



## Deliverable information

WP NO.	WP5
DEL. REL. NO.	D5.2
DEL. NO.	D14
TITLE	Repository of use cases and slides in artificial intelligence
DESCRIPTION	Repository of use cases and slides (including paper, data and code) in artificial intelligence shared during the Suptech and RegTech workshops
NATURE	Website
EST. DEL. DATE	31 May 2020

## Document information

DATE	06 July 2020
WRITTEN BY	Paolo Pagnottoni
APPROVED BY	Paolo Giudici

## **Actions and achieved results**

The University of Pavia, with the support of all the project partners, has developed the repository of use cases and slides for the Artificial Intelligence part. The repository includes research and teaching material, in the form of use cases, regarding AI, Robo Advice and Market Risk. The material has been developed first gathering contributions from the consortium and, then, improving them according to the received feedbacks from all project's participants: partners, regulators, supervisors, banks, fintechs and international advisors.

The material has been prepared, shared and disseminated during:

1. The Regtech sessions (one in Winthertur and one in Vienna) in which FinTech companies and banks had the chance to replicate the use cases, with the provided software code;
2. The Suptech sessions, in which each partner discussed the developed use cases with their national Supervisors;

The material was firstly prepared by the consortium partners under the form of proposal. Secondly, the proposal has been then evaluated by the WP 3 leader, together with the Project Coordinator. Thirdly, consortium partners submitted their feedbacks on the material which has been consequently updated. Fourthly partners from the consortium have collected feedbacks from their national Supervisors during the Suptech sessions organised in all European Union countries, and material has been updated. Fifthly, we have gathered feedbacks from participants during the Regtech sessions, which comprised mostly IT developers and risk management experts from FinTechs and banks, and updated the material. Finally, the material has been finalized taking into account all feedbacks. All intermediate outputs and the final one (attached to this report) have been made available to all the stakeholders on the project web platform.

Specifically, the final repository contains:

1. Final open source papers regarding each use case
2. Final codes related to each use case
3. Final data required for the empirical application of each use case
4. Final set of slides, gathering all use cases into one comprehensive training session.

From a technical viewpoint, the use cases selected are three. A summary description of each of them is reported below.

### **USE CASE 1: NETWORK MODELS TO ENHANCE AUTOMATED CRYPTOCURRENCY PORTFOLIO MANAGEMENT**

(GIUDICI P.- UNIPV, PAGNOTTONI P.- UNIPV, POLINESI G.- UNIVERSITA POLITECNICA DELLE MARCHE)

Published in: Front. Artif. Intell., 24 April 2020 | <https://doi.org/10.3389/frai.2020.00022>

The usage of cryptocurrencies, together with that of financial automated consultancy, is widely spreading in the last few years. However, automated consultancy services are not yet exploiting the potentiality of this nascent market, which represents a class of innovative financial products that can be proposed by robo-advisors. For this reason, we propose a novel approach to build efficient portfolio allocation strategies involving volatile financial instruments such as cryptocurrencies. In other words, we develop an extension

of the traditional Markowitz model which combines Random Matrix Theory and network measures, in order to achieve portfolio weights enhancing portfolios' risk-return profiles. The results show that overall our model overperforms several competing alternatives, maintaining a relatively low level of risk.

Keywords: cryptocurrencies, correlation networks, network centrality, portfolio optimization, random matrix theory, minimal spanning tree

## USE CASE 2: SENTIMENT ANALYSIS OF EUROPEAN BONDS 2016–2018

(SCHWENDNER P.- ZHAW, SCHÜLE M.- ZHAW AND HILLEBRAND M.- ESM LUXEMBOURG)

Published in: Front. Artif. Intell., 15 October 2019 | <https://doi.org/10.3389/frai.2019.00020>

We revisit the discussion of market sentiment in European sovereign bonds using a correlation analysis toolkit based on influence networks and hierarchical clustering. We focus on three case studies of political interest. In the case of the 2016 Brexit referendum, the market showed negative correlations between core and periphery only in the week before the referendum. Before the French presidential elections in 2017, the French bond spread widened together with the estimated Le Pen election probability, but the position of French bonds in the correlation blocks did not weaken. In summer 2018, during the budget negotiations within the new Italian coalition, the Italian bonds reacted very sensitively to changing political messages but did not show contagion risk to Spain or Portugal for several months. The situation changed during the week from October 22 to 26, as a spillover pattern of negative sentiment also to the other peripheral countries emerged.

Keywords: sovereign bonds, contagion, sentiment, European sovereign bond crisis, correlation, correlation influence, networks

## USE CASE 3: EXPLAINABLE AI IN FINTECH RISK MANAGEMENT

(BUSSMANN N.- FIRAMIS and UNIPV, GIUDICI P.- UNIPV, MARINELLI D. - FIRAMIS, PAPENBROCK J.- FIRAMIS)

Published in: Front. Artif. Intell., 24 April 2020 | <https://doi.org/10.3389/frai.2020.00026>

The paper proposes an explainable AI model that can be used in fintech risk management and, in particular, in measuring the risks that arise when credit is borrowed employing peer to peer lending platforms. The model employs Shapley values, so that AI predictions are interpreted according to the underlying explanatory variables. The empirical analysis of 15,000 small and medium companies asking for peer to peer lending credit reveals that both risky and not risky borrowers can be grouped according to a set of similar financial characteristics, which can be employed to explain and understand their credit score and, therefore, to predict their future behavior.

Keywords: credit risk management, explainable AI, financial technologies, peer to peer lending, logistic regression, predictive models

## Attached documents

The material concerning the repository in Artificial Intelligence is attached as follows:

1. Slides used during the Suptech workshop
2. The three open access papers chosen as use cases
3. Screenshots of the R codes related to the chosen use cases

Please note that – given the deliverable format constraints - full replication codes and the datasets are attached in .zip folders.



This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 825215. All material presented here reflects only the authors' view.

The European Commission is not responsible for any use that may be made of the information it contains.

# Artificial Intelligence Use Cases

FIN-TECH HO2020

July 3, 2020

# Table of contents

Use Case I: Convergence and Divergence in European Bond Correlations (Peter Schwendner, Martin Schüle and Martin Hillebrand, ZHAW and European Stability Mechanism)

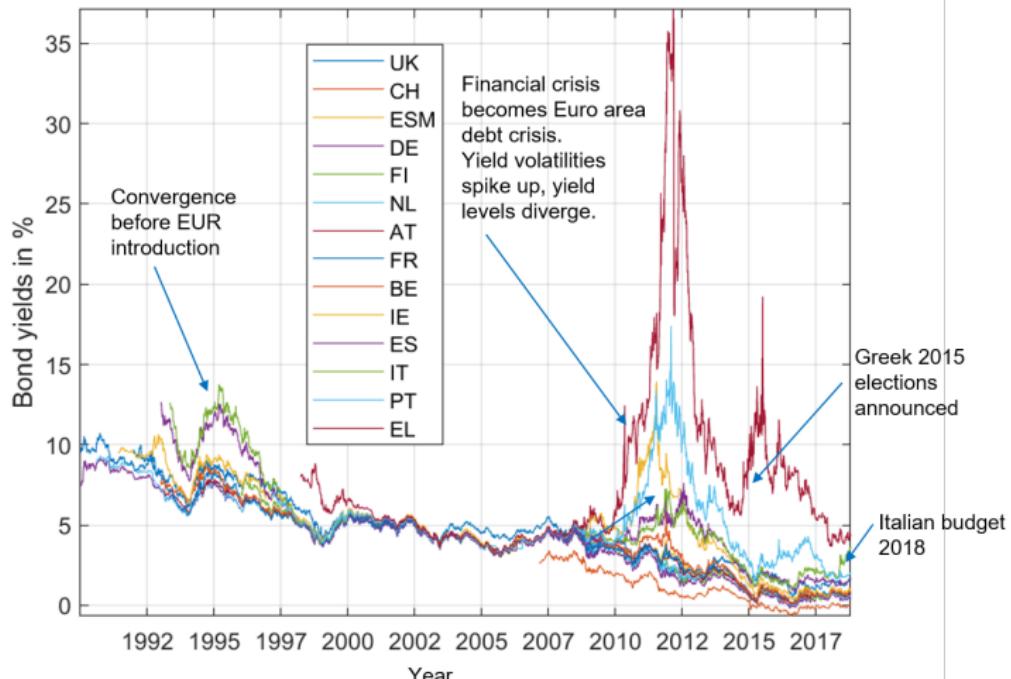
Use Case II: Network models to enhance automated cryptocurrency portfolio management (Paolo Giudici, Paolo Pagnottoni, Gloria Polinesi, UNIPV and POLITECNICA)

Use Case III: eXplainable AI in credit scoring and portfolio construction (Dimitri Marinelli, Jochen Papenbrock, Niklas Bussmann, Paolo Giudici; FIRAMIS and UNIPV)

Use Case I: Convergence and Divergence in European Bond Correlations (Peter Schwendner, Martin Schüle and Martin Hillebrand, ZHAW and European Stability Mechanism)

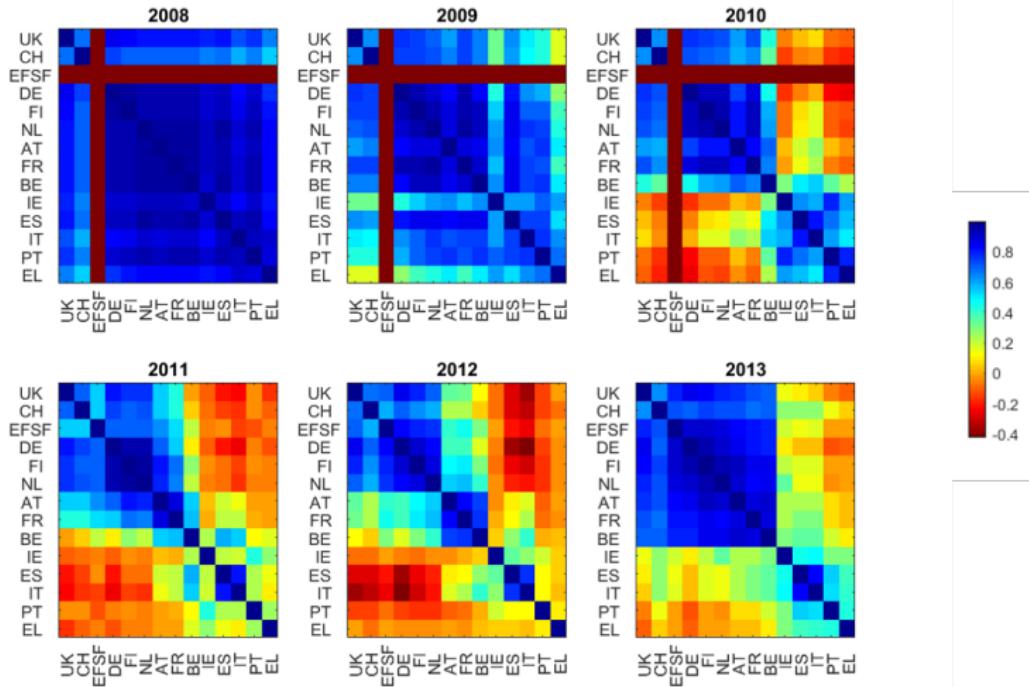
# European Bond Yields (daily Bloomberg data)

- ▶ Euro convergence for bonds yields during end of 90s.
- ▶ Wide spreads during European sovereign debt crisis 2010-2012.
- ▶ Since 2015, bond spreads primarily signal political divergence.



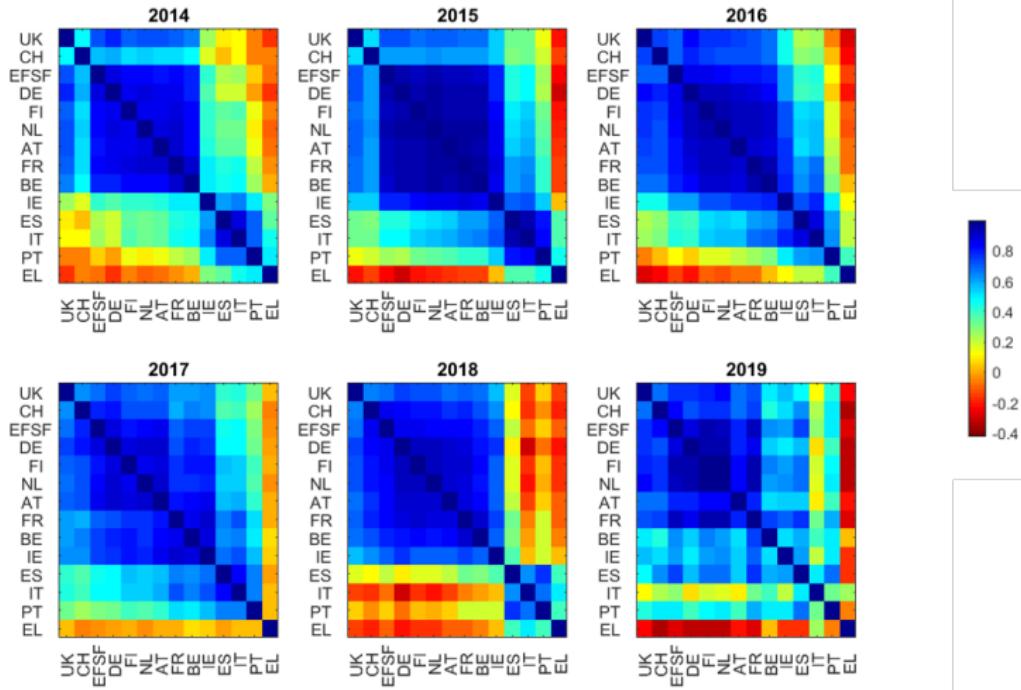
# European Bond Return Correlations 2008 - 2013

- Containment of the 2010 sovereign bond crisis



# European Bond Return Correlations 2014-2019

- ▶ From financial crisis to political divergence



# Problems with correlations

- ▶ They are unstable in time
- ▶ Common factors may lead to spurious correlations
- ▶ Too many links: each market is correlated to any other market. Who is driving what?
- ▶ Idea: "Correlation influence" shows driving factors of correlations. Bootstrap resampling to simulate statistical noise in return blocks of random length ("wild bootstrap").

Original return matrix

UK	CH	...	PT	EL
1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5
6	6	6	6	6
7	7	7	7	7

One of 10.000 bootstrap resamples

UK	CH	...	PT	EL
3	3	3	3	3
5	5	5	5	5
2	2	2	2	2
3	3	3	3	3
6	6	6	6	6
1	1	1	1	1
2	2	2	2	2

## Correlation influence Network

- ▶ The partial correlation measure is defined as

$$\rho_{ij:k} = \frac{C_{ij} - C_{ik} C_{kj}}{\sqrt{1 - C_{ik}^2} \sqrt{1 - C_{kj}^2}}. \quad (1)$$

- ▶ Correlation influence is defined as

$$d_{i,j:k} = C_{ij} - \rho_{ij:k}. \quad (2)$$

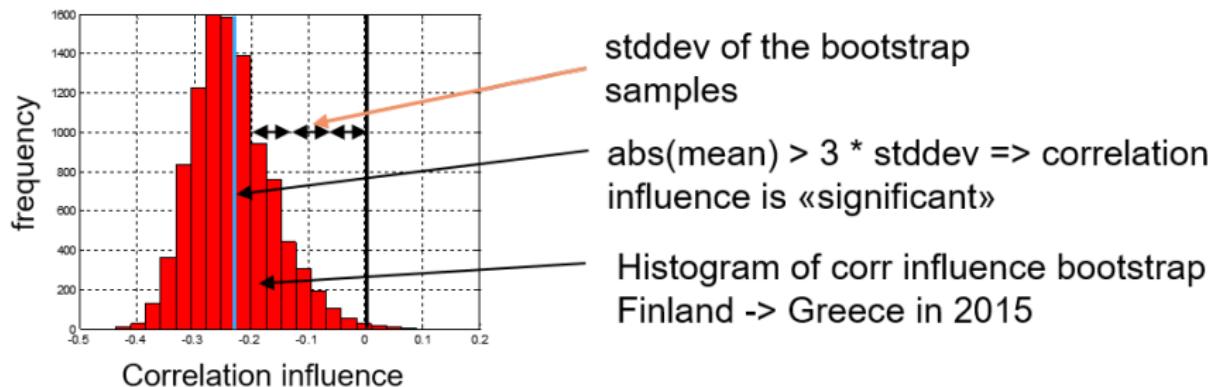
- ▶ The average correlation influence is defined as

$$d_{i:k} = \text{mean}(d_{i,j:k} \mid j \neq i, k). \quad (3)$$

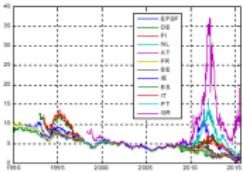
This is a directed arrow from market  $k$  pointing to market  $i$ .

## Bootstrap filter

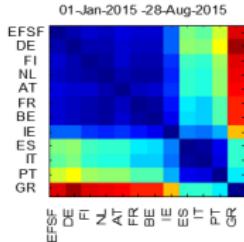
- ▶ For each resample, we compute the average correlation influence matrix.
- ▶ The standard deviation across all resamples is a measure for the noise in the correlation influence.
- ▶ We filter out correlation influences with a threshold of three standard deviations.



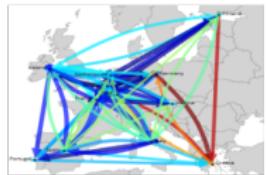
# Overview: Generate Filtered Correlation Influence Network



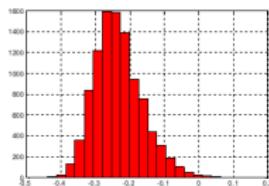
Bond yield time series



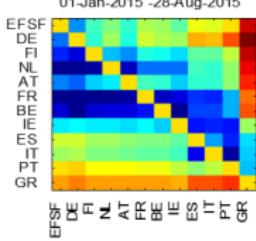
Correlation matrix of yield changes



Filtered influence network



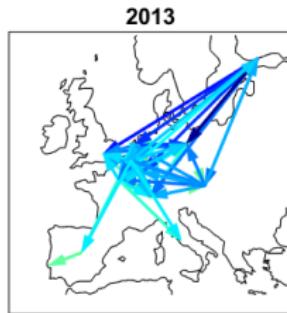
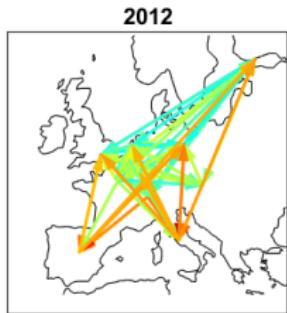
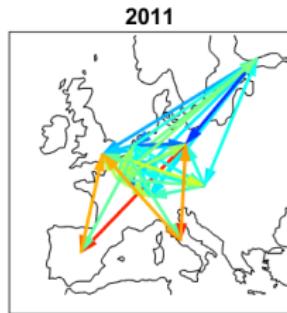
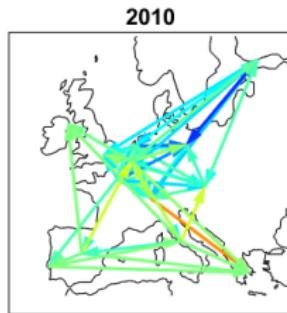
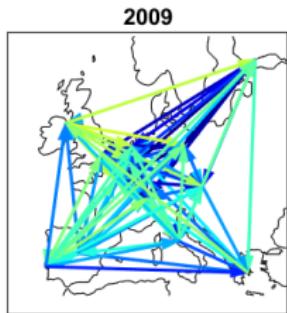
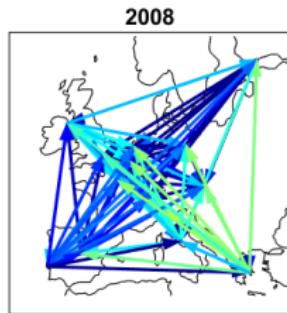
Bootstrap filter



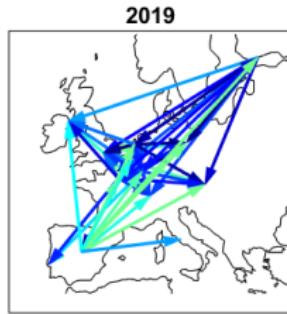
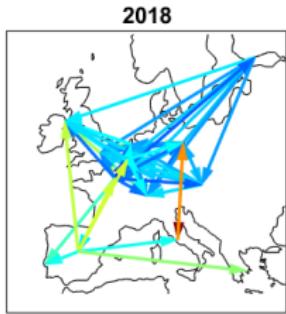
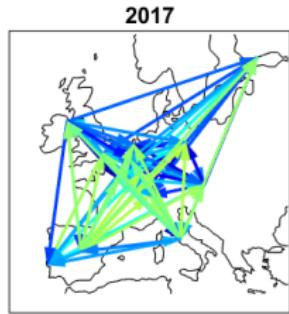
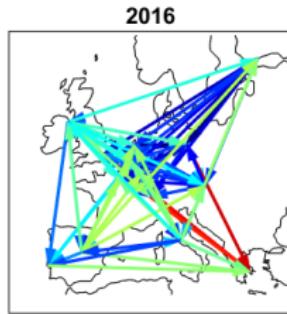
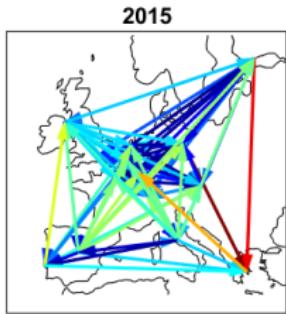
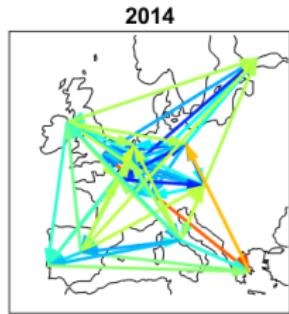
Correlation influence

Positive correlation influences: **blue arrows**  
Negative correlation influences: **red arrows**

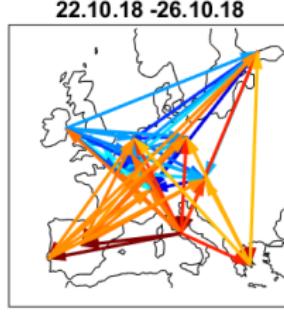
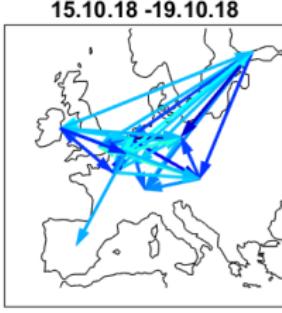
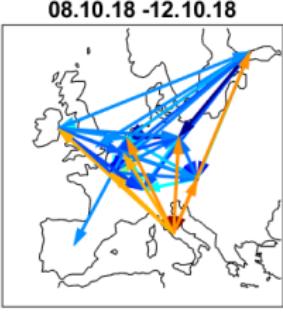
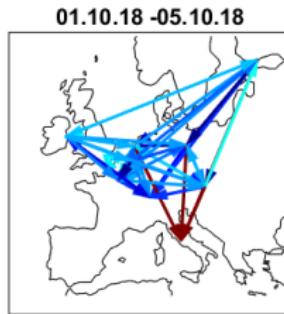
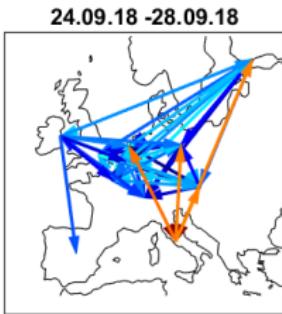
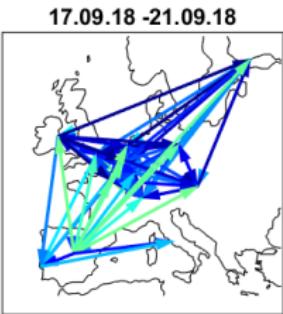
# Filtered Correlation Influence Networks 2008 - 2013



# Filtered Correlation Influence Networks 2014 - 2019

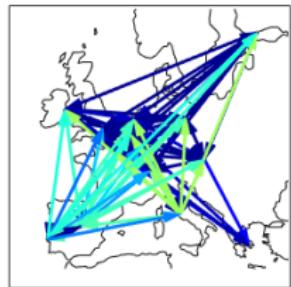


# Filtered Correlation Influence Networks October 2018

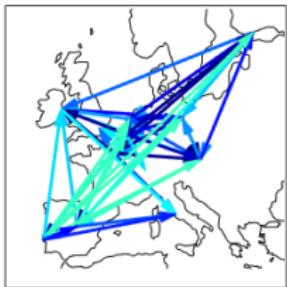


# Filtered Correlation Influence Networks August 2019

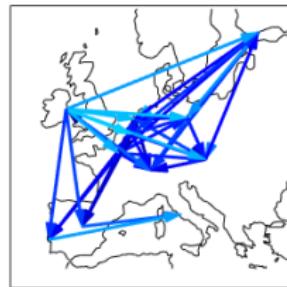
22.07.19 -26.07.19



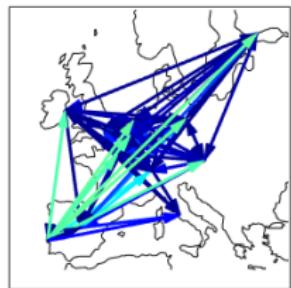
29.07.19 -02.08.19



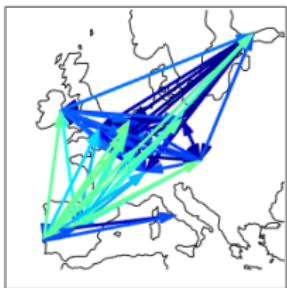
05.08.19 -09.08.19



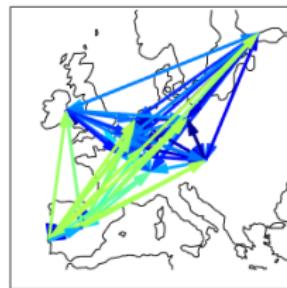
12.08.19 -16.08.19



19.08.19 -23.08.19

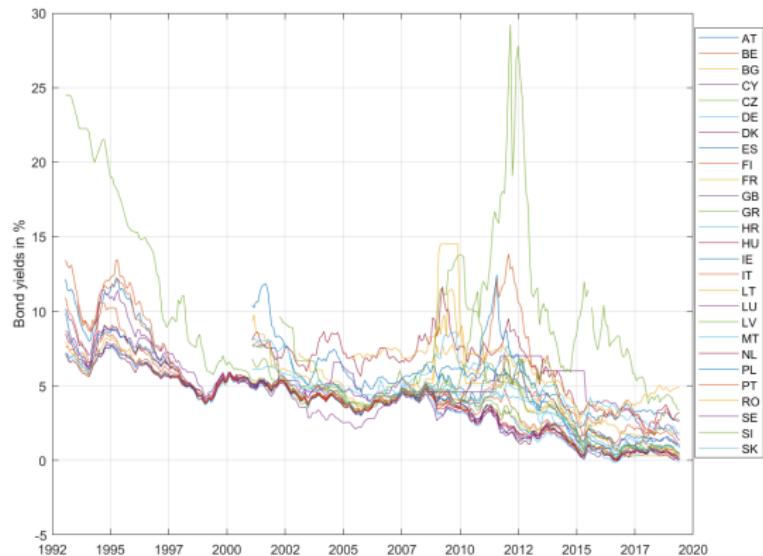


26.08.19 -30.08.19



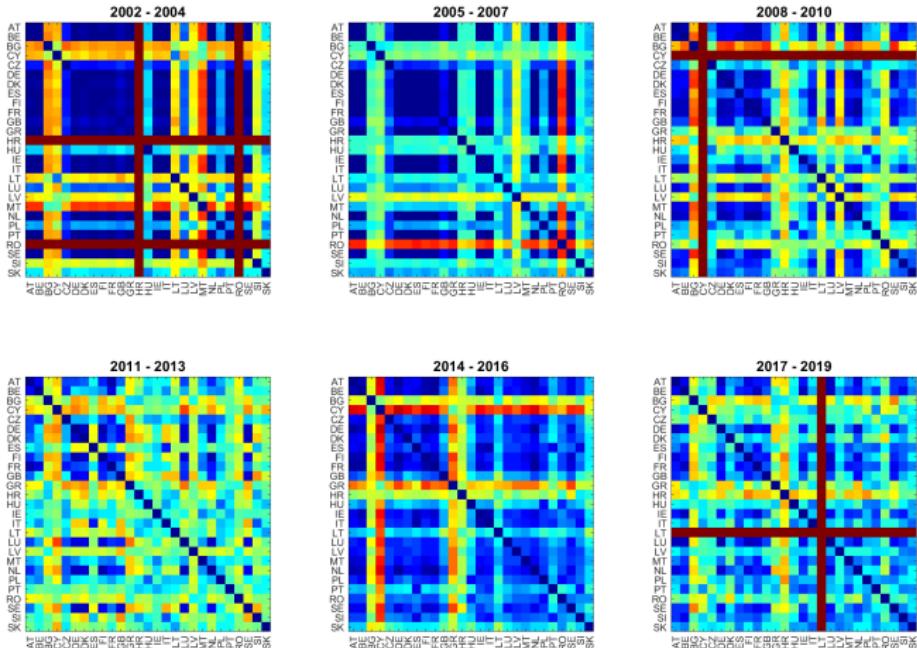
# Publicly available European Bond Yield data

- ▶ Source: ECB <https://sdw.ecb.europa.eu>
- ▶ Only monthly, but 27 EU countries (all but Estonia)



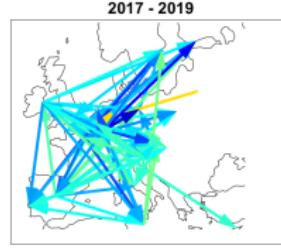
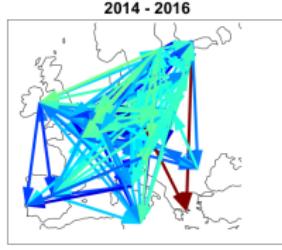
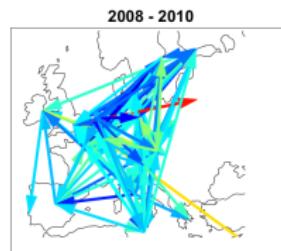
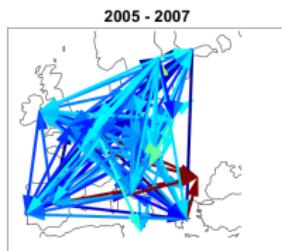
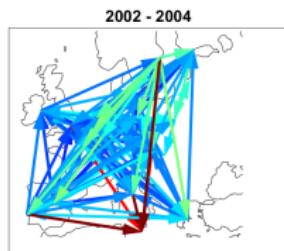
# European Bond Return Correlations 2002 - 2019

- We define 3-year-windows as we only have monthly data



# Filtered Correlation Influence Networks 2002 - 2019

- ▶ Also with monthly data, the networks replicate the core-periphery dynamics



## Conclusions

- ▶ Since 2010, European bonds cluster into core and periphery groups according to their return correlations. We use filtered correlation influence networks to show the most significant drivers of convergence and divergence.
- ▶ During the European sovereign debt crisis 2010 - 2012, negative correlation influences between the core and periphery groups are the dominating force. Since 2013, the situation improved a lot.
- ▶ In 2015 during the negotiations between Greece and the Eurogroup and in 2018 during the Italian budget negotiations, the warning signals of negative correlation influences reappeared for short periods, although the absolute level of spreads is substantially smaller than during 2010 - 2012.
- ▶ The findings point to markets becoming more politically driven.

Full paper: ESM Working Paper #8 and JNTF (2015), "Sentiment Analysis of European Bonds 2016 - 2018" , Frontiers in AI (2019).

Use Case II: Network models to enhance  
automated cryptocurrency portfolio management  
(Paolo Giudici, Paolo Pagntonni, Gloria Polinesi,  
UNIPV and POLITECNICA)

## Robot advisors, intro

- ▶ **FinTech** innovations are increasing exponentially, for the evolving technology on the supply side and for the shifting of consumer preferences on the demand side
- ▶ The total masses managed by the automatic consultancy are estimated around 980 billion dollars in 2019, and 2,552 billion in 2023

# Robot advisors and financial automation, Pros&Cons

- ▶ **Advantages:**

- ▶ Improved financial inclusion
- ▶ Lower fees
- ▶ High speed of service
- ▶ Customized user experience

- ▶ **Disadvantages:**

- ▶ User may not understand portfolio construction
- ▶ Portfolio models may be too simple
- ▶ Contagion between asset returns increases
- ▶ Portfolio allocation may not be complaint with investors' risk profile

## Our contribution

- ▶ Build **similarity network models** from the available asset return data
- ▶ Models that can incorporate multiple correlations (contagion) between asset returns in portfolio allocation.
- ▶ The ultimate goal is to improve portfolio allocation and risk compliance, taking systemic risk into account

Two main original contributions

- ▶ We extend the application of similarity networks from stock returns to **Exchange Traded Fund** returns
- ▶ We propose **an extension to Markowitz' portfolio allocation** that takes network centrality and, therefore, contagion, explicitly into account

## The Random Matrix approach

- ▶ RMT separates the “sistematic” part of a signal embedded into a return correlation matrix from the “noise”
- ▶ Tests the eigenvalues of the correlation matrix:  
 $\lambda_k < \lambda_{k+1}; k = 1, \dots, n$ , against the null hypothesis that they are from a random Wishart matrix  $\mathbf{R} = \frac{1}{T} \mathbf{A} \mathbf{A}^T$

Let  $r_i$ , for  $i = 1, \dots, n$ , be a time series of **Cryptocurrency returns** and  $\mathbf{C}$  be their correlation matrix. The RMT matrix is given by:

$$\mathbf{C}^* = \mathbf{V} \mathbf{L} \mathbf{V}^T, \quad (4)$$

where  $\mathbf{V}$  is the eigenvector matrix and

$$\mathbf{L} = \begin{cases} 0 & \lambda_i < \lambda_+ \\ \lambda_i & \lambda_i \geq \lambda_+ \end{cases}$$

## Similarity Network

- ▶ In a similarity network **nodes** represent asset returns and **edges** the distance between adjacent nodes.
- ▶ There exist different metrics to build **distances** between nodes: we apply the Euclidean distance

$$d_{ij} = \sqrt{2 - 2c'_{ij}},$$

- ▶ There exist different algorithms to simplify a similarity network: we apply the **Minimum Spanning Tree**, that reduces the number of edges from  $N*(N-1)/2$  to  $N-1$ .
- ▶ In the MST, at each step, two cluster nodes  $I_i$  and  $I_j$  are merged into a single cluster if:

$$d(I_i, I_j) = \min \{d(I_i, I_j)\}$$

with the distance between clusters being defined as:

$$d(I_i, I_j) = \min \{d_{rq}\}$$

with  $r \in I_i$  and  $q \in I_j$ .

## Centrality measures

- ▶ To measure the importance of each node, we can use the **eigenvector centrality**.
- ▶ The importance of a node depends on the importance of the nodes to which it is connected:

$$x_i = \frac{1}{\lambda} \sum_{j=1}^N \hat{d}_{i,j} x_j \quad (5)$$

## Portfolio Construction

- ▶ Differently from previous works which employ centrality measures as an alternative measure of diversification risk, we extend Markowitz' approach using RMT and MST in the optimisation function itself:

$$\min_{\mathbf{w}} \mathbf{w}^T \mathbf{C}^* \mathbf{w} + \gamma \sum_{i=1}^n x_i w_i$$

subject to

$$\begin{cases} \sum_{i=1}^n w_i = 1 \\ \mu_P \geq \frac{\sum_{i=1}^n \mu_i}{n} \\ w_i \geq 0 \end{cases}$$

- ▶ A high risk propensity (represented by a high value of  $\gamma$ ) translates in a portfolio composed by more systemically risky assets, that lay in the central body of the network, avoiding isolated cryptocurrencies.

## Application

- ▶ The data contains 10 time series of returns referred to cryptocurrencies traded over the period 14 September 2017 - 17 October 2019 (764 daily observations)
- ▶ Cryptocurrencies were selected in terms of market capitalization
- ▶ Portfolio returns are computed using the last month of each time window
- ▶ We use eleven months of observations as a look-back period computing asset centrality and the consequent portfolio weights
- ▶ Then we calculate the return of each portfolio over the next month rebalancing cryptocurrencies with the retrieved weights. Finally we connect each monthly portfolio performances from January 2018 to October 2019

## Summary statistics

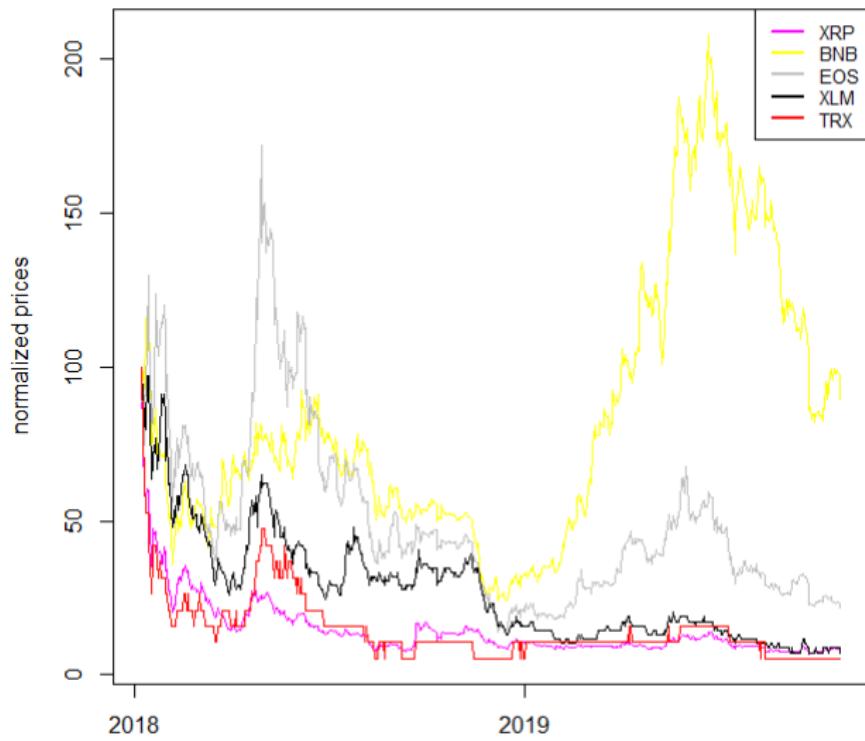
	mean	std.	kurtosis
BTC	0.0009	0.04	3.35
ETH	-0.0007	0.05	2.90
XRP	0.0004	0.07	15.73
USDT	0.0000	0.01	4.28
BCH	-0.0011	0.08	6.47
LTC	-0.0003	0.06	8.02
BNB	0.0033	0.07	7.74
EOS	0.0017	0.07	3.93
XLM	0.0021	0.10	26.19
TRX	0.0021	0.15	13.15

Cryptocurrency summary statistics over the period 14 September 2017 - 17 October 2019

# Prices - I



## Prices - II



# MST networks

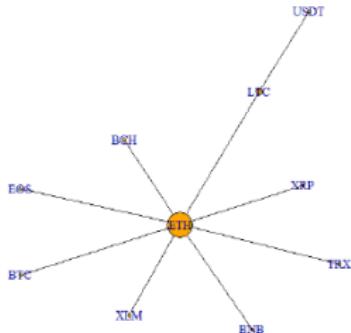


Figure 1: MST September 2017- January 2018. The figure shows the MST representation relative to the period of the speculative bubble.

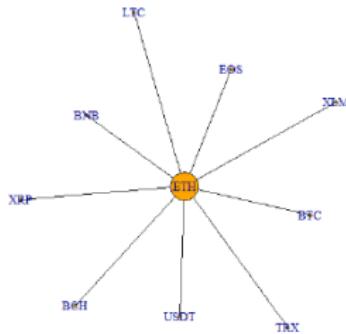


Figure 2: MST June 2019- October 2019. The figure shows the MST relative to the period June 2019- October 2019.

# Portfolio Results - I

Period	CRIX	GM	EW	CM	NW	$\gamma = 0.005$	$\gamma = 0.025$	$\gamma = 0.05$	$\gamma = 0.15$	$\gamma = 0.7$	$\gamma = 1$
Jan-2018	-0.14	-0.13	-0.16	0.04	-0.22	-0.21	-0.26	-0.27	-0.36	-0.43	-0.43
May-2018	-0.67	-0.62	-0.60	-0.12	-0.79	-0.78	-0.73	-0.66	-0.83	-1.08	-1.10
Sep-2018	-1.37	-1.37	-1.43	-0.88	-0.83	-1.02	-1.24	-1.23	-1.40	-1.60	-1.64
Jan-2019	-1.85	-1.78	-1.78	-1.32	-0.87	-1.50	-1.86	-1.98	-2.19	-2.29	-2.31
May-2019	-1.35	-1.25	-1.27	-1.01	-0.74	-1.22	-1.33	-1.29	-1.44	-1.55	-1.57
Sep-2019	-0.99	-1.45	-1.49	-1.02	-0.54	-1.19	-1.34	-1.44	-1.86	-2.13	-2.15

Cumulative profit & losses

## Portfolio Results - II

Period	GM	EW	CM	NW	$\gamma = 0.005$	$\gamma = 0.025$	$\gamma = 0.05$	$\gamma = 0.15$	$\gamma = 0.7$	$\gamma = 1$
Jan-2018	0.74	0.75	0.63	0.64	0.69	0.77	0.79	0.78	0.77	0.99
May-2018	0.73	0.75	0.95	0.83	0.74	0.77	0.83	0.87	0.87	0.55
Sep-2018	0.81	0.84	0.87	0.61	0.80	0.75	0.76	0.80	0.80	0.48
Jan-2019	1.16	1.11	1.47	1.24	1.34	1.36	1.39	1.40	1.40	1.26
May-2019	0.80	0.80	1.05	0.97	0.93	0.84	0.75	0.72	0.72	0.98
Sep-2019	0.75	0.78	1	1.14	0.43	0.38	0.38	0.38	0.37	0.78

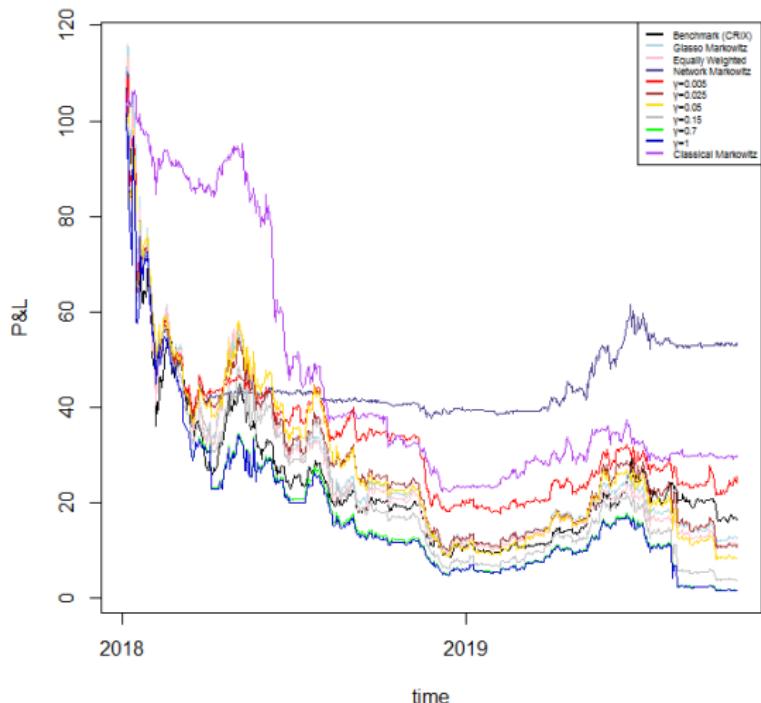
Rachev ratio

## Portfolio Results - III

Period	CRIX	EW	NW	GM	CM
Jan-2018	0.11	0.13	0.15	0.14	0.03
May-2018	0.04	0.05	0.02	0.05	0.03
Sep-2018	0.11	0.11	0.10	0.12	0.02
Jan-2019	0.07	0.10	0.05	0.07	0.01
May-2019	0.04	0.02	0.03	0.02	0.04
Sep-2019	0.05	0.05	0.02	0.05	0.01

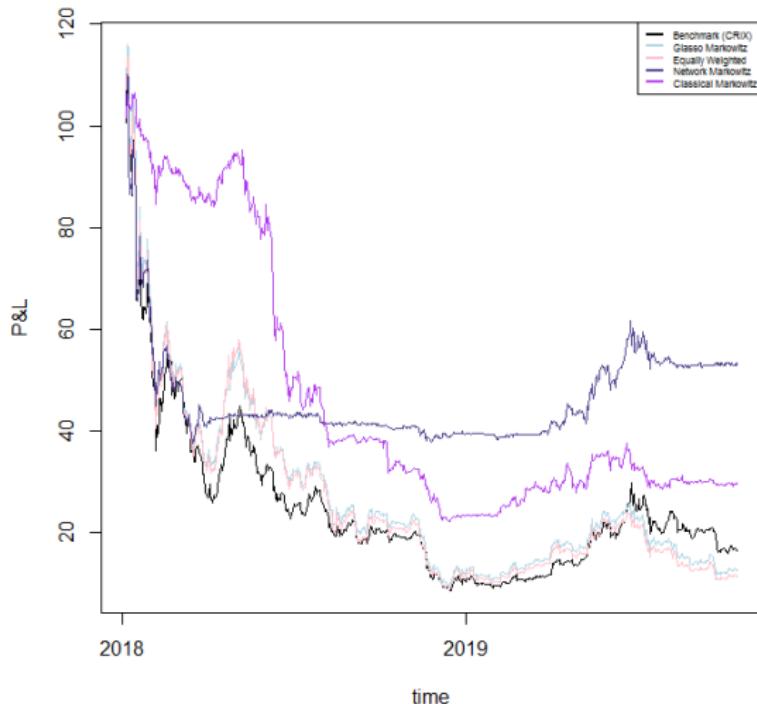
Value at Risk (VaR)

# Portfolio Results - IV



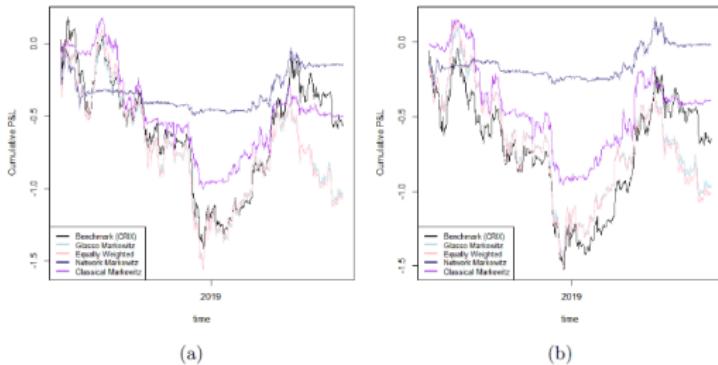
Portfolio returns

# Portfolio Results - V



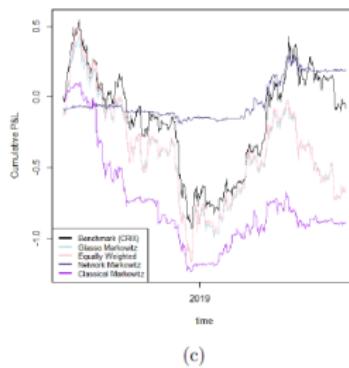
Highlight of portfolio returns

# Sensitivity



(a)

(b)



(c)

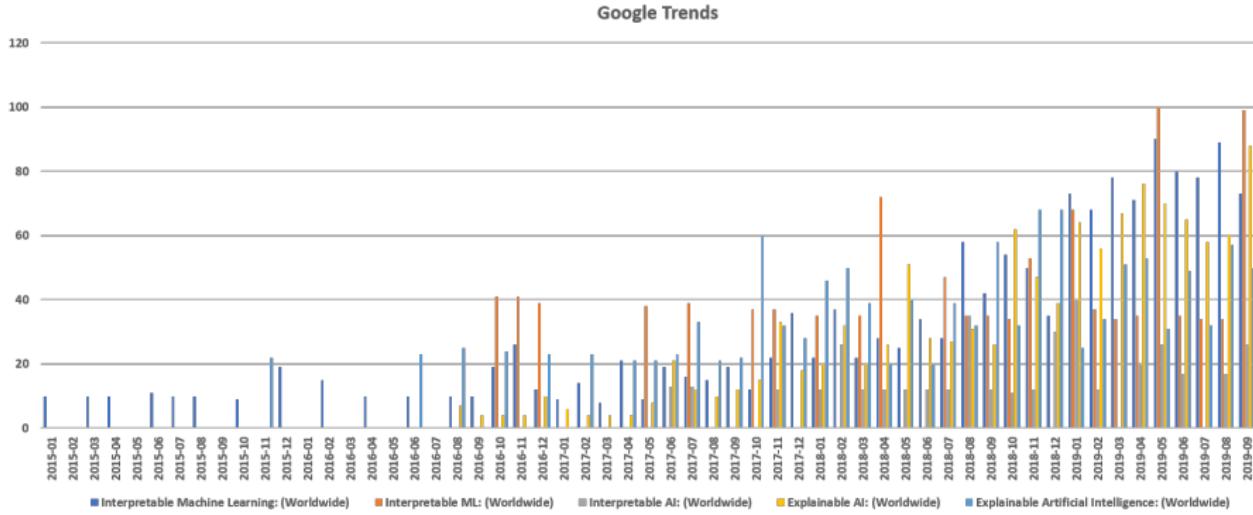
Sensitivity analysis with respect to different rolling windows

Use and distribution of this material is allowed only for purposes concerning the Fin-Tech HO2020 project

Use Case III: eXplainable AI in credit scoring and portfolio construction (Dimitri Marinelli, Jochen Papenbrock, Niklas Bussmann, Paolo Giudici; FIRAMIS and UNIPV)

## AI methods in credit scoring

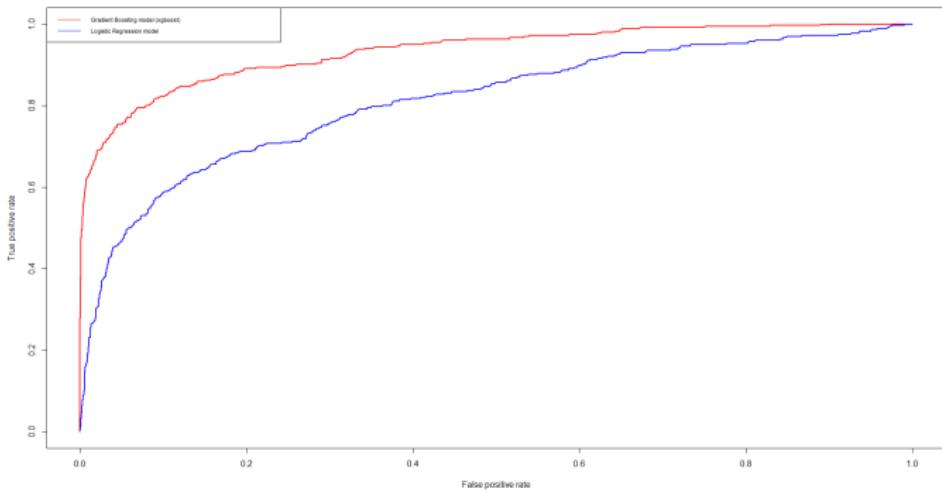
- ▶ Computationally intensive AI models can beat classic logistic regression scoring models
- ▶ However, they are not interpretable (black-boxes)
- ▶ Explainable AI models can help interpretability maintaining high predictive accuracy



Explainable AI and Interpretable AI are trending

## AI methods in credit scoring: application

- ▶ We consider about 15,000 SME companies, which have received a loan, out of which about 11% have defaulted. The data contains 20 explanatory variables
- ▶ We apply the XGboost algorithm on a training dataset (80%) and compare the predictions with the best logistic regression model on the test set (20%)

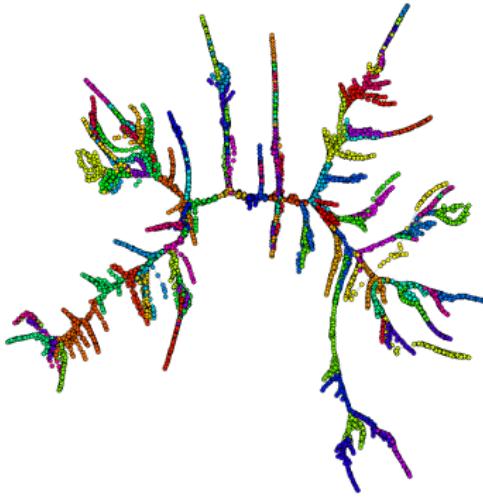


The AUROC improves from 0.81 to 0.93.

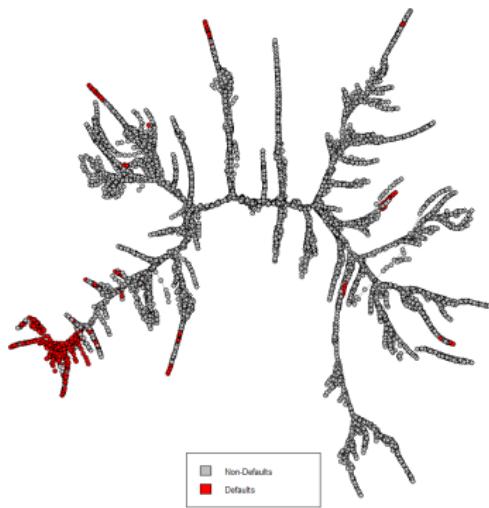
To interpret the model, we propose to apply to Shapley values:

$$\phi_i(f, x) = \sum_{z' \subseteq x'} \frac{|z'|!(M - |z'| - 1)!}{M!} [f_x(z') - f_x(z' \setminus i)] \quad (6)$$

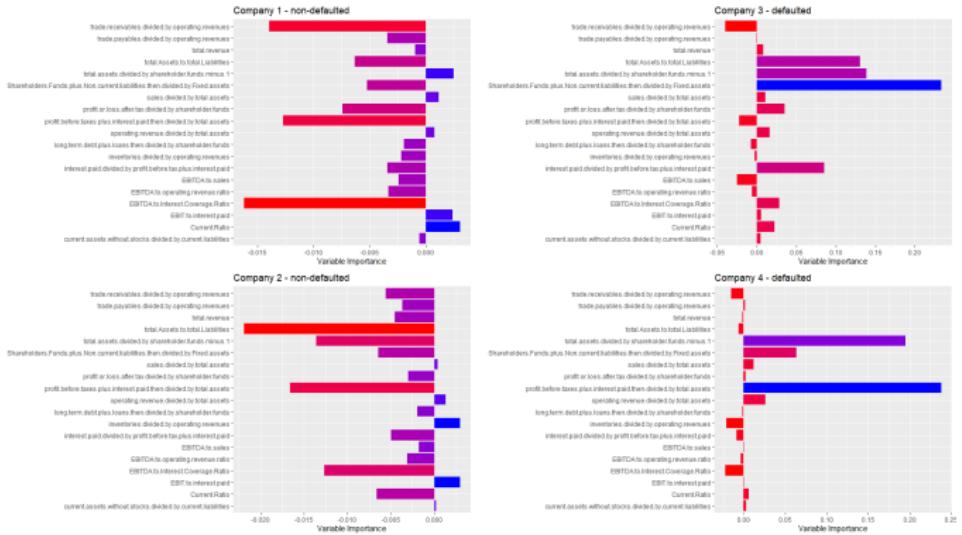
correlation network models obtained from minimum spanning tree.



**Figure 1:** Minimal Spanning Tree representation of the borrowing companies. Clustering has been performed using the standardized Euclidean distance between institutions. Companies are colored according to their cluster of belonging.



**Figure 2:** Minimal Spanning Tree representation of the borrowing companies. Clustering has been performed using the standardized Euclidean distance between institutions. Companies are colored according to their default status: red= defaulted; grey= not defaulted.



**Figure 3:** Contribution of each explanatory variable to the Shapley's decomposition of four predicted default probabilities, for two defaulted and two non defaulted companies. A red color indicates a low variable importance, and a blue color a high variable importance.



# Network Models to Enhance Automated Cryptocurrency Portfolio Management

**Paolo Giudici<sup>1\*</sup>, Paolo Pagnoncelli<sup>1\*</sup> and Gloria Polinesi<sup>2</sup>**

<sup>1</sup> Department of Economics and Management, University of Pavia, Pavia, Italy, <sup>2</sup> Department of Economics and Social Sciences, Università Politecnica delle Marche, Ancona, Italy

The usage of cryptocurrencies, together with that of financial automated consultancy, is widely spreading in the last few years. However, automated consultancy services are not yet exploiting the potentiality of this nascent market, which represents a class of innovative financial products that can be proposed by robo-advisors. For this reason, we propose a novel approach to build efficient portfolio allocation strategies involving volatile financial instruments, such as cryptocurrencies. In other words, we develop an extension of the traditional Markowitz model which combines Random Matrix Theory and network measures, in order to achieve portfolio weights enhancing portfolios' risk-return profiles. The results show that overall our model overperforms several competing alternatives, maintaining a relatively low level of risk.

## OPEN ACCESS

### Edited by:

Shatha Qamhieh Hashem,  
An-Najah National University, Palestine

### Reviewed by:

Audrius Kabasinskas,  
Kaunas University of Technology,  
Lithuania  
Jirakom Sirisrisakulchai,  
Chiang Mai University, Thailand

### \*Correspondence:

Paolo Giudici  
paolo.giudici@unipv.it  
Paolo Pagnoncelli  
paolo.pagnoncelli01@universitadipavia.it

### Specialty section:

This article was submitted to  
Artificial Intelligence in Finance,  
a section of the journal  
*Frontiers in Artificial Intelligence*

**Received:** 06 November 2019

**Accepted:** 24 March 2020

**Published:** 24 April 2020

### Citation:

Giudici P, Pagnoncelli P and Polinesi G  
(2020) Network Models to Enhance  
Automated Cryptocurrency Portfolio  
Management. *Front. Artif. Intell.* 3:22.  
doi: 10.3389/frai.2020.00022

## 1. INTRODUCTION

FinTech innovations are rapidly expanding nowadays, with applications including payments, lending, insurance and asset management, among others. Two technical reports from the Financial Stability Board (FSB) (FSB, 2017a,b)—establish several key drivers for FinTech, i.e., the shift of consumer preferences on the demand side, the change of financial regulations on the supply side and the technology evolution.

In this context, services of automated financial consulting are widely spreading and, in particular robo-advisors<sup>1</sup>. They are supposed to match the investors' risk profile with specific class of financial assets and thereby build an efficient portfolio allocation for each specific client. However, the mechanisms underlying the portfolio construction are often obscure, as well as they arguably do not properly take into account for multivariate dependencies across securities which are key to achieve diversification and, therefore, mitigate financial risk. This is particularly true when dealing with peculiarly volatile markets, such as the cryptocurrency one, which could be one of the future target market of robo-advisors, given its rapidly growing influence in the financial world.

Indeed, after its introduction by Nakamoto (2008), Bitcoin was launched online in 2009 and paved the way for many other cryptocurrencies. As a matter of fact, as of 17 October 2019, the cryptocurrency market capitalization amounts to ~220 billion USD, with a daily trading volume of roughly 52 billion USD.

Along with descriptive and qualitative studies, many researches dealt with quantitative analysis applied to the cryptocurrency market. In particular, a stream of research focuses on price discovery on Bitcoin markets, aiming to determine which are the leaders and followers of the Bitcoin

<sup>1</sup>An article published on "Statista" in 2019 states that assets under management in the robo-advisory segment amounts to roughly 981 billion USD, as well as that they are expected to grow at an annual growth rate (CAGR 2019–2023) of 27% (source: <https://www.statista.com/outlook/337/100/robo-advisors/worldwide>).

price formation process (see Brandvold et al., 2015; Pagnoncelli and Dimpfl, 2018; Giudici and Abu-Hashish, 2019). Other related researches studied the interconnectedness and spillover in the cryptocurrency market (such as Corbet et al., 2018b; Giudici and Pagnoncelli, 2019a,b). Another important area regards the study of Bitcoin derivatives—i.e., options and futures written on Bitcoin, with studies conducted by Corbet et al. (2018a), Baur and Dimpfl (2019), Giudici and Polinesi (2019), and Pagnoncelli (2019).

From a methodological viewpoint, we base our analysis on an important stream of literature, which focuses on stock and financial networks built on correlation matrices. The seminal paper by Mantegna (1999) uses correlation matrices to infer the hierarchical structure of stock markets, deriving a distance measure based on correlation matrices and building the so called Minimal Spanning Tree (MST), a graphical representation able to connect assets which are similar in terms of returns in a pairwise manner. After that, a research by Tola et al. (2008) uses the Random Matrix Theory (RMT) together with several clustering techniques and show that this significantly lowers portfolio risks. Subsequently, other papers about portfolio construction involving the network structure of financial assets followed (see Zhan et al., 2015; León et al., 2017; Raffinot, 2017; Ren et al., 2017).

To the best of our knowledge, there are no papers yet that exploit network topologies to build portfolios composed by cryptocurrencies. We fill this gap proposing a model that exploits the network structure of cryptocurrencies to provide a portfolio asset allocation that well compares with traditional ones. Following Mantegna (1999) we use Markowitz' asset allocation as a benchmark, and we check whether our proposal is able to improve on it, in terms of risk/return profile.

Indeed, the originality of the current paper is 2-fold. From a methodological point of view, we improve the traditional (Markowitz, 1952) portfolio allocation strategy by means of RMT and MST and by taking network centralities specifically into account. Moreover, throughout this technique we are able to set a parameter of systemic risk aversion that investors can tune to better match their investment strategies with their own risk profile. From an empirical viewpoint, we apply our methodology to data coming from a nascent and highly volatile market, i.e., the cryptocurrency one. This is particularly interesting, as the cryptocurrency market is rapidly expanding and its opportunities due to the high uncertainty (and volatility) around it are quite appealing, and thus a greater number of investors will likely enter it in the short run.

Our empirical findings confirm the effectiveness of our model in achieving better cumulative portfolio performances, while keeping a relatively low level of risk. In particular, we show that our proposed model which employs RMT, MST and centrality measures rapidly adapts to market conditions, and is able to yield satisfactory performances during bull market periods. During bear market periods—instead—our Network Markowitz model employing RMT and MST realizes the best performances, protecting investors from relatively high losses which are instead generated by many other asset allocation strategies tested. Furthermore, the riskiness of our strategy is

still lower than most of the competing model we analyze. These outcomes suggest that a sound combination of the proposed models should be employed in order to achieve an efficient cryptocurrency allocation strategy, which could be also used as robo-advisory toolboxes to improve automated financial consultancy.

The paper proceeds as follows. Section 2 presents our methodology and, particularly, the Random Matrix Theory, the Minimal Spanning Tree and the portfolio construction. Section 3 illustrates our empirical results. Section 4 concludes.

## 2. METHODOLOGY

### 2.1. Random Matrix Theory

Random Matrix Theory (RMT) is widely employed in several fields, such as quantum mechanics (Beenakker, 1997), condensed matter physics (Guhr et al., 1998), wireless communications (Tulino et al., 2004), as well as economics and finance (Potters et al., 2005). This technique is able to remove the noise component from the pure signal which is embedded into correlation matrices.

The algorithm tests subsequent empirical eigenvalues of the correlation matrix:  $\lambda_k < \lambda_{k+1}$ ;  $k = 1, \dots, n$ , against the null hypothesis that they are equal to the eigenvalues of a random Wishart matrix  $\mathbf{R} = \frac{1}{T}\mathbf{A}\mathbf{A}^T$  of the same size, being  $\mathbf{A}$  a  $N \times T$  matrix containing  $N$  time series of length  $T$ . The elements of  $\mathbf{A}$  are *i.i.d.* random variables, with zero mean and unit variance.

Marchenko and Pastur (1967) show that as  $N \rightarrow \infty$  and  $T \rightarrow \infty$ , and the ratio  $Q = \frac{T}{N} \geq 1$  is fixed, there is convergence of the sample eigenvalues' density to:

$$f(\lambda) = \frac{T}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}, \quad (1)$$

with  $\lambda \in (\lambda_-, \lambda_+)$ ,  $\lambda_{\pm} = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}$ .

Provided that, if  $\lambda_k > \lambda_+$  the null hypothesis is rejected from the  $k$ -th eigenvalue onwards. Hence, through a singular value decomposition the RM approach builds up a filtered correlation matrix (see Eom et al., 2009).

In our specific case, consider the continuous log return time series  $r_i$  of a generic cryptocurrency  $i$  at any time point  $t$ . i.e.,

$$r_i^t = \log P_i^t - \log P_i^{t-1}, \quad (2)$$

where  $P_i^t$  is the price of the cryptocurrency  $i$  at time  $t$ .

Considering a bunch of  $N$  cryptocurrency return time series, let  $\mathbf{C}$  be the  $N \times N$  correlation matrix of the cryptocurrency return time series. The random matrix approach filters the correlation matrix, thus obtaining a new matrix  $\mathbf{C}^*$  as:

$$\mathbf{C}^* = \mathbf{V}\Lambda\mathbf{V}^T, \quad (3)$$

with

$$\Lambda = \begin{cases} 0 & \lambda_i < \lambda_+ \\ \lambda_i & \lambda_i \geq \lambda_+ \end{cases}$$

and  $\mathbf{V}$  being the matrix of the deviating eigenvectors linked to the eigenvalues which are larger than  $\lambda_+$ .

## 2.2. The Minimal Spanning Tree

In order to simplify the relationships given by the filtered correlation matrix  $\mathbf{C}^*$  obtained from the random matrix approach, we apply the Minimal Spanning Tree representation of the cryptocurrency return time series. This is consistent with the literature on stock similarities (i.e., Mantegna and Stanley, 1999; Bonanno et al., 2003; Spelta and Araújo, 2012).

Given the filtered correlation matrix obtained in the step above, we may derive an Euclidean distance for each pairwise correlation element in the matrix, i.e.,

$$d_{ij} = \sqrt{2 - 2c_{ij}^*}, \quad (4)$$

where  $c_{ij}^*$  is a generic element  $(i, j)$  of the matrix  $\mathbf{C}^*$ , with  $i, j = 1, \dots, N$ . Each pairwise distance can be inserted in the so-called distance matrix  $\mathbf{D} = \{d_{ij}\}$ . The MST algorithm is able to reduce the number of links between the assets from  $\frac{N(N-1)}{2}$  to  $N - 1$  linking each node to its closest neighbor. In particular, we initially consider  $N$  clusters associated to the  $N$  cryptocurrencies and, at each subsequent step, we merge two generic clusters  $l_i$  and  $l_j$  if:

$$d(l_i, l_j) = \min \{d(l_i, l_j)\},$$

with the distance between clusters being defined as:

$$\hat{d}(l_i, l_j) = \min \{d_{pq}\},$$

being  $p \in l_i$  and  $q \in l_j$ . This procedure is iteratively repeated until we remain with just one cluster at hand.

Moreover, with the aim of explaining the evolution of relationships evolve over time, Spelta and Araújo (2012) proposed the so-called residuosity coefficient, which compares the relative strength of the connections above and below a threshold distance value, i.e.,

$$R = \frac{\sum_{d_{ij} > L} d_{ij}^{-1}}{\sum_{d_{ij} \leq L} d_{ij}^{-1}} \quad (5)$$

with  $L$  being the highest threshold distance value ensuring connectivity of the MST. Intuitively, the residuosity coefficient  $R$  increases when the number of links increases—meaning the network becomes more sparse, and viceversa lowers with decreasing number of links.

## 2.3. Network Centrality Measures

In this paper we employ of centrality measures in order to develop a portfolio allocation that takes into account the centrality of a node (cryptocurrency) in the system. Network theory includes several centrality measures, such as the degree centrality, counting how many neighbors a node has, as well centrality measures based on the spectral properties of graphs (see Perra and Fortunato, 2008). Among the spectral centrality measures we remark Katz's centrality (see Katz, 1953), PageRank (Brin and Page, 1998), hub and authority centralities (Kleinberg, 1999), and the eigenvector centrality (Bonacich, 2007).

In this paper we employ of the eigenvector centrality, as it measures the importance of a node in a network by assigning relative scores to all nodes in the network. Relative scores are based on the principle that being connected to few high scoring nodes contributes more to the score of the node in question than equal connections to low scoring nodes. In other words, considering a generic node  $i$ , the centrality score is proportional to the sum of the scores of all nodes which are connected to it, i.e.,

$$x_i = \frac{1}{\lambda} \sum_{j=1}^N \hat{d}_{ij} x_j \quad (6)$$

where  $x_j$  is the score of a node  $j$ ,  $\hat{d}_{ij}$  is the element  $(i, j)$  of the adjacency matrix of the network,  $\lambda$  is a constant. The equation from above can be rewritten in a compact form as:

$$\hat{\mathbf{D}}\mathbf{x} = \lambda \mathbf{x} \quad (7)$$

where  $\hat{\mathbf{D}}$  is the adjacency matrix,  $\lambda$  is the eigenvalue of the matrix  $\hat{\mathbf{D}}$ , with associated eigenvector  $\mathbf{x}$ , a vector of scores of dimension  $N$ , meaning one element for each node. Note that as our networks are based on distances between returns, the higher the centrality measure associated to a node, the more the node behaves dissimilarly with respect to the other nodes in the network.

## 2.4. Portfolio Construction

Asset correlations are key items in investment theory and risk measurement, in particular for optimization problems as in the case of the widely known portfolio theory described by Markowitz (1952). As a consequence, correlation based graphs are useful tool to build optimal investment strategies. In this subsection we show how portfolio construction can be enhanced by means of a combination of the RMT, MST, and network centrality measures described above.

Several researches have investigated the relationship between the network structure of financial assets and portfolio strategies. The study (Onnela et al., 2003) shows how a portfolio constructed via Markowitz theory is mainly composed by assets that lie in the periphery of the asset network structure, i.e., outer node assets, and not in its core. Pozzi et al. (2013) find that peripheral assets in the network yield to better performances and lower portfolio risk with respect to central ones. Peralta and Zareei (2016) show that the centrality of assets within a network are negatively related with the optimal weights obtained through the Markowitz technique. Building on that, Výrost et al. (2018) conclude that asset allocation strategies including the network structure of financial asset are able to improve a portfolio's risk-return profile.

Another stream of literature focused on proposing alternative portfolio allocation strategies based on the network structure of financial assets. To illustrate, Plerou et al. (2002) and Conlon et al. (2007) use the random matrix theory to filter the correlation matrix to be inserted in the Markowitz minimization problem, while Tola et al. (2008) add the MST obtaining improvements with respect to the raw model.

In the present context we aim to study the differences in the risk-return profiles of our strategy, which includes topological

measures in the optimization problem, with respect to the traditional Markowitz model, possibly yielding to better risk-return characteristics of the portfolios. The originality of our approach builds on the fact that we do not only use RMT and MST as alternative approaches to quantify risk diversification, but we employ an extension of the traditional Markowitz method by including these techniques in the minimization problem. Indeed, in the present case we want to solve the following problem:

$$\min_{\mathbf{w}} \mathbf{w}^T \Sigma^* \mathbf{w} + \gamma \sum_{i=1}^n x_i w_i \quad (8)$$

subject to

$$\begin{cases} \sum_{i=1}^n w_i = 1 \\ \mu_P \geq \frac{\sum_{i=1}^n \mu_i}{n} \\ w_i \geq 0 \end{cases}$$

where  $\mathbf{w}$  is the vector of portfolio weights, being  $w_i$  the weight associated to the cryptocurrency  $i$ ,  $\Sigma^*$  is the filtered variance-covariance matrix with generic element  $(i,j)$  represented by  $\sigma_i \sigma_j c_{i,j}^*$ ,  $\gamma$  is the parameter representing the risk aversion of the investor,  $x_i$  is the eigenvector centrality associated with the cryptocurrency  $i$ ,  $\mu_P$  indicates the return of the portfolio and  $\mu_i$  the return of the generic cryptocurrency  $i$ .

Generally speaking, portfolios built upon the traditional Markowitz theory are such that the risk is minimized for a given expected return, using as input the raw variance-covariance matrix of returns. In our case, the methodological improvement is 2-fold. Firstly, we modify the input variance-covariance matrix, which is filtered by both RMT and MST. Secondly, we add a component derived from the MST structure which relates to an extra risk component the investor may want to control for. Indeed, by modulating  $\gamma$  the investor can set its own level of risk aversion toward systemic risk specifically, and not just to the portfolio risk as in the Markowitz framework. As a matter of fact, being centralities inversely related with distances, a small value of  $\gamma$  yields to portfolios composed by less systemically risky cryptocurrencies, which generally lie in the peripheral part of the network. Conversely, a large value of  $\gamma$  makes the algorithm select more systemically relevant cryptocurrencies, meaning those who are in the center of the network structure. For the sake of completeness, we will test different values of the systemic risk aversion parameter in the course of the current application.

Starting from the cryptocurrency return time series, the steps of the algorithm can be summarized as follows:

1. Estimation of the filtered correlation matrix  $\mathbf{C}^*$  by RMT
2. Reduction of the number of links in the filtered correlation matrix  $\mathbf{C}^*$  by MST
3. Computation of the filtered variance-covariance matrix  $\Sigma^*$  associated to the filtered correlation matrix  $\mathbf{C}^*$  in step 2
4. Computation of the eigenvector centralities  $x_i$

5. Computation of the portfolio weights by solving the minimization problem:

$$\min_{\mathbf{w}} \mathbf{w}^T \Sigma^* \mathbf{w} + \gamma \sum_{i=1}^n x_i w_i \text{ s.t. } \begin{cases} \sum_{i=1}^n w_i = 1 \\ \mu_P \geq \frac{\sum_{i=1}^n \mu_i}{n} \\ w_i \geq 0 \end{cases}$$

The weights calculation finally yields to the portfolio returns which we use to evaluate the performance of our allocation method.

### 3. EMPIRICAL FINDINGS

#### 3.1. Data Description and Network Topology Analysis

In our empirical application we consider 10 time series of returns referred to cryptocurrencies traded over the period 14 September 2017–17 October 2019 (764 daily observations). In particular, we consider the first 10 cryptocurrencies in terms of market capitalization as of 17 October 2019<sup>2</sup>. To be precise, we analyze the return time series of the following cryptocurrencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Tether (USDT), Bitcoin Cash (BCH), Litecoin (LTC), Binance Coin (BNB), Eos (EOS), Stellar (XLM), Tron (TRX).

We provide some basic descriptive statistics of our data in **Table 1**. From **Table 1** one may notice that average daily returns are all close to zero, in line with the general economic theory regarding asset returns. However, the 10 cryptocurrencies exhibit different standard deviations, meaning that the variability in returns differs quite strongly among cryptocurrencies. To illustrate, USDT is the one showing the lowest relative variability; this is in line with the fact that this cryptocurrency is classified as stable coin, therefore its price should not deviate too much on a daily basis. On the other hand, TRX is the one showing the highest standard deviation; indeed, this particular cryptocurrency witnessed a period of high fluctuations during the considered sample period. As far as kurtosis is concerned, most of the cryptocurrencies exhibit values which reflects the non-Gaussian and heavy tailed behavior of their associated distribution. This is particularly true for XLM and XRP, whose kurtosis are relatively larger than the ones of the other time series.

To better understand the dynamics of the cryptocurrency time series, we plot the normalized price series in **Figures 1, 2**<sup>3</sup>. The two figures confirm well-known features of cryptocurrencies, such as their overall high volatility (with TRX being the most volatile), the stability of the stable coin (USDT) as well as the low liquidity that some of them exhibit (such as TRX).

In order to apply the filter through RMT, we divide the dataset into consecutive overlapping windows having a width  $T = 120$  (4 trading months). We set the window step length to 1 week (7 trading days), which makes up a total of 93 weekly 4-months windows.

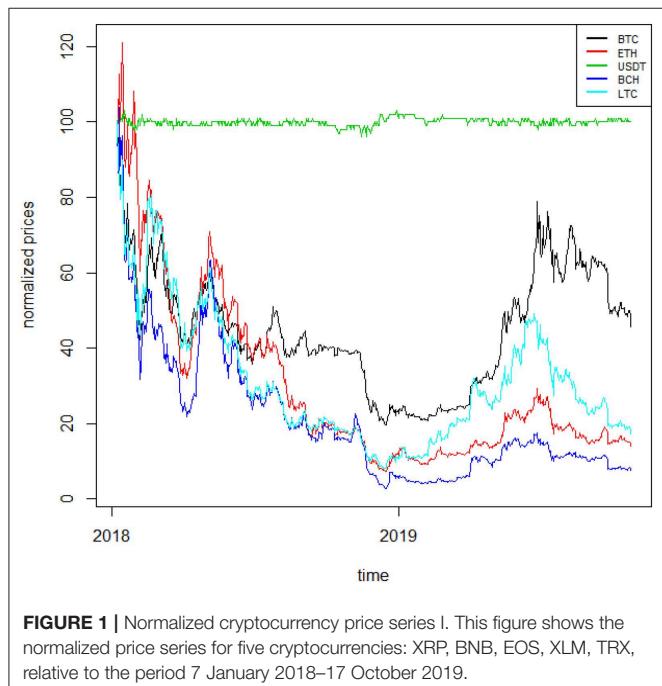
<sup>2</sup>We exclude Bitcoin SV (BSV) in order to achieve a sufficiently large timespan, meaning a more than 2-years time period.

<sup>3</sup>We split the plot in two different figures for scale reasons.

**TABLE 1 |** Summary statistics.

	<b>Mean</b>	<b>Std</b>	<b>Kurtosis</b>	<b>Skewness</b>
BTC	0.0009	0.04	3.35	-0.07
ETH	-0.0007	0.05	2.90	-0.33
XRP	0.0004	0.07	15.73	1.80
USDT	0.0000	0.01	4.28	0.22
BCH	-0.0011	0.08	6.47	0.49
LTC	-0.0003	0.06	8.02	0.66
BNB	0.0033	0.07	7.74	0.78
EOS	0.0017	0.07	3.93	0.60
XLM	0.0021	0.10	26.19	2.03
TRX	0.0021	0.15	13.15	0.66

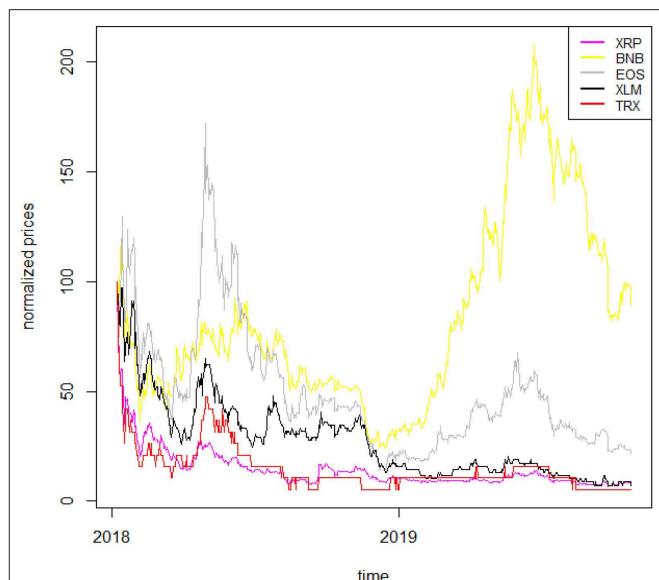
The table shows relevant summary statistics for the 10 cryptocurrencies considered related to the whole sample period, i.e., 13 September 2017–10 October 2019.



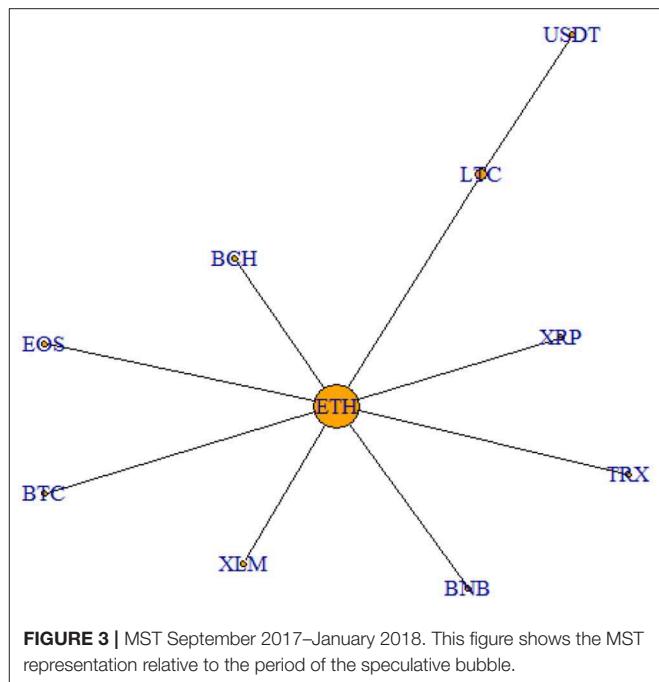
**FIGURE 1 |** Normalized cryptocurrency price series I. This figure shows the normalized price series for five cryptocurrencies: XRP, BNB, EOS, XLM, TRX, relative to the period 7 January 2018–17 October 2019.

For each time window considered, we use 15 weeks of daily observations to estimate the model, while the last week is used for validation purposes. In other words, we compute 93 correlation matrices between the 10 cryptocurrency return time series, each one based on 15 weeks of daily returns and then filter them by means of the Random Matrix approach. Applying the Random Matrix filtering, correlation matrices are rebuilt considering only the eigenvectors corresponding to the deviating eigenvalues.

In order to have a better understanding of the links existing between cryptocurrencies, the filtered correlation matrices are then used to derive the MST representation over two main periods of interest. In particular, we plot the MST structure emerging from the period of the cryptocurrency price hype (September 2017–January 2018) in **Figure 3**, while the MST



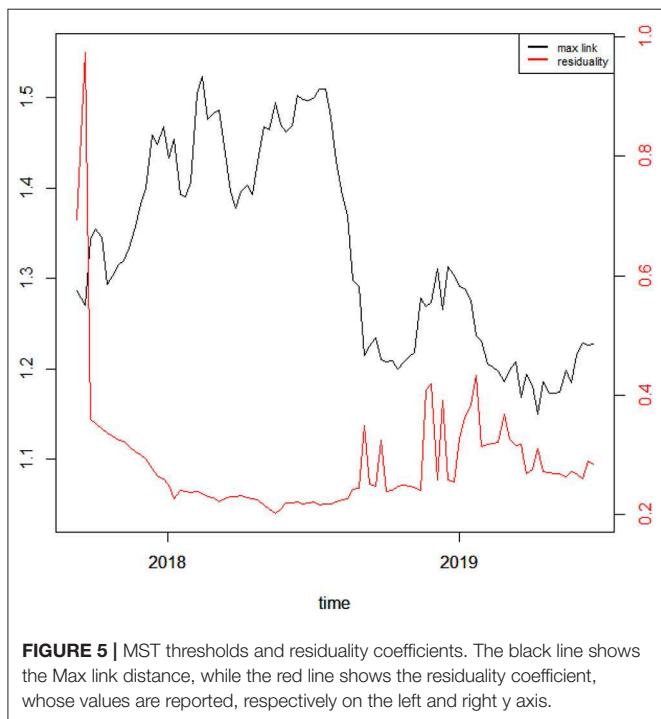
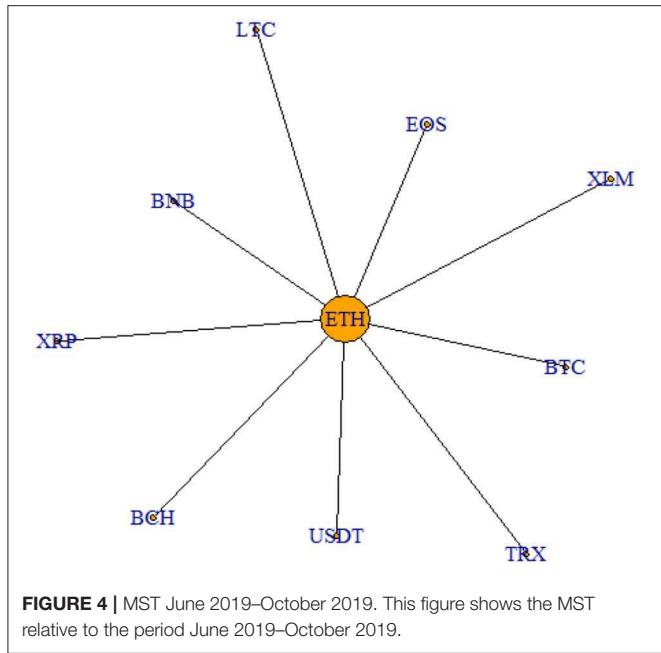
**FIGURE 2 |** Normalized cryptocurrency price series II. The figure shows the normalized price series for five cryptocurrencies: BTC, ETH, USDT, BCH, LTC, relative to the period 7 January 2018–17 October 2019.



**FIGURE 3 |** MST September 2017–January 2018. This figure shows the MST representation relative to the period of the speculative bubble.

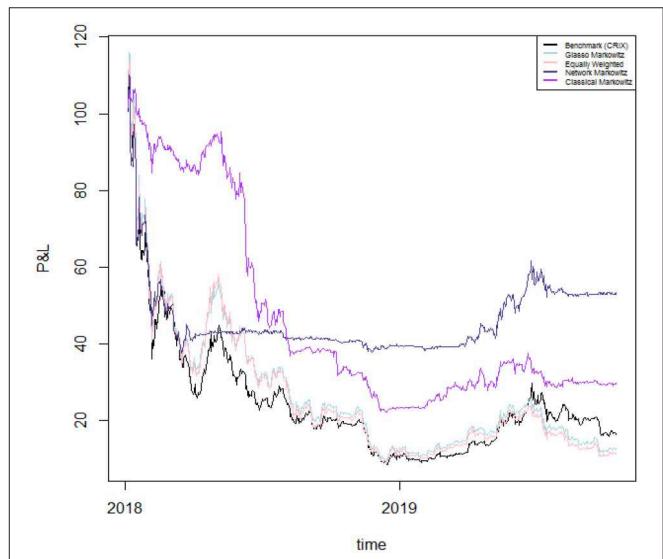
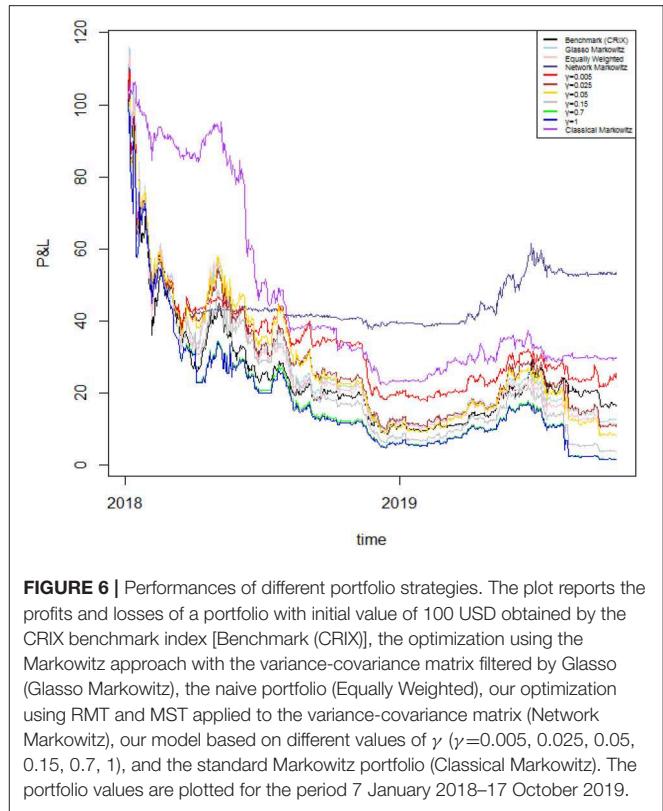
structure related to the latest trading period analyzed (June 2019–October 2019) in **Figure 4**.

As it is clear from the graph, the two networks show quite similar features. Indeed, ETH is the cryptocurrency which always lies in the center of the structure, indicating its central role in the cryptocurrency market. The only difference between the two graphical representations concerns USDT, which during the price hype is not connected directly to ETH as the other cryptocurrencies, but to LTC. This is linked to the fact that USDT



is a stable coin and, therefore, behaves dissimilarly from the other cryptocurrencies considered, being it much less volatile. However, this difference in behavior levels out during the latest period, as it emerges from **Figure 4**.

To better understand the dynamics of the MST among cryptocurrencies, we investigate the evolution of the links over time. Indeed, we compute two different measures: the Max link, i.e., the value of the maximum distance between two pairs of



**TABLE 2 |** Cumulative Profits and Losses.

Period	CRIX	GM	EW	CM	NW	$\gamma = 0.005$	$\gamma = 0.025$	$\gamma = 0.05$	$\gamma = 0.15$	$\gamma = 0.7$	$\gamma = 1$
Jan-2018	-0.14	-0.13	-0.16	0.04	-0.22	-0.21	-0.26	-0.27	-0.36	-0.43	-0.43
May-2018	-0.67	-0.62	-0.60	-0.12	-0.79	-0.78	-0.73	-0.66	-0.83	-1.08	-1.10
Sep-2018	-1.37	-1.37	-1.43	-0.88	-0.83	-1.02	-1.24	-1.23	-1.40	-1.60	-1.64
Jan-2019	-1.85	-1.78	-1.78	-1.32	-0.87	-1.50	-1.86	-1.98	-2.19	-2.29	-2.31
May-2019	-1.35	-1.25	-1.27	-1.01	-0.74	-1.22	-1.33	-1.29	-1.44	-1.55	-1.57
Sep-2019	-0.99	-1.45	-1.49	-1.02	-0.54	-1.19	-1.34	-1.44	-1.86	-2.13	-2.15

This table shows the cumulative 4-months Profits and Losses of portfolios under different strategies. Particularly, Profits & Losses are computed for the CRIX benchmark index (CRIX), the Glasso Markowitz (GM), the naive portfolio (EW), the Network Markowitz (NW), the classical Markowitz (CM), and the proposed models with different values of  $\gamma$  ( $\gamma = 0.005, 0.025, 0.05, 0.15, 0.7, 1$ ). All values are expressed in percentage terms.

nodes in the tree, and the residuosity coefficient, meaning the ratio between the number of links which are dropped and the number of those who are kept by the MST algorithm. The two metrics, computed over the whole sample period, are illustrated in **Figure 5**.

From **Figure 5** one may notice that the Max link increases during the Bitcoin price hype and fluctuates around relatively large values until roughly mid 2018, meaning that during this period correlations between cryptocurrency returns are strongly misaligned. After that, the index bounces back toward its previous values and even below, suggesting that cryptocurrency returns start to behave more similarly during the latest period. Furthermore, the residuosity coefficient increases during the very beginning of the sample period, while it sharply declines during the price hype phase. After the decrease, the coefficient stays quite stable and then gently increases not without fluctuations from mid 2018 to the end of the sample period. This suggests that the number of links until mid 2018 was quite limited, and therefore, returns misaligned, whereas the same number started to increase after that phase, meaning there were more connections and thus more synchronicity across cryptocurrency returns.

### 3.2. Portfolio Construction

In this subsection we illustrate the results related to the proposed portfolio strategies. The optimal portfolio weights are obtained through the constrained minimization of the objective function in Equation 8. For the sake of completeness, we use different values of the systemic risk aversion parameter  $\gamma$ , meaning  $\gamma = 0.005, 0.025, 0.05, 0.15, 0.7, 1$ . These values have been chosen, without loss of generality, to be representative of different aversion profiles. While  $\gamma = 0$  indicates no aversion,  $\gamma = 1$  indicates a high aversion, with systemic risk being given the same importance as non-systemic one.

We use fifteen weeks, i.e., to compute the optimal portfolio weights as described in section 2. We then use the last week associated to each window to evaluate the out-of-sample performance of our technique, meaning to compute the portfolio returns and, therefore, the resulting Profit & Losses. We then compute portfolio returns for the period 7 January 2018–17 October 2019, accounting for rebalancing costs, which are supposed to amount to 10 basis points.

In **Figure 6** we plot the returns of our investment strategies for the different values of  $\gamma$  mentioned above as well as for  $\gamma =$

**TABLE 3 |** VaR.

Period	CRIX	EW	NW	GM	CM
Jan-2018	0.11	0.13	0.15	0.14	0.03
May-2018	0.04	0.05	0.02	0.05	0.03
Sep-2018	0.11	0.11	0.10	0.12	0.02
Jan-2019	0.07	0.10	0.05	0.07	0.01
May-2019	0.04	0.02	0.03	0.02	0.04
Sep-2019	0.05	0.05	0.02	0.05	0.01

This table shows the 4-months Value at Risk of portfolios under different strategies for a confidence interval of 95%. In particular, the VaR is computed for the CRIX benchmark index (CRIX), the naive portfolio (EW), the Network Markowitz (NW), the Glasso Markowitz (GM), and the classical Markowitz (CM). All values are expressed in absolute terms multiplied by a scale factor of 100.

0 (Network Markowitz), meaning the results of the Markowitz portfolio strategy using the variance-covariance matrix filtered by RMT and MST. In doing so, we plot portfolio performances under the hypothesis of investing 100 USD at the beginning of the period, and examining how much is lost along time. The results of our strategies are compared with the performance of several strategies and indicators: the benchmark portfolio (CRIX<sup>4</sup>), the Markowitz portfolio with variance-covariance matrix filtered by the Glasso<sup>5</sup> technique (Glasso Markowitz), the naive portfolio (Equally Weighted) and the traditional Markowitz portfolio (Classical Markowitz). To better highlight the results of our best proposed model, we plot the results only for a selection of portfolio strategies in **Figure 7**. To complement this information, we report the 4-months cumulative Profits and Losses of each of the considered strategy in **Table 2**.

Overall, we are considering a period in which the cryptocurrency market witnesses a down period—except for the first part of our analyzed timespan and several short periods consequently occurring. Therefore, as the market is not profitable during the studied period, we aim to achieve through

<sup>4</sup>The CRIX is a cryptocurrency market index following the Laspeyres methodology for the construction of indexes. More information about CRIX can be found at <https://thecrix.de/>

<sup>5</sup>The sparsity parameter  $\rho$  has been set to 0.01, as in the reference paper by Friedman et al. (2008).

**TABLE 4 | Sharpe ratio.**

Period	GM	EW	CM	NW	$\gamma = 0.005$	$\gamma = 0.025$	$\gamma = 0.05$	$\gamma = 0.15$	$\gamma = 0.7$	$\gamma = 1$
Jan-2018	-0.05	-0.05	-0.03	-0.13	-0.12	-0.08	-0.06	-0.08	-0.09	-0.10
May-2018	-0.14	-0.14	-0.19	-0.03	-0.04	-0.08	-0.09	-0.08	-0.07	-0.07
Sep-2018	-0.10	-0.09	-0.17	-0.04	-0.17	-0.17	-0.20	-0.20	-0.18	-0.17
Jan-2019	0.10	0.09	0.11	0.09	0.06	0.08	0.11	0.12	0.12	0.12
May-2019	-0.02	-0.02	0.01	0.08	0.02	0.02	-0.00	-0.03	-0.04	-0.04
Sep-2019	-0.06	-0.06	-0.03	0.03	0.07	-0.11	-0.14	-0.14	-0.14	-0.14

This table shows the 4-months values of Sharpe ratio of portfolios under different strategies. In particular, the SR is computed for the Glasso Markowitz (GM), the naive portfolio (EW), the classical Markowitz (CM), the Network Markowitz (NW), and for all the value of  $\gamma$ .

**TABLE 5 | Rachev ratio.**

Period	GM	EW	CM	NW	$\gamma = 0.005$	$\gamma = 0.025$	$\gamma = 0.05$	$\gamma = 0.15$	$\gamma = 0.7$	$\gamma = 1$
Jan-2018	0.74	0.75	0.63	0.64	0.69	0.77	0.79	0.78	0.77	0.99
May-2018	0.73	0.75	0.95	0.83	0.74	0.77	0.83	0.87	0.87	0.55
Sep-2018	0.81	0.84	0.87	0.61	0.80	0.75	0.76	0.80	0.80	0.48
Jan-2019	1.16	1.11	1.47	1.24	1.34	1.36	1.39	1.40	1.40	1.26
May-2019	0.80	0.80	1.05	0.97	0.93	0.84	0.75	0.72	0.72	0.98
Sep-2019	0.75	0.78	1	1.14	0.43	0.38	0.38	0.38	0.37	0.78

This table shows the 4-months values of Rachev Ratio (RR) of portfolios under different strategies. In particular, the RR is computed for the Glasso Markowitz (GM), the naive portfolio (EW), the classical Markowitz (CM), the Network Markowitz (NW), and for all the value of  $\gamma$ .

our allocation strategies losses which are lower than those yielded by other competing methodologies.

On the one hand, during a first phase which lasts roughly until mid 2018, the traditional Markowitz portfolio seems to overperform the other portfolio allocation strategies. Indeed, the allocation by Markowitz' technique yields to positive (cumulative) returns until January 2018 and just slightly negative ones until May 2018, however still lower than the losses provided by the other strategies in absolute terms.

On the other hand, from September 2018 onwards all portfolios start providing strong negative returns. Indeed, the returns yielded by the portfolio constructed via Markowitz start to decline dramatically, together with those of the model including the systemic risk aversion parameter. This is because the latter model takes into account the centrality of the cryptocurrencies in the network and is therefore more adaptive to market conditions, regardless of whether they are favorable or not. Indeed it can be noticed that—overall—during bull market periods our model taking into account for risk aversion reacts very fast to upward movements and yields to good cumulative performances; conversely, during down market periods, the same model yields to worse relative performances due to declining market conditions.

However, during the second half of our sample period our proposed model with the systemic risk aversion parameter  $\gamma$  set to 0 (Network Markowitz) clearly overwhelms the other portfolio allocation strategies. To illustrate, if we look at the cumulative performance of the above mentioned method, we can see that it more than halves losses with respect to the equally weighted portfolio, to the Glasso Markowitz portfolio and to all portfolios including a risk aversion parameter  $\gamma > 0$ . Moreover, it almost halves the losses with respect to the benchmark index (CRIX) and

to the traditional Markowitz methodology. This suggests that this model is capable to provide a stronger coverage for losses in case of down market periods with respect to all other considered asset allocation strategies<sup>6</sup>.

In Table 3 we compute the 4-months Value at Risk (VaR) with a confidence level of 0.05% for the benchmark index (CRIX), the equally weighted portfolio, our Network Markowitz portfolio, the Glasso Markowitz and the traditional Markowitz portfolios. This is done in order to compare, together with cumulative returns, the potential riskiness of our strategy with respect to the alternative portfolio allocation methods considered.

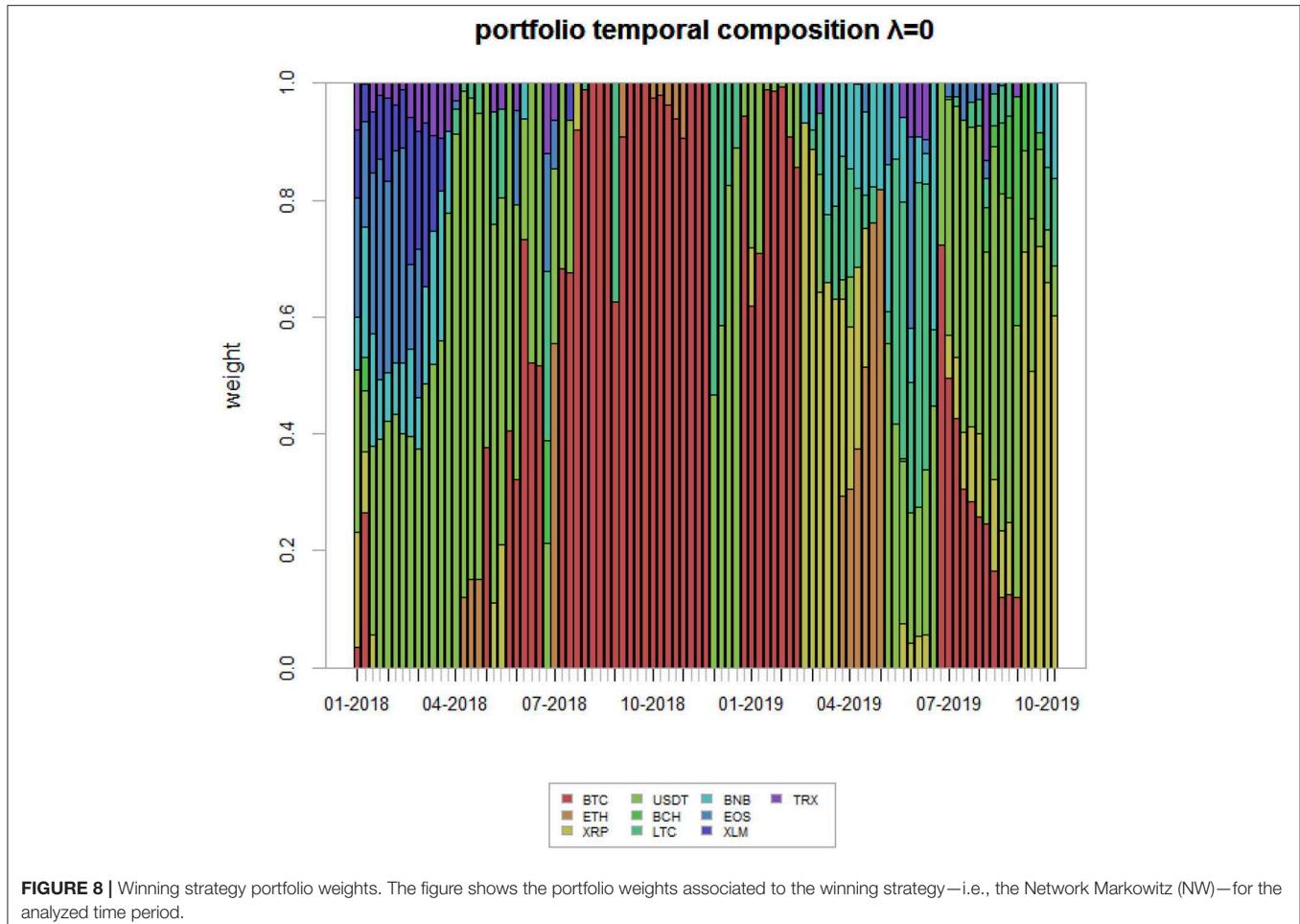
Table 3 shows that, except for the price hype period, our proposed Network Markowitz approach generally yields to lower values at risk with respect to the benchmark index (CRIX), the naive portfolio and the Glasso Markowitz. The aforementioned model is instead more risky than the traditional Markowitz model, although the latter, overall, yields too far way larger negative returns. In general, the riskiness of our strategy seems to be quite satisfactory with respect to the alternative allocation strategies analyzed.

To further support our conclusions, Table 4 presents the Sharpe ratio under the different strategies.

Table 4 gives further evidence to support our conclusions: the proposed Network Markowitz approach yields better Sharpe Ratios.

To strengthen the robustness of our conclusions, Table 5 presents the Rachev ratio, with a confidence level of 10%, under the different strategies. The Rachev ratio is a useful supplement of

<sup>6</sup>A sensitivity analysis reported in the Appendix confirms that results are robust with respect to different choices of the starting points and rolling estimation windows.



the Sharpe ratio, when data is non-symmetric, as in our context. It is calculated as the ratio between an extreme gain and an extreme loss.

**Table 5** shows that the Network Markowitz approach yields the best performances in the initial and final periods, and the Classic Markowitz in all other periods. The other strategies generally show worse performances. This is consistent with our previous findings, and with the fact that the Rachev ratio takes higher values during periods characterized by decreasing returns, such as the quarter preceding January 2019.

Overall, we cannot say that the proposed model overperforms traditional approach (such as Glasso Markowitz and Classical Markowitz). It does so in certain periods and according to certain risk aversion parameterizations.

For the sake of completeness, we plot the portfolio weights of the winning strategy over the evaluation time horizon in **Figure 8**. As one can clearly see, the composition of the portfolio varies quite much over time. Indeed, during the first period of the sample, approximately until February 2018, the portfolio is composed by various assets, with USDT gaining a high share over time, being it the most stable across all. After that, BTC is the cryptocurrency which is mostly selected by our algorithm, roughly until October 2018 (with some exceptions), as it is considered a proxy of the whole market. Finally, the algorithm

selects different cryptocurrency compositions until the end of the sample, being the latter a highly uncertain period for the market.

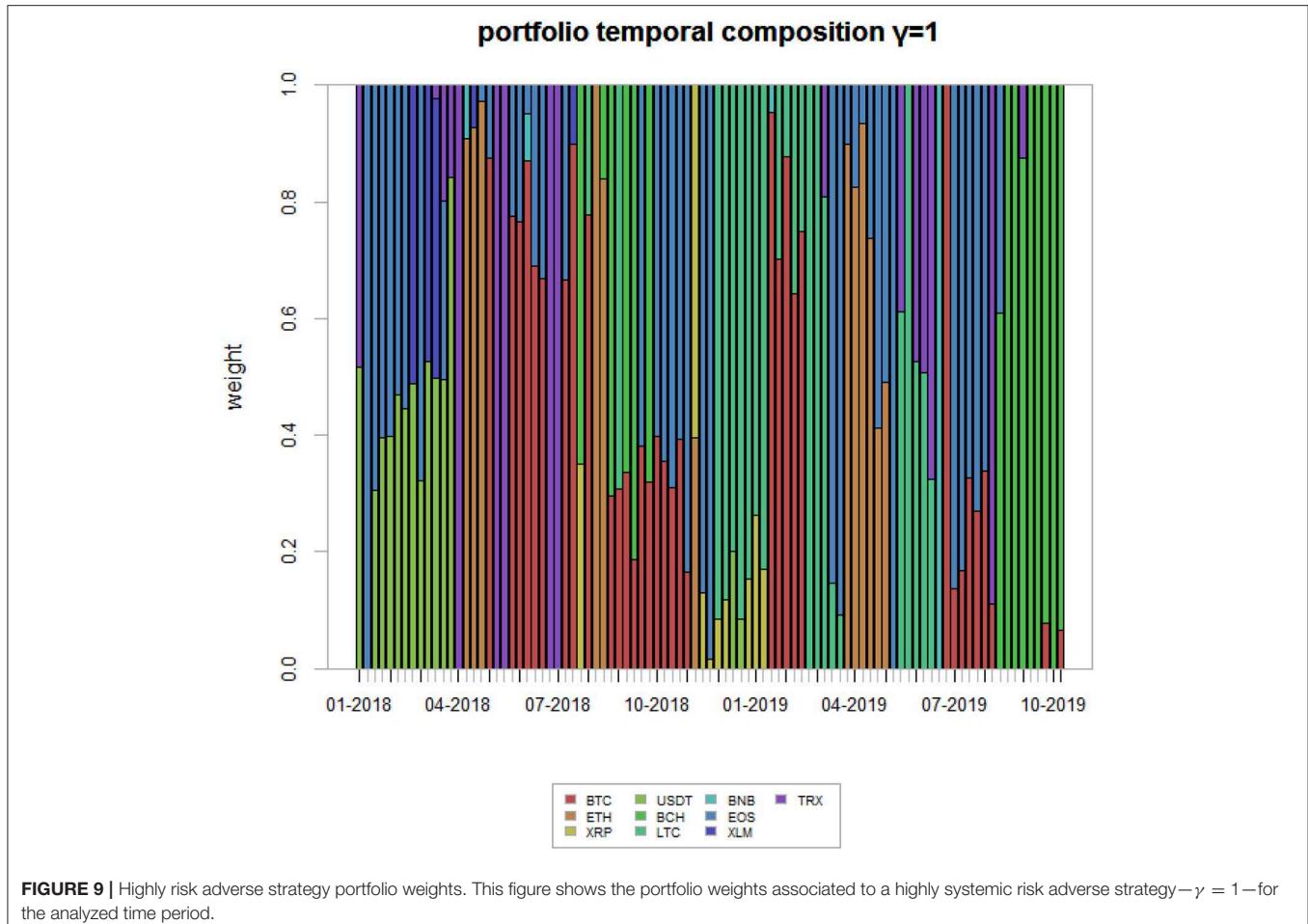
Last, we present, for comparison purposes, the portfolio weights associated with  $\gamma = 1$ .

While **Figure 8** gives the weights relative to the situation of no systemic risk aversion, **Figure 9** gives the weight corresponding to a very high systemic risk aversion, in which it has the same importance as non-systemic risk.

## 4. CONCLUSIONS

In this paper we have proposed a methodology that aims to build an allocation strategy which is suitable for highly volatile markets, such as cryptocurrency ones. In particular, we have applied our models to a set of 10 cryptocurrency return time series, selected in terms of market capitalization. We have shown that the use of network models can enhance portfolios' risk-return profiles and mitigate losses during down market periods.

We have demonstrated how the use of centrality measures, together with tuning an investor's systemic risk aversion, is a suitable methodology to make profits during bull market periods, as this method is rapidly adaptive to market conditions. We have also shown that, to protect investors from losses during bear market periods, the combination of Random Matrix Theory



and Minimal spanning trees can yield to acceptable risk-return profiles and/or mitigate losses.

Our empirical findings show that, overall, the proposed method is acceptable, even during downturn periods. However, we cannot claim that this proposed model should always be used in automated consultancy. It should always be compared with competing alternatives, according to different market conditions and different risk aversions.

We strongly believe that the proposed model should be further tested in different contexts. For this purpose, we provide at <https://www.fintech-ho2020.eu> a link to the used data and software, so the proposed methods can be fully reproduced. The software is written in the R language, and allows the methods to be extended to other data and contexts.

Further research should involve, besides the application to other contexts, the consideration of different base portfolio allocation models. We have used Markowitz' as is the most employed by robot advisory platforms.

## DATA AVAILABILITY STATEMENT

The datasets analyzed for this study can be found in Coinmarketcap (<https://coinmarketcap.com/>). The software used

in this article is not publicly available because based on a proprietary source. Requests to access the datasets should be directed to Gloria Polinesi ([gropol@hotmail.it](mailto:gropol@hotmail.it)).

## AUTHOR CONTRIBUTIONS

The paper was written in close collaboration between the authors. However, sections 2 and 4 have been written by GP, section 1 and 3 by PP, while PG has supervised and coordinated the work.

## FUNDING

This research has received funding from the European Union's Horizon 2020 research and innovation program FIN-TECH: A Financial supervision and Technology compliance training programme under the grant agreement No 825215 (Topic: ICT-35-2018, Type of action: CSA).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frai.2020.00022/full#supplementary-material>

## REFERENCES

- Baur, D. G. and Dimpfl, T. (2019). Price discovery in bitcoin spot or futures?. *J. Fut. Markets.* 39.7, 803–817. doi: 10.2139/ssrn.3171464
- Beenakker, C. W. (1997). Random-matrix theory of quantum transport. *Rev. Mod. Phys.* 69:731. doi: 10.1103/RevModPhys.69.731
- Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Soc. Netw.* 29, 555–564. doi: 10.1016/j.socnet.2007.04.002
- Bonanno, G., Caldarelli, G., Lillo, F., and Mantegna, R. N. (2003). Topology of correlation-based minimal spanning trees in real and model markets. *Phys. Rev. E* 68:046130. doi: 10.1103/PhysRevE.68.046130
- Brandvold, M., Molnár, P., Vagstad, K., and Valstad, O. C. A. (2015). Price discovery on bitcoin exchanges. *J. Int. Financ. Markets Inst. Money* 36, 18–35. doi: 10.1016/j.intfin.2015.02.010
- Brin, S., and Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Comput. Netw. ISDN Syst.* 30, 107–117. doi: 10.1016/S0169-7552(98)00110-X
- Conlon, T., Ruskin, H. J., and Crane, M. (2007). Random matrix theory and fund of funds portfolio optimisation. *Phys. A Stat. Mech. Appl.* 382, 565–576. doi: 10.1016/j.physa.2007.04.039
- Corbet, S., Lucey, B., Peat, M., and Vigne, S. (2018a). Bitcoin futures—what use are they? *Econ. Lett.* 172, 23–27. doi: 10.1016/j.econlet.2018.07.031
- Corbet, S., Meegan, A., Larkin, C., Lucey, B., and Yarovaya, L. (2018b). Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Econ. Lett.* 165, 28–34. doi: 10.1016/j.econlet.2018.01.004
- Eom, C., Oh, G., Jung, W.-S., Jeong, H., and Kim, S. (2009). Topological properties of stock networks based on minimal spanning tree and random matrix theory in financial time series. *Phys. A Stat. Mech. Appl.* 388, 900–906. doi: 10.1016/j.physa.2008.12.006
- Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9, 432–441. doi: 10.1093/biostatistics/kxm045
- FSB (2017a). *Financial Stability Implications From Fintech: Supervisory and Regulatory Issues That Merit Authorities' Attention*. Basel: FSB.
- FSB (2017b). *Fintech Credit*. Financial Stability Board Report (27 June, 2017).
- Giudici, P., and Abu-Hashish, I. (2019). What determines bitcoin exchange prices? A network var approach. *Finance Res. Lett.* 28, 309–318. doi: 10.1016/j.frl.2018.05.013
- Giudici, P., and Pagnoncelli, P. (2019a). High frequency price change spillovers in bitcoin markets. *Risks* 7:111. doi: 10.3390/risks7040111
- Giudici, P., and Pagnoncelli, P. (2019b). Vector error correction models to measure connectedness of bitcoin exchange markets. *Appl. Stochast. Models Bus. Ind.* 36, 95–109. doi: 10.1002/asmb.2478
- Giudici, P., and Polinesi, G. (2019). Crypto price discovery through correlation networks. *Ann. Oper. Res.* 2. doi: 10.1007/s10479-019-03282-3
- Guhr, T., Müller-Groeling, A., and Weidenmüller, H. A. (1998). Random-matrix theories in quantum physics: common concepts. *Phys. Rep.* 299, 189–425. doi: 10.1016/S0370-1573(97)00088-4
- Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika* 18, 39–43. doi: 10.1007/BF02289026
- Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *J. ACM* 46, 604–632. doi: 10.1145/324133.324140
- León, D., Aragón, A., Sandoval, J., Hernández, G., Arévalo, A., and Nino, J. (2017). Clustering algorithms for risk-adjusted portfolio construction. *Proc. Comput. Sci.* 108, 1334–1343. doi: 10.1016/j.procs.2017.05.185
- Mantegna, R. N. (1999). Hierarchical structure in financial markets. *Eur. Phys. J. B Condens. Matter Complex Syst.* 11, 193–197. doi: 10.1007/s100510050929
- Mantegna, R. N., and Stanley, H. E. (1999). *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge: Cambridge University Press.
- Marchenko, V. A., and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Matematich. Sbornik* 114, 507–536. doi: 10.2307/2975974
- Markowitz, H. (1952). Portfolio selection. *J. Finance* 7, 77–91. doi: 10.1111/j.1540-6261.1952.tb01525.x
- Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Available online at: www.bitcoin.org
- Onnela, J.-P., Chakraborti, A., Kaski, K., Kertesz, J., and Kanto, A. (2003). Dynamics of market correlations: taxonomy and portfolio analysis. *Phys. Rev. E* 68:056110. doi: 10.1103/PhysRevE.68.056110
- Pagnoncelli, P. (2019). Neural network models for bitcoin option pricing. *Front. Artif. Intell.* 2:5. doi: 10.3389/frai.2019.00005
- Pagnoncelli, P., and Dimpfl, T. (2018). Price discovery on bitcoin markets. *Digital Finance* 1, 1–23. doi: 10.1007/s42521-019-00006-x
- Peralta, G., and Zareei, A. (2016). A network approach to portfolio selection. *J. Empir. Finance* 38, 157–180. doi: 10.1016/j.jempfin.2016.06.003
- Perra, N., and Fortunato, S. (2008). Spectral centrality measures in complex networks. *Phys. Rev. E* 78:036107. doi: 10.1103/PhysRevE.78.036107
- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T., and Stanley, H. E. (2002). Random matrix approach to cross correlations in financial data. *Phys. Rev. E* 65:066126. doi: 10.1103/PhysRevE.65.066126
- Potters, M., Bouchaud, J.-P., and Laloux, L. (2005). Financial applications of random matrix theory: old lace and new pieces. *arXiv* 0507111.
- Pozzi, F., Di Matteo, T., and Aste, T. (2013). Spread of risk across financial markets: better to invest in the peripheries. *Sci. Rep.* 3:1665. doi: 10.1038/srep01665
- Raffinot, T. (2017). Hierarchical clustering-based asset allocation. *J. Portfolio Manag.* 44, 89–99. doi: 10.3905/jpm.2018.44.2.089
- Ren, F., Lu, Y.-N., Li, S.-P., Jiang, X.-F., Zhong, L.-X., and Qiu, T. (2017). Dynamic portfolio strategy using clustering approach. *PLoS ONE* 12:e0169299. doi: 10.1371/journal.pone.0169299
- Spelta, A., and Araújo, T. (2012). The topology of cross-border exposures: beyond the minimal spanning tree approach. *Phys. A Stat. Mech. Appl.* 391, 5572–5583. doi: 10.1016/j.physa.2012.05.071
- Tola, V., Lillo, F., Gallegati, M., and Mantegna, R. N. (2008). Cluster analysis for portfolio optimization. *J. Econ. Dyn. Control* 32, 235–258. doi: 10.1016/j.jedc.2007.01.034
- Tulino, A. M., and Verdú, S. (2004). Random matrix theory and wireless communications. *Found. Trends Commun. Inform. Theory* 1, 1–182. doi: 10.1561/0100000001
- Výrost, T., Lyócsa, Š., and Baumöhl, E. (2018). Network-based asset allocation strategies. *N. Am. J. Econ. Finance* 47, 516–536. doi: 10.1016/j.najef.2018.06.008
- Zhan, H. C. J., Rea, W., and Rea, A. (2015). An application of correlation clustering to portfolio diversification. *arXiv* 1511.07945.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Giudici, Pagnoncelli and Polinesi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Sentiment Analysis of European Bonds 2016–2018

Peter Schwendner<sup>1\*</sup>, Martin Schüle<sup>2</sup> and Martin Hillebrand<sup>3</sup>

<sup>1</sup> School of Management and Law, Center for Asset Management, Zurich University of Applied Sciences, Winterthur, Switzerland, <sup>2</sup> School of Life Sciences and Facility Management, Institute for Applied Simulation, Zurich University of Applied Sciences, Wädenswil, Switzerland, <sup>3</sup> European Stability Mechanism, Luxembourg, Luxembourg

## OPEN ACCESS

### Edited by:

Dror Y. Kenett,  
Johns Hopkins University,  
United States

### Reviewed by:

Aparna Gupta,

Rensselaer Polytechnic Institute,  
United States

German Gonzalo Creamer,  
Stevens Institute of Technology,  
United States

### \*Correspondence:

Peter Schwendner  
scwp@zhaw.ch

### Specialty section:

This article was submitted to  
Artificial Intelligence in Finance,  
a section of the journal  
*Frontiers in Artificial Intelligence*

Received: 01 November 2018

Accepted: 20 September 2019

Published: 15 October 2019

### Citation:

Schwendner P, Schüle M and  
Hillebrand M (2019) Sentiment  
Analysis of European Bonds  
2016–2018. *Front. Artif. Intell.* 2:20.  
doi: 10.3389/frai.2019.00020

We revisit the discussion of market sentiment in European sovereign bonds using a correlation analysis toolkit based on influence networks and hierarchical clustering. We focus on three case studies of political interest. In the case of the 2016 Brexit referendum, the market showed negative correlations between core and periphery only in the week before the referendum. Before the French presidential elections in 2017, the French bond spread widened together with the estimated Le Pen election probability, but the position of French bonds in the correlation blocks did not weaken. In summer 2018, during the budget negotiations within the new Italian coalition, the Italian bonds reacted very sensitively to changing political messages but did not show contagion risk to Spain or Portugal for several months. The situation changed during the week from October 22 to 26, as a spillover pattern of negative sentiment also to the other peripheral countries emerged.

**Keywords:** sovereign bonds, contagion, sentiment, European sovereign bond crisis, correlation, correlation influence, networks

## INTRODUCTION

In this empirical study, we discuss the short-term impact of three specific political situations relevant to the European Union on the return correlations between its sovereign bond markets in 2016, 2017, and 2018. We focus on effects happening at the same time in these markets and interpret the correlation patterns on an hourly timescale in non-overlapping weekly time windows as an expression of the sentiment of market makers regarding a potential risk spillover. Forbes and Rigobon (2002) and Rigobón (2019) present a precise differentiation of “spillover,” “contagion,” and “interdependence” phenomena.

To illustrate our interpretation of “sentiment,” we point out that positioning decisions of large investors happen at a slower pace than quote changes generated by quote machines of bond market makers. Quote machines need to make sure that market makers cannot get “arbitraged” by external traders who have access to all public market information. Therefore, market maker quotes need to include current market information, even information inferred from other markets. These “cross-sectional” quotation models can enable correlation patterns in the quoted time series. For example, negative news concerning a specific country may trigger a spread widening of bonds of this country and also of bonds of other, similar countries even before many actual trades happen. The changes in observed quotes then may have an impact on the trading decisions of speculative traders who might follow a momentum trading rule. On a longer time scale, these quote changes can also have an impact on the positioning of long-term investors who might be forced to cut their positions as the need to comply with a stop-loss or value-at-risk rule.

The Euro-denominated sovereign bond markets within the European Union are a very specific universe, as the yield levels across countries significantly converged already before the introduction

of the Euro in 1999 and diverged during the European sovereign debt crisis from 2010 to 2014, accompanied by a pronounced block structure in the correlation matrix reflecting the “core-periphery” dichotomy. At the peak of the crisis between 2010 and 2012, the correlations between core European and periphery bonds have even been negative as only the core bonds acted as “safe havens,” but not the periphery bonds, inducing capital flows from the weaker to the stronger bond markets.

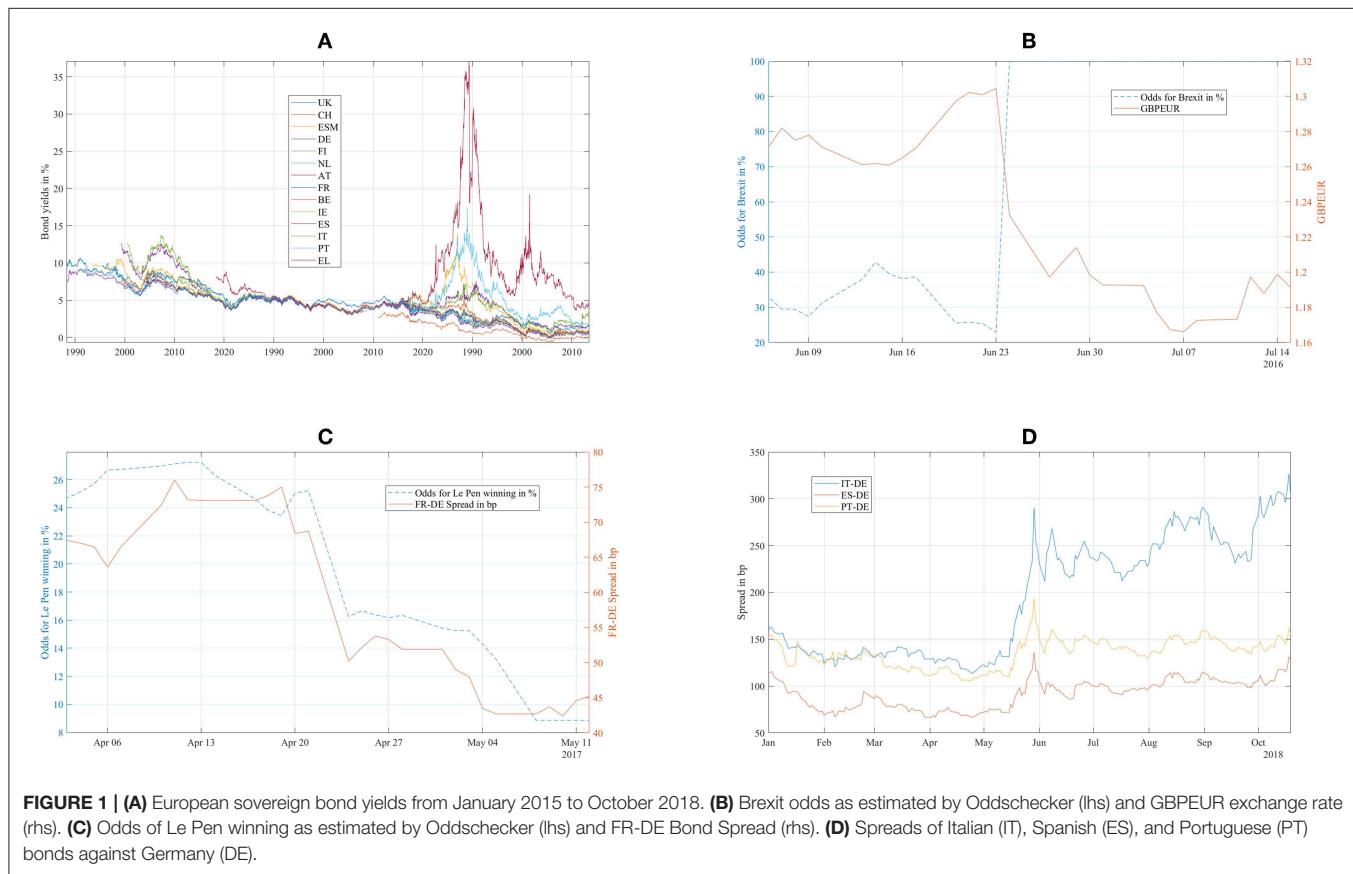
The spread increase in Euro area bonds from 2010 to 2012 has been discussed thoroughly by academia as well as by central bank research and related European institutions, for example by Beirne and Fratscher (2013) and Tola and Waelti (2015). D’Agostino and Ehrmann (2014) pointed to an overreaction of the market given the change in fundamentals and thus to a structural change in longer-term risk perception. Gross and Kok (2013), Alter and Beyer (2014), Broner et al. (2014), Glover and Richards-Shubik (2014), Shoesmith (2014), Erce (2015), Li and Waterworth (2016), Lange et al. (2017) discussed the relationships between private and public sector bonds, between sovereign bonds and credit derivatives, and the transmission channels between bank risk and sovereign risk. Gerlach-Kristen (2015), Blasques et al. (2016), Ehrmann and Fratzscher (2017), Moessner (2018), Arakelian et al. (2019) confirmed the stabilizing impact of ECB measures on bond spreads after 2012.

Many of these authors use variations of the Diebold and Yilmaz (2014) variance-decomposition framework that allows

applying network theory to interpret the time-lagged variance contributions as variance spillover effect between markets.

Schwendner et al. (2015) applied a correlation influence approach from Kenett et al. (2010). This approach does not employ a time lag structure and therefore, does not address realized variance spillover across time, but the current perception regarding spillover risk reflected in bond correlations. In contrast to correlations, the concept of correlation influence is a directed measure from a market *A* to another market *B* that explains correlations between market *B* and all other markets. A noise filter using a bootstrap scheme allows dropping the less significant correlation influences and thus to identify the markets that have the highest explanatory power regarding the correlation matrix. The authors found positive correlations dominating the European bond markets from 2004 to 2009. Between 2010 and 2012, negative correlations between the core and periphery markets had the highest explanatory power for the European bond market correlations. The situation normalized in 2013 and 2014, but negative correlations between core and periphery and negative correlation influences reappeared during the negotiations between Greece and the Eurogroup in the first half of 2015. Contagion risk and a possible breakup of the Euro area was no more an abstract risk but even used as negotiation leverage.

After the agreement to the third ESM-funded Euro area financial assistance program in July 2015, bond spreads and contagion risk declined substantially. Media focus switched to the increasing influx of refugees from Syria, Afghanistan, Iraq,



**FIGURE 1 | (A)** European sovereign bond yields from January 2015 to October 2018. **(B)** Brexit odds as estimated by Oddschecker (lhs) and GBPEUR exchange rate (rhs). **(C)** Odds of Le Pen winning as estimated by Oddschecker (lhs) and FR-DE Bond Spread (rhs). **(D)** Spreads of Italian (IT), Spanish (ES), and Portuguese (PT) bonds against Germany (DE).

and African countries to Europe that peaked in October 2015 and a wave of terrorist attacks after that. Populist parties gained substantially since then by stressing anti-immigration positions even more than anti-austerity and anti-EU postures.

Before the Brexit referendum on June 23, 2016, most studies warned of the negative economic consequences of a potential Brexit (Boettcher, 2016; EIU, 2016; Kierzenkowski et al., 2016). The unexpected Brexit outcome was explained afterwards by immigration fears and distrust in established media being more convincing than abstract rational economic arguments. The impact on bond markets was small as a decline of the British pound relative to the Euro absorbed the Brexit shock.

In the Dutch general elections on March 15, 2017, the right-wing PVV gained grounds, but finally, a four-party conservative-social-liberal coalition formed a new government in October 2017. During the presidential elections in France in spring

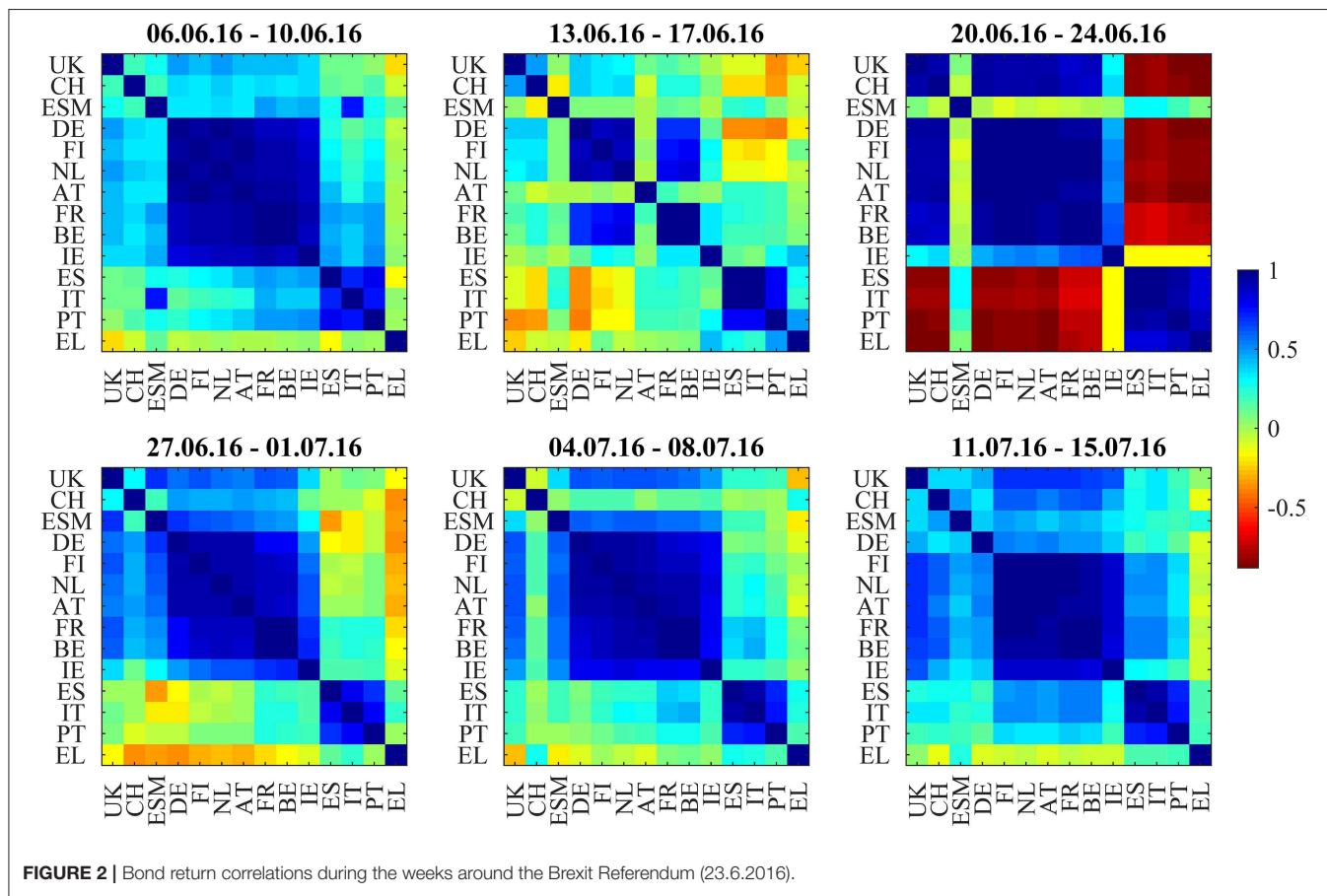
2017, the most important topics were the relationship toward the EU and immigration. The spread between French and German bonds closely followed the odds of the right-wing Marie Le Pen winning in the second round (Bird and Sindreu, 2017; Macintosh, 2017). After Emmanuel Macron won the second round on May 7, 2017, Europe embraced a wave of positive mood, and sovereign spreads declined (Whittall, 2017). The next risk scenario highlighted by the financial press (Marriage and Jennifer, 2017) was a Eurosceptic government in Italy after the next elections and a potential exit from the EU (Kelly et al., 2015).

The Italian general elections on March 4, 2018, indeed resulted in gains for the populist Five Stars Movement and the right-wing Lega, but not immediately in a new government or a sharp reaction of financial markets. Sandhu (2018) noted a large demand for Euro-denominated sovereign bonds from Asian investors who have a very low funding rate. The BTP-Bund spread widened and whipsawed during the formation phase of the new government until the end of May. Giuseppe Conte took office as a new prime minister on June 1st and confirmed increased spending commitments. During July and August, the spread lowered slightly. Italian bonds showed increasing volatility as the negotiations for the 2019 budget proceeded (O'Brien, 2018) and both parties postured against the Maastricht criteria. However, in contrast to the 2015 situation

**TABLE 1 |** Average silhouette widths for hierarchical and k-means clustering.

<i>k</i>	2	3	4	5	6	Avg
Hierarchical	0.65	0.62	0.67	0.70	0.75	0.68
k-means	0.65	0.62	0.66	0.69	0.70	0.66

The *p*-value of the *t*-test for the mean difference of the average silhouette width between the hierarchical clustering and k-means for each *k* is  $\leq 1\%$ .



with Greece, the spillover to other peripheral countries in the form of increasing Bund spreads was limited (Macintosh, 2018), despite the larger size of the Italian economy and bond market. The limited spillover was reasoned with increasing economic resilience in those countries (Pascual et al., 2018) and contrasted a 2015 “Eurozone Meltdown” risk scenario developed by Kelly et al. (2015).

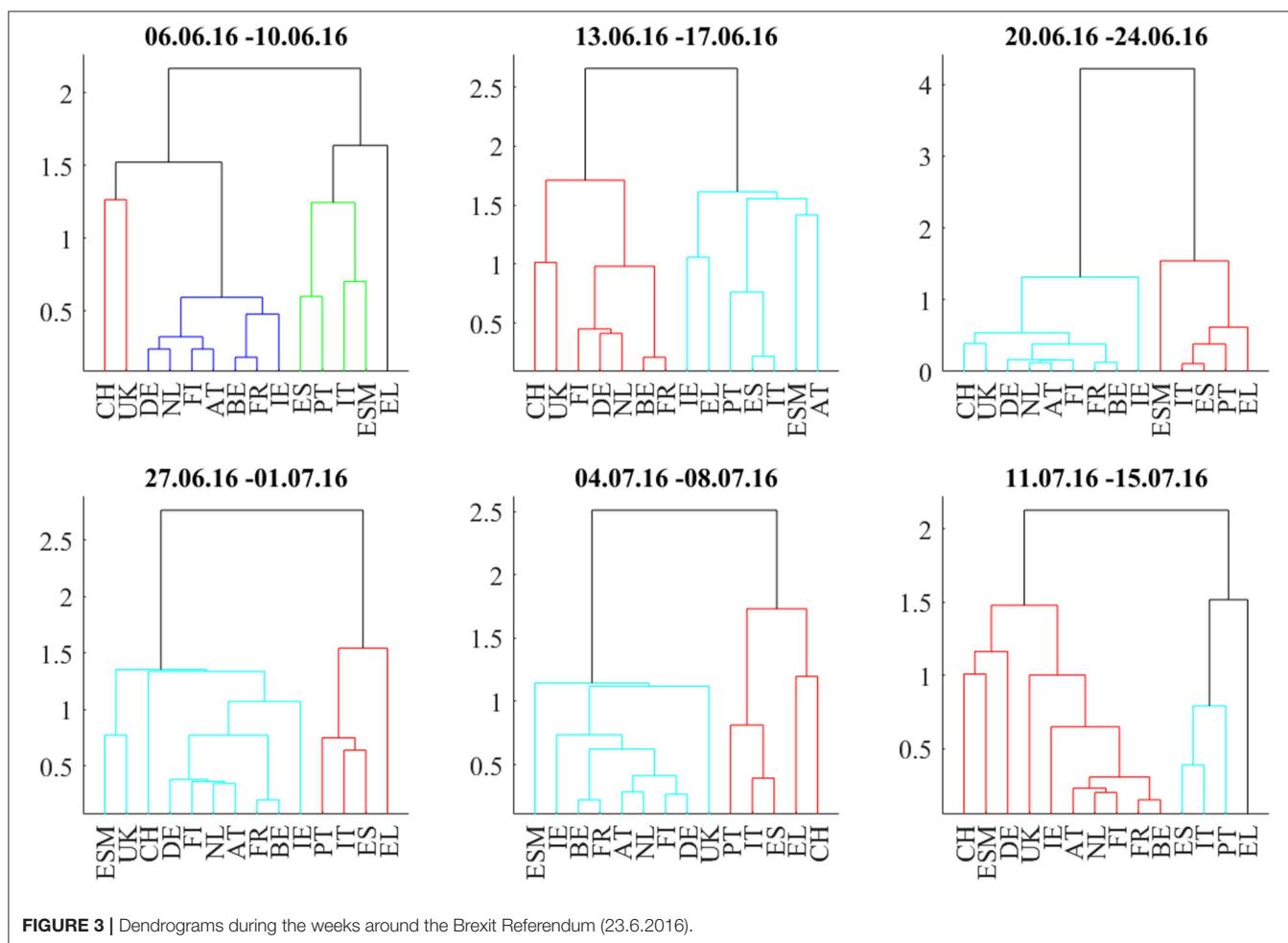
## DATA AND METHODS

The 10y bonds are the most liquid “benchmark bonds” in the sovereign bond market. For the larger European bond markets (UK, Germany, France, Italy, Spain, and Switzerland), the ICE and EUREX derivatives exchanges offer bond futures as a risk management, hedging and speculation instrument. The open interest of bond futures is much lower than the outstanding volume of bond issues, but bond futures trade at lower bid-ask spreads than bonds and don't require full funding of their market value, so they are the preferred tool for fast intraday trading. EUREX introduced the Spanish BONO bond futures as recently as 2015 as the Spanish bond yields deviated from the Italian bond yields that were previously often used as a proxy for Spanish sovereign risk (EUREX,

2018). Bond market makers often link their bond quotes to the higher-frequency bond futures market to capture short-term market movements in their bond quotes (Allen, 2018; Stafford and Allen, 2018). Therefore, the trading of bond futures instruments can have an impact on the quotes of the much larger bond market.

For this paper, we use a dataset of hourly generic 10y bond yields (**Figure 1**) from Bloomberg for UK, Switzerland (CH), ESM, Germany (DE), Finland (FI), the Netherlands (NL), Austria (AT), France (FR), Belgium (BE), Ireland (IE), Spain (ES), Italy (IT), Portugal (PT), and Greece (EL). In contrast to our 2015 paper, we added the UK to discuss the Brexit impact and Switzerland to have another non-EUR denominated reference beyond the UK. To get intraday ESM bond yields, we use the current 10y ESM benchmark price quote and compute the yields from those.

From the proprietary EFSF/ESM primary and secondary market databases (source: ESM, 2018), we got insight into the net flows of specific investor types into EFSF and ESM bonds (**Supplementary Table 3**) to investigate if risk-on/off signals that we see in the correlation patterns have corresponding flow patterns in the trade data. The flows from Asian investors are especially interesting to get an external view on the risk



**FIGURE 3 |** Dendograms during the weeks around the Brexit Referendum (23.6.2016).

and reward perception of the Euro area, even though FX dynamics may add some noise on the data. Two mechanisms let risk-reward perception having an impact on secondary market flows: the first mechanism is a so-called “flight-to-safety” reaction that lets investors shift bond positions within the Euro area, into the safe assets. For EFSF/ESM bonds, this means net bond inflows. The second mechanism is the reaction to the decision to reduce exposure to the Euro area bond market as a whole. For EFSF/ESM bonds, this means net bond outflows. These mechanisms may happen at the same time and then partially neutralize each other, meaning that some investors are shifting EUR bond exposure to EFSF/ESM and some investors are reducing their overall EUR bond exposure, including EFSF/ESM.

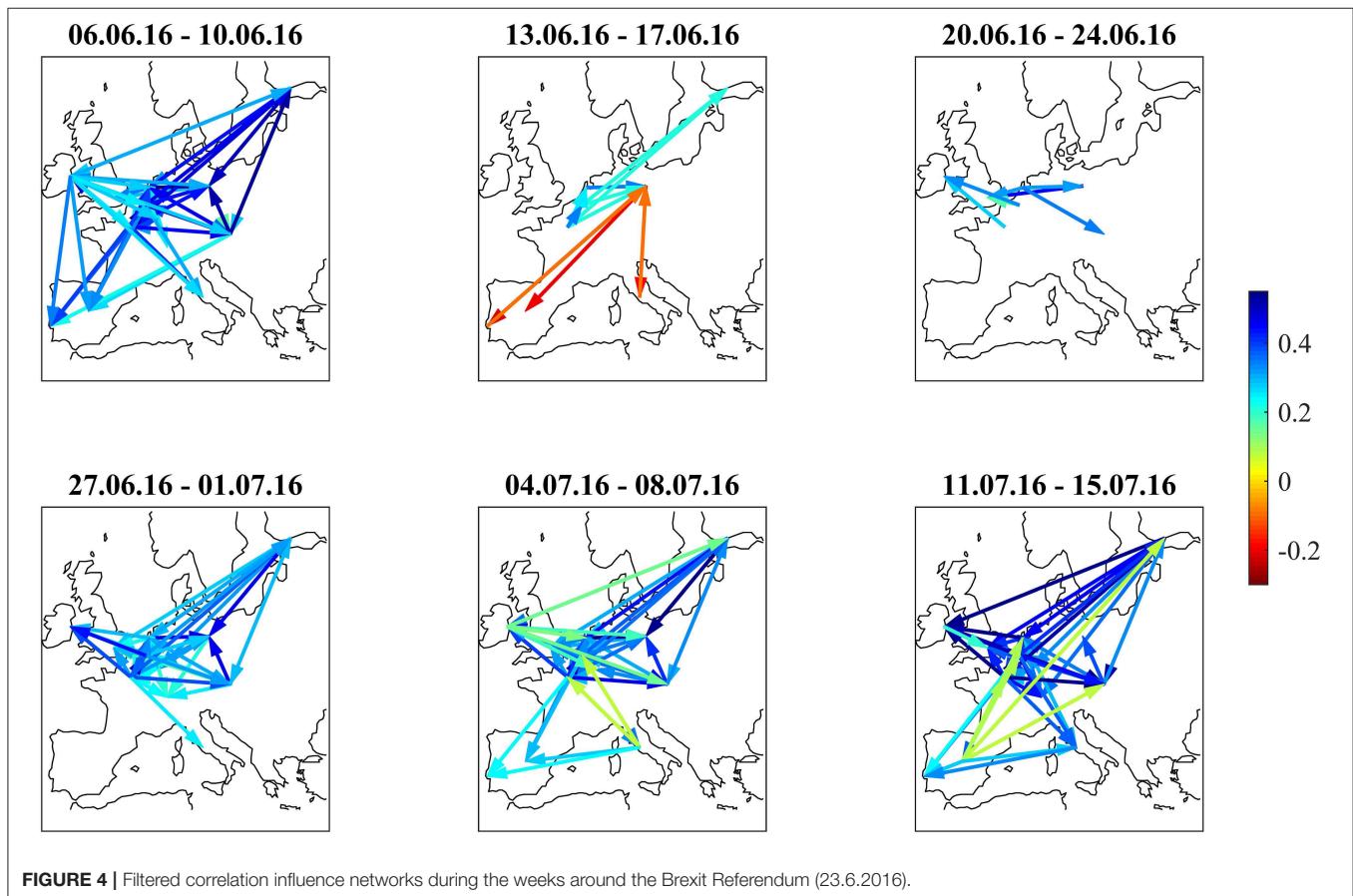
The three political situations in Europe relevant for bond markets that gained the most public interest after 2015 were the 2016 Brexit referendum, the 2017 French presidential elections, and the 2018 Italian budget negotiations. For a detailed quantitative analysis, we picked a time window of 6 weeks for each of these three situations:

- 2016 Brexit referendum: June 6, 2016, to July 15, 2016. The actual day of the referendum was on June 23, 2016.
- 2017 French presidential elections: April 3, 2017, to May 12, 2017. The first round of the elections took place on April 23, the second round on May 7.

- 2018 Italian budget negotiations: September 17, 2018, to October 26, 2018. The deadline to submit the Italian budget to the EU commission was October 15.

Following Schwendner et al. (2015), we use the Pearson correlation coefficient  $C_{ij} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sigma_i \sigma_j}$  of the bond return time series  $r_i^t$  and  $r_j^t$  between two markets  $i$  and  $j$  for 50 hourly bond returns during a window of 1 week, sampled from 08:00 to 17:00 CET. To transform the bond yield time series  $y_i^t$  into a bond return time series  $r_i^t$ , we apply a duration approximation:  $r_i^t \sim -D_i^t(y_i^t - y_i^{t-1})$  with duration  $D_i^t$  for bond  $i$  at time  $t$ .

To extract the correlation influence  $d_{i,j:k}$  from one market  $k$  to the correlations of another market  $i$  to all other markets  $j$ , we employ a definition of correlation influence  $d_{i,j:k} = C_{ij} - \rho_{ij:k}$  from Kenett et al. (2010) based on partial correlations  $\rho_{ij:k} = \frac{C_{ij} - C_{ik} C_{kj}}{\sqrt{1 - C_{ik}^2} \sqrt{1 - C_{kj}^2}}$ . If the correlation influence is positive, the return time series of market  $k$  has a positive, converging influence on the correlations between the return time series of markets  $i$  and  $j$ . If the correlation influence is negative, the return time series of market  $k$  has a negative and diverging influence on the correlation between the returns of markets  $i$  and  $j$ . We average across market  $j$  to get the average correlation influence  $d_{i,k} = \langle d_{i,j:k} \rangle_{j \neq i,k}$ . This asymmetric matrix reflects a directed graph from  $k$  to  $i$ .



To reduce the number of directed links in the resulting correlation influence network, we employ a bootstrap (Efron, 1979) filter that only retains the directed links  $k \rightarrow i$  if and only if  $|d_{i,k}| > Q \times \sigma_{bootstrap}(d_{i,k})$  with a parameter  $Q = 3$ .  $Q$  is not a convergence parameter, as it only filters out more links at a higher  $Q$ . We compute  $\sigma_{bootstrap}(d_{i,k})$  with a resampling with the synchronous replacement of the cross-section of bond returns. Following Politis and Romano (1992), we draw the block length from a uniform distribution between 1 and 10 h for each sample to account for serial correlation.

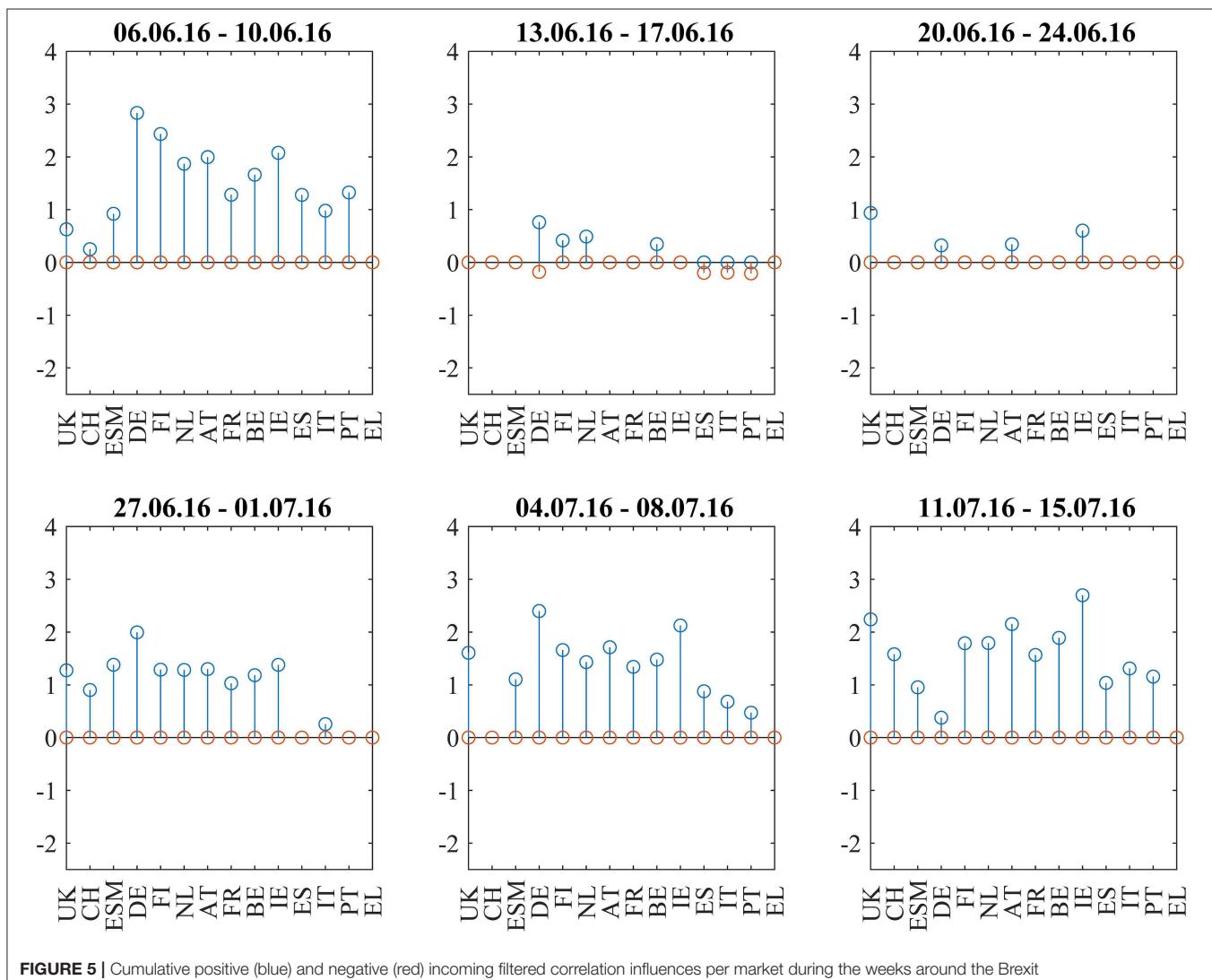
This method does not involve a time lag between the time series of the respective markets and thus addresses only synchronous effects. In contrast to Beetsma et al. (2017) and Van Der Heijden et al. (2018), the news events themselves are not explicitly part of the model.

Partial correlations have also been employed by Saroyan and Popayan (2017) to analyse risk spillover between European bank and sovereign credit risk. They find contagion from other

countries to the correlations between the CDS spreads of banks and the sovereign bonds of their home country and recommend non-zero risk weights for sovereign bond holdings of banks.

Giudici and Parisi (2018) integrated partial correlation networks into a structural VAR model, labeled CoRisk approach. They find high contagion risk for peripheral countries from other peripheral countries, but low contagion risk between core and periphery. These findings confirm our results of a strong core-periphery segmentation, visible in the persistent block structure of the bond return correlation matrices.

To enable a more detailed discussion of this block structure, we analyse the blocks using a non-parametric clustering method. We apply a hierarchical clustering method (Ward, 1963) using the distance matrix metric  $G_{ij} = \sqrt{2(1 - C_{ij})}$  as a function of the bond return correlation matrix  $C_{ij}$  according to Gower (1971). This choice of the distance metric preserves the sign of the correlation coefficients, which is important as we specifically want to discriminate positive from negative correlations. In



**FIGURE 5 |** Cumulative positive (blue) and negative (red) incoming filtered correlation influences per market during the weeks around the Brexit Referendum (23.6.2016).

contrast to the standard portfolio management literature, negative correlations are not an opportunity for diversification, but a warning signal in the specific case of this dataset as they appear between Euro area sovereign bonds that should be benchmark instruments without default risk.

To assess the quality of the hierarchical clustering compared to a simpler  $k$ -means clustering algorithm, we employ the “average silhouette width” criterion as suggested by Rousseeuw (1987). According to Rousseeuw, a higher number for the average silhouette width points to a more appropriate clustering. **Table 1** shows a comparison of the average silhouette widths of the hierarchical and the  $k$ -means clustering for different values of  $k$ . For larger values of  $k$ , hierarchical clustering shows higher average silhouette widths. The null hypothesis of hierarchical clustering not leading to higher average silhouette widths than  $k$ -means clustering could be rejected with a  $p$ -value of 1% for the dataset given by the three discussed time periods and  $k$  values from 2 to 6.

From the viewpoint of the specific application domain of European bonds, hierarchical clustering has the additional advantage of making overlapping correlation blocks visible. Following Gower and Ross (1969) and Mantegna (1999), we present the membership of the various bond markets to a hierarchy of clusters using a dendrogram. The clusters at the lowest levels of the dendrogram correspond to the most pronounced blocks in a correlation matrix. We found almost

the same clusters using “complete linkage” or “single linkage” methods instead of Ward’s method.

The advantage of a dendrogram compared to a heatmap is the objective representation of the clusters, as they are sorted in clusters according to the distance metric, whereas the visual impression of a correlation matrix as a heatmap depends on the predefined ordering. This ordering may depend on subjective beliefs or a market practice to sort issuers into a tiered hierarchy.

## DISCUSSION

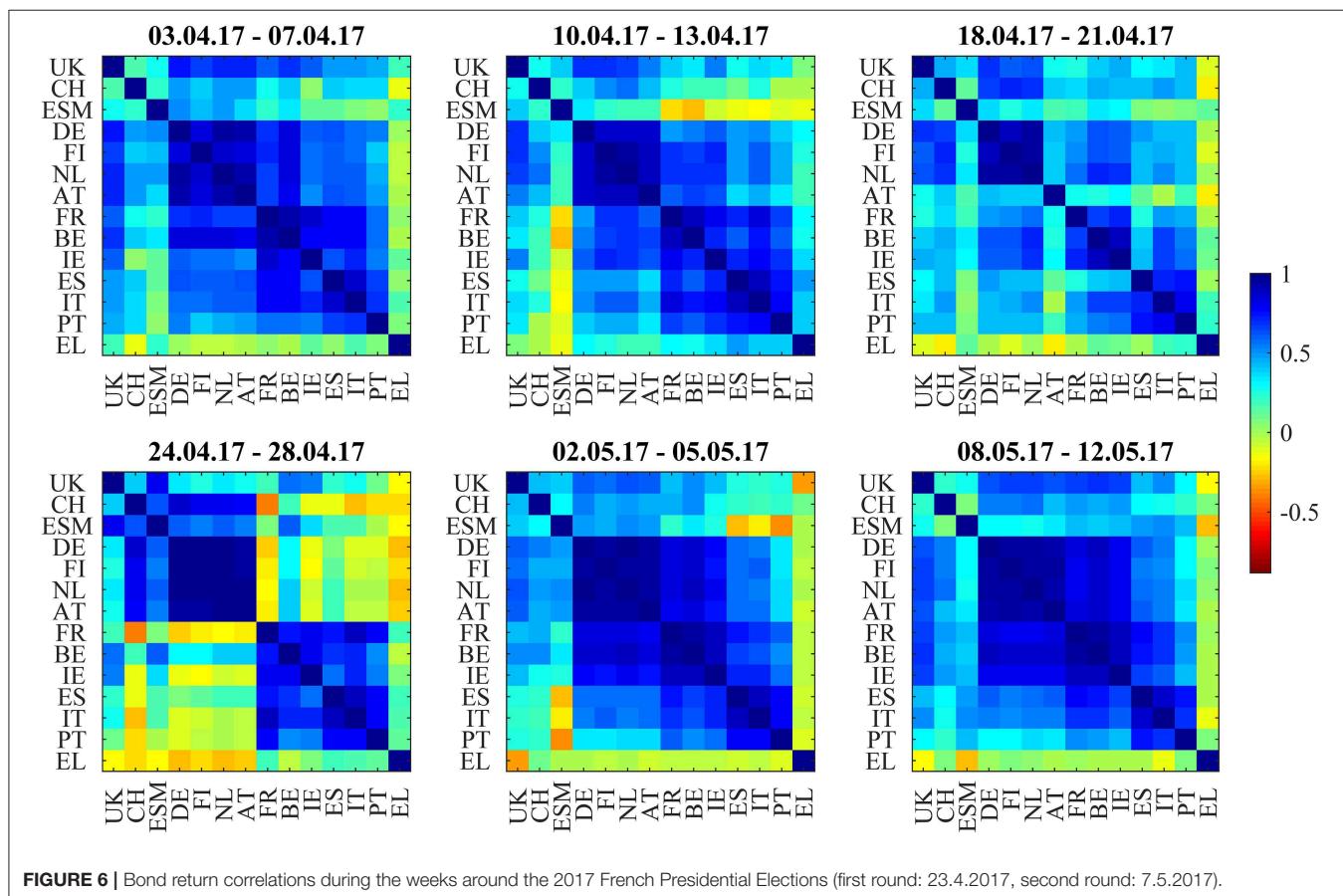
In the Discussion section, we discuss the bond return correlation matrices, hierarchical clusters and filtered correlation influence networks for the three political situations “Brexit referendum,” “French presidential elections,” and “Italian budget negotiations” as main results. A supplementary spreadsheet offers more technical details:

**Supplementary Table 1** shows the correlation matrices as numbers.

**Supplementary Table 2** shows the filtered average correlation influences as numbers.

**Supplementary Table 3** shows investor flows in EFSF/ESM bonds.

**Supplementary Figure 1** shows the results of  $k$ -means clustering with  $k = 4$ .



**Supplementary Figure 2** shows silhouette widths for  $k$ -means and hierarchical clustering for different values of  $k$  to compare the performance of both clustering methods.

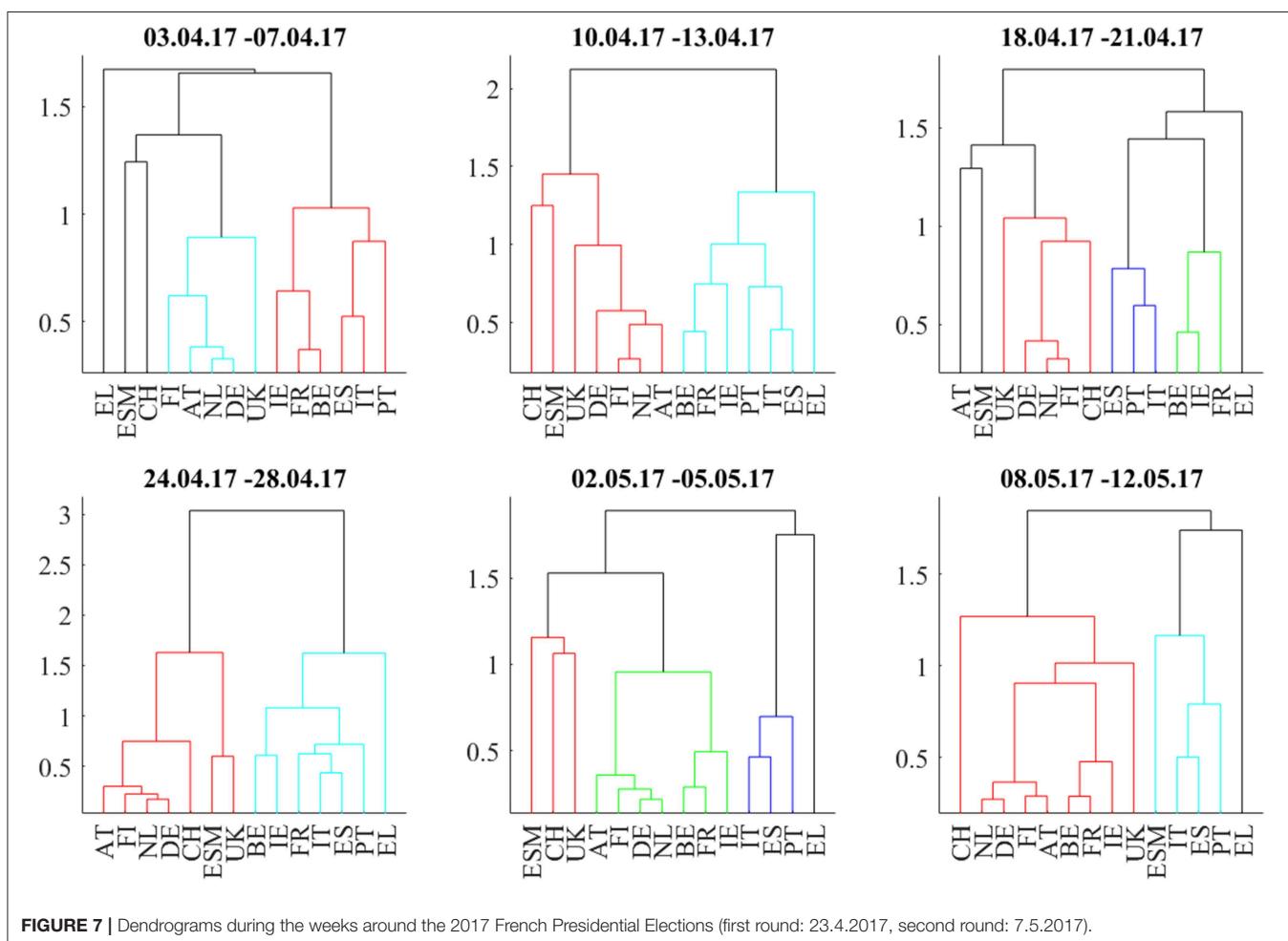
**Supplementary Figure 3** shows the cumulative outgoing filtered correlation influences per market.

## Brexit Referendum

We discuss the first situation describing the weeks around the 2016 Brexit referendum using **Figures 1B, 2–5**: **Figure 1B** shows the odds of the “leave” outcome as estimated by the British bookmaker odds comparison service “Oddschecker” (Bloomberg ticker: ODCHLEAV Index) and the GBP exchange rate. In the weeks before the referendum, the odds for “leave” hovered in a range between 23 and 43%. The British pound exchange rate against the Euro inversely mirrored these odds. After the referendum, the odds massively underestimated the outcome and jumped from 23 to 100%, with the British pound losing almost 9% against the Euro in 2 days. **Figure 2** shows the correlation matrix of hourly bond returns during the weeks before, during and after the referendum (June 23). **Figure 3** shows the results of Wards’ hierarchical clustering as dendograms. **Figure 4** presents the filtered correlation influence networks during the same weeks on geographical maps. **Figure 5** shows the cumulative positive

(blue) and negative (red) *incoming* filtered correlation influences per market. The *outgoing* filtered correlation influences are shown in **Supplementary Figure 3**.

Two weeks before (June 6–June 10) the referendum, the correlation matrix showed strong positively correlated core/semi-core and periphery blocks, and positive to neutral correlations between core/semi-core and periphery. UK bonds show weak positive correlations to the European core and semi-core. The core/semi-core block has only a very weak substructure. Irish bonds belong to the core/semi-core block. The dendrogram for this week confirms the block structure. The  $k$ -means clustering assigns a discrete cluster number from 1 to 4 to each of the bond markets but does not relate the four clusters to each other. The  $k$ -means cluster assignments are roughly consistent with the results from the hierarchical clusters but deliver a more “binary” view. For example, Italy belongs to the ESM cluster in both clusterings, but only the hierarchical clustering shows the tight coupling of Italy to Spain and Portugal one hierarchy level above. Throughout the 6 weeks with very few exceptions, we see, in the dendograms, Greece, Portugal, Spain, and Italy as main constituents of the periphery block, Germany, Netherlands, Finland, and Austria as main “core” countries and Belgium, France, and Ireland as main “semi-core” countries.



Interestingly, UK stays very close to the core block, as well as Switzerland. ESM is also part of the core block except for the Brexit week where it was hierarchically part of the periphery. It moved back to the core a week later, after worries about the further European integration had quickly calmed down.

The correlation influence network shows strong connections within and between core and periphery.

During the week directly before (June 13–June 17) the referendum, the smaller issuers ESM, Austria, and Ireland decorrelate. Spain, Italy, and Portugal develop slightly negative correlations to Germany. Portugal also shows slightly negative correlations to British and Swiss bonds. The dendrogram for this week shows members leaving the clusters compared to the week before. The network (**Figure 4**, second panel) shows negative filtered correlation influences between Germany and the three peripheral countries Spain, Italy, and Portugal. These negative influences are statistically significant, as they pass the noise filter, but of small amplitude (**Figure 5**, second panel). Only a few core countries are affected by positive correlation influences.

The week of the referendum (June 20–June 24) induced strong positive correlations within the core and periphery blocks, and very strong negative correlations between core and periphery. UK and Swiss bonds were highly correlated to the “core Europe” block and thus also negatively correlated to the Euro area periphery (Spain, Italy, Portugal, Greece). The British currency absorbed the negative shock of the referendum to the UK

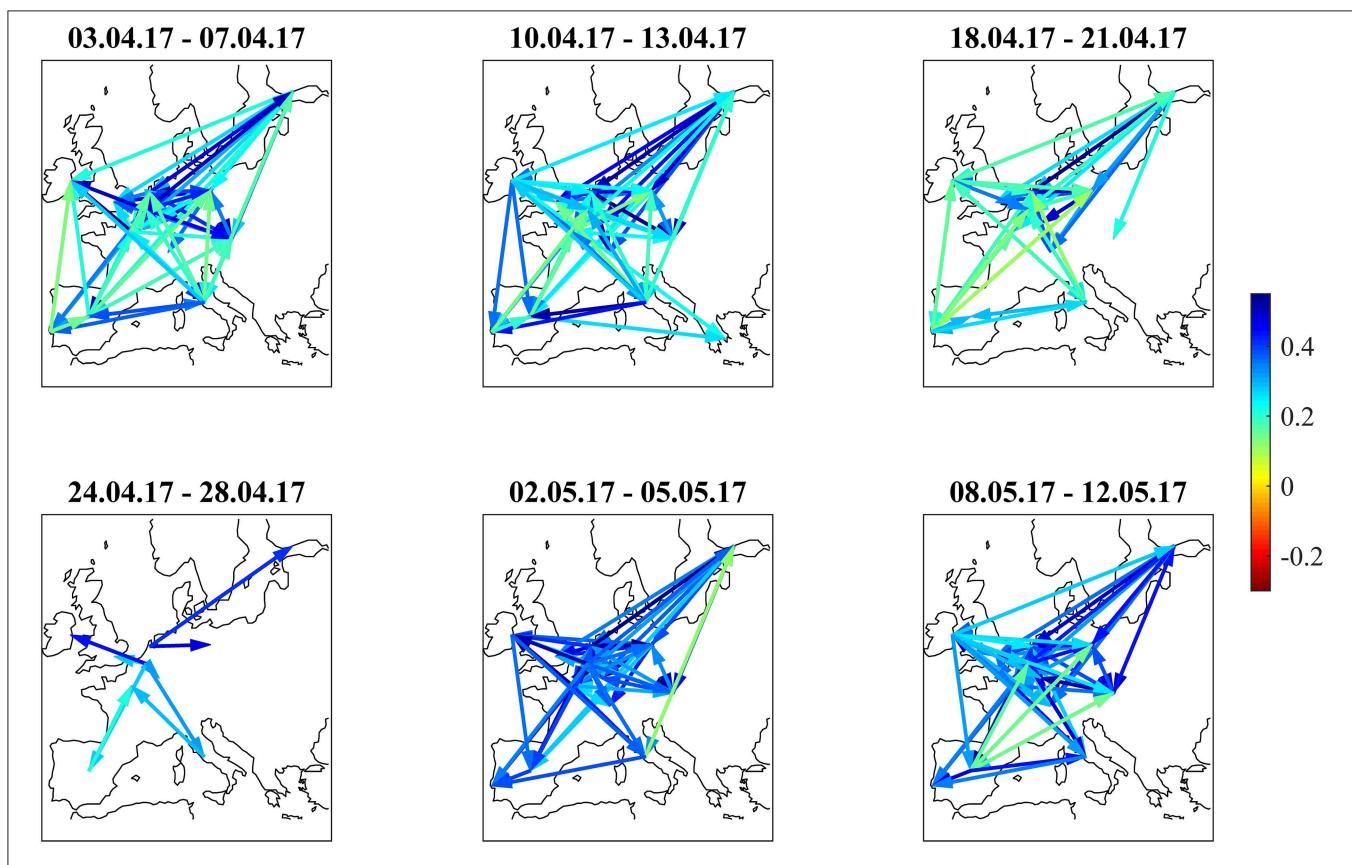
economy. British bonds even gained in market value, consistent with the core Euro area bonds. The dendrogram of this week confirms the strong core-periphery segmentation. The network shows only a few connections that pass the noise filter.

During the 3 weeks after the referendum (June 27–July 15), correlations returned to the first week in the panel. Irish bonds return to the core/semi-core block. The first and the last week of the correlation matrix panel look very similar, also the dendrograms and networks.

From June 6 to July 1, the net flows from Asian investors into EFSF/ESM bonds were balanced. Two weeks and 3 weeks after the referendum (July 4–July 8 and July 11–July 15), net flows were negative at about  $-0.5$  bn EUR, respectively. These flows after the referendum may be completely independent of the political event, or they may be a reversed flight-to-safety reaction (i.e., outflows from the safe haven when the political situation normalizes).

## French Presidential Elections

The second situation begins 3 weeks before the first round of the 2017 French presidential elections and ends 1 week after the second round. **Figure 1C** shows the odds of Le Pen winning from Oddschecker (Bloomberg: ODCHFRML Index) together with the spread of 10Y French bonds vs. 10Y German bonds. The spread decreases from 73 bp with the Le Pen odds until 50 bp at the first round (April 23, resulting in the second round between Le Pen

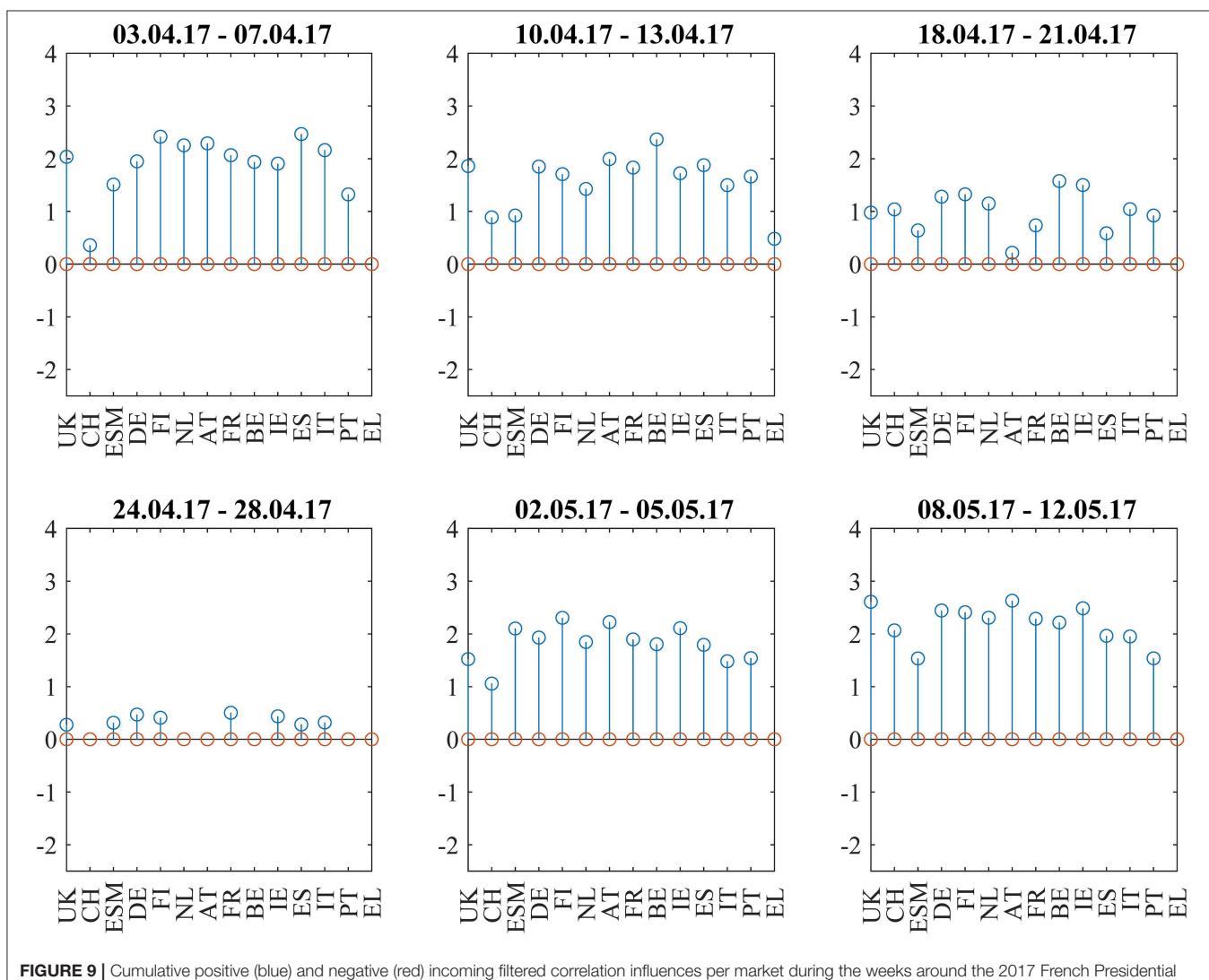


and Macron) and then further until 43 bp at the second round (May 7, resulting in the victory of Macron).

**Figure 6** shows the bond return correlations as heatmaps. As the political position of France within the EU was an important topic of the elections, the position of French bonds within the European tier structure was a trading topic. The market challenged the usual structure of a “core” block (DE, FI, NL, AT), a “semi-core” block (FR, BE, IE), and a “periphery” block (ES, IT, PT). Especially in the week immediately after the first round (April 24–28), France was part of a “semi-core plus periphery” block (FR, BE, IE, ES, IT, PT) and showed slightly negative correlations to Swiss bonds. After that, the block structure normalized. The dendograms in **Figure 7** confirm the “semi-core plus periphery” block in a corresponding hierarchy. A similar hierarchy is already visible in the second panel (April 10–April 13) of **Figure 7**. The dendograms hence show that the uncertainty around France was affecting the “semi-core” block as a whole. Uncertainty stopped 1 week after the first round when the other candidates endorsed Macron such that it became likely

that he would win the second round. The correlation influence networks in **Figure 8** confirm the weakening of the established block structure until April 28 and recovery to an almost fully positively connected network afterwards. In contrast to the 2015 Greek negotiations and the 2016 Brexit referendum, there are no negative correlation influences during these 6 weeks (**Figure 9**).

The net flows of Asian investors into EFSF/ESM bonds are substantially positive (+384 mln EUR) in the week from April 3 to April 7 and in the week after the first round (+251 mln EUR from April 24 to April 28). The net selling in this week is most probably a technical flow: investors swap old bonds to the new issuance. Important is here the positive net volume, showing additional buying of the issued volume.). After that, they are negative during the weeks before and after the second round (−166 mln EUR from May 2 to May 5 and −133 mln EUR from May 8 to May 12). We interpret the data as a flight-to-safety movement with a reversal after the result from the second round: Asian investors were, in sight of a political event with an uncertain outcome, increasing their “core block” exposure



(where the correlations clearly show that EFSF/ESM belong to) at the cost of peripheral bonds. Consistent with this interpretation, French bonds traded at a 30 bp risk premium to the yield of ESM bonds at the beginning of 2017. This spread decreased to zero at the end of the second quarter of 2017, as it did with respect to other core block bonds such as Bunds.

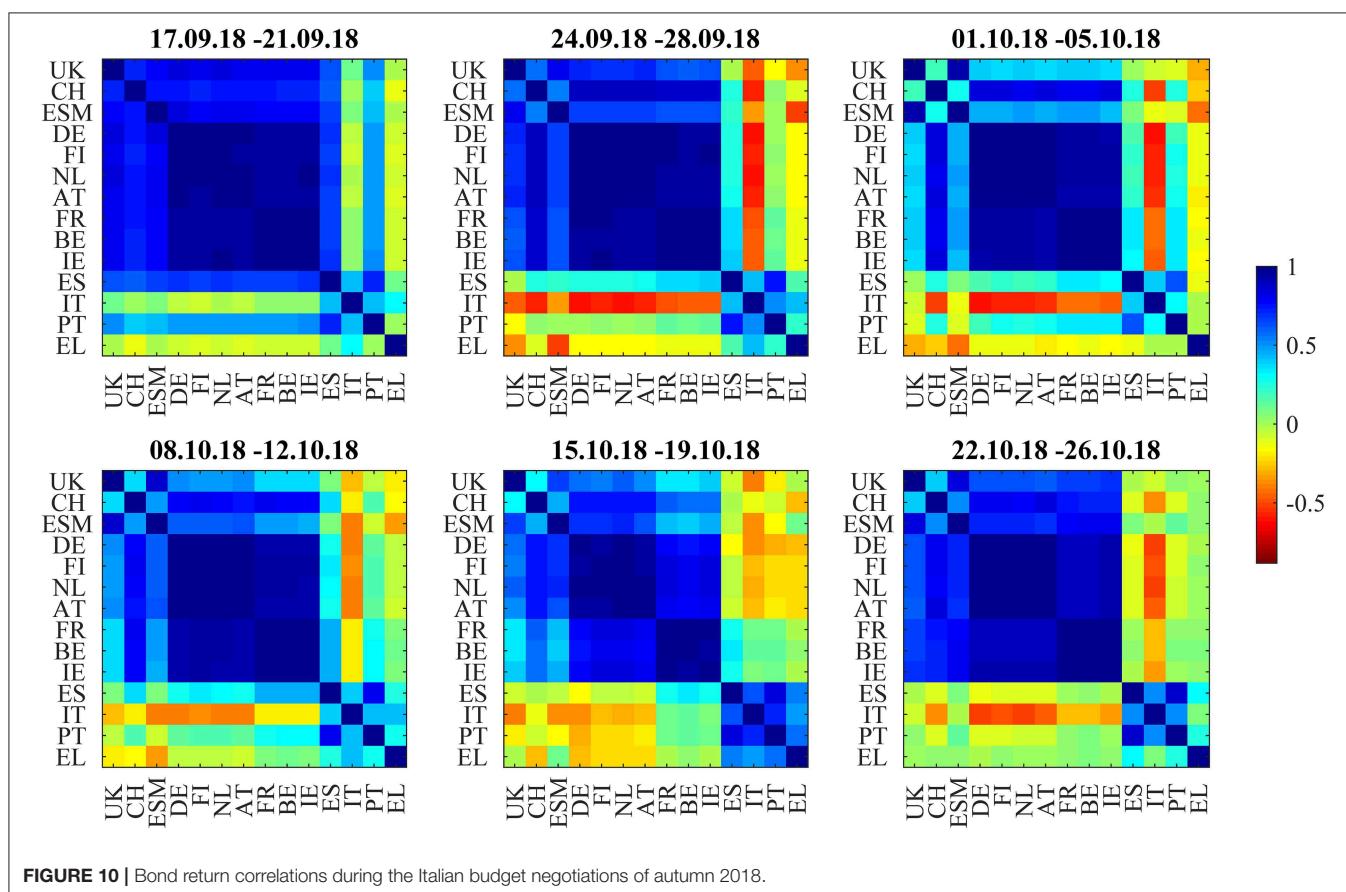
## Italian Budget Negotiations

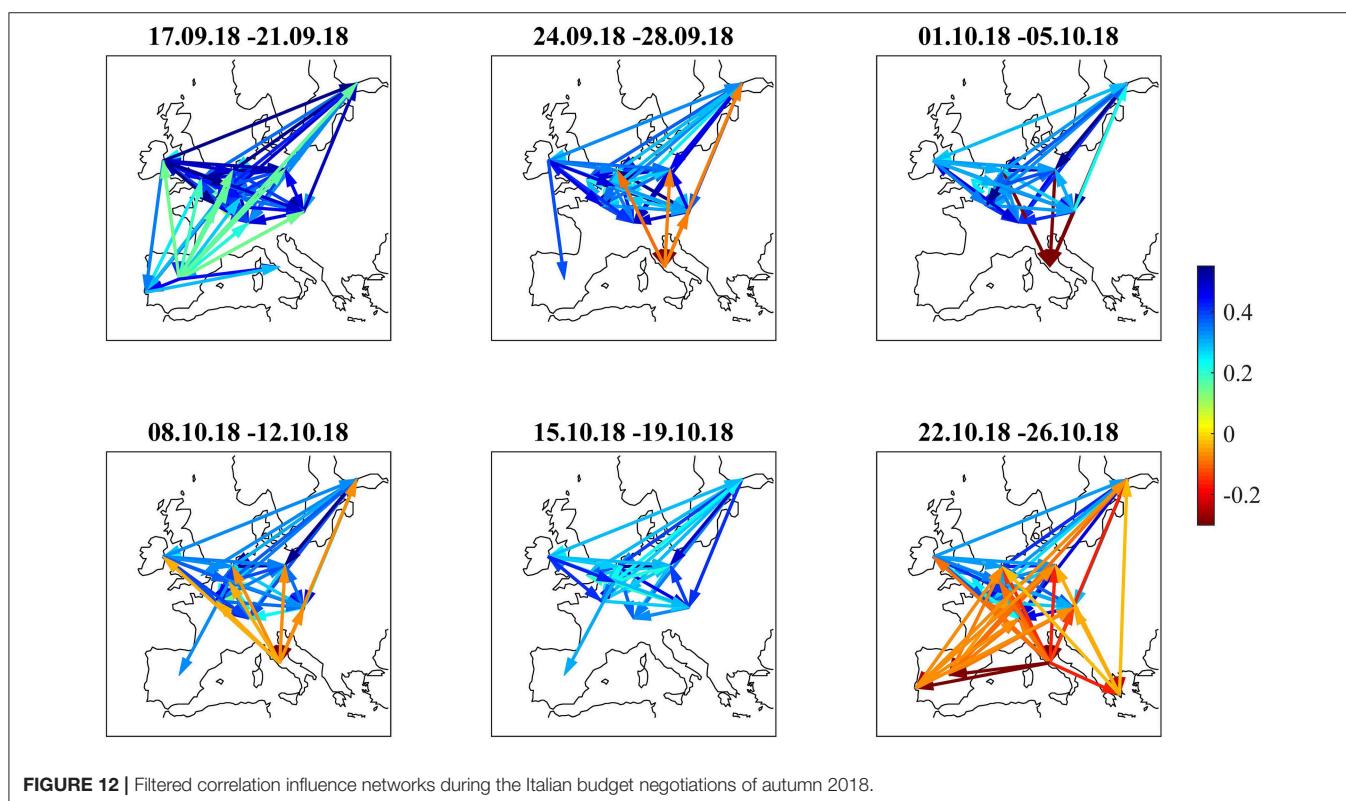
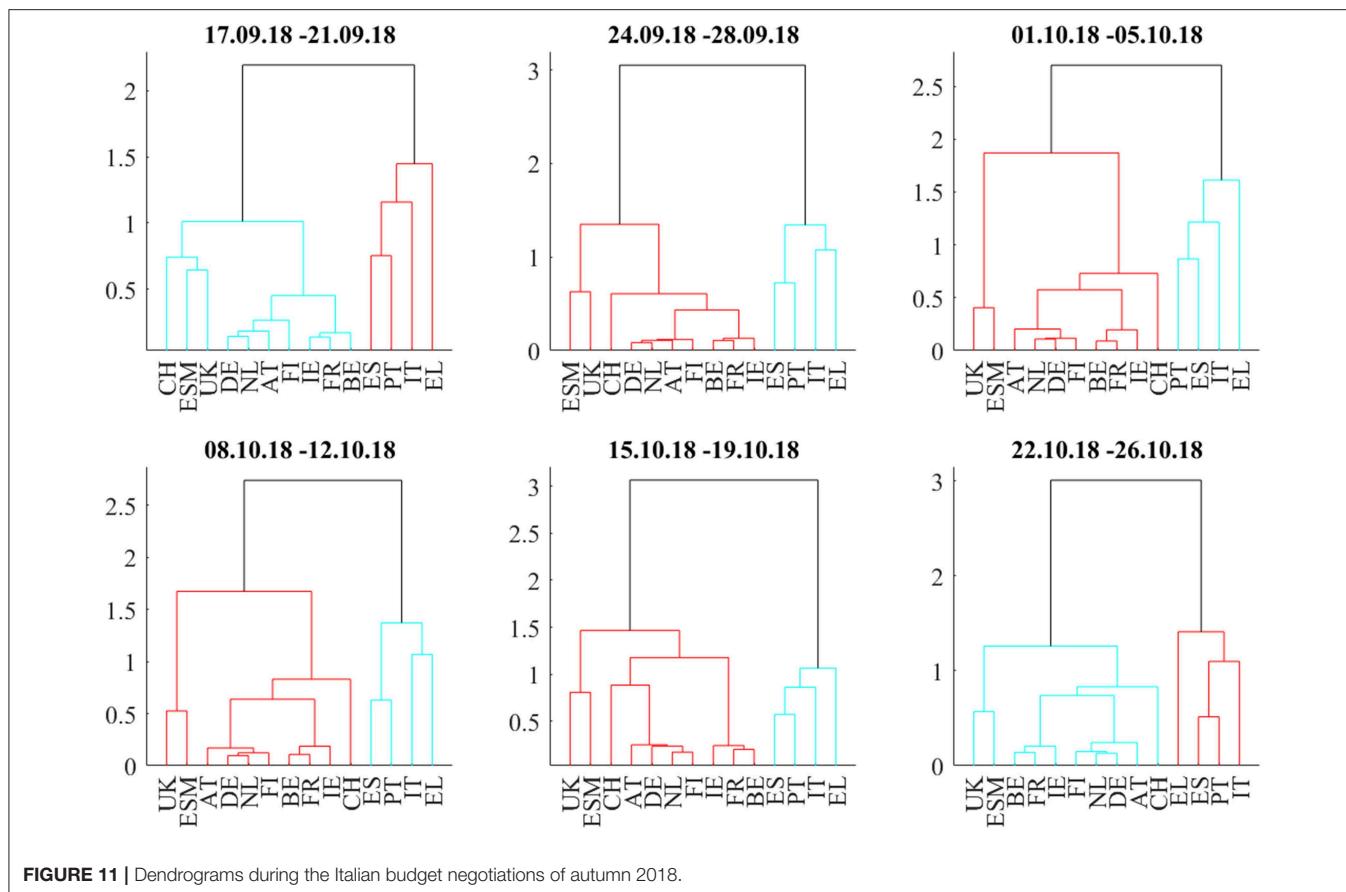
**Figure 1D** shows the main observable of Italian fiscal and EU political discussions, the spread between Italian and German 10y bonds (IT-DE) from January to October 2018. At the beginning of the year, the spread was at 150 bp on par with the spread of Portuguese bonds (PT-DE) and about 50 bp higher than the spread of Spanish bonds (ES-DE). After the electoral success of Five Stars and Lega in early March, the Italian spread decorrelated from Portugal and Spain. As the new government was set up at the end of May, the spread widened by an additional 100 bp. During the negotiations within the new government about the budget given the electoral promises to increase spending and frequent postures against the EU budget rules, the spread showed increased volatility in several waves until October 19 when it reached 336 bp. Portuguese and Spanish bonds traded in much lower ranges, showing only mild contagion.

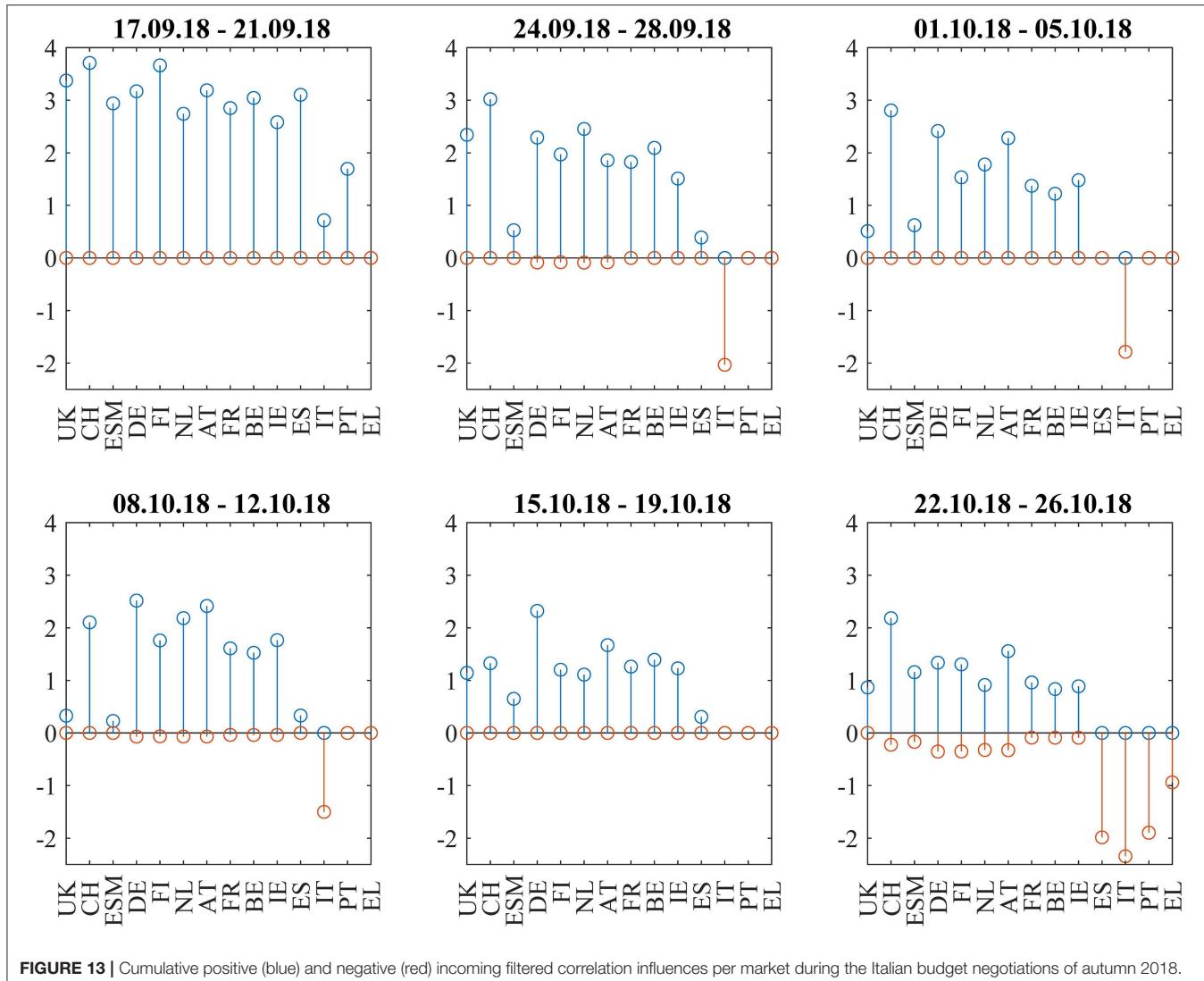
In **Figure 10**, the correlation heatmaps show positive correlations within and between the core and semi-core blocks and positive correlations to the ESM bonds and the non-Euro

denominated UK and Swiss bonds throughout the full 6-week period from September 17 to October 26. The boundary between the core and semi-core block is barely visible but consistent. The correlations of the two peripheral countries, Spain and Portugal, to the semi-core countries are between neutral and strongly positive. The correlations between Italy and the core (AT, DE, FI, NL) and semi-core (BE, FR, IE) are between neutral and strongly negative. Greece (EL) decouples and sometimes shows negative correlations to ESM, CH, and UK bonds. The dendrograms in **Figure 11** confirm the consistent core and semi-core blocks, the strong coupling between Spain and Portugal and the isolated role of Italian bonds until the third week. During the fourth week (October 8–12), Italy forms a cluster with Greece. In the fifth week (October 15–19), a periphery block with Spain, Italy, Portugal, and Greece is visible both in the correlation matrix and in the dendrogram. This block weakens in the last week (October 22–26). It is noteworthy that the block structure “core,” “semi-core,” and “periphery” remained constant through the observation period in the dendrograms. The intact block structure means that every yield movement on the Italian bond market affected the other peripheral markets more as markets belonging to the other blocks. In other words, while the level of correlation and influence changed within the observation period, the fundamental structure remained unchanged.

The correlation influence graphs in **Figure 12** show strong positive influences between the core and semi-core







**FIGURE 13 |** Cumulative positive (blue) and negative (red) incoming filtered correlation influences per market during the Italian budget negotiations of autumn 2018.

countries and toward Spain and Portugal in the first week, whereas Italian bonds couple positively to Spain. In the second week (September 24–28), all core countries develop negative correlation influences toward Italy. This sentiment improvement is confirmed in the third week (October 1–5). During the fourth week (October 8–12), there are negative correlation influences between Italy and all core and semi-core countries. Spain recoupled to the semi-core in the fourth week. During the week from October 15 to October 19, positive correlation influences within core and semi-core bonds passed the noise filter. The budget submitted by the Italian government on October 16 was rejected on October 18 by the EU commission.

During the last week (October 22–26), Equities sold off as the EU commission formally requested the Italian government to revise their budget within 3 weeks. Negative correlation influences were visible between the core European block and all peripheral countries and from Italy to the rest of the periphery.

The amplitudes of these negative correlation influences are larger (**Figure 13**) than during the Brexit referendum and French election cases. This pattern echoes the frequent spillover patterns during the 2015 negotiations between the Eurogroup and Greece (Schwendner et al., 2015).

The net flows of Asian investors into EFSF/ESM bonds were close to zero in the period from September 17 to October 5. In the week from October 8 to October 12, net selling of 187 mln EUR was overcompensating primary purchases of 136 mln EUR. In the week from October 15 to October 19, there was net buying of more than 1 bln EUR, more than 90% of it on the primary market. In the week after that, we saw only little net inflows of 91 mln EUR on the secondary market. These flows reflect the increasing buying from Asian investors in the fourth quarter of the year; still the volume in the time window of this case study was above average. On the background of the political scenery, the inflows may be attributed to steady investment in quality, if not even be interpreted as flight-to-safety, taking into

account the above-average volume. Flight-to-safety movements usually happen at a higher pace than the reverse ones since risk protection usually has more urgency than the relaxation of risk protection measures. Also, there has not been any strong political signal letting investors move toward a “risk-off” mode. Hence, we do not see any reverse flight-to-safety in the observation period of 6 weeks.

## CONCLUSION

In an empirical study, we discussed the European bond market return correlations in three prominent events during 2016–2018. In contrast to the frequent spillover patterns that happened during the negotiation between the Eurogroup and Greece in 2015 (Schwendner et al., 2015) about the third financial assistance programme, the patterns around the 2016 Brexit referendum, the 2017 French presidential election and the 2018 budget negotiations in Italy were different.

The 2016 Brexit referendum only caused a muted warning signal in the form of negative correlation influences from German to Spanish, Italian, and Portuguese bonds in the week before the referendum and stronger core-periphery distortions with volatile correlations during the week of the referendum due to the unexpected result. The pattern of negative correlation sentiment reversed quickly. However, the devaluation of the British pound remained.

The 2017 French presidential elections showed a merge between the semi-core correlation block and the periphery correlation block before the second round, but no negative correlations or correlation influences between core and periphery. The French bond spreads improved after the second round.

Finally, the Italian budget negotiations in autumn 2018 showed increased spreads for Italian bonds and negative correlation influences between core Europe and Italy. During the last week from October 22 to 26, a significant pattern of negative correlation influences from core Europe and Italy to the rest of the periphery was visible.

Interpreting the primary and secondary market aggregated net flows of Asian investors in the context of euro area bond correlations, we observe an interesting relation: we saw flight-to-safety patterns into ESM bonds in the two case studies where ESM was, in terms of correlations, part of the core block. In

contrast, during the week of the Brexit referendum, the ESM correlations did not show significant relations, and the flows did not show clear patterns. With the quick calming down of the markets, the normal core structure with ESM being part of it was visible again.

## AUTHOR CONTRIBUTIONS

MS implemented the analytics. PS wrote the main parts of the paper and produced the figures. MH contributed to the discussion section. All authors are accountable for the content of the work together.

## FUNDING

This research has received funding from the European Unions Horizon 2020 research and innovation program FIN-TECH: A Financial supervision and Technology compliance training programme under the grant agreement No. 825215 (Topic: ICT-35-2018, Type of action: CSA).

## ACKNOWLEDGMENTS

We thank Siegfried Ruhl, Rolf Strauch, Peter Dattels, Daniel Hardy, Juan Rojas, Aitor Erce, Dimple Bhawsar, Roel Beetsma, and Paolo Giudici for inspiring discussions. We are also grateful for valuable comments received from participants of the 9th Financial Risks International Forum at Bachelier Institute (Paris, March 2016), the ADEMU Risk-Sharing Mechanisms for the European Union Workshop at European University Institute (Florence, May 2016), the Mathfinance conference (Frankfurt, April 2017), the 3rd COST Conference for Artificial Intelligence in Finance and Industry (Winterthur, September 2018) and two Financial Evolution—Sentiment Analysis, AI, and Machine Learning UNICOM conferences (Zurich, October 2018 and London, June 2019) and from the two reviewers.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frai.2019.00020/full#supplementary-material>

## REFERENCES

- Allen, K. (2018, September 25). European sovereign bonds change hands less often. *Financial Times*. Available online at: <https://www.ft.com/content/c5875f50-bd9c-11e8-94b2-17176fb93f5>
- Alter, A., and Beyer, A. (2014). The dynamics of spillover effects during the european sovereign debt turmoil. *J. Bank. Finan.* 42, 134–153. doi: 10.1016/j.jbankfin.2014.01.030
- Arakelian, V., Dellaportas, P., Savona, R., and Vezzoli, M. (2019). Sovereign risk zones in Europe during and after the debt crisis. *Quant. Finance* 20:1562197. doi: 10.1080/14697688.2018.1562197
- Beetsma, R., de Jong, F., Giuliodori, M., and Widijanto, D. (2017). Realized (co)variances of eurozone sovereign yields during the crisis: the impact of news and the Securities Markets Programme. *J. Int. Money Finance* 75, 14–31. doi: 10.1016/j.intmonfin.2017.04.003
- Beirne, J., and Fratscher, M. (2013). The pricing of sovereign risk and contagion during the european sovereign debt crisis. *J. Int. Money Finance*. 34, 60–82. doi: 10.1016/j.intmonfin.2012.11.004
- Bird, M., and Sindreu, J. (2017, February 25). Europe's periphery debt market welcomes a new member: France. *Wall Street Journal*. Available online at: <https://www.wsj.com/articles/europe-s-periphery-debt-market-welcomes-new-member-france-1487604548>
- Blasques, F., Koopman, S. J., Lucas, A., and Schaumburg, J. (2016). Spillover dynamics for systemic risk measurement using spatial financial time series models. *J. Econ.* 195, 211–223. doi: 10.1016/j.jeconom.2016.09.001

- Boettcher, W. (2016). *Logic Dictates*. London: Colliers International.
- Broner, F., Erce, A., Martin, A., and Ventura, J. (2014). Sovereign debt markets in turbulent times: creditor discrimination and crowding-out effects. *J. Monetary Eco.* 61, 114–142. doi: 10.1016/j.jmoneco.2013.11.009
- D'Agostino, A., and Ehrmann, M. (2014). The pricing of G7 sovereign bond spreads—the times, they are a-changin. *J. Bank. Finance* 47, 155–176. doi: 10.1016/j.jbankfin.2014.06.001
- Diebold, F. X., and Yilmaz, K. (2014). On the network topology of variance decompositions: measuring the connectedness of financial firms. *J. Econ.* 182, 119–134. doi: 10.1016/j.jeconom.2014.04.012
- Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Stat.* 7, 1–26. doi: 10.1214/aos/1176344552
- Ehrmann, M., and Fratzscher, M. (2017). Euro area government bonds – Fragmentation and contagion during the sovereign debt crisis. *J. Int. Money Finance* 70, 26–44. doi: 10.1016/j.intmonfin.2016.08.005
- EIU (2016). *Out and Down. Mapping the Impact of Brexit*. The Economist Intelligence Unit.
- Erce, A. (2015). “Bank and sovereign risk feedback loops,” in *ESM Working Paper* (Luxembourg).
- ESM (2018). “Proprietary primary and secondary market database for EFSF/ESM bonds,” in *European Stability Mechanism* (Luxembourg).
- EUREX (2018). *Euro-BONO Futures: Spanish Government Bond Futures*. Frankfurt am Main: Eurex Frankfurt AG.
- Forbes, K. J., and Rigobon, R. (2002). No contagion, only interdependence: measuring stock market comovements. *J. Finance* 57, 2223–2261. doi: 10.1111/0022-1082.00494
- Gerlach-Kristen, P. (2015). The impact of ECB crisis measures on euro-area CDS spreads. *Financ. Markets Portf. Manage.* 29, 149–168. doi: 10.1007/s11408-015-0249-1
- Giudici, P., and Parisi, L. (2018). CoRisk: credit risk contagion with correlation network models. *Risks* 6:95. doi: 10.3390/risks6030095
- Glover, B., and Richards-Shubik, S. (2014). “Contagion in the European sovereign debt crisis,” in *NBER Working Paper No. 20567* (Cambridge, MA).
- Gower, J.C. (1971). A general coefficient of similarity and some of its properties. *Biometrics* 27, 857–871. doi: 10.2307/2528823
- Gower, J.C., and Ross, G. J. S. (1969). Minimum spanning trees and single linkage cluster analysis. *J. R. Stat. Soc. Ser. C Appl. Stat.* 18, 54–64 doi: 10.2307/2346439
- Gross, M., and Kok, C. (2013). “Measuring contagion potential among sovereigns and banks using a mixed-cross-section GVAR,” in *ECB Working Paper Series* (Frankfurt am Main).
- Kelly, S., Chaplin, A., Coburn, A. W., Copic, J., Evan, T., Neduv, E., et al. (2015). *Stress Test Scenario: Eurozone Meltdown*. Cambridge Risk Framework Series. Cambridge: Centre for Risk Studies, University of Cambridge.
- Kenett, D. Y., Tumminello, M., Madi, A., Gur-Gershgoren, G., Mantegna, R. N., and Ben-Jacob, E. (2010). Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market. *PLoS ONE* 5:e15032. doi: 10.1371/journal.pone.0015032
- Kierzenkowski, R., Pain, N., Rusticelli, E., and Zwart, S. (2016). “The Economic consequences of Brexit: a taxing decision,” in *OECD Economic Policy Paper No. 16* (Paris).
- Lange, R.-J., Lucas, A., and Siegmann, A. (2017). “5 - Score-driven systemic risk signaling for European sovereign bond yields and CDS spreads,” in *Systemic Risk Tomography*, eds M. Billio, L. Pelizzon, and R. Savona (Elsevier), 129–150. doi: 10.1016/B978-1-78548-085-0.50005-4
- Li, Y., and Waterworth, J. (2016). *Eurozone Network Connectedness During Calm and Crisis: Evidence From the MTS Platform for Interdealer Trading of European Sovereign Debt*. Available online at: <https://ssrn.com/abstract=2778291>
- Macintosh, J. (2017, February 10). Treat French debt like Italy, but don't worry about Le Pen. *Wall Street Journal*. Available online at: <https://www.wsj.com/articles/treat-france-like-italy-but-dont-worry-about-le-pen-1486667644>
- Macintosh, J. (2018, October 4). You should worry more about Italy's bond market. *Wall Street Journal*. Available online at: <https://www.wsj.com/articles/you-should-worry-more-about-italys-bond-market-1538672646>
- Mantegna, R. (1999). Hierarchical structure in financial markets. *Eur. Phys. J. B* 11, 193–197 doi: 10.1007/s100510050929
- Marriage, M., and Jennifer, T. (2017, May 14). Italian populism unnerves investors in the eurozone. *Financial Times*. Available online at: <https://www.ft.com/content/ff14351c-3572-11e7-99bd-13beb0903fa3>
- Moessner, R. (2018). Effects of asset purchases and financial stability measures on term premia in the euro area. *Appl. Econ.* 50, 4617–4631 doi: 10.1080/00036846.2018.1458199
- O'Brien, F. (2018, September 25). The long and winding road to Italy's budget. *Bloomberg News*. Available online at: <https://www.bloomberg.com/news/articles/2018-09-25/the-long-and-winding-road-to-italy-s-budget>
- Pascual, A. G., Gudin, P., and Aksu, C. (2018). *Contagion: This Time Is Different*. Barclays Economic Research, 1–15.
- Politis, D., and Romano, J. (1992). “A circular block-resampling procedure for stationary data,” in *Exploring the Limits of Bootstrap*, eds R. LePage and L. Billard (Wiley), 263–270.
- Rigobón, R. (2019). Contagion, spillover and interdependence. *Economía* 19, 69–99. Available online at: <https://www.muse.jhu.edu/article/722873>
- Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20, 53–65. doi: 10.1016/0377-0427(87)90125-7
- Sandhu, T. (2018, March 19). European yields on steroids may tempt Japanese, U.S. Investors. *Bloomberg News*. Available online at: <https://www.bloomberg.com/news/articles/2018-03-19/european-yields-on-steroids-may-tempt-japanese-u-s-investors>
- Saroyan, S., and Popoyan, L. (2017). “Bank-sovereign ties against interbank market integration: the case of the Italian segment,” in *LEM Working Paper Series, No. 2017/02, Scuola Superiore Sant'Anna, Laboratory of Economics and Management (LEM)* (Pisa).
- Schwendner, P., Schuele, M., Ott, T., and Hillebrand, M. (2015). European government bond dynamics and stability policies: taming contagion risks. *J. Netw. Theory Finance* 1, 1–24 doi: 10.21314/JNTF.2015.012
- Shoesmith, G. L. (2014). A time-series postmortem on eurozone financial integration and the debt crisis: modeling and policy implications. *Int. J. Business* 19, 113–131.
- Stafford, P., and Allen, K. (2018, June 12). Italian debt strife reveals pitfalls of Europe's market structure. *Financial Times*. Available online at: <https://www.ft.com/content/a111f158-6d65-11e8-92d3-6c13e5c92914>
- Tola, A., and Waelti, S. (2015). Deciphering financial contagion in the euro area during the crisis. *Q. Rev. Econ. Finance* 55, 108–123. doi: 10.1016/j.qref.2014.09.009
- Van Der Heijden, M., Beetsma, R., and Romp, W. (2018). “Whatever it takes” and the role of Eurozone news. *Appl. Econ. Lett.* 25, 1166–1169. doi: 10.1080/13504851.2017.1403555
- Ward, J. H. Jr. (1963). Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58, 236–244. doi: 10.1080/01621459.1963.10500845
- Whittall, C. (2017, June 12). Fading populism boosts European bonds. *Wall Street Journal*. Available online at: <https://www.wsj.com/articles/fading-populism-boosts-european-bonds-1497270185>

**Disclaimer:** The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Stability Mechanism.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Schwendner, Schüle and Hillebrand. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Explainable AI in Fintech Risk Management

Niklas Bussmann<sup>1,2</sup>, Paolo Giudici<sup>3\*</sup>, Dimitri Marinelli<sup>1</sup> and Jochen Papenbrock<sup>1</sup>

<sup>1</sup> FIRAMIS, Frankfurt, Germany, <sup>2</sup> Department of Economics and Management, University of Pavia, Pavia, Italy, <sup>3</sup> Fintech Laboratory, Department of Economics and Management, University of Pavia, Pavia, Italy

The paper proposes an explainable AI model that can be used in fintech risk management and, in particular, in measuring the risks that arise when credit is borrowed employing peer to peer lending platforms. The model employs Shapley values, so that AI predictions are interpreted according to the underlying explanatory variables. The empirical analysis of 15,000 small and medium companies asking for peer to peer lending credit reveals that both risky and not risky borrowers can be grouped according to a set of similar financial characteristics, which can be employed to explain and understand their credit score and, therefore, to predict their future behavior.

**Keywords:** credit risk management, explainable AI, financial technologies, peer to peer lending, logistic regression, predictive models

## OPEN ACCESS

### Edited by:

Shatha Qamieh Hashem,  
An-Najah National University, Palestine

### Reviewed by:

Paolo Barucca,  
University College London,  
United Kingdom  
Bertrand Kian Hassani,  
University College London,  
United Kingdom

### \*Correspondence:

Paolo Giudici  
giudici@unipv.it

### Specialty section:

This article was submitted to  
Artificial Intelligence in Finance,  
a section of the journal  
*Frontiers in Artificial Intelligence*

**Received:** 18 December 2019

**Accepted:** 30 March 2020

**Published:** 24 April 2020

### Citation:

Bussmann N, Giudici P, Marinelli D  
and Papenbrock J (2020) Explainable  
AI in Fintech Risk Management.  
*Front. Artif. Intell.* 3:26.  
doi: 10.3389/frai.2020.00026

## 1. INTRODUCTION

Black box Artificial Intelligence (AI) is not suitable in regulated financial services. To overcome this problem, Explainable AI models, which provide details or reasons to make the functioning of AI clear or easy to understand, are necessary.

To develop such models, we first need to understand what “Explainable” means. During this year, some important benchmark definitions have been provided, at the institutional level. We report some of them, in the context of the European Union.

For example, the Bank of England (Joseph, 2019) states that “Explainability means that an interested stakeholder can comprehend the main drivers of a model-driven decision.” The Financial Stability Board (FSB, 2017) suggests that “lack of interpretability and auditability of AI and ML methods could become a macro-level risk.” Finally, the UK Financial Conduct Authority (Croxson et al., 2019) establishes that “In some cases, the law itself may dictate a degree of explainability.”

The European GDPR (EU, 2016) regulation states that “the existence of automated decision-making, should carry meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.” Under the GDPR regulation, the data subject is therefore, under certain circumstances, entitled to receive meaningful information about the logic of automated decision-making.

Finally, the European Commission High-Level Expert Group on AI presented the Ethics Guidelines for Trustworthy Artificial Intelligence in April 2019. Such guidelines put forward a set of seven key requirements that AI systems should meet in order to be deemed trustworthy. Among them three related to XAI, and are the following.

- Human agency and oversight: decisions must be informed, and there must be a human-in-the-loop oversight.
- Transparency: AI systems and their decisions should be explained in a manner adapted to the concerned stakeholder. Humans need to be aware that they are interacting with an AI system.
- Accountability: AI systems should develop mechanisms for responsibility and accountability, auditability, assessment of algorithms, data and design processes.

Following the need to explain AI models, stated by legislators and regulators of different countries, many established and startup companies have started to embrace Explainable AI (XAI) models.

From a mathematical viewpoint, it is well-known that, while “simpler” statistical learning models, such as linear and logistic regression models, provide a high interpretability but, possibly, a limited predictive accuracy, “more complex” machine learning models, such as neural networks and tree models provide a high predictive accuracy at the expense of a limited interpretability.

To solve this trade-off, we propose to boost machine learning models, that are highly accurate, with a novel methodology, that can explain their predictive output. Our proposed methodology acts in the post-processing phase of the analysis, rather than in the preprocessing part. It is agnostic (technologically neutral) as it is applied to the predictive output, regardless of which model generated it: a linear regression, a classification tree or a neural network model.

More precisely, our proposed methodology is based on Shapley values (see Lundberg and Lee, 2017 and references therein). We consider a relevant application of AI in financial technology: peer to peer lending.

We employ Shapley values to predict the credit risk of a large sample of small and medium enterprises which apply for credit to a peer to peer lending platform. The obtained empirical evidence shows that, while improving the predictive accuracy with respect to a standard logistic regression model, we maintain and, possibly, improve, the interpretability (explainability) of the results.

In other words, our results confirm the validity of this approach in discriminating between defaulted and sound institutions, and it shows the power of explainable AI in both prediction accuracy and in the interpretation of the results.

The rest of the paper is organized as follows: section 2 introduces the proposed methodology. Section 3 shows the results of the analysis in the credit risk context. Section 4 concludes.

## 2. METHODOLOGY

### 2.1. Credit Risk in Peer to Peer Lending

Credit risk models are useful tools for modeling and predicting individual firm default. Such models are usually grounded on regression techniques or machine learning approaches often employed for financial analysis and decision-making tasks.

Consider  $N$  firms having observation regarding  $T$  different variables (usually balance-sheet measures or financial ratios). For each institution  $n$  define a variable  $\gamma_n$  to indicate whether such institution has defaulted on its loans or not, i.e.,  $\gamma_n = 1$  if company defaults,  $\gamma_n = 0$  otherwise. Credit risk models develop relationships between the explanatory variables embedded in  $T$  and the dependent variable  $\gamma$ .

The logistic regression model is one of the most widely used method for credit scoring. The model aims at classifying the dependent variable into two groups, characterized by different

status (defaulted vs. active) by the following model:

$$\ln\left(\frac{p_n}{1-p_n}\right) = \alpha + \sum_{t=1}^T \beta_t x_{nt} \quad (1)$$

where  $p_n$  is the probability of default for institution  $n$ ,  $\mathbf{x}_i = (x_{i,1}, \dots, x_{i,T})$  is the  $T$ -dimensional vector of borrower specific explanatory variables, the parameter  $\alpha$  is the model intercept while  $\beta_t$  is the  $t$ -th regression coefficient. It follows that the probability of default can be found as:

$$p_n = (1 + \exp(\alpha + \sum_{t=1}^T \beta_t x_{nt}))^{-1} \quad (2)$$

### 2.2. Machine Learning of Credit Risk

Credit risk can be measured with very different Machine Learning (ML) models, able to extract non-linear relations among the financial information in the balance sheets. In a standard data science life cycle, models are chosen to optimize the predictive accuracy. In highly regulated sectors, like finance or medicine, models should be chosen balancing accuracy with explainability (Murdoch et al., 2019). We improve the choice selecting models based on their predictive accuracy, and employing *a posteriori* an explanations algorithm. This does not limit the choice of the best performing models.

To exemplify our approach we consider, without loss of generality, the XGBoost model, one of the most popular and fast algorithm (Chen and Guestrin, 2016), that implements gradient tree boosting learning models.

### 2.3. Learning Model Comparison

For evaluating the performance of each learning model, we employ, as a reference measure, the indicator  $\gamma \in \{0, 1\}$ , a binary variable which takes value one whenever the institutions has defaulted and value zero otherwise. For detecting default events represented in  $\gamma$ , we need a continuous measurement  $p \in [0, 1]$  to be turned into a binary prediction  $B$  assuming value one if  $p$  exceeds a specified threshold  $\tau \in [0, 1]$  and value zero otherwise. The correspondence between the prediction  $B$  and the ideal leading indicator  $\gamma$  can then be summarized in a so-called confusion matrix.

From the confusion matrix we can easily illustrate the performance capabilities of a binary classifier system. To this aim, we compute the receiver operating characteristic (ROC) curve and the corresponding area under the curve (AUC). The ROC curve plots the false positive rate (FPR) against the true positive rate (TPR), as follows:

$$FPR = \frac{FP}{FP + TN} \quad (3)$$

$$TPR = \frac{TP}{TP + FN} \quad (4)$$

The overall accuracy of each model can be computed as:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

and it characterizes the proportion of true results (both true positives and true negatives) among the total number of cases.

## 2.4. Explaining Model Predictions

We now explain how to exploit the information contained in the explanatory variables to localize and cluster the position of each individual (company) in the sample. This information, coupled with the predicted default probabilities, allows a very insightful explanation of the determinant of each individual's creditworthiness. In our specific context, information on the explanatory variables is derived from the financial statements of borrowing companies, collected in a vector  $\mathbf{x}_n$  representing the financial composition of the balance sheet of institution  $n$ .

We propose calculate the Shapley value associated with each company. In this way we provide an agnostic tool that can interpret in a technologically neutral way the output from a highly accurate machine learning model. As suggested in Joseph (2019), the Shapley values of a model can be used as a tool to transfer predictive inferences into a linear space, opening a wide possibility of using the toolbox of econometrics, hypothesis testing, and network analysis.

We develop our Shapley approach using the SHAP (Lundberg and Lee, 2017) computational framework, which allows to express each single prediction as a sum of the contributions of the different explanatory variables.

More formally, the Shapley explanation model for each prediction  $\phi(\hat{f}(x_i))$  is obtained by an additive feature attribution method, which decomposes them as:

$$\phi(\hat{f}(\mathbf{x}_i)) = \phi_0 + \sum_{k=1}^M \phi_k(x_i). \quad (6)$$

where  $M$  is the number of available explanatory variables,  $\phi \in \mathbb{R}^M$ ,  $\phi_k \in \mathbb{R}$ . The local functions  $\phi_k(x_i)$  are called Shapley values.

Indeed, Lundberg and Lee (2017) prove that the only additive feature attribution method that satisfies the properties of *local accuracy*, *missingness*, and *consistency* is obtained attributing to each feature  $x_k, k = 1, \dots, M$ , a SHapley Additive exPlanation (SHAP) defined by

$$\phi_k(x_i) = \sum_{x' \subseteq \mathcal{C}(x) \setminus x_k} \frac{|x'|!(M - |x'| - 1)!}{M!} [\hat{f}(x' \cup x_k) - \hat{f}(x')], \quad (7)$$

where  $\mathcal{C}(x) \setminus x_k$  is the set of all the possible models excluding variable  $x_k$  (with  $m = 1, \dots, M$ ),  $|x'|$  denotes the number of variables included in model  $x'$ ,  $M$  is the number of the available variables,  $\hat{f}(x' \cup x_k)$  and  $\hat{f}(x')$  are the predictions associated with all the possible model configurations including variable  $x_k$  and excluding variable  $x_k$ , respectively.

The quantity  $\hat{f}(x' \cup x_k) - \hat{f}(x')$  defines the contribution of variable  $x_k$  to each individual prediction.

## 3. APPLICATION

### 3.1. Data

We test our proposed model to data supplied by European External Credit Assessment Institution (ECAI) that specializes in credit scoring for P2P platforms focused on SME commercial lending. The data is described by Giudici et al. (2019a) to which we refer for further details. In summary, the analysis relies on a dataset composed of official financial information (balance-sheet variables) on 15,045 SMEs, mostly based in Southern Europe, for the year 2015. The information about the status (0 = active, 1 = defaulted) of each company 1 year later (2016) is also provided. Using this data, Giudici (2018), Ahelegbey et al. (2019), and Giudici et al. (2019a,b) have constructed logistic regression scoring models that aim at estimating the probability of default of each company, using the available financial data from the balance sheets and, in addition, network centrality measures that are obtained from similarity networks.

Here we aim to improve the predictive performance of the model and, for this purpose, we run an XGBoost tree algorithm (see e.g., Chen and Guestrin, 2016). To explain the results from the model, typically highly predictive, we employ Shapley values.

The proportion of defaulted companies within this dataset is 10.9%.

### 3.2. Results

We first split the data in a training set (80%) and a test set (20%).

We then estimate the XGBoost model on the training set, apply the obtained model to the test set and compare it with the optimal logistic regression model. The ROC curves of the two models are contained in **Figure 1** below.

From **Figure 1** note that the XGBoost clearly improves predictive accuracy. Indeed the calculation of the AUROC of the two curves indicate an increase from 0.81 (best logistic regression model) to 0.93 (best XGBoost model).

We then calculate the Shapley values for the companies in the test set.

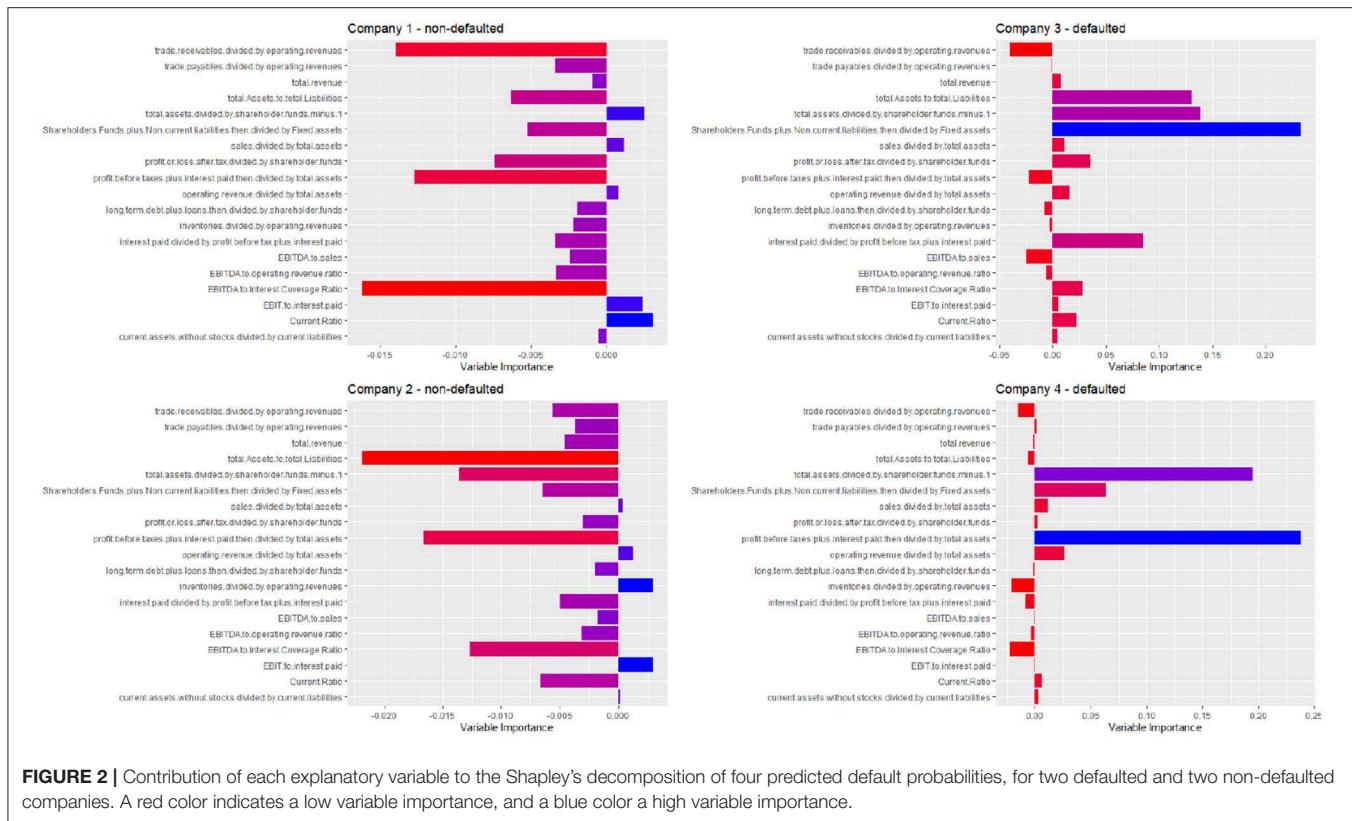
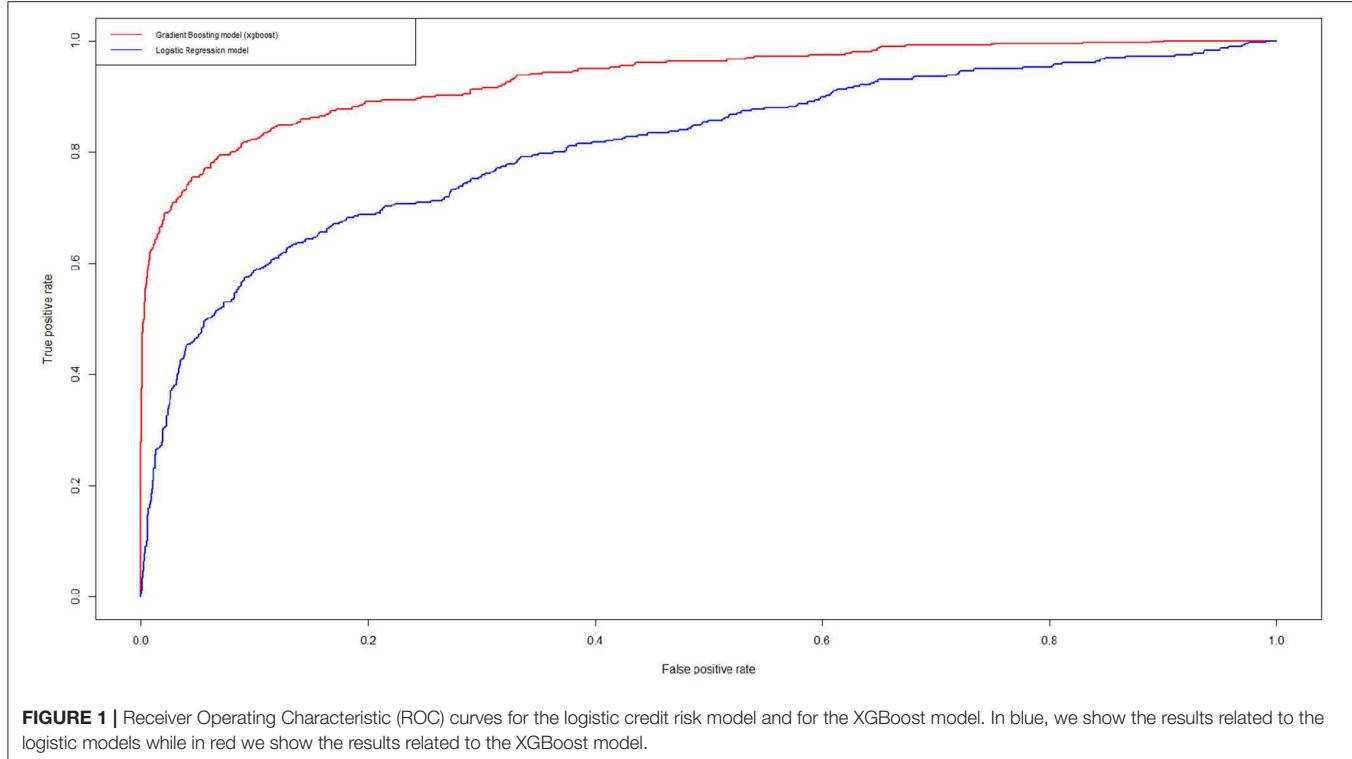
To exemplify our results, **Figure 2** we provide the interpretation of the estimated credit scoring of four companies: two that default and two that do not default.

**Figure 2** clearly shows the advantage of our explainable model. It can indicate which variables contribute more to the prediction. Not only in general, as is typically done by feature selection models, but differently and specifically for each company in the test set. Note how the explanations are rather different ("personalized") for each of the four considered companies.

## 4. CONCLUSIONS

The need to leverage the high predictive accuracy brought by sophisticated machine learning models, making them interpretable, has motivated us to introduce an agnostic, post-processing methodology, based on Shapley values. This allows to explain any single prediction in terms of the potential contribution of each explanatory variable.

Future research should include a better understanding of the predictions through clustering of the Shapley values. This can



be achieved, for example, using correlation network models. A second direction would be to extend the approach developing model selection procedures based on Shapley values, which would require appropriate statistical testing. A last extension would be to develop a Shapley like measure that applies also to ordinal response variables.

Our research has important policy implications for policy makers and regulators who are in their attempt to protect the consumers of artificial intelligence services. While artificial intelligence effectively improve the convenience and accessibility of financial services, they also trigger new risks, and among them is the lack of model interpretability. Our empirical findings suggest that explainable AI models can effectively advance our understanding and interpretation of credit risks in peer to peer lending.

Future research may involve further experimentation and the application to other case studies.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## REFERENCES

- Ahelegbey, D., Giudici, P., and Hadji-Misheva, B. (2019). Latent factor models for credit scoring in P2P systems. *Phys. A Stat. Mech. Appl.* 522, 112–121. doi: 10.1016/j.physa.2019.01.130
- Chen, T., and Guestrin, C. (2016). “Xgboost: a scalable tree boosting system,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, NY: ACM), 785–794.
- Croxson, K., Bracke, P., and Jung, C. (2019). *Explaining Why the Computer Says ‘no’*. London, UK: FCA-Insight.
- EU (2016). Regulation (EU) 2016/679—general data protection regulation (GDPR). *Off. J. Eur. Union*.
- FSB (2017). *Artificial Intelligence and Machine Learning in Financial Services—Market Developments and Financial Stability Implications*. Technical report, Financial Stability Board.
- Giudici, P. (2018). “Financial data science,” in *Statistics and Probability Letters*, Vol. 136 (Elsevier), 160–164.
- Giudici, P., Hadji Misheva, B., and Spelta, A. (2019a). Correlation network models to improve P2P credit risk management. *Artif. Intell. Finance*.
- Giudici, P., Hadji-Misheva, B., and Spelta, A. (2019b). Network based credit risk models. *Qual. Eng.* 32, 199–211. doi: 10.1080/08982112.2019.1655159
- Joseph, A. (2019). *Shapley Regressions: A Framework for Statistical Inference on Machine Learning Models*. Resrepot 784, Bank of England.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

## FUNDING

This research has received funding from the European Union’s Horizon 2020 research and innovation program FIN-TECH: A Financial supervision and Technology compliance training programme under the grant agreement No 825215 (Topic: ICT-35-2018, Type of action: CSA), and from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 750961.

## ACKNOWLEDGMENTS

In addition, the authors thank ModeFinance, a European ECAI, for the data; the partners of the FIN-TECH European project, for useful comments and discussions.

- Lundberg, S., and Lee, S.-I. (2017). “A unified approach to interpreting model predictions,” in *Advances in Neural Information Processing Systems 30*, eds I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, et al. (New York, NY: Curran Associates, Inc.), 4765–4774.
- Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., and Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proc. Natl. Acad. Sci. U.S.A.* 116, 22071–22080. doi: 10.1073/pnas.1900654116

**Conflict of Interest:** NB, DM, and JP are employed by the company FIRAMIS.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Bussmann, Giudici, Marinelli and Papenbrock. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## Screenshot of use case I code

```
rm(list = ls())

library(PerformanceAnalytics)
library(xts)
library(quantmod)
library(timeSeries)
library(xtable)
library(igraph)
library(tcltk2)
library(MTS)
library(matrixcalc)
library(Matrix)
library(fPortfolio)
library(IntroCompFinR) #install.packages("IntroCompFinR", repos="http://R-Forge.R-project.org")
require(quadprog)
library(pracma)
library(glasso)

#setwd("/home/rstudio")

prices<-read.table("def.csv", header=TRUE, sep=",", dec=".") 
ZOO <- zoo(prices[,-1], order.by=as.Date(as.character(prices$date)), format='%m/%d/%Y')) 
head(prices)

#to plot normalized prices

norm_prices<-read.table("def1.csv", header=TRUE, sep=",", dec=".") 
ZOO_norm<- zoo(norm_prices[,-1], order.by=as.Date(as.character(norm_prices$date)), format='%m/%d/%Y')) 
#dev.new()
plot(ZOO_norm[,c(1,2,4,5,6)], ylab="normalized prices",xlab="time", plot.type= "single", col=c(1:5))
legend("topright", c("BTC","ETH","USDT","BCH","LTC"), col=c(1:5), cex=0.80, lwd=2)
#dev.new()
plot(ZOO_norm[,c(3,7,8,9,10)], ylab="normalized prices",xlab="time", plot.type= "single", col=c(6:10))
legend("topright", c("XRP","BNB","EOS","XLM","TRX" ), col=c(6:10), cex=0.80, lwd=2)
|
return<- Return.calculate(ZOO, method="log")
```

## Screenshot of use case II code

```
## AI Market Risk Use Case "Convergence and Divergence in European Bond Correlations"
#
# computes filtered average correlation influences for monthly returns of 10-year European bonds in three-year windows
#
# Peter Schwendner, Martin Schuele and Martin Hillebrand
# Zurich University of Applied Sciences, 1.9.2019
# for EU Horizon 2020 FIN-TECH project
#
# implementation and application reference:
# Schwendner, Peter; Schuele, Martin; Ott, Thomas; Hillebrand, Martin:
# European government bond dynamics and stability policies : taming contagion risks.
# Journal of network theory in finance. 1(4), S. 1-25, 2015.
# and ESM Working Paper #8, 2015.

# original correlation influence network reference:
# Kenett, D. Y., Tumminello, M., Madi, A., Gur-Gershgoren, G., Mantegna,
# R. N., and Ben-Jacob, E. (2010). Dominating clasp of the financial
# sector revealed by partial correlation analysis of the stock market.
# Plos one, 5(12):e15032, 2010.

setwd("/home/rstudio")

#rm(list=ls())
library(HH)
library(seriation)
library(purrr)
library(magclass)
library(ppcor)

options(warn=-1)

## read ECB SDW data
# Dataset name Interest rate statistics (2004 EU Member States & ACCBs)
# Frequency Monthly
# Interest rate type Long-term interest rate for convergence purposes
# Transaction type Debt security issued
```

### Screenshot of use case III code

```
#### D I S C L A I M E R ####

# This code supplements the paper https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3506274.
# It is only allowed to use it for evaluation purposes in the EU Horizon2020 project
# FIN-TECH (www.fintech-h2020.eu) under grant agreement No 825215.
#
# It is strictly prohibited to use the code outside the project or even distribute it.

# Clean the environment
graphics.off()
rm(list = ls(all = TRUE))

library("ggpubr")
library("caTools")
library("grDevices")
library("e1071")
library("tidyverse")
library("readxl")
library("caret")
library("xgboost")
library("ROCR")
library("mlr")
library("mlrMBO")
#library("devtools")
library("rgenoud")
library("emstreeR")
library("igraph")
library("graphlayouts")
library("grid")
library("autoxgboost") #devtools::install_github("ja-thomas/autoxgboost")
library("SHAPforxgboost")
library("e1071")

# setwd("/home/rstudio")
```