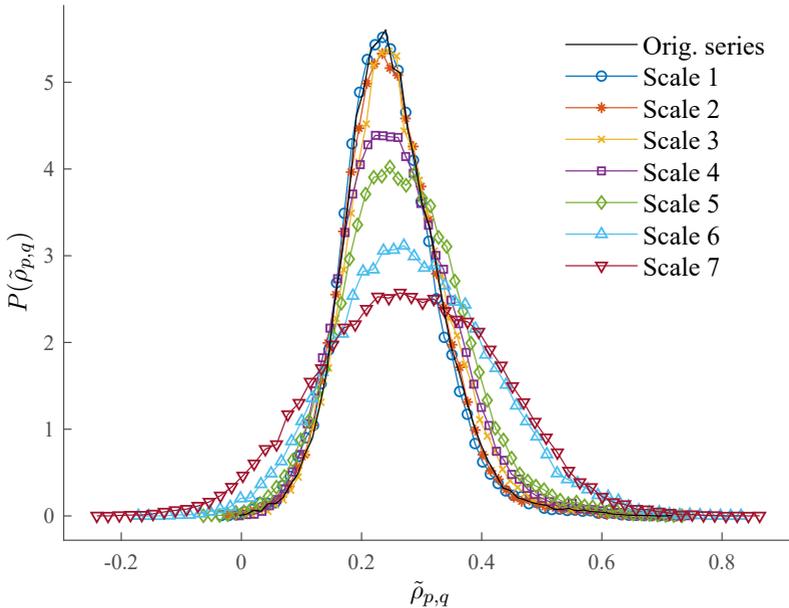
This section applies wavelet decomposition to determine whether correlation structures in US stocks vary over different time horizons. I start describing the general distribution of (wavelet) correlations for different timescales. Next, I study interactions between stocks on individual timescales in more detail: In a first step, I use a representative subsample of the stocks to gain a general understanding of the timescale-dependency of correlations (term-structure of correlations). In a second step, the analysis of the term-structure of correlations is extended to the full data sample.

Aste, Shaw and Di Matteo (2010) showed that correlations of untransformed times series vary between crisis and non-crisis periods. If correlations are sensitive to the prevailing market condition, it is likely that the timescale structure of correlations also differs across market states. Therefore, I also divide the observation period into crisis and non-crisis periods to study changes in the term-structure of correlations during different market states. Finally, I examine the overall structure of the wavelet correlation matrix using multivariate analysis.

### 4.1 Statistics of Multiscale Correlation

Studying the distribution of (wavelet) correlations  $\tilde{\rho}_{p,q}(\lambda_j)$  allows identifying possible changes in the relation between stock correlations across time horizons. The

(wavelet) correlation matrix for each timescale  $\lambda_j$  consists of  $M(M-1)/2$  unique correlation coefficients  $\{\tilde{\rho}_{p,q}(\lambda_j); p \neq q\}$  where  $M$  specifies the number of assets. For the given sample of 268 stocks, this amounts to 35,778 coefficients per timescale. Figure I.1 presents the distribution  $P(\tilde{\rho}_{p,q}(\lambda_j))$  of these elements for each timescale  $\{\lambda_j; j = 0, \dots, 7\}$ . The distributions of the correlations of the original, untransformed time series  $P(\tilde{\rho}_{p,q}(\lambda_0))$  and those of wavelet-decomposed series of the lowest timescale ( $P(\tilde{\rho}_{p,q}(\lambda_1))$ ; 2–4 days) exhibit similar shapes. Both distributions are centered around a positive correlation value, positively skewed, and relatively peaked. However, with increasing timescale, the distribution of the correlation coefficients flattens, showing higher variance and a lower level of kurtosis.



**Figure I.1:** Distribution of correlation coefficients of the original (untransformed) return series  $\{\tilde{\rho}_{p,q}(\lambda_0); p \neq q\}$  and of wavelet transformed return series  $\{\tilde{\rho}_{p,q}(\lambda_j); p \neq q; j = 1, \dots, 7\}$  (scale levels 1 to 7). Notes: A Daubechies least asymmetric MODWT filter of length 8 was used to decompose returns; correlations were calculated between 268 stocks, covering the period June 30, 1980 to June 30, 2018.

Interestingly, mainly the right tail of the correlation distribution extends with higher timescales: The probability of higher correlations grows with the time horizon. This observation is consistent with the general belief that stock prices are increasingly driven by common systematic factors at longer time horizons and that macroeconomic news exert greater influence at these timescales. The greater impact of those common systematic factors induces average correlations to increase. This intuition is also supported by previous research. For example, Handa, Kothari and Wasley (1989) found that the capital asset pricing model (CAPM) gains in explanatory power with increasing return interval.

**Table I.1:** Descriptive statistics of correlation coefficients of the original (untransformed) return series  $\{\tilde{\rho}_{p,q}(\lambda_0); p \neq q\}$  and of wavelet transformed return series  $\{\tilde{\rho}_{p,q}(\lambda_j); p \neq q; j = 1, \dots, 7\}$  (scale levels 1 to 7).

	Mean	Std. dev.	Min.	Max.	Skew.	Kurt.	JB-stat.
Orig. series	0.2526	0.0803	-0.0089	0.7670	0.6314	4.4343	5,525***
Scale 1	0.2467	0.0808	-0.0346	0.7226	0.5753	4.5284	5,538***
Scale 2	0.2512	0.0814	-0.0274	0.7779	0.5422	4.2952	4,318***
Scale 3	0.2617	0.0846	0.0026	0.8624	0.6680	4.2971	5,246***
Scale 4	0.2619	0.0921	-0.0019	0.9113	0.5884	3.8284	3,134***
Scale 5	0.2707	0.0998	-0.0682	0.9471	0.3798	3.5579	1,344***
Scale 6	0.2897	0.1276	-0.1754	0.9610	0.2339	3.0011	331***
Scale 7	0.2819	0.1453	-0.2417	0.9717	0.0239	2.7698	84***

Notes: JB-stat. shows the Jarque-Bera test statistics for the null hypothesis of normality in the correlation coefficient distribution; a Daubechies least asymmetric MODWT filter of length 8 was used to decompose returns; correlations were calculated between 268 stocks, covering the period June 30, 1980 to June 30, 2018. \*, \*\*, and \*\*\* indicate rejection of the null hypothesis of the test statistics at the 1%, 5%, and 10% level of significance, respectively.

Table I.1 shows the descriptive statistics corresponding to the correlation distributions presented in Figure I.1. Along with Figure I.1, the table exemplifies four characteristics of correlations at different timescales in more detail:

First, for the period under study, US stocks generally exhibit positive comovements in their price dynamics at all timescales. While negative correlations exist, they are relatively scarce.

Second, the average correlation increases with increasing timescale. The mean rises from 0.2467 at scale level 1 to 0.2897 at scale level 6 with a slight reduction to 0.2819

at scale level 7. While more extreme correlations are reported for larger timescales, the increase in probability of positive correlations outweighs the increase in negative correlations.

Third, skewness and kurtosis decrease with increasing time horizon. While the distributions feature fat right-tails over shorter horizons, they become more normally distributed with increasing timescale. Nevertheless, they remain non-normal even for higher time horizons. The Jarque-Bera statistic rejects the null hypothesis of a Gaussian distribution at all timescales (including the untransformed data) at a 1% level of significance.

Fourth, the standard deviations of the correlation distributions increase and the range of possible correlations expands with the timescale. Correlations of the US stock market become more diverse with longer time horizon. Visual inspection of the distributions in Figure I.1 reveals that this increase in standard deviations — in combination with higher average correlations — results in higher probabilities of more positive correlations. While more positive correlations are coherent with the presence of systematic factors, the increase in width of the distribution is less consistent with this general expectation, and results in a dichotomous interpretation. One would anticipate correlations to consolidate due to a reduction of idiosyncratic risk (randomness) relative to the risk incurred by systematic factors.

A possible explanation of this conundrum may be the generally high level of randomness present in the correlation structure (in combination with estimation errors). Even when the contribution of the systematic (informative) component increases, random components still define a large part of the correlation (see Laloux, Cizeau, Bouchaud & Potters, 1999; Plerou et al., 2002). These random components lead to assigning a large probability mass to correlation coefficients in the region around zero (random correlation matrix). On the other hand, the systematic components may result in a larger probability mass being assigned to high, positive correlations. The combination of these random and systematic components in the overall correlation structure may therefore lead to comparatively wide distributions. Due to the growing contribution of the systematic component to the general correlation structure (and the resulting rise in the level of correlation), this effect is expected to intensify for increasing timescales.

In fact, Figure I.1 indicates a possible bimodality in the correlation distribution for timescales  $\{\lambda_j; j = 6, 7\}$ . This observation supports the assumption regarding the timescale-dependent composition of the correlation structure based on random and systematic components.

Another reason for the widening of the distribution might be the non-stationarity of correlations. The correlation estimates used to derive the distribution in Figure I.1 and the statistics in Table I.1 are smoothed over the entire observation period. However, correlations may change both with different market states and with temporal evolution (see, e.g., Fenn et al., 2011; Sandoval & Franca, 2012). Similarly, it is likely that the way in which correlations change with respect to market states varies across timescales. For example, macroeconomic information is likely to differently impact short- and long-term correlations in calm and distressed market periods. The combination of the correlation distributions of the different market phases may cause an expansion of the overall distribution. However, following general intuition, one would expect the distribution of short-term rather than of long-term correlations to expand.

Section 4.3 examines the effects of different market conditions on the term-structure of correlations in more detail. Analysis reveals different timescale-dependency structures in non-crisis and crisis periods. Consistent with our expectation, correlations exhibit larger differences at lower timescales. Consequently, the non-stationarity of correlations seems less effective in explaining the widening of the distribution at higher timescales (as seen in Figure I.1).

To summarize, correlation measures exhibit varying statistical properties across different timescales. Wavelet correlation coefficients at the lowest timescales feature similar distributional shapes as the correlations of the original, untransformed time series. In contrast, correlation distributions at longer time horizons display largely different distributional characteristics. Both the mean correlation and the width of the distribution increase with timescale. These results indicate that stock returns show varying degrees of interaction over different time horizons and justify additional statistical tests.

## 4.2 Multiscale Wavelet Correlation Pairs

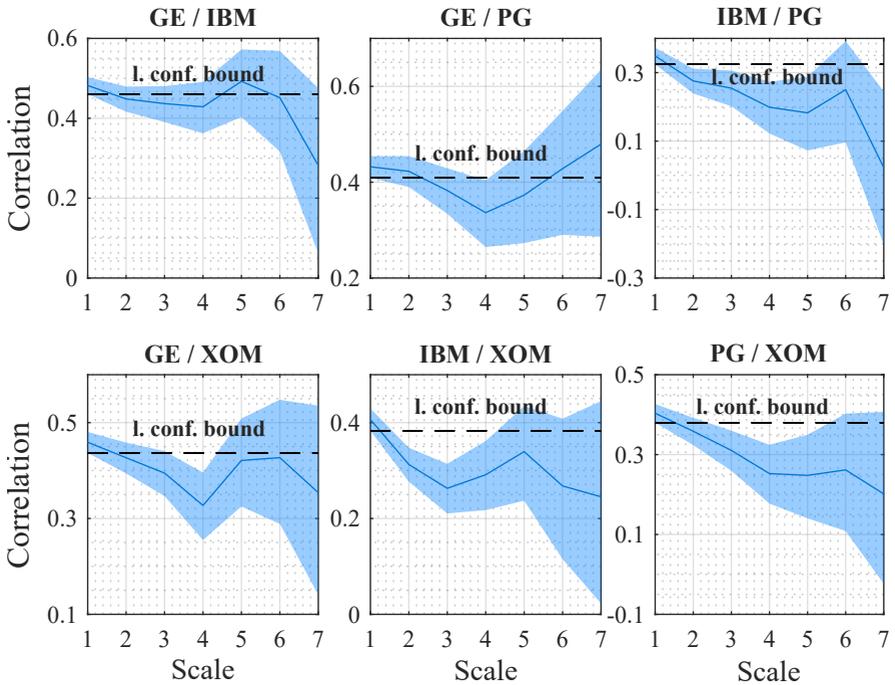
This section examines individual wavelet correlations to identify the underlying characteristics of correlations and their properties over different timescales in more detail. In contrast to the previous analysis, individual features of correlations might be preserved and not obscured by the conflation of distributions and by pooling observations.

Section 4.2.1 first only considers a subset of four stocks out of the full sample of 268 stocks. The focus on this smaller sample helps to reduce analytical complexity. Section 4.2.2 then extends the analysis to the entire data sample. Finally, section 4.2.3 divides the sample into non-crisis and crisis periods and analyzes timescale correlations for both phases. This progression enables understanding correlation properties at different timescales and deciphering potential enigmatic features in correlation structure.

### 4.2.1 Subsample Analysis

The subset studied in this section consists of four stocks (resulting in a total of six mutual interactions), which serve as a proxy for stock interrelation over different timescales. It comprises General Electric (GE), International Business Machines (IBM), Exxon Mobil (XOM), and Procter & Gamble (PG). These stocks are part of the S&P 500 and are also constituents of the Dow Jones Industrial Average (DJIA). At the same time, they belong to different sectors: Industrials, Information Technology, Energy, and Consumer Staples. The subset can be understood as an indicative and diverse sample of the US stock market.

Figure I.2 displays wavelet correlations for the subset of stocks over different timescales. The timescale spectrum ranges from scale level 1, which encompasses fluctuations of 2–4 days, up to scale level 7, which spans oscillations within 128–256 days. The blue shaded region in each graph represents the 95% confidence intervals of the correlations (as defined in formula 8). The dashed horizontal line in each subfigure represents the lower confidence bound of the wavelet correlations of scale level 1, extrapolated across all timescales. This reference line enables assessing the deviations of higher timescale correlations from those correlations of the first timescale.



**Figure I.2:** Wavelet correlations between four selected stocks: Exxon Mobil (XOM), General Electric (GE), International Business Machines (IBM), and Procter & Gamble (PG), covering the full period June 30, 1980 to June 30, 2018. Notes: The blue shaded region in each graph represents the 95% confidence intervals; the horizontal dashed line extrapolates the lower confidence bound of wavelet correlation at scale level 1 to all scale levels; a Daubechies least asymmetric MODWT filter of length 8 was used to decompose returns.

Figure I.2 reveals significant deviations in correlations between different timescales. For each pair of stocks, the reference line characterizing the lower confidence bound of the scale level 1 correlation is undercut by at least one observation at longer time horizons. For example, the correlation between the stock pair IBM/XOM at scale level 1 (2–4 days) is 0.41, with a lower 95% confidence bound of 0.38 (reference line). The correlation at scale level 2 (4–8 days) already decreases to a value of 0.31 with an upper 95% confidence bound of 0.35, i.e., well below the 0.38 lower confidence bound of the first scale level. This decrease in correlation continues up to

scale level 3, for which the measure further reduces to 0.26 (upper confidence bound of 0.31). Thus, the correlation between IBM and XOM declines by almost 35% from a bi-daily (scale level 1) to a weekly/bi-monthly time horizon (scale level 3). Figure I.2 also shows similar observations for the remaining stock pairs. Note that the deviations in correlation are more pronounced, the greater the spacing between timescales. This is in line with the general intuition that daily correlations more strongly resemble weekly rather than monthly correlations.

Overall, the correlation varies with the time horizon over which it is considered. This divergence presents a first indication of scale-dependent relations between stock price processes — at least within this very limited sample.

These basic results of heterogeneity in correlation suggest that there are different degrees of diversification across short-, medium-, and long-term investment horizons. As initially mentioned, possible explanations of this disparity may be grounded in heterogeneous agent and market theories. These theories imply the presence of non-homogeneous market agents with different life-cycle preferences and heterogeneous interpretations of information. Perceptual differences make short-, medium-, and long-term investors react differently to the advent of new information and these differences thus result in diverse investment decisions. Eventually, this heterogeneous interpretation of information affects the price dynamics of financial assets and their mutual interaction.

Another explanation of scale-dependent correlations may be the presence of market frictions (e.g., transaction costs, illiquidity, or non-synchronous trading). These frictions may expose stocks to lead-lag-effects and serial correlation. Both effects could impart cross-correlations between assets across different timescales. However, given that this study uses daily closing prices of large US stocks, at least the exemplified market frictions seem less explanatory for timescale-varying cross-correlations. First, all stocks are highly liquid with low transaction costs. Second, the closing prices of US stocks are largely synchronized at daily sampling frequencies. In addition, all stocks are traded on the NYSE or the NASDAQ stock exchange, both of which have the same closing and opening hours (full stock sample).

Nevertheless, numerous other market frictions significantly impact the correlation

dynamics between stocks at different timescales (e.g., insider trading or non-heterogeneous dissemination of information). Consequently, market frictions cannot be completely discarded as a cause of heterogeneous timescale correlations.

Independent of the two above hypotheses, the core finding of Figure I.2 remains the same: There is a reasonable degree of certainty that correlations vary over different timescales for this subsample of stocks.

To investigate the significance of the above results, I introduce a statistical test for the equality of scale correlations. This enables quantitatively assessing differences in stock correlations between different timescales. To the best of my knowledge, this is the first study to apply this form of statistical comparison to wavelet correlations. The null hypothesis of the test statistic states that correlations measured for two dissimilar timescales are identical (i.e.  $\tilde{\rho}_{p,q}(\lambda_j) = \tilde{\rho}_{p,q}(\lambda_k)$ , where  $j \neq k$ ). The alternative hypothesis states that correlations are different (i.e.  $\tilde{\rho}_{p,q}(\lambda_j) \neq \tilde{\rho}_{p,q}(\lambda_k)$ , where  $j \neq k$ ). The corresponding test statistic is given by  $(z_j - z_k) / \sqrt{z_{se}}$ , where  $z_{se} = \frac{1}{\hat{N}_j - 3} + \frac{1}{\hat{N}_k - 3}$ ,  $z_j$  and  $z_k$  are the Fisher  $z$ -transformed scale correlation coefficients  $\tilde{\rho}_{p,q}(\lambda_j)$ , and  $\tilde{\rho}_{p,q}(\lambda_k)$  for scale level  $j$  and  $k$ , respectively. Further,  $\hat{N}_j$  and  $\hat{N}_k$  refer to the number of wavelet coefficients dictated by the decorrelating DWT. This test is applied to every combination of timescales and to all pairwise stock interactions in the specified subsample.<sup>11</sup>

Table I.2 presents the results for the statistical test of equality between correlations of different timescales. Unlike the graphical analysis in Figure I.2, these results provide a quantitative measure with exact  $p$ -values for testing the hypotheses. Note that  $p$ -values are derived under the same distributional assumptions as adopted in formula 8. The first column in Table I.2 specifies the two timescales for which correlations are compared. Given that the original time series is decomposed into 7 timescales, a total of 21 unique tests for the equality of timescale correlation results for each asset pair. The  $z$ -values reported in Table I.2 relate to the difference between two timescale correlation coefficients and correspond to the nominator of the test statistic. Hence, the sign of the  $z$ -value indicates the direction of deviations between

---

<sup>11</sup>The same test will be used in the next section to analyze the full sample.

correlation measures. If the  $z$ -value is positive (negative), the first scale correlation of the comparison is greater (smaller) than the second scale correlation. For example, the positive  $z$ -value of 2.639 for GE/PG between the correlations at scale level 1 and scale level 4 signifies reduced correlation from the former to the latter. Columns two and three in Table I.2 further list the relevant DWT coefficients used to derive the test statistic (denominator).<sup>12</sup>

In accordance with Figure I.2, Table I.2 shows that significant deviations in scale correlations mainly appear between lower and medium timescales. Deviations between bi-daily (scale 1) and monthly (scale 4) correlations of the investigated stock pairs — GE/PG, IBM/PG, GE/XOM, IBM/XOM, and PG/XOM — are all highly significant. With the exception of GE/PG, all of these stocks also demonstrate significant deviations between daily (scale 1), weekly (scale 2), and bi-weekly (scale 3) correlations. Given that the correlation measures are based on the same underlying stocks, i.e., the same price co-dynamics, these significances are unexpectedly high. The results strongly indicate that correlations vary over different time horizons. Moreover, all corresponding  $z$ -values of the significant deviations are positive, indicating decreasing correlations from bi-daily to monthly periods. Interestingly, this implies that those deviating stocks provide greater diversification benefits at weekly, bi-weekly, and monthly rather than at bi-daily time horizons.

While Figure I.2 and Table I.2 exemplify that correlations decrease from short- to medium-term horizons, there is no clear trend in their behavior from short- to long-term horizons. Five stock pairs show reduced correlation from scale level 1 to scale level 7, while one stock pair exhibits increasing correlation. However, only deviations of GE/IBM, IBM/PG, and PG/XOM with positive  $z$ -values are significant, moreover only at a relatively modest level of significance of 6%.

---

<sup>12</sup>Note that the quantity  $\hat{N}_j$  significantly decreases for higher scale levels. This reduction in coefficients is responsible for the expansion of the confidence region in Figure I.2 with increasing scale levels. As a result, it becomes more difficult to draw statistical inferences about the relations of correlations at higher timescales. Unfortunately, this problem cannot be resolved without introducing other limitations. For example, extending the observation period may reduce the confidence intervals. However, this would simultaneously further smooth non-stationarities in the time series and reduce the size of the stock sample (full time series requirement).

**Table I.2:** Test of equality between correlations of different timescales, covering correlations between four stocks: Exxon Mobil (XOM), General Electric (GE), International Business Machines (IBM), and Procter & Gamble (PG).

	$\hat{N}_j$	$\hat{N}_k$	GE/IBM		GE/PG		IBM/PG		GE/XOM		IBM/XOM		PG/XOM	
			$z$	$p$	$z$	$p$	$z$	$p$	$z$	$p$	$z$	$p$	$z$	$p$
Scale 1 / Scale 2	4,967	2,483	1.749	<b>0.080</b>	0.473	0.636	3.293	<b>0.001</b>	1.635	0.102	4.393	<b>0.000</b>	2.146	<b>0.032</b>
Scale 1 / Scale 3	4,967	1,241	1.812	<b>0.070</b>	1.873	<b>0.061</b>	3.261	<b>0.001</b>	2.478	<b>0.013</b>	5.106	<b>0.000</b>	3.362	<b>0.001</b>
Scale 1 / Scale 4	4,967	620	1.579	0.114	2.639	<b>0.008</b>	3.806	<b>0.000</b>	3.664	<b>0.000</b>	3.081	<b>0.002</b>	3.979	<b>0.000</b>
Scale 1 / Scale 5	4,967	310	-0.230	0.818	1.202	0.230	3.051	<b>0.002</b>	0.801	0.423	1.325	0.185	2.970	<b>0.003</b>
Scale 1 / Scale 6	4,967	155	0.473	0.636	0.059	0.953	1.317	0.188	0.485	0.628	1.905	<b>0.057</b>	1.942	<b>0.052</b>
Scale 1 / Scale 7	4,967	77	2.018	<b>0.044</b>	-0.508	0.611	2.916	<b>0.004</b>	1.079	0.281	1.546	0.122	1.909	<b>0.056</b>
Scale 2 / Scale 3	2,483	1,241	0.418	0.676	1.376	0.169	0.650	0.516	1.107	0.268	1.557	0.119	1.553	0.120
Scale 2 / Scale 4	2,483	620	0.542	0.588	2.246	<b>0.025</b>	1.811	<b>0.070</b>	2.583	<b>0.010</b>	0.522	0.602	2.603	<b>0.009</b>
Scale 2 / Scale 5	2,483	310	-0.934	0.350	0.976	0.329	1.627	0.104	0.114	0.909	-0.497	0.619	2.015	<b>0.044</b>
Scale 2 / Scale 6	2,483	155	-0.048	0.961	-0.081	0.935	0.329	0.742	-0.003	0.997	0.585	0.559	1.283	0.200
Scale 2 / Scale 7	2,483	77	1.639	0.101	-0.603	0.546	2.208	<b>0.027</b>	0.730	0.465	0.619	0.536	1.448	0.148
Scale 3 / Scale 4	1,241	620	0.200	0.842	1.079	0.281	1.194	0.232	1.576	0.115	-0.623	0.533	1.279	0.201
Scale 3 / Scale 5	1,241	310	-1.115	0.265	0.175	0.861	1.189	0.234	-0.496	0.620	-1.322	0.186	1.065	0.287
Scale 3 / Scale 6	1,241	155	-0.216	0.829	-0.636	0.525	0.056	0.955	-0.451	0.652	-0.062	0.951	0.618	0.536
Scale 3 / Scale 7	1,241	77	1.494	0.135	-0.995	0.320	1.988	<b>0.047</b>	0.398	0.691	0.158	0.875	0.975	0.329
Scale 4 / Scale 5	620	310	-1.158	0.247	-0.601	0.548	0.243	0.808	-1.565	0.118	-0.767	0.443	0.069	0.945
Scale 4 / Scale 6	620	155	-0.314	0.754	-1.191	0.234	-0.596	0.551	-1.286	0.198	0.280	0.779	-0.109	0.913
Scale 4 / Scale 7	620	77	1.373	0.170	-1.400	0.162	1.455	0.146	-0.244	0.807	0.403	0.687	0.436	0.663
Scale 5 / Scale 6	310	155	0.529	0.597	-0.664	0.507	-0.715	0.474	-0.072	0.942	0.796	0.426	-0.149	0.882
Scale 5 / Scale 7	310	77	1.929	<b>0.054</b>	-1.005	0.315	1.251	0.211	0.612	0.541	0.797	0.426	0.377	0.706
Scale 6 / Scale 7	155	77	1.392	0.164	-0.454	0.650	1.644	0.100	0.610	0.542	0.171	0.864	0.449	0.654

*Notes:* The null hypothesis of the test states that for stocks  $p$  and  $q$ , timescale-correlations are equal  $\tilde{\rho}_{p,q}(\lambda_j) = \tilde{\rho}_{p,q}(\lambda_k)$  for scale level  $j \neq k$ ; tests are conducted for all possible combinations of scale levels 1–7; for a given stock pair, the test statistics is defined by  $z/\sqrt{(\hat{N}_j - 3)^{-1} + (\hat{N}_k - 3)^{-1}}$ . The variables  $\hat{N}_j$  and  $\hat{N}_k$  refer to the number of DWT coefficients for scale (levels)  $j$  and  $k$ . The variable  $z = z_j - z_k$  denotes the difference between the Fisher  $z$ -transformed correlation coefficients  $z_j$  and  $z_k$  for different scale (levels) but the same stock pair;  $p$  denotes the  $p$ -value of the test; a Daubechies least asymmetric MODWT filter of length 8 was used to decompose returns; correlations are calculated for the period June 30, 1980 to June 30, 2018. Bold  $p$ -values indicate rejection of the null hypothesis at least at the 10% level of significance.

Note that differences between correlations are predominantly traceable at lower timescales due to narrow confidence bands. The probability of a type 1 or type 2 error at these levels of significance are relatively low. Understandably, fewer observations exist for longer cycles. This implies that the precision of estimates decreases, and that confidence intervals widen with increasing timescales. Consequently, the possibility of statistical inference generally deteriorates for tests that rely on measures at higher timescales. Nevertheless, deviations between correlations at bi-daily

(scale level 1) and half-yearly/yearly time horizons (scale level 7) for the stock pairs GE/IBM, IBM/PG, and PG/XOM are significant.

The declining correlations with increasing timescales contrast with previous observations. For the full sample presented in Table I.1, an increasing trend in average correlation was observed from scale levels 1 to 6. Previous studies also reported increasing correlations with the timescale in stock markets (Fernández-Macho, 2012; Gallegati, 2005; Conlon et al., 2018). Given the small sample size of only four assets with six mutual interactions (out of 35'778 possible interactions for the full sample of 268 stocks), the declining trend might be a distinctive peculiarity observed for these few stocks. Thus, extending the analysis to all 268 stocks is needed to give further insights into the overall behavior of the stock market across different timescales.

#### 4.2.2 Full Sample Analysis

Table I.3 presents the aggregated results of testing equality between timescale correlations for the full data sample of 268 stocks. For each stock pair, I derive the test statistics for the equality of scale correlations and retrieve the corresponding  $p$ -value. Table I.3 shows three different levels of significance (10%, 5%, and 1%). If the  $p$ -value retrieved from the test is below the specified level of significance, the observation is considered a significant deviation. The percentage of significant deviations for the full stock sample is given by the sum of the significant deviations over the total number of analyzed correlations.

Table I.3 further subdivides the significant observations into positive and negative deviations. This separation helps to understand whether correlations increase or decrease with timescale. For example, the subdivision reveals that 22.32% of the significant deviations (at 5%) are due to correlations decreasing from bi-daily (scale level 1) to bi-weekly timescales (scale level 3). In contrast, 77.68% of the significant deviations result from correlations increasing from bi-daily to bi-weekly timescales. The bottom row in Table I.3 summarizes the stock pairs, which exhibit at least one significant deviation between their timescale correlations. If none of the deviations are significant, the stock pair is considered timescale-independent. If at least one scale correlation deviates from the group, the stock pair correlation is classified as

timescale-dependent.<sup>13</sup> In a nutshell, the last row in Table I.3 summarizes those stocks with a non-flat correlation curve.

Table I.3 reveals a considerable number of significant deviations in correlations for different timescales. Multiple stock pairs show at least one timescale correlation that deviates from the general correlation level (67.47% at 10% level of significance). Even at a more stringent level of significance of 1%, approximately 23.00% of stock pairs demonstrate timescale-varying correlations. This amount is non-negligible. Yet, these results need to be considered with caution. Combining the individual tests inflates the number of scale-dependent stock pairs (experimentwise error rate). Each individual test has an error margin of 1%. Therefore, the probability of finding significant deviations between stock pairs increases considerably.<sup>14</sup> Moreover, individual tests are not strictly independent. Therefore, the tests for correlations between individual timescales are considered next. In addition, section 4.3 introduces the a method (wavelet multiple correlation), to investigate the overall structure of the correlation matrix, which is not subject to these limitations.

Analyzing the individual relations between scale correlations reveals that significant deviations are mainly observed among low timescale correlations. This finding is consistent with the results documented for the subsample in section 4.2.1. Unfortunately, the test design does not enable concluding whether this occurs due to the equivalence of longer-term correlations or due to the limited number of observations at higher timescales.

Between scale levels 1 and 6, the share of significant negative deviations is 87.30% (at the 5% level of significance). This implies that correlations for these stocks increase from bi-daily to quarterly/half-yearly time horizons. In other words, stocks exhibit higher interdependencies with increasing timescale. This finding further supports the initial assumption that correlation structure is increasingly determined by a systematic component at higher timescales (see section 4.1).

<sup>13</sup>For example, while the correlations between GE/PG at scale levels 1 and 2 show no significant deviations, correlations at scale levels 1 and 3 are significantly different. Therefore, the relationship between the stock pair is considered timescale-dependent.

<sup>14</sup>Pairwise comparison of scale correlations across all timescales  $J = 7$  increases the probability of finding a significant deviation between stock pairs to  $1 - (1 - \alpha)^{21} = 0.1903$  for a level of significance of  $\alpha = 1\%$ .

**Table I.3:** Test of equality between wavelet correlations of different timescales, covering the complete data sample of 268 stocks.

	$p < 0.10$				$p < 0.05$				$p < 0.01$			
	Num. of sig. dev.	Perc. of sig. dev.	Perc. of pos. sig. dev.	Perc. of neg. sig. dev.	Num. of sig. dev.	Perc. of sig. dev.	Perc. of pos. sig. dev.	Perc. of neg. sig. dev.	Num. of sig. dev.	Perc. of sig. dev.	Perc. of pos. sig. dev.	Perc. of neg. sig. dev.
Scale 1 / Scale 2	6,167	17.24	38.17	61.83	3,684	10.30	39.14	60.86	1,210	3.38	43.31	56.69
Scale 1 / Scale 3	9,330	26.08	23.44	76.56	6,482	18.12	22.32	77.68	2,859	7.99	22.04	77.96
Scale 1 / Scale 4	9,070	25.35	28.07	71.93	6,280	17.55	27.56	72.44	2,741	7.66	26.89	73.11
Scale 1 / Scale 5	7,331	20.49	22.34	77.66	4,622	12.92	20.58	79.42	1,648	4.61	19.42	80.58
Scale 1 / Scale 6	8,769	24.51	15.49	84.51	5,938	16.60	12.70	87.30	2,447	6.84	7.60	92.40
Scale 1 / Scale 7	6,647	18.58	23.30	76.70	4,044	11.30	20.05	79.95	1,190	3.33	15.13	84.87
Scale 2 / Scale 3	3,499	9.78	17.38	82.62	1,853	5.18	14.30	85.70	491	1.37	11.61	88.39
Scale 2 / Scale 4	5,489	15.34	28.38	71.62	3,245	9.07	25.45	74.55	1,049	2.93	22.50	77.50
Scale 2 / Scale 5	4,892	13.67	21.67	78.33	2,745	7.67	18.98	81.02	768	2.15	13.28	86.72
Scale 2 / Scale 6	7,053	19.71	14.08	85.92	4,498	12.57	10.87	89.13	1,701	4.75	4.59	95.41
Scale 2 / Scale 7	5,373	15.02	23.39	76.61	3,095	8.65	20.26	79.74	799	2.23	16.27	83.73
Scale 3 / Scale 4	1,707	4.77	48.80	51.20	684	1.91	49.27	50.73	115	0.32	53.91	46.09
Scale 3 / Scale 5	3,266	9.13	33.34	66.66	1,599	4.47	31.64	68.36	315	0.88	32.38	67.62
Scale 3 / Scale 6	5,410	15.12	17.21	82.79	3,226	9.02	13.30	86.70	991	2.77	6.26	93.74
Scale 3 / Scale 7	4,449	12.44	29.09	70.91	2,392	6.69	27.22	72.78	540	1.51	27.04	72.96
Scale 4 / Scale 5	1,239	3.46	33.17	66.83	458	1.28	32.75	67.25	53	0.15	20.75	79.25
Scale 4 / Scale 6	4,150	11.60	15.04	84.96	2,256	6.31	10.99	89.01	623	1.74	6.10	93.90
Scale 4 / Scale 7	3,545	9.91	27.67	72.33	1,768	4.94	26.13	73.87	360	1.01	22.50	77.50
Scale 5 / Scale 6	2,384	6.66	14.72	85.28	1,163	3.25	9.20	90.80	305	0.85	1.64	98.36
Scale 5 / Scale 7	2,756	7.70	36.57	63.43	1,245	3.48	34.62	65.38	222	0.62	30.18	69.82
Scale 6 / Scale 7	1,310	3.66	60.99	39.01	505	1.41	68.51	31.49	85	0.24	72.94	27.06
Any sig. dev. for given asset pairs	24,140	67.47			17,916	50.08			8,230	23.00		

*Notes:* The null hypothesis of the test states that for stocks  $p$  and  $q$ , timescale-correlations are equal  $\bar{\rho}_{p,q}(\lambda_j) = \bar{\rho}_{p,q}(\lambda_k)$  for scale level  $j \neq k$ ; for a given stock pair, the test statistics is defined by  $(z_j - z_k) / \sqrt{(\hat{N}_j - 3)^{-1} + (\hat{N}_k - 3)^{-1}}$ , where the variables  $\hat{N}_j$  and  $\hat{N}_k$  refer to the number of DWT coefficients and the coefficients  $z_j$  and  $z_k$  are the Fisher  $z$ -transformed correlations at different scale levels ( $j \neq k$ ); tests are conducted for all possible combinations of scale levels 1–7 and all possible combinations of stock pairs. For a given scale level combination  $j$  and  $k$ , "Number of sig. dev." summarizes the number of tests between all possible combinations of stock pairs that reject the null hypothesis  $\mathcal{R}(\lambda_j, \lambda_k) \equiv \sum_{(p < q)} \mathbb{I}[\{\bar{\rho}_{p,q}(\lambda_j) \neq \bar{\rho}_{p,q}(\lambda_k)\}_H]$ . "Perc. of sig. dev." determines the proportion of these significant deviations to total number of correlation coefficients  $\frac{M(M-1)}{2} = 35,778$ . "Perc. of pos. sig. dev." and "Perc. of neg. sig. dev." designate the proportion of positive  $\left(\sum_{(p < q)} \mathbb{I}[\{\bar{\rho}_{p,q}(\lambda_j) > \bar{\rho}_{p,q}(\lambda_k)\}_H]\right)$  and negative deviations  $\left(\sum_{(p < q)} \mathbb{I}[\{\bar{\rho}_{p,q}(\lambda_j) < \bar{\rho}_{p,q}(\lambda_k)\}_H]\right)$  of all significant deviations  $\mathcal{R}(\lambda_j, \lambda_k)$ ; the bottom row summarizes the number of stock pairs with at least one significant deviation over all scale level combinations  $\left(\sum_{(p < q)} \mathbb{I}\left[\sum_{(j < k)} \mathbb{I}[\{\bar{\rho}_{p,q}(\lambda_j) \neq \bar{\rho}_{p,q}(\lambda_k)\}_H] > 0\right]\right)$  and its proportion to total number of correlation coefficients; tests are conducted for level of significance  $p$  of 1%, 5%, and 10%; a Daubechies least asymmetric MODWT filter of length 8 was used to decompose returns; correlations are calculated for the period June 30, 1980 to June 30, 2018.

Unlike the observations for the subsample (Table I.2), most significant deviations between correlations at lower timescales are negative (i.e., indicative of increasing correlations). For example, weekly correlations (scale level 2) are higher than bi-daily correlations (scale level 1) in 60.86% of the cases for which significant deviations are detected (at the 5% level). In contrast, only 39.14% of the significant deviations can be attributed to higher correlations at the first two scale levels. This contradicts the exemplary results in section 4.2.1, for which a reduced correlation was observed from scale level 1 to scale level 2 for all stocks. However, the different proportion of positive and negative deviations is relatively balanced. It should not be ignored that many stocks pairs also display reduced correlations with increasing time horizon.

This might indicate that the mechanisms underlying the correlations exhibit more complex structural changes over timescales than a simple increase in the global level of correlation. For example, with increasing timescale, correlations between stock groups (e.g., sectors) may evolve differently from within-group correlations. The stocks in the subsample of section 4.2.1 all belong to different sectors. The declining trend in correlations for those stocks may therefore represent an overall change in the relationship between their corresponding sectors.<sup>15</sup>

Another cause might be the presence of non-stationarities in the time series. It is natural to assume that correlations of different time horizons show different sensitivities regarding the dominant market state. In other words, short-term correlations behave differently to market changes than long-term correlations. Therefore, the synopsis of observations from different market states might lead to a mixture of different timescale correlation structures. For this reason, section 4.2.3 examines correlations with regard to their evolution in different market states.

### 4.2.3 Non-Crisis and Crisis Period Analysis

To further investigate the term-structure of correlations, I subdivide the observation period into market periods of non-crisis and crisis. For the period of analysis, I

---

<sup>15</sup>I leave analyzing the changes in the correlation substructure to future research as considering these relationships would go beyond the scope of this study.

qualitatively identify seven economic crises that significantly impacted the US stock market.<sup>16</sup>

Note that specifying crisis-period duration is undermined by a trade-off between long-term and short-term frequency characterization (besides the basic difficulty of defining a start and an end date of a crisis period). If crisis duration is defined as too short, longer-term cycles may be blended. In turn, wider time-window specification may lead to incorporating price dynamics outside the scope of a particular crisis. As a result, specifying the duration of crisis periods is clearly subject to debate. To ensure the robustness of the results, I used longer and shorter time intervals for different crisis and non-crisis periods. However, correlation structures remained mostly comparable for these adjusted periods.

Figure I.3 illustrates wavelet correlations for the subsample during crisis (red) and non-crisis (blue) market states. Similar to the observations for the full time series in Figure I.2, during crisis periods, correlations display an inverting curve from scale levels 1 to 4. Correlations are high for bi-daily frequencies and decrease when approaching monthly periods. While another high is observed for most stocks at scale level 5, correlations once again decrease for longer-term horizons (scale level 7).<sup>17</sup>

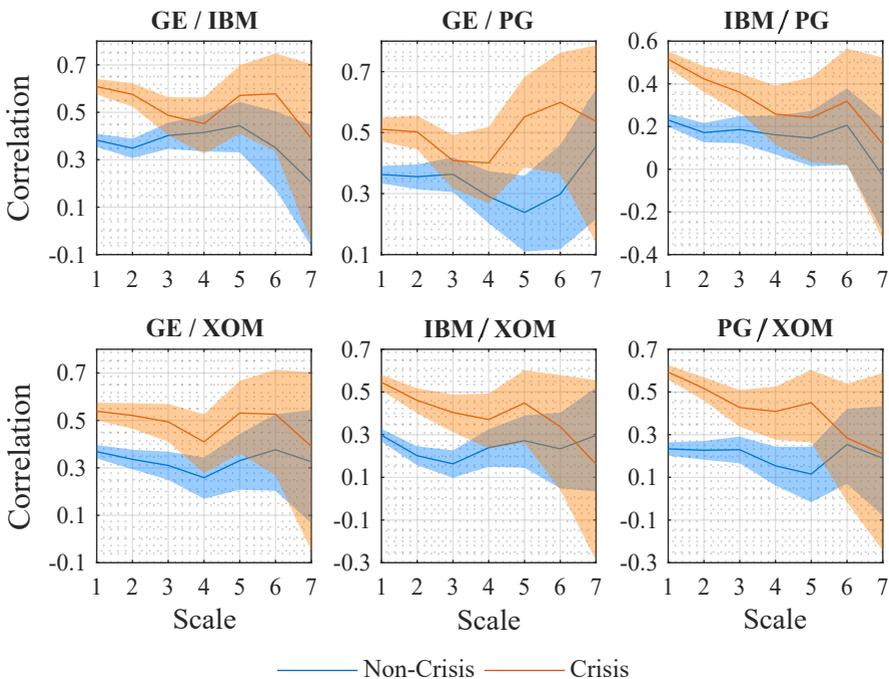
The correlation curves for non-crisis periods show no clear trend either. Some correlations increase with the scale level, others decrease, and some remain flat. However, unlike the correlation curve of crisis periods, there is little indication that correlations differ during non-crisis periods. Interestingly, the term-structure of correlations for crisis periods strongly resembles that of the full observation period (Figure I.2). Thus, the correlation structure for the full observation period seems to be largely influenced by the correlation structures of crisis periods. This pattern emerges despite the fact that market crises cover a shorter time span than non-crisis market phases. These findings underline the necessity of dividing the observation period into crisis

---

<sup>16</sup>The following time periods were used to classify crisis market phases: the US savings and loan crisis and Latin American crisis (Jan. 1980–Jun. 1980; Jan. 1981–Dec. 1982), the market phase after Black Monday (Oct. 1987–Apr. 1989), the early 1990s recession (Feb. 1989–Mar. 1991), the Asian crisis (Apr. 1997–Dec. 1998), the early 2000s recession (Mar. 2000–Dec. 2002), the global financial crisis of 2007 (Jul. 2007–Jun. 2009), the European sovereign debt crisis and the time period after the downgrading of America's credit rating (Jun. 2011–May 2012). The remaining time periods were classified as non-crisis periods.

<sup>17</sup>Again, no direct conclusion can be drawn about the shape of the correlation curve for longer timescales due to the now even broader confidence bands.

and non-crisis market states. The separation enables examining the non-crisis correlation structures in more detail while excluding the dominant influence of crisis period correlations.



**Figure I.3:** Wavelet correlations between four selected stocks: Exxon Mobil (XOM), General Electric (GE), International Business Machines (IBM), and Procter & Gamble (PG) during non-crisis and crisis periods. Notes: The blue and orange shaded regions in each graph represent the 95% confidence intervals; a Daubechies least asymmetric MODWT filter of length 8 was used to decompose returns.

Figure I.3 also allows comparing the correlations between non-crisis and crisis market states. Correlations at a timescale of 2–4 days (scale level 1) are significantly different during crisis and non-crisis market phases. Correlations in crisis states surpass correlations in non-crisis states for all six stock pairs. This deviation between correlations for different market states is consistent with the general expectation of

higher correlations during distressed market phases. Interestingly, correlations of non-crisis and crisis periods move closer together for all stocks as the timescale increases from bi-daily (scale level 1) to monthly frequencies (scale level 4). With the exception of GE/PG, this convergence in correlations roughly continues up to the largest scale level 7. These exemplary results suggest that medium- to long-term correlations are less exposed to short-term market perturbations (positive and negative shocks). Also, a common level of long-term correlation may be assumed.

Nevertheless, the construction of the test statistics does not allow drawing conclusions about the equality of correlations. Therefore, it can only be stated that no significant differences can be found between correlations of crisis and non-crisis market states at medium- to long-term frequencies.

Table I.4 extends the analysis of multiscale correlation structures for the different market phases to the complete stock sample. Time series are divided into the same non-crisis and crisis market periods. The level of significance of equality in correlations is set at 5% (two-sided test). Considering the different market phases for the overall sample directly complements the analysis of the complete observation period in Table I.3.

Compared to the results in Table I.3, fewer significant relative deviations exist between scale correlations. The measure accounting for differences in correlations at any timescale decreases from 50.08% for the full sample period to 38.22% (33.36%) for the non-crisis (crisis) period (bottom row in Table I.4). While this reduction is profound, it is most likely attributable to the broader confidence bounds resulting from subdividing the sample period.

Individually comparing timescale correlations provides more detailed information on those timescale correlations between which significant differences occur. Again, significant differences are only observed between correlations of the lowest timescales (scale levels 1 or 2) and those of the remaining timescales. No significant differences can be found between correlations at higher timescales. This holds for both market states.

**Table I.4:** Test of equality between correlations of different timescales during non-crisis and crisis periods, covering the complete data sample of 268 stocks.

	Non-Crisis						Crisis					
	Num. of sig. dev.	Perc. of sig. dev.	Positive sig. dev.	Perc. of pos. sig. dev.	Negative sig. dev.	Perc. of neg. sig. dev.	Num. of sig. dev.	Perc. of sig. dev.	Positive sig. dev.	Perc. of pos. sig. dev.	Negative sig. dev.	Perc. of neg. sig. dev.
Scale 1 / Scale 2	2,065	5.77	174	8.43	1,891	91.57	2,425	6.78	1,426	58.80	999	41.20
Scale 1 / Scale 3	6,265	17.51	198	3.16	6,067	96.84	3,585	10.02	1,847	51.52	1,738	48.48
Scale 1 / Scale 4	5,790	16.18	167	2.88	5,623	97.12	3,080	8.61	1,725	56.01	1,355	43.99
Scale 1 / Scale 5	3,342	9.34	337	10.08	3,005	89.92	1,637	4.58	464	28.34	1,173	71.66
Scale 1 / Scale 6	2,400	6.71	260	10.83	2,140	89.17	3,110	8.69	446	14.34	2,664	85.66
Scale 1 / Scale 7	1,547	4.32	219	14.16	1,328	85.84	1,418	3.96	314	22.14	1,104	77.86
Scale 2 / Scale 3	1,427	3.99	54	3.78	1,373	96.22	958	2.68	421	43.95	537	56.05
Scale 2 / Scale 4	2,468	6.90	104	4.21	2,364	95.79	1,867	5.22	971	52.01	896	47.99
Scale 2 / Scale 5	1,793	5.01	289	16.12	1,504	83.88	1,087	3.04	226	20.79	861	79.21
Scale 2 / Scale 6	1,607	4.49	267	16.61	1,340	83.39	2,547	7.12	265	10.40	2,282	89.60
Scale 2 / Scale 7	1,109	3.10	228	20.56	881	79.44	1,138	3.18	227	19.95	911	80.05
Scale 3 / Scale 4	361	1.01	100	27.70	261	72.30	436	1.22	254	58.26	182	41.74
Scale 3 / Scale 5	1,100	3.07	531	48.27	569	51.73	785	2.19	207	26.37	578	73.63
Scale 3 / Scale 6	973	2.72	355	36.49	618	63.51	2,053	5.74	150	7.31	1,903	92.69
Scale 3 / Scale 7	753	2.10	270	35.86	483	64.14	1,008	2.82	239	23.71	769	76.29
Scale 4 / Scale 5	331	0.93	207	62.54	124	37.46	246	0.69	39	15.85	207	84.15
Scale 4 / Scale 6	705	1.97	345	48.94	360	51.06	1,642	4.59	78	4.75	1,564	95.25
Scale 4 / Scale 7	580	1.62	267	46.03	313	53.97	862	2.41	187	21.69	675	78.31
Scale 5 / Scale 6	244	0.68	93	38.11	151	61.89	726	2.03	50	6.89	676	93.11
Scale 5 / Scale 7	513	1.43	192	37.43	321	62.57	493	1.38	172	34.89	321	65.11
Scale 6 / Scale 7	135	0.38	65	48.15	70	51.85	208	0.58	171	82.21	37	17.79
Any sig. dev. for given asset pairs	13,673	38.22					11,937	33.36				

Notes: The null hypothesis of the test states that for stocks  $p$  and  $q$ , timescale-correlations are equal  $\tilde{\rho}_{p,q}(\lambda_j) = \tilde{\rho}_{p,q}(\lambda_k)$  for scale level  $j \neq k$ ; for a given stock pair, the test statistics is defined by  $(z_j - z_k) / \sqrt{(\hat{N}_j - 3)^{-1} + (\hat{N}_k - 3)^{-1}}$ , where the variables  $\hat{N}_j$  and  $\hat{N}_k$  refer to the number of DWT coefficients and the coefficients  $z_j$  and  $z_k$  are the Fisher  $z$ -transformed correlations at different scale levels ( $j \neq k$ ); tests are conducted for all possible combinations of scale levels 1–7 and all possible combinations of stock pairs. For a given scale level combination  $j$  and  $k$ , "Number of sig. dev." summarizes the number of tests between all possible combinations of stock pairs that reject the null hypothesis  $\mathcal{R}(\lambda_j, \lambda_k) \equiv \sum_{(p < q)} \mathbb{I}[\{\tilde{\rho}_{p,q}(\lambda_j) \neq \tilde{\rho}_{p,q}(\lambda_k)\}_H]$ . "Perc. of sig. dev." determines the proportion of these significant deviations to total number of correlation coefficients  $\frac{M(M-1)}{2} = 35,778$ . "Positive sig. dev." and "Negative sig. dev." denote the number of positive  $\left(\sum_{(p < q)} \mathbb{I}[\{\tilde{\rho}_{p,q}(\lambda_j) > \tilde{\rho}_{p,q}(\lambda_k)\}_H]\right)$  and negative deviations  $\left(\sum_{(p < q)} \mathbb{I}[\{\tilde{\rho}_{p,q}(\lambda_j) < \tilde{\rho}_{p,q}(\lambda_k)\}_H]\right)$ . "Perc. of pos. sig. dev." and "Perc. of neg. sig. dev." designate the proportion of these deviations over all significant deviations  $\mathcal{R}(\lambda_j, \lambda_k)$ ; the bottom row summarizes the number of stock pairs with at least one significant deviation over all scale level combinations  $\left(\sum_{(p < q)} \mathbb{I}\left[\sum_{(j < k)} \mathbb{I}[\{\tilde{\rho}_{p,q}(\lambda_j) \neq \tilde{\rho}_{p,q}(\lambda_k)\}_H] > 0\right]\right)$  and its proportion to overall combinations; tests are conducted for level of significance  $p$  of 5%; a Daubechies least asymmetric MODWT filter of length 8 was used to decompose returns; (wavelet) correlations were calculated between 268 stocks for the period June 30, 1980 to June 30, 2018; for selected crisis and non-crisis periods, see Figure I.3.

Interestingly, compared to the full sample period in Table I.3, the share of negative deviations greatly increased for *non-crisis periods*. For example, the percentage of negative deviations rose from 60.86% to 91.57% for the relation scale 1/scale 2 and from 74.55% to 95.79% for the relation scale 2/scale 4. These negative deviations imply that almost all significant correlations increase from lower to higher timescales. In other words, correlations between these stocks increase over longer time horizons during regular market periods. Unlike the analysis of the reduced sample in Figure I.3, a clear trend is therefore observable in deviating correlations across timescales with this separation of market states.

In contrast, correlations in *crisis periods* are more equally balanced between positive and negative deviations at low timescales. Compared to the non-crisis period, correlations more often decrease from short-term to medium-term time horizons. The relations between lower timescale correlations generally even point to a slight inversion of the timescale-correlation curve. For example, the relation scale 1/scale 4 exhibits a 56.01% share in positive significant deviations. However, the imbalance between negative and positive deviations is not particularly pronounced.

Comparing short- and long-term correlations reveals that negative deviations again outnumber positive deviations. For example, for the relation scale 1/scale 6 the share of negative deviations rises to 85.66%. In accordance with the observation for non-crisis periods, this indicates a positive trend in the term-structure of deviating correlations for longer timescales. Note, however, that the results for the crisis period must be considered with caution because the overall number of significantly deviating correlations is comparably low.

Overall, there are multiple observations for changes in the term-structure of correlations for different market states. While correlations increase over timescales during regular market periods, this trend may be (slightly) inverted during distressed market phases. During these market periods, short-term correlations surge while medium- to long-term correlations experience a smaller relative increase or even remain stable.

Heterogeneous market hypotheses can again be used to explain the divergent characteristics of correlations in non-crisis and crisis periods. According to these hy-

potheses, differences in correlations result from market agents with different investment horizons internalizing information heterogeneously. On this premise, these agents are likely to react different to information depending on the prevailing market state. In particular, newsbreaks may have different impacts in crises and calm market states.

Following this intuition, in non-crisis periods, short-term correlations are influenced to a higher degree by idiosyncratic price movements of individual stocks than long-term correlations. However, as the time horizon increases, long-term macroeconomic trends become more relevant. As a result, overall stock correlation increases and a positive trend in correlations is observed with increasing time horizon.

By contrast, during distressed market phases, general market news equally affect all market agents. Newsbreaks dominate the price behavior of all stocks and penetrate all timescale levels. As a consequence, stock prices move in tandem and are susceptible to the same general market risk. This induces correlations over all time horizons to increase. Thus, the positive trend in the term-structure of correlations weakens or even disappears. This effect is particularly pronounced for the short-term time horizon.

This correlation mechanism can be compared to the term-structure of interest rates (yield curve). Similar to the yield curve, correlation structures may be influenced by expectations about future interrelations in the stock market and steered by different factor compositions both at the short- and long-term horizon and in different market states.

In summary, this section has shown that the term-structure of correlations varies in different market phases. While in times of crisis correlations increase with the time horizon, the trend is less clear in non-crisis periods. At shorter timescales, the term-structure of correlations is evenly balanced between positive and negative deviations. This study argues that these differences may be attributed to the heterogeneity of market agents and consequently to the heterogeneity of information.

Despite many significant differences in correlations, the majority of timescale correlation pairs show no statistically significant differences. Thus, no heterogeneity is detected for numerous timescale correlations. Moreover, the test design generally

inflates type I errors due to the experimentwise error rate when considering the union of the deviating correlations and the results in this section should be viewed with a certain degree of caution.

Consequently, the analysis should be extended to considering the multivariate structure of correlation to legitimate the above results. This allows extending and generalizing findings from individual correlation analysis to the overall correlation structure. If analyzing the overall correlation structure reveals similar properties, this underpins the findings of individual correlation analysis. The following section conducts this multivariate analysis.

### 4.3 Overall Multiscale Correlation

Fernández-Macho (2012) introduced a statistical tool for studying the multiscale properties of correlations in a multivariate dataset. The tool enables deriving a measure that summarizes all stock correlations in a single metric. This unified measure — referred to as wavelet multiple correlation (WMC) — serves as a general proxy of the overall relationship between stocks at a particular timescale. Thus, the method provides a description of the overall structural behavior of correlations on a scale-by-scale basis. Consolidating timescale correlations into a single measure enables a holistic and simplified interpretation of the timescale dependency structure between stocks.

According to Fernández-Macho (2012), the method not only simplifies interpreting the complex stock market system but also offers other benefits beyond the analysis of individual correlations. As explained, bundling tests for equality in correlations of different timescales may cause type I errors to increase. The WMC method overcomes this deficiency. In addition, it prevents misidentifying spurious correlations, which only emerge due to possible hidden variables in the multivariate dataset. Therefore, WMC-based correlation analysis is a useful addition to the previous examination of individual correlations (see previous section).<sup>18</sup>

---

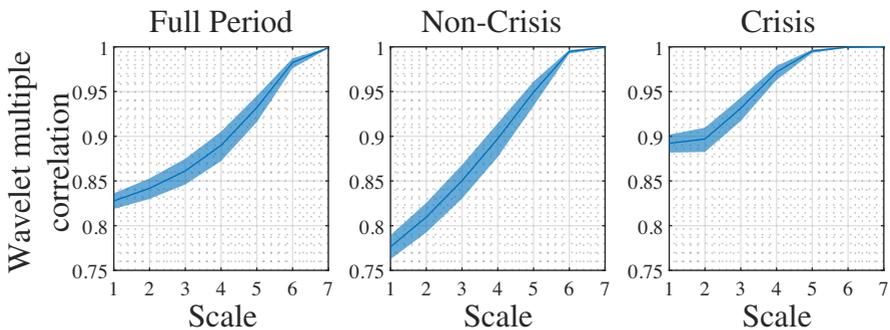
<sup>18</sup>However, investigating both individual correlations and the overall correlation structure is needed to obtain a complete picture of the timescale behavior of stocks. Thus, WMC-based analysis does not replace investigating individual correlations.

The wavelet multiple correlation  $\bar{\rho}(\cdot)$  for timescale  $\lambda_j$  is defined by

$$\bar{\rho}(\lambda_j) = \sqrt{1 - \frac{1}{\max \{diag(\mathbf{C}(\lambda_j)^{-1})\}}}, \quad (9)$$

where  $\mathbf{C}(\lambda_j)^{-1}$  refers to the inverse of the wavelet correlation matrix  $\mathbf{C}(\lambda_j) = \{\tilde{\rho}_{p,q}(\lambda_j); p = 1, \dots, M; q = 1, \dots, M\}$  at timescale  $\lambda_j$  and where the  $\max \{diag(\cdot)\}$  operator extracts the largest element from the diagonal of the matrix. Similar to wavelet correlation coefficients, wavelet multiple correlation coefficients lie between  $-1$  and  $+1$ . Moreover, confidence bounds for the WMC are given by the same interval as in formula 8. The wavelet correlation coefficient  $\tilde{\rho}_{p,q}(\lambda_j)$  simply needs to be replaced by the WMC coefficient  $\bar{\rho}(\lambda_j)$ .

Figure I.4 displays the WMC coefficients and the corresponding 95% confidence bounds across timescales and for different market periods. The results generally show growing stock correlation with increasing timescale. This trend can be observed for the full period, as well as for the non-crisis and the crisis periods.



**Figure I.4:** Wavelet multiple correlation of the 268 stocks for the full period and the non-crisis and crisis periods. Notes: The blue-shaded region in each graph represents the 95% confidence intervals; the full period covers June 30, 1980 to June 30, 2018; a Daubechies least asymmetric MODWT filter of length 8 was used to decompose returns.

The WMC is quite high for all market phases and timescales. For the *full period*, the WMC starts at approximately 0.83 at scale level 1 (bi-daily time horizon) and increases with the timescale. The WMC approaches a value close to 1 at the highest scale level 7 (half-yearly/yearly time horizon). This implies that stocks are highly interconnected, and that this interrelationship even strengthens for longer time horizons.

For the *non-crisis periods*, the correlation at scale level 1 is slightly lower than for the full period. However, similar to the full period, the WMC coefficient rises with the time horizon and approaches a value close to 1 at the largest timescale. As a result, the trend in the term-structure of correlations for non-crisis periods is slightly more positive than for the full period.

These results are consistent with the findings from analyzing the timescale characteristics of individual correlations (see previous section, Table I.4). However, the previous analysis only allowed interpreting those individual relationships that significantly differ between the timescale correlations. Considering the WMC measure enables extending this interpretation to describe the overall (global) change in correlations. Analyzing the WMC coefficients demonstrates that, in non-crisis periods, stock market correlations tend to increase with an increasing time horizon.

Interestingly, for *crisis periods*, we also observe higher correlations with increasing time horizon. However, Figure I.4 shows a flat section of the curve and thus highly similar WMC coefficients at the lowest two timescales. In other words, the WMC graph shows no clear trend for changes between correlations of the shortest time horizons. This observation concurs with the findings of the previous analysis of individual correlations (section 4.2), where also no clear trend was found between correlations at the lowest timescales.<sup>19</sup> Similarly, in both analyses increasing correlations are only detected at larger timescales.

The analyses of the individual correlations and of the WMC coefficients in crisis periods are generally consistent except for some slight differences. The emerging

---

<sup>19</sup>The balance between the positive and negative deviations of the individual timescale correlations in Table I.4 seems to translate to a static WMC metric. This can be explained by the fact that the WMC reflects the overall correlation structure. It summarizes the behavior of all correlations. Because the positive and negative deviations of the individual timescale correlations in Table I.4 approximately balance, the WMC for the respective timescales therefore remains more or less static.

ascent of the WMC curve starts at a lower timescale (scale level 3) than the corresponding increase in the individual correlations (scale level 5) reported in Table I.4.

One possible reason for the discrepancy between the analyses could be a greater similarity of correlations in times of crisis. These smaller differences complicate identifying significant deviations in the analysis of individual correlations. This does not necessarily affect the WMC metric to an equal extent as this metric considers an aggregate of all correlations. Hence, the WMC method may be better suited to detecting changes in the overall structure of correlations.

Conversely, analyzing individual correlations allows for better identifying finer changes in the correlation structure. Because the WMC metric measures the general degree of correlation, it might ignore more subtle changes (smoothing) and harbors the risk of masking finer details of the structural interrelationship between stocks: For example, with an increasing time horizon, correlations may change in the market as a whole (global). However, they may also change within (intra-group) and between (inter-group) certain groups of stocks. In other words, the structure of the correlation itself may change with timescale. Therefore, the discrepancies between individual correlation analysis and WMC analysis may point to deeper structural differences in the correlations, independently of the evolution of the general correlation level.

Overall, the WMC graph of crisis periods exhibits a high level of correlation across all timescales. Even at low timescales, correlations are pronounced.<sup>20</sup> Compared to non-crisis periods, this results in a flatter term-structure of correlations with a less positive trend. In other words, correlations across different timescales approach a similarly high level. Again, this is consistent with the findings in section 4.2, which also reported a less significant trend in the term-structure of correlations at short- to medium-term time horizons. In periods of financial distress, stocks are highly synchronized and share a common behavior at all timescales. This discovery has important practical implications: In times of crisis, the benefits of diversification

---

<sup>20</sup>In fact, correlations of the lowest timescales rise the most. This can also be explained by the correlation metric being limited to a maximum value of 1. Correlations of timescales already showing high values may therefore only marginally increase. In contrast, low correlations have enough space to rise.

disappear at all time horizons. The possibility of diversifying a portfolio reduces, regardless of the investment horizon under consideration.

To summarize, analyzing the WMC coefficients generalizes and supplements the findings in section 4.2. Together, they point to a dissimilar structure in correlations for different market phases.

For the non-crisis period, a general increase in correlations is observed. The stock market thus shows a stronger collective behavior. This suggests that macroeconomic factors gain in relevance with increasing time horizon. Consequently, in non-crisis periods, the diversification potential decreases with increasing time horizon. In contrast, in crisis periods, the general level of correlation is high at all timescales. Stocks show a strong collective behavior over all time horizons. Only minor differences can be observed, in particular for low timescales. However, analyzing the WMC and the individual correlations reveals slightly different correlation dynamics at lower timescales.

A major weakness of the WMC metric is its sensitivity to estimation errors in the correlation matrix. This sensitivity arises because the WMC metric relies on the inverse of the correlation matrix for its derivation (see formula 9). As less information is available for longer-term processes (fewer observations), the measurement error of correlations generally increases with timescale. As a result, the reliability of the WMC measure tends to decline with increasing time horizon.

Laloux et al. (1999), as well as Plerou, Gopikrishnan, Rosenow, Amaral and Stanley (1999), showed that correlations already contain a high degree of randomness, regardless of the timescale considered. Therefore, even for low timescales, the WMC measure may already be noticeably influenced by random deflections in the time series. Ultimately, the individual correlation estimates are not free of this problem either. This deficiency raises the question about finding a more robust method for extracting the structural properties of stock correlations.

Advanced filtration methods such as RMT and graph theory could be interesting alternatives for studying the structure of correlations. Specifically, RMT enables filtration of random components from the signal and hence would be less susceptible to the aforementioned measurement errors. However, the applicability of RMT to

wavelet correlations requires further investigation. In particular, it must first be clarified whether the filtered components enjoy reasonable economic interpretation at different time horizons. Similarly, considering substructures via graph theory needs further tests, which would go beyond the scope of this study.

#### 4.4 Robustness

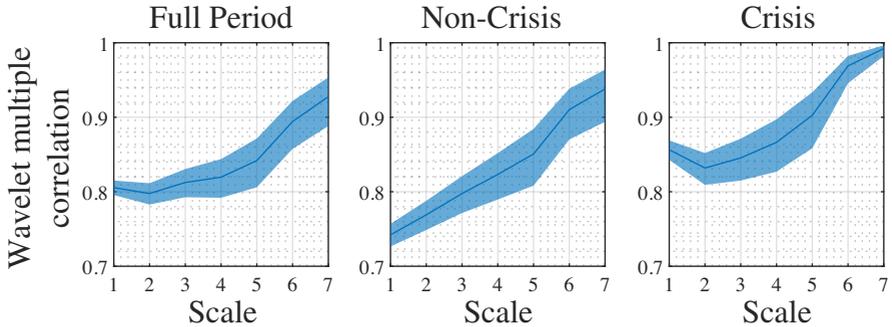
This section substantiates the previous findings using different robustness tests. These include changing the analyzed time interval and considering different wavelet filters.

Figure I.5 shows the WMC graphs for the extended period from December 31, 1960 to June 30, 2018. Hence, the observation period was extended by approximately 19 years compared to the analysis in the main section.<sup>21</sup> However, as a full time series history is required, this significantly reduces the breadth of the data sample to a mere 65 stocks.

The results for this extended period generally confirm the previous findings regarding timescale-variant correlation behavior. Again, in non-crisis periods, the graph indicates increasing global correlation with the time horizon. In contrast, the results for crisis periods slightly deviate from those in section 4.3. Particularly, correlations of the weekly time horizon (scale level 2) fall below those of the bi-daily time horizon (scale level 1). However, this deviation is neither particularly large nor significant. Even if the deviation were significant, the observation would confirm the previous assumption: In times of crisis, the trend in the average correlation is less pronounced for increasing timescales and correlations of different timescales generally converge. Further, Table I.7 (see appendix) shows the corresponding statistics for the comparison of individual correlations. These results are also largely consistent with the findings of the analysis from June 30, 1980 to June 30, 2018.

---

<sup>21</sup>The recession in the years 1960–61 (Apr. 1960–Feb. 1961), the 1973 oil crisis (Oct. 1973–Mar. 1974), and the 1979 energy crisis (Jan. 1979–Dec. 1980) were considered crisis market phases for these additional years.



**Figure I.5:** Wavelet multiple correlation of the 65 stocks for the full period and the non-crisis and crisis periods, covering the extended sample period Dec. 31, 1960 to Jun. 30, 2018. Notes: The blue-shaded region in each graph represents the 95% confidence intervals; additional crisis periods considered: 1960–61 recession (Apr. 1960–Feb. 1961), 1973 oil crisis (Oct. 1973–Mar. 1974), 1979 energy crisis (Jan. 1979–Dec. 1980); a Daubechies least asymmetric MODWT filter of length 8 was used to decompose returns.

Next, I study the timescale characteristics of correlations for the shorter period between December 31, 1999 to June 30, 2018. The shortening of the observation period makes it possible to examine a larger number of stocks and thus to better reflect the overall stock market behavior. In total, the reduction of the observation period allows for an expansion of the sample to 505 stocks. However, the shortening of the observation period means that correlation coefficients at higher timescales can no longer be estimated reliably. Hence, Table I.5 only compares correlations up to scale level 5 (bi-monthly time horizons).

For non-crisis periods, Table I.5 again illustrates an increase in significant correlations for bi-daily to longer time horizons. However, the trend is less evident and even reverses for higher timescales. This finding disagrees with the results for the longer observation period in the main analysis. However, the lower accuracy of the correlation estimates is likely to cause these contradictions at higher timescales. Nevertheless, the presence of a survivorship bias cannot be ruled out completely. The deviating results might also be caused by a recent change in the general market structure.

**Table I.5:** *Test of equality between correlations of different timescales for the extended data sample of 505 stocks, covering the period December 31, 1999 to June 30, 2018.*

	Non-Crisis						Crisis					
	Num. of sig. dev.	Perc. of sig. dev.	Positive sig. dev.	Perc. of pos. sig. dev.	Negative sig. dev.	Perc. of neg. sig. dev.	Num. of sig. dev.	Perc. of sig. dev.	Positive sig. dev.	Perc. of pos. sig. dev.	Negative sig. dev.	Perc. of neg. sig. dev.
Scale 1 / Scale 2	7,181	5.69	1,804	25.12	5,377	74.88	8,091	6.41	7,379	91.20	712	8.80
Scale 1 / Scale 3	13,733	10.88	3,773	27.47	9,960	72.53	8,430	6.68	5,512	65.39	2,918	34.61
Scale 1 / Scale 4	8,378	6.64	2,621	31.28	5,757	68.72	11,811	9.36	10,837	91.75	974	8.25
Scale 1 / Scale 5	7,302	5.78	5,736	78.55	1,566	21.45	4,253	3.37	3,547	83.40	706	16.60
Scale 2 / Scale 3	3,918	3.10	1,757	44.84	2,161	55.16	1,998	1.58	466	23.32	1,532	76.68
Scale 2 / Scale 4	4,833	3.83	2,191	45.33	2,642	54.67	5,304	4.20	4,343	81.88	961	18.12
Scale 2 / Scale 5	6,537	5.18	5,650	86.43	887	13.57	2,363	1.87	1,558	65.93	805	34.07
Scale 3 / Scale 4	2,054	1.63	1,043	50.78	1,011	49.22	2,490	1.97	2,299	92.33	191	7.67
Scale 3 / Scale 5	6,032	4.78	5,293	87.75	739	12.25	2,321	1.84	1,746	75.23	575	24.77
Scale 4 / Scale 5	2,997	2.37	2,697	89.99	300	10.01	630	0.50	158	25.08	472	74.92
Any sig. dev. for given asset pairs	41,450	32.83					32,299	25.58				

*Notes:* See description in Table I.4. However, the total number of correlation coefficients differs:  $\frac{M(M-1)}{2} = 128, 778$ ; selected crisis periods: early 2000s recession (Mar. 2000–Dec. 2002), global financial crisis (Jul. 2007–Jun. 2009), European sovereign debt/American credit risk crisis (Jun. 2011–May 2012). All remaining time periods were classified as non-crisis periods.

Lastly, I test the results of this study by employing wavelet filters of different families and using different wavelet filter lengths to calculate wavelet correlations. Besides a Daubechies least asymmetric filter of length 16 LA(16), I also use a symmetric Daubechies filter of length 4 DB(4) (different filter family). The greater length of the LA filter allows capturing longer term cycles in the time series. However, it also results in a loss of time resolution. On the other hand, the shorter length of the DB filter leads to a loss in the coverage of longer cycle characteristics. It does, however, allow for better time resolution, which is important for classifying non-crisis and crisis periods. Moreover, the symmetry of the DB filter allows replicating and uncovering different aspects of the signal.

Table I.6 presents the analysis for the LA(16) filter. The analysis of the DB(4) filter is displayed in Table I.8 (appendix). Results are largely consistent with the findings of the main section. However, LA(16) filter analysis generally shows higher significances compared to the analysis in Table I.4. Applying the longer filter thus even reinforces the main results of this study.

**Table I.6:** *Test of equality between correlations of different timescales using an alternative LA(16) wavelet filter, covering the period June 30, 1980 to June 30, 2018.*

	Non-Crisis						Crisis					
	Num. of sig. dev.	Perc. of sig. dev.	Positive sig. dev.	Perc. of pos. sig. dev.	Negative sig. dev.	Perc. of neg. sig. dev.	Num. of sig. dev.	Perc. of sig. dev.	Positive sig. dev.	Perc. of pos. sig. dev.	Negative sig. dev.	Perc. of neg. sig. dev.
Scale 1 / Scale 2	2,788	7.79	290	10.40	2,498	89.60	3,382	9.45	2,092	61.86	1,290	38.14
Scale 1 / Scale 3	7,104	19.86	266	3.74	6,838	96.26	4,142	11.58	1,966	47.46	2,176	52.54
Scale 1 / Scale 4	6,283	17.56	229	3.64	6,054	96.36	3,657	10.22	2,152	58.85	1,505	41.15
Scale 1 / Scale 5	3,947	11.03	436	11.05	3,511	88.95	1,930	5.39	659	34.15	1,271	65.85
Scale 1 / Scale 6	2,908	8.13	335	11.52	2,573	88.48	4,007	11.20	555	13.85	3,452	86.15
Scale 1 / Scale 7	1,908	5.33	364	19.08	1,544	80.92	1,717	4.80	393	22.89	1,324	77.11
Scale 2 / Scale 3	2,227	6.22	105	4.71	2,122	95.29	1,651	4.61	628	38.04	1,023	61.96
Scale 2 / Scale 4	2,998	8.38	183	6.10	2,815	93.90	2,443	6.83	1,299	53.17	1,144	46.83
Scale 2 / Scale 5	2,285	6.39	409	17.90	1,876	82.10	1,409	3.94	369	26.19	1,040	73.81
Scale 2 / Scale 6	2,056	5.75	358	17.41	1,698	82.59	3,426	9.58	324	9.46	3,102	90.54
Scale 2 / Scale 7	1,440	4.02	397	27.57	1,043	72.43	1,437	4.02	281	19.55	1,156	80.45
Scale 3 / Scale 4	792	2.21	273	34.47	519	65.53	926	2.59	603	65.12	323	34.88
Scale 3 / Scale 5	1,571	4.39	781	49.71	790	50.29	1,131	3.16	415	36.69	716	63.31
Scale 3 / Scale 6	1,308	3.66	488	37.31	820	62.69	2,741	7.66	245	8.94	2,496	91.06
Scale 3 / Scale 7	1,083	3.03	485	44.78	598	55.22	1,253	3.50	331	26.42	922	73.58
Scale 4 / Scale 5	713	1.99	413	57.92	300	42.08	610	1.70	141	23.11	469	76.89
Scale 4 / Scale 6	1,017	2.84	486	47.79	531	52.21	2,461	6.88	123	5.00	2,338	95.00
Scale 4 / Scale 7	889	2.48	453	50.96	436	49.04	1,185	3.31	244	20.59	941	79.41
Scale 5 / Scale 6	538	1.50	191	35.50	347	64.50	1,540	4.30	150	9.74	1,390	90.26
Scale 5 / Scale 7	804	2.25	367	45.65	437	54.35	722	2.02	225	31.16	497	68.84
Scale 6 / Scale 7	358	1.00	202	56.42	156	43.58	539	1.51	435	80.71	104	19.29
Any sig. dev. for given asset pairs	16,525	46.19					15,156	42.36				

Notes: See description in Table I.4.

## 5 Conclusion

Analysis of frequency dynamics in US stocks has shown that the stock relationships change over different time horizons and market states. Specifically, in *non-crisis periods*, correlations between stocks pairs are generally lower at shorter than at longer time horizons (timescales). Although examining individual correlations could only significantly establish this relationship for a minority of correlation pairs, analyzing wavelet multiple correlation (WMC) coefficients has revealed that this relation is equally valid for the overall stock market.

A different structure emerges for correlations in *crisis periods*. Fewer significant deviations between correlations of different timescales are detected. Moreover, the few significant deviations relate to small differences between bi-daily and monthly time horizons. Only for longer-term time horizons are correlations once again found to increase with timescale. A similar behavior is observed for the change in overall correlation structure (WMC) in crisis periods. However, the WMC analysis signals increasing correlations at a lower timescale than the analysis of individual correlations suggests. Regardless of these differences, the results indicate that after correlations at shorter time horizons initially stagnate, the general level of correlation rises with increasing time horizon.

These findings have important implications for risk management and portfolio decision-making. For example, they suggest that different risk assessments or portfolio allocations may be necessary regarding an investor's time horizon. These strategies might need to be adjusted depending on the prevalent market state. Further, the results also provide intuitions as to why some market models deliver dissimilar results for the consideration of different sampling intervals.

A possible explanation for timescale-variant correlations was found in heterogeneous market theories. Based on these theories, the increasing correlations in non-crisis times can be attributed to the growing influence of macroeconomic factors. On the other hand, at low timescales, idiosyncratic factors have a higher impact on stock interrelations. This relationship changes in times of crisis, during which all stocks react collectively irrespectively of the time horizon. In this market state, the market

factor is dominant across all timescales. However, investigating a subsample has revealed that relationships between individual stock pairs may run counter to the general trend. This suggests that the underlying structure is much more complex and may not be explained by a single factor. Methods such as graph theory might allow more efficiently deciphering the underlying timescale-variant dynamics in the complex system of correlations.

A major limitation of this study is the large proportion of randomness hidden in the correlations. This complicates identifying underlying trends and influences the accuracy of the measures used. Advanced filtration methods could separate relevant from non-relevant information in correlations. In this context, random matrix theory in particular presents a promising avenue of research. In addition to filtering noise from correlations, it also allows for identifying underlying factors in the correlation structure. Thus, the method is likely to enable more adequately reflecting the complexities in correlations described above.

This study underpins the existence of timescale-variant correlations and implies the presence of a term-structure of correlations. As such, it has laid the foundation for further investigating the timescale-dependent mechanisms underlying stock correlations.

## Appendix

**Table I.7:** *Test of equality between correlations of different timescales for the extended data sample of 65 stocks, covering the period December 31, 1960 to June 30, 2018.*

	Non-Crisis						Crisis					
	Num. of sig. dev.	Perc. of sig. dev.	Positive sig. dev.	Perc. of pos. sig. dev.	Negative sig. dev.	Perc. of neg. sig. dev.	Num. of sig. dev.	Perc. of sig. dev.	Positive sig. dev.	Perc. of pos. sig. dev.	Negative sig. dev.	Perc. of neg. sig. dev.
Scale 1 / Scale 2	208	10.00	13	6.25	195	93.75	119	5.72	70	58.82	49	41.18
Scale 1 / Scale 3	549	26.39	36	6.56	513	93.44	215	10.34	124	57.67	91	42.33
Scale 1 / Scale 4	471	22.64	28	5.94	443	94.06	205	9.86	130	63.41	75	36.59
Scale 1 / Scale 5	273	13.13	11	4.03	262	95.97	87	4.18	15	17.24	72	82.76
Scale 1 / Scale 6	222	10.67	27	12.16	195	87.84	203	9.76	34	16.75	169	83.25
Scale 1 / Scale 7	110	5.29	33	30.00	77	70.00	79	3.80	10	12.66	69	87.34
Scale 2 / Scale 3	156	7.50	9	5.77	147	94.23	40	1.92	19	47.50	21	52.50
Scale 2 / Scale 4	201	9.66	11	5.47	190	94.53	133	6.39	86	64.66	47	35.34
Scale 2 / Scale 5	141	6.78	7	4.96	134	95.04	79	3.80	12	15.19	67	84.81
Scale 2 / Scale 6	146	7.02	27	18.49	119	81.51	152	7.31	16	10.53	136	89.47
Scale 2 / Scale 7	76	3.65	33	43.42	43	56.58	58	2.79	5	8.62	53	91.38
Scale 3 / Scale 4	22	1.06	4	18.18	18	81.82	35	1.68	15	42.86	20	57.14
Scale 3 / Scale 5	81	3.89	20	24.69	61	75.31	51	2.45	9	17.65	42	82.35
Scale 3 / Scale 6	72	3.46	29	40.28	43	59.72	127	6.11	6	4.72	121	95.28
Scale 3 / Scale 7	49	2.36	31	63.27	18	36.73	41	1.97	0	0.00	41	100.00
Scale 4 / Scale 5	30	1.44	8	26.67	22	73.33	18	0.87	0	0.00	18	100.00
Scale 4 / Scale 6	50	2.40	30	60.00	20	40.00	101	4.86	3	2.97	98	97.03
Scale 4 / Scale 7	34	1.63	26	76.47	8	23.53	38	1.83	1	2.63	37	97.37
Scale 5 / Scale 6	14	0.67	9	64.29	5	35.71	31	1.49	4	12.90	27	87.10
Scale 5 / Scale 7	41	1.97	31	75.61	10	24.39	23	1.11	3	13.04	20	86.96
Scale 6 / Scale 7	11	0.53	9	81.82	2	18.18	0	0.00	0	0.00	0	0.00
Any sig. dev. for given asset pairs	1,007	48.41					737	35.43				

**Table I.7:** (continued)

*Notes:* The null hypothesis of the test states that for stocks  $p$  and  $q$ , timescale-correlations are equal  $\tilde{\rho}_{p,q}(\lambda_j) = \tilde{\rho}_{p,q}(\lambda_k)$  for scale level  $j \neq k$ ; for a given stock pair, the test statistics is defined by  $(z_j - z_k) / \sqrt{(\hat{N}_j - 3)^{-1} + (\hat{N}_k - 3)^{-1}}$ , where the variables  $\hat{N}_j$  and  $\hat{N}_k$  refer to the number of DWT coefficients and the coefficients  $z_j$  and  $z_k$  are the Fisher  $z$ -transformed correlations at different scale levels ( $j \neq k$ ); tests are conducted for all possible combinations of scale levels 1–7 and all possible combinations of stock pairs. For a given scale level combination  $j$  and  $k$ , "Number of sig. dev." summarizes the number of tests between all possible combinations of stock pairs that reject the null hypothesis  $\left( \mathcal{R}(\lambda_j, \lambda_k) \equiv \sum_{(p < q)} \mathbb{I} \left[ \{ \tilde{\rho}_{p,q}(\lambda_j) \neq \tilde{\rho}_{p,q}(\lambda_k) \}_H \right] \right)$ . "Perc. of sig. dev." determines the proportion of these significant deviations to total number of correlation coefficients  $\frac{M(M-1)}{2} = 2,080$ . "Positive sig. dev." and "Negative sig. dev." denote the number of positive  $\left( \sum_{(p < q)} \mathbb{I} \left[ \{ \tilde{\rho}_{p,q}(\lambda_j) > \tilde{\rho}_{p,q}(\lambda_k) \}_H \right] \right)$  and negative deviations  $\left( \sum_{(p < q)} \mathbb{I} \left[ \{ \tilde{\rho}_{p,q}(\lambda_j) < \tilde{\rho}_{p,q}(\lambda_k) \}_H \right] \right)$ . "Perc. of pos. sig. dev." and "Perc. of neg. sig. dev." designate the proportion of these deviations over all significant deviations  $\mathcal{R}(\lambda_j, \lambda_k)$ ; the bottom row summarizes the number of stock pairs with at least one significant deviation over all scale level combinations  $\left( \sum_{(p < q)} \mathbb{I} \left[ \sum_{(j < k)} \mathbb{I} \left[ \{ \tilde{\rho}_{p,q}(\lambda_j) \neq \tilde{\rho}_{p,q}(\lambda_k) \}_H \right] > 0 \right] \right)$  and its proportion to overall combinations; tests are conducted for level of significance  $p$  of 5%; selected crisis periods: 1960–61 recession (Apr. 1960–Feb. 1961), 1973 oil crisis (Oct. 1973–Mar. 1974), 1979 energy crisis (Jan. 1979–Dec. 1980), US savings and loan/Latin American crisis (Jan. 1980–Jun. 1980; Jan. 1981–Dec. 1982), Black Monday (Oct. 1987–Apr. 1989), early 1990s recession (Feb. 1989–Mar. 1991), Asian crisis (Apr. 1997–Dec. 1998), early 2000s recession (Mar. 2000–Dec. 2002), global financial crisis (Jul. 2007–Jun. 2009), European sovereign debt/American credit risk crisis (Jun. 2011–May 2012). All remaining time periods were classified as non-crisis periods; a Daubechies least asymmetric MODWT filter of length 8 was used to decompose returns.

**Table I.8:** *Test of equality between correlations of different timescales using an alternative DB(4) wavelet filter, covering the period June 30, 1980 to June 30, 2018.*

	Non-Crisis							Crisis						
	Num. of sig. dev.	Perc. of sig. dev.	Positive sig. dev.	Perc. of pos. sig. dev.	Negative sig. dev.	Perc. of neg. sig. dev.	Num. of sig. dev.	Perc. of sig. dev.	Positive sig. dev.	Perc. of pos. sig. dev.	Negative sig. dev.	Perc. of neg. sig. dev.		
Scale 1 / Scale 2	1,292	3.61	72	5.57	1,220	94.43	1,413	3.95	784	55.48	629	44.52		
Scale 1 / Scale 3	4,890	13.67	122	2.49	4,768	97.51	2,579	7.21	1,397	54.17	1,182	45.83		
Scale 1 / Scale 4	4,855	13.57	83	1.71	4,772	98.29	2,251	6.29	1,263	56.11	988	43.89		
Scale 1 / Scale 5	2,601	7.27	187	7.19	2,414	92.81	1,215	3.40	311	25.60	904	74.40		
Scale 1 / Scale 6	1,770	4.95	178	10.06	1,592	89.94	2,079	5.81	321	15.44	1,758	84.56		
Scale 1 / Scale 7	1,121	3.13	98	8.74	1,023	91.26	973	2.72	194	19.94	779	80.06		
Scale 2 / Scale 3	562	1.57	8	1.42	554	98.58	385	1.08	198	51.43	187	48.57		
Scale 2 / Scale 4	1,673	4.68	41	2.45	1,632	97.55	1,056	2.95	572	54.17	484	45.83		
Scale 2 / Scale 5	1,227	3.43	149	12.14	1,078	87.86	717	2.00	131	18.27	586	81.73		
Scale 2 / Scale 6	1,071	2.99	164	15.31	907	84.69	1,542	4.31	175	11.35	1,367	88.65		
Scale 2 / Scale 7	740	2.07	96	12.97	644	87.03	750	2.10	151	20.13	599	79.87		
Scale 3 / Scale 4	90	0.25	17	18.89	73	81.11	112	0.31	62	55.36	50	44.64		
Scale 3 / Scale 5	529	1.48	206	38.94	323	61.06	374	1.05	74	19.79	300	80.21		
Scale 3 / Scale 6	554	1.55	187	33.75	367	66.25	1,162	3.25	80	6.88	1,082	93.12		
Scale 3 / Scale 7	423	1.18	102	24.11	321	75.89	630	1.76	135	21.43	495	78.57		
Scale 4 / Scale 5	58	0.16	37	63.79	21	36.21	51	0.14	5	9.80	46	90.20		
Scale 4 / Scale 6	341	0.95	173	50.73	168	49.27	774	2.16	32	4.13	742	95.87		
Scale 4 / Scale 7	281	0.79	97	34.52	184	65.48	470	1.31	95	20.21	375	79.79		
Scale 5 / Scale 6	55	0.15	23	41.82	32	58.18	194	0.54	8	4.12	186	95.88		
Scale 5 / Scale 7	210	0.59	62	29.52	148	70.48	213	0.60	77	36.15	136	63.85		
Scale 6 / Scale 7	17	0.05	5	29.41	12	70.59	15	0.04	14	93.33	1	6.67		
Any sig. dev. for given asset pairs	9,869	27.58					7,795	21.79						

*Notes:* See description in Table I.7. However, the sample covers 268 stocks and the total number of correlation coefficients differs:  $\frac{M(M-1)}{2} = 35,778$ . Selected crisis periods: US savings and loan/Latin American crisis (Jan. 1980–Jun. 1980; Jan. 1981–Dec. 1982), Black Monday (Oct. 1987–Apr. 1989), early 1990s recession (Feb. 1989–Mar. 1991), Asian crisis (Apr. 1997–Dec. 1998), early 2000s recession (Mar. 2000–Dec. 2002), global financial crisis (Jul. 2007–Jun. 2009), European sovereign debt/American credit risk crisis (Jun. 2011–May 2012). All remaining time periods were classified as non-crisis periods.



# Chapter II

## Multiscale Analysis of the Underlying Structures in US Stock Correlations: A Wavelet-Based Random Matrix Theory Approach

*Christian Vial*

### 1 Introduction

Stock markets are *complex systems* characterized by non-trivial interactions between many components. The structures in the interdependencies between stocks are typically assumed to be invariant across different time horizons (timescales). In other words, correlations and their defining structures are expected to remain identical irrespective of whether we study daily or monthly price intervals. However, empirical evidence on multiscale asset dependencies indicates that degrees of correlation change across time horizons (see, e.g., Epps, 1979; Tumminello, Di Matteo, Aste & Mantegna, 2007; Borghesi, Marsili & Micciché, 2007). Similarly, the distributions of stock returns generally exhibit a multi-scaling behavior (see, e.g., Di Matteo, 2007; Mantegna & Stanley, 2004).

The notion of timescale-variant correlations agrees well with economic intuition and is often connected to the underlying structural properties. For example, long-term dynamics in the correlation structure are frequently explained by changes in common systematic (macroeconomic) factors. In contrast, short-term changes in correlations are often associated with the occurrence of singular events or with idiosyncratic price movements (Conlon et al., 2018).

The present study aims to identify changes in the interdependencies between stocks and the structure of cross-correlations across different time horizons. To this end, it develops a method for filtering relevant information from the complexities of the

system across different timescales. These timescales have common mechanisms, which help to coherently explain the dependency structure between stocks.

Understanding the factors that govern the system of stocks and their influence across different time horizons is crucial for many practical applications (e.g., portfolio and risk management). Information about the structural dynamics of stock correlations can help to assess the benefits of diversification across different timescales. In practice, investors have different decision-making time horizons and operate on different investment timescales. At one end of the investment process are fundamentalist investors who focus on long-term investment horizons. At the other end are day traders and intraday traders with short-term investment horizons (Gençay et al., 2010; In & Kim, 2013; Müller et al., 1993). Understanding the timescale-dependent structural relationship between stocks helps both types of investors to synchronize their investment decisions in accordance with their planned investment horizon. Similarly, insights gained from analyzing structural timescale dynamics enable better understanding the macro- and microeconomic forces driving stock prices.

This study uses wavelet transformation to decompose return series into the components of different timescales (scale-by-scale decomposition). From these components, I derive the cross-correlation matrix to examine the interbehavior of stocks over different time horizons. This decomposition helps to distinguish short- from long-term comovements between stocks.

The correlation matrix is a useful metric for extracting meaningful information about the interaction between stocks. However, analyzing and interpreting cross-correlation is a complex task (Plerou, Gopikrishnan, Rosenow, Amaral & Stanley, 2000). It may be difficult to identify apparent timescale-dependent structures in the complex web of interactions with a large number of stocks. Similarly, estimating empirical cross-correlation is beset by many complications. Market conditions change over time and as a result cross-correlation between stocks may be non-stationary (Fenn et al., 2011; Plerou et al., 2002). Moreover, the finite length of the time series introduces measurement noise into cross-correlation estimates. In particular, fewer observations (lower signal content) are available for estimating cross-correlations for longer time horizons. Hence, the problem of measurement noise is even more pronounced for

cross-correlation estimates at higher timescales.<sup>1</sup> These limitations introduce randomness to cross-correlations (Plerou et al., 2001, 2002, 2000).

It is therefore important to devise methods able to extract relevant information from the system and to filter the non-random statistical fluctuations inherent in empirical cross-correlation matrices, specifically when considering longer timescale dynamics (Conlon, Crane & Ruskin, 2008; Conlon, Ruskin & Crane, 2009; Laloux et al., 1999). Accordingly, this study combines wavelet decomposition with factorization and random matrix theory (RMT) to explore the multi-scaling and multivariate cross-correlation properties of financial time series. It is unclear which components of the cross-correlation matrix provide genuine information about the correlation structure. RMT has emerged as a useful tool to assess the non-random properties in a high-dimensional multivariate system (Laloux et al., 1999; Laloux, Cizeau, Potters & Bouchaud, 2000; Plerou et al., 1999).<sup>2</sup> RMT describes the statistical properties of matrices with independent random elements and provides many theoretical results for these so-called random matrices. Comparing the statistics of the empirical cross-correlation matrix with the properties of these random matrices provides information about randomness in the system. Therefore, contrasting the RMT's theoretical prediction with the empirical observed correlations allows separating information (incongruity with RMT) from noise (conformity with RMT) in the empirical (wavelet) correlation matrix.

Additionally, factorization of the cross-correlation matrix identifies a reduced number of common components, which largely explain variation in the system. Thus, the suggested procedure parsimoniously represents the underlying structures in a cross-correlation matrix. This low-dimensional representation of the multivariate data helps to detect latent common factors in stock returns. Factorization thus enables a condensed and streamlined analysis of the underlying structural relationships.

---

<sup>1</sup>Denosing techniques help to distinguish signal from noise and measurement errors (see Gençay et al., 2002). The technique employed in this study may be seen as a special form of denosing technique for multivariate data.

<sup>2</sup>Originally, RMT was designed by Wigner (1951b, 1951a, 1955), Dyson (1962), Mehta and Dyson (1963), Dyson and Mehta (1963), and Mehta (2004), and others to describe the statistics of energy levels of complex nuclei (eigenvalues) in many-body quantum systems. Instead of predicting the detailed sequence of energy levels in the nucleus, they focused on the general distribution of energy levels (Conlon et al., 2009; Plerou et al., 2002).

As a result, structures in stock return interactions may be uncovered that might otherwise have remained hidden to the observer.

This study contributes to existing research in four important ways. First, it analyzes the structural changes of correlations in the US stock market across different time horizons. While the structure of US stock correlations has been investigated before, analysis has mostly been limited to one timescale (e.g., Plerou et al., 2002; Laloux et al., 2000) or some few assets (see, e.g., Gençay et al., 2002). The present study closes this gap and provides a multivariate analysis of the underlying structural relationship in the correlations for a large set of US stocks across different timescales. Results show significant changes in the dependency structure in the US stock market across different time horizons.

Second, this study uses RMT to analyze the correlation matrices of wavelet decomposed time series to handle high-dimensional data and to identify relevant structural components. To the best of my knowledge, this is the first study to analyze the combination of RMT and wavelet decomposition in more detail. Using a series of simulations, I show that the theoretical predictions of RMT can also be applied to wavelet-decomposed time series. The assumptions underlying the adjustment of parameters in the theoretical distributions are shown to match the simulation outcomes. A third contribution involves using the eigendecomposition (factorization) of the wavelet correlation matrices to identify latent structures in the correlation matrix. This study differs from previous research by investigating changes in those latent structures across different timescales. By studying the timescale dynamics of correlations, I show that the largest eigenvalue (which is found to be associated with a general market factor) changes with the time horizon. However, analysis of different market states reveals that only in normal market periods the largest eigenvalue increases with timescale. This relation breaks down during turbulent market periods, where no significant differences in the largest eigenvalue can be observed. These findings have important implications for both investment management and the general understanding of financial markets.

Finally, this study contributes to ongoing research by applying a broad set of statistical tests to underpin the findings obtained from factorization of the correlation ma-

trix. While factorization (e.g., Principal Component Analysis or Explanatory Factor Analysis) is used abundantly in financial research, statistical tests are rarely reported. For example, the confidence intervals of eigenvalues are often missing. This is particularly critical in view of the general publication standards, which mandate that confidence intervals always ought to be reported (Larsen & Warne, 2010). Similarly, the classical criteria for determining the number of factors to be retained from factorization are often arbitrary in nature (e.g., Kaiser-Guttman (1954) criterion and scree test Cattell (1966)). In contrast, RMT uses theoretical results to define the number of factors with relevant information.

The remainder of this study is organized as follows. Section 2 reviews the literature on multiscale analysis of correlations structures. Section 3 briefly outlines wavelet theory and the methodology for deriving multiscale correlations. This theory is complemented by describing factorization and by introducing random matrix theory. These methods allow filtering relevant information from the wavelet correlation matrices. Section 4 describes the empirical data, namely, the time series of the historical constituents of the S&P 500 with a full track history. Extending RMT to wavelet correlation matrices is tested in section 5. Based on these results, I analyze the structural dependencies in the US stock market across different time horizons. Section 6 provides a summary and gives inputs for future research.

## 2 Literature Review

The analysis of correlation structures in stock markets using RMT dates back to Laloux et al. (1999, 2000) and Plerou et al. (1999, 2000, 2001, 2002). Laloux et al. (1999, 2000) studied the eigenvalue distribution of correlations for an S&P 500 sample of 406 stocks. They analyzed this sample for the period 1991–1996 based on daily return observations. Laloux et al. (1999, 2000) demonstrated that the eigenvalue and eigenvector distribution of the empirical observations strongly agree with theoretical predictions from RMT. On the premise of RMT, they concluded that less than 6% of eigenvalues appear to carry information. However, these eigenvalues were found to be responsible for 26% of the volatility in the correlation structure.

Similarly, Plerou et al. (1999, 2000, 2001, 2002) analyzed US stock correlations of 30-min and daily returns using RMT and detected comparable properties for eigenvalue and eigenvector distributions.<sup>3</sup> In accordance with Laloux et al.'s (1999, 2000) findings, they observed only a small portion of the largest eigenvalues (approximately 2%) to deviate from RMT predictions. Plerou et al. (1999, 2000, 2001, 2002) substantiated these results by investigating additional statistical properties of the empirical eigenvalues for which they also found good agreement with RMT. According to their results, the correlation matrix of stock returns contains a large amount of randomness.

Gopikrishnan, Rosenow, Plerou and Stanley (2000, 2001) showed that the largest eigenvalues have meaningful economic interpretation. More precisely, they reported that the largest eigenvalue globally influences all stocks and thus represents a collective market mode. In contrast, they found that the next largest eigenvalues (eigenvectors) are associated with conventional business sectors (see also Kim & Jeong, 2005).

Several studies confirmed these results and interpretations for other stock markets: among others, for the Brazilian market (Sandoval, Bruscatto & Venezuela, 2012), the German market (Drozd, Grümmer, Górski, Ruf & Speth, 2000; Drozd, Grümmer, Ruf & Speth, 2001; Kwapień, Drozd & Speth, 2003), the Indian market (Kulkarni & Deo, 2007; Pan & Sinha, 2007), the Japanese market (Utsugi, Ino & Oshikawa, 2004), and the South African market (Wilcox & Gebbie, 2004, 2007).

Gopikrishnan et al. (2000, 2001) also investigated the scaling properties of the time series obtained from projecting the original time series on the eigenvectors. They found the autocorrelation function of those time series to exhibit power-law decay. This indicates that correlations persist over long timescales. However, they did not examine these results in more detail.

Timescale characteristics in stocks were further explored for intraday dynamics and different markets (see, e.g., Borghesi et al., 2007). Kwapień, Drozd and Speth (2004) used high-frequency tick-by-tick data for the American and the German stock

---

<sup>3</sup>Plerou et al. (1999) analyzed the cross-correlation matrix of the 30-min returns of 1,000 US stocks for the period 1994–1995. Plerou et al. (2000, 2001, 2002) extended this analysis to 30-min returns of 881 US stocks for the period 1996–1997 and daily returns of 422 US stocks for the period 1962–1996.

markets and examined changes in magnitude of the largest eigenvalue for timescales ranging from seconds up to two days. Their results indicated an increase in the dominant eigenvalue with timescale. Correlations were thus found to be timescale-variant at low data frequencies.<sup>4</sup> Kwapien et al. (2004) further discovered that the market factor (collective market behavior) already emerges at short timescales — such as minutes or even intra-minute timescales.

Characteristics of stock correlations of higher timescales were explored by Nakayama and Iyetomi (2009) using daily price data for the Japanese stock market. They applied Fourier transformation to this daily data to construct correlation matrices at each frequency and analyzed those matrices using RMT. They found that the collective behavior of stock prices appears for timescales longer than one day and that eigenvalues vary over different timescales.

In a related approach, Conlon et al. (2009) used wavelet decomposition to obtain correlation matrices for different timescales. They applied this decomposition to a rolling window of daily and intraday price data of the Dow Jones Euro Stoxx 50. This analysis allowed them to study the evolution of the largest eigenvalue (retrieved from correlation matrices) for timescales of 3 to 11 days. Similar to Nakayama and Iyetomi (2009), Conlon et al. (2009) reported correlations to depend on both time and timescale.

The present study differs from Conlon et al. (2009) in many ways: First, it investigates a larger number of stocks and a different stock market (US stock market). Second, it also explores the properties of subdominant eigenvalues besides the characteristics of the largest eigenvalues.<sup>5</sup> Third, it analyzes correlation structures for longer time horizons (up to 128–256 days) and directly compares eigenvalues across different timescales. Fourth, it also introduces statistical tests and investigates the applicability of RMT to wavelet correlation in more detail. Finally, it explores the characteristics of eigenvalues for non-crisis and crisis market states.

Previously, Sharkasi, Crane, Ruskin and Matos (2006) studied the behavior of eigenvalue dynamics in crisis and non-crisis periods for mature and emerging markets

---

<sup>4</sup>This timescale-variant behavior of correlation accords with the well-reported Epps (1979) effect.

<sup>5</sup>Conlon (2009) extended this analysis and provided a rudimentary description of the dynamics of the second- and third-largest eigenvalues. However, the present study more closely examines the characteristics of these subdominant eigenvalues.

(stock market indices). The authors reported different dynamics between the largest and the subdominant eigenvalues and concluded that the second and third largest eigenvalues provide additional information on market comovement (especially for emerging markets). They also found that the dynamics of and between those eigenvalues differ in crisis and non-crisis periods. Sharkasi et al. (2006) assumed that these changing dynamics are due to a heterogeneous behavior of market agents. Further, they suggested that the largest eigenvalues may help explain crash dynamics. However, the authors did not interpret the eigenvalues in more detail. Although they initially decomposed the time series, the evolution of eigenvalues of different timescales was not explored.<sup>6</sup>

In a similar study, Fenn et al. (2011) investigated the temporal evolution of eigenvalues for correlations between stock market indices. They also found changing relative contributions of different eigenvalues in crisis periods. In particular, the authors reported that the variance proportion, which the largest eigenvalue explains, increases in the event of a crisis.

More recently, RMT has been combined with other filtering methods to study the timescale characteristics of stock market correlations. Such methods include detrended cross-correlation analysis (Wang, Xie, Chen, Yang & Yang, 2013), multifractal detrended (cross-correlation/fluctuation) analysis (Kumar & Deo, 2012; Lin, Shang & Zhou, 2014), and topological approaches (Eom, Oh, Jung, Jeong & Kim, 2009). Most of these studies reported timescale-variant correlations in stock markets and the same macroeconomic structures underlying these correlations.

### 3 Methodology

To analyze the structure of the cross-correlations in the US stock market across different time horizons, I combine wavelet decomposition with random matrix theory. The first part of this section briefly presents wavelet theory and introduces the basic

---

<sup>6</sup>They recomposed the filtered time series before studying the eigenvalues.

concepts of wavelet decomposition.<sup>7</sup> The second part explores the statistical properties of correlation matrices in the context of RMT.

### 3.1 Wavelet Theory

Wavelet transformation is a mathematical tool for studying the multiscale properties of time series, i.e., to investigate the properties of time series across different time horizons (timescales). It decomposes a time series into hierarchical sets of components that relate to a certain timescale. Each element of a set is associated with a location  $t$  and a timescale  $\lambda_j \equiv 2^{j-1}$  where  $j$  refers to the level of decomposition. Hence, the decomposition simultaneously represents a time series in the time and in the timescale domain and enables separating short- from long-term characteristics in the time series.

Wavelet transformation is based on a discrete wavelet filter  $\{h_l; l = 0, \dots, L - 1\}$  and scaling filter  $\{g_l; l = 0, \dots, L - 1\}$  in  $\mathbb{R}^L$  where  $L$  refers to the filter width.<sup>8</sup> The wavelet filter  $\{h_l\}$  is a high-pass filter. It passes dynamics at low timescales (high frequencies) and attenuates dynamics at high timescales (low frequencies). The filter is defined so as to fulfill three basic conditions: It must sum to zero ( $\sum_{l=0}^{L-1} h_l = 0$ ), have unit energy ( $\sum_{l=0}^{L-1} h_l^2 = 1$ ), and be orthogonal to its even shifts ( $\sum_{l=-\infty}^{\infty} h_l h_{l+2n} = 0$ ) for all integers  $n \neq 0$ . These conditions show that the wavelet filter resembles a differencing operator that identifies changes in the data.<sup>9</sup>

---

<sup>7</sup>For a more detailed discussion of the mathematical properties of wavelets, see Daubechies (1992), Percival and Walden (2000), or Ramsey (2002). Gençay et al. (2002), In and Kim (2013), and Gallegati, Gallegati, Ramsey and Semmler (2014) have applied wavelet transformation to study various problems in economics and finance. Finally, Ramsey (1999) as well as Chakrabarty, De, Gunasekaran and Dubey (2015) have reviewed pertinent literature on application of wavelet transformation in economics and finance.

<sup>8</sup>The length of the filter  $L$  must be even. Further, we define the filters for  $l < 0$  and  $l \geq L$  such that  $g_l = h_l = 0$ .

<sup>9</sup>Several functions fulfill the necessary conditions of a wavelet filter. This study uses the Daubechies Least-Asymmetric (symmlet) wavelet filter of length  $L = 8$ . This wavelet filter is generally considered as an appropriate filter for the decomposition of financial time series and is often used in studies on interdependencies in financial assets (see, e.g., Gençay et al., 2001a; Gallegati, 2005; Ranta, 2010; Dajčman, 2013; Wang, Xie & Chen, 2017; Conlon et al., 2018). Similarly, the filter length  $L = 8$  is generally considered adequate to reflect the timescale features of the signal (Najeeb, Bacha & Masih, 2015). Smaller filter lengths may lead to leakages and produce misleading results (see Percival & Walden, 2000).

In contrast, the scaling filter  $\{g_l\}$  is a high pass-filter. The filter attenuates dynamics at low timescales (high frequencies) and passes dynamics at high timescales (low frequencies). It is given by the quadrature mirror relationship  $g_l = (-1)^{l+1} h_{L-1-l}$ , from which three basic properties immediately follow:  $\left| \sum_{l=0}^{L-1} g_l \right| = \sqrt{2}$ ,  $\sum_{l=0}^{L-1} g_l^2 = 1$ , and  $\sum_{l=-\infty}^{\infty} g_l g_{l+2n} = 0$  for all integers  $n \neq 0$ .<sup>10</sup> These conditions exemplify that the scaling filter is a local averaging operator that captures long-term data variations (Gençay et al., 2010).

These wavelet and scaling filters can be used to decompose a time series into components related to variations at certain timescales. This is achieved through the so-called pyramid algorithm formulated by Mallat (1989). Below, I first present the discrete wavelet transform (DWT) followed by the maximal overlap discrete wavelet transform (MODWT).

Let  $\{r_t; t = 0, \dots, N-1\}$  in  $\mathbb{R}^N$  be a vector of returns with dyadic length  $N$ . With each iteration  $j$ , the pyramid algorithm filters (convolves) the return series with the wavelet and scaling filters to obtain the DWT wavelet  $\{W_{j,t}; t = 0, \dots, N_{j-1}-1\}$  and scaling coefficients  $\{V_{j,t}; t = 0, \dots, N_{j-1}-1\}$  for timescale  $\lambda_j$ :

$$W_{j,t} = \sum_{l=0}^{L-1} h_l V_{j-1, 2t+1-l \bmod N_{j-1}}, \quad V_{j,t} = \sum_{l=0}^{L-1} g_l V_{j-1, 2t+1-l \bmod N_{j-1}}, \quad (1)$$

where  $N_j \equiv N 2^{-j}$  and  $V_{0,t} \equiv \{r_t\}$  for  $t = 0, \dots, N_j - 1$ . The wavelet coefficients  $W_{j,t}$  capture high frequency dynamics, whereas the scaling coefficients  $V_{j,t}$  represent the long-term trends in the time series (low frequency dynamics). This filtering operation may be applied  $J$  times where  $J = \log_2 N$ . Note that the time series is downsampled at each timescale  $\lambda_j$ , i.e., every second data point is removed with an iteration in the algorithm.

The modulus operator in formula 1 is necessary to deal with the boundary of the finite time series. The operation results in circular filtering. This, however, introduces a

---

<sup>10</sup>These conditions imply that the gain function of the wavelet filter  $\mathcal{H}(f)$  and the scaling filter  $\mathcal{G}(f)$  have to fulfill  $\mathcal{H}(f) + \mathcal{G}(f) = 2$  for all frequencies  $f$ . If the wavelet filter represents a high-pass filter, the scaling filter therefore forms a low-pass filter.

bias to the boundary coefficients as they are mixed with the circulated values. This bias needs to be considered in estimating statistical moments.<sup>11</sup>

Percival and Mofjeld (1997) have shown that the discrete wavelet transform approximately decorrelates a wide variety of time series (Percival & Mofjeld, 1997). This decorrelation property is of critical relevance for many statistical tests. It is used below to combine RMT results with wavelet theory (section 5.2).

This study specifically applies the maximal overlap discrete wavelet transform, which is a modification of the DWT. In contrast to the DWT, the MODWT retains all values and does not downsample the filtered output. Although the MODWT introduces redundancy, gives up orthogonality, and no longer decorrelates the time series, the method has many useful properties. Three of these properties are specifically relevant for this study: First, MODWT can handle any sample size and is not similarly restricted to dyadic sample sizes  $N$ . This feature is important because I analyze a non-dyadic time series. Second, with MODWT the wavelet and scaling coefficients can be aligned with events and features in the original time series. This property is needed to divide the transformed time series into different market phases and to ensure the proper alignment of the wavelet coefficients with these periods. Third, MODWT provides an asymptotically more efficient variance estimator than the DWT (Gençay et al., 2002; Percival & Mofjeld, 1997; Percival & Walden, 2000). The MODWT algorithm, which is used to retrieve the wavelet coefficients  $\tilde{W}_{j,t}$  and scaling coefficients  $\tilde{V}_{j,t}$  for time  $t$  and scale level  $j$ , is given by

$$\tilde{W}_{j,t} = \sum_{l=0}^{L-1} \tilde{h}_l \tilde{V}_{j-1,t-2^{j-1}l \bmod N}, \quad \tilde{V}_{j,t} = \sum_{l=0}^{L-1} \tilde{g}_l \tilde{V}_{j-1,t-2^{j-1}l \bmod N} \quad (2)$$

for  $t = 0, \dots, N - 1$  and where  $\tilde{h}_l \equiv h_l/\sqrt{2}$  and  $\tilde{g}_l \equiv g_l/\sqrt{2}$  are rescaled wavelet and scaling filters and  $\tilde{V}_{0,t} \equiv \{r_t\}$ . The filters are rescaled to preserve the energy in the system. This is necessary because the filter output is no longer downsampled.

---

<sup>11</sup>In this study, I use a reflection boundary, i.e., the time series is reflected at the last observation. This procedure produces continuity in the function and allows us to avoid assuming periodicity in the time series. However, the reflection does not alter the sample mean or variance (Gençay et al., 2002; In & Kim, 2013; Percival & Walden, 2000).

Let  $\{r_{p,t}\}$  and  $\{r_{q,t}\}$  be two return time series and  $\tilde{W}_{p,j,t}$  and  $\tilde{W}_{q,j,t}$  the corresponding MODWT wavelet coefficients for timescale  $\lambda_j$ . The MODWT estimator for the wavelet variance, covariance, and correlation at timescale  $\lambda_j$  are then given by

$$\tilde{v}_x(\lambda_j) = \frac{1}{\tilde{N}_j} \sum_{t=L_j-1}^{N-1} [\tilde{W}_{x,j,t}]^2 \quad x \in \{p, q\}, \quad (3)$$

$$\tilde{v}_{p,q}(\lambda_j) = \frac{1}{\tilde{N}_j} \sum_{t=L_j-1}^{N-1} \tilde{W}_{p,j,t} \tilde{W}_{q,j,t}, \quad (4)$$

$$\tilde{\rho}_{p,q}(\lambda_j) = \frac{\tilde{v}_{p,q}(\lambda_j)}{\tilde{v}_p(\lambda_j) \tilde{v}_q(\lambda_j)}. \quad (5)$$

where  $L_j \equiv (2^j - 1)(L - 1) + 1$  represents the wavelet filter length at scale level  $j$ , and  $\tilde{N}_j \equiv N - L_j + 1$  refers to the number of coefficients that are not impaired by the boundary.

The wavelet correlation coefficient  $\tilde{\rho}_{p,q}(\lambda_j)$  describes the relationship between the return series  $\{r_{p,t}\}$  and  $\{r_{q,t}\}$  on a scale-by-scale basis. Analogous to the usual correlation coefficients,  $\tilde{\rho}_{p,q}(\lambda_j)$  lies in the interval  $[-1, 1]$ . The wavelet correlation coefficients between  $M$  assets for timescale  $\lambda_j$  can be collected in a correlation matrix  $\mathbf{C}(\lambda_j) = \{\tilde{\rho}_{p,q}(\lambda_j); p = 1, \dots, M; q = 1, \dots, M\}$  describing the interactions between  $M$  stocks at timescale  $\lambda_j$ . This study uses daily returns and employs seven levels of decomposition  $J = 7$ . Therefore, the wavelet correlation matrices reflect stock price interdependencies on time horizons of 2–4 days ( $\mathbf{C}(\lambda_1)$ ), 4–8 days ( $\mathbf{C}(\lambda_2)$ ), 8–16 days ( $\mathbf{C}(\lambda_3)$ ), 16–32 days ( $\mathbf{C}(\lambda_4)$ ), 32–64 days ( $\mathbf{C}(\lambda_5)$ ), 64–128 days ( $\mathbf{C}(\lambda_6)$ ), and 128–256 days ( $\mathbf{C}(\lambda_7)$ ).

## 3.2 Random Matrix Theory

Random matrix theory describes the universal statistical properties of random matrices. In the context of financial applications, RMT is used to identify and filter relevant information from empirical cross-correlations. To this end, the universal

properties of random correlation matrices (i.e., a correlation matrix from mutually uncorrelated time series) are contrasted with the statistical properties of the empirical cross-correlation matrix. Congruency of the properties of the empirical correlation matrix with the RMT's universal predictions indicate that the empirical correlation matrix is defined by randomness. In contrast, deviations of the properties of the empirical cross-correlation matrix from those of random correlation matrices point to non-random (informative) characteristics in the dependency structure. Comparing the statistical properties of empirical cross-correlations with RMT's universal properties thus helps to differentiate random components from genuine contributions in multivariate data (Fenn et al., 2011; Plerou et al., 2002). Consequently, RMT analysis allows extracting information and identifying structures that might otherwise have been obscured by the presence of random noise in the data.

Many RMT results relate to the statistical properties of correlation matrices in terms of their eigenvalues and eigenvectors. Therefore, the following section first describes factorization of the correlation matrix through eigendecomposition. This is followed by the presentation of some fundamental results from RMT.

### 3.2.1 Eigendecomposition

As introduced in section 3.1, let  $\mathbf{C}(\lambda_j)$  be the  $M \times M$  empirical (wavelet) cross-correlation matrix collecting the (wavelet) correlation coefficients  $\{\rho_{p,q}(\lambda_j); p = 1, \dots, M; q = 1, \dots, M\}$  for  $M$  stocks and timescale  $\lambda_j$ . For the sake of simplicity, the timescale variable  $\lambda_j$  is dropped below unless reference is made to a specific timescale. Further, let  $\mathbf{C}(\lambda_0)$  refer to the correlation matrix of the original untransformed time series.<sup>12</sup> The eigenvalues  $\{\mathcal{E}_k; k = 1, \dots, M\}$  and the standardized eigenvectors  $\{\mathbf{u}_k; k = 1, 2, \dots, M\}$  of the empirical cross-correlation matrix  $\mathbf{C}$  are then given by

$$\mathbf{C}\mathbf{u}_k = \mathcal{E}_k\mathbf{u}_k. \quad (6)$$

These eigenvalues and eigenvectors are then rank-ordered with respect to the eigenvalues' magnitude,  $\mathcal{E}_1 \geq \mathcal{E}_2 \geq \dots \geq \mathcal{E}_M$ .

<sup>12</sup>The correlation between any two stocks for the untransformed return series is calculated with Pearson's correlation coefficient. For higher timescales, formula 5 provides the respective wavelet correlation coefficients.

### 3.2.2 Eigenvalue Distribution of Random Matrices

Statistical properties of random matrices — specifically those of Wishart matrices — are well known (Chatterjee & Chakrabarti, 2006; Dyson & Mehta, 1963; Rojkova & Kantardzic, 2007; Sengupta & Mitra, 1999). A Wishart matrix is a random matrix  $\mathbf{R} = \frac{1}{N} \mathbf{A} \mathbf{A}^T$  where  $\mathbf{A}$  is a  $M \times N$  matrix of  $M$  mutually uncorrelated time series of length  $N$ . The elements of these time series are independent and identically Gaussian distributed real random variables with zero mean and unit variance.

The Marčenko-Pastur distribution provides a limiting distribution of the eigenvalues of a Wishart matrix. Particularly, as  $N \rightarrow \infty$  and  $M \rightarrow \infty$  such that  $Q \equiv \frac{M}{N} \geq 1$  is fixed, the probability density function  $P_{\mathbf{R}}(\mathcal{E})$  of the eigenvalues  $\mathcal{E}$  of the random matrix  $\mathbf{R}$  follows the distribution (Marčenko & Pastur, 1967; Mehta, 2004):

$$P_{\mathbf{R}}(\mathcal{E}) = \frac{Q}{2\pi} \frac{\sqrt{(\mathcal{E}_+ - \mathcal{E})(\mathcal{E} - \mathcal{E}_-)}}{\mathcal{E}}, \quad (7)$$

for  $\mathcal{E}$  within  $\mathcal{E}_- \leq \mathcal{E} \leq \mathcal{E}_+$ , where  $\mathcal{E}_+$  and  $\mathcal{E}_-$  are the lower and upper eigenvalue bounds given by

$$\mathcal{E}_{\pm} = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}. \quad (8)$$

Note that equations 7 and 8 are only valid in the limit  $N \rightarrow \infty$  and  $M \rightarrow \infty$ . Hence, for finite  $N$  or  $M$ , there is a non-zero probability of finding eigenvalues larger than  $\mathcal{E}_+$  and smaller than  $\mathcal{E}_-$  (Fenn et al., 2011).

The eigenvalue distribution  $P_{\mathbf{C}}(\mathcal{E})$  of the empirical cross-correlation matrix  $\mathbf{C}$  can now be compared to the eigenvalue distribution  $P_{\mathbf{R}}(\mathcal{E})$  of the random matrix. Large deviations between the empirical eigenvalues and the theoretical boundaries indicate the existence of informative cross-correlations and the presence of distinct cross-dependency patterns in the empirical time series.<sup>13</sup> Eigenvalues that appear within

<sup>13</sup>The method can also be used as an alternative to traditional techniques for deciding on the number of factors to retain in a factor analysis, such as the Kaiser-Guttman (1954) criterion or the scree test (Cattell, 1966).

the maximum and minimum eigenvalue interval  $\mathcal{E}_- \leq \mathcal{E} \leq \mathcal{E}_+$  are generally referred to as the *bulk of the eigenvalue spectrum* (Laloux et al., 1999; Plerou et al., 1999; Rosenow, Plerou, Gopikrishnan & Stanley, 2002).

Analyzing and comparing eigenvalues further requires estimating confidence intervals. These confidence intervals help to identify ranges of plausible values for the eigenvalue estimates. Larsen and Warne (2010) provided theoretical results for the confidence intervals around individual eigenvalues (by assuming Wishart distribution):

$$\mathcal{E}_k \pm \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \cdot \left( \sqrt{\frac{2\mathcal{E}_k^2}{N}} \right), \quad (9)$$

where  $\Phi^{-1}$  refers to the probit function and  $\alpha$  to the level of significance.

### 3.2.3 Distribution of the Bulk of the Eigenvalue Spectrum

Plerou et al. (2002) conjectured that comparing the eigenvalue distribution of the empirical cross-correlation with that of random matrices is not sufficient to measure randomness in the eigenvalue spectrum. Random matrices with similar eigenvalue distributions can have significantly different eigenvalue correlations. Conversely, matrices with different eigenvalue distributions may show similarities in the structure of eigenvalue correlations (Guhr, Müller-Groeling & Weidenmüller, 1998; Plerou et al., 2002).

Hence, assessing randomness requires additional tests: Correlations of eigenvalues of the empirical observations should be compared to those correlations of random matrices. These correlations are described by the distribution of the spacings between adjacent rank-ordered eigenvalues (eigenvalue spacings). Comparison requires constructing the random matrices in accordance with the properties of the empirical cross-correlation matrix  $\mathbf{C}$ . The latter is symmetric and consists of real elements. Therefore, a random matrix must also be real and symmetric. The off-diagonal elements of such a random matrix, however, can be chosen from an arbitrary distribution with zero mean.

As the order of the matrix  $M \rightarrow \infty$ , the eigenvalue correlations (in terms of local mean eigenvalue spacings) of such a random matrix display the universal properties of the so-called Gaussian orthogonal ensemble (GOE) (Mehta, 2004; Plerou et al., 2002; Sinha, Chatterjee, Chakraborti & Chakrabarti, 2010).<sup>14,15</sup> These universal properties require uniform average spacing between adjacent rank-ordered eigenvalues throughout the eigenvalue spectrum. However, local intervals between eigenvalues vary as a function of the magnitude of the eigenvalues. This necessitates applying a transformation to the eigenvalues, which ensures uniformity of the eigenvalue spacing. This transformation is known as *unfolding* and maps the eigenvalues  $\mathcal{E}_k$  to new variables, so-called *unfolded eigenvalues*  $\tilde{\mathcal{E}}_k$  (Brody et al., 1981; Guhr et al., 1998).

Deriving unfolded eigenvalues requires defining the cumulative distribution function, which specifies the number of eigenvalues in the interval  $\mathcal{E}_i \leq \mathcal{E}$ :

$$\eta(\mathcal{E}) = M \int_{-\infty}^{\mathcal{E}} P_{\mathbf{E}}(\mathcal{E}') d\mathcal{E}', \quad (10)$$

where  $P_{\mathbf{E}}(\mathcal{E}')$  describes the probability density of eigenvalues, and where  $M$  corresponds to the total number of eigenvalues. This function  $\eta(\mathcal{E})$  can then be separated into an average and a fluctuating component:

$$\eta(\mathcal{E}) = \eta_{av}(\mathcal{E}) + \eta_{fluc}(\mathcal{E}). \quad (11)$$

Given that the probability density of the fluctuating part is zero on average ( $P_{fluc} \equiv d\eta_{fluc}(\mathcal{E})/d\mathcal{E} = 0$ ), the average eigenvalue density is given by  $d\eta_{av}(\mathcal{E})/d\mathcal{E}$ . The

---

<sup>14</sup>The GOE describes an ensemble of random symmetric matrices. The entries of these matrices are statistically independent (up to the symmetricity constraint) and distributed according to a Gaussian probability measure. The variances of the entries are defined such that the ensemble is invariant under conjugation by orthogonal matrices. The cross-correlation matrix  $\mathbf{R}$  (introduced in the previous section) is not strictly a GOE-type matrix. However, the eigenvalue correlations of  $\mathbf{R}$  in the bulk of the spectrum can be shown to be generally consistent with those of the standard GOE.

<sup>15</sup>This relation holds true irrespective of the actual distribution of the matrix elements. This may be viewed as analogous to the central limit theorem.

dimensionless, unfolded eigenvalues are then defined as

$$\tilde{\mathcal{E}}_k \equiv \eta_{av}(\mathcal{E}_k). \quad (12)$$

With this transformation, distances between eigenvalues are rescaled with respect to the local mean eigenvalue spacings. As a result, the distribution of these unfolded eigenvalues is uniform.

The mean component  $\eta_{av}(\mathcal{E}_k)$  in formula 12 can be approximated using a series of Gaussian functions in a kernel density estimation. Accordingly, the eigenvalue distribution  $P_{\mathbf{E}}(\mathcal{E})$  is expressed as a superposition of  $\delta$ -functions about each eigenvalue

$$P_{\mathbf{E}}(\mathcal{E}) = \frac{1}{M} \sum_{k=1}^M \delta(\mathcal{E} - \mathcal{E}_k), \quad (13)$$

where the  $\delta$ -function about each eigenvalue is approximated by a Gaussian distribution (Sinha et al., 2010, p. 69). This distribution is centered around the respective eigenvalue with standard deviation  $(\mathcal{E}_{k+a} - \mathcal{E}_{k-a})/2$ , where  $2a$  is the broadening window size. Integration of equation 13 approximates  $\eta_{av}(\mathcal{E})$ , which reflects the unfolded eigenvalues  $\tilde{\mathcal{E}}_k$  (see equation 12).

Deriving the unfolded eigenvalues  $\tilde{\mathcal{E}}$  enables studying two universal properties of eigenvalue spacings of GOE-type matrices: i) the distribution of the nearest-neighbor eigenvalue spacings and ii) the distribution of the next-nearest-neighbor eigenvalue spacings (Plerou et al., 2002). These properties are used to assess the randomness of the eigenvalue spectrum.

### Nearest-Neighbor Eigenvalue Spacing Distribution

The distribution of the nearest-neighbor eigenvalue spacings for GOE-type random matrices, i.e., the eigenvalue spacings for successive rank-ordered (unfolded) eigen-

values  $s_1 \equiv \tilde{\mathcal{E}}_{k+1} - \tilde{\mathcal{E}}_k$ , is defined as

$$P_{GOE}(s_1) = \frac{\pi s}{2} \exp\left(-\frac{\pi}{4}s_1^2\right) \quad (14)$$

(Brody et al., 1981; Guhr et al., 1998), also known as Wigner surmise.

This probability density declines with decreasing spacings between eigenvalues. Hence, two eigenvalues are unlikely to be close to each other. This phenomenon is referred to as eigenvalue level repulsion. A consequence of this repulsion is that the eigenvalues of random matrices must be correlated. Uncorrelated eigenvalues would otherwise follow a Poisson distribution  $P_P(s_1) = \exp(-s_1)$

Given this intuition, the agreement of the empirical nearest-neighbor eigenvalue spacing distribution  $P_{nn}(s_1)$  with the assumption of RMT can be further tested. This test can be conducted by fitting  $P_{nn}(s_1)$  to the one-parameter Brody distribution

$$P_{Br}(s_1) = B(1 + \beta)s_1^\beta \exp(-Bs_1^{1+\beta}), \quad (15)$$

where  $B \equiv \left\{ \Gamma\left(\frac{\beta+2}{\beta+1}\right) \right\}^{1+\beta}$ . A parameter value close to  $\beta = 1$  indicates that eigenvalue spacings are distributed according to the theoretical expectation for GOE-type matrices. Conversely, a parameter close to  $\beta = 0$  implies that eigenvalue spacings follow a Poisson distribution and that eigenvalues show no correlation (Brody et al., 1981).

### Next-Nearest-Neighbor Eigenvalue Spacing Distribution

Investigating the distribution of the empirical next-nearest-neighbor eigenvalue spacings  $P_{nnn}(s_2)$  provides an alternative independent test for assessing the randomness of the bulk of the eigenvalue spectrum. Here,  $s_2$  is defined as  $s_2 \equiv (\tilde{\mathcal{E}}_{k+1} - \tilde{\mathcal{E}}_k) + (\tilde{\mathcal{E}}_{k+1} - \tilde{\mathcal{E}}_{k-1}) = \tilde{\mathcal{E}}_{k+1} - \tilde{\mathcal{E}}_{k-1}$ . Thus, the next-nearest-neighbor eigenvalue spacing for the unfolded eigenvalues corresponds to the sum of gaps of an eigenvalue  $\mathcal{E}_k$  to its two adjacent rank-ordered eigenvalues. According to Mehta and Dyson's (1963)

theorem, the distribution of the next-nearest-neighbor spacings for matrices of GOE-type is equivalent to the distribution of the nearest-neighbor spacings of a Gaussian symplectic ensemble (GSE), i.e., symmetric square matrices composed of quaternions (Brody et al., 1981; Guhr et al., 1998). Therefore, the distribution of the next-nearest-neighbor eigenvalue spacings of the unfolded eigenvalue can be described by

$$P_{GSE}(s_1) = \frac{2^{18}}{3^6 \pi^3} s_1^4 \exp\left(-\frac{64}{9\pi} s_1^2\right). \quad (16)$$

### 3.2.4 Eigenvector Distribution

The eigendecomposition in formula 6 provides not only eigenvalues but also eigenvectors of the empirical cross-correlation matrix. While an eigenvalue constitutes a factor that is common to all stocks, eigenvector components specify the exposure of individual stocks to this factor. Eigenvector analysis helps to understand the meaning of a factor and its influence on individual stocks or groups of stocks.

Besides providing theoretical distributions for eigenvalues of random matrices (see section 3.2.2), RMT also provides theoretical distributions for the eigenvector components of random matrices. For a random correlation matrix  $\mathbf{R}$ , the theoretical distribution of the eigenvector components  $\{u_k^l; l = 1, \dots, M\}$  of eigenvector  $\mathbf{u}_k$  corresponds to a Gaussian distribution with zero mean and unit variance (Guhr et al., 1998; Plerou et al., 2002):

$$P_{\mathbf{R}}(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right). \quad (17)$$

Unless a factor is equally common to all stocks, deviations of  $P_{\mathbf{E}}(\mathcal{E})$  from the random matrix eigenvalue distribution  $P_{\mathbf{R}}(\mathcal{E})$  should also be observable in the statistics of the corresponding eigenvector components. Consequently, comparing the empirical eigenvector distribution  $P_{\mathbf{E}}(u)$  with  $P_{\mathbf{R}}(u)$  can provide additional details about the interrelation between stocks.

## 4 Data

The dataset in this study comprises daily price data for 268 constituents of the S&P 500 from June 1980 to June 2018, i.e., for a total of 9,935 days. Time series data were obtained from the Center of Research in Security Prices (CRSP). Information on index affiliation and sector classification was acquired from Compustat.

The S&P 500 is a market capitalization stock market index consisting of 500 stocks listed on US stock exchanges. The subset of 268 stocks results from filtering the S&P 500 for those constituents with full track history of the analyzed period.<sup>16</sup>

Restricting data to this smaller subset is motivated by several factors: First, correlation dynamics between stocks can be reliably estimated only if there is a sufficient period of simultaneous observations between pairs of stocks. Second, an appropriate observation period is required to capture long-term dynamics in stocks' cross-correlations. Third, sufficiently long time series are necessary for consistently estimating the (wavelet) correlation matrix and its subsequent eigendecomposition.

In this context, the choice of time series length for estimating empirical cross-correlation matrices underlies a decisive tradeoff. The longer the time series used for the estimation of correlations, the more information can be obtained about stock interrelations. Thus, incorporating longer time series can help to reduce measurement noise. However, stocks are also affected by constant changes in market conditions. Consequently, correlations between pairs of stocks are not necessarily stationary. Estimating correlations over a longer time period may conceal non-stationarities in the signal and suppress dynamics in stock correlations. As a result, empirical cross-correlation will contain random contributions.

A similar tradeoff exists between covering a broad spectrum of the market using a large representative set of stocks and the accurate representation of interrelations in the system of stocks. Samples that are too small may not sufficiently reflect the market. Conversely, estimating a correlation matrix for a large set of stocks may result in a matrix that is close to singular and contains only insufficient information about the system's dynamics.

---

<sup>16</sup>This is in accordance with previous research, such as Plerou et al. (2002), where a full survival period of 35 years was required for each stock.

The chosen observation period reflects an optimal choice between the *length of the time series* and the *number of stocks with full track history*.

## 5 Empirical Results

This section combines RMT with wavelet decomposition to study the correlation structure in the US stock market on a scale-by-scale basis. It investigates whether this structure changes over different time horizons. Initially, I employ wavelet transformation to construct empirical (wavelet) cross-correlations of return fluctuations at different timescales. Then, I inspect these correlation matrices for changes over different time horizons. Ultimately, I use RMT to study the statistical properties of these correlation matrices. This allows isolating structural relationships and identifying common behavior in the multivariate data. If correlation matrices exhibit timescale-varying structures, it is likely that these variations are also reflected in the common components. In addition, RMT allows filtering noise from the system and revealing relevant information about the correlation matrix (Nguyen, Tran & Nguyen, 2018). This is especially important regarding the high noise content of wavelet correlations for longer time horizons. Consequently, this section provides a deeper understanding of the underlying mechanics in stock markets and their possible timescale-variant behavior.

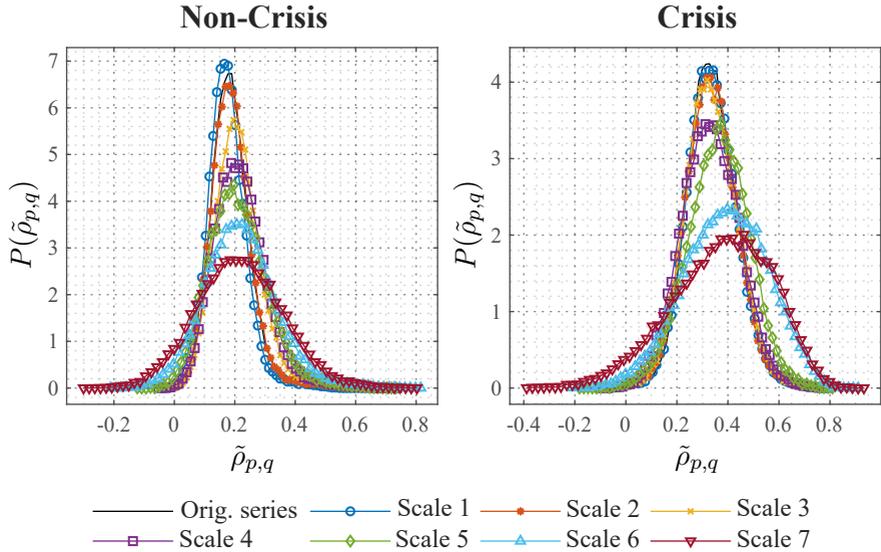
### 5.1 Cross-Correlation Statistics

Before analyzing the eigenvalues and eigenvectors of the empirical (wavelet) correlation matrix  $\mathbf{C}$ , I examine the distribution  $P_{\mathbf{E}}(\tilde{\rho}_{p,q}(\lambda_j))$  of its elements  $\{\tilde{\rho}_{p,q}(\lambda_j); p \neq q\}$  for different timescales  $\{\lambda_j; j = 0 \dots, 7\}$ .<sup>17</sup> However, correlations may not

---

<sup>17</sup>The correlation coefficients of timescale  $\lambda_0$  relate to the original (untransformed) return series.

only vary across timescales but also with regard to the prevalent market state (Fenn et al., 2011; Zheng, Podobnik, Feng & Li, 2012).<sup>18</sup>



**Figure II.1:** Distribution of correlation coefficients  $\{\tilde{\rho}_{p,q}(\lambda_j); p \neq q\}$  at different timescales  $\{\lambda_j; j = 0, \dots, 7\}$  in non-crisis and crisis periods. Notes: Correlation coefficients of timescale  $\lambda_0$  relate to the original (untransformed) return series; correlations were estimated between 268 stocks, covering the period June 30, 1980 to June 30, 2018.

Figure II.1 presents the distribution of the wavelet cross-correlation for crisis and non-crisis periods. While the correlation distribution in the non-crisis period exhibits an excess kurtosis at the lowest timescale (see Table II.1), it gradually approaches the normal distribution as the timescale increases. However, the null hypothesis of normality is still rejected at the 1% level of significance (Jarque-Bera test). The

<sup>18</sup>I subdivide the observation period into non-crisis and crisis periods. The crisis periods encompass the global economic recession of the early 1980s (Jan. 1980–Jun. 1980; Jan. 1981–Dec. 1982), the aftermath following Black Monday 1987 (Oct. 1987–Apr. 1989), the early 1990s recession (Feb. 1989–Mar. 1991), the Asian crisis (Apr. 1997–Dec. 1998), the early 2000s recession (Mar. 2000–Dec. 2002), the global financial crisis of 2007 (Jul. 2007–Jun. 2009), and the European Debt crisis, which was accompanied by the downgrading of America’s credit rating (Jun. 2011–May 2012). All remaining periods are considered to be non-crisis states.

gradual assimilation of the empirical to the normal distribution might indicate that the informational content of correlations decreases over increasing time horizons. The distributions of the correlation coefficients in crisis periods exhibit a similar reduction in kurtosis with increasing timescales. Thus, the correlation distribution displays a higher peak for lower than for higher timescales.

**Table II.1:** *Descriptive statistics of cross-correlation coefficients  $\{\tilde{\rho}_{p,q}(\lambda_j); p \neq q\}$  at timescales  $\{\lambda_j; j = 0, \dots, 7\}$  for non-crisis and crisis periods.*

	Orig. series	Scale 1	Scale 2	Scale 3	Scale 4	Scale 5	Scale 6	Scale 7
<b>Panel A: Non-Crisis</b>								
Mean	0.1879	0.1746	0.1886	0.2097	0.2181	0.2108	0.2187	0.2172
Std. dev.	0.0677	0.0656	0.0685	0.0779	0.0877	0.1001	0.1205	0.1432
Max.	0.7104	0.6876	0.7295	0.7386	0.7400	0.7372	0.8182	0.8074
Min.	-0.0047	-0.0427	-0.0034	-0.0254	-0.0511	-0.1290	-0.2160	-0.3058
Skewness	0.9565	0.8280	0.9250	0.8155	0.7269	0.5690	0.3670	0.1605
Kurtosis	5.6720	5.4756	5.5865	4.9839	4.4321	3.8819	3.6971	3.0794
JB-stat.	16,099***	13,224***	15,075***	9,834***	6,208***	3,090***	1,528***	163***
Observations	35,778	35,778	35,778	35,778	35,778	35,778	35,778	35,778
<b>Panel B: Crisis</b>								
Mean	0.3365	0.3356	0.3326	0.3354	0.3312	0.3607	0.3896	0.3741
Std. dev.	0.0966	0.0995	0.1019	0.1024	0.1135	0.1213	0.1634	0.1966
Max.	0.7826	0.8052	0.7668	0.8011	0.7764	0.8169	0.9145	0.9402
Min.	-0.0571	-0.0716	-0.0867	-0.0374	-0.1003	-0.1897	-0.2172	-0.3916
Skewness	0.1361	0.0837	0.0194	0.1796	0.0725	-0.1591	-0.1766	-0.4017
Kurtosis	3.4985	3.6004	3.4385	3.2263	2.9770	3.1246	2.7427	2.8584
JB-stat.	481***	579***	289***	269***	32***	174***	285***	992***
Observations	35,778	35,778	35,778	35,778	35,778	35,778	35,778	35,778

*Notes:* JB-stat. shows the Jarque-Bera test statistics for the null hypothesis of normality in correlation coefficient distribution; the correlation coefficients of timescale  $\lambda_0$  relate to the original (untransformed) return series.

\*, \*\*, and \*\*\* indicate rejection of the null hypothesis of the test statistics at the 1%, 5%, and 10% level of significance, respectively.

The distributions in both market states also demonstrate a surge in mean correlations with longer time horizons. This observation is consistent with the intuition that macroeconomic variables exert a greater influence on stocks at higher timescales.

Stocks collectively react to market stimuli and become more correlated at longer time horizons.

Despite these similarities, the skewness in the two market phases adjusts differently across timescales. In crisis periods, the correlation distribution shows an increasing left skew over longer time horizons. Contrastingly, in non-crisis periods, the distribution shows no similar change in skewness.

Correlations are also significantly higher in crisis periods.<sup>19</sup> This observation agrees well with the general intuition that stocks exhibit higher interrelation in distressed market states. During these periods, stocks react jointly to news and display strong collective behavior.

These initial results suggest that stocks show varying degrees of interaction over different time horizons (timescales) and across different market states. They call for investigating the timescale properties and underlying structures of stock correlations in more detail.

## 5.2 Theoretical Eigenvalue Distribution of Wavelet Correlation Matrix

RMT has previously been used in financial research to analyze statistical fluctuations of empirical cross-correlations. However, the method has not yet been applied to analyze wavelet filtered time series. This raises the question about whether the theoretical RMT assumptions can be generalized to analyze these time series. This requires first investigating whether RMT laws also apply to wavelet correlation matrices. Hence, what follows studies the properties of wavelet correlation in the context of RMT. Specifically, I define the theoretical assumptions about the distribution of eigenvalues of random wavelet correlation matrices and the corresponding bounds of the distributions.

Wavelet decomposition is an excellent tool for filtering and isolating characteristics of the underlying time series while preserving full signal information. However,

---

<sup>19</sup>These results are consistent with the findings of previous studies (e.g., Aste et al., 2010).

even if wavelet decomposition is applied, a basic problem of timescale analysis remains: The finiteness of the data sample and the limited observation period hinder adequately describing long-term dynamics.<sup>20</sup> Hence, estimations of the eigenvalue distribution are expected to be less representative at higher timescales.

The limited observation period most likely also impairs the theoretical bounds of the eigenvalue distribution of wavelet correlation matrices at each timescale. One might unwarily assume that the length of the original time series is suitable for determining the eigenvalue bounds in formula 8, irrespective of the timescale under consideration. However, due to the limited observation period, it is unlikely that this specification correctly reflects the true eigenvalue distribution at higher timescales. Hence, the measure  $Q$  (see section 3.2.2) should be adjusted to account for the limited information. Instead of the length of the full data sample, I therefore use the number of DWT coefficients  $N_j$  (see section 3.1) to derive both the maximum and the minimum eigenvalue bounds  $\mathcal{E}_+$  and  $\mathcal{E}_-$ .

The dataset in this study consists of  $N = 9,935$  daily return observations and  $M = 268$  constituents of the S&P 500. I assume that the Marčenko-Pastur distribution holds for the distribution of the eigenvalues of (wavelet) correlation matrices. The maximum and minimum theoretical eigenvalue bounds for the untransformed time series are then given by  $\mathcal{E}_+ = 1.356$  and  $\mathcal{E}_- = 0.699$ . To evaluate the proposed eigenvalue distribution of the wavelet correlation matrix, I dyadically reduce the measure  $N$ , which accounts for the number of observations in  $Q$ , with each timescale. This generates the number of corresponding DWT coefficients  $N_j$ . For example, as a result of this reduction, the maximum and minimum eigenvalues at scale level 4 change to  $\mathcal{E}_+ = 2.756$  and  $\mathcal{E}_- = 0.116$ . Thus, the adjusted specification of the theoretical distribution widens the eigenvalue interval with increasing timescale.

Note that there is no theoretical proof justifying the use of the number of DWT coefficients rather than the full time series length for evaluating the theoretical dis-

---

<sup>20</sup>A simple analogy to this problem may be found in an example from meteorology. While the observation of daily temperatures over a period of one year may provide good descriptions of daily, weekly, or even monthly weather cycles, it is unlikely that the same descriptive quality is achieved for yearly cycle periods.

tribution. However, the dyadic increase in the length of the period interval used in wavelet analysis suggests that the same decimation procedure should be applied to calibrating the eigenvalue distribution. Furthermore, RMT relies on the distribution of eigenvalues of independent random matrices. The DWT is an approximately decorrelating transformation. Consequently, the length of the DWT also appears to be a suitable measure for determining the eigenvalue bounds. Lastly, using the number of DWT coefficients neither induces higher testing errors nor inflects fallacious conclusions for identifying deviating eigenvalues. In fact, the theoretical bounds widen with a reduced number of observations. Hence, the likelihood of identifying eigenvalues that deviate from the theoretical distribution decreases. Under this more stringent specification, the testing procedure is thus more restrictive and findings are more robust. The only risk is that some eigenvalues with significant deviation might not be detected.

The previous discussion has defined the theoretical assumptions about the Marčenko-Pastur distribution of random wavelet correlation matrices. We can now test the consistency of the bounds with the eigenvalue distribution obtained from simulated wavelet correlation matrices. Specifically, I generate *simulated random return series* and *randomly reshuffled empirical return series*. Next, I compare the eigenvalue distributions of correlation matrices obtained from these simulations with the previously introduced distributional assumptions.

For the simulated random correlation matrix, I generate  $M = 268$  mutually uncorrelated Gaussian distributed return series of length  $N = 9,935$ . This corresponds to the number of observations in the empirical dataset. The simulated returns are then transformed using MODWT decomposition up to scale level 7. From these decomposed series, I derive a correlation matrix for each timescale. The newly generated random correlation matrices reflect the interrelations of random wavelet time series for timescales ranging from 2–4 days (scale level 1) up to 128–256 days (scale level 7). Next, I derive the eigenvalue spectrum of these random correlation matrices at each timescale. Generation of these random correlation matrices and deriving the eigenvalue spectrum is repeated 1,000 times.

For the correlation matrix of the reshuffled returns, the empirical time series of each

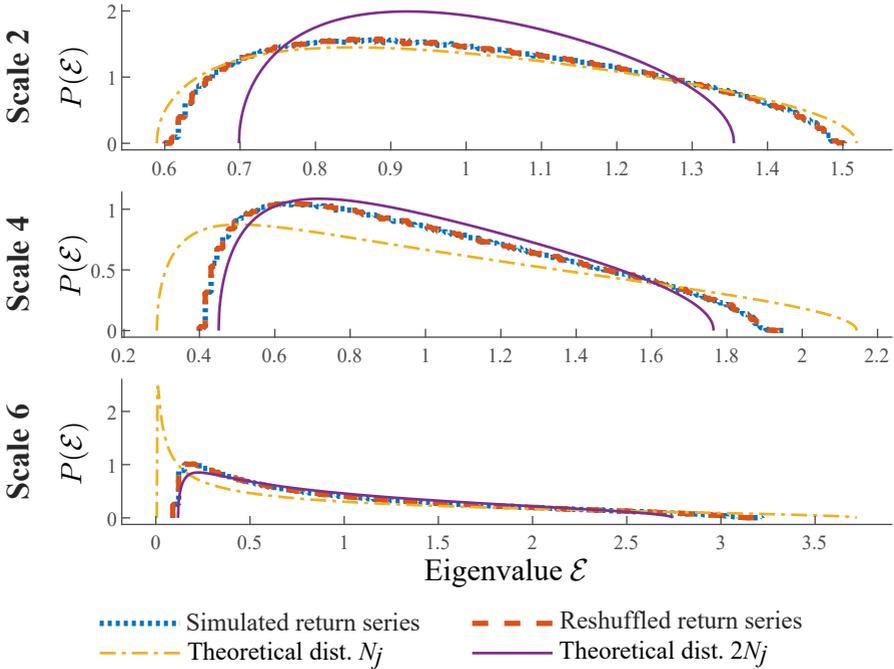
stock is randomly reshuffled, then MODWT wavelet transformed, and the correlation matrix of the decorrelated series evaluated (1,000 independent repetitions).

There are two important reasons for introducing the random reshuffling of time series in addition to the random simulation of time series: First, reshuffling destroys possible correlations within and among stock returns, while retaining the distributional characteristics of the empirical time series (e.g., power-law of the tails). Hence, comparing the eigenvalues of the theoretical distribution with those obtained from the randomly reshuffled time series allows assessing the conformity of the theoretical predictions from random matrices with empirical observations. Second, the matrices of the randomly reshuffled time series help to verify that possible outliers in observations are due neither to measurement noise nor to the finiteness of the data sample. Merely considering simulated returns does not suffice to draw similar conclusions, due to the artificial construction of the time series.

Figure II.2 displays the simulated, reshuffled, and theoretical distributions for different timescales. It shows the theoretical eigenvalue distribution that is reflective of the number of DWT coefficients of the respective timescale  $N_j$ . By comparison, Figure II.2 also provides the theoretical distribution accounting for the number of coefficients of the next lower timescale  $2N_j$ .

For low timescales, the eigenvalue distributions of the simulated and reshuffled return series generally agree well with the surmised theoretical distribution  $N_j$ . However, at high timescales, the generated distributions more closely resemble the theoretical predictions obtained from calibrating the eigenvalue distribution with the number of DWT coefficients of the next lower timescale  $2N_j$  (resulting in narrower eigenvalue intervals). Therefore, the proposed specification of the theoretical eigenvalue distribution (i.e., using the number of DWT coefficients of the corresponding timescale) does not fully account for the effective relation between the eigenvalue spectrum across different timescales. Nevertheless, the simulated and reshuffled distributions are both embedded within the two theoretical distributions. Therefore, the Marčenko-Pastur law generally adequately describes the distribution of random, wavelet-decomposed time series at different timescales. This is the case besides the

proposed distribution specification not being fully consistent with observations of the simulated series.



**Figure II.2:** Simulated, reshuffled, and theoretical (surmised) eigenvalue distributions at different timescales  $\{\lambda_j; j = 2, 4, 6\}$ . Notes: The variable  $N_j$  refers to the number of coefficients in the DWT transformation of scale level  $j$ . The theoretical Marčenko-Pastur distribution was derived for scale level  $j$  ( $N_j$ ) and for the next lower scale level  $j - 1$  ( $2N_j$ ).

Note that the proposed theoretical bounds ( $N_j$ ) consistently exceed the respective eigenvalue intervals of the generated series. These more restrictive bounds might prevent detecting eigenvalues containing information. However, the risk of erroneously classifying eigenvalues of random noise as relevant is also reduced. Therefore, results are more robust with this specification of the theoretical bounds. Using the proposed theoretical distribution  $N_j$  seems appropriate for identifying deviating eigenvalues.

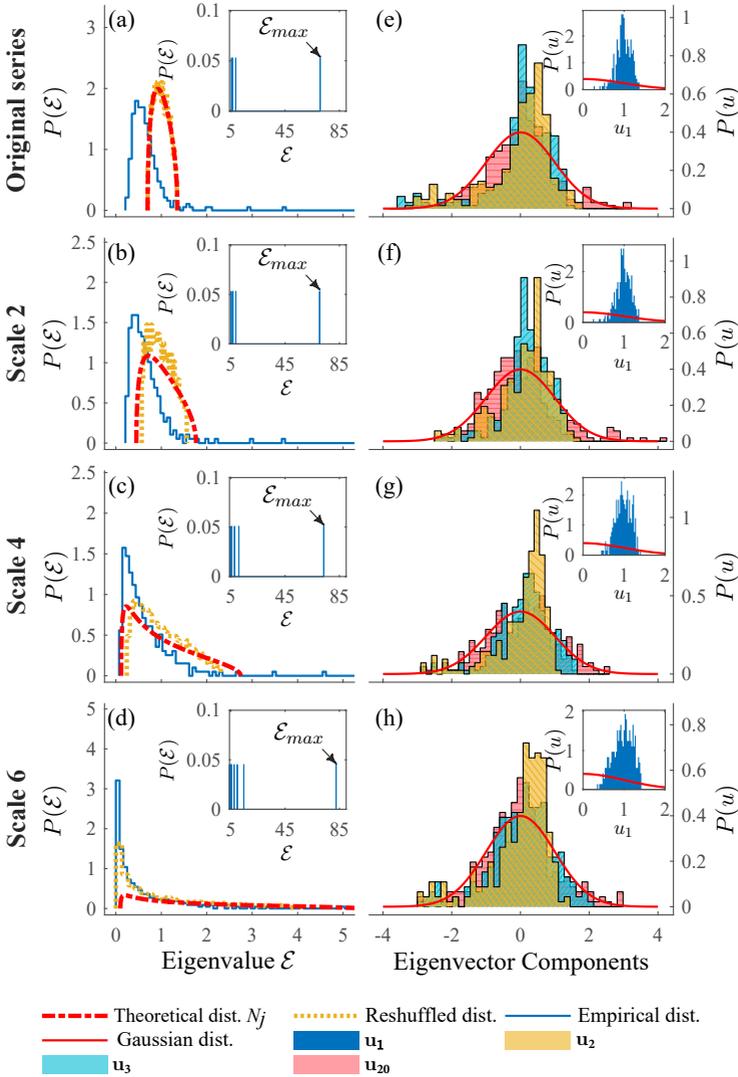
### 5.3 Analysis of the Empirical Eigenvalue and Eigenvector Distribution

The eigenvalue spectrum of the empirical correlation matrices can now be studied by considering the theoretical and the generated eigenvalue bounds obtained in the previous section. Comparing the theoretical predictions with the empirical observations allows distinguishing that part of the correlation matrix that agrees with RMT (random correlations) from the deviations from RMT (genuine information). Figure II.3 illustrates the distributions of the eigenvalues (left column) and the eigenvectors (right column) obtained from the empirical correlation matrices of the untransformed and the wavelet-decomposed daily return observations. For reasons of space, Figure II.3 only displays the eigenvalue spectrum of wavelet correlation matrices for timescales of 4–8 days (scale 2), 16–32 days (scale 4), and 64–128 days (scale 6).<sup>21</sup>

Figure II.3 (a–d) exemplifies a well-defined bulk of eigenvalues that fall within the theoretical bounds  $[\mathcal{E}_-, \mathcal{E}_+]$ . The eigenvalue distribution  $P_{\mathbf{E}}(\mathcal{E})$  of the empirical correlation matrix is therefore in good agreement with the properties dictated by RMT. However, several eigenvalues also appear outside the theoretical bounds. This dispersion of the deviating eigenvalues remains relatively stable across time horizons. However, with increasing timescale, some of the eigenvalues begin to fall below the maximum eigenvalue bound of the theoretical distribution  $\mathcal{E}_+$  and into the bulk of the eigenvalue spectrum. This is most likely caused by the dilation of the theoretical bounds with increasing time horizons (finite data sample) and by the resulting loss in statistical significance. Nevertheless, even if the loss in significance is due to the smaller sample size, the corresponding eigenvalues should no longer be considered non-random.

---

<sup>21</sup>Prior inspection of the eigenvalue distributions revealed that changes in the distributions from one timescale to the next are incremental. For the sake of brevity and without significant loss of information, results are therefore not shown for all timescales.



**Figure II.3:** (a–d) Empirical, reshuffed, and theoretical (surmised) eigenvalue distributions at different timescales ( $\{\lambda_j; 0, 2, 4, 6\}$ ). The insets show the largest eigenvalues. (e–h) Empirical eigenvector distribution of eigenvectors corresponding to the three largest eigenvalues ( $u_1, u_2, u_3$ ), and an eigenvalue from the bulk ( $u_{20}$ ). These distributions are compared to the theoretical Gaussian distribution. The inset displays the eigenvector distribution for the eigenvector corresponding to the largest eigenvalue.

The inset in Figure II.3 provides a microscopic view of the largest eigenvalues by limiting the domain of the eigenvalue axis. This inset enables observing that on the bi-daily horizon (scale level 2) the largest eigenvalue  $\mathcal{E}_1$  of the group of outliers is roughly 40 times larger than the maximum eigenvalue  $\mathcal{E}_+$ . The second and third largest eigenvalues also fall well outside the bounds of the theoretical distribution. For higher timescales, fewer eigenvalues deviate from theoretical bounds. However, the magnitude of the deviating eigenvalues generally increases. Under the assumptions of RMT, all these deviating eigenvalues must be interpreted as carrying information about the correlation structure and as "genuine indicators of correlated movement among the stocks" (Sinha et al., 2010, p. 57).

Eigenvalue deviations from the RMT are most likely also reflected in the statistics of the corresponding eigenvector components (Laloux et al., 1999; Plerou et al., 2001). According to RMT, the eigenvector components of random correlation matrices follow a Gaussian distribution with zero mean and unit variance (see section 3.2.4). Comparing the empirical distribution of eigenvectors with the Gaussian distribution therefore allows isolating structural differences between empirical observations and random correlation matrices.

Note that analysis of eigenvectors is less indicative than that of eigenvalues. Eigenvectors may be normally distributed even if the corresponding eigenvalues exhibit high deviations from the theoretical distribution. Hence, eigenvector coherence with the normal distribution does not necessarily imply random behavior. In addition, it is more difficult to identify significant deviations of the eigenvector distribution from the theoretical distribution. However, the combined analysis of the eigenvalue and the eigenvector distribution enables drawing inferences about the presence of collective modes, about the contribution of single stocks to overall correlation, and about the general structure of the correlation matrix. Together, the two analyses complement each other and can be used to isolate the non-random components, which explain the variance of the correlation matrix.

The right-hand side of Figure II.3 shows the distribution of eigenvectors  $\mathbf{u}_1$ ,  $\mathbf{u}_2$ ,  $\mathbf{u}_3$ , and  $\mathbf{u}_{20}$ , which complement eigenvalues  $\mathcal{E}_1$ ,  $\mathcal{E}_2$ ,  $\mathcal{E}_3$ , and  $\mathcal{E}_{20}$ .<sup>22</sup> The three eigenvalues

---

<sup>22</sup>Note that the eigenvectors are normalized such that  $\sum_{l=1}^M (u_k^l)^2 = M$ .

$\mathcal{E}_1$ ,  $\mathcal{E}_2$ , and  $\mathcal{E}_3$  belong to the group of deviating eigenvalues, which fall outside the theoretical bounds predicted by RMT. In contrast, eigenvalue  $\mathcal{E}_{20}$  is chosen from the bulk of the eigenvalue spectrum  $\{\mathcal{E}_- \leq \mathcal{E}_k \leq \mathcal{E}_+\}$ . Analyzing the eigenvector distribution of  $\mathbf{u}_{20}$  shows good agreement with the Gaussian distribution. Similar results are also observed for the remaining eigenvector distributions, whose corresponding eigenvalues belong to the bulk of the eigenvalue spectrum (not shown). These findings are consistent with the eigenvector distribution predicted by RMT and add to the results of the previous analysis of the eigenvalue components.

In contrast, the distributions of eigenvectors  $\mathbf{u}_1$ ,  $\mathbf{u}_2$ , and  $\mathbf{u}_3$  (whose respective eigenvalues are larger than the maximum eigenvalue  $\mathcal{E}_+$ ) systematically and significantly deviate from the Gaussian distribution. The eigenvector distributions of  $\mathbf{u}_2$  and  $\mathbf{u}_3$  are both heavily skewed and point to a non-random nature of its components. Specifically, the eigenvector distribution  $\mathbf{u}_1$ , which corresponds to the largest eigenvalue ( $\mathcal{E}_1$ ), profoundly differs from the theoretical eigenvector distribution for random matrices. All of the components in  $\mathbf{u}_1$  have positive signs and are located around unity. This suggests that the corresponding eigenvalue represents a common factor affecting all stocks (Laloux et al., 1999; Plerou et al., 2002). The non-random nature of the eigenvectors related to the largest eigenvalues is also observed at all other timescales. Indeed, the shape of most eigenvector distributions for the largest eigenvalue exhibit only modest changes over different time horizons.

Table II.2 quantitatively assesses the number of eigenvalues and eigenvectors conforming and disagreeing with the theoretical distributional assumptions. To identify deviating eigenvalues, I use the theoretical maximum and minimum values obtained from the Marčenko-Pastur distribution and extract the three sets  $\mathcal{E} < \mathcal{E}_-$ ,  $\mathcal{E}_- \leq \mathcal{E} \leq \mathcal{E}_+$ , and  $\mathcal{E} > \mathcal{E}_+$ . In accordance with Wang et al. (2013), I then derive the cardinality of each set such that:

$$\begin{aligned} M_{\mathcal{E} < \mathcal{E}_-} &= \#\{\mathcal{E}; \mathcal{E} < \mathcal{E}_-\}, \\ M_{\mathcal{E}_- \leq \mathcal{E} \leq \mathcal{E}_+} &= \#\{\mathcal{E}; \mathcal{E}_- \leq \mathcal{E} \leq \mathcal{E}_+\}, \\ M_{\mathcal{E} > \mathcal{E}_+} &= \#\{\mathcal{E}; \mathcal{E} > \mathcal{E}_+\}, \end{aligned}$$

where  $\#$  gives the number of observations. The contribution of each set, in relation to all eigenvalues, is then obtained by division with the total number of eigenvalues  $M$ .

Similarly, I classify the eigenvector components into two sets. This involves comparing the empirical eigenvector distribution with the theoretical Gaussian distribution. I apply the Kolmogorov-Smirnov test at a 5% level of significance to test for equality in distribution. Results classification leads to the two sets  $M_{KS < \alpha}$  and  $M_{KS \geq \alpha}$ , which specify the number of significant and insignificant test observations. Similar to the previous test statistics, the relative share of each set compared to all observations is obtained by division with the total number of eigenvalues  $M$ .

Table II.2 illustrates the cardinality sets for the eigenvalue and eigenvector tests. At scale level 1, approximately 48.89% of eigenvalues fall inside the interval predicted by RMT (39.26% for test with reshuffled data). With increasing timescale, this amount increases to almost 97.04% at scale level 4 (68.15% for test with reshuffled data). Tests show that this trend continues. The share of eigenvalues belonging to the bulk of the spectrum increases with timescale.

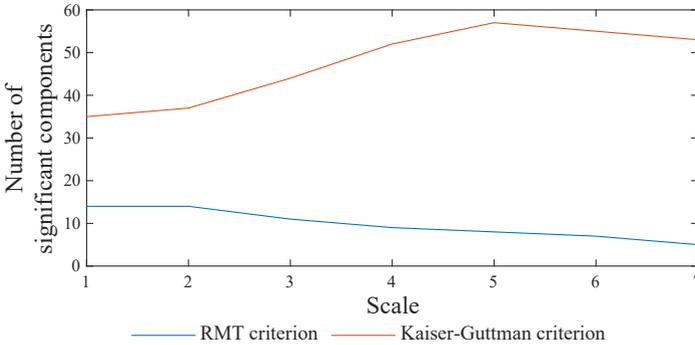
In contrast, only 4.44% of eigenvalues exceed the maximum eigenvalue bound at scale level 1 and 3.33% at scale level 4. Nevertheless, this amounts to 12 (6) eigenvalues at scale level 1 (4) with non-random characteristics. All these eigenvalues may be considered to contain relevant, non-random information. This applies in particular if deviations of the eigenvalues from the maximum eigenvalue bound are large.

**Table II.2:** Number and percentage of eigenvalues and eigenvectors disagreeing with the theoretical assumptions from RMT.

	Eigenvalue						Eigenvector											
	$\mathcal{E} < \mathcal{E}_-$		$\mathcal{E}_- \leq \mathcal{E} \leq \mathcal{E}_+$		$\mathcal{E} \geq \mathcal{E}_+$		$M_{K S \geq \alpha}; H_0 : \mathbf{u}_k \sim N(0, 1)$				$M_{K S < \alpha}; H_0 : \mathbf{u}_k \not\sim N(\mathbf{0}, \mathbf{1})$							
	$M_{\mathcal{E} < \mathcal{E}_-}$	$P_{\mathcal{E} < \mathcal{E}_-}$	$M_{\mathcal{E}_- \leq \mathcal{E} \leq \mathcal{E}_+}$	$P_{\mathcal{E}_- \leq \mathcal{E} \leq \mathcal{E}_+}$	$M_{\mathcal{E} \geq \mathcal{E}_+}$	$P_{\mathcal{E} \geq \mathcal{E}_+}$	$M_F$	$P_F$	$M_L$	$P_L$	$M_U$	$P_U$	$M_F$	$P_F$	$M_L$	$P_L$	$M_U$	$P_U$
Scale 1	151 (126)	55.93% (46.67%)	106 (132)	39.26% (48.89%)	13 (12)	4.81% (4.44%)	221	81.85%	257	95.19%	234	86.67%	49	18.15%	13	4.81%	36	13.33%
Scale 2	137 (81)	50.74% (30.00%)	120 (180)	44.44% (66.67%)	13 (9)	4.81% (3.33%)	235	87.04%	260	96.30%	245	90.74%	35	12.96%	10	3.70%	25	9.26%
Scale 3	104 (38)	38.52% (14.07%)	155 (225)	57.41% (83.33%)	11 (7)	4.07% (2.59%)	246	91.11%	261	96.67%	255	94.44%	24	8.89%	9	3.33%	15	5.56%
Scale 4	77 (2)	28.52% (97.04%)	184 (262)	68.15% (97.04%)	9 (6)	3.33% (2.22%)	260	96.30%	265	98.15%	265	98.15%	10	3.70%	5	1.85%	5	1.85%
Scale 5	48 (0)	17.78% (0.00%)	214 (265)	79.26% (98.15%)	8 (5)	2.96% (1.85%)	266	98.52%	267	98.89%	269	99.63%	4	1.48%	3	1.11%	1	0.37%
Scale 6	31 (-)	11.48% (-)	232 (-)	85.93% (-)	7 (-)	2.59% (-)	266	98.52%	268	99.26%	268	99.26%	4	1.48%	2	0.74%	2	0.74%
Scale 7	25 (-)	9.26% (-)	240 (-)	88.89% (-)	5 (-)	1.85% (-)	267	98.89%	268	99.26%	269	99.63%	3	1.11%	2	0.74%	1	0.37%

*Notes:* The group "Eigenvalue" lists the number ( $M_{\mathcal{E} < \mathcal{E}_-}$ ,  $M_{\mathcal{E}_- \leq \mathcal{E} \leq \mathcal{E}_+}$ ,  $M_{\mathcal{E} > \mathcal{E}_+}$ ) and percentage ( $P_{\mathcal{E} < \mathcal{E}_-}$ ,  $P_{\mathcal{E}_- \leq \mathcal{E} \leq \mathcal{E}_+}$ ,  $P_{\mathcal{E} > \mathcal{E}_+}$ ) of eigenvalues that lie below ( $\{\mathcal{E} < \mathcal{E}_-\}$ ), within ( $\{\mathcal{E}_- \leq \mathcal{E} \leq \mathcal{E}_+\}$ ), or above ( $\{\mathcal{E} > \mathcal{E}_+\}$ ) the theoretical Marčenko-Pastur eigenvalue bounds. The numbers without parentheses illustrate results for tests with reshuffled data for scale level 1–5; the numbers in parentheses illustrate results for tests with theoretical assumptions for scale level 1–5; the group "Eigenvector" shows the number ( $M_F$ ) and percentage ( $P_F$ ) of insignificant deviations from the null hypothesis  $M_{K S \geq \alpha}(H_0 : \mathbf{u}_k \sim N(0, 1))$  and the number ( $M_F$ ) and percentage ( $P_F$ ) of significant deviations from the null hypothesis  $M_{K S < \alpha}(H_0 : \mathbf{u}_k \not\sim N(0, 1))$  using a Kolmogorov-Smirnov test. Test results are split for lower (number:  $M_L$ ; percentage:  $P_L$ ) and upper (number:  $M_U$ ; percentage:  $P_U$ ) eigenvector groups (i.e., eigenvectors corresponding to eigenvalues in the lower and upper half of the rank-ordered eigenvalue spectrum).

Eigenvector analysis yields similar results. Almost 18.15% of eigenvectors exhibit significant deviations from the Gaussian distribution at scale level 1. The set of eigenvectors was split further to determine whether these differences occur for those eigenvectors that correspond either to the largest or to the smallest eigenvalue. For this split, eigenvectors were ordered by the size of the corresponding eigenvalues. Eigenvectors corresponding to the upper half of the ordered eigenvalue list are subsumed in  $M_U$ . Eigenvectors of the lower half of the ordered eigenvalue list are summarized in  $M_L$ . Results show that almost 13.33% of those eigenvectors that correspond to the largest eigenvalues ( $M_U$ ) significantly deviate from the Gaussian distribution at the 5% level of significance. However, for scale level 4, this value reduces to 1.85%. Note that the corresponding eigenvalues are still outside the theoretical bounds predicted by RMT for some observations with random eigenvector characteristics. However, as mentioned, eigenvectors may display random characteristics, whereas eigenvalues imply non-random behavior.



**Figure II.4:** Number of significant components (eigenvalues) as a function of the timescale identified using the RMT and the Kaiser-Guttman criterion.

The number of deviating eigenvalues that can be isolated with this method is relatively restrictive compared to classical criteria such as the scree test or the Kaiser-Guttman criterion. Whereas the scree test is rather subjective, the Kaiser-Guttman criterion is quantifiable. The Kaiser-Guttman criterion allows dropping all eigenvalues falling below unity. Any component above this value is considered to con-

dense more information than the original variable. Figure II.4 shows the number of eigenvalue components deemed relevant by this classical criterion against the isolated eigenvalues from RMT. The Kaiser-Guttman criterion is far less restrictive than the RMT criterion and retains far more eigenvalues. Interestingly, the differences between these methods even increase with timescale.

To summarize, combining eigenvalue and eigenvector analyses indicates that eigenvalues mostly belong to the bulk of the eigenvalue spectrum. This agreement of the eigenvalue statistics of the empirical cross-correlation matrix  $\mathbf{C}$  with RMT implies that considerable randomness exists in  $\mathbf{C}$  (Plerou et al., 2000). Only few eigenvalues and eigenvectors exhibit significant deviations. However, those deviations are relatively large. For example, the largest eigenvalue exceeds the predicted maximum eigenvalue of the theoretical distribution between 52 to 15 times from timescale  $\lambda_0$  to  $\lambda_6$ . Considering the change in eigenvalues across different timescales, eigenvalues outside the bulk of the eigenvalue spectrum decrease in size. Nevertheless, the largest eigenvalues still remain far outside the theoretical maximum Marčenko-Pastur bound. Their sizes even tend to increase for longer timescales. In addition, the reduction in the number of deviating eigenvalues most likely results from the widening of the theoretical distribution.

These results indicate that the most relevant information is contained only in the largest few eigenvalues. This pattern prevails across all time horizons and even intensifies with increasing timescale. Accordingly, correlation matrices across different timescales are characterized by underlying structures and common behavior — both are incompatible with the assumption of random correlation matrices. These findings point to an underlying factor structure that defines the structure of correlations. Moreover, the influence of these factors can change across timescales. In general, RMT is thereby far more conservative in the number of eigenvalues (factors) to be retained compared to classical criteria.

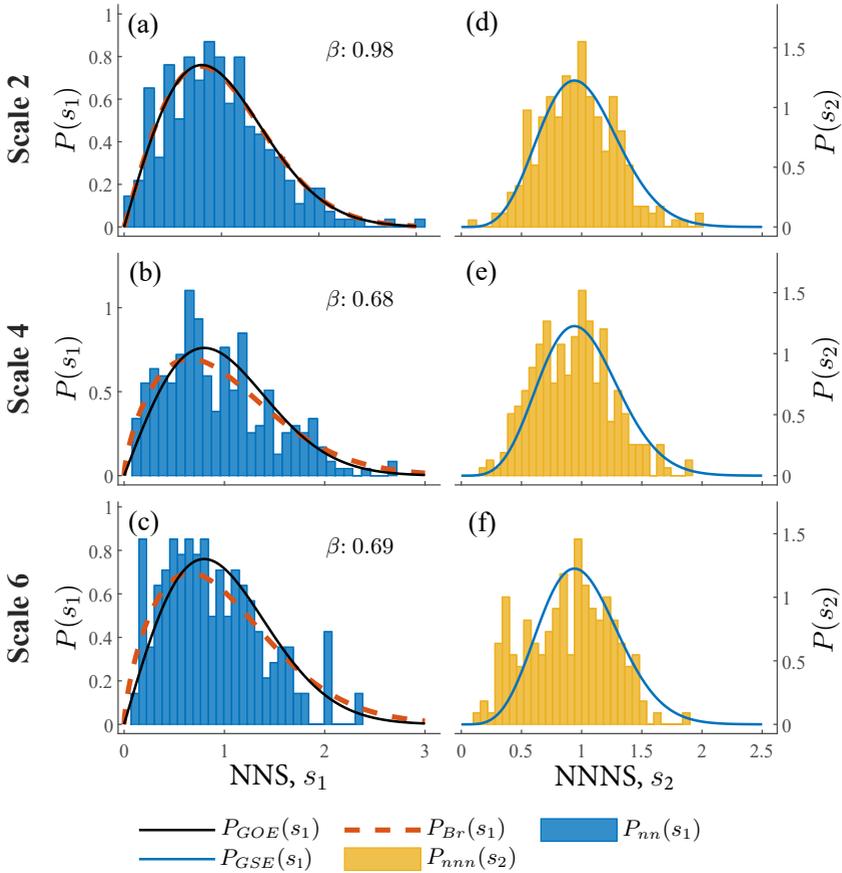
## 5.4 Consistency of the Bulk of the Eigenvalue Spectrum with RMT

As elucidated in section 3.2.3, analyzing eigenvalue and eigenvector distributions is not sufficient to evaluate the randomness of the bulk of the eigenvalue spectrum. It is thus premature to draw conclusions about the suitability of RMT for evaluating the statistical properties of the correlation matrix at different timescales. Additional tests of the randomness of the bulk of the eigenvalue spectrum across different time horizons are needed to interpret the eigenvalues and eigenvectors. These tests involve investigating eigenvalue correlations by analyzing the nearest-neighbor and the next-nearest-neighbor eigenvalue spacing distributions. These eigenvalue spacing distributions (qua representations of the eigenvalue correlations) are compared to the universal features of eigenvalue correlations displayed by real symmetric random matrices. This comparison reveals the consistency of the empirical eigenvalue correlations with RMT (Plerou et al., 2002).

Figure II.5 presents the nearest-neighbor (NNS, left column) and the next-nearest-neighbor (NNNS, right column) eigenvalue spacing distributions of the empirical cross-correlation matrix for different timescales. The empirical distributions are supplemented by the corresponding theoretical distributions. These include the distribution of the nearest-neighbor eigenvalue spacings of a GOE-type matrix  $P_{GOE}(s_1)$  for  $s_1 \equiv \xi_{k+1} - \xi_k$  (Wigner surmise) and the distribution of the nearest-neighbor eigenvalue spacings of a GSE-type matrix  $P_{GSE}(s_1)$  (for the next-nearest-neighbor eigenvalue spacings  $s_2 \equiv \xi_{k+2} - \xi_k$  of a GOE-type matrix). Figure II.5 also displays the fitted one-parameter Brody distribution for the nearest-neighbor eigenvalue spacings and the corresponding parameter estimates.

Figure II.5 (a, b) shows that for lower timescales, the distribution of the next-nearest eigenvalue spacings accords with the theoretical distribution  $P_{GOE}(s_1)$  predicted by RMT. However, on an increasing time horizon, the empirical distribution starts to deviate from the theoretical distribution (Figure II.5 c). This pattern is also confirmed by the parameter estimation from the Brody distribution. The  $\beta = 0.98$  of the first timescale indicates a close correspondence of the empirical distribution with the GOE eigenvalue spacing distributions. In contrast, the parameter values  $\beta = 0.68$

and  $\beta = 0.68$  observed for scale levels 4 and 6 present mixed results. The parameter estimates lie between a value signaling a GOE-type distribution ( $\beta = 1$ ) and a value indicating a Poisson distribution ( $\beta = 0$ ) of the eigenvalue spacings.



**Figure II.5:** (a, b, c) Nearest-neighbor eigenvalue-spacing distribution and (d, e, f) next-nearest-neighbor eigenvalue-spacing distribution for different timescales  $\{\lambda_j; j = 2, 4, 6\}$ . Notes: The variable  $\beta$  refers to the parameter obtained from fitting the one-parameter Brody distribution  $P_{Br}(s_1)$  to the empirical eigenvalue-spacing distribution  $P_{nn}(s_1)$ , where  $s_1$  is the next-nearest eigenvalue-spacing (NNS); the theoretical distribution of  $s_1$  is shown by  $P_{GOE}(s_1)$ ; the empirical distribution  $P_{nnn}(s_2)$  of the next-nearest neighbor eigenvalue-spacing  $s_2$  (NNNS) is compared to the theoretical distribution  $P_{GSE}(s_1)$  of  $s_1$ .

Analyzing the empirical distribution of next-nearest eigenvalue spacings  $P_{nnn}(s_2)$  reveals a similar picture (Figure II.5 d, e, f). At lower timescales, the empirical distribution agrees well with the theoretical distribution from RMT. However, the distribution begins to gradually differ on an increasing time horizon.

These results show that, at lower timescales, the eigenvalue-spacing distribution of the empirically measured cross-correlation matrix is in general consistent with RMT predictions. However, at higher timescales, the results are less conclusive and should be viewed more cautiously. The difference in the eigenvalue-spacing distributions at high timescales may be partially attributable to the presence of non-random characteristics in the bulk of the eigenvalue spectrum. This would be in line with the observation that the critical eigenvalue bounds expand on an increasing time horizon. Thus, non-random components may be incorrectly considered to be part of the bulk of the eigenvalue spectrum. Nevertheless, these findings indicate, that genuine information is predominantly contained in those eigenvalues that deviate from the Marčenko-Pastur bounds. Therefore, RMT predictions and the RMT criterion for specifying relevant eigenvalues are appropriate for analyzing correlation structures across different timescales.

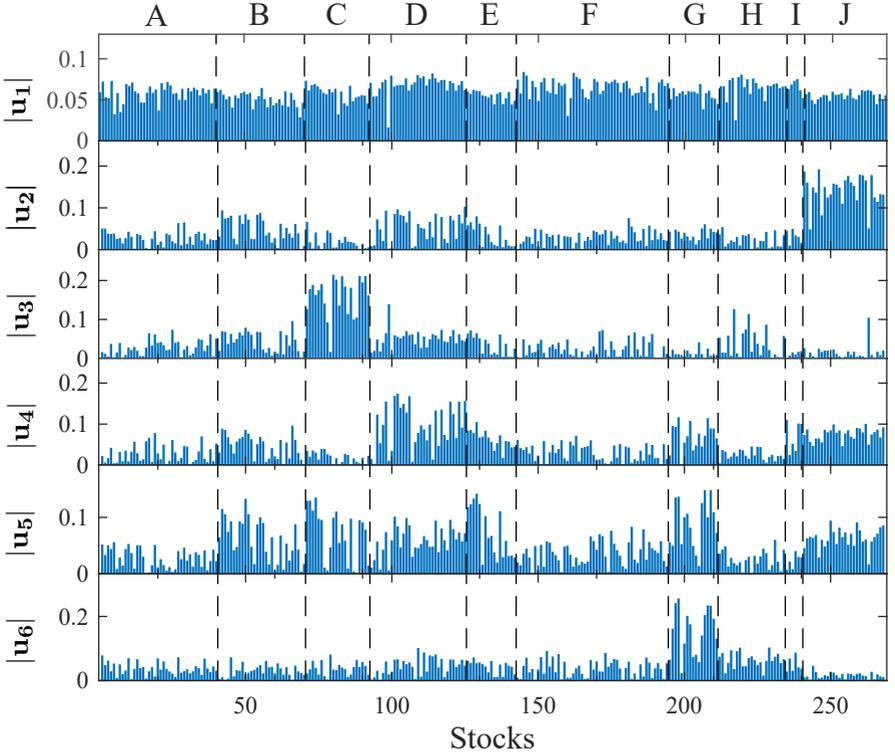
## 5.5 Interpretation of Correlation Statistics Across Timescales

Having established RMT's validity and suitability for evaluating (wavelet) cross-correlation statistics, I now interpret the eigenvalues and eigenvectors. Does the interpretation of eigenvalues remain consistent across different time horizons? Or does the underlying structure instead change for different timescales? Eigendecomposition makes no assumptions about its components and their relation to certain economic factors. Consequently, we examine the eigenvalues and eigenvectors to gain an economic understanding of their meaning, and ultimately to define the driving forces of cross-correlations between stocks across different time horizons. It is important to note, however, that the observed eigenvalues and eigenvectors do not

necessarily lend themselves to economic interpretation. This makes it even more imperative to examine the components in detail.

First, I study the single eigenvectors corresponding to the deviating eigenvalues in order to gain insight into the structure of correlation statistics and their possible interpretation. Figure II.6 displays the absolute eigenvectors associated with the largest six eigenvalues obtained for the correlation matrix of the original return series. The eigenvectors corresponding to the largest eigenvalue  $\mathcal{E}_1$  exhibit a relatively homogeneous distribution. Hence, all stocks show a more or less uniform exposure to the largest eigenvalue. This is indicative of an influence that is common to all stocks and that affects the entire stock market. It describes the part of the variation in the correlation matrix that is due to the system responding collectively to external information (news). Therefore, the largest eigenvalue is often associated with a general market factor (Gopikrishnan et al., 2001; Laloux et al., 1999; Pan & Sinha, 2007; Sinha et al., 2010).

In contrast, the eigenvectors of the second largest eigenvalue and all the following eigenvalues are highly localized. Only few eigenvector components (i.e., stocks) significantly contribute to each of these modes. The stocks shown in Figure II.6 are ordered by their GICS sector classification. This grouping reveals that stocks contributing significantly to a certain eigenvector mostly belong to similar or related sectors. For example, for the second largest eigenvalue, the corresponding eigenvector indicates significant contributions to stocks that are part of the Utilities sector (J). In contrast, stocks in the Energy (C) and Consumer Staples (B) sector display the highest loading for eigenvectors corresponding to the third and fourth largest eigenvalues. Similar findings are also obtained for the remaining eigenvectors. These results are consistent with the findings in previous studies (see Gopikrishnan et al., 2000, 2001; Kim & Jeong, 2005; Liu et al., 1999).

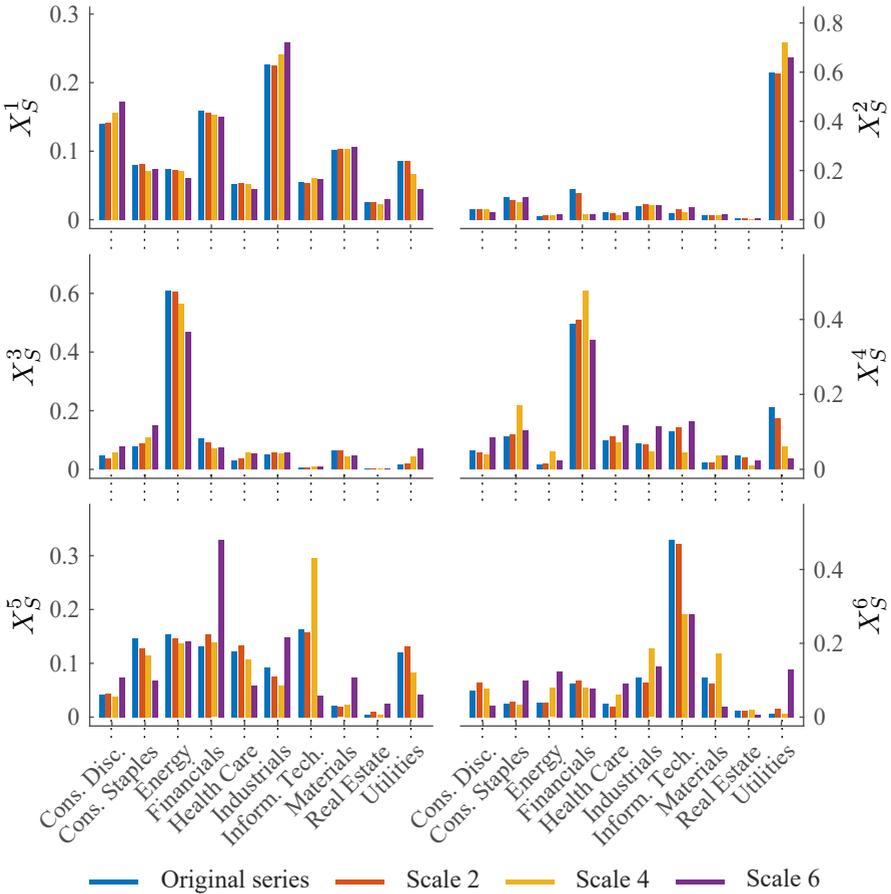


**Figure II.6:** Sector classification of original (untransformed) return data using absolute eigenvector components  $\{|u_k^l|; k = 1, \dots, 6\}$  of stock  $l$  corresponding to the six largest eigenvalues  $\{\mathcal{E}_k; k = 1, \dots, 6\}$ . Notes: Stocks are grouped by sectors (dashed lines) where A: Consumer Discretionary, B: Consumer Staples, C: Energy, D: Financials, E: Health Care, F: Industrials, G: Information Technology, H: Materials, I: Real Estate, and J: Utilities.

Next, I investigate whether this localization property remains consistent across time-scales. Following Gopikrishnan et al. (2001), I construct a map  $\mathcal{M}_{S,l}$  for each stock  $l$  indicating its affiliation to a sector  $S$  such that  $\mathcal{M}_{S,l} = \begin{cases} 1 & \text{if } l \in S \\ 0 & \text{otherwise} \end{cases}$ . An aggregate measure for the industry affiliation of a given eigenvector is then given by  $X_S^k = \sum_{l=1}^M \mathcal{M}_{S,l} [u_k^l]^2$ . This mapping is derived for each timescale.

Figure II.7 illustrates the aggregate measure  $X_S^k(\lambda_j)$  for the six largest eigenvalues

and for selected timescales  $\{\lambda_j; j = 0, 2, 4, 6\}$ . Similar to Figure II.6, the measure  $X_S^1$  (which corresponds to the largest eigenvalue) exhibits comparable loading across all sectors. This indicates a common behavior across all sectors. Thus, the largest eigenvalue is indeed related to a market factor. This association holds across all timescales.



**Figure II.7:** Contribution  $X_S^k$  of the eigenvector  $u_k$  to sector  $S$  over different timescales.

Similarly, the second and third largest eigenvalues are well localized. They can be associated with the Utilities sector and the Energy sector respectively. For the next larger eigenvalues, results are less indicative. Interpretation is less straightforward as it is more difficult to relate the eigenvalues to one specific sector. Nevertheless, the eigenvalues mostly associate with a group consisting of one or two sectors.

Figure II.7 further reveals that the sector association of a given eigenvector persists across different timescales. The second largest eigenvalue is still associated with the Utilities sector irrespective of the timescale under consideration. This consistency of eigenvalue interpretation indicates that the general structure of correlation matrices remains relatively stable across timescales. This is an important observation. It shows that the correlation is driven by the same fundamental factors and that the influence of these factors remains stable across different timescales. The eigenvalue order seems to vary only between timescales for observations of the fifth and sixth largest eigenvalues. For example, in contrast to the observation at lower timescales, the fifth largest eigenvalue is no longer associated with the Information Technology sector and the Consumer Staples sector at scale level 6.

Note that a more refined subgrouping might unearth even more detailed structural relations between stocks. However, due to the relatively small sample of only 268 stocks, results for more granular subgroupings would most likely be hard to interpret. For example, associating eigenvector components with a certain industry is difficult to justify if there are only three members in that particular group. Other approaches, which do not use predefined exogeneous group classifications (e.g., graph theory or clustering), also provide more data-driven methods for isolating subgroups and for identifying clusters. On the other hand, these methods may lack reasonable economic interpretation.

To confirm the assumption that the largest eigenvalue and the corresponding eigenvectors are representative of a market factor across different timescales, an additional test was conducted. I derive the projection  $PF_{k,t}(\lambda_j)$  of the wavelet transformed returns  $\tilde{W}_{l,j,t}$  of stock  $l$  on the eigenvector  $\mathbf{u}_k$  given by

$$PF_{k,t}(\lambda_j) \equiv \sum_{l=1}^M u_k^l \tilde{W}_{l,j,t}. \quad (18)$$

Hence, this projection  $PF_{k,t}(\lambda_j)$  can be interpreted as the return of a portfolio with weighting  $\mathbf{u}_k$ . This (normalized) portfolio is now regressed onto the normalized wavelet transformed returns of the S&P 500 index  $\check{W}_{j,t}^{\mathcal{G}}$  (qua representation of the US stock market):

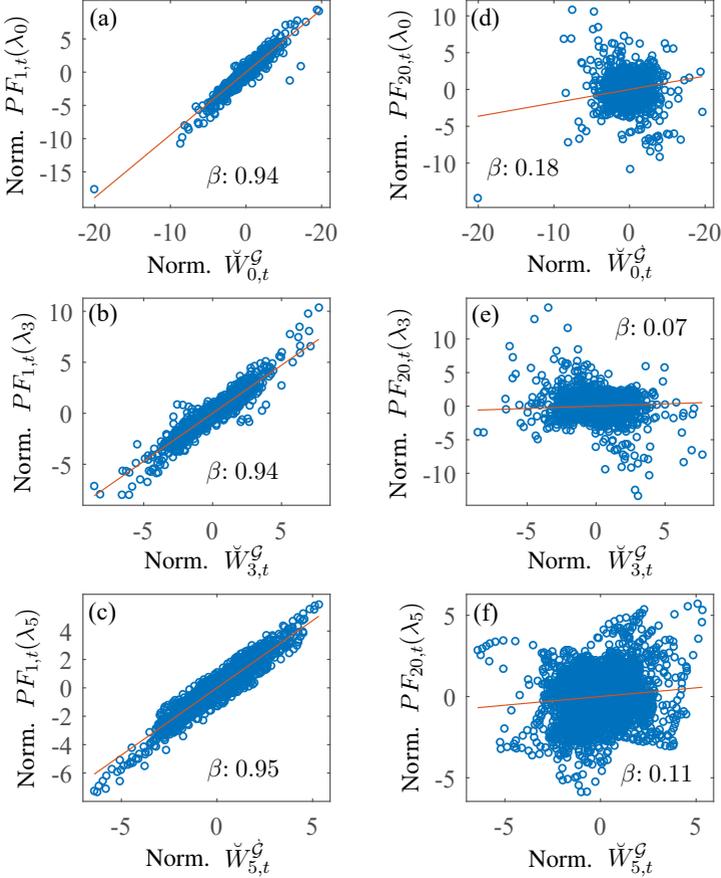
$$PF_{k,t}(\lambda_j) = \alpha_{k,j} + \beta_{k,j} \check{W}_{j,t}^{\mathcal{G}} + \varepsilon_{k,t}(\lambda_j), \quad (19)$$

where  $\varepsilon_{k,t}(\cdot)$  is an independently and identically distributed idiosyncratic noise. The value of the slope  $\beta$  is indicative of the relation between the projection and the market index. A value close to unity (with little scattering) signifies a high correspondence between the two time series. In contrast, a value of zero indicates the absence of a linear relationship.

This regression is applied to portfolios  $PF_{1,t}(\lambda_j)$  and  $PF_{20,t}(\lambda_j)$ , which are the projections of the time series onto eigenvectors  $\mathbf{u}_1$  and  $\mathbf{u}_{20}$ , respectively. Hence, the former relates to the largest eigenvalue whereas the latter relates to the bulk of the eigenvalue spectrum. Figure II.8 (a-c) displays the regressions of portfolio  $PF_{1,t}(\lambda_j)$  on the market index  $\check{W}_{j,t}^{\mathcal{G}}$  for different timescales  $\{\lambda_j; j = 2, 4, 6\}$ . For the regression in Figure II.8 (d-f), the regressand is replaced by portfolio  $PF_{20,t}(\lambda_j)$ .

The results for portfolio  $PF_{1,t}(\lambda_j)$  show close correspondence between the market index  $\check{W}_{j,t}^{\mathcal{G}}$  and the projection at all timescales. Slope  $\beta$  is close to unity and the datapoints scatter narrowly around the linear fit. Previously, Plerou et al. (2002) reported a slope of 0.85 for 30-min returns of the largest 1,000 US stocks for the period 1994–1995 and claimed similar results for daily returns series. The slope obtained here for the untransformed series (daily returns) is  $\beta = 0.94$  and thus even higher than the value reported by Plerou et al. (2002). Similar values for the slope are found across all timescales. Therefore, I conclude that the largest eigenvalue and the corresponding eigenvectors are indeed representative of the market mode. This interpretation is consistent for all timescales. In contrast, the slopes for the portfolio

$PF_{20,t}(\lambda_j)$  are close to zero for all time horizons. Hence, the eigenvalues and the corresponding eigenvectors are independent of the market index  $\check{W}_{j,t}^G$ .<sup>23</sup> This finding accords with predictions for the bulk of the eigenvalue spectrum in RMT.



**Figure II.8:** (a, b, c) Eigenvector portfolio  $PF_{1,t}(\lambda_j)$ , corresponding to largest eigenvalue, regressed on the normalized S&P 500 (wavelet) return series  $\check{W}_{j,t}^G$ , for different timescales  $\{\lambda_j; 0, 3, 5\}$ . (d, e, f) Eigenvector portfolio  $PF_{20,t}(\lambda_j)$ , corresponding to 20<sup>th</sup> largest eigenvalue, regressed on the S&P 500 (wavelet) return series  $\check{W}_{j,t}^G$ , for different timescales  $\{\lambda_j; 0, 3, 5\}$ . Notes: The variable  $\beta$  refers to the slope of the linear regression.

<sup>23</sup>Note that the portfolio return series indicates a high degree of autocorrelation. However, this does not affect the coefficient estimator, which is still unbiased.

The results also demonstrate that the sample of stocks chosen for this study sufficiently represents general market structures. Otherwise, no similarly high value would be achieved for the slope of the regression in Figure II.8 (a-c). Consequently, inflection due to the presence of a survivorship bias may be at least partially rejected.

### Inverse Participation Ratio

Following Plerou et al. (1999, 2002), we next study the number of components participating in a given eigenvector. This analysis indicates how far the influence of a specific eigenvalue spreads across the stock universe. Further, it offers an additional method for testing the deviations of eigenvector components from RMT assumptions. To obtain a quantitative measure of the number of eigenvector components that contribute to a given eigenmode, the notion of the inverse participation ratio (IPR) is introduced (Fyodorov & Mirlin, 1992, 1993). The IPR for the  $k$ -th eigenvector  $\mathbf{u}_k$  is defined as

$$I_k \equiv \sum_{l=1}^M [u_k^l]^4, \quad (20)$$

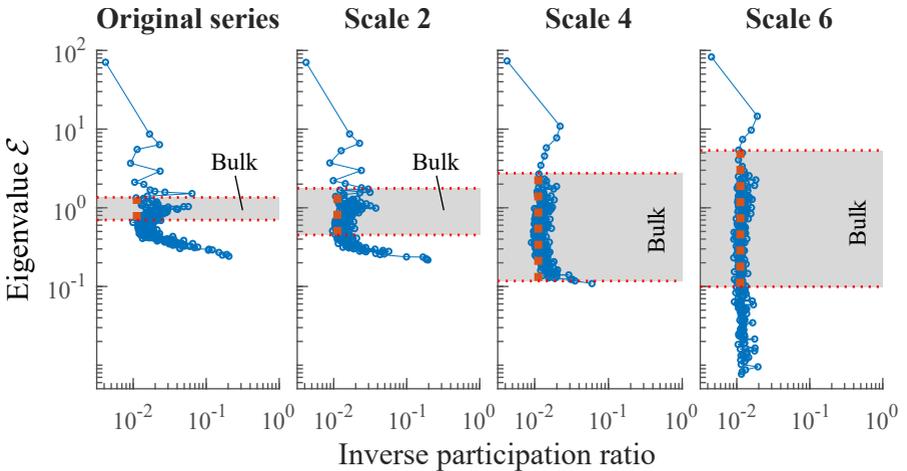
where  $u_k^l$  is the  $l$ -th component of eigenvector  $\mathbf{u}_k$ . The meaning of IPR is best exemplified by considering two limiting cases. For an eigenvector that consists of components with equal contributions (i.e.  $u_k^l = \frac{1}{\sqrt{M}} \forall l$ ), the IPR is  $I_k = 1/M$ . Conversely, for an eigenvector with a single dominant stock (e.g.  $u_k^1 = 1$  and  $u_k^l = 0 \forall l \neq 1$ ), the IPR is  $I_k = 1$ . In RMT, the expected value for IPR of a random correlation matrix is given by

$$\langle I_k \rangle = M \int_{-\infty}^{\infty} [u_k^l]^4 \frac{1}{\sqrt{2\pi M}} \exp\left(-\frac{[u_k^l]^2}{2M}\right) du_k^l = \frac{3}{M} \quad (21)$$

(Utsugi et al., 2004).

Figure II.9 presents the IPR for different timescales  $\{\lambda_j; j = 0, 2, 4, 6\}$ . For the eigenvalues lying outside the bulk of the eigenvalue spectrum, the IPR deviates from

the theoretical expectation (formula 21). In particular, the IPR of the largest eigenvalue is below the expected value. To fathom this observation, we move to the more intuitive interpretation of the IPR given by its reciprocal. This reciprocal quantifies the number of eigenvector components making a significant contribution (Plerou et al., 2002). At timescale  $\lambda_0$ , the reciprocal of the IPR for the largest eigenvalue is  $1/I_1(\lambda_0) \approx 240$ . With increasing timescale, the IPR shows a decreasing trend and reduces to  $1/I_1(\lambda_6) \approx 223$  at timescale  $\lambda_6$ . These results show that a significant number of stocks contribute to the largest eigenvector at all timescales. Again, this is consistent with the interpretation of the largest eigenvalue representing the market mode. The intermediate eigenvalues (i.e., the eigenvalues lying between the largest eigenvalue and the bulk of the eigenvalue spectrum) display a more localized behavior. Hence, fewer stocks contribute to these eigenvectors. Again, this is consistent with the association of these eigenvectors with group or sector characteristics. However, this localization seems to weaken slightly with increasing timescale.



**Figure II.9:** Log-log plot of inverse participation ratio (IPR) as a function of the eigenvalue for original (untransformed) stock returns and wavelet transformed returns at different scale levels. Notes: The grey-shaded region marks the bulk of the eigenvalue spectrum; the thick red-dotted line indicates the expected value of the IPR of a random matrix.

In contrast, the eigenvalues belonging to the bulk of the eigenvalue spectrum closely correspond to the expected value in formula 21. With increasing timescale, a gradual convergence towards the expected value  $\langle I_k \rangle$  may be observed. Hence, at higher timescales, the bulk of the eigenvalue spectrum agrees better with the expectations of RMT. To some extent, this contradicts the findings for the distribution of the eigenvalue spacings in section 5.4, for which a lower conformity with RMT was observed at higher timescales.

The analysis in this section plainly shows that eigenvalues and eigenvectors have an economic interpretation. While the eigenvector that corresponds to the largest eigenvalue can be described as a market component, those eigenvectors that correspond to the intermediate eigenvalues relate to different market groups (sectors). The remaining eigenvalues are associated with random deflections and idiosyncratic risks in the correlation matrix. These idiosyncratic risks are company-specific characteristics of a stock's price dynamics, which are not necessarily related to the price dynamics of other stocks. This interpretation remains mostly consistent across the different timescales and holds irrespective of the time horizon. Thus, over different time horizons, the fundamental structure of the correlation matrix is driven by the same underlying factors.

These results also show that eigenvalues and eigenvectors can be closely linked to market models such as the Capital Asset Pricing Model (CAPM) and the Arbitrage Pricing Theory (APT). Therefore, RMT-based analysis is closely coupled to economic methods and theories (Chamberlain & Rothschild, 1983; Connor & Korajczyk, 1986; Nguyen et al., 2018).

## 5.6 Timescale Characteristics of Correlation

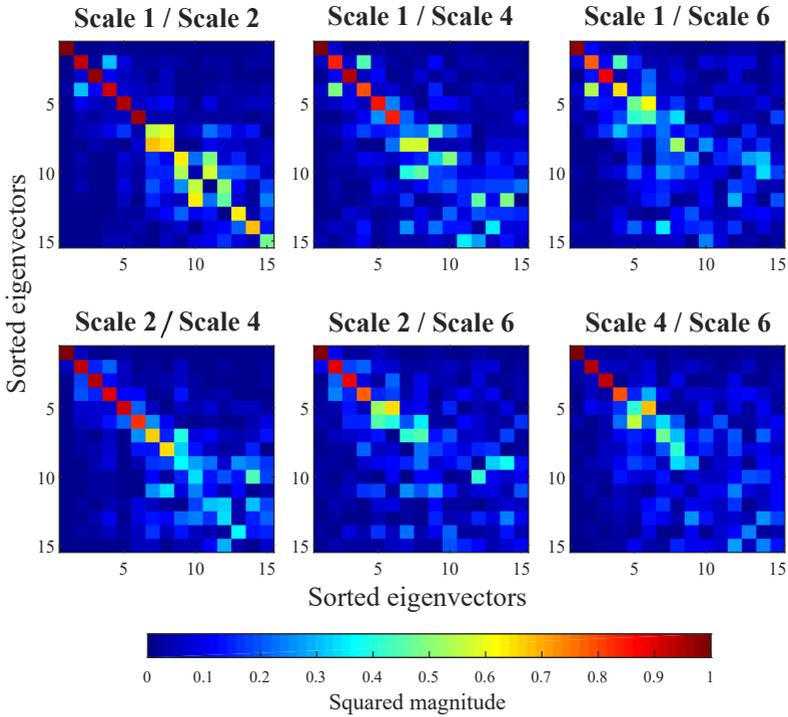
The fact that the correlation matrix is driven by the same underlying factors raises the question about whether correlations actually change over different time horizons. Therefore, this section explores the timescale properties of correlation statistics in more detail.

To assess the stability of the eigenvector structure (i.e., the exposure of individual stocks to a certain factor) across different timescales, I introduce a new metric. I choose those eigenvectors that correspond to the 15 largest eigenvalues and derive the  $15 \times 268$  eigenvector matrix  $\mathbf{U}_{\lambda_j}$  for every timescale  $\{\lambda_j; j = 1, \dots, 7\}$ . Next, I use these eigenvector matrices to construct an overlap matrix  $\mathbf{O}(\lambda_i, \lambda_j) = \mathbf{U}_{\lambda_i} \mathbf{U}_{\lambda_j}^T$  where  $\lambda_i$  and  $\lambda_j$  refer to the  $i$ -th and  $j$ -th timescale. If the eigenvectors are absolutely stable over the different timescales,  $\mathbf{O}$  is equivalent to the identity matrix.

In contrast to the analysis in Figure II.7, this metric is detached from interpreting eigenvalues and eigenvectors. The only problem with the metric arises when the rank-order of the eigenvalues changes. In this case, different eigenvector components are compared. However, the results in Figure II.7 show that the eigenvalue order is generally stable (at least for the five lowest timescales).

Figure II.10 displays the  $\mathbf{O}$ -metric for the given sample using a heat-map. For the largest four eigenvalues, the corresponding eigenvectors show no significant changes across different timescales. A modest decrease in eigenvector relations occurs only at scale level 6. Therefore, results indicate a high degree of stability across different timescales for eigenvectors corresponding to the largest eigenvalues.

However, for the remaining eigenvectors, the  $\mathbf{O}$ -metric only reveals minor coherence between the individual contributions of the eigenvector components at different time horizons. Therefore, the remaining eigenvectors are less stable. Again, these results need to be considered with caution given that the eigenvalue order may change between the different timescales. As a result, two different eigenvectors might be compared. Nevertheless, shifts in the eigenvalue order imply changes in the relative contribution of the eigenvalues. This would also be an indication of a less consistent correlation structure across timescales.



**Figure II.10:** Consistency of the eigenvector components (corresponding to the 15 largest eigenvalues) across different timescales ( $\mathcal{O}$ -metric). Notes: Eigenvectors are rank-ordered with respect to the eigenvalues (1=largest eigenvalue); the color code ranges from blue (low consistency, squared magnitude) to red (high consistency, squared magnitude).

Because the contribution of the eigenvectors remains more or less stable across different timescales, only eigenvalues may lead to time-inconsistent correlations (see section 5.1). Therefore, we now explore the timescale characteristics of the eigenvalues. The basic properties of eigenvalues link them to the total variance of the correlation matrix. The ratio of the  $k$ -th largest eigenvalue  $\mathcal{E}_k$  to the number of stocks  $M$  describes the proportion of variation in the correlation matrix that is explained by the eigenvalue. Accordingly, the highest eigenvalue explains the highest amount of variance. This property allows studying the contribution of an eigenvalue

to total system variance across different timescales. This analysis is linked to the investigation of the eigenvalue spectrum in section 5.3. However, it provides a more refined picture of the actual contribution of those eigenvalues that deviate from RMT predictions.

Panel A in Table II.3 shows the part of variance in the correlation matrix that can be explained either by a single eigenvalue or by a group of eigenvalues. For scale level 1, 25.86% of the variation in the correlation matrix is associated with the largest eigenvalue. With increasing timescale, this percentage even tends to increase. As a result, the largest eigenvalue roughly explains 30.23% of the variance in the correlation matrix at scale level 7. The second largest eigenvalue already exhibits significantly less explanatory power with only 3.07% at scale level 1. However, this value increases to 6.27% for scale level 7 (half-yearly to yearly timescale). Similar increasing trends in explanatory power can be observed for the remaining eigenvalues in Table II.3.

Panel B in Table II.3 presents the cumulative explained variance. For scale level 1, 34.33% of the variance is explained by the 5 largest eigenvalues. Due to the aforementioned increase in explained variance with an increase in timescale, the cumulative percentage of explained variance even rises to 47.77% at scale level 7. The contribution of the largest eigenvalues to the total variation of the correlation matrix is profound. Nevertheless, the majority of variation is still associated with the remaining eigenvalues  $\{\mathcal{E}_k; k = 6, \dots, 268\}$ .<sup>24</sup>

---

<sup>24</sup>Note that I present only the five largest eigenvalues separate in Table II.3 because this corresponds to the number of eigenvalues that are detected as deviating from the theoretical RMT bounds (see Table II.2).

**Table II.3:** *Percentage and cumulative percentage variance contribution of eigenvalues at different timescales.*

	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$	$\mathcal{E}_5$	$\mathcal{E}_{6,\dots,268}$
<b>Panel A: Percentage variance contribution</b>						
Scale 1	25.86%	3.07%	2.13%	1.95%	1.31%	65.67%
Scale 2	26.25%	3.22%	2.45%	1.97%	1.38%	64.73%
Scale 3	27.26%	3.48%	2.67%	2.18%	1.50%	62.91%
Scale 4	27.38%	4.04%	2.87%	2.15%	1.69%	61.87%
Scale 5	28.34%	4.57%	2.91%	2.15%	1.71%	60.32%
Scale 6	30.82%	5.42%	3.60%	2.75%	2.01%	55.41%
Scale 7	30.23%	6.27%	5.22%	3.38%	2.67%	52.23%
<b>Panel B: Cumulative percentage variance contribution</b>						
Scale 1	25.86%	28.93%	31.07%	33.02%	34.33%	100.00%
Scale 2	26.25%	29.47%	31.92%	33.89%	35.27%	100.00%
Scale 3	27.26%	30.74%	33.41%	35.59%	37.09%	100.00%
Scale 4	27.38%	31.42%	34.29%	36.45%	38.13%	100.00%
Scale 5	28.34%	32.91%	35.82%	37.97%	39.68%	100.00%
Scale 6	30.82%	36.24%	39.84%	42.58%	44.59%	100.00%
Scale 7	30.23%	36.50%	41.72%	45.10%	47.77%	100.00%

*Notes:* Percentage variance contribution (Panel A) is obtained by dividing the eigenvalue  $\mathcal{E}_k$  by the number of stocks  $M$ ; cumulative percentage variance is derived by summing over all eigenvalues that are larger than  $\mathcal{E}_k$  and dividing by the number of stocks  $M$ .

Consequently, a significant part of the correlation matrix can be explained by merely some few eigenvalues. Further, the analysis over different timescales reveals that the largest eigenvalues contribute more to the total variance with increasing timescale. Therefore, correlation structure is subject to significant changes with the investment horizon.

A fundamental property of eigenvalue decomposition is that the sum of all eigenvalues must equal the trace of the original correlation matrix. Consequently, if some eigenvalues increase, others must decrease, to compensate for these effects. The observation that the largest eigenvalues increase with the timescale implies that the variation that is explained by lower eigenvalues must in turn decrease with the timescale. In fact, Table II.3 shows a decreasing trend for eigenvalues  $\lambda_6$  to  $\lambda_{268}$ . Hence, the

increase in the explanatory power of the largest eigenvalue is mostly covered by a reduction in the smaller eigenvalues of the bulk. The considerations in the previous section have shown that the deviating eigenvalues can be interpreted as market or sector factors. Eigenvalues from the bulk of the eigenvalue spectrum, on the other hand, are associated with random noise. The increase in the explanatory power of the largest eigenvalues indicates that the systematic risk factor, as well as factors relating to the stock's sector affiliation, become more important for describing the variation in the correlation matrix. Therefore, we can conjecture that the structure of the correlation matrix becomes more systematic with increasing timescale.

Handa et al. (1989) and Gençay et al. (2005) studied the timescale properties of the systematic risk using the Sharpe-Lintner CAPM model and reached a similar conclusion. Both studies found that the relationship between the return of a portfolio and the market becomes stronger with increasing timescale. Therefore, they stated that beta is horizon-inconsistent and that the systematic risk factor explains more of the variation in the covariance matrix over longer time horizons. Similarly, Fama (1980, 1981) reported that the explanatory power of macroeconomic variables increases with increasing time length.

One way of interpreting these results is by turning to heterogeneous market theories. High-frequency fluctuations in stocks are assumed to be induced by short-term traders with short-term investment perspectives. These investors rely on idiosyncratic, i.e., company-specific, news to determine their investment strategy while paying less attention to systematic market news. On the other hand, long-term investors base their investment decisions on long-term systematic and general market information. As a result, the explanatory power of beta and of the other factors (largest eigenvalues) becomes more pronounced when considering long-term horizons.

The previous analysis has shown that the largest eigenvalues, and thus the amount of variance that is explained by the deviating eigenvalues, increases with the timescale. However, no evidence exists that these deviations between eigenvalues of different timescales are significant.

Earlier studies on the multiscale properties of stocks have often omitted providing statistical tests for eigenvalues and were more descriptive in nature. However,

the present study introduces the necessary tests for evaluating the deviations between eigenvalues and thus provides a more quantitative assessment of differences in timescale correlation structures.

Moreover, this study presents the properties of eigenvalues for different market periods (namely crisis and non-crisis market states). Similar to correlations tending to surge during distressed market periods (see section 5.1), it is likely that the overall correlation structure also changes during those market states. These changes would be directly reflected in the eigenvalues of the correlation matrix. Consequently, I subdivide the observation period into non-crisis and crisis periods.

Table II.4 presents the three largest eigenvalues and the corresponding confidence bounds. Thereby, each eigenvalue is ascertained for the full periods and the non-crisis and crisis periods. The theoretical bounds are obtained from formula 9. In contrast, the empirical bounds are estimated using a block bootstrapping method. For this simulation, 1,000 random variable block samples were drawn from the original wavelet time series with replacements at each timescale. The correlation matrix and the corresponding eigenvalues were calculated for each new sample. Finally, the confidence interval was retrieved from the respective quantiles of the empirical eigenvalue distribution.

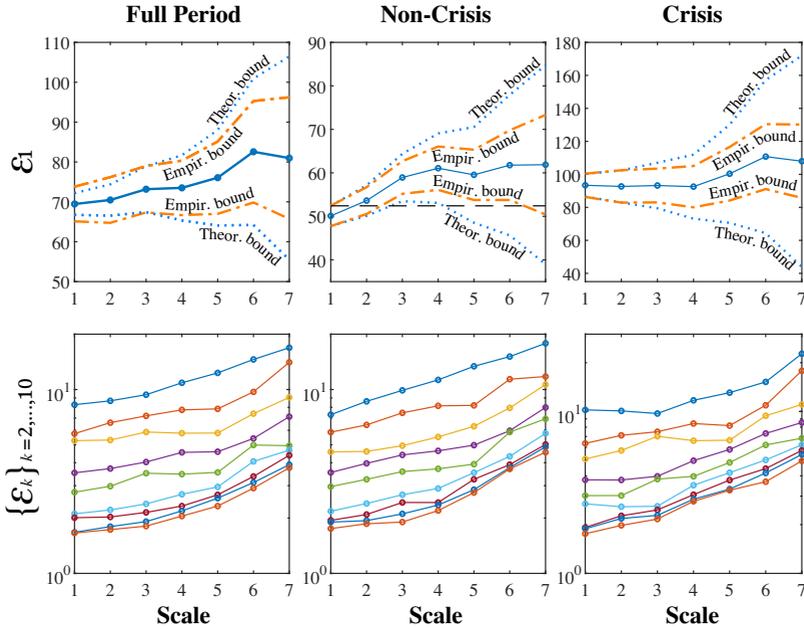
The block bootstrapping technique is necessary to account for the respective serial-correlations in the time series. Unlike the DWT, the MODWT is a non-decorrelating transformation (see section 3.1). Hence, neighboring MODWT coefficients share some interrelations that should not be discarded by the resampling procedure. The block size is set to correspond to twice the length of the respective timescale  $\lambda_j = [2^j, 2^{j+1}]$ . I have chosen this block size because it corresponds to twice the distance between coefficients of the decorrelating DWT. Unfortunately, a block bootstrap provides no stationary observations. To test for the validity of the results, I therefore also used a stationary bootstrap to specify the confidence bounds (not shown). However, the stationary bootstrap generally showed smaller standard errors, which is why the block bootstrap method using specified block sizes was used instead. This guarantees the robustness of the results.

**Table II.4:** Analysis of the three largest eigenvalues across different timescales for the full period and for the non-crisis and crisis periods.

	Eigenvalue $\mathcal{E}_1$			Eigenvalue $\mathcal{E}_2$			Eigenvalue $\mathcal{E}_3$		
	Full Period	Non-Crisis	Crisis	Full Period	Non-Crisis	Crisis	Full Period	Non-Crisis	Crisis
Scale 1	69.46 (66.73, 72.19) (65.11, 73.80)	50.10 (47.79, 52.42) (47.74, 52.46)	93.39 (86.47, 100.31) (86.34, 100.45)	8.27 (7.94, 8.59) (7.39, 9.15)	7.30 (6.96, 7.64) (6.88, 7.71)	10.22 (9.46, 10.98) (8.05, 12.39)	5.75 (5.53, 5.98) (5.41, 6.09)	5.86 (5.59, 6.13) (5.52, 6.20)	6.35 (5.88, 6.82) (5.60, 7.10)
Scale 2	70.47 (66.55, 74.39) (64.74, 76.20)	53.62 (50.12, 57.13) (50.54, 56.71)	92.68 (82.97, 102.40) (82.87, 102.50)	8.68 (8.19, 9.16) (7.91, 9.45)	8.62 (8.06, 9.18) (8.04, 9.20)	10.08 (9.02, 11.13) (8.46, 11.69)	6.60 (6.23, 6.97) (6.15, 7.06)	6.42 (6.00, 6.84) (6.04, 6.80)	7.12 (6.37, 7.87) (6.47, 7.77)
Scale 3	73.17 (67.42, 78.93) (67.35, 79.00)	58.94 (53.49, 64.38) (55.18, 62.69)	93.24 (79.42, 107.06) (83.00, 103.47)	9.37 (8.63, 10.11) (8.68, 10.06)	9.89 (8.98, 10.81) (9.16, 10.63)	9.71 (8.27, 11.14) (8.06, 11.35)	7.19 (6.62, 7.75) (6.50, 7.88)	7.47 (6.78, 8.16) (6.64, 8.30)	7.49 (6.38, 8.60) (6.14, 8.85)
Scale 4	73.51 (65.34, 81.69) (66.68, 80.34)	61.07 (53.09, 69.05) (56.09, 66.05)	92.55 (73.15, 111.95) (80.02, 105.07)	10.88 (9.67, 12.09) (9.69, 12.06)	11.28 (9.81, 12.75) (10.09, 12.47)	11.71 (9.25, 14.16) (9.22, 14.20)	7.75 (6.89, 8.61) (6.92, 8.58)	8.15 (7.09, 9.22) (7.17, 9.14)	8.42 (6.65, 10.18) (6.74, 10.09)
Scale 5	76.06 (64.09, 88.02) (66.99, 85.13)	59.54 (48.54, 70.55) (53.76, 65.33)	100.38 (70.62, 130.13) (84.14, 116.61)	12.31 (10.37, 14.25) (10.80, 13.82)	13.37 (10.90, 15.85) (11.79, 14.96)	13.06 (9.19, 16.93) (9.35, 16.77)	7.85 (6.61, 9.08) (6.67, 9.03)	8.20 (6.68, 9.71) (6.84, 9.55)	8.17 (5.75, 10.60) (5.49, 10.86)
Scale 6	82.56 (64.20, 100.93) (69.84, 95.29)	61.78 (45.63, 77.92) (53.76, 69.79)	110.76 (64.33, 157.20) (91.08, 130.44)	14.58 (11.33, 17.82) (12.25, 16.90)	15.11 (11.16, 19.06) (12.44, 17.77)	15.22 (8.84, 21.60) (9.61, 20.82)	9.71 (7.55, 11.87) (7.94, 11.47)	11.38 (8.40, 14.35) (8.79, 13.97)	10.92 (6.34, 15.50) (7.07, 14.76)
Scale 7	80.98 (55.50, 106.46) (65.73, 96.23)	61.88 (39.00, 84.75) (50.43, 73.32)	107.98 (43.96, 171.99) (85.75, 130.20)	16.88 (11.57, 22.19) (11.71, 22.06)	17.83 (11.24, 24.42) (12.96, 22.71)	22.71 (9.25, 36.17) (11.47, 33.95)	14.06 (9.64, 18.48) (9.76, 18.36)	11.77 (7.42, 16.12) (6.53, 17.01)	17.83 (7.26, 28.40) (11.37, 24.29)

Notes: The 95% confidence interval for the theoretical (top values) and the block bootstrap (bottom values) eigenvalue distributions are depicted in parentheses; block bootstrap confidence bounds are obtained using 1,000 random simulations.

For the full sample period, all eigenvalues in Table II.4 show an increasing trend with increasing timescale. This is in accordance with the observations in Table II.3. However, unlike the second and third largest eigenvalues, the differences between the largest eigenvalue  $\mathcal{E}_1$  at different timescales are too small to be significant at a 5% level of significance. For example, the largest eigenvalue at timescale 1 (69.46) is well within the theoretical and empirical bounds of all higher timescales. Therefore, no significant deviations between timescale realizations of the largest eigenvalue can be stated for the full sample period.



**Figure II.11:** Evolution of eigenvalues across different timescales for the full period and for the non-crisis and crisis periods. Notes: "Theoretical bound" indicates the 95% confidence interval of the theoretical eigenvalue distribution; "Empirical bound" describes the 95% confidence interval estimated using a block bootstrapping with 1,000 random simulations; the horizontal dashed line (non-crisis period) extrapolates the lower confidence bound of wavelet correlation at scale level 1 across all scale levels; the y-axis for the eigenvalues  $\mathcal{E}_3, \dots, \mathcal{E}_{10}$  is log-transformed.

Dividing the time series into non-crisis and crisis periods allows for a more detailed interpretation of the eigenvalue structure. In non-crisis periods, the largest eigenvalue  $\mathcal{E}_1$  now exhibits significant variation across timescales  $\lambda_j$ . While the eigenvalue is 50.10 at scale level 1, the same eigenvalue increases to 61.07 at scale level 4. For a better understanding of this trend, Figure II.11 visualizes the eigenvalues for the different timescales. The first row of subfigures represents the largest eigenvalue for the full period and for the non-crisis and crisis periods. The blue- and orange-dashed lines serve as reference for comparing eigenvalues at different scale levels. Jointly examining this graph and the results in Table II.4 shows that the largest eigenvalue is indeed horizon-inconsistent during non-crisis periods. Both the theoretical and empirical eigenvalue bounds at scale level 1 and scale level 4 do not overlap. This implies that the null hypothesis of equal eigenvalues is at least rejected at a 2.5% level of significance. If the empirical bounds are considered, deviations are even significant between the shortest and longest timescales.

Analyzing the largest eigenvalue in crisis periods draws a different picture. No significant differences can be observed between the largest eigenvalue of different timescales. Indeed, Table II.4 indicates that the largest eigenvalue even slightly decreases from scale levels 1 to 4.

Consequently, the influence of the market factor on the general correlation structure markedly differs between non-crisis and crisis periods. While the impact of the systematic factor becomes stronger with increasing timescale during non-crisis periods, this is not the case during crisis periods.

One possible interpretation of these results may again be found in heterogeneous market theories. During regular market periods, idiosyncratic risks have a higher impact on the correlation structure at shorter frequencies (low timescales) while the market factor makes a smaller relative contribution. However, over longer timescales, the impact of macroeconomic variables becomes stronger and thus both the largest eigenvalue and the share of explained variance increase.

In contrast, during crisis periods, news (shocks) affect all stocks similarly over all investment horizons. Hence, both long-term and short-term investors are directly exposed to market risks during these periods. As a result, stocks are strongly cor-

related and feature a collective mode across all time horizons. The market factor becomes the universal element and hence equally affects stocks at all timescales (Billio, Getmansky, Lo & Pelizzon, 2012; Zheng et al., 2012). By contrast, the impact of idiosyncratic risks on the correlation structure reduces. The assimilation of the largest eigenvalue across timescales results from this harmonization of risk across time horizons.

Complementary to this analysis of the largest eigenvalues, the second row of subfigures in Figure II.11 presents results for the next nine largest eigenvalues (log-scale plot). Interestingly, the next higher eigenvalues show a distinct behavior. Almost all eigenvalues increase with the timescale and show a positive trend irrespective of the market state (significant at the 1–5% level for non-crisis periods and at the 5–10% level for crisis periods). As shown in the previous section, these eigenvalues are typically associated with factors referring to sector or industry classifications. Hence, these findings imply that the sectorial factors become more relevant for long-term cycles in crisis and non-crisis periods. In other words, correlations between stocks in the same sector increase with increasing timescale regardless of the market state. However, the relative changes of these eigenvalues, between crisis and non-crisis periods, are far smaller than for the market factor. This can be explained by *eigenvalue repulsion*. The market factor (largest eigenvalue) becomes predominant during distressed market periods. This increase in the largest eigenvalue must be compensated by decreases in the magnitude of the remaining eigenvalues. Even though this compensation mainly stems from the bulk of the eigenvalue spectrum, it leaves less room for the intermediate eigenvalues to expand. Hence, the smaller relative change results from the declining relative importance of the sector factors compared to the market factor.

Having established timescale-inconsistent eigenvalue behavior, I now study the effects on the structure of the correlation matrix. The above results suggest the existence of a market-induced structure in correlations that changes across timescales. Similarly, the correlation structure is found to be driven by timescale-dependent sector characteristics. Therefore, I used filtering methods to separate the correlation matrix from these effects. The correlation matrix is thus decomposed into a market

( $\mathbf{C}_M$ ), a group ( $\mathbf{C}_G$ ), and a random ( $\mathbf{C}_R$ ) component:

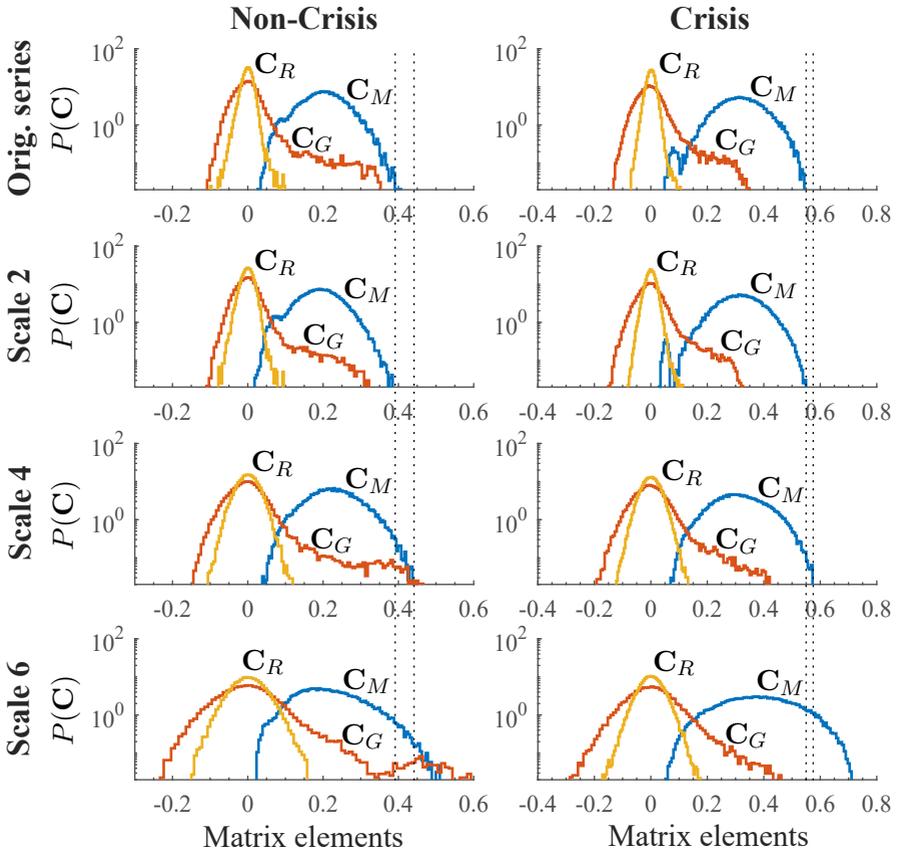
$$\mathbf{C} = \varepsilon_1 \mathbf{u}_1 \mathbf{u}_1^T + \sum_{k=2}^{M_g} \varepsilon_k \mathbf{u}_k \mathbf{u}_k^T + \sum_{k=M_g+1}^M \varepsilon_k \mathbf{u}_k \mathbf{u}_k^T = \mathbf{C}_M + \mathbf{C}_G + \mathbf{C}_R, \quad (22)$$

where  $M_g$  refers to the number of eigenvalues that are used to define the collection of group correlations. Although the diagonal of the filtered components does not equal unity, these matrices can be interpreted and are referred to as correlation matrices. The market component  $\mathbf{C}_M$  relies on the largest eigenvalue and the corresponding eigenvectors. The group correlation matrix  $\mathbf{C}_G$  is the filtered correlation matrix of stock groups. In accordance with the findings in section 5.3, I specify  $M_g = 9$ . This value corresponds to roughly the average number of deviating eigenvalues across all timescales (minus the largest eigenvalue).<sup>25</sup> Finally, the random component  $\mathbf{C}_R$  is obtained from those eigenvalues that belong to the bulk of the eigenvalue spectrum. Again, this decomposition can be derived for different time horizons.

Figure II.12 shows the results for the decomposed components of the correlation matrix for non-crisis and crisis market states across different timescales  $\{\lambda_j; j = 0, 2, 4, 6\}$ . For non-crisis states, the right tail of the distribution of the market correlation matrix (blue line) gradually increases from the daily time horizon ( $\lambda_0$ ) to the half-yearly time horizon ( $\lambda_6$ ). In contrast, the left tail of the distribution remains roughly equal across all timescales. This finding is consistent with the trend in the largest eigenvalue (more extreme values) and the expectations of increasingly positive correlations for longer time horizons. Surprisingly, a similar pattern can be observed for the distribution of the market correlation matrix during crisis periods. However, the broadening of the distribution is mainly present at the largest timescale. This suggests that macroeconomic factors become more relevant at large timescales during crisis periods than initially presumed. Contrary to the findings in Figure II.11, correlations may thus also show timescale-varying properties during crisis periods.

---

<sup>25</sup>No profound alteration in results was registered for small changes in  $M_g$ .



**Figure II.12:** Distribution of filtered correlation matrix for non-crisis and crisis periods at different timescales. Notes: Correlations is filtered into market ( $C_M$ ), group ( $C_G$ ), and random ( $C_R$ ) components. The vertical lines indicate the largest element of the market correlation matrix for the original (untransformed) time series and the wavelet time series at scale level 4; the y-axis is log-transformed.

Similarly, the group component (red line) demonstrates non-trivial characteristics. For non-crisis periods, the distribution of the group correlation shows an expanding right tail with increasing timescale. At timescale  $\lambda_6$ , the inter-group correlations even surpass those of the market component. This indicates that correlations become more systematic and inter-group correlations are likely to gain in significance over

longer time horizons. While the distribution of the correlation matrix also widens with increasing timescales during crisis periods, this expansion is far less pronounced than for non-crisis periods. Notably, the probability of high inter-group correlations is in fact greater during non-crisis periods. Again, this is most likely the result of *eigenvalue repulsion*. During crisis periods, common market-wide behavior dictates stock correlations. Thus, the market component represses the contribution of the inter-group component to overall correlation. In contrast, during non-crisis periods, the market component is less dominant and inter-group correlations have more leeway to unfold.

The random component (yellow line) shows only minor changes over different time horizons and is comparable between the two market states. This agrees with theoretical expectations for the random component. In particular, the finding that the distribution of the random component does not change for different market states reinforces the assumption that these correlations are in fact random.

Consequently, the structure of the correlation matrix is determined by systematic factors that change over time and with respect to market conditions. It is defined by a subtle interplay of timescale-variant collective interconnectedness and inter-group relations. These correlations are overlaid by idiosyncratic noise in the form of random perturbations.

Generally, the structure of the correlation becomes more systematic with increasing timescale. However, in times of crisis, the correlations for all time horizons are equally characterized by a common behavior. While the market factor is overall the dominant component, group correlations gain in significance in non-crisis periods and for longer time horizons.

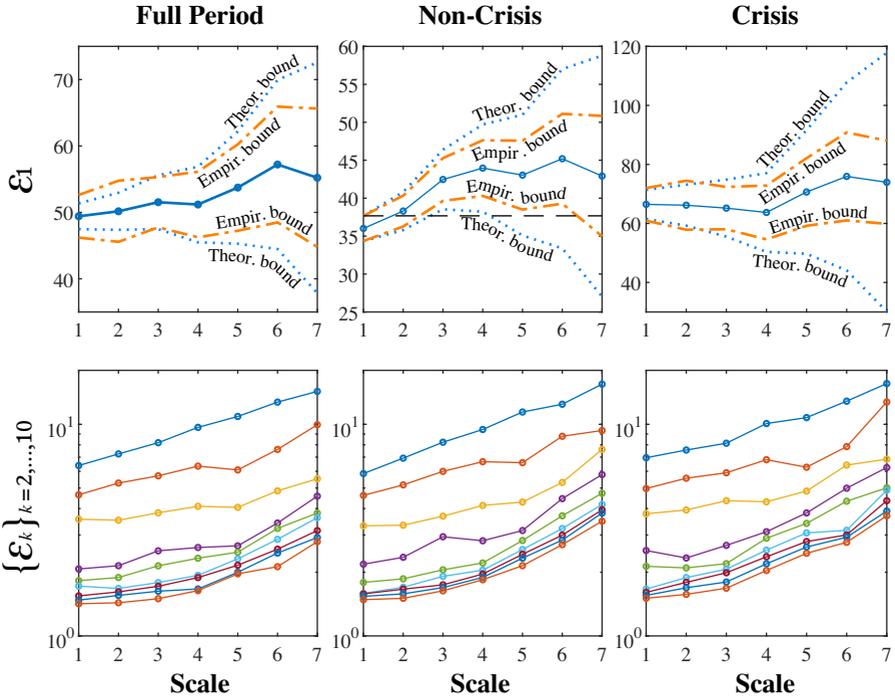
The results that correlations are determined by fundamental structural relationships that change across timescales and market periods has important implications in practice. For example, portfolio and risk management decisions may have different impacts on short- and long-term investment horizons. Thus, considering the timescale properties of correlations is crucial to assessing potential diversification benefits. Similarly, the findings presented here help to resolve certain contradictions in existing research. Studies using data from different time horizons produce diverging and

sometimes even inconsistent results. The present study shows that such contradictions might result from timescale-variant correlations. Finally, the obtained results provide new insights into the functioning of stock markets.

## 5.7 Robustness

This section tests the reported results for robustness using an extended time interval (June 30, 1970 to June 30, 2018). This longer measurement period enables reducing measurement noise. At the same time, the required full track history reduces the sample to 185 stocks. This describes the tradeoff between the depth and width of the sample that was mentioned at the beginning of this study (in section 4).

Figure II.13 shows the ten largest eigenvalues for this increased sample period. We note a reduction in magnitude of the eigenvalues. However, this reduced magnitude results from the smaller sample size and is not due to changes in the relevance of the respective components. Rather, it is a consequence of the property of eigenvalues, which requires the trace of a matrix to equal the sum of its eigenvalues  $tr(\mathbf{C}) = \sum_{k=1}^M \mathcal{E}_k$ . If we therefore ignore the magnitude of the eigenvalues, almost no change is evident compared to the results of the previous section. The eigenvalue curve shows a similar trend. In fact, deviations of the largest eigenvalues even slightly increase in significance.



**Figure II.13:** Evolution of eigenvalues across different timescales for the full period and for the non-crisis and crisis periods for an extended time series encompassing June 30, 1970 to June 30, 2018. Notes: "Theoretical bound" indicates the 95% confidence interval of the theoretical eigenvalue distribution; "Empirical bound" describes the 95% confidence interval estimated using a block bootstrapping with 1,000 random simulations; the horizontal dashed line (non-crisis period) extrapolates the lower confidence bound of wavelet correlation at scale level 1 across all scale levels; the y-axis for the eigenvalues  $\mathcal{E}_3, \dots, \mathcal{E}_{10}$  is log-transformed.

## 6 Conclusion

Analyzing the structure of correlations reveals that stock interdependencies are driven by a set of fundamental factors. This study has shown that the meaning and the ranking of these factors remains comparable across different time horizons. Only a few changes in the general structures have been observed over different timescales. This research shows that the market factor (largest eigenvalue) is the principal determinant of the correlation structure regardless of whether daily or monthly stock correlations are considered. Similar findings emerge for the remaining key factors (intermediate eigenvalues). These factors retain their significance for explaining the correlation structure across different time horizons.

It is noteworthy that stock correlations at low timescales show strong sectoral structures given that idiosyncratic movements are often believed to account for the correlation structure of stocks over short-term periods (lower explanatory power of CAPM for daily data). In contrast, previous research generally attaches greater significance to macroeconomic factors to describe longer-term dependency structures. The results of this study relativize this view to some degree. The present findings indicate that the short- and longer-term dependency structures of the stock market are determined by the same fundamental systematic factors.

However, this study has also revealed that the strength of these underlying factors varies over different time horizons. For example, factors reflecting sectoral affiliation become increasingly important for longer-term periods. While the same factors determine the dependency structure, they have a timescale-variant impact on correlations. Hence, assuming timescale-variant correlation structures nonetheless seems legitimate. Idiosyncratic risk may more strongly influence short-term correlations, while longer-term dynamics are increasingly driven by structural factors.

Consideration of the market factor leads to a more diffuse interpretation. While the market factor becomes increasingly important for longer time horizons during non-crisis periods, this is not the case during crisis periods. This can be explained by the fact that the market factor exerts even greater influence on correlation structure in times of crisis. As the predominant factor, it determines the short- and longer-term structures in the correlation matrix during these periods.

Therefore, the results show that the stock market can be described in terms of structural factors and of how their influence changes over different time horizons. The complex interplay of these factors produces timescale-variant correlations. As a result, this study gives new insights into the mechanisms of stock markets and into the rationale underlying heterogeneous timescale-dependent correlation structures.

These results provide new perspectives for understanding stock markets. For example, they help to explain conflicting empirical results for tests of capital market models (e.g., CAPM). These models often display different explanatory power under consideration of different return intervals (daily, monthly, yearly). The assumption of changing correlation structures due to timescale-variant influencing factors (in combination with heterogeneous market theories) at least partially helps to fathom these contradictions.

Moreover, the present findings also have central implications for portfolio and risk management decisions. They suggest that investors with different investment horizons should adapt their investment strategy to the correlation structures of the corresponding timescale. This insight lays the foundation for future research into timescale-optimized portfolio strategies.

Finally, the methodological approach of combining wavelet transformation with random matrix theory has possible applications in other academic and practical disciplines. Specifically, the method could be used in image processing. Wavelet transformation allows filtration of images into different granularities. Random matrix theory could be used to identify common components between decomposed images and filter non-random information from this data. Thus, the approach could for example be used in face recognition algorithms. Similarly, this methodological procedure could be adopted in medical research for the joint analysis of MRI images (see, for example, Conlon, 2009). This could make it possible to isolate characteristics that may contribute to disease detection. Many other possible applications across other academic and practical disciplines are conceivable.



# Chapter III

## Portfolio Optimization Under Heterogeneous Investment Horizons

*Christian Vial*

### 1 Introduction

Financial markets are *complex interacting systems* of heterogeneous agents. These agents differ in their risk profile, beliefs, investment constraints, consumption requirements, as well as in their perception and interpretation of information. Despite these observable heterogeneities, many economic models neglect some of them and rely upon the simplifying assumption of homogeneous agents. Economists have often viewed this assumption as an oversimplification and have proposed alternative models that incorporate heterogeneities. Among others, Dacorogna et al. (1998), Lux and Marchesi (1999), as well as Nekhili, Altay-Salih and Gençay (2002) showed that the hypothesis of heterogeneous agents helps to explain stylized facts of financial time series (e.g., fat-tails, long-memory, long-range dependence, and excess kurtosis).

A central assumption of these heterogeneous agent theories is the idea of agents operating at different (heterogeneous) investment horizons. Following this intuition of heterogeneity, financial markets consist of a diverse group of agents including but not limited to intraday traders, day-traders, fund managers, central banks, insurance companies, and pension funds. These agents have different life-cycle preferences and base their decisions on heterogeneous investment horizons (timescales).<sup>1</sup> For example, agents with shorter investment horizons may perceive new information (news) as negative, while agents with longer investment horizons consider the same

---

<sup>1</sup>Peters (1994) and Müller et al. (1993, 1997) formalized the theory of heterogeneous market agents with the Fractal Market Hypothesis (FMH) and the Heterogeneous Market Hypothesis (HMH).

information as a buying opportunity. In turn, agents may also adapt investment horizons according to their perception of information and the relative attractiveness of a short-term over a long-term strategy.

The assumption of timescale heterogeneity thereby implies both the *risk of financial assets* and their *interdependencies* to be timescale-dependent. These two dimensions are central to portfolio allocation decisions (Chakrabarty et al., 2015). A portfolio allocation suitable for a short-term investor might therefore not be optimal for a long-term investor (and vice versa) even though both possibly share identical objectives and risk tolerance levels. For example, a pension fund might ignore short-term risk due to the fund's longer-term life-cycle preferences and investment horizon. In contrast, a day trader, who operates on a shorter timescale, might consider long-term risk irrelevant. Consequently, it seems necessary for heterogeneous investors to optimize their portfolios over their respective investment horizon.

This study aims to demonstrate that taking into account investment timescales benefits portfolio allocation decisions. The main hypothesis is that a portfolio strategy can be formed that minimizes the risk at a specific investment timescale. Accordingly, investors with different life-cycle preferences can minimize their risk over their respective investment horizon.

To test this hypothesis, I proceed in two steps: First, I examine stocks for timescale dynamics in terms of their dependency structures and variances. Second, I study the implications of those timescale properties of stock prices for portfolio optimization. Specifically, I develop portfolio strategies that aim to minimize portfolio variance at a specific time horizon (timescale). Further, I examine the ability of these portfolio strategies for minimizing risk at a particular timescale. For this examination, out-of-sample volatilities of those timescale-optimized portfolios are compared with each other and with the volatility of a portfolio obtained from classical time domain methods.<sup>2</sup>

Researchers have investigated heterogeneity in asset prices before by using long-term and short-term time series analysis (see, for example, Burns & Mitchell, 1946).

---

<sup>2</sup>Classical time domain methods refer to the derivation of portfolios using covariance matrices that are obtained from untransformed daily return observations. This study adopts different estimation windows sizes for calculating these covariance matrices.

In contrast, I apply wavelet decomposition to study timescale dynamics in stock prices and to create portfolios optimized for a specific investment horizon.<sup>3</sup> Wavelet analysis enables analyzing finer frequency granularities by using a more sophisticated filtering technique.<sup>4</sup>

My research contributes to existing literature in four ways. First, it analyzes the dynamics of individual assets and their dependency structures in the time- and spectral-domain. Second, it extends existing portfolio optimization research by introducing timescale considerations to portfolio allocation decisions. Thus, it constructs and analyzes investment strategies in the context of heterogeneous investment horizons using wavelet decomposition. Third, it introduces wavelet variance equality test statistics to test differences in the variance of these timescale-optimized portfolios. Fourth, it presents a new method for portfolio optimization, namely, multiscale portfolio optimization.

Chakrabarty et al. (2015) observed that earlier wavelet studies in finance were predominantly exploratory. The present study bridges the gap between heterogeneous investment analysis and practical application by constructing horizon-heterogeneous portfolio strategies. It is structured as follows: Section 2 reviews the existing literature on investment horizon-based portfolio optimization. Section 3 describes the wavelet methodology and its application to portfolio optimization. Section 4 introduces the general properties and assumptions of the optimization method. Section 5 provides a brief data overview. Section 6 presents the main findings of the timescale analysis for individual stocks, as well as the results of the analysis of timescale-optimized portfolio strategies. Section 7 introduces multiscale portfolio optimization as a basis for considering multi-horizon preferences in portfolio optimization. Section 8 summarizes the results of this study and provides input for future research.

---

<sup>3</sup>Wavelet theory percolated into the sphere of economics and finance with the advent of so-called econophysics (Iacobucci, 2003).

<sup>4</sup>Unlike the more familiar Fourier transformation, wavelet analysis permits simultaneous representation of the signal in the time and the timescale domain.

## 2 Literature Review

Research on the impact of the investment horizon on asset prices and portfolio decisions is not new in finance. In 1972, Levy demonstrated that incorrect assumptions about an investor's holding period can lead to systematic biases in performance and risk measures. He conjectured that an empirical study using daily data might yield different results than the same study using a monthly sampling interval. This bias results from differences between the assumed investment horizon and an investor's actual behavior.

In a later study, Gressis, Philippatos and Hayya (1976) analyzed the effects of heterogeneous investment horizons on portfolio selection using a multiperiod mean-variance approach. They found that the investment horizon affects portfolio choice. Two investors sharing the same utility function but underlying different investment horizons will choose different portfolios. Consequently, the authors concluded that the utility function is not sufficient for determining portfolio choice. Gressis et al. (1976) suggested that the investment horizon should thus become an additional parameter of an investor's utility function.

These findings fostered further research on investment horizon effects and on how different return intervals impact performance and risk measurement (see Levhari & Levy, 1977; Gilster, 1983; Handa et al., 1989; Handa, Kothari & Wasley, 1993; Brailsford & Faff, 1997; Bjornson, Kim & Lee, 1999). Following this reevaluation, researchers began characterizing financial markets as *dynamic systems* consisting not only of one, but of many agents with heterogeneous investment horizons operating at different frequencies (Corsi, 2009; Dacorogna et al., 2001; Gençay et al., 2002, 2005; In & Kim, 2013; Ramsey, 2002).

Spectral theory — and more recently wavelet theory — have laid the foundation for analyzing frequency characteristics of financial time series. The earliest adoption of wavelet theory to economics and finance can be attributed to Ramsey and Zhang (1997), Ramsey and Lampart (1998a, 1998b), and Gençay et al. (2002). Since then, several studies investigating heterogeneity in risk and dependency structures between

financial assets have appeared.<sup>5</sup> However, only few studies have explored timescale-based (i.e., frequency-based) portfolio strategies.

Orlov and Äijö (2015) investigated carry trade diversification under consideration of timescale dynamics for the eight most liquid currencies and for LIBOR rates. At each rebalancing date, they grouped the two assets with the lowest correlation (across all horizons) into one portfolio. The authors concluded that the constructed wavelet portfolios exhibit improved Sharpe ratios compared to equally weighted portfolios and the S&P 500. However, Orlov and Äijö (2015) did not observe outperformance for all timescales.

In a recent empirical study, Berger (2016) used wavelet decomposed US stock returns to form timescale-optimized portfolios. He showed that portfolio strategies concentrating on short-term frequencies outperform strategies using raw return data. However, long-term timescale-optimized portfolios incur higher average losses than benchmark portfolios and short-term optimized portfolio analogs. Even though Berger (2016) constructs portfolio strategies based on wavelet decomposed returns, he evaluates performance and risk metrics only for daily portfolio returns. Thus, he leaves unconsidered optimization effects at higher timescales or increasing sampling intervals.

In a recent working paper, Chaudhuri and Lo (2016) introduced spectral portfolio theory using Fourier transformation. Their analysis highlights that asset prices change across time and frequency. Similarly, portfolio performance evaluation depends on the investment horizon under consideration. They concluded that the frequency dimension can be useful in portfolio design (specifically, if portfolio goals differ across time horizons).

### 3 Wavelet Methodology

Wavelet analysis is a method for studying the frequency (timescale) characteristics of a signal as a function of time. In contrast to Fourier transformation, wavelet

---

<sup>5</sup>It is referred to Chakrabarty et al. (2015) for a comprehensive review of pertinent literature about wavelet analysis in economics and finance.

analysis thereby enables simultaneously representing the signal in the time and in the timescale domain. This makes wavelet analysis a particularly useful instrument for examining the time evolution of spectral characteristics in financial time series. Specifically, it allows analyzing time series that feature non-stationary power at many frequencies (Daubechies, 1990; Torrence & Compo, 1998).

This study applies the discretized continuous wavelet transform (CWT) to analyze the wavelet squared coherence spectrum. Further, it applies the Maximal Overlap Discrete Wavelet Transform (MODWT) to construct scale-based portfolio strategies.

### 3.1 Continuous Wavelet Transform

Wavelet transformation employs an elementary function, the so-called wavelet function  $\psi(\cdot)$ , to decompose a time series into the time-timescale-dimension (Rua & Nunes, 2012). This wavelet function can be real- or complex-valued and has support on the real axis. The general definition of the wavelet function is relatively simplistic: A wavelet function must i) integrate to zero ( $\int_{-\infty}^{\infty} \psi(t) dt = 0$ ) and ii) be square integrable to unity ( $\int_{-\infty}^{\infty} \psi^2(t) dt = 1$ ). While the second condition imposes that the wavelet function is non-zero within a certain interval, the first condition implies that the excursions from zero must cancel out overall. These excursions are imposed to be localized, i.e., the non-zero values of the function are mostly limited to a finite interval (Percival & Walden, 2000). Accordingly, wavelet functions (wavelets) resemble small waves that vacillate around zero (hence their name). In contrast to their trigonometric counterparts in Fourier transformation, they do so only in a limited range.

The wavelet function is used to decompose a time series into time-timescale-components, the so-called continuous wavelet transforms. Let  $r_t$  specify a real-valued time series of an independent time variable  $t$ .<sup>6</sup> The continuous wavelet transforms  $\{\mathcal{W}(\lambda, \tau); \lambda > 0, -\infty < \tau < \infty\}$  are then obtained from the convolution of  $r_t$

---

<sup>6</sup>In fact, other units are also possible for the signal. It does not necessarily have to be a time series. In this study, however, the signal always represents a return series (time series).

with the wavelet function  $\psi(\cdot)$

$$\mathcal{W}(\lambda, \tau) \equiv \int_{-\infty}^{\infty} r_t \psi_{\lambda, \tau}^*(t) dt, \quad (1)$$

where

$$\psi_{\lambda, \tau}(t) \equiv \frac{1}{\sqrt{\lambda}} \psi\left(\frac{t - \tau}{\lambda}\right), \quad (2)$$

(\*) indicates the complex conjugate,  $\tau$  refers to the translation (location) parameter,<sup>7</sup> and  $\lambda$  describes the dilation (timescale) parameter of the wavelet function.<sup>8</sup> These translation and dilation parameters shift and scale the wavelet function  $\psi(\cdot)$ , respectively.<sup>9</sup> Thus, the wavelet function  $\psi(\cdot)$  serves as a prototype for the translated and shifted derivatives of itself  $\psi_{\lambda, \tau}(\cdot)$  (Madaleno & Pinho, 2012).

As the translation parameter  $\tau$  is altered, the wavelet function is shifted in time. This shifting procedure allows obtaining wavelet transforms for different times. As a result, wavelet transforms provide time resolution (time information).

Similarly, by varying the dilation parameter  $\lambda$ , the wavelet function is stretched (compressed). Thereby, a large timescale parameter results in a broad support of the wavelet function in the time domain. This large window enables isolating coarse features in the signal, i.e., low-frequencies characteristics. In contrast, a small scale parameter induces a narrow support of the wavelet function in the time domain. This small window allows extracting finer features in the time series, i.e., high-frequency characteristics (Madaleno & Pinho, 2012).<sup>10</sup>

Applying the transformation to a continuum of location and scale parameters pro-

---

<sup>7</sup>The translation parameter  $\tau$  and the time variable  $t$  are directly connected and describe the same dimension. However, a separate variable is used for the sake of clarity. In fact, the two values may differ if a non-symmetric wavelet function is used for transformation (circular shifting).

<sup>8</sup>Deriving the wavelet transforms for a discrete stock return series requires discretizing the integral in formula 1.

<sup>9</sup>A scale parameter  $\lambda > 1$  dilates, whereas a scale parameter  $\lambda < 1$  compresses the wavelet function. Typically, only positive scale factors  $\lambda > 0$  are considered.

<sup>10</sup>A wavelet function with a high scale parameter (narrow support) delivers good time resolution, but bad timescale (frequency) resolution. In contrast, a wavelet function with a low scale parameter (broad support) provides good timescale (frequency) resolution, but bad time resolution. This trade-off between time and frequency localization describes the well-known Heisenberg uncertainty principle in signal processing.

vides a joint representation of the original signal in the time-timescale-domain. In other words, the method simultaneously describes the high- and low-frequency content of the original time series and the change of these frequency characteristics over time. For the analysis of stock returns series, this means that different frequency (timescale) dynamics in stock returns can be extracted and studied for different time periods.

Note that the two necessary conditions (previously introduced) for the wavelet functions are met by many different functions. For analytical requirements and practical reasons, further conditions are typically imposed on the wavelet functions: among others, admissibility and regularity<sup>11</sup> conditions. The admissibility condition requires that  $C_\Psi \equiv \int_0^\infty \frac{|\Psi(f)|^2}{|f|} df$  satisfies  $0 < C_\psi < \infty$ , where  $f$  denotes the frequency, and  $\Psi(f)$  defines the Fourier transform of  $\psi(\cdot)$  given by  $\Psi(f) \equiv \int_{-\infty}^\infty \psi(t) e^{-i2\pi ft} dt$ .<sup>12</sup> This condition ensures that a square-integrable signal  $r_t$ , i.e.,  $\int_{-\infty}^\infty r_t dt$ , can be reconstructed<sup>13</sup> from the wavelet transforms  $\mathcal{W}(\lambda, \tau)$  such that

$$r_t = \frac{1}{C_\Psi} \int_0^\infty \int_{-\infty}^\infty \mathcal{W}(\lambda, \tau) \frac{1}{\sqrt{\lambda}} \psi\left(\frac{t-\tau}{\lambda}\right) d\tau \frac{d\lambda}{\lambda^2}. \quad (3)$$

Moreover, the condition guarantees conservation of energy:<sup>14</sup>

$$\int_{-\infty}^\infty |r_t|^2 dt = \frac{1}{C_\Psi} \int_0^\infty \left[ \int_{-\infty}^\infty |\mathcal{W}(\lambda, \tau)|^2 d\tau \right] \frac{d\lambda}{\lambda^2}, \quad (4)$$

<sup>11</sup>The regularity condition requires the wavelet function to be local in both the time and the frequency domain. To ascertain good localization in the frequency domain, the wavelet transforms  $\mathcal{W}(\lambda, \tau)$  needs to converge rapidly to zero with a decrease of the scale parameter  $\lambda$ . This speed of convergence is determined by the number of non-zero moments of the wavelet function (Sheng, 2000).

<sup>12</sup>Note that in order to guarantee  $C_\psi < \infty$ , it must hold that  $\Psi(0) = 0$ . It is straightforward to show that this restriction reproduces the two necessary conditions of a wavelet function, i.e.,  $\int_{-\infty}^\infty \psi(u) du = 0$  and  $\int_{-\infty}^\infty \psi^2(u) du = 1$ .

<sup>13</sup>This reconstruction property is evident once the wavelet transforms are transformed into the Fourier domain. This condition is directly related to the equivalence relationship in Parseval's equality.

<sup>14</sup>The energy conservation property is directly linked to Parseval's energy relation in Fourier transformation.

where  $|W(\lambda, \tau)|^2$  is the wavelet power spectrum (Grossmann & Morlet, 1984; Mallat, 2009). This property implies that the total variance of the series is preserved. Furthermore, it illustrates that variance (energy) can be decomposed on a scale-by-scale basis. The quantity  $|W(\lambda, \tau)|^2/\lambda^2$  effectively specifies a density function of the local variance of different time and timescales (Percival & Walden, 2000).

### 3.2 Maximal Overlap Discrete Wavelet Transform

The maximal overlap discrete wavelet transform is used below to derive timescale-optimized portfolio strategies (see section 4). The MODWT differs from the CWT in that it performs a (coarser) discretization of the scale parameter  $\lambda$ .<sup>15</sup> In the MODWT, wavelet transforms are calculated only for dyadic timescales  $\lambda = 2^{-j}$ , where  $j$  is a positive integer. As timescale-optimized portfolios are only evaluated for these discrete timescales, the coarser discretization of the MODWT reduces the number of investigated portfolios to a manageable amount. At the same time, it allows constructing portfolios that are optimized over a broader timescale spectrum. Finally, it enables deriving wavelet variance and wavelet covariance measures that closely resemble conventional moment statistics (see section 3.3). In the following,  $\lambda_j = 2^{-j}$  is used to refer to a timescale of a specific scale level  $j$ .

The MODWT is usually implemented by an iterated filter bank. The respective algorithm is referred to as pyramid algorithm and goes back to Mallat (1989).

Let  $\{\tilde{h}_l; l = 0, \dots, L-1\}$  in  $\mathbb{R}^L$  be a high-pass filter (wavelet filter) with even filter width  $L$  that, nonetheless, represents an infinite sequence such that  $\tilde{h}_l = 0$  for  $l < 0$  and  $l > L$ . Three properties characterize this filter: It must sum to zero ( $\sum_{l=0}^{L-1} \tilde{h}_l = 0$ ), have half-unit energy ( $\sum_{l=0}^{L-1} \tilde{h}_l^2 = \frac{1}{2}$ ), and be orthogonal

---

<sup>15</sup>The wavelet function for the so-called maximal overlap discrete wavelet transform (MODWT) can be described by  $\psi_{j,k}(t) = \lambda_c^{-j/2} \psi(t\lambda_c^{-j} - kt_c\lambda_c^{-j})$  where  $\lambda_c = 2$  refers to a dyadic scaling factor,  $t_c = 1$  specifies the time spacing, and  $k$  and  $j$  are integers of discrete dilations and translations. The scaling parameter of the CWT in formula 1 is thus defined as  $\lambda = 2^{-j}$ . The translation parameter from formula 1 is expressed as  $u = kt_c$ . Due to the discreteness of the return series, the translation factor remained the same for both the CWT and the DWT.

to its even shifts ( $\sum_{l=-\infty}^{\infty} \tilde{h}_l \tilde{h}_{l+2n} = 0$ ).<sup>16</sup> This wavelet filter measures deviations from the trend in the data and resembles a differencing operator. It captures high-frequency (low timescale) and attenuates low-frequency (high timescale) characteristics.

Similar to different wavelet functions satisfying the wavelet function conditions in section 3.1, different wavelet filters fulfill these filter conditions. The choice of a wavelet filter should be based on the characteristics of the underlying data. This study uses a Least Asymmetric wavelet filter of length 8, LA(8). The LA(8) filter was shown to match the characteristics of stock returns and has been widely used in earlier studies (Gençay et al., 2010, 2002, 2005).

The low-pass filter (scaling filter)  $\{\tilde{g}_l; l = 0, \dots, L-1\}$  in  $\mathbb{R}^L$  complements the wavelet filter  $\tilde{h}_l$ . In contrast to the wavelet filter, it averages consecutive values and thus can be considered an averaging operator. It is obtained from the quadrature mirror relationship:  $\tilde{g}_l = (-1)^{l+1} \tilde{h}_{L-1-l}$ . This so-called scaling filter allows capturing the remaining low frequency spectrum, which is left out by the wavelet filter.

These two filters enable decomposing the time series into components that relate to timescale  $\lambda_j$ . Let  $\{r_t; t = 0, \dots, N-1\}$  in  $\mathbb{R}^N$  specify a vector of returns of length  $N$ . Then, the wavelet and scaling coefficients at timescale  $\lambda_j$  are obtained from convolution of the filters with the return series in step ( $j = 1$ ) and with the scaling coefficients in the subsequent steps ( $j > 1$ ):

$$\tilde{W}_{j,t} = \sum_{l=0}^{L-1} \tilde{h}_l \tilde{V}_{j-1,t-2^j-1l \bmod N}, \quad \tilde{V}_{j,t} = \sum_{l=0}^{L-1} \tilde{g}_l \tilde{V}_{j-1,t-2^j-1l \bmod N}, \quad (5)$$

for  $t = 0, 1, \dots, N-1$  and  $\tilde{V}_{0,t} \equiv r_t$ . The modulus operator is necessary because the return series is finite. However, this operation results in wavelet and scaling coefficients at the edges being flawed (boundary conditions). This edge-effect needs to be accounted for in the derivation of unbiased wavelet moments (section 3.3).<sup>17</sup>

<sup>16</sup>These filter conditions are directly related to the conditions of the wavelet function in section 3.1.

<sup>17</sup>Note that the modulus operator imposes the signal to be cyclical. In the transition from the end to the start value, this may lead to non-stationarity. To avoid this effect, the time series is reflected at the last observation. This reflection has no effect on the estimated moments, but allows circumventing non-stationarities in the time series.

The wavelet coefficients  $\tilde{W}_{j,t}$  capture the high frequency content of the signal. They are directly related to the wavelet transforms in formula 1. With increasing scale level  $j$ , oscillations of higher period lengths are covered. The relation between scale level  $j$  of the wavelet transforms and the associated frequency band is given by  $[1/2^{(j+1)} < f \leq 1/2^j]$ . For daily data, this implies that scale level 1 wavelet transforms describe oscillations in the signal for periods of 2–4 days (frequency band:  $[1/4, 1/2]$ ) and scale level 2 captures oscillations in the signal for periods of 4–8 days (frequency band:  $[1/8, 1/4]$ ). This study uses a maximum scale level of decomposition of  $J = 7$ . Consequently, the highest scale level 7 covers cycle period lengths of 128–256 days (frequency band:  $[1/256, 1/128]$ ).

The scaling coefficients  $\tilde{V}_{j,t}$  describe the remaining high-timescale dynamics and contain the trend of the original signal. However, the scaling coefficients are not further investigated in this study. They are only used to derive the wavelet coefficients in the iterative algorithm of formula 5.

### 3.3 Wavelet Variance and Covariance

Wavelet and scaling coefficients (transforms) are associated with changes at a particular timescale. Formula 4 shows that this decomposition is energy preserving in the CWT. An analog relation holds for the MODWT such that:

$$\begin{aligned} \|r\|^2 &= \sum_{t=0}^{N-1} r_t^2 = \sum_{j=1}^J \sum_{t=0}^{N-1} \tilde{W}_{j,t}^2 + \sum_{t=0}^{N-1} \tilde{V}_{J,t}^2 \\ &= \sum_{j=1}^J \|\tilde{\mathbf{W}}_j\|^2 + \|\tilde{\mathbf{V}}_J\|^2, \end{aligned} \tag{6}$$

where  $\tilde{\mathbf{W}}_j = \{W_{j,t}; t = 0, \dots, N-1\}$  and  $\tilde{\mathbf{V}}_J = \{V_{J,t}; t = 0, \dots, N-1\}$ . This implies that the variance of the original signal can be decomposed on a scale-by-scale basis. As a result, it is possible to calculate a variance for a specific timescale

$\lambda_j$ . An *unbiased estimator of wavelet variance* for scale  $\lambda_j$  is given by

$$\tilde{v}^2(\lambda_j) = \frac{1}{\tilde{N}_j} \sum_{t=L_j-1}^{N-1} \tilde{W}_{j,t}^2, \quad (7)$$

where  $\tilde{W}_{j,t}^2$  corresponds to the wavelet coefficient of scale level  $j$  for time series  $\{r_t\}$ ,  $\tilde{N}_j \equiv N - L_j + 1$  specifies the coefficients unaffected by the boundary conditions (see section 3.2), and  $L_j \equiv (2^j - 1)(L - 1) + 1$  refers to the length of a filter at scale level  $j$ .

The analogous definition of the *unbiased estimator of wavelet covariance* between the stochastic processes  $r_{p,t}$  and  $r_{q,t}$  is obtained from

$$\tilde{v}_{p,q}(\lambda_j) = \frac{1}{\tilde{N}_j} \sum_{t=L_j-1}^{N-1} \tilde{W}_{p,j,t} \tilde{W}_{q,j,t}, \quad (8)$$

where  $\tilde{W}_{p,j,t}$  and  $\tilde{W}_{q,j,t}$  are the wavelet coefficients of the respective time series  $r_{p,t}$  and  $r_{q,t}$ . These covariances can be compiled in a wavelet covariance matrix  $\Sigma(\lambda_j) = \{\tilde{v}_{p,q}(\lambda_j); p = 1, \dots, M; q = 1, \dots, M\}$  for a particular timescale  $\lambda_j$ , where  $M$  refers to the number of stocks. Timescale-optimized portfolios in section 4 are derived based on these covariance matrices.

### 3.4 Standard Errors and Test Statistics of Wavelet Variance

Comparing the wavelet variances of different scale portfolios necessitates formulating test statistics. These test statistics require adopting distributional assumptions and deriving standard errors of variances.

Let  $\bar{W}_{j,t}$  refer to those wavelet coefficients that are not affected by the boundary conditions in formula 5 (edge-effect). It is assumed that  $\bar{W}_{j,t}$  is a Gaussian stationary process with mean zero and spectral density  $S_j(\cdot)$ .<sup>18</sup> Percival (1983, 1995) showed

<sup>18</sup>Let  $\{r_t\}$  be a second-order discrete stationary process. Further, let  $\{s_\tau; \tau = \dots, -1, 0, 1, \dots\}$

that the estimator for the variance  $\hat{v}^2(\lambda_j)$  of these coefficients is asymptotically normally distributed with mean  $v^2(\lambda_j)$  and large sample variance  $2A_j/\tilde{N}_j$ , where  $A_j$  is defined as  $A_j \equiv \int_{-1/2}^{1/2} S_j^2(f) df$ . This result holds if the spectrum  $S_j(f)$  is greater than zero almost everywhere and if  $A_j$  is square integrable.<sup>19</sup> For large  $\tilde{N}_j$  it thus follows that

$$\frac{\tilde{N}_j^{1/2} (\hat{v}^2(\lambda_j) - v^2(\lambda_j))}{(2A_j)^{1/2}} \sim \mathcal{N}(0, 1). \quad (9)$$

An approximately unbiased estimator for  $A_j$  is given by

$$\hat{A}_j = \frac{(\hat{s}_{j,0}^{(p)})^2}{2} + \sum_{\tau=1}^{\tilde{N}_j-1} (\hat{s}_{j,\tau}^{(p)})^2 = \frac{\hat{v}^4(\lambda_j)}{2} + \sum_{\tau=1}^{\tilde{N}_j-1} (\hat{s}_{j,\tau}^{(p)})^2, \quad (10)$$

where  $\hat{s}_{j,\tau}^{(p)}$  is a biased estimator of the sample autocovariance sequence of  $\bar{W}_{j,t}$ , defined as  $\hat{s}_{j,\tau}^{(p)} \equiv \frac{1}{\tilde{N}_j} \sum_{t=L_j-1}^{N-1-|\tau|} \bar{W}_{j,t} \bar{W}_{j,t+|\tau|}$  for  $0 \leq |\tau| \leq \tilde{N}_j - 1$  and  $\hat{s}_{j,\tau}^{(p)} \equiv 0$  for  $|\tau| \geq \tilde{N}_j$  (Percival, 1995). It is straightforward to derive confidence intervals and test statistics of wavelet variance if  $A_j$  in formula 9 is replaced by the estimator  $\hat{A}_j$ . The test statistics for comparing wavelet variances of two dissimilar timescales  $\hat{v}^2(\lambda_j)$  and  $\hat{v}^2(\lambda_k)$  is then given by:

$$\frac{\hat{v}^2(\lambda_j) - \hat{v}^2(\lambda_k)}{\sqrt{\frac{2\hat{A}_j}{\tilde{N}_j} + \frac{2\hat{A}_k}{\tilde{N}_k}}}, \quad (11)$$

where  $j \neq k$ .

Even if the underlying process is non-Gaussian, the asymptotic distribution of the

---

be the autocovariance sequence, where the autocovariance between the components  $r_t$  and  $r_{t+\tau}$  is given by  $s_\tau = \text{cov}\{r_t, r_{t+\tau}\}$  for all integers  $t$  and  $\tau$ . If this autocovariance sequence is square summable ( $\sum_{\tau=-\infty}^{\infty} s_\tau^2 < \infty$ ), then the spectral density is given by  $S(f) = \sum_{\tau=-\infty}^{\infty} s_\tau e^{-i2\pi f\tau}$  for  $|f| \leq \frac{1}{2}$  (under the assumption of unit time-spacing).

<sup>19</sup>Due to Parseval's theorem and the stationarity of  $\{\bar{W}_{j,t}\}$ , the latter condition is equivalent to implying that the autocovariance sequence of  $\{\bar{W}_{j,t}\}$  dies down fast enough so that it is square summable.

wavelet variance may be known. For a broad class of non-Gaussian and non-linear processes, the estimator  $\hat{v}^2(\lambda_j)$  is asymptotically Gaussian-distributed, with similar properties as listed above (formula 9). However, the spectral density may no longer have the same convenient form (Gençay et al., 2002).

The normal distribution assumption does not prevent the lower confidence limits of wavelet variance from being negative. Therefore, Percival (1995) provided an alternative approach to characterizing the distribution of wavelet variance using the assumption of chi-square distributed wavelet variance. The distribution is adjusted with respect to the degrees of freedom, in order to account for possible correlations in the underlying variables (see Priestley, 1981, p. 466). For this approach, the following approximation can be derived:

$$\frac{\xi \hat{v}^2(\lambda_j)}{v^2(\lambda_j)} \sim \chi_{\xi}^2, \quad (12)$$

where  $\xi$  is known as the equivalent degrees of freedom. It can be estimated by

$$\hat{\xi} = \frac{\tilde{N}_j \hat{v}^4(\lambda_j)}{\hat{A}_j}, \quad (13)$$

The respective *F-test* for equality of variance is readily derived from this definition.

## 4 Wavelet Portfolio Strategy

The adoption of timescale-optimized portfolio strategies centers around the assumption of stock returns exhibiting distinct variance and covariance dynamics at different time horizons. Accordingly, a portfolio can be optimized for dynamics at a specific timescale (risk optimization). In out-of-sample analysis, these portfolios should show smaller risk at the optimized timescale compared to other portfolios (optimized at other timescales or for the untransformed data). If test results confirm this assump-

tion, it would imply that investors with different life-cycle preferences can optimize their portfolio over their respective investment horizon. This section outlines the portfolio formation and risk analysis techniques for constructing and assessing these timescale-optimized portfolio strategies (henceforth scale portfolio strategy).

It is important to note that investors (specifically those with a long-term perspective) are likely to operate at multiple timescales rather than at a single timescale. However, to study the effectiveness of scale portfolio strategies, I initially employ the simplifying assumption of a single-timescale investor. This investor operates at a specific timescale and strives to reduce risk over a specific timescale (fixed investment horizon). The single-timescale assumption is relaxed in section 7 and a multiscale optimization is introduced.

## 4.1 Optimization Method

I use minimum variance portfolio optimization to construct scale portfolios. This choice is made for analytical reasons and because of the implicit restrictions imposed by the properties of wavelet analysis. First, mean-variance optimization is not applicable or at best subjective in nature. All wavelet coefficients have an expected value of zero. Only scaling coefficients have non-zero expected values. Hence, it is not possible to associate a mean with a specific timescale. Second, the minimum variance portfolio optimization method only relies on the covariance matrix estimate. For this reason, it is possible to attribute changes in portfolio allocation solely to variations of the covariance matrix. Finally, minimum variance optimization is less sensitive to estimation errors than other optimization approaches, such as mean-variance optimization (see Chan, Karceski & Lakonishok, 1999; Jagannathan & Ma, 2003). Among others, Michaud (1989), Broadie (1993), Best and Grauer (1991), Green and Hollifield (1992), Britten-Jones (1999), and DeMiguel, Garlappi, Nogales and Uppal (2009) have shown that small errors in estimated mean returns can cause substantial changes in portfolio allocation. These estimation errors transfer directly to errors in portfolio weights. In contrast, covariance matrices can be estimated more precisely (see Merton, 1980).

## 4.2 Weight Restrictions

Minimum variance portfolio optimization is also not immune to estimation errors (see Chan et al., 1999; Jagannathan & Ma, 2003). Small measurement errors in the covariance matrix can lead to unstable portfolio weights with high risk concentrations in only few assets. Several approaches have been suggested to make portfolio weights more stable. One strand of the literature has focused on the deduction of robust covariance matrix estimates using shrinkage (see Ledoit & Wolf, 2004). However, modifying the wavelet covariance structure makes it more difficult to study the effects of scale-based portfolio optimization in isolation. Another strand of the literature has suggested imposing portfolio weighting restrictions using thresholds (see, for example, DeMiguel et al., 2009). This method limits portfolio weights rather than modifying the moments of asset returns. It can thus be considered less intrusive than the first approach. Consequently, I use this weighting methodology below to study the effects of scale-based portfolio optimization.

Weighting restrictions, however, limit the range of portfolio weights and constrain system flexibility (especially if the asset universe is small). As a result, it can be more difficult to identify the superiority of a certain portfolio strategy. For this reason, I use two different weighting restrictions for portfolio formation: The maximum weighting set  $A_{mw}$  and the less restrictive long-only weighting set  $A_{lo}$ . Analysis using the *long-only weighting approach* is more flexible, whereas that using the *maximum weighting approach* guarantees more stable results.

Let  $\alpha \equiv (\alpha_1, \dots, \alpha_M)^T$  specify the portfolio weight vector for  $M$  possible investment opportunities. The admissible set of strategies for the long-only portfolio is then given by  $A_{lo} = \left\{ \alpha \in \mathbb{R}^M \mid \alpha \in [0, 1]^M, \mathbf{1}^T \alpha = 1 \right\}$ . Similarly, the admissible set for the maximum-weight portfolio is defined as  $A_{mw} = \left\{ \alpha \in \mathbb{R}^M \mid \alpha \in [0, b]^M, \mathbf{1}^T \alpha = 1 \right\}$ , where  $b < 1$  is the threshold value. Hence, short selling and borrowing additional capital is restricted.

No-short and no-leverage conditions prevent the emergence of extreme portfolio allocations. For example, some assets might otherwise be shorted extensively to fi-

nance long positions in others.<sup>20</sup> Moreover, short selling is accompanied by high (borrowing) costs and is subject to several restrictions in practice. It is therefore reasonable to exclude short selling and the borrowing of additional capital from the analysis of scale portfolios.

### 4.3 Portfolio Formation

Scale-based minimum variance portfolio formation is conducted in two steps: First, wavelet transformation is applied to past return observations for every stock and every rebalancing date to derive wavelet transforms. Second, these transformed series are used to obtain the covariance matrix for a specific timescale and estimation window. Thus, the covariance matrix estimate derives from the comovement between stock price processes at a certain time and timescale.

Let  $\Sigma_t(\lambda_j)$  specify the wavelet covariance matrix at timescale  $\lambda_j$  and time  $t$  (with a rolling estimation window ending at  $t$ ).<sup>21</sup> Then, the scale-based minimum variance portfolio weights for the respective rebalancing date  $t + 1$  are obtained by

$$\alpha_{t+1}(\lambda_j)^* = \underset{\alpha_{t+1} \in \mathbf{A}}{\operatorname{argmin}} \alpha_{t+1}(\lambda_j)^T \Sigma_t(\lambda_j) \alpha_{t+1}(\lambda_j), \quad (14)$$

where  $\mathbf{A}$  refers either to the long-only ( $A_{lo}$ ) or to the maximum-weight portfolio weight set ( $A_{mw}$ ). The scale-based covariance matrix is re-estimated at every rebalancing date. Finally, the collection of portfolio weights for all rebalancing periods is used to derive out-of-sample portfolio returns.

---

<sup>20</sup>The presence of estimation errors can lead to high-risk concentrations.

<sup>21</sup>Different lag-days (2, 3, and 4) have been tested for implementation of the portfolio weights. Results remained largely consistent.

#### 4.4 Performance & Risk Assessment

To test the main hypothesis of this study, portfolio performance and risk need to be assessed and compared. Analyzing risk measures reveals whether, for a given timescale, the lowest out-of-sample variance is indeed observed for the portfolio of matching scale. Portfolio performance and risk can be assessed either with conventional time domain methods or with wavelet decomposition.

Conventional time domain methods refer to analysis using simple portfolio returns. Most studies have restricted themselves to the analysis of one uniform sampling interval (e.g., daily, monthly, or yearly data). However, several different sampling intervals have to be considered to assess the effectiveness of scale-optimized portfolios in reducing variance at a particular timescale. Below, I thus analyze portfolio statistics for multiple timescales. This is achieved by compounding discrete return observations so that they conform with the targeted periodicity (temporal aggregation). For example, daily return data is compounded for each month so that a monthly return series is obtained.

While conventional time domain analysis is intuitive and appealing in practice, it is less accurate than wavelet decomposition. Higher frequency perturbations (e.g., daily variations) can superimpose longer cycle variations in returns. This might mask the effectiveness of optimization by distorting the volatility estimates of conventional sampling methods (using daily/weekly/monthly data).

Wavelet decomposition is a more advanced filtering method and allows for more effectively isolating optimization effects for particular timescales. Moreover, the wavelet transformation of portfolio returns provides a scale-by-scale decomposition of variance that precisely matches the portfolio timescale. This is the case because it is the same method as the one used initially to construct the scale portfolios. Therefore, decomposing portfolio returns using wavelet transformation allows more accurately testing the hypothesis of scale-based variance minimization.

## 5 Data & Data Analysis

This study uses a subset of stocks temporarily or permanently listed on the Dow Jones Industrial Average (DJIA) during the period from January 1, 1969 to December 30, 2016. Stocks not covering the full sample period were excluded, reducing the number of securities from a total of 48 to 23. All stock prices are quoted in US dollars (no exchange rate effects) and retrieved at the same closing time (no asynchronous trading effects). Data for daily stock prices and the DJIA constituents list were obtained from the Center of Research and Security Prices (CRSP) and Compustat.

The condition of full data coverage for the period under study imposes restrictions on analyzing and interpreting the data sample: First, a selection bias is present given that some stocks are considered part of the sample before they were historically included in the DJIA. Stocks that are subject to index inclusion are typically characterized by sustained growth. Second, stocks that are no longer quoted at an exchange are not considered in the sample (survivorship bias). Third, assets with a shorter track history, and having decisive and unique price dynamics (substituting or complementing certain industries), might be excluded from the sample (e.g., Microsoft or Apple). Due to these biases, certain stock price characteristics might not be uncovered that would otherwise have been detected in a non-restricted sample.

While the loss of information due to excluding stocks cannot be offset, the impact of the biases on the analysis and comparison of portfolio strategies is minimized. This is achieved by comparing wavelet portfolio results with benchmark portfolios comprising the same subsample of stocks. In contrast, comparison with the DJIA index itself is limited: The index consists of different time-dependent constituents and uses a different portfolio construction mechanism.

The restriction to a relatively small sample size of only 23 stocks is a compromise between the stability of a covariance matrix estimate and the benefits of diversification.<sup>22</sup>

---

<sup>22</sup>Evans and Archer (1968) conjectured that approximately 15 stocks are required to obtain a sufficiently diversified portfolio. More recent studies conclude that the number of stocks necessary to eliminate non-systematic risk has risen in the recent past and more stocks are required (see Elton & Gruber, Martin, 1977; Newbould & Poon, 1993; Campbell, Lettau, Malkiel & Xu, 2001).

**Table III.1:** Descriptive statistics of daily log-returns, covering the period January 1, 1969 to December 30, 2016.

	Ticker	Ann. mean	Ann. std.	Skew.	Kurt.	Min.	Max.	SR	JB-stat. (1,000)	LB(20)	LB <sup>2</sup> (20)	
	3M	MMM	7.27%	23.50%	-0.70	21.01	-0.30	0.11	0.31	156.7***	66.5***	701.3***
	Alcoa	AA.3	2.13%	35.60%	-0.23	11.35	-0.28	0.21	0.06	33.3***	108.9***	7,593.5***
	Altria Group	MO	14.92%	26.79%	-0.49	14.54	-0.26	0.15	0.56	64.3***	74.7***	588.3***
	Boeing	BA	9.98%	33.19%	0.10	7.81	-0.19	0.21	0.30	11.0***	29.2***	1,385.6***
	Caterpillar	CAT	6.82%	30.51%	-0.29	9.45	-0.24	0.14	0.22	20.0***	70.4***	2,032.2***
	Chevron	CVX	7.30%	26.42%	-0.04	9.26	-0.18	0.19	0.28	18.6***	78.8***	6,141.5***
	Coca-Cola	KO	8.91%	24.64%	-0.39	18.43	-0.28	0.18	0.36	113.6***	69.9***	2,632.9***
	Du Pont	DD	4.81%	26.81%	-0.18	8.34	-0.20	0.11	0.18	13.8***	21.3	3,303.8***
	Exxon Mobil	XOM	7.87%	23.36%	-0.38	19.42	-0.27	0.16	0.34	128.0***	187.2***	3,758.0***
	General Electric	GE	7.61%	26.91%	-0.07	10.85	-0.19	0.18	0.28	29.4***	72.4***	7,740.6***
	Goodyear	GT	1.63%	39.45%	-0.45	12.28	-0.34	0.18	0.04	39.8***	57.1***	4,739.0***
	HP	HPQ	8.62%	36.86%	-0.26	9.60	-0.23	0.19	0.23	21.2***	59.7***	830.4***
	Honeywell Intl.	HON	7.12%	31.63%	-0.29	21.35	-0.35	0.27	0.23	158.7***	35.8**	1,785.0***
	IBM	IBM	5.03%	26.07%	-0.27	14.80	-0.26	0.12	0.19	68.4***	30.6*	919.3***
	Intl. Paper	IP	3.54%	31.98%	-0.28	14.68	-0.31	0.20	0.11	65.5***	84.5***	5,522.8***
	JPMorgan Chase	JPM	4.77%	34.99%	-0.09	17.33	-0.32	0.22	0.14	96.7***	24.6	4,995.3***
	Johnson & Johnson	JNJ	10.76%	23.69%	-0.24	10.31	-0.20	0.12	0.45	25.8***	101.0***	2,066.5***
	McDonald's	MCD	13.58%	28.17%	-0.21	10.60	-0.22	0.15	0.48	27.7***	83.0***	2,707.4***
	Merck	MRK	8.52%	26.22%	-0.70	18.71	-0.31	0.12	0.33	119.8***	74.7***	299.6***
	Pfizer	PFE	9.09%	28.20%	-0.16	6.87	-0.19	0.11	0.32	7.1***	81.1***	2,477.9***
	Procter & Gamble	PG	9.02%	23.16%	-2.25	68.16	-0.38	0.20	0.39	2,043.1***	81.1***	386.7***
	United Tech.	UTX	8.39%	28.53%	-0.52	16.40	-0.33	0.13	0.29	85.9***	51.2***	681.2***
	Walt Disney	DIS	12.37%	32.53%	-0.57	16.22	-0.34	0.17	0.38	84.8***	48.8***	1,294.3***

*Notes:* Mean and standard deviation are annualized; JB-stat. shows the Jarque-Bera test statistics for the null hypothesis of normality in sample return distribution. The statistics is reported in 1,000s; LB(20) and LB<sup>2</sup>(20) refer to the Ljung-Box test of autocorrelation of 20-order lags for changes and squared changes in stock returns; SR defines the Sharpe ratio.

\*, \*\*, and \*\*\* indicate rejection of the null hypothesis of the test statistics at the 1%, 5%, and 10% level of significance.

Table III.1 presents the descriptive statistics of the 23 stocks' annualized daily log-returns for the period from 1969 to 2016. Annualized mean returns lie between 1.63% for Goodyear and 14.92% for Altria Group. While Procter & Gamble displays the most volatile returns among all stocks (in terms of standard deviation), it also exhibits the highest negative return, highest kurtosis, and lowest skewness. Its mean

return of 9.02% is approximately equal to Pfizer's mean return. However, Pfizer shows a higher standard deviation of 28.20% and less pronounced tails.

The return series are thus characterized by differing distributional characteristics. However, all return series exhibit leptokurtic behavior and all series (except for Boeing) are negatively skewed. The Jarque-Bera (JB) test statistic strongly rejects the hypothesis of normally distributed returns for all 23 stocks. To test for the absence of autocorrelation, the Ljung-Box ( $LB$ ) statistic was employed with a 20-order lag. Results are significant for almost all return series. Only Du Pont and JPMorgan Chase show no proof of serial correlation, while the test statistic is rejected for Boeing and IBM at the 10% level of significance. The test to determine the absence of serial correlation in squared returns ( $LB^2$ ) is rejected for all return series.

Table III.2 displays the correlation matrix of the 23 stocks for the period from March 29, 1986 to December 30, 2016.<sup>23</sup> The lower triangular matrix represents daily return correlations, while coefficients in the upper triangular matrix are calculated based on monthly returns.

In general, daily return correlations experienced weaker linear dependence than the same correlations estimated for monthly return data. Absolute correlations seem to increase with higher sampling intervals.<sup>24</sup>

To more quantitatively estimate the deviations between correlations, differences between daily and monthly sampling interval correlations were tested using Fisher  $z$ -transformation. The analysis revealed significantly lower correlation coefficients for 19 stocks at a monthly sampling period with 5% level of significance, whereas 13 stocks showed a higher correlation. In total, about 13% of the correlations displayed significant differences (together with significant Jennrich test). Thus, Table III.2 indicates a potential presence of scale-dependent correlation structures.

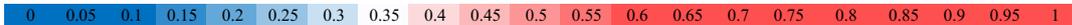
---

<sup>23</sup>The period of analysis is reduced since the analysis of dependency structures in the following sections only adopt this shorter time period. Data before April 1986 is only used for added precision in the calculation of the wavelet coefficients.

<sup>24</sup>Note that the color code in Table III.2 is very narrowly specified with respect to the size of the correlations. The difference between daily and monthly correlations may therefore appear more pronounced.

**Table III.2:** Analysis of unconditional correlations covering the period March 29, 1986 to December 30, 2016. Lower triangular matrix: daily log-return correlations; upper triangular matrix: monthly log-return correlations.

Daily and monthly sampling correlations																							
	MMM	AA.3	MO	BA	CAT	CVX	KO	DD	XOM	GE	GT	HPQ	HON	IBM	IP	JPM	JNJ	MCD	MRK	PFE	PG	UTX	DIS
MMM	1.00	0.51	0.29	0.47	0.53	0.37	0.34	0.59	0.40	0.54	0.47	0.35	0.43	0.30	0.57	0.32	0.38	0.31	0.27	0.32	0.38	0.55	0.47
AA.3	0.44	1.00	0.17	0.38	0.66	0.38	0.22	0.57	0.33	0.48	0.48	0.47	0.48	0.40	0.61	0.41	0.25	0.25	0.20	0.20	0.16	0.51	0.42
MO	0.32	0.24	1.00	0.17	0.12	0.18	0.36	0.22	0.20	0.26	0.24	0.21	0.27	0.21	0.20	0.15	0.36	0.30	0.30	0.29	0.35	0.24	0.27
BA	0.40	0.39	0.24	1.00	0.40	0.34	0.36	0.50	0.34	0.48	0.41	0.26	0.46	0.17	0.48	0.29	0.33	0.35	0.33	0.31	0.28	0.60	0.47
CAT	0.47	0.52	0.26	0.40	1.00	0.47	0.24	0.57	0.38	0.49	0.48	0.42	0.50	0.36	0.56	0.42	0.28	0.35	0.25	0.28	0.22	0.54	0.45
CVX	0.42	0.45	0.27	0.34	0.40	1.00	0.23	0.47	0.75	0.42	0.32	0.27	0.42	0.28	0.39	0.33	0.30	0.34	0.19	0.31	0.17	0.40	0.32
KO	0.43	0.28	0.37	0.33	0.31	0.34	1.00	0.31	0.30	0.37	0.21	0.19	0.30	0.11	0.25	0.28	0.52	0.43	0.47	0.44	0.52	0.31	0.38
DD	0.55	0.52	0.29	0.41	0.51	0.45	0.39	1.00	0.46	0.60	0.51	0.35	0.49	0.33	0.67	0.45	0.39	0.37	0.32	0.32	0.29	0.57	0.45
XOM	0.47	0.43	0.30	0.36	0.40	0.75	0.43	0.47	1.00	0.45	0.28	0.30	0.40	0.28	0.37	0.29	0.34	0.36	0.25	0.32	0.29	0.37	0.31
GE	0.53	0.46	0.31	0.44	0.48	0.41	0.44	0.52	0.45	1.00	0.50	0.40	0.49	0.37	0.57	0.53	0.43	0.43	0.37	0.45	0.36	0.58	0.52
GT	0.42	0.44	0.21	0.36	0.44	0.33	0.28	0.44	0.33	0.45	1.00	0.39	0.50	0.29	0.51	0.41	0.25	0.33	0.21	0.23	0.22	0.48	0.42
HPQ	0.34	0.33	0.21	0.30	0.34	0.28	0.27	0.34	0.29	0.40	0.31	1.00	0.37	0.47	0.34	0.41	0.22	0.26	0.21	0.24	0.19	0.43	0.49
HON	0.48	0.43	0.26	0.45	0.46	0.40	0.36	0.48	0.43	0.51	0.42	0.35	1.00	0.33	0.52	0.42	0.24	0.43	0.28	0.28	0.33	0.58	0.44
IBM	0.38	0.34	0.24	0.31	0.36	0.30	0.30	0.37	0.35	0.45	0.31	0.47	0.38	1.00	0.30	0.34	0.22	0.27	0.13	0.18	0.07	0.30	0.35
IP	0.50	0.52	0.26	0.37	0.51	0.40	0.32	0.55	0.40	0.50	0.45	0.31	0.45	0.33	1.00	0.39	0.33	0.32	0.32	0.34	0.31	0.53	0.45
JPM	0.43	0.41	0.25	0.37	0.42	0.36	0.32	0.45	0.36	0.56	0.43	0.37	0.42	0.40	0.43	1.00	0.23	0.36	0.21	0.33	0.18	0.42	0.47
JNJ	0.41	0.27	0.34	0.31	0.29	0.35	0.47	0.36	0.41	0.44	0.25	0.26	0.33	0.31	0.30	0.32	1.00	0.41	0.55	0.58	0.48	0.38	0.37
MCD	0.36	0.28	0.30	0.32	0.31	0.29	0.40	0.34	0.33	0.40	0.27	0.26	0.33	0.30	0.30	0.33	0.36	1.00	0.37	0.39	0.44	0.40	0.48
MRK	0.35	0.29	0.31	0.31	0.28	0.32	0.39	0.35	0.36	0.40	0.26	0.25	0.30	0.28	0.30	0.34	0.53	0.31	1.00	0.56	0.38	0.32	0.33
PFE	0.37	0.30	0.32	0.32	0.31	0.34	0.40	0.35	0.38	0.43	0.28	0.26	0.32	0.31	0.31	0.34	0.53	0.32	0.56	1.00	0.32	0.42	0.43
PG	0.44	0.26	0.34	0.30	0.30	0.32	0.51	0.39	0.38	0.42	0.27	0.24	0.35	0.29	0.32	0.31	0.47	0.39	0.39	0.39	1.00	0.31	0.30
UTX	0.49	0.45	0.26	0.49	0.48	0.39	0.35	0.48	0.40	0.51	0.40	0.35	0.52	0.36	0.44	0.43	0.34	0.34	0.32	0.34	0.34	1.00	0.51
DIS	0.44	0.41	0.30	0.39	0.41	0.37	0.38	0.43	0.40	0.51	0.38	0.38	0.44	0.39	0.42	0.45	0.37	0.36	0.34	0.36	0.35	0.43	1.00



Correlation

Notes: The color code of the correlation ranges from blue (low correlation) to red (high correlation).

## 6 Empirical Analysis

This section explores the timescale dynamics of different stocks and their implications for portfolio optimization. Section 6.1 uses coherence analysis to study how stock prices co-move over time and across different time horizons. Section 6.2 investigates the timescale characteristics of variance in more detail. The variance of stock returns and the dependency structure of these stocks are key factors for portfolio optimization. Therefore, the analysis of these factors provides important insights for interpreting results of scale-based portfolio optimization.

Finally, section 6.3 examines the main hypothesis of this study — namely that portfolio variance can be minimized for a specific timescale with scale-based portfolio strategies — and presents the results for these timescale-optimized portfolios.

### 6.1 Coherence Analysis

Coherence analysis transforms two time series into the two-dimensional time-frequency plane in order to depict their dependency structures across different periodicities (see Appendix A for mathematical details). Wavelet magnitude-squared coherence is a bivariate linear dependency measure in the time-timescale-domain. The measure is bounded between 0 and 1 such that  $0 \leq R_{p,q}^2(\lambda, \tau) \leq 1$ . High wavelet magnitude-squared coherence provides evidence for strong dependence and low wavelet magnitude squared coherence indicates weak dependence. Hence, wavelet magnitude-squared coherence resembles a squared correlation coefficient in linear regression (Vacha & Barunik, 2012). It allows capturing time- and timescale-varying features of correlation between two stocks (Rua & Nunes, 2009).<sup>25</sup>

---

<sup>25</sup>Coherence analysis has previously been used to examine the timescale dependency structures in and across commodity (Aguar-Conraria, Rodrigues & Soares, 2014; Akoum, Graham, Kivihaho, Nikkinen & Omran, 2012; Barunik & Vacha, 2009; Bekiros, Nguyen, Uddin & Sjö, 2016; Madaleno & Pinho, 2014; Vacha & Barunik, 2012), foreign exchange (Andries, Ihnatov & Tiwari, 2014; Reboredo & Rivera-Castro, 2013), and stock markets (Aloui & Hkiri, 2014; Graham, Kiviahho & Nikkinen, 2012; Graham, Kiviahho, Nikkinen & Omran, 2013; Loh, 2013; Madaleno & Pinho, 2012; Rua & Nunes, 2009, 2012). Furthermore, it has been employed to study macroeconomic interrelationships (Aguar-Conraria, Azevedo & Soares, 2008; Aguiar-Conraria & Joana Soares, 2011; Gallegati et al., 2014). However, it was rarely used for the investigation of comovements between individual stocks.

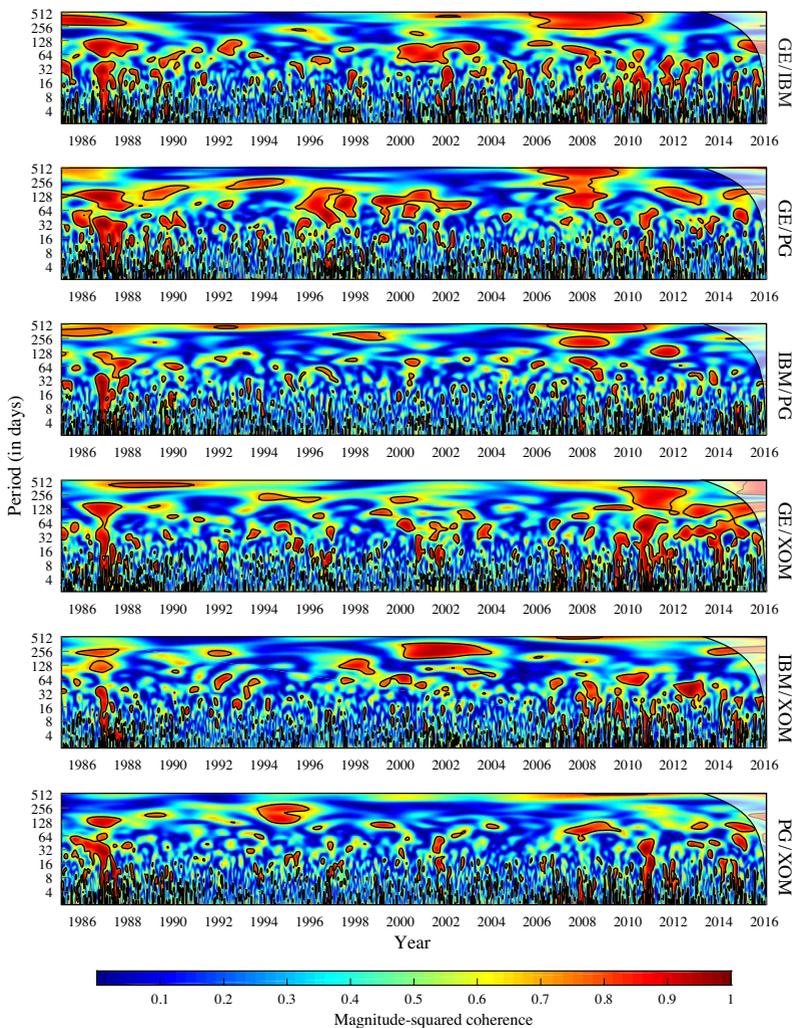
Figure III.1 displays contour graphs for the wavelet magnitude-squared coherence of four selected stocks: Exxon Mobil (XOM), General Electric (GE), IBM, and Procter & Gamble (PG). These stocks have been selected due to their heterogeneous dependency structures and diverse sectoral affiliation (Energy, Industrials, Information Technology, and Consumer Staples). A Morlet wavelet is used for decomposition.<sup>26</sup> The color code for power (strength of relationship) in Figure III.1 ranges from blue, indicating low magnitude-squared coherence, to red, specifying high magnitude-squared coherence. The vertical axis depicts timescales in days.<sup>27</sup> The so-called "cone of influence", which encircles the region affected by boundary effects, is depicted by the thick solid black line (see Appendix A). Wavelet coefficients were calculated based on the complete dataset. As a result, no cone of influence exists in the left part of the figure.

The coherence analysis in Figure III.1 draws a more detailed picture of the dependencies between stocks compared to the correlation analysis in Table III.2. Dependency structures vary profoundly for different asset pairs and frequently change over time and across timescales. For example, while IBM and General Electric exhibit high coherence between 2008 and 2010 for periods around 256 days, this comovement is not observed for IBM and Exxon Mobil. On the other hand, all stock pairs exhibit high coherence across all timescales for the period after the stock market crash of October 19, 1987 (Black Monday).

---

<sup>26</sup>The Morlet wavelet is defined as  $\psi(t) = \pi^{-(1/4)} \left( e^{i\kappa_0 t} - e^{-\kappa_0^2/2} \right) e^{-t^2/2}$ . The parameter  $\kappa_0$  defines the number of oscillations within the Gaussian envelope. It directly relates to the frequency-/time-localization of the wavelet function. Here,  $\kappa_0$  is set to 6.

<sup>27</sup>Note that the scale  $\lambda_j$  is a unitless standardized measure. A physical meaningful unit is only obtained if the sampling interval  $t_c$  is considered such that the physical scale  $\lambda_j t_c$  is obtained. In this analysis  $t_c$  corresponds to one day and thus scales correspond to an actual physical unit of the same magnitude.



**Figure III.1:** Wavelet squared coherence between four selected stocks: Exxon Mobil (XOM), General Electric (GE), International Business Machines (IBM), and Procter & Gamble (PG), covering the period March 29, 1986 to December 30, 2016. Notes: The horizontal axis shows time and the vertical axis illustrates the timescale (period) in days; the color code of the wavelet magnitude-squared coherence ranges from blue (low power) to red (high power); the cone of influence (COI), which indicates the region affected by edge-effects, is displayed in lighter shade; black contour lines designate the 5% significance level obtained from Monte Carlo simulations using randomized surrogate series; a Morlet wavelet is used for decomposition.

The coherence contour graph of General Electric and IBM further reveals that the global financial crisis of 2008 severely impacted correlations at periods of 256 days and higher. The increase in comovement at these periodicities lasted for several years. At low timescales (high frequencies), correlations show high variability. However, specific periods of high and persistent correlations are evident for the highest frequencies at the end of 2008 and in the first quarter of 2010. Interestingly, these increased comovements do not transfer to coherence dynamics in the period interval between 64 and 256 days.

Another interesting observation is that the burst of the dot-com bubble in 2001 had a higher relative impact on correlations at timescales corresponding to period lengths of 64 days than at other timescales. In contrast, the global financial crisis of 2008 was more pronounced for oscillations with greater period lengths. The 2008 crisis even had a higher impact on long-term correlation dynamics than the dot-com bubble of 2001. In the years after 2012, high magnitude-squared coherence is again evident at frequencies in the interval of 32–64 days. The coherence graph for General Electric and Procter & Gamble shows similar patterns to those of General Electric and IBM. However, for the period after 2012, a high concentration of power of the General Electric/Procter & Gamble stock pair is located at lower frequencies with cycle lengths of approximately 32 days.

Analyzing the interaction between Procter & Gamble and Exxon Mobil reveals a low degree of comovement for frequency bands with cycle lengths greater than 16 days. Thus, Exxon Mobil provides high benefits of diversification at these higher timescales.<sup>28</sup> One of the sole exceptions occurred from the end of 2011 to mid-2012, with the advent of the European sovereign debt crisis and the downgrading of America's credit risk. High magnitude-squared coherence can be observed within the frequency interval for cycles from 2 to 64 days (power at higher timescales is still relatively small). Exxon Mobil also displays unique characteristics in relation to other assets. For example, the coherence analysis for IBM/Exxon Mobil after the dot-com bubble of 2001 indicates a strong comovement at periodicities of 256 days

---

<sup>28</sup>Note that coherence is an absolute measure and thus "combines" positive and negative correlations. Negative correlations would provide even better diversification. However, analysis of wavelet correlations revealed almost no negative correlations between assets (at all timescales).

and only modest comovement for timescales between 16 to 64 days. At short periodicities of 2–16 days, power is locally concentrated and switches from high to low regimes. In contrast, the global financial crisis mainly influenced comovements at low timescales (2–4 days) and timescales between 16 to 64 days. However, the crisis did not increase the power at lower frequencies (higher timescales). This contrasts sharply with the observations for General Electric and Exxon Mobil. Their correlations surged for all timescales in the aftermath of the financial crisis of 2008 and remained at high levels across all timescales for the entire period thereafter.

To summarize, analyzing the relations between this selective subset of stocks has revealed that stocks are characterized by a varying degree of interaction over time and across frequencies.

## 6.2 Analysis of Variance

Estimation of variance is central to portfolio optimization and risk management decisions. As outlined (section 3.3), variance can be decomposed into different timescale components. Below, an equally-weighted index is used as a proxy to study timescale characteristics of variance. Comparing the wavelet decomposition of single stocks and an equally-weighted index revealed a similar pattern of energy distribution. Table III.3 demonstrates the MODWT variances for the equally-weighted portfolio of the subsample of DJIA stocks. Further, it shows the variance contribution of each scale level, i.e., the portion of overall variance, which is explained by dynamics at a certain timescale. Finally, Table III.3 displays the same metrics for non-crisis and crisis market states to examine how the variance structure changes during these periods.<sup>29</sup>

Table III.3 indicates that most of the variance of daily data is explained by short timescale dynamics, corresponding to bi-daily, weekly, and bi-weekly periodicities

---

<sup>29</sup>Market phases are classified qualitatively and encompass the following five economic crises: Black Monday (Oct. 1987–Dec. 1988), Asian crisis (Jul. 1997–Dec. 1998), dot-com bubble (Jan. 2001–Dec. 2002), global financial crisis (Jul. 2007–Jun. 2009), and European debt crisis/American credit risk downgrading (Jun. 2001–May 2012). All remaining time periods are categorized as non-crisis states. Note that crisis periods cover a relatively long period to guarantee filtration of longer-term cycle dynamics.

(see columns 3 and 4). These scale variances contribute approximately 89% to total variance (scale level 1 to 3). In contrast, longer timescales only account for a relatively small portion of total variance. For example, the variance at monthly periodicities (scale level 4) is only about 10% of the size of the variance at bi-daily periodicities (scale level 1).

**Table III.3:** *MODWT variance decomposition ( $J = 7$ ) of equally-weighted, daily portfolio returns for the full period and for the non-crisis and crisis periods (March 29, 1986–December 30, 2016).*

Scale Level	Period interval	Full period		Non-Crisis		Crisis	
		Variance cont.	Ann. var. $\times 10^4$	Variance cont.	Ann. var. $\times 10^4$	Variance cont.	Ann. var. $\times 10^4$
J1	2 - 4	51.53%	0.688	48.95%	0.369	53.06%	1.432
J2	4 - 8	25.65%	0.343	25.48%	0.192	25.96%	0.701
J3	8 - 16	12.19%	0.163	13.38%	0.101	11.34%	0.306
J4	16 - 32	5.34%	0.071	6.36%	0.048	4.70%	0.127
J5	32 - 64	2.74%	0.037	3.14%	0.024	2.56%	0.069
J6	64 - 128	1.45%	0.019	1.51%	0.011	1.40%	0.038
J7	128 - 256	1.10%	0.015	1.18%	0.009	0.99%	0.027

*Notes:* The equally-weighted portfolio is composed of the 23 DJIA stocks (sample);  $J_1, J_2, \dots, J_7$  refer to scale levels 1–7; period intervals designate the periods (in days) corresponding to a certain scale level; scale variance is annualized and multiplied by a factor of  $10^4$ ; variance contribution is obtained by dividing the variance at a particular scale by the total variance.

The classification of time periods into different market phases shows that overall variance of the index significantly increases during crisis states. However, energy distribution between individual scale levels remains relatively similar in the two market phases. While a modest increase in variance contribution can be observed for lower timescales (scale levels 1 and 2) during crisis periods, variance contribution at higher timescales (scale levels 5, 6, and 7) does not markedly change. Increases in variance contribution for bi-daily and weekly periodicities are predominantly offset by decreases for bi-monthly and monthly periodicities. One possible interpretation is that monthly investors react more frequently to shocks and adopt a shorter invest-

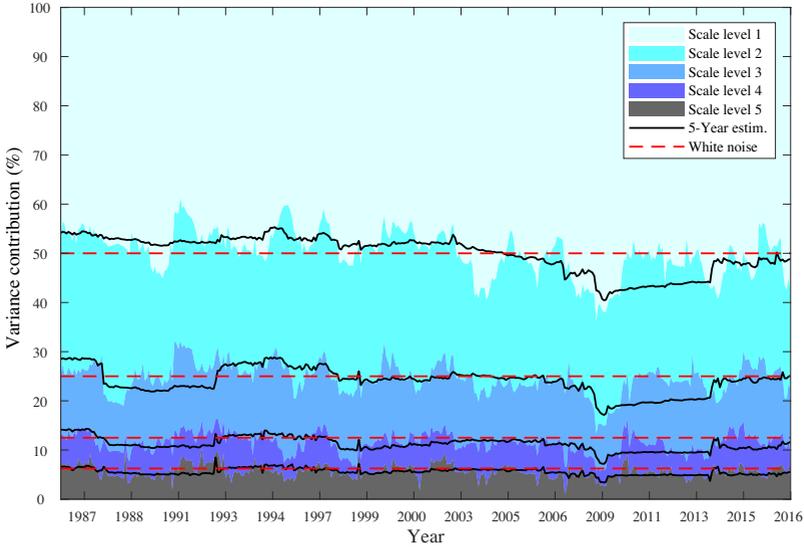
ment horizon. In contrast, longer-term investors do not show the same sensitivity and remain with their initial investment timescale.<sup>30</sup>

Figure III.2 depicts the 1-year (colored areas) and the 5-year (black lines) rolling window estimations of the sample variance contribution of daily returns for the period between January 1986 and December 2016. Red-dashed lines illustrate the variance contribution of a decomposed white-noise process. In analogy to Chaudhuri and Lo (2016), these bounds are constructed by generating random permutations of the order of the return series. Wavelet decomposition is then applied to the resulting serially uncorrelated data. The white-noise boundaries (red lines) are obtained by averaging the variance contributions of 10,000 of these simulated white-noise processes. Because variance contribution decreases exponentially with increasing timescale, Figure III.2 only shows variance decomposition up to scale level 5. Variance contribution of higher timescales (scale level 5 and higher) is subsumed in the last level of decomposition.

Figure III.2 illustrates that the variance contribution of both the 1-year and 5-year rolling window deviates from the expectations of a serially uncorrelated process during certain periods. While the variance contribution of scale level 1 is smaller at the beginning of analysis, it gradually approaches the white-noise band by 1999. Scale levels 2 and 3 absorb most of the energy during this period. Strong serial correlations of weekly and bi-weekly returns during this period are a potential reason for higher fluctuations at these scale levels. These findings are in line with test results for 20-order lag Ljung-Box statistics for serial correlation of individual stocks (see Table III.1).

---

<sup>30</sup>Effectively, agents operate at several timescales at once; there is no unique stream of investment behavior. Observations in Table III.3 thus might be even better explained by an information theoretic approach. Adapting this view, short-term news (shocks) exhibit a higher impact during crisis states, prompting adjustments in the beliefs and investment timescale of investors.



**Figure III.2:** MODWT-decomposed variance contribution of 1-year and 5-year rolling window equally-weighted daily portfolio returns, covering the period March 29, 1986 to December 30, 2016. Notes: The equally-weighted portfolio is composed of the 23 DJIA stocks (sample); colored regions indicate variance contribution using a 1-year rolling window for estimation, whereas black lines depict variance contribution using a 5-year rolling window for estimation; for both estimation windows, variance contribution is shown for scale level 1 (period of 2–4 days) to 5 (period of 32–64 days). Variance contribution of higher timescales is subsumed in scale level 5; red dashed lines specify white-noise boundaries of scale variance contribution obtained from simulation of 10,000 independent Gaussian-distributed returns.

For the interval between 2003 and 2006, variance contribution closely resembles the energy distribution of a white-noise process. However, the advent of the global financial crisis of 2007–08 increased the energy contribution of lower timescales (to almost 94% for the first three scales combined). This observation is in line with higher serial correlations of stock returns during market downturns.

Similar to the findings in Table III.3, Figure III.2 demonstrates that lower timescales account for most of the variation in daily stock returns (specifically during crisis periods). Higher scales rarely show variance contributions exceeding the bounds of a white-noise process. However, these observations do not render lower frequency pro-

cesses irrelevant for three reasons: First, variance contribution is a relative measure. Hence, congruency of scale energy between the white-noise process and the index at high scale levels does not necessarily imply the absence of long-term fluctuations. Second, fluctuations at low timescales might smooth or cancel out over the course of longer investment periods. Therefore, these oscillations might not be perceived as a risk by long-term investors, for whom long-term fluctuations are more relevant. For example, a pension fund assesses the risk of intra-day fluctuations differently than an inter-day trader. Finally, correlation is a key factor in all portfolio formation. The interaction of scale variance with scale correlation is thus an important factor to be considered (covariance). As shown in the previous section, interdependence between stocks shows high variation for different timescales.

To summarize, in accordance with findings in previous literature, the highest timescales have been found to contribute most to the variation of daily returns. However, this does not necessarily render higher timescales irrelevant, specifically if fluctuations smooth over the course of a longer investment period.

### 6.3 Wavelet-Based Portfolio Optimization

Given the previous results for coherency and timescale variance, I now investigate the main hypothesis that it is possible to construct portfolios that minimize volatility for a targeted timescale. This section tests this hypothesis by introducing wavelet-based scale portfolios and by comparing their relative performances. A daily rebalancing period is used for the rolling recalibration of the portfolio weights. The first rebalancing date, and thus the beginning of the out-of-sample period (the following day), is chosen as February 28, 1986. This allows for a relatively large estimation (calibration) period, which goes back to January 1, 1969 and continuously extends with each rebalancing day.<sup>31</sup>

---

<sup>31</sup>Note that this long estimation period from January 1, 1969 to March 1, 1986 is required for the evaluation of the dynamic estimation-window-strategy in section 6.4.

### Portfolio Construction

The general structure and testing procedure for the portfolio strategies relies upon the specifications outlined in section 4. In accordance with these specifications, the detailed construction of wavelet-based scale portfolios is as follows: i) At each rebalancing date, the return series is subsampled to only contain past data, i.e., data up until the specific rebalancing date. ii) Wavelet filtering is then applied to the reduced dataset, returning a time series for every scale level. Return series of single stocks are decomposed into seven scale components corresponding to periodicities of 2–4 (scale level 1), 4–8 (scale level 2), 8–16 (scale level 3), 16–32 (scale level 4), 32–64 (scale level 5), 64–128 (scale level 6), and 128–256 (scale level 7) days. For example, the transformed time series at scale level 4 (corresponding to cycle lengths of 16–32 days) approximately covers monthly stock price periodicities. iii) Subsequently, the transformed series are circularly shifted to align the wavelet coefficients with the effective sequence of events. iv) After shifting, the wavelet transforms are subsampled to obtain the 1,250 data points with closest proximity to the rebalancing date (corresponding to an estimation window of approximately five years). v) Finally, the covariance matrix for a particular scale is derived from this subsampled series. Portfolio weights are obtained using the scale covariance matrix as input to the minimum variance optimization problem in formula 14. Optimal portfolio weights are derived for each individual timescale. This results in portfolio weights for seven different scale portfolios at each rebalancing date (seven levels of decomposition).

It is important to highlight and discuss some critical aspects of this procedure in more detail:

First, I use a daily rebalancing period because this periodicity corresponds to the sampling rate of the original data. This periodicity lies outside the frequency band for which portfolio dynamics are studied here (Nyquist frequency). Therefore, daily rebalancing (at least partially) prevents the dynamics of the rebalancing procedure from being mixed with portfolio dynamics. Nevertheless, a bias will remain due to the effects of aliasing<sup>32</sup>.

---

<sup>32</sup>In signal processing, aliasing describes the effect of misidentifying a signal frequency because of

Second, limiting the raw data sample prior to applying wavelet filtering prevents future observations entering wavelet coefficients, i.e., observations made after the rebalancing date (filtration). At the same time, no observations prior to the rebalancing date have been excluded. Hence, the entire set of past observations (information) is available for deriving the wavelet coefficients.

Third, applying circular shifting is necessary because a Least Asymmetric (LA) wavelet filter is used for wavelet decomposition. The LA filter is a non-zero phase filter. As a result, coefficients are not properly aligned with the effective sequence of events after transformation. Circular shifting ensures that coefficients, and thus covariance estimates, are in sync. Note that the circular shift is only necessary because a subset rather than the full wavelet series is used to estimate the covariance matrix.

Finally, the presented procedure improves estimates of the scale covariance matrix. Instead of reducing the time series before wavelet transformation, it applies wavelet transformation first and then shortens the time series to match the estimation window (used to estimate the covariance matrix). This procedure allows incorporating additional information to derive wavelet coefficients. As a result, the coefficients at the beginning of the estimation window are no longer affected by boundary effects (see section 3.2). This procedure improves the covariance estimation compared to an approach where the two steps are applied in reverse order (i.e., first subsampling data and subsequently using wavelet transformation). Despite this improvement in estimation, coefficients near a rebalancing date remain subject to boundary effects. This bias has to be taken into account in the unbiased covariance estimator of formula 8.

Note that an even better covariance estimate can be achieved if all coefficients are retained for estimating the covariance matrix. However, to compare the models and to account for non-stationarities, the size of the estimation window is fixed to 1,250 data points. While the presented technique for deriving scale portfolios is rather complex, it ensures that no future information enters the estimation of scale-based covariance matrices (out-of-sample restriction). At the same time, it minimizes the covariance matrix estimation bias.

---

an insufficient sampling rate. As a result, higher frequency components of the signal cannot be captured accurately (leading to distortion or errors).

### Comparison of Scale-Decomposed Variance

The returns of the scale portfolios can be derived directly from the portfolio weights obtained in the previous steps. Let  $P_{Jj}$  refer to the portfolio optimized for dynamics at timescale  $\lambda_j$ . For example, portfolio  $P_{J1}$  minimizes variance at timescale  $\lambda_1$ . Similarly, portfolio  $P_{J2}$  is optimized for stock price fluctuations at timescale  $\lambda_2$ . In total, seven portfolios up to timescale  $\lambda_7$  are constructed, i.e.  $\{P_{Jj}; j = 1, \dots, 7\}$ . The returns of these individual scale portfolios need to be decomposed by wavelet decomposition once again (section 4.4). Using the same decomposition level  $J = 7$  as before yields seven scale variances  $\{\tilde{v}_{P_{Jj}}^2(\lambda_k); j = 1, \dots, 7; k = 1, \dots, 7\}$  for each individual scale portfolio  $\{P_{Jj}; j = 1, \dots, 7\}$ . These variances enable testing whether scale portfolios are effective in minimizing variance (volatility) at a targeted timescale. Comparing scale variances among portfolios now allows assessing the validity of the main hypothesis of this study, i.e., that scale portfolio optimization reduces variance for the respective timescale. Following this hypothesis, I expect to observe the lowest variances (volatilities) where the scale of the portfolio and the scale of the variance measure coincide, i.e. where  $j = k$ .

Table III.4 presents scale-decomposed variances for the individual scale portfolios. Note that scale variance  $\tilde{v}_{P_{Jj}}^2(\lambda_k)$  is normalized by multiplication with the scale-level-dependent factor  $2^{j-1}$ .<sup>33</sup> Multiplication with the scaling factor has no implications for interpretation but allows for better visual comparison of results.

The description "Wavelet Covariance Estimator Portfolios" groups the seven scale portfolio strategies  $\{P_{Jj}; j = 1, \dots, 7\}$ . The subscript of each portfolio corresponds to the scale level return dynamics for which the portfolio is optimized. For example, scale portfolio  $P_{J1}$  minimizes variance at scale level 1.

---

<sup>33</sup>This multiplication factor derives from the scaling properties of a Brownian motion.

**Table III.4:** MODWT-decomposed variance of conventional time domain and scale-based (wavelet-based) minimum variance portfolios, covering the period March 29, 1986 to December 30, 2016 (daily rebalancing).

Scale Level	Period interval	Conv. covariance estim. portfolios			Wavelet covariance estimator portfolios						
		$P_{0.5y}$	$P_{1y}$	$P_{5y}$	$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$
J1	2 - 4	43.92 (0.028)	44.29 (0.028)	48.76 (0.031)	48.80 (0.031)	50.60 (0.032)	52.08 (0.033)	54.78 (0.035)	57.96 (0.038)	60.57 (0.039)	64.45 (0.042)
J2	4 - 8	44.17 (0.024)	44.73 (0.024)	49.79 (0.027)	51.90 (0.028)	50.89 (0.028)	51.02 (0.027)	51.67 (0.028)	55.46 (0.030)	56.73 (0.030)	59.54 (0.032)
J3	8 - 16	42.79 (0.025)	42.78 (0.024)	44.64 (0.025)	48.07 (0.027)	45.13 (0.025)	44.11* (0.024)	44.34* (0.025)	47.22 (0.026)	50.10 (0.028)	55.38 (0.031)
J4	16 - 32	39.60 (0.022)	39.41 (0.022)	39.33 (0.022)	43.37 (0.024)	40.73 (0.022)	37.68** (0.021)	37.15** (0.020)	38.40* (0.021)	38.77* (0.020)	44.05 (0.024)
J5	32 - 64	41.58 (0.023)	42.15 (0.024)	40.66 (0.022)	43.93 (0.024)	42.28 (0.023)	39.45 (0.022)	37.21* (0.020)	39.95 (0.022)	44.08 (0.025)	48.83 (0.028)
J6	64 - 128	43.68 (0.023)	43.71 (0.023)	38.83 (0.020)	41.76 (0.021)	38.58 (0.020)	37.38 (0.020)	37.91 (0.020)	37.40 (0.019)	47.35 (0.024)	45.62 (0.026)
J7	128 - 256	45.35 (0.028)	46.05 (0.028)	40.61 (0.025)	45.42 (0.027)	37.59 (0.023)	37.77 (0.024)	34.56 (0.019)	36.75 (0.020)	42.04 (0.023)	49.69 (0.029)

Notes: Conventional time domain portfolios  $P_{0.5y}$ ,  $P_{1y}$ , and  $P_{5y}$  are constructed using half-yearly, yearly, and 5-yearly estimation windows, respectively; portfolio  $P_{J1}, P_{J2}, \dots, P_{J7}$  refer to the scale portfolios optimized for scale variance at scale level 1–7 and are calculated using an estimation window of approximately 5 years; each row specifies the scale variance of these portfolios at a specific scale level, where  $J1, J2, \dots, J7$  refer to scale levels 1–7. Scale variance is multiplied by the scale-level-dependent factor  $2^{j-1} \times 10^6$ ; period intervals designate the periods (in days) corresponding to a certain scale level; grey-shaded areas indicate lowest scale variances; the null hypothesis of the test statistics states that  $\tilde{v}_{P_{J1}}^2(\lambda_k) = \tilde{v}_{P_{Jj}}^2(\lambda_k)$  for  $j \neq k$  (two-sided test). However, significances are only reported where  $\tilde{v}_{P_{J1}}^2(\lambda_k) > \tilde{v}_{P_{Jj}}^2(\lambda_k)$ ; standard errors are multiplied by the scale-level-dependent factor  $\sqrt{2^{(j-1)} \times 10^9}$  and depicted in parentheses.

\*, \*\*, and \*\*\* indicate rejection of the null hypothesis of the test statistics at the 1%, 5%, and 10% level of significance, respectively.

Table III.4 compares variances of these scale portfolios at each individual scale level (row-by-row). The grey boxes indicate the lowest variances among the group of scale portfolios at a particular timescale. In addition to this qualitative comparison, I further test for variance equality. The variance of portfolio  $P_{J1}$  serves as a benchmark against which the variances of all other portfolios are tested. Thus, the null hypothesis of the test statistics states that — at timescale  $\lambda_k$  — the variance of portfolio  $P_{J1}$  optimized for timescale  $\lambda_1$  is identical to the variance of another portfolio

$P_{J_j}$  optimized for timescale  $\lambda_j$  (i.e.  $\tilde{v}_{P_{J_1}}^2(\lambda_k) = \tilde{v}_{P_{J_j}}^2(\lambda_k)$ , where  $j \neq 1$ ). The alternative hypothesis states that the variances between these portfolios are different (i.e.  $\tilde{v}_{P_{J_1}}^2(\lambda_k) \neq \tilde{v}_{P_{J_j}}^2(\lambda_k)$ , where  $j \neq 1$ ). The choice of comparing all scale portfolios to  $P_{J_1}$  is motivated by the fact that the highest variance contribution stems from dynamics at the lowest timescale (see section 6.2). If stock returns do not experience different dynamics across timescales, there should be no differences among the variances of the different portfolios.

The test for variance equality uses the Gaussian standard errors introduced in section 3.4. In general, tests for which the Gaussian distributional assumption was employed showed less significant results than the corresponding tests using the chi-square distributional assumption. Thus, Table III.4 only displays the more stringent (robust) results using the test in formula 11.

Table III.4 illustrates that portfolio  $P_{J_1}$  exhibits the lowest variance among all wavelet covariance estimator portfolios at scale level 1 (row 1). These results are consistent with the findings of Berger (2016). Comparison of the variance against the remaining portfolios shows that this difference is highly significant. For the sake of readability, Table III.4 only presents test results where variances are found to be significantly lower than the benchmark portfolio.<sup>34</sup> Hence, even though portfolio  $P_{J_1}$  exhibits significantly lower variances than the remaining portfolios, these results are not reported in Table III.4.

Note that portfolio  $P_{J_1}$  does not dominate other strategies at higher scale levels. At these timescales, portfolios optimized with respect to lower scale dynamics prevail. For example, at scale level 4, variances of portfolio  $P_{J_3}$  and  $P_{J_4}$  are significantly lower (at 5% level of significance) compared to the variance of portfolio  $P_{J_1}$ . In general, the lowest variance is observed for portfolios optimizing fluctuations within the same frequency band over which variance is measured. This is where the scale of the portfolio and the scale of analysis coincide (diagonal elements). This finding demonstrates that it is possible to construct portfolios that minimize volatility at a certain timescale and thus supports the main hypothesis of this study.

---

<sup>34</sup>If results were reported for both sides of the test, nearly the entire upper triangular part of Table III.4 would show significances. Listing all these significances would make it more difficult to interpret the results. Consequently, the test in Table III.4 is a two-sided test, where only significant deviations at left tail of the benchmark distribution are reported.

The fact that there is no dominating strategy in Table III.4 substantiates the assumption that individual stocks follow timescale (cross-)dynamics. If stock behavior or interaction was invariant over different timescales, either all portfolios would show the same variances (due to equal portfolio weights), or the variance ranking of portfolio strategies would remain constant over different scale levels (and thus produce the aforementioned dominant strategy).

However, the results must be considered with caution. First, only some variances exhibit significant deviations and only with a relatively modest confidence level of 5%. Nevertheless, these results are relatively strong given that portfolios were optimized with respect to the same small set of underlying stocks. Second, results are less conclusive for observations beyond scale level 5. At these frequencies, the lowest variance is no longer observed for portfolios of approximately matching timescale. This ambiguity might be due to two reasons:

First, covariance matrix estimates at higher scales are less representative due to fewer observations of long-cycle periods. For example, scale level 6 covers quarterly to half-yearly frequencies. Given that the estimation window for calibrating the covariance matrix comprises approximately five years of data, there might be too few observations for extracting long frequency dynamics.

Second, stock prices contain less information at higher scales and show weak timescale-dynamics within these frequency bands. As a result, portfolio optimization might produce less reliable strategies. This would agree with the above variance analysis (section 6.2), which found that most variance contribution can be attributed to the first four timescales. Higher timescale-dynamics only contribute little to overall variance.<sup>35</sup> Compared to shorter-term dynamics, long-term fluctuations appear to be less pronounced and more random.

While the first explanation claims that estimation errors are responsible for the inconclusive results at higher timescales, the second explanation implies that the energy of timescale dynamics is limited to shorter frequencies. Section 6.4 shows that extending the estimation window reduces variances at higher scale levels. This indicates

---

<sup>35</sup>Even for normalized scale variance (with respect to its scale level) the relative variance contribution is generally subpar.

that the inconclusive results for higher timescales are more likely caused by higher estimation errors.

The results in Table III.4 thus support the assumption that stock prices exhibit time and frequency characteristics and are consistent with the above findings (sections 6.1 and 6.2). Remarkably, wavelet covariance estimator portfolios exhibit significantly different scale variances even though they all employ the same optimization methodology and the same underlying dataset (only the timescale of the data varies). The fact that scale variance differs significantly between the portfolios despite using the same data strongly indicates the existence of frequency dynamics in stock prices.

Accordingly, a stock should not be assessed solely based on its time dimension. Frequency characteristics should be given equal consideration. A stock can have properties that are desired at one timescale but undesired at others. Investors' preferences define how these different timescale features are related. Consequently, the investment horizon should be considered as an additional dimension in an investor's utility function. This study thus comes to a similar conclusion as Gressis et al. (1976).

Besides scale portfolios, Table III.4 shows the scale-by-scale decomposition of portfolio variance using conventional covariance estimates for optimization ("Conventional Covariance Estimator Portfolios"). These portfolios rely on the covariance estimates obtained from undecomposed stock returns. Subscripts in portfolio notation refer to the size of the estimation window used to derive the covariance matrix. Correspondingly, the covariance matrices studied in Table III.4 are based on half-yearly ( $P_{0.5y}$ ), yearly ( $P_{1y}$ ), and five-yearly ( $P_{5y}$ ) estimation intervals.

At lower timescales, conventional estimator portfolios outperform wavelet covariance estimator portfolios in terms of risk optimization. Among all tested portfolios, these portfolios exhibit the lowest variance. But even for higher scale levels, conventional portfolios perform reasonably well. Variance lies somewhere between the values of short- and long-term scale portfolios.

These low variances can be explained by the structure of conventional portfolios. A portfolio optimization strategy using conventional covariance estimators is tantamount to an optimization method minimizing variance across all timescales. The

dynamics of each timescale are combined and effectively weighted by their contribution to overall variance. This method can therefore be interpreted as a *variance-contribution-weighted portfolio strategy*. This helps to explain why conventional covariance estimator portfolios perform reasonably well across several scale levels.

Within-group comparison of conventional covariance estimator portfolios reveals that the size of the estimation window impacts portfolio variance at different scale levels. The smallest scale variances are achieved for portfolios using short estimation windows at low scale levels and long estimation windows at high scale levels. A widening of the estimation window seems to allow capturing longer cycle periods. The reason for observing decreasing portfolio variance at higher timescales for portfolios with longer estimation intervals can most likely be attributed to the (implicit) filtering incurred by the extension of the estimation window. Long-term effects are included (attenuated) with longer estimation windows while non-stationarities are smoothed. Therefore, a longer estimation window helps to construct more stable covariance estimates and to include long-term dynamics. This explains why portfolios with short estimation intervals outperform portfolios with long estimation intervals at low timescales, whereas portfolios with longer estimation intervals perform better at higher timescales.

It is reasonable to assume that results for scale portfolios might be similarly affected by the specified estimation window. To exclude the possibility of falsely attributing the effects of variance reduction to the effects of scale optimization (rather than the calibration of the estimation window), several different sizes of estimation windows were also tested for the scale portfolios. However, findings remained unchanged irrespective of the estimation window (see section 6.4). Lowest scale-variances were generally observed for portfolios optimized for dynamics of matching timescale. Consequently, the estimation window has no effect on the general relationship between scale variance and the ranking of portfolio optimization strategies. The size of the estimation window can thus be excluded as the main driving factor of timescale variance minimization.

### **Adjustment of Sampling Intervals**

Wavelet decomposition of scale portfolio returns is a powerful technique for analyzing the impact of scale portfolio optimization. However, analysis of scale portfolio returns with wavelet decomposition also shows some deficiencies.

One caveat of the analysis in Table III.4 is that portfolio strategies are decomposed with the same filtering technique as that used to construct scale portfolios. Note that portfolio strategies were tested out-of-sample and no data-generating processes were imposed by wavelet decomposition. However, it might be argued that the observed scale variance structures emerge due to the lacking separation between the analysis and the construction method (model testing bias). In addition, the analysis in Table III.4 is also complex and the concept of wavelet variance is difficult to interpret. An alternative method of multiscale analysis is to study portfolio returns using different sampling intervals (e.g., daily, weekly, or monthly return intervals). The choice of sampling interval highlights fluctuations at a particular timescale (with a certain contamination). This approach helps to substantiate the findings obtained from wavelet decomposition in Table III.4.

There are three main benefits of reviewing the results in Table III.4 with this technique:

First, the underlying data is not filtered or transformed. Only the sampling interval over which returns are compiled is changed. Hence, the analysis is independent of the frequency decomposition used to construct the portfolio strategies. If findings for the sampling interval technique are consistent with results from wavelet analysis, the presence of a model bias can be refuted.

Second, the effects of scale portfolio optimization can be studied for aggregated frequency dynamics. This is relevant because variance reductions at a particular timescale are negligible if they are superimposed by other frequency dynamics. Analyzing scale portfolio returns with different sampling intervals reveals whether possible optimization effects are strong enough to also capitalize on reducing sampling interval volatility.

Third, the method is more intuitive. Researchers and practitioners alike are familiar with measuring variance over different sampling intervals in the time domain. For

example, in risk management, variance and other risk measures are often considered over multiple horizons. Therefore, the method provides a more intuitive representation of risk at different timescales. Moreover, it is a heuristic approach to analyzing the effects of scale portfolio optimization.

Despite these useful properties, the sampling interval approach also has some major deficiencies (which is why wavelet decomposition is used in the first place). In general, it is a less powerful technique than wavelet decomposition as it does not allow for a similarly seamless time-timescale decomposition. It also adds some fundamental flaws to the analysis (see Gençay, Selçuk & Whitcher, 2001b):

First, an increasing sampling interval significantly reduces the number of available observations. Hence, information for analyzing higher timescales is lost. Second, the estimated variance critically depends on the specific sampling date used. A changing sampling date may alter the results of analysis (non-shift invariance). Third, observations at a particular sampling frequency may be blended by periodicities outside the targeted frequency interval (aliasing).

In light of these advantages and disadvantages, the results in Table III.4 will subsequently be reviewed using the time domain method (temporal aggregation).

Table III.5 presents descriptive statistics of scale portfolios using varying data sampling intervals of raw stock returns (daily/weekly/monthly/quarterly/yearly data). The analysis shows that scale portfolio volatility varies with respect to the sampling interval employed. For a given sampling interval, the lowest volatility is generally observed for the portfolio for which the scale approximately matches the sampling interval. For example, for daily data, the lowest volatility is reported for the scale portfolio optimized with respect to processes of scale level 1 (portfolio  $P_{J1}$ ). This scale level corresponds to periodicities of two to four days. On the other hand, lowest standard deviation at a monthly sampling interval is observed for portfolios optimized with respect to monthly scale dynamics (portfolio  $P_{J4}$ ).



**Table III.5:** (continued)

	Conv. covariance estim. portfolios			Wavelet covariance estimator portfolios						
	$P_{0.5y}$	$P_{1y}$	$P_{5y}$	$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$
<b>Panel D: Quarterly sampling interval</b>										
Mean (%)	10.13	10.29	11.41	11.77	10.79	10.20	10.85	10.77	10.27	10.98
Volatility (%)	14.75	14.77	14.82	15.35	14.56	14.53	14.40	14.35	14.44	16.19
Sharpe ratio	0.69	0.70	0.77	0.77	0.74	0.70	0.75	0.75	0.71	0.68
Skewness	-0.88	-0.88	-0.58	-0.65	-0.51	-0.37	-0.50	-0.67	-0.79	-0.62
Kurtosis	5.84	6.26	4.67	4.78	4.75	4.18	4.39	4.52	5.26	7.38
Obs.	126	126	126	126	126	126	126	126	126	126
<b>Panel E: Yearly sampling interval</b>										
Mean (%)	9.95	10.18	11.19	11.50	10.57	9.90	10.68	10.57	10.03	10.65
Volatility (%)	14.24	14.84	13.30	13.64	13.25	12.60	13.11	13.21	12.32	13.11
Sharpe ratio	0.70	0.69	0.84	0.84	0.80	0.79	0.81	0.80	0.81	0.81
Skewness	-0.29	0.24	-0.37	-0.27	-0.37	-0.43	-0.45	0.10	-0.94	0.01
Kurtosis	3.66	3.13	3.12	2.97	3.12	3.71	3.25	2.82	4.20	3.41
Obs.	31	31	31	31	31	31	31	31	31	31

Notes: Conventional time domain portfolios  $P_{0.5y}$ ,  $P_{1y}$ , and  $P_{5y}$  are constructed using half-yearly, yearly, and 5-yearly estimation windows, respectively; portfolios  $P_{J1}, P_{J2}, \dots, P_{J7}$  refer to the scale portfolios optimized for scale variance at scale level 1–7 and are calculated using an estimation window of approximately 5 years; mean and volatility are annualized.

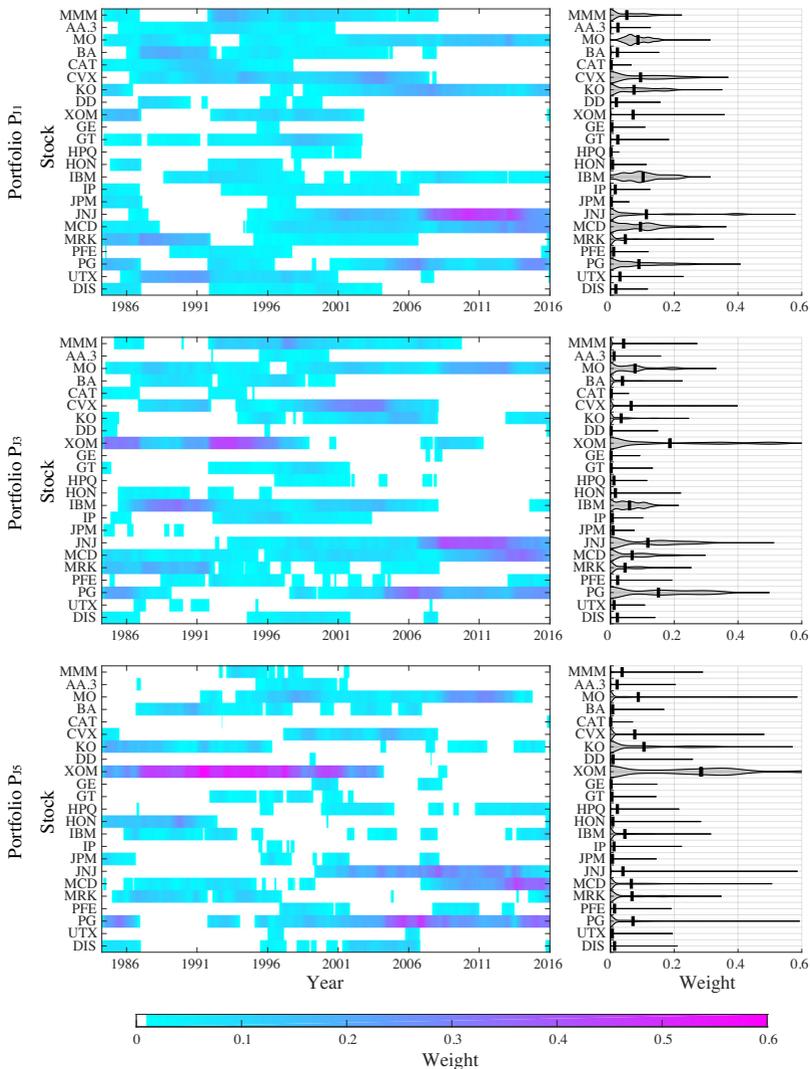
In correspondence with previous findings, scale portfolio strategies mainly minimize variance at the timescales for which portfolios are optimized. However, similar to analysis using wavelet transformation in Table III.4, results become less conclusive for higher sampling intervals such as quarterly or yearly estimations (approximately scale levels 5/6 and 7). Again, this could be due to a lower number of observations or because longer-term timescale dynamics contain less information. In general, results therefore support the previous findings from wavelet analysis (Table III.4).

Further, Table III.5 shows that portfolio  $P_{J_1}$  demonstrates the highest Sharpe ratios compared to the other scale portfolio strategies at all sampling intervals. It could be argued that no benefit is gained from scale-based investment strategies because the high volatility of portfolio  $P_{J_1}$  for lower sampling rates might be compensated for by a return premium.

However, this interpretation ignores four important aspects: First, minimum variance portfolios are optimized with respect to variance. Therefore, variance (volatility) should be the main consideration when assessing the effectiveness of a portfolio strategy. Second, the estimation of mean returns is accompanied by higher measurement errors. The Sharpe ratio might thus be biased. Third, the return premium might be due to lower skewness and higher kurtosis. Both characteristics are observed for portfolio  $P_{J_1}$  at higher sampling frequencies. Finally, deviations in Sharpe ratio are insignificant. Therefore, findings for the Sharpe ratio do not contradict the effectiveness of scale-based minimum variance portfolio optimization.

### **Analysis of Portfolio Weights**

Next, I study the weights of stocks in the different scale portfolios. Figure III.3 demonstrates the portfolio composition of portfolio strategies  $P_{J_1}$ ,  $P_{J_3}$ , and  $P_{J_5}$ . The left-hand side of the figure shows the portfolio weights of the different strategies over the observation period. The right-hand side presents the distribution of portfolio weights over the entire period. This distribution is approximated by a kernel density estimation.



**Figure III.3:** Portfolio weights of scale-based (wavelet-based) minimum variance portfolio strategies optimized for scale level 1 ( $P_{J_1}$ ), scale level 3 ( $P_{J_3}$ ), and scale level 5 ( $P_{J_5}$ ), covering the period March 29, 1986 to December 30, 2016 (daily rebalancing). Notes: Left panel: portfolio weights over time; the color code of portfolio weights ranges from white (small weight) to magenta (large weight); right panel: the portfolio weights distribution approximated by kernel density; vertical lines indicate mean weights.

Composition between the three portfolios differs significantly at the beginning of the observation period. While portfolio  $P_{J_1}$  shows a relatively diverse investment in different stocks, portfolio  $P_{J_5}$  is highly concentrated in only few stocks. For example, in the period from 1990 to 2002, portfolio  $P_{J_5}$  shows a particularly high exposure to Exxon Mobil with almost 60% of the portfolio weight being allocated to that stock. In contrast, Portfolio  $P_{J_1}$  displays no similar concentration in any stock. Portfolio  $P_{J_3}$  lies somewhere in between these two portfolios. While portfolio  $P_{J_3}$  exhibits lower exposure to Exxon Mobil than scale portfolio  $P_{J_5}$ , it has a significantly higher exposure than portfolio  $P_{J_1}$ . This indicates that the benefit of diversification contributed by Exxon Mobil gradually increases with the investment horizon.

Interestingly, the composition of the three portfolios becomes more similar at the end of the observation period. This suggests that the differences in stock price processes across different timescales are less pronounced in the second part of the sample period.

Another considerable observation is that  $P_{J_5}$  displays less stable portfolio weights than its lower-scale counterparts. Indeed, analysis of average turnover (not shown) revealed a substantial increase in turnover from portfolio  $P_{J_1}$  (1.39%) to portfolio  $P_{J_5}$  (4.78%). These findings contradict general intuition, that portfolios focusing on long-term cycle periods are more stable. The increase in average turnover is most likely due to a lower signal-to-noise ratio for higher timescales. As a result of higher noise contamination, the covariance matrix of the high-scale portfolio is likely to be less stable than its lower timescale counterparts.

In order to ensure the consistency of results, I tested two different approaches to increasing the stability of the covariance matrix for high-scale portfolios: *soft thresholding* and *extension of the estimation interval*.

*Soft thresholding* smooths wavelet coefficients and sets lower wavelet coefficients to zero see Gençay et al. (2002). This method is comparable to applying shrinkage to the covariance matrix. In contrast, for the *extension of the estimation interval*, a larger estimation window is used to derive the covariance matrix at higher timescales (see section 6.4). Both approaches resulted in more stable covariances and smaller turnover ratios. Average turnover for the portfolio  $P_{J_5}$  with *soft thresholding* and

*estimation interval extension* reduced to 2.94% and 3.68%, respectively.<sup>36</sup> These findings support the hypothesis that the relative instability of portfolio weights is due to higher measurement errors at higher timescales. It is important to view results in light of the larger amount of noise that is contained in the signal at higher timescales. Covariance at higher timescales should thus not be considered without proper stabilization procedures or sufficiently large estimation intervals.

### **Crisis and Non-Crisis Periods**

Sections 6.1 and 6.2 have shown that stocks are characterized by varying time and frequency characteristics during crisis and non-crisis periods. Therefore, Table III.6 separates the observation period and presents the variance of the different portfolios during crisis and non-crisis market states.<sup>37</sup> Again, tests statistics indicate whether a certain scale portfolio exhibits significantly lower volatility compared to the benchmark portfolio  $P_{J1}$ . An  $F$ -test is applied to test for equality of variances.

The sampling volatility structure of wavelet covariance estimator portfolios for both states are consistent with observations in Table III.5. Portfolios whose scale closely resembles the frequency of the sampling interval exhibit comparably low volatility (diagonal structure). While differences in volatility are significant for observations in non-crisis periods (Panel B), results for crisis periods (Panel A) are less conclusive. However, the insignificance in results for crisis periods may be due to the smaller number of observations.

---

<sup>36</sup>Note that the latter method is generally prioritized since it is considered a less intrusive method.

<sup>37</sup>The same time intervals as in section 6.2 are thereby used for the division of the time series into crisis and non-crisis periods.

**Table III.6:** Annualized volatility of conventional time domain and scale-based (wavelet-based) minimum variance portfolios for crisis and non-crisis periods, covering the period March 29, 1986 to December 30, 2016 (daily rebalancing).

Sampling interval	Obs.	Conv. covariance estim. portfolios (%)			Wavelet covariance estimator portfolios (%)						
		$P_{0.5y}$	$P_{1y}$	$P_{3y}$	$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$
<b>Panel A: Crisis</b>											
Daily	2,431	19.87	20.03	20.88	20.97	21.36	21.58	21.95	22.23	22.68	23.52
Bidaily	1,215	20.18	20.43	21.30	21.39	21.86	21.87	22.05	22.22	22.31	22.34
Weekly	486	19.20	19.26	19.26	19.79	19.60	19.22	19.78	20.18	21.15	21.56
Biweekly	243	18.15	18.53	17.57	18.33	17.69	17.34	17.20	17.56	18.46	20.50
Monthly	121	18.42	18.49	17.50	18.30	17.63	17.34	17.17	16.93	18.53	20.60
Bimonthly	59	17.73	18.23	16.52	17.04	16.37	16.01	16.21	15.94	17.69	19.67
Quarterly	39	19.18	19.37	17.33	18.44	16.88	16.74	16.58	15.47	18.75	19.48
Half-yearly	19	16.44	18.21	14.95	15.57	15.30	15.36	14.56	12.42	15.44	18.22
Yearly	9	16.95	18.88	12.14	13.55	11.46	11.85	12.83	10.78	13.16	10.28
<b>Panel B: Non-Crisis</b>											
Daily	5,835	11.78	11.72	12.16	12.52	12.18**	12.14**	12.30	12.96	13.29	13.69
Bidaily	2,917	11.87	11.78	12.33	12.87	12.34**	12.11***	12.06***	12.81	12.98	13.52
Weekly	1,167	11.54	11.40	11.67	12.30	11.71*	11.45**	11.22***	11.62*	11.97	12.85
Biweekly	583	11.61	11.20	11.36	11.95	11.38	11.06*	10.75**	11.27	11.50	12.25
Monthly	291	11.63	11.20	11.16	11.79	11.20	10.98	10.62*	10.97	11.72	12.17
Bimonthly	142	11.98	11.77	11.56	11.97	11.56	11.73	11.10	11.37	11.78	12.27
Quarterly	94	11.09	10.84	11.54	12.29	11.34	11.01	10.40	10.76	10.14*	10.78
Half-yearly	46	12.36	12.14	11.17	12.28	11.33	10.62	10.36	10.65	9.10	10.32
Yearly	23	12.42	13.31	13.73	15.01	13.56	12.36	11.22	11.53	9.87*	10.95

Notes: Conventional time domain portfolios  $P_{0.5y}$ ,  $P_{1y}$ , and  $P_{3y}$  are constructed using half-yearly, yearly, and 5-yearly estimation windows, respectively; portfolios  $P_{J1}$ ,  $P_{J2}$ , ...,  $P_{J7}$  refer to the scale portfolios optimized for scale variance at scale level 1–7 and are calculated using an estimation window of approximately 5 years; grey-shaded areas indicate lowest volatilities; the null hypothesis of the test statistics states that the variance of portfolio  $P_{J1}$  is equal to the variance of portfolio  $P_{Jk}$  for  $k = 2, \dots, 7$  (two-sided  $F$ -test). However, significances are only reported where the volatility (variance) of a portfolio  $\{P_{Jk}; k = 2, \dots, 7\}$  is lower compared to the volatility (variance) of  $P_{J1}$ .

\*, \*\*, and \*\*\* indicate rejection of the null hypothesis of the test statistics at the 1%, 5%, and 10% level of significance, respectively.

The volatility structure in panel B of Table III.6 appears to be slightly shifted in favor of higher scale portfolios (right shift). For daily, bi-daily, and weekly periodicities, the lowest volatility is observed for portfolios with higher timescale than the main hypothesis of this study has suggested (for portfolios  $P_{J3}$ ,  $P_{J4}$ , and  $P_{J5}$ , respectively). Higher-scale portfolios seem to compete with lower-scale strategies in non-crisis states. This shift might be explained by the fact that long-term trends become more important during calm market periods. As a result, variance reductions at

higher timescales could surpass reductions at lower timescales. However, the analysis of wavelet variance (Table III.9 in Appendix B) shows no similar shift in the diagonal volatility structure. Thus, the shift is more likely due to measurement inaccuracies. Nevertheless, we cannot completely rule out that the reduction in variance is caused by an increased relevance of long-term dynamics. Panel B in Table III.6 further indicates that the diagonal volatility structure extends to higher sampling intervals (e.g., the lowest volatility is observed for portfolio  $P_{J6}$  at a yearly sampling rate with 5% level of significance). However, when the same analysis is conducted using wavelet decomposition, findings are still only significant for low timescales (Table III.9 in Appendix B). For example, the scale level 5 portfolio exhibited no significantly lower variance at the corresponding scale level. Consequently, Table III.6 needs to be considered with caution for higher timescales. As for the analysis of the full period, the effects of scale minimization cannot be confirmed with a reasonable degree of scientific certainty at higher timescales. Nevertheless, the effectiveness of scale optimization at lower scales is further substantiated by both methods of analysis (wavelet decomposition and conventional sampling interval).

Again, the discrepancy between analysis using wavelet decomposition and the sampling interval technique can be explained by the fact that wavelets are more accurate filters than conventional time domain methods. Aliasing and the reduction of the sample dilute the results of the sampling interval technique. However, the general structure is consistent with previous results (Tables III.4 and III.5). The lowest volatilities lie approximately on the diagonal. Furthermore, this diagonal structure seems to prevail irrespective of the market period. However, results for the crisis period are not significant.

### **Summary and Implications of Findings**

This section shows that scale portfolio optimization effectively reduces variance at the corresponding timescale. While it seems trivial at first that portfolio variance can be minimized for short-, mid-, or long-term processes, these findings have far-reaching consequences:

First, they further support the assumption that stock price processes exhibit some form of scale-dependency. Stock prices are likely to be driven by information heterogeneity, by investment heterogeneity, or by both.

Second, performance evaluation and the assessment of portfolio results are timescale-dependent. Similarly, the diversification potential of an individual stock critically depends on the investment perspective and on the evaluation approach employed by an investor. A stock having undesirable characteristics for a portfolio at shorter frequencies might provide high diversification benefits at longer frequencies and vice versa.

Therefore, investment decisions should consider the timescales over which performance is measured. A pension fund for which performance is assessed over a long-term perspective should factor long-term processes into its investment decision-making process (the effects of scale optimization are only established up to monthly frequencies with significance). In contrast, a day-trader might focus on short-term dynamics. Both investors can achieve better results with respect to their idiosyncratic investment perspective. This insight might help to explain the enigma of why both short- and long-term portfolio strategies can be observed in practice.

Third, the results imply that an investor's utility function should include an additional dimension in the form of the investment horizon. Note that stock timescale processes cannot be traded in isolation. A portfolio strategy reducing variance at one timescale can simultaneously increase variance at other timescales. Utility gains due to risk reductions at a certain timescale might therefore be offset by losses in utility due to an increase in variance at other timescales. Thus, the benefit of scale portfolio strategies is difficult to assess in the absence of a utility function that accounts for the investment horizon.

Fourth, in contrast to general intuition, high scale portfolios exhibit less stable portfolio weights. This instability is most likely due to a lower signal-to-noise ratio. However, increasing the estimation window or applying thresholding methods can improve the stability of portfolio weights.

Finally, the diagonal volatility structure prevails irrespective of the market phase. However, differences are not significant for crisis periods.

## 6.4 Robustness

The portfolio optimization discussed in the previous section is only subject to no-short and no-leverage restrictions. However, the effects of scale portfolio optimization might only be valid under extreme portfolio allocations. To test for robustness of results, this section introduces more restrictive portfolio weighting constraints applying the *maximum weighting approach*  $A_{mw} = \left\{ \alpha \in \mathbb{R}^M \mid \alpha \in [0, b]^M, \mathbf{1}^T \alpha = 1 \right\}$  from section 4.2.

Table III.7 analyzes constrained portfolios where maximum weights of individual constituents are restricted to  $b = 20\%$  (columns 3–9) and  $b = 10\%$  (columns 10–16). While panel A displays wavelet decomposed portfolio variances, panel B shows volatilities derived from sampling interval returns.

Even with weighting restrictions imposed, the same diagonal volatility (variance) structure as previously stated can be observed. Neither the analysis with wavelet decomposition (panel A), nor with varying sampling intervals (panel B), deviates from previous results. Portfolio weight restrictions do not change the interpretation regarding the effectiveness of scale-based portfolio optimization. Further, differences in volatility (variance) between portfolios with 20% and 10% thresholds are generally relatively small. Consequently, the above findings for timescale-optimized portfolios can be considered to be robust and independent of extreme portfolio allocations.

**Table III.7:** MODWT-decomposed variance and annualized volatility of conventional time domain and scale-based (wavelet-based) minimum variance portfolios using different weighting restrictions, covering the period March 29, 1986 to December 30, 2016 (daily rebalancing).

<b>Panel A: Wavelet variance decomposition</b>																
Scale Level	Period interval	Minimum variance portfolio with 20% threshold							Minimum variance portfolio with 10% threshold							
		$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$	$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$	
J1	2 - 4	4.87	50.80	51.91	53.24	55.36	55.71	61.00	51.72	52.70	53.03	55.15	55.82	56.68	60.35	
J2	4 - 8	5.16	50.75	50.45	50.51	53.79	53.11	55.88	53.51	52.48	51.78	52.86	54.00	54.15	56.59	
J3	8 - 16	4.80	45.15	44.15*	44.23*	45.84	47.52	51.65	49.60	47.36	46.33*	46.72	48.63	48.85	52.57	
J4	16 - 32	4.30	40.84	38.17**	37.70**	38.90*	39.05*	44.54	44.38	42.76	40.36*	39.99*	42.20	41.98	44.70	
J5	32 - 64	4.38	42.50	39.86	37.68*	40.96	45.17	49.43	45.37	44.91	41.81	41.53	44.19	47.30	49.59	
J6	64 - 128	4.16	38.30	38.14	38.31	38.72	48.38	44.99	43.63	42.37	39.97	40.50	42.07	48.49	46.15	
J7	128 - 256	4.49	37.10	37.75	36.55	37.00	41.70	45.37	42.86	37.77	38.06	36.52	37.48	41.74	43.44	

Table III.7: (continued)

Panel B: Sampling interval decomposition															
Scale Level	Obs.	Minimum variance portfolio with 20% threshold (%)							Minimum variance portfolio with 10% threshold (%)						
		$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$	$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$
Daily	8,283	15.46	15.46	15.49	15.60	15.96	16.05	16.67	15.83	15.78	15.73	15.94	16.13	16.21	16.66
Bidaily	4,141	15.29	14.87*	14.82**	14.90*	15.28	15.54	16.13	15.64	15.32	15.21*	15.38	15.64	15.76	16.13
Weekly	1,656	14.67	14.18*	13.90**	14.03*	14.29	14.74	15.36	14.91	14.62	14.27*	14.40	14.74	14.88	15.39
Biweekly	828	13.76	13.30	13.12**	13.11*	13.48	14.26	14.64	14.05	13.80	13.45*	13.57	13.87	14.42	14.64
Monthly	380	14.66	14.06	13.98	13.60	13.57	14.49	15.14	14.70	14.33	14.06	13.97	14.06	14.46	14.70
Bimonthly	190	14.65	13.77	13.88	13.49	13.25	14.31	14.72	14.82	14.12	14.00	13.89	13.99	14.21	14.38
Quarterly	126	15.33	14.56	14.73	14.41	14.15	14.07	15.81	15.56	14.95	14.86	14.92	15.11	14.68	15.45
Half-yearly	63	14.91	14.25	14.12	14.08	13.68	12.81	14.87	14.92	14.31	14.22	14.42	14.19	13.28	14.07
Yearly	31	13.72	13.24	12.92	12.97	12.78	12.14	12.56	14.28	14.04	13.47	14.49	13.95	12.97	13.72

Notes: Conventional time domain portfolios  $P_{0.5y}$ ,  $P_{1y}$ , and  $P_{5y}$  are constructed using half-yearly, yearly, and 5-yearly estimation windows, respectively; portfolios  $P_{J1}, P_{J2}, \dots, P_{J7}$  refer to the scale portfolios optimized for scale variance at scale level 1–7 and are calculated using an estimation window of approximately 5 years; Panel A: each row specifies the scale variance of these portfolios at a specific scale level, where  $J1, J2, \dots, J7$  refer to scale levels 1–7. Scale variance is multiplied by the scale-level-dependent factor  $2^{j-1} \times 10^6$ ; period intervals designate the periods (in days) corresponding to a certain scale level; grey-shaded areas indicate lowest scale variances; the null hypothesis of the test statistics states that  $\tilde{v}_{P_{J1}}^2(\lambda_k) = \tilde{v}_{P_{Jj}}^2(\lambda_k)$  for  $j \neq k$  (two-sided test). However, significances are only reported where  $\tilde{v}_{P_{J1}}^2(\lambda_k) > \tilde{v}_{P_{Jj}}^2(\lambda_k)$ . Panel B: each row specifies portfolio volatility obtained from using a different sampling interval; the null hypothesis of the test statistics states that the variance of portfolio  $P_{J1}$  is equal to the variance of portfolio  $P_{Jk}$  for  $k = 2, \dots, 7$  (two-sided  $F$ -test). However, significances are only reported where the volatility (variance) of a portfolio  $\{P_{Jk}; k = 2, \dots, 7\}$  is lower compared to the volatility (variance) of  $P_{J1}$ .

\*, \*\*, and \*\*\* indicate rejection of the null hypothesis of the test statistics at the 1%, 5%, and 10% level of significance, respectively.

Additional robustness tests were carried out. These included i) altering the analysis base date, ii) changing portfolio rebalancing intervals, iii) varying the rebalancing date, and iv) adjusting the covariance matrix estimation window:

1. Results for the analysis of a return series using the sampling interval technique depend on the chosen base date (non-shift-invariance). Therefore, several base dates were tested for each analysis in which a sampling interval approach was applied. Irrespective of the chosen alternative base date, the relation of volatiles between portfolios (ranking) and the sizes of volatilities display no major differences (Table III.10 in Appendix B). Therefore, the results for the sampling interval technique are consistent with respect to the analysis base date.
2. In addition to a daily portfolio rebalancing period, rebalancing intervals of 2, 4, 8, 16, 32, 128, and 256 days were tested (Table III.11 in Appendix B shows the results for 4, 32, 64, and 256 days whereas the remain days are not shown). For almost all portfolios, variance increases with the size of the rebalancing period. However, variance does not increase proportionally for all portfolios. Portfolios whose scale most closely resemble the rebalancing period exhibit a smaller relative increase in variance.<sup>38</sup> For example, if a monthly rebalancing period is established, the portfolio optimized for monthly frequencies ( $P_{J4}$ ) exhibits a smaller relative increase in variance than the remaining portfolios.

This observation can be explained by the fact that portfolio rebalancing itself inflicts frequency dynamics on portfolio returns. As a result, scale portfolios optimizing dynamics matching the periodicity of the rebalancing interval more effectively reduce variations at the given periodicity. The relative increase in variance is smaller for these portfolios.

For an unbiased study of scale portfolio strategies, this feedback effect should be excluded from analysis. For this reason, this study resorts to using a daily rebalancing period to assess relations between portfolio strategies.

---

<sup>38</sup>However, no change in the overall variance ranking of portfolios was observed. Lowest sampling variance was still recorded for the portfolio which closely resembled the periodicity of the sampling interval (rather than the rebalancing interval).

3. Altering the portfolio rebalancing date shows no notable change in results compared to the previous analyses (not shown). This can be explained by the fact that wavelet decomposition is shift-invariant and not prone to altered starting dates. Because scale portfolios are constructed using wavelet decomposed series, altering the rebalancing date thus only exerts a minor impact on portfolio optimization. In any case, altering the rebalancing date has no influence on results when applying a daily rebalancing frequency.
4. Several different estimation window sizes were tested for the wavelet covariance matrix (results not shown; compare Table III.8). With a smaller estimation window, portfolio variance was found to decrease at low timescales. This reduction in variance can most likely be attributed to better time resolution and to taking into account local characteristics in the time series (e.g., non-stationarities). At the same time, variance at higher timescales increased for all scale portfolios. This increase in variance at higher timescales is most likely due to less representative and less stable covariance estimates.

In contrast, extending the estimation interval generally reduces variances at longer timescales. Portfolios considering long-term dynamics are more likely to suffer from ill-conditioned covariance estimates because of fewer observations. An extension of the estimation window helps to lessen the impact of these measurement errors and thus reduces volatility. At the same time, local features of the signal may be lost.

This sensitivity of the covariance matrix estimate with respect to the estimation window has some similarities with the considerations of the Heisenberg uncertainty principle.<sup>39</sup> A larger estimation window allows deriving more stable variance estimates but simultaneously leads to a loss of local features (e.g., non-stationarities).

Wavelet analysis adapts to the restrictions dictated by the Heisenberg uncertainty principle by dynamically partitioning the time-frequency plane. I suggest a simi-

---

<sup>39</sup>In quantum mechanics, the Heisenberg uncertainty principle asserts that complementary variables, such as position and moment, of a particle cannot be measured with absolute precision. The more accurately the position of the particle is determined, the less accurately the momentum of the particle is known and vice versa. In signal processing, it refers to the fact that it is not possible to simultaneously improve time and frequency resolution (Gençay et al., 2002, p. 99).

lar specification for the wavelet covariance matrix by using a dynamic estimation window approach. To this end, I define the size of the estimation window by introducing an exponential function with base 128 and an exponent  $j$  (corresponding to the scale level of analysis). For scale level  $j$ , the estimation window thus comprises  $128^j$  business days. The relation between the width of the wavelet filter and the estimation interval remains constant (so-called constant  $Q$ -factor).

Table III.8 presents the wavelet variance of scale portfolios with these dynamically adjusted estimation windows. Note that the estimation window for the covariance matrix of scale portfolio  $P_{J7}$  needs to be reduced due to an insufficient number of observations. Hence, the highest scale portfolio only uses a shorter estimation window of 4,096 data points.

Lowest variance is again observed for portfolios optimized with respect to price periodicities (scale) closely matching the scale of the variance estimate (grey diagonal). The relation between scale investment strategies and the scale variance estimate is even more evident when applying this adaptive-estimation-window approach. While the diagonal variances at low timescales only marginally change compared to the variances in Table III.4, portfolio variance at higher timescales decreases significantly.<sup>40</sup> These reductions in variance even suffice for portfolios  $P_{J5}$ ,  $P_{J6}$ , and  $P_{J7}$  to display lowest variance among all portfolios at scales 5 to 7. The diagonal structure for minimal variance (which was previously only observed at low timescales) is extended to also include higher timescales. Scale-based portfolio optimization thus effectively reduces variance at longer timescales if a sufficiently large estimation window is used. This property aligns well with the main hypothesis of this study. Thus, the results for this adaptive-estimation-window approach further support the findings in the previous section and illustrate the effectiveness of timescale-optimized investment strategies. I previously used a constant-time-window approach because of its simplicity and heuristic nature (section 6.3). However, the dynamic-estimation-window approach presented in Table III.8 generally achieves better results.

---

<sup>40</sup>In contrast, variances increase for some off-diagonal elements.

**Table III.8:** MODWT-decomposed variance of adaptive-estimation-window scale portfolio approach, covering the period March 29, 1986 to December 30, 2016 (daily rebalancing).

Scale Level	Period interval	$P_{J_1}$	$P_{J_2}$	$P_{J_3}$	$P_{J_4}$	$P_{J_5}$	$P_{J_6}$	$P_{J_7}$
J1	2 - 4	46.65 (0.031)	48.51 (0.032)	50.62 (0.033)	53.81 (0.035)	59.95 (0.038)	68.11 (0.039)	721.69 (0.042)
J2	4 - 8	47.54 (0.028)	49.01 (0.028)	49.72 (0.027)	51.56 (0.028)	57.38 (0.030)	62.99 (0.030)	65.37 (0.032)
J3	8 - 16	47.84 (0.027)	45.79 (0.025)	44.63 (0.024)	44.68 (0.025)	48.38 (0.026)	55.86 (0.028)	56.93 (0.031)
J4	16 - 32	45.27 (0.024)	41.49 (0.022)	39.13** (0.021)	37.86*** (0.020)	38.09** (0.021)	44.15 (0.020)	43.08 (0.024)
J5	32 - 64	44.89 (0.024)	43.14 (0.023)	41.70 (0.022)	39.47 (0.020)	38.72 (0.022)	46.74 (0.025)	44.30 (0.028)
J6	64 - 128	47.43 (0.021)	44.09 (0.020)	39.90 (0.020)	39.66 (0.020)	36.48** (0.019)	40.80 (0.024)	40.36 (0.026)
J7	128 - 256	54.61 (0.027)	44.18 (0.023)	38.35* (0.024)	36.34* (0.019)	34.16** (0.020)	35.43** (0.023)	35.06** (0.029)
Estim. window		128	256	512	1,024	2,048	4,096	4,096

Notes: Portfolios  $P_{J_1}, P_{J_2}, \dots, P_{J_7}$  refer to the scale portfolios optimized for scale variance at scale level 1–7 and are calculated using an adaptive estimation window; each row specifies the scale variance of the constructed portfolios at a specific scale level, where  $J_1, J_2, \dots, J_7$  refer to scale levels 1–7; scale variance is multiplied by the scale-level-dependent factor  $2^{j-1} \times 10^6$ ; period intervals designate the periods (in days) corresponding to a certain scale level; grey-shaded areas indicate lowest scale variances; the null hypothesis of the test statistics states that  $\tilde{v}_{P_{J_1}}^2(\lambda_k) = \tilde{v}_{P_{J_j}}^2(\lambda_k)$  for  $j \neq k$  (two-sided test). However, significances are only reported where  $\tilde{v}_{P_{J_1}}^2(\lambda_k) > \tilde{v}_{P_{J_j}}^2(\lambda_k)$ ; standard errors are multiplied by the scale-level-dependent factor  $\sqrt{2^{(j-1)} \times 10^9}$  and depicted in parentheses.

\*, \*\*, and \*\*\* indicate rejection of the null hypothesis of the test statistics at the 1%, 5%, and 10% level of significance, respectively.

The dynamic-estimation-window and the wavelet decomposition seem to complement each other. A narrower estimation window highlights local features, whereas a wider estimation window smooths singularities and non-stationarities. The size of the estimation window acts in a similar way as a filter. Arguably, the estimation window has a comparable impact on optimization as wavelet filtering. However, even if this were true, it does not change the general conclusion of this study: Portfolio risk at a specific timescale can be reduced with proper optimization. Moreover, the

observation that wider estimation windows lead to better results at higher timescales similarly insinuates the presence of timescale dynamics in stock returns.

## 7 Multiscale Portfolio Optimization

So far, I have used the simplifying assumption that investors operate at a specific timescale and strive to reduce risk over fixed investment horizons. However, a more realistic assumption is that investors optimize and assess risk over several different investment timescales. For example, daily variations can be conceived as posing considerable risk by long-term investors, even if these variations cancel out over the course of their investment period.

Investors often face heterogeneity in their decision-making process. For example, a portfolio manager may have to simultaneously satisfy investors with short- and long-term investment perspectives. Similarly, risk management requires identifying, evaluating, and prioritizing risk over several investment periods. Consequently, a portfolio optimization method focusing on a single timescale interval might be unsuitable for practical applications. Considering that investors have diverse timescale preferences, the portfolio optimization problem becomes more complex. The effects of portfolio adjustments must be considered across multiple timescales. Variance minimization over one frequency band changes risk at all other timescales and vice versa. Consequently, I introduce a multiscale portfolio optimization strategy that facilitates simultaneous risk diversification across multiple timescales.

Whitcher, Guttorp and Percival (1999) demonstrated that for appropriate stationary processes  $\{r_{p,t}\}$  and  $\{r_{q,t}\}$ , the scale-decomposed covariance is given by  $Cov\{r_{p,t}, r_{q,t}\} = Cov\{\tilde{V}_{p,J,t}, \tilde{V}_{q,J,t}\} + \sum_{j=1}^{J-1} \tilde{v}_{p,q}(\lambda_j)$  where  $\tilde{V}_{p,J,t}$  and  $\tilde{V}_{q,J,t}$  are the scaling coefficients of  $\{r_{p,t}\}$  and  $\{r_{q,t}\}$  at scale level  $J$ , respectively. This property allows excluding covariances between scale processes and reduces the dimensionality of a multiscale portfolio optimization problem. Consequently, a multiscale minimum variance portfolio strategy — which optimizes risk over several timescales — can be

expressed as:

$$\begin{aligned}\alpha^* &= \operatorname{argmin}_{\alpha \in A} \sum_j^J 2^{j-1} \theta_j \alpha(\lambda_j)^T \Sigma(\lambda_j) \alpha(\lambda_j) \\ &= \operatorname{argmin}_{\alpha \in A} \alpha^T \left( \sum_j^J 2^{j-1} \theta_j \Sigma(\lambda_j) \right) \alpha,\end{aligned}\tag{15}$$

where  $\Sigma(\lambda_j)$  is the scale covariance matrix and  $\alpha(\lambda_j)$  the portfolio weights at scale  $\lambda_j$ .<sup>41</sup> The upper limit of the summation  $J$  specifies the level of decomposition. The higher the level of decomposition, the more accurate the optimization for lower frequency characteristics of the signal.<sup>42</sup> The factor  $2^{j-1}$  serves as a means of normalizing scale variance and corresponds to a standardization factor derived from the energy distribution of a white-noise process (scaling law).

However, it is not possible to invest in a single-frequency process. Only stocks (which constitute superpositions of individual frequency processes) are investable. Hence, the portfolio weights in formula 15 apply equally to every scale covariance matrix and hence  $\alpha(\lambda_j) = \alpha$ . The dimension of the weight vector  $\alpha$  remains the same for both the multiscale and the single-scale portfolio optimization. Hence, the same admissible sets as outlined in section 4.2 can be used.<sup>43</sup>

Finally,  $\theta_j = \left\{ \theta \in \mathbb{R}^J \mid \theta \in [0, 1]^J, \mathbf{1}^T \theta = 1 \right\}$  specifies the relative energy that is attributed to a particular scale within the optimization. This concentration measure incorporates the multiscale utility function and characterizes the risk relationship between different timescales in accordance with an investor's preferences. The method allows minimizing risk at several timescales and can be adjusted to fit an investor's individual scale-risk preferences. The concentration measure thereby serves as a weighting function for covariance matrices at different timescales. For example, an investor preferring risk minimization at scale levels 1 and 5 can allocate 50% of rel-

<sup>41</sup>For simplification, the time index for the covariance matrix and the portfolio weights is omitted.

<sup>42</sup>At scale level  $J$ , the covariance matrix estimation needs to be replaced by the covariance matrix of the respective scaling coefficients in order to cover the full frequency spectrum.

<sup>43</sup>In accordance with previous notation,  $A$  either represents the long-only or the maximum-weight admissible set.

ative energy to scale level 1 (i.e.,  $\theta_1 = 0.5$ ) and 50% of relative energy to scale level 5 (i.e.,  $\theta_5 = 0.5$ ). If the concentration measure is specified as  $\theta_j = 2^{-(j-1)}$ , the optimization in formula 15 corresponds to a minimum variance strategy applied to untransformed returns. If relative energy is attributed to only one timescale, a single-scale portfolio optimization results (as applied in section 6.3).

Note that other functional forms are possible for multiscale portfolio optimization. For example, in contrast to the approach in formula 15, a regularization term could be added to the multiscale minimum variance objective function. This term would serve as a cost function penalizing high risk concentrations in certain timescales. However, the formulation in formula 15 allows integrating both conventional minimum-variance optimization and single-scale portfolio optimization into a generalized model.

The approach presented here lays the foundation for introducing multiscale investment strategies. This method helps to include individual timescale preferences in portfolio decision-making processes. While other research has studied the investment horizon in multi-period optimization problems, this section provides a new approach to optimizing the multi-horizon portfolio choice problem.

## 8 Conclusion

This study has investigated timescale dynamics of stock returns and the implications of these dynamics for an investor's portfolio formation process. It has applied wavelet analysis to decompose the time series into its individual scale components, which reflect stock return dynamics at a particular timescale. Decomposition has enabled examining the behavior of stock return interdependencies and variances at particular timescales. Based on this decomposition, I have developed a portfolio optimization method that allows minimizing risk at a specific investment timescale.

My empirical results suggest that the scale-based portfolio optimization strategy minimizes portfolio variance at a targeted timescale. Similarly, results indicate that stock prices unfold varying timescale characteristics. Nevertheless, these findings

must be treated with caution because the significance of the presented results is mostly limited to 5%–10%.

One important implication of the above findings is that a stock's diversification potential depends on an investor's investment perspective and evaluation approach. Hence, the investment timeframe is an important factor for assessing and pricing stocks. It is thus essential to simultaneously analyze financial assets in the time- and spectral-domain.

Similarly, the investment timescale is central to portfolio construction and should be included in deriving the optimal portfolio choice. Portfolio managers should optimize their portfolios in accordance with their investment horizon perspectives and the horizons over which performance/risk is measured. For example, it might be sensible for a long-term portfolio manager — whose performance is assessed on a long-term basis — to more strongly weigh long-term information. In contrast, it might be more beneficial for a short-term trader — whose performance is measured on a day-to-day basis — to invest with respect to daily timescale information. Consequently, the investment horizon should be considered as an additional dimension in an investor's utility function.

In practice, it is unlikely that market agents optimize their portfolios only over a particular timescale. Instead, they base their investment decisions on different time horizons and diversify their portfolios over multiple timescales. This study has shown that the conventional covariance matrix estimate already provides some (decent) diversification across different timescales. However, I have presented a more flexible multiscale portfolio strategy, one which can incorporate investors' selective timescale preferences. This strategy extends existing multi-horizon portfolio choice approaches.

## Appendix A

### Wavelet Squared Coherence

After the decomposition of variance in formula 4, a natural next step involves deriving the cross-wavelet transform and the cross-wavelet power. Let  $r_{p,t}$  and  $r_{q,t}$  specify two time series with wavelet transforms  $\mathcal{W}_p(\lambda, \tau)$  and  $\mathcal{W}_q(\lambda, \tau)$ , respectively. The cross-wavelet transform is defined as  $\mathcal{W}_{p,q}(\lambda, \tau) \equiv \mathcal{W}_p(\lambda, \tau) \mathcal{W}_q(\lambda, \tau)^*$ . The cross-wavelet power spectrum then follows from  $|\mathcal{W}_{p,q}(\lambda, \tau)|$ . This measure indicates common power between two series in the time-timescale-space. Analogously to the wavelet power spectrum, this quantity can be interpreted as a local covariance measure of the time series for different time and timescales.<sup>44</sup>

When examining stock return comovements in the time domain, the relationship between stocks is often considered independently of variance. Similarly, it can be of interest to examine the interaction patterns between two stocks independently of cross-wavelet power. The wavelet coherence measure is a useful tool for this purpose. It can be considered as a localized correlation coefficient in the time-timescale-domain. Following Grinsted, Moore and Jevrejeva (2004), as well as Torrence and Webster (1999), wavelet coherence is computed as

$$R_{p,q}^2(\lambda, \tau) \equiv \frac{|\mathcal{S}(\lambda^{-1} \mathcal{W}_{p,q}(\lambda, \tau))|^2}{\mathcal{S}(\lambda^{-1} |\mathcal{W}_p(\lambda, \tau)|^2) \mathcal{S}(\lambda^{-1} |\mathcal{W}_q(\lambda, \tau)|^2)}, \quad (16)$$

where  $\mathcal{S}(\cdot)$  refers to a smoothing operator and  $0 \leq R_{p,q}^2(\lambda, \tau) \leq 1$  (Rua & Nunes, 2009; Torrence & Webster, 1999).<sup>45</sup> The smoothing operation is successively applied to the time and to the scale dimension of the wavelet coherence coefficients

<sup>44</sup>If the wavelet function is complex, the resulting wavelet transforms are also complex. As a result, the wavelet transforms can be divided into a real and an imaginary part. In addition to the amplitude, this property allows deriving the phase for different timescales  $\tan^{-1}(\mathcal{I}\{\mathcal{W}_{p,q}(\lambda, \tau)\} / \mathcal{R}\{\mathcal{W}_{p,q}(\lambda, \tau)\})$ .

<sup>45</sup>Following Torrence and Webster (1999), a suitable smoothing operator for the Morlet wavelet is  $S_{time}(\mathcal{W})|_{\lambda} = \left( \mathcal{W}(\lambda, \tau)^* \varrho_1^{-t^2/2\lambda^2} \right) \Big|_{\lambda}$  and  $S_{scale}(\mathcal{W})|_{\tau} = (\mathcal{W}(\lambda, \tau)^* \varrho_2 \Pi(0, 6\lambda))|_{\tau}$ , where  $\varrho_1$ , and  $\varrho_2$  are normalization constants, and  $\Pi$  is the rectangle function. In practice, the normalization constants are estimated numerically. The constant 0.6 refers to the scale decorrelation length of the Morlet wavelet (Grinsted et al., 2004; Torrence & Webster, 1999).

such that  $\mathcal{S}_{scale}(\mathcal{S}_{time}(\cdot))$ , where  $\mathcal{S}_{time}(\cdot)$  refers to the smoothing operation in the time dimension, and  $\mathcal{S}_{scale}$  to the smoothing operation in the timescale dimension.<sup>46</sup> Following Grinsted et al. (2004) and Torrence and Compo (1998), the statistical significance of the wavelet squared coherence can be obtained by applying Monte Carlo simulation (generation of surrogate matrices).

This study uses finite length return observations. When deriving the wavelet transforms for finite data, the signal is implicitly assumed to be cyclical. As a result, wavelet transforms at the beginning and at the end of the series are erroneous. This problem can be mitigated by padding the time series with zeros (Torrence & Compo, 1998). An alternative approach is to reflect the signal at the last data point. This study uses the second approach.

Nevertheless, the boundary values of the wavelet transforms remain less reliable. Note that the support of the wavelet function increases with the scale parameter  $\lambda$ . As a result, the number of wavelet transforms suffering from these edge-effects increases with  $\lambda$  (Madaleno & Pinho, 2014). The cone of influence (COI) defines the region in the wavelet squared coherence spectrum that is affected by these edge-effects (boundary conditions) (Torrence & Compo, 1998). This COI must be interpreted with caution when analyzing wavelet squared coherence.

---

<sup>46</sup>This application of a smoothing operation is analogous to the derivation of coherency in Fourier transformation. The operation is necessary in order that the wavelet squared coherence is not simply unity.

## Appendix B

**Table III.9:** MODWT-decomposed variance of scale-based (wavelet-based) minimum variance portfolios for crisis and non-crisis periods, covering the period March 29, 1986 to December 30, 2016.

Scale level	Conv. covariance estim. portfolios (%)			Wavelet covariance estimator portfolios (%)						
	$P_{0.5y}$	$P_{1y}$	$P_{5y}$	$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$
<b>Panel A: Crisis</b>										
J1	82.78	84.86	95.82	94.02	101.48	104.87	109.46	111.22	114.99	127.82
J2	84.82	86.13	94.67	95.81	99.16	101.58	102.09	105.68	106.40	108.15
J3	76.20	78.11	80.12	84.19	82.33	80.72	82.51	86.64	94.93	98.61
J4	68.69	69.58	66.20	72.90	68.57	62.55	64.46	66.27	65.48	72.93
J5	61.51	65.84	60.74	65.73	62.79	58.55	58.06	62.55	66.24	67.85
J6	56.62	64.47	55.87	60.45	54.93	51.37	48.43	49.97	59.38	68.00
J7	21.55	23.03	23.77	27.64	21.14	22.85	24.27	26.33	25.74	37.46
<b>Panel B: Non-Crisis</b>										
J1	27.70	27.37	29.13	29.93	29.39	30.09	31.99	35.73	37.92	38.05
J2	84.82	27.53	31.15	33.67	30.83**	30.01***	30.71**	34.57	36.12	39.35
J3	76.20	28.14	29.93	33.10	29.70*	28.93**	28.50***	30.84	31.48	37.39
J4	68.69	26.89	28.18	31.13	29.16	27.36*	25.80**	26.84**	27.69	32.02
J5	61.51	32.03	32.12	34.69	33.57	31.31	28.33*	30.25	34.48	40.60
J6	56.62	35.12	31.78	34.03	31.81	31.58	33.55	32.17	42.36	36.23
J7	21.55	55.68	47.63	52.84	44.44	44.00	38.84	41.06	48.88	54.84

*Notes:* Conventional time domain portfolios  $P_{0.5y}$ ,  $P_{1y}$ , and  $P_{5y}$  are constructed using half-yearly, yearly, and 5-yearly estimation windows, respectively; portfolios  $P_{J1}, P_{J2}, \dots, P_{J7}$  refer to the scale portfolios optimized for scale variance at scale level 1–7 and are calculated using an estimation window of approximately 5 years; each row specifies the scale variance of these portfolios at a specific scale level, where  $J1, J2, \dots, J7$  refer to scale levels 1–7. Scale variance is multiplied by the scale-level-dependent factor  $2^{j-1} \times 10^6$ ; period intervals designate the periods (in days) corresponding to a certain scale level; grey-shaded areas indicate lowest scale variances; the null hypothesis of the test statistics states that  $\tilde{v}_{P_{J1}}^2(\lambda_k) = \tilde{v}_{P_{Jj}}^2(\lambda_k)$  for  $j \neq k$  (two-sided test). However, significances are only reported where  $\tilde{v}_{P_{J1}}^2(\lambda_k) > \tilde{v}_{P_{Jj}}^2(\lambda_k)$ .

\*, \*\*, and \*\*\* indicate rejection of the null hypothesis of the test statistics at the 1%, 5%, and 10% level of significance, respectively.

**Table III.10:** Annualized volatility of conventional time domain and scale-based (wavelet-based) minimum variance portfolios using different base date, covering the period March 29, 1986 to December 30, 2016 (daily rebalancing).

Sampling interval	Obs.	Conv. covariance estim. portfolios (%)			Wavelet covariance estimator portfolios (%)						
		$P_{0.5y}$	$P_{1y}$	$P_{5y}$	$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$
<b>Panel A: 50 day lag</b>											
Daily	8,283	14.63	14.66	15.24	15.49	15.45	15.52	15.76	16.24	16.60	17.17
Bidaily	4,141	14.36	14.37	14.76	15.29	14.84	14.79	14.88	15.36	15.82	16.64
Weekly	1,656	14.30	14.15	14.06	14.69	14.14	13.76	13.99	14.45	14.98	15.64
Biweekly	828	13.53	13.49	13.17	13.74	13.25	12.99	13.05	13.42	14.32	15.09
Monthly	380	14.56	14.57	14.22	14.91	14.22	13.94	13.72	13.74	14.40	15.48
Bimonthly	190	14.91	14.76	14.13	14.75	13.92	13.92	13.83	13.64	14.38	15.63
Quarterly	126	15.21	15.32	14.49	15.21	14.26	14.26	13.98	14.29	14.46	16.56
Half-yearly	63	15.49	16.05	15.70	16.23	15.66	15.72	15.19	15.08	15.30	17.30
Yearly	31	16.56	17.23	17.72	18.13	17.71	17.52	16.44	16.51	16.45	18.99
<b>Panel B: 100 day lag</b>											
Daily	8,283	14.81	14.83	15.43	15.67	15.64	15.71	15.95	16.44	16.80	17.37
Bidaily	4,141	14.36	14.37	14.76	15.29	14.84*	14.79**	14.88*	15.36	15.82	16.64
Weekly	1,656	14.31	14.32	14.46	15.05	14.52*	14.35**	14.22***	14.64	15.19	16.26
Biweekly	828	14.16	14.18	14.02	14.63	14.07	13.91	13.63**	13.99	14.46	15.24
Monthly	380	14.53	14.61	14.33	14.87	14.34	14.12	13.79*	14.05	14.81	15.61
Bimonthly	190	13.77	13.96	13.58	14.12	13.43	13.42	12.84	12.98	14.01	15.10
Quarterly	126	14.61	14.62	14.11	14.80	13.97	13.67	12.93	13.13	13.63	14.74
Half-yearly	63	15.19	15.12	14.14	15.01	14.18	13.53	13.28	13.15	13.40	14.77
Yearly	31	14.13	15.02	13.27	13.90	13.12	13.01	13.28	13.01	13.12	11.90

Notes: Conventional time domain portfolios  $P_{0.5y}$ ,  $P_{1y}$ , and  $P_{5y}$  are constructed using half-yearly, yearly, and 5-yearly estimation windows, respectively; portfolios  $P_{J1}, P_{J2}, \dots, P_{J7}$  refer to the scale portfolios optimized for scale variance at scale level 1–7 and are calculated using an estimation window of approximately 5 years; grey-shaded areas indicate lowest scale variances; the null hypothesis of the test statistics states that the variance of portfolio  $P_{J1}$  is equal to the variance of portfolio  $P_{Jk}$  for  $k = 2, \dots, 7$  (two-sided  $F$ -test). However, significances are only reported where the volatility (variance) of a portfolio  $\{P_{Jk}; k = 2, \dots, 7\}$  is lower compared to the volatility (variance) of  $P_{J1}$ .

\*, \*\*, and \*\*\* indicate rejection of the null hypothesis of the test statistics at the 1%, 5%, and 10% level of significance, respectively.

**Table III.11:** *MODWT-decomposed variance of scale-based (wavelet-based) minimum variance portfolios using different rebalancing intervals, covering the period March 29, 1986 to December 30, 2016.*

Scale Level	Period interval	Rebalancing interval: 4 days							Rebalancing interval: 32 days						
		$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$	$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$
J1	2 - 4	4.88	50.57	52.37	55.01	58.21	60.81	64.13	49.26	50.51	52.28	55.43	59.25	61.20	65.98
J2	4 - 8	5.27	51.62	51.38	51.88	55.61	56.72	59.99	53.31	51.61	51.23	52.38	57.02	57.21	63.21
J3	8 - 16	4.81	45.28	44.11**	44.23**	47.13	49.41	55.00	48.43	45.46	43.94**	44.41**	47.40	50.22	56.97
J4	16 - 32	4.30	40.44	37.71**	36.76**	38.11**	38.52*	44.02	42.65	39.97	37.10**	36.36**	37.73**	38.13*	43.56
J5	32 - 64	4.33	41.75	39.22	37.02*	39.42	44.01	48.62	43.69	42.10	39.28	37.26*	38.27*	41.63	47.31
J6	64 - 128	4.11	37.69	37.13	38.14	36.98	46.36	43.30	41.01	38.02	37.83	37.03	35.44	41.44	42.34
J7	128 - 256	4.42	36.42	37.24	34.77	37.28	40.00	44.05	46.20	38.53	40.12	36.14	37.28	39.64	39.19

Table III.11: (continued)

Scale Level	Period interval	Rebalancing interval: 64 days							Rebalancing interval: 256 days						
		$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$	$P_{J1}$	$P_{J2}$	$P_{J3}$	$P_{J4}$	$P_{J5}$	$P_{J6}$	$P_{J7}$
J1	2 - 4	4.96	51.06	51.53	55.33	60.51	61.28	66.35	51.78	53.07	53.49	56.37	60.85	60.91	66.19
J2	4 - 8	5.34	51.67	50.67*	52.18	57.82	57.52	62.74	54.21	52.80	52.05	53.00	58.53	58.18	65.05
J3	8 - 16	4.84	45.14	43.79**	44.27**	47.83	50.51	56.61	49.00	46.50	45.89*	45.38*	49.10	50.22	58.00
J4	16 - 32	4.26	40.01	37.19**	36.54**	38.11*	37.74**	42.36	41.30	39.48	37.51*	37.31*	38.89	38.21	45.36
J5	32 - 64	4.37	42.04	39.12	37.48*	38.74	41.58	46.42	44.46	42.33	40.03	38.32*	38.62*	41.81	48.20
J6	64 - 128	4.08	37.90	37.74	36.85	35.49	40.44	42.75	41.19	38.48	38.17	36.55	36.01	38.84	43.96
J7	128 - 256	4.56	38.14	39.45	35.68	37.89	39.95	40.03	43.97	37.66	38.41	34.66	36.75	35.17	34.50

Notes: Portfolios  $P_{J1}, P_{J2}, \dots, P_{J7}$  refer to the scale portfolios optimized for scale variance at scale levels 1–7 and are calculated using an estimation window of approximately 5 years; number of observations reduce by the size of the rebalancing window (not shown); each row specifies the scale variance of these portfolios at a specific scale level, where  $J1, J2, \dots, J7$  refer to scale levels 1–7. Scale variance is multiplied by the scale-level-dependent factor  $2^{j-1} \times 10^6$ ; period intervals designate the periods (in days) corresponding to a certain scale level; grey-shaded areas indicate lowest scale variances; the null hypothesis of the test statistics states that  $\tilde{v}_{P_{J1}}^2(\lambda_k) = \tilde{v}_{P_{Jj}}^2(\lambda_k)$  for  $j \neq k$  (two-sided test). However, significances are only reported where  $\tilde{v}_{P_{J1}}^2(\lambda_k) > \tilde{v}_{P_{Jj}}^2(\lambda_k)$ . \*, \*\*, and \*\*\* indicate rejection of the null hypothesis of the test statistics at the 1%, 5%, and 10% level of significance, respectively.



# Bibliography

- Aguiar-Conraria, L., Azevedo, N. & Soares, M. J. (2008). Using Wavelets to Decompose the Time-Frequency Effects of Monetary Policy. *Physica A: Statistical Mechanics and its Applications*, 387(12), 2863–2878. doi: 10.1016/j.physa.2008.01.063
- Aguiar-Conraria, L. & Joana Soares, M. (2011). Business Cycle Synchronization and the Euro: A Wavelet Analysis. *Journal of Macroeconomics*, 33(3), 477–489. doi: 10.1016/j.jmacro.2011.02.005
- Aguiar-Conraria, L., Rodrigues, T. M. & Soares, M. J. (2014). Oil Shocks and the Euro as an Optimum Currency Area. In M. Gallegati & W. Semmler (Eds.), *Wavelet applications in economics and finance* (pp. 143–156). Cham, Switzerland: Springer International Publishing.
- Akoum, I., Graham, M., Kivihaho, J., Nikkinen, J. & Omran, M. (2012). Co-Movement of Oil and Stock Prices in the GCC Region: A Wavelet Analysis. *Quarterly Review of Economics and Finance*, 52(4), 385–394. doi: 10.1016/j.qref.2012.07.005
- Aloui, C. & Hkiri, B. (2014). Co-Movements of GCC Emerging Stock Markets: New Evidence from Wavelet Coherence Analysis. *Economic Modelling*, 36, 421–431. doi: 10.1016/j.econmod.2013.09.043
- Andries, A. M., Ihnatov, I. & Tiwari, A. K. (2014). Analyzing Time-Frequency Relationship Between Interest Rate, Stock Price and Exchange Rate Through Continuous Wavelet. *Economic Modelling*, 41, 227–238. doi: 10.1016/j.econmod.2014.05.013
- Ang, A. & Chen, J. (2002). Asymmetric Correlations of Equity Portfolios. *Journal of Financial Economics*, 63(3), 443–494. doi: 10.1016/S0304-405X(02)00068-5
- Anufriev, M., Bottazzi, G. & Pancotto, F. (2006). Equilibria, Stability and Asymptotic Dominance in a Speculative Market With Heterogeneous Traders. *Journal of Economic Dynamics and Control*, 30(9-10), 1787–1835. doi: 10.1016/j.jedc.2005.10.012

- Arthur, W. B., Holland, J. H., LeBaron, B., Palmer, R. & Tayler, P. (1997). Asset Pricing under Endogenous Expectations in an Artificial Stock Market. In B. W. Arthur, S. N. Durlauf & D. A. Lane (Eds.), *The economy as an evolving complex system* (2nd ed., pp. 15–44). Boca Raton, FL: CRC Press.
- Aste, T., Shaw, W. & Di Matteo, T. (2010, August). Correlation Structure and Dynamics in Volatile Markets. *New Journal of Physics*, 12(8), 085009. doi: 10.1088/1367-2630/12/8/085009
- Bachelier, L. (1900). Théorie de la spéculation. *Annales scientifiques de l'École normale supérieure*, 17, 21–86. doi: 10.24033/asens.476
- Barunik, J. & Vacha, L. (2009, October). *Wavelet Analysis of Central European Stock Market Behaviour During the Crisis* (Vol. Working Papers IES; Working Papers IES No. 2009/23). Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies. Retrieved from [https://ideas.repec.org/p/fau/wpaper/wp2009\\_23.html](https://ideas.repec.org/p/fau/wpaper/wp2009_23.html)
- Bekiros, S., Nguyen, D. K., Uddin, G. S. & Sjö, B. (2016). On the Time Scale Behavior of Equity-Commodity Links: Implications for Portfolio Management. *Journal of International Financial Markets, Institutions and Money*, 41, 30–46. doi: 10.1016/j.intfin.2015.12.003
- Berger, T. (2016). Wavelet Decomposition and Applied Portfolio Management. *Journal of Risk*, 18(4), 53–77.
- Best, M. J. & Grauer, R. R. (1991). On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results. *Review of Financial Studies*, 4(2), 315–342. doi: 10.1093/rfs/4.2.315
- Billio, M., Getmansky, M., Lo, A. W. & Pelizzon, L. (2012). Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors. *Journal of Financial Economics*, 104(3), 535–559. Retrieved from <http://dx.doi.org/10.1016/j.jfineco.2011.12.010> doi: 10.1016/j.jfineco.2011.12.010
- Bjornson, B., Kim, H. S. & Lee, K. (1999). Low and High Frequency Macroeconomic Forces in Asset Pricing. *The Quarterly Review of Economics and Finance*, 39(1), 77–100.
- Borghesi, C., Marsili, M. & Miccichè, S. (2007). Emergence of time-horizon in-

- variant correlation structure in financial returns by subtraction of the market mode. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 76(2). doi: 10.1103/PhysRevE.76.026104
- Bouchaud, J.-P. & Potters, M. (2003). *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management* (2nd ed.). Cambridge: Cambridge University Press.
- Brailsford, T. J. & Faff, R. W. (1997). Testing the Conditional CAPM and the Effect of Intervaling: A Note. *Pacific-Basin Finance Journal*, 5(5), 527–537. doi: 10.1016/S0927-538X(97)00018-8
- Britten-Jones, M. (1999). The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights. *The Journal of Finance*, 54(2), 655–671.
- Broadie, M. (1993). Computing Efficient Frontiers Using Estimated Parameters. *Annals of Operations Research*, 45(1), 21–58. doi: 10.1007/BF02282040
- Brock, W. A. & Hommes, C. H. (1998). Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model. *Journal of Economic Dynamics and Control*, 22(8-9), 1235–1274. doi: 10.1016/s0165-1889(98)00011-6
- Brody, T. A., Flores, J., French, J. B., Mello, P. A., Pandey, A. & Wong, S. S. M. (1981, jul). Random-Matrix Physics: Spectrum and Strength Fluctuations. *Reviews of Modern Physics*, 53(3), 385–479. doi: 10.1103/RevModPhys.53.385
- Bruce, A. & Gao, H. Y. (1996). *Applied Wavelet Analysis with S-Plus*. New York, NY: Springer.
- Buiter, W. H. (2003). James Tobin: An Appreciation of His Contribution to Economics. *Economic Journal*, 113(491). doi: 10.1046/j.0013-0133.2003.00170.x
- Burns, A. F. & Mitchell, W. C. (1946). *Measuring Business Cycles*. New York, NY: National Bureau of Economic Research.
- Campbell, J. Y., Lettau, M., Malkiel, B. G. & Xu, Y. (2001). Have Individual Stocks Become More Volatile. An Empirical Exploration of Idiosyncratic Risk. *The Journal of Finance*, 56(1), 1–43.
- Candelon, B., Piplack, J. & Straetmans, S. (2008). On Measuring Synchronization of Bulls and Bears: The Case of East Asia. *Journal of Banking and Finance*,

- 32(6), 1022–1035. doi: 10.1016/j.jbankfin.2007.08.003
- Cattell, R. B. (1966, apr). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, 1(2), 245–276. doi: 10.1207/s15327906mbr0102{\\_}10
- Chakrabarty, A., De, A., Gunasekaran, A. & Dubey, R. (2015). Investment Horizon Heterogeneity and Wavelet: Overview and Further Research Directions. *Physica A: Statistical Mechanics and its Applications*, 429, 45–61. doi: 10.1016/j.physa.2014.10.097
- Chamberlain, G. & Rothschild, M. (1983, sep). Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets. *Econometrica*, 51(5), 1281. doi: 10.2307/1912275
- Chan, L. K. C., Karceski, J. & Lakonishok, J. (1999). On Portfolio Pptimization: Forecasting Covariannccs and Choosing the Risk Model. *Review of Financial Studies*, 12(5), 937–974.
- Chatterjee, A. & Chakrabarti, B. K. (2006). *Econophysics of Stock and other Markets*. doi: 10.1007/978-88-470-0502-0
- Chaudhuri, S. E. & Lo, A. W. (2016). Spectral Portfolio Theory. *SSRN Electronic Journal*. Retrieved from <https://www.ssrn.com/abstract=2788999> doi: 10.2139/ssrn.2788999
- Chen, W. D. & Li, H. C. (2016). Wavelet Decomposition of Heterogeneous Investment Horizon. *Journal of Economics and Finance*, 40(4), 714–734. doi: 10.1007/s12197-015-9321-y
- Conlon, T. (2009). *Alternative Risk Management : Correlation and Complexity Signed* (Ph.D. Thesis). Dublin City University.
- Conlon, T., Cotter, J. & Gençay, R. (2018). Long-Run Wavelet-Based Correlation for Financial Time Series. *European Journal of Operational Research*, 271(2), 676–696. doi: 10.1016/j.ejor.2018.05.028
- Conlon, T., Crane, M. & Ruskin, H. J. (2008). Wavelet Multiscale Analysis for Hedge Funds: Scaling and Strategies. *Physica A: Statistical Mechanics and its Applications*, 387(21), 5197–5204. doi: 10.1016/j.physa.2008.05.046
- Conlon, T., Ruskin, H. J. & Crane, M. (2009). Cross-Correlation Dynamics in Financial Time Series. *Physica A: Statistical Mechanics and its Applications*,

- 388(5), 705–714. doi: 10.1016/j.physa.2008.10.047
- Connor, G. & Korajczyk, R. a. (1986, mar). Performance Measurement With the Arbitrage Pricing Theory. *Journal of Financial Economics*, 15(3), 373–394. doi: 10.1016/0304-405X(86)90027-9
- Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2), 174–196. doi: 10.1093/jfinec/nbp001
- Dacorogna, M. M., Gençay, R., Müller, U., Olsen, R. B. & Pictet, O. V. (2001). *An Introduction to High-Frequency Finance*. San Diego, CA: Academic Press.
- Dacorogna, M. M., Müller, U. A., Pictet, O. V. & Olsen, R. B. (1998). Modelling Short-Term Volatility with GARCH and HARCH Models. In C. L. Dunis & B. Zhou (Eds.), *Nonlinear modelling of high frequency financial time series*. Chichester: John Wiley & Sons.
- Dajčman, S. (2013). Interdependence Between Some Major European Stock Markets - A Wavelet Lead/Lag Analysis. *Prague Economic Papers*, 22(1), 28–49. doi: 10.18267/j.pep.439
- Daubechies, I. (1990). The Wavelet Transform: Time-Frequency Localization and Signal Analysis. *IEEE Transactions on Information Theory*, 36(5), 961–1005.
- Daubechies, I. (1992). *Ten Lectures on Wavelets*. Pennsylvania, PA: Society for Industrial and Applied Mathematics.
- DeMiguel, V., Garlappi, L., Nogales, F. J. & Uppal, R. (2009). A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms. *Management Science*, 55(5), 798–812. doi: 10.1287/mnsc.1080.0986
- Di Matteo, T. (2007). Multi-Scaling in Finance. *Quantitative Finance*, 7(1), 21–36. doi: 10.1080/14697680600969727
- Drozd, S., Grümmner, F., Górski, A. Z., Ruf, F. & Speth, J. (2000). Dynamics of Competition Between Collectivity and Noise in the Stock Market. *Physica A: Statistical Mechanics and its Applications*, 287(3-4), 440–449. doi: 10.1016/S0378-4371(00)00383-6
- Drozd, S., Grümmner, F., Ruf, F. & Speth, J. (2001). Towards Identifying the World Stock Market Cross-Correlations: DAX Versus Dow Jones. *Physica A:*

- Statistical Mechanics and its Applications*, 294(1-2), 226–234. doi: 10.1016/S0378-4371(01)00119-4
- Dyson, F. J. (1962, jan). Statistical Theory of the Energy Levels of Complex Systems. I. *Journal of Mathematical Physics*, 3(1), 140–156. doi: 10.1063/1.1703773
- Dyson, F. J. & Mehta, M. L. (1963, may). Statistical Theory of the Energy Levels of Complex Systems. IV. *Journal of Mathematical Physics*, 4(5), 701–712. doi: 10.1063/1.1704008
- Elton, E. J. & Gruber, Martin, J. (1977). Risk Reduction and Portfolio Size: An Analytical Solution. *The Journal of Business*, 50(4), 415–437.
- Eom, C., Oh, G., Jung, W. S., Jeong, H. & Kim, S. (2009). Topological Properties of Stock Networks Based on Minimal Spanning Tree and Random Matrix Theory in Financial Time Series. *Physica A: Statistical Mechanics and its Applications*, 388(6), 900–906. doi: 10.1016/j.physa.2008.12.006
- Epps, T. W. (1979). Comovements in Stock Prices in the Very Short Run. *Journal of the American Statistical Association*, 74(366), 291. doi: 10.2307/2286325
- Evans, J. L. & Archer, S. H. (1968). Diversification and the Reduction of Dispersion: An Empirical Analysis. *The Journal of Finance*, 23(5), 761–767. doi: 10.2307/2325905
- Fama, E. F. (1980). Stock Returns, Expected Returns, and Real Activity. *Journal of Finance*, 45(4), 1089–1108.
- Fama, E. F. (1981). Stock Returns, Real Activity, Inflation, and Money. *The American Economic Review*, 71(4), 545–565.
- Fenn, D. J., Porter, M. A., Williams, S., McDonald, M., Johnson, N. F. & Jones, N. S. (2011). Temporal Evolution of Financial-Market Correlations. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 84(2), 1–13. doi: 10.1103/PhysRevE.84.026109
- Fernández-Macho, J. (2012). Wavelet Multiple Correlation and Cross-Correlation: A Multiscale Analysis of Eurozone Stock Markets. *Physica A: Statistical Mechanics and its Applications*, 391(4), 1097–1104. doi: 10.1016/j.physa.2011.11.002
- Fyodorov, Y. V. & Mirlin, A. D. (1992, aug). Analytical Derivation of the Scal-

- ing Law for the Inverse Participation Ratio in Quasi-One-Dimensional Disordered Systems. *Physical Review Letters*, 69(7), 1093–1096. doi: 10.1103/PhysRevLett.69.1093
- Fyodorov, Y. V. & Mirlin, A. D. (1993, jul). Level-to-Level Fluctuations of the Inverse Participation Ratio in Finite Quasi 1D Disordered Systems. *Physical Review Letters*, 71(3), 412–415. doi: 10.1103/PhysRevLett.71.412
- Gallegati, M. (2005). A Wavelet Analysis of MENA Stock Markets. *Macroeconomic Dynamics*(April), 1–18.
- Gallegati, M., Gallegati, M., Ramsey, J. B. & Semmler, W. (2014). Does Productivity Affect Unemployment? A Time-Frequency Analysis for the US. In M. Gallegati & W. Semmler (Eds.), *Wavelet applications in economics and finance* (pp. 23–46). Cham, Switzerland: Springer International Publishing.
- Gençay, R., Gradojevic, N., Selçuk, F. & Whitcher, B. (2010). Asymmetry of Information Flow Between Volatilities Across Time Scales. *Quantitative Finance*, 10(8), 895–915. doi: 10.1080/14697680903460143
- Gençay, R., Selçuk, F. & Whitcher, B. (2001a, jan). Differentiating Intraday Seasonalities Through Wavelet Multi-Scaling. *Physica A: Statistical Mechanics and its Applications*, 289(3-4), 543–556. doi: 10.1016/S0378-4371(00)00463-5
- Gençay, R., Selçuk, F. & Whitcher, B. (2001b). Scaling Properties of Foreign Exchange Volatility. *Physica A: Statistical Mechanics and its Applications*, 289(1-2), 249–266. doi: 10.1016/S0378-4371(00)00456-8
- Gençay, R., Selçuk, F. & Whitcher, B. (2002). *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. New York, NY: Academic Press.
- Gençay, R., Selçuk, F. & Whitcher, B. (2003). Systematic Risk and Timescales. *Quantitative Finance*, 3(2), 108–116. doi: 10.1088/1469-7688/3/2/305
- Gençay, R., Selçuk, F. & Whitcher, B. (2005). Multiscale Systematic Risk. *Journal of International Money and Finance*, 24(1), 55–70. doi: 10.1016/j.jimonfin.2004.10.003
- Gilster, J. E. (1983). Capital Market Equilibrium with Divergent Investment Horizon Length Assumptions. *Journal of Financial and Quantitative Analysis*, 18(2), 257–268. doi: 10.2307/2330922

- Gopikrishnan, P., Rosenow, B., Plerou, V. & Stanley, H. E. (2000, November). Identifying Business Sectors from Stock Price Fluctuations. *arXiv.org, Quantitative Finance Papers*, 2, 997–1000.
- Gopikrishnan, P., Rosenow, B., Plerou, V. & Stanley, H. E. (2001). Quantifying and Interpreting Collective Behavior in Financial Markets. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 64(3), 4. doi: 10.1103/PhysRevE.64.035106
- Graham, M., Kiviaho, J. & Nikkinen, J. (2012). Integration of 22 Emerging Stock Markets: A Three-Dimensional Analysis. *Global Finance Journal*, 23(1), 34–47. doi: 10.1016/j.gfj.2012.01.003
- Graham, M., Kiviaho, J., Nikkinen, J. & Omran, M. (2013). Global and Regional Co-Movement of the MENA Stock Markets. *Journal of Economics and Business*, 65, 86–100. doi: 10.1016/j.jeconbus.2012.09.005
- Green, R. C. & Hollifield, B. (1992). When Will Mean-Variance Efficient Portfolios Be Well Diversified. *The Journal of Finance*, 47(5), 1785–1809.
- Gressis, N., Philippatos, G. C. & Hayya, J. (1976). Multiperiod Portfolio Analysis and the Inefficiency of the Market Portfolio. *Journal of Finance*, 31(4), 1115–1126.
- Grinsted, A., Moore, J. C. & Jevrejeva, S. (2004). Application of the Cross Wavelet Transform and Wavelet Coherence to Geophysical Time Series. *Nonlinear Processes in Geophysics*, 11(5/6), 561–566. doi: 10.5194/npg-11-561-2004
- Grossmann, A. & Morlet, J. (1984, jul). Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape. *SIAM Journal on Mathematical Analysis*, 15(4), 723–736. doi: 10.1137/0515056
- Guhr, T., Müller-Groeling, A. & Weidenmüller, H. A. (1998). Random-Matrix Theories in Quantum Physics: Common Concepts. *Physics Report*, 299(4-5), 189–425. doi: 10.1016/S0370-1573(97)00088-4
- Guttman, L. (1954, jun). Some Necessary Conditions for Common-Factor Analysis. *Psychometrika*, 19(2), 149–161. doi: 10.1007/BF02289162
- Handa, P., Kothari, S. P. & Wasley, C. (1989). The Relation Between the Return Interval and Betas. Implications for the Size Effect. *Journal of Financial Economics*, 23(1), 79–100. doi: 10.1016/0304-405X(89)90006-8

- Handa, P., Kothari, S. P. & Wasley, C. (1993, September). Sensitivity of Multivariate Tests of the Capital Asset-Pricing Model to the Return Measurement Interval. *Journal of Finance*, 48(4), 1543–1551.
- Härdle, W., Kerkycharian, G., Picard, D. & Tsybakov, A. (1998). *Wavelets, Approximation, and Statistical Applications. (Lecture Notes in Statistics, Volume 129)*. New York, NY: Springer.
- Hommes, C. H. (2006). Heterogeneous Agent Models in Economics and Finance. In L. Tesfatsion & L. Judd, Kenneth (Eds.), *Handbook of computational economics: Agent-based computational economics volume 2* (2nd ed., pp. 1109–1186). Amsterdam: Elsevier B.V.
- Iacobucci, A. (2003). *Spectral Analysis of Economic Time Series*.
- In, F. & Kim, S. (2006). Multiscale Hedge Ratio Between the Australian Stock and Futures Markets: Evidence From Wavelet Analysis. *Journal of Multinational Financial Management*, 16(4), 411–423. doi: 10.1016/j.mulfin.2005.09.002
- In, F. & Kim, S. (2013). *An Introduction to Wavelet Theory in Finance. A Wavelet Multiscale Approach*. Singapore: World Scientific Publishing.
- Jagannathan, R. & Ma, T. (2003). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraint Helps. *Journal of Finance*, 58(4), 1651–1684. doi: 10.1111/1540-6261.00580
- Jammazi, R. & Aloui, C. (2012). Crude Oil Price Forecasting: Experimental Evidence From Wavelet Decomposition and Neural Network Modeling. *Energy Economics*, 34(3), 828–841. doi: 10.1016/j.eneco.2011.07.018
- Kim, D. H. & Jeong, H. (2005). Systematic Analysis of Group Identification in Stock Markets. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 72(4), 1–8. doi: 10.1103/PhysRevE.72.046133
- Kirman, A. (2006). Heterogeneity in Economics. *Journal of Economic Interaction and Coordination*, 1(1), 89–117. doi: 10.1007/s11403-006-0005-8
- Kulkarni, V. & Deo, N. (2007). Correlation and Volatility in an Indian Stock Market: A Random Matrix Approach. *European Physical Journal B*, 60(1), 101–109. doi: 10.1140/epjb/e2007-00322-1
- Kumar, S. & Deo, N. (2012). Correlation and Network Analysis of Global Financial Indices. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*,

- 86(2), 1–32. doi: 10.1103/PhysRevE.86.026101
- Kwapien, J., Drozd, S. & Speth, J. (2003). Alternation of Different Fluctuation Regimes in the Stock Market Dynamics. *Physica A: Statistical Mechanics and its Applications*, 330(3-4), 605–621. doi: 10.1016/j.physa.2003.09.012
- Kwapien, J., Drozd, S. & Speth, J. (2004). Time Scales Involved in Emergent Market Coherence. *Physica A: Statistical Mechanics and its Applications*, 337(1-2), 231–242. doi: 10.1016/j.physa.2004.01.050
- Laloux, L., Cizeau, P., Bouchaud, J. P. & Potters, M. (1999). Noise Dressing of Financial Correlation Matrices. *Physical Review Letters*, 83(7), 1467–1470. doi: 10.1103/PhysRevLett.83.1467
- Laloux, L., Cizeau, P., Potters, M. & Bouchaud, J.-P. (2000, jul). Random Matrix Theory and Financial Correlations. *International Journal of Theoretical and Applied Finance*, 03(03), 391–397. doi: 10.1142/S0219024900000255
- Larsen, R. & Warne, R. T. (2010). Estimating Confidence Intervals for Eigenvalues in Exploratory Factor Analysis. *Behavior Research Methods*, 42(3), 871–876. doi: 10.3758/BRM.42.3.871
- LeBaron, B. (2006). Agent-Based Computational Finance. In L. Tesfatsion & K. L. Judd (Eds.), *Handbook of computational economics: Agent-based computational economics volume 2* (2nd ed., pp. 1187–1234). Amsterdam: Elsevier B.V.
- Ledoit, O. & Wolf, M. (2004). A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices. *Journal of Multivariate Analysis*, 88(2), 365–411. doi: 10.1016/S0047-259X(03)00096-4
- Levhari, D. & Levy, H. (1977). The Capital Asset Pricing Model and the Investment Horizon. *Review of Economics & Statistics*, 59(1), 92–104.
- Lin, A., Shang, P. & Zhou, H. (2014). Cross-Correlations and Structures of Stock Markets Based on Multiscale MF-DXA and PCA. *Nonlinear Dynamics*, 78(1), 485–494. doi: 10.1007/s11071-014-1455-5
- Liu, Y., Gopikrishnan, P., Cizeau, Meyer, Peng & Stanley, H. E. (1999). Statistical Properties of the Volatility of Price Fluctuations. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 60(2), 1390–1400. doi: 10.1103/PhysRevE.60.1390

- Lo, A. W. (2004, jan). The Adaptive Markets Hypothesis. *The Journal of Portfolio Management*, 30(5), 15–29. doi: 10.3905/jpm.2004.442611
- Lo, A. W. & MacKinlay, A. C. (1999). *A Non-Random Walk Down Wall Street*. Princeton, NJ: Princeton University Press. doi: 10.1111/1467-6419.00091
- Lo, Andrew, W. (2005). Reconciling Efficient Markets with Behavioral Finance: The Adaptive Markets Hypothesis. *Journal of Investment Consulting*, 7, 21–44.
- Loh, L. (2013). Co-Movement of Asia-Pacific with European and US Stock Market Returns: A Cross-Time-Frequency Analysis. *Research in International Business and Finance*, 29(1), 1–13. doi: 10.1016/j.ribaf.2013.01.001
- Longin, F. & Solnik, B. (2001). Extreme Correlation of International Equity Markets. *Journal of Finance*, 56(2), 649–676. doi: 10.1111/0022-1082.00340
- Lux, T. & Marchesi, M. (1999). Scaling and Criticality in a Stochastic Multi-Agent Model of a Financial Market. *Nature*, 397(6719), 498–500.
- Madaleno, M. & Pinho, C. (2012, jan). International Stock Market Indices Comovements: a New Look. *International Journal of Finance & Economics*, 17(1), 89–102. doi: 10.1002/ijfe.448
- Madaleno, M. & Pinho, C. (2014). Wavelet Dynamics for Oil-Stock World Interactions. *Energy Economics*, 45, 120–133. doi: 10.1016/j.eneco.2014.06.024
- Mallat, S. (2009). *A Wavelet Tour of Signal Processing* (3rd ed.). Burlington, MA: Elsevier Inc.
- Mallat, S. G. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674–693.
- Mandelbrot, B. B. (1962). Sur certains prix spéculatifs: faits empiriques et modèle basé sur les processus stables additifs de Paul Lévy. *Comptes Rendus*, 254, 3968–3970.
- Mandelbrot, B. B. (1963a). New Methods in Statistical Economics. *Journal of Political Economy*, 71(5), 421–440. doi: 10.1086/258792
- Mandelbrot, B. B. (1963b). The Variation of Certain Speculative Prices. *The Journal of Business*, 36(4), 394. doi: 10.1086/294632
- Mandelbrot, B. B. (1965). Une classe de processus stochastiques homothétiques

- à soi. Application à la loi climatologique de H. E. Hurst. *Comptes Rendus (Paris)*, 260, 3274–3277.
- Mandelbrot, B. B. (1967). The Variation of Some Other Speculative Prices. *The Journal of Business*, 40(4), 393–413.
- Mandelbrot, B. B. (1997). *Fractals and Scaling in Finance*. New York, NY: Springer.
- Mantegna, R. N. & Stanley, E. H. (2004). *An Introduction to Econophysics. Correlations and Complexity in Finance*. Cambridge: Cambridge University Press.
- Marčenko, V. A. & Pastur, L. A. (1967). Distribution of Eigenvalues for Some Sets of Random Matrices. *Mathematics of the USSR-Sbornik*, 1(4), 457–483. doi: 10.1070/sm1967v001n04abeh001994
- Masih, M., Alzahrani, M. & Al-Titi, O. (2010). Systematic Risk and Time Scales: New Evidence from an Application of Wavelet Approach to the Emerging Gulf Stock Markets. *International Review of Financial Analysis*, 19(1), 10–18. doi: 10.1016/j.irfa.2009.12.001
- Masset, P. (2008). Analysis of Financial Time-series using Fourier and Wavelet Methods. *Faculty of Economics and Social Science*, 1–36.
- Mehta, M. L. (2004). *Random Matrices* (3rd ed.). Boston: Academic Press.
- Mehta, M. L. & Dyson, F. J. (1963, may). Statistical Theory of the Energy Levels of Complex Systems. V. *Journal of Mathematical Physics*, 4(5), 713–719. doi: 10.1063/1.1704009
- Merton, R. C. (1980). On Estimating the Expected Return on the Market: An Exploratory Investigation. *Journal of Financial Economics*, 8(4), 323–361. doi: 10.1016/0304-405X(80)90007-0
- Michaud, R. O. (1989). The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal*, 45(1), 31–42. doi: 10.2469/faj.v45.n1.31
- Müller, U. a., Dacorogna, M. M., Davé, R. D., Olsen, R. B., Pictet, O. V. & von Weizsäcker, J. E. (1997). Volatilities of Different Time Resolutions . Analyzing the Dynamics of Market Components. *Journal of Empirical Finance*, 4(2-3), 213–239. doi: 10.1016/S0927-5398(97)00007-8
- Müller, U. a., Dacorogna, M. M., Davé, R. D., Pictet, O. V., Olsen, R. B. & Ward, J. R. (1993). *Fractals and Intrinsic Time: A Challenge to Econometricians*.

- Zurich. Retrieved from <http://finance.martinsewell.com/stylized-facts/scaling/\-Mueller-etal1993.pdf>
- Najeeb, S. F., Bacha, O. & Masih, M. (2015). Does Heterogeneity in Investment Horizons Affect Portfolio Diversification? Some Insights Using M-GARCH-DCC and Wavelet Correlation Analysis. *Emerging Markets Finance and Trade*, 51(1), 188–208. doi: 10.1080/1540496X.2015.1011531
- Nakayama, Y. & Iyetomi, H. (2009). Random Matrix Theory of Dynamical Cross Correlations in Financial Data. *Progress of Theoretical Physics Supplement*(179), 60–70. doi: 10.1143/PTPS.179.60
- Nekhili, R., Altay-Salih, A. & Gençay, R. (2002). Exploring Exchange Rate Returns at Different Time Horizons. *Physica A: Statistical Mechanics and its Applications*, 313(3-4), 671–682. doi: 10.1016/S0378-4371(02)00986-X
- Newbould, G. D. & Poon, P. S. (1993). The Minimum Number of Stocks Needed for Diversification. *Financial Practice and Education*, 3(2), 85–87.
- Nguyen, H. T., Tran, P. N. U. & Nguyen, Q. (2018). An Analysis of Eigenvectors of a Stock Market Cross-Correlation Matrix. In L. Anh, L. Dong, V. Kreinovich & N. Thach (Eds.), *Studies in computational intelligence* (pp. 504–513). Cham: Springer. doi: 10.1007/978-3-319-73150-6{\\_}40
- Oppenheim, A. V. & Schaffer, R. W. (2009). *Discrete-Time Signal Processing* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Orlov, V. & Äijö, J. (2015). Benefits of Wavelet-Based Carry Trade Diversification. *Research in International Business and Finance*, 34, 17–32. doi: 10.1016/j.ribaf.2014.11.002
- Osborne, M. F. M. (1959). Operations Research. *Operations Research*, 7(2), 145–173.
- Pan, R. K. & Sinha, S. (2007). Collective Behavior of Stock Price Movements in an Emerging Market. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 76(4). doi: 10.1103/PhysRevE.76.046116
- Percival, D. B. (1983). *The Statistics of Long Memory Processes* (Unpublished doctoral dissertation). University of Washington, Seattle.
- Percival, D. B. (1995). On Estimation of the Wavelet Variance. *Biometrika*, 82(3), 619–631. doi: 10.1214/aoms/1177698699

- Percival, D. B. & Mofjeld, H. O. (1997, sep). Analysis of Subtidal Coastal Sea Level Fluctuations Using Wavelets. *Journal of the American Statistical Association*, 92(439), 868–880. doi: 10.1080/01621459.1997.10474042
- Percival, D. B. & Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge, England: Cambridge University Press.
- Peters, E. E. (1994). *Fractal Market Analysis. Applying Chaos Theory to Investment and Economics*. New York, NY: Wiley.
- Peters, E. E. (1996). *Chaos and Order in the Capital Markets: A New View of Cycles, Prices, and Market Volatility*. New York: John Wiley & Sons.
- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A., Guhr, T. & Stanley, H. E. (2001). A Random Matrix Approach to Cross Correlations in Financial Data. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 65(6). doi: 10.1103/PhysRevE.65.066126
- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A., Guhr, T. & Stanley, H. E. (2002). Random Matrix Approach to Cross Correlations in Financial Data. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 65(6), 1–18. doi: 10.1103/PhysRevE.65.066126
- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. & Stanley, H. E. (1999). Universal and nonuniversal properties of cross correlations in financial time series. *Physical Review Letters*, 83(7), 1471–1474. doi: 10.1103/PhysRevLett.83.1471
- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. & Stanley, H. E. (2000). A Random Matrix Theory Approach to Financial Cross-Correlations. *Physica A: Statistical Mechanics and its Applications*, 287(3-4), 374–382. doi: 10.1016/S0378-4371(00)00376-9
- Priestley, M. B. (1981). *Spectral Analysis and Time Series*. London: Academic Press.
- Ramsey, J. B. (1999). The Contribution of Wavelets to the Analysis of Economic and Financial Data. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 357(1760), 2593–2606.
- Ramsey, J. B. (2002). Wavelets in Economics and Finance: Past and Future. *Studies in Nonlinear Dynamics & Econometrics*, 6(3), 1–27. doi: 10.2202/1558-3708

.1090

- Ramsey, J. B. & Lampart, C. (1998a). Decomposition of Economic Relationships by Timescale Using Wavelets. *Macroeconomic Dynamics*, 2(May), 49–71.
- Ramsey, J. B. & Lampart, C. (1998b). The Decomposition of Economic relationships by time scale using wavelets: Expenditure and income. *Studies in Nonlinear Dynamics & Econometrics*, 3(1), 23–42.
- Ramsey, J. B. & Zhang, Z. (1997). The Analysis of Foreign Exchange Data Using Waveform Dictionaries. *Journal of Empirical Finance*, 4(4), 341–372. doi: 10.1016/S0927-5398(96)00013-8
- Ranta, M. (2010). *Wavelet Multiresolution Analysis of Financial Time Series* (Ph.D. Thesis). Wasaensis.
- Reboredo, J. C. & Rivera-Castro, M. A. (2013). A Wavelet Decomposition Approach to Crude Oil Price and Exchange Rate Dependence. *Economic Modelling*, 32(1), 42–57. doi: 10.1016/j.econmod.2012.12.028
- Rojkova, V. & Kantardzic, M. (2007). Analysis of Inter-Domain Traffic Correlations: Random Matrix Theory Approach. *CoRR*, abs/0706.2520(May). Retrieved from <http://arxiv.org/abs/0706.2520>
- Rosenow, B., Plerou, V., Gopikrishnan, P. & Stanley, H. E. (2002). Portfolio Optimization and the Random Magnet Problem. *Europhysics Letters*, 59(4), 500–506. doi: 10.1209/epl/i2002-00135-4
- Routledge, B. R. (1999). Adaptive Learning in Financial Markets. *Review of Financial Studies*, 12(5), 1165–1202. doi: 10.1093/rfs/12.5.1165
- Routledge, B. R. (2001). Genetic Algorithm Learning to Choose and Use Information. *Macroeconomic Dynamics*, 5(2), 303–325. doi: 10.1017/S1365100501019083
- Rua, A. & Nunes, L. C. (2009). International Comovement of Stock Market Returns: A Wavelet Analysis. *Journal of Empirical Finance*, 16(4), 632–639. doi: 10.1016/j.jempfin.2009.02.002
- Rua, A. & Nunes, L. C. (2012). A Wavelet-Based Assessment of Market Risk: The Emerging Markets Case. *Quarterly Review of Economics and Finance*, 52(1), 84–92. Retrieved from <http://dx.doi.org/10.1016/j.qref.2011.12.001> doi: 10.1016/j.qref.2011.12.001

- Sandoval, L., Bruscato, A. & Venezuela, M. K. (2012, January). *Building Portfolios of Stocks in the São Paulo Stock Exchange Using Random Matrix Theory* (Papers No. 1201.0625). arXiv.org. doi: 10.1016/j.physa.2014.05.006
- Sandoval, L. & Franca, I. D. P. (2012). Correlation of Financial Markets in Times of Crisis. *Physica A: Statistical Mechanics and its Applications*, 391(1-2), 187–208. doi: 10.1016/j.physa.2011.07.023
- Sengupta, A. M. & Mitra, P. P. (1999, sep). Distributions of Singular Values for Some Random Matrices. *Physical Review E*, 60(3), 3389–3392. doi: 10.1103/PhysRevE.60.3389
- Shah, A., Deo, M. & King, W. (2016). What Econo-Physics Can Tell Us About Korean Equity Market Co-Movements? *Journal of Economic Studies*, 43(4), 549–573. doi: 10.1108/JES-04-2015-0058
- Sharkasi, A., Crane, M., Ruskin, H. J. & Matos, J. A. (2006). The Reaction of Stock Markets to Crashes and Events: A Comparison Study Between Emerging and Mature Markets Using Wavelet Transforms. *Physica A: Statistical Mechanics and its Applications*, 368(2), 511–521. doi: 10.1016/j.physa.2005.12.048
- Sheng, Y. (2000). *Wavelet Transform* (2nd ed.; A. D. Poularikas, Ed.). Boca Raton: CRC Press. doi: doi:10.1201/9781420036756.ax5
- Sinha, S., Chatterjee, A., Chakraborti, A. & Chakrabarti, B. K. (2010). *Econophysics*. Weinheim: Wiley-VCH.
- Torrence, C. & Compo, G. P. (1998). A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, 79(1), 61–78. doi: 10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2
- Torrence, C. & Webster, P. J. (1999). Interdecadal Changes in the ENSO-Monsoon System. *Journal of Climate*, 12(8), 2679–2690. doi: 10.1175/1520-0442(1999)012<2679:ICITEM>2.0.CO;2
- Tumminello, M., Di Matteo, T., Aste, T. & Mantegna, R. N. (2007). Correlation Based Networks of Equity Returns Sampled at Different Time Horizons. *European Physical Journal B*, 55(2), 209–217. doi: 10.1140/epjb/e2006-00414-4
- Utsugi, A., Ino, K. & Oshikawa, M. (2004). Random Matrix Theory Analysis of Cross Correlations in Financial Markets. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 70(2), 11. doi:

10.1103/PhysRevE.70.026110

- Vacha, L. & Barunik, J. (2012). Co-Movement of Energy Commodities Revisited: Evidence From Wavelet Coherence Analysis. *Energy Economics*, 34(1), 241–247. doi: 10.1016/j.eneco.2011.10.007
- Wang, G. J., Xie, C. & Chen, S. (2017). Multiscale Correlation Networks Analysis of the US Stock Market: A Wavelet Analysis. *Journal of Economic Interaction and Coordination*, 12(3), 561–594. doi: 10.1007/s11403-016-0176-x
- Wang, G. J., Xie, C., Chen, S., Yang, J. J. & Yang, M. Y. (2013). Random Matrix Theory Analysis of Cross-Correlations in the US Stock Market: Evidence From Pearson's Correlation Coefficient and Detrended Cross-Correlation Coefficient. *Physica A: Statistical Mechanics and its Applications*, 392(17), 3715–3730. doi: 10.1016/j.physa.2013.04.027
- Weron, A. & Weron, R. (2000). Fractal Market Hypothesis and Two Power-Laws. *Chaos, solitons and fractals*, 11(1), 289–296. doi: 10.1016/S0960-0779(98)00295-1
- Whitcher, B., Guttorp, P. & Percival, D. B. (1999). *Mathematical Background for Wavelet Estimators of Cross-Covariance and Cross-Correlation* (Tech. Rep. No. 38). Seattle: National Research Center for Statistics and the Environment.
- Whitcher, B. J. (1998). *Assessing Nonstationary Time Series Using Wavelets* (Doctoral dissertation, University of Washington). Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.-29.5644{\&}rep=rep1{\&}type=pdf>
- Wigner, E. P. (1951a, jan). On a Class of Analytic Functions from the Quantum Theory of Collisions. *The Annals of Mathematics*, 53(1), 36. doi: 10.2307/1969342
- Wigner, E. P. (1951b). On the Statistical Distribution of the Widths and Spacings of Nuclear Resonance Levels. *Mathematical Proceedings of the Cambridge Philosophical Society*, 47(4), 790–798. doi: 10.1017/S0305004100027237
- Wigner, E. P. (1955, nov). Characteristic Vectors of Bordered Matrices With Infinite Dimensions. *The Annals of Mathematics*, 62(3), 548. doi: 10.2307/1970079
- Wilcox, D. & Gebbie, T. (2004). On the Analysis of Cross-Correlations in South African Market Data. *Physica A: Statistical Mechanics and its Applications*,

344(1-2), 294–298. doi: 10.1016/j.physa.2004.06.138

Wilcox, D. & Gebbie, T. (2007). An Analysis of Cross-Correlations in an Emerging Market. *Physica A: Statistical Mechanics and its Applications*, 375(2), 584–598. doi: 10.1016/j.physa.2006.10.030

Xu, Z. & Gençay, R. (2003). Scaling, Self-Similarity and Multifractality in FX Markets. *Physica A: Statistical Mechanics and its Applications*, 323, 578–590. doi: 10.1016/S0378-4371(03)00030-X

Zheng, Z., Podobnik, B., Feng, L. & Li, B. (2012). Changes in Cross-Correlations as an Indicator for Systemic Risk. *Scientific Reports*, 2, 1–8. doi: 10.1038/srep00888

# Curriculum Vitae

Christian Vial, born on February 24, 1989, in Zug, Switzerland

## Education

- 2013 - 2019      **University of St. Gallen**, Switzerland  
Doctoral Student in Finance
- 2010 - 2013      **University of St. Gallen**, Switzerland  
Master of Arts in Banking and Finance (M.A. HSG)
- 2007 - 2010      **University of St. Gallen**, Switzerland  
Bachelor of Arts in Business Administration (B.A. HSG)
- 2003 - 2007      **Gymnasium Immensee**, Switzerland  
Grammar School

## Work Experience

- 2020 -            **Finreon**  
*Zürich, Switzerland*  
Head of Innovation & Research
- 2016 - 2017      **Kraus Partner Investment Solutions**  
*Zürich, Switzerland*  
Senior Quantitative Analyst, GTAA
- 2013 - 2016      **Vescore** (former 1741 Asset Management)  
*Zürich, Switzerland*  
Quantitative Analyst, Indexing & Quantitative Research
- 2013              **Notenstein Private Bank**  
*St. Gallen, Switzerland*  
Quantitative Analyst, Quantitative Strategies and Research
- 2011              **Credit Suisse Asset Management**  
*Zürich, Switzerland*  
Internship, Multi Asset Class Solutions Investment Risk