

The background of the cover is a complex, abstract design. It features a large, solid purple circle in the center. Surrounding this circle are various geometric and organic shapes in shades of blue, green, and orange. These include thin, radiating lines, wavy lines, and small squares. The overall effect is a sense of dynamic energy and technological complexity.

AI AND FINANCIAL TECHNOLOGY

EDITED BY: Paolo Giudici, Jochen Papenbrock, Peter Schwendner,
Ronald Hochreiter and Joerg Osterrieder
PUBLISHED IN: Frontiers in Artificial Intelligence



frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-88963-341-8

DOI 10.3389/978-2-88963-341-8

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

AI AND FINANCIAL TECHNOLOGY

Topic Editors:

Paolo Giudici, University of Pavia, Italy

Jochen Papenbrock, Independent researcher, Germany

Peter Schwendner, Zurich University of Applied Sciences, Switzerland

Ronald Hochreiter, Vienna University of Economics and Business, Austria

Joerg Osterrieder, Zurich University of Applied Sciences, Switzerland

Citation: Giudici, P., Papenbrock, J., Schwendner, P., Hochreiter, R., Osterrieder, J., eds. (2020). AI and Financial Technology. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88963-341-8

Table of Contents

04	<i>Editorial: AI and Financial Technology</i> Paolo Giudici, Ronald Hochreiter, Jörg Osterrieder, Jochen Papenbrock and Peter Schwendner
06	<i>Spatial Regression Models to Improve P2P Credit Risk Management</i> Arianna Agosto, Paolo Giudici and Tom Leach
12	<i>Applied Machine Learning for Stochastic Local Volatility Calibration</i> Jürgen Hakala
24	<i>Network Based Scoring Models to Improve Credit Risk Management in Peer to Peer Lending Platforms</i> Paolo Giudici, Branka Hadji-Misheva and Alessandro Spelta
32	<i>An Artificial Intelligence Approach to Regulating Systemic Risk</i> Sharyn O'Halloran and Nikolai Nowaczyk
46	<i>Factorial Network Models to Improve P2P Credit Risk Management</i> Daniel Felix Ahelegbey, Paolo Giudici and Branka Hadji-Misheva
55	<i>Validation of PARX Models for Default Count Prediction</i> Arianna Agosto and Emanuela Raffinetti
62	<i>Neural Network Models for Bitcoin Option Pricing</i> Paolo Pagnottoni
71	<i>Artificial Intelligence (AI) in the Financial Sector—Potential and Public Strategies</i> Stephan Bredt
76	<i>Sentiment Analysis of European Bonds 2016–2018</i> Peter Schwendner, Martin Schüle and Martin Hillebrand



Editorial: AI and Financial Technology

Paolo Giudici^{1*}, Ronald Hochreiter², Jörg Osterrieder³, Jochen Papenbrock⁴ and Peter Schwendner⁵

¹ Department of Economics and Management, University of Pavia, Pavia, Italy, ² Department of Business and Management, Webster Vienna Private University, Vienna, Austria, ³ School of Engineering, Zurich University of Applied Sciences, Winterthur, Switzerland, ⁴ Firamis GmbH, Frankfurt, Germany, ⁵ Center for Asset Management, School of Management and Law, Zurich University of Applied Sciences, Winterthur, Switzerland

Keywords: FinTech, SupTech, RegTech, AI, machine learning, P2P lending, Blockchain

Editorial on the Research Topic

AI and Financial Technology

The Financial Stability Board defines FINancial TECHnology as “technologically enabled financial innovation that could result in new business models, applications, processes, or products with an associated material effect on financial markets and institutions and the provision of financial services.” While innovation in Finance is not a new concept, the focus on technological innovations and its pace have increased significantly. Fintech solutions that make use of Big Data analytics, Artificial Intelligence, and Blockchain technologies are currently introduced at an unprecedented rate. These new technologies are changing the nature of the financial industry, creating opportunities for Fintechs startups to offer more inclusive access to financial services. The advantages notwithstanding, Fintech solutions leave the door open for many challenges such as underestimation of creditworthiness, market volatility, cyber attacks, fraud and money laundering which represent central points of interest for regulators and supervisory bodies.

In this context, a key issue becomes identifying the desired level of trade-off between innovation incentives on one hand, and mitigation of risks on the other. The European regulatory framework should enable Fintech companies operating in their jurisdiction to benefit from innovations in Technology and Finance while at the same time ensuring both a high level of protection for consumers and investors and resilience of the financial system. This point has been framed by the current European Commissioner for the Euro and Social Dialogue and Vice-President of the European Commission, Valdis Dombrovskis: “Across the board, we are working to strike the right balance between risks and opportunities; so that Europe can benefit fully from new technologies in the financial services sector.”

There is a strong need to improve the competitiveness of the European Fintech sector, introducing a framework for a common regulatory approach across all countries that can supervise Fintech companies without stifling their economic potential. Such a framework should support both Fintechs as well as supervisors: on one hand, Fintech firms that want to grow and scale-up across Europe require a neutral technology and proportional regulatory compliance as well as advice on how to identify opportunities for innovation procurement, e.g., in advanced regulatory technology (RegTech) solutions; on the other hand, the supervisory bodies’ ability to monitor innovative financial products proposed by Fintechs is limited and advanced supervisory technology (SupTech) solutions are required.

The Horizon 2020 project FIN-TECH (Financial Supervision and Technological Compliance)—funded by the European Commission for the period 2019–2020—conducts research on Fintech risk management models to be shared with European regulators, Fintechs as well as banks. These models are evaluated on a global level which helps to close the gap between technical and regulatory expertise, in particular providing risk management procedures

OPEN ACCESS

Edited and reviewed by:

Thomas Hartung,
Johns Hopkins University,
United States

*Correspondence:

Paolo Giudici
paolo.giudici@unipv.it

Specialty section:

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

Received: 24 October 2019

Accepted: 30 October 2019

Published: 15 November 2019

Citation:

Giudici P, Hochreiter R, Osterrieder J,
Papenbrock J and Schwendner P
(2019) Editorial: AI and Financial
Technology. *Front. Artif. Intell.* 2:25.
doi: 10.3389/frai.2019.00025

common to both sides and uniform across countries. It will eventually lead to the development of a regulatory framework that encourages innovations in Big Data analytics, Artificial Intelligence, and Blockchain technologies which, at the same time, satisfies supervisory concerns to apply regulations in an effective and efficient way, that well protects consumers and investors. In particular, the FIN-TECH project aims to create a European training program aimed at shared risk management solutions that automatize compliance of Fintech companies (RegTech) and, at the same time, increases the efficiency of supervisory activities (SupTech). In other words, the project aims at connecting FINancial supervision with TECHnological compliance.

This special issue contains the first contributions from this European project. Some of them are research papers that evolved into use cases of the project and are shared as well as used by regulators, banks, and Fintechs. Other papers are based on extensive talks given by external speakers that participated at specific events organized by the project. This collection of papers is discussing public policy viewpoints as well as AI applications to measure market risks and credit risks especially in the areas of Robo Advisory and Peer to Peer (P2P) lending.

The paper by Bredt as well as the paper by O'Halloran and Nowaczyk present and discuss public policy strategies aimed at addressing financial innovations brought by disruptive technologies, capturing their opportunities while mitigating the related risks.

Furthermore, the paper by Schwendner et al., the paper by Hakala and the paper by Pagnottoni show how Machine Learning methods and Artificial Intelligence solutions can be employed to

develop new asset management practices addressing risks. While Schwendner et al. focus on European Bonds, Hakala considers modeling volatilities and Pagnottoni works on Blockchain based bitcoin transactions.

The paper by Giudici et al. the paper by Ahelegbey et al., and the paper by Agosto et al. all consider how the measurement of network effects arising from Peer to Peer (P2P) lending platforms can improve the measurement of credit risk of borrowers. They apply different models but use the same database with the common goal to provide applicable use cases for the FIN-TECH European project to monitor and control credit risk arising from the application of Big Data analytics.

Finally, the paper by Agosto and Raffinetti focuses on building appropriate model comparison tools for credit risk modeling.

AUTHOR CONTRIBUTIONS

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

Conflict of Interest: JP was employed by the company Firamis GmbH.

The remaining 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 Giudici, Hochreiter, Osterrieder, Papenbrock and Schwendner. 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.



Spatial Regression Models to Improve P2P Credit Risk Management

Arianna Agosto*, Paolo Giudici and Tom Leach

Department of Economics and Management, University of Pavia, Pavia, Italy

Calabrese et al. (2017) have shown how binary spatial regression models can be exploited to measure contagion effects in credit risk arising from bank failures. To illustrate their methodology, the authors have employed the Bank for International Settlements' data on flows between country banking systems. Here we apply a binary spatial regression model to measure contagion effects arising from corporate failures. To derive interconnectedness measures, we use the World Input-Output Trade (WIOT) statistics between economic sectors. Our application is based on a sample of 1,185 Italian companies. We provide evidence of high levels of contagion risk, which increases the individual credit risk of each company.

Keywords: credit risk, systemic risk, contagion, spatial autoregressive models, binary data

OPEN ACCESS

Edited by:

Jiancheng Jiang,
University of North Carolina at
Charlotte, United States

Reviewed by:

Jianan Peng,
Acadia University, Canada
Laura Vana,
Vienna University of Economics and
Business, Austria

*Correspondence:

Arianna Agosto
arianna.agosto@unipv.it

Specialty section:

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

Received: 04 March 2019

Accepted: 30 April 2019

Published: 16 May 2019

Citation:

Agosto A, Giudici P and Leach T
(2019) Spatial Regression Models to
Improve P2P Credit Risk
Management. *Front. Artif. Intell.* 2:6.
doi: 10.3389/frai.2019.00006

1. INTRODUCTION

In recent years, the emergence of financial technologies (fintechs) is redefining the roles of financial intermediaries and introducing many opportunities for consumers and investors. In particular, peer-to-peer (P2P) online lending platforms allow private individuals to directly make small and unsecured loans to private borrowers.

P2P lending business models vary in scope and structure: a comprehensive review is provided by Claessens et al. (2018). Here we specifically refer to the platforms that lend to small and medium enterprises (SME).

While both classic banks and P2P platforms rely on credit scoring models for the purpose of estimating the credit risk of their loans, the incentive for model accuracy may differ significantly.

In a bank, credit risk assessment is conducted by the financial institution itself which, being the actual entity that assumes the risk, is interested to have the most accurate possible model. In a P2P lending platform, credit risk is determined by the platform but the risk is fully borne by the lender. In other words, P2P platforms allow for direct matching between borrowers and lenders.

A factor that penalizes the accuracy of P2P credit scoring models is that they often do not have access to borrowers' data usually employed by banks, such as account transaction data, financial data and credit bureau data. For these reasons, the accuracy of credit risk estimates provided by P2P lenders may be poor. However, P2P platforms involve their users and, in particular, the borrowers, in a continuous networking activity. Data from such activity can be leveraged not only for commercial purposes, as it is customarily done, but also to improve credit risk accuracy.

We believe that networking information can offset the lack of financial and credit behavioral data and improve credit risk measurement accuracy of P2P lenders, but also of banks. There are indeed cases in which also traditional financial intermediaries face lack of information about the borrower. Consider, for example, credit granting to new customers, for whom internal behavioral data—known to be the most predictive in rating models—are not available.

When financial networks are backed by statistical models, inferential statements can be obtained. Important contributions in this framework are Billio et al. (2012); Diebold and Yilmaz (2014); Hautsch et al. (2015); Ahelegbey et al. (2016); Giudici and Spelta (2016), and Giudici and Parisi (2018), who propose measures of connectedness based on similarities, Granger-causality tests, variance decompositions and partial correlations between market price variables.

We improve these contributions, extending them to the P2P context and linking network models, that are often merely descriptive, with econometric models, thus providing a predictive framework.

More specifically, we suggest to use spatial econometrics to study the interconnectedness in the corporate sector. Spatial econometrics incorporates dependence among observations that are in any kind of proximity, not only geographical.

In particular, the model we apply is a logit Spatial Autoregressive model based on an exogenously defined network. The main advantage of this approach over the traditional network analysis is that it can be used as both an early warning model, to forecast the failure of a given company, and as a stress testing technique taking systemic effects into account.

The paper is organized as follows. Section 2 explains the econometric methodology. Section 3 presents the results obtained by applying the proposed methodology to data collected from a European P2P lending information provider. Section 4 concludes.

2. METHODOLOGY

2.1. Spatial Logit

The model we use in this paper has a binary spatial autoregressive structure, whereby the dependent variable is binary and a spatial autoregressive structure is assumed in the underlying latent variable. Taking the latent underlying quantity to be represented by a continuous variable y_i^* , we consider the observation mechanism as

$$y_i = \begin{cases} 1, & y_i^* > 0 \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

with $i = 1, 2, \dots, n$.

We implement the spatial structure with an autoregressive model specification, such that

$$Y^* = \rho WY^* + X\beta + \epsilon, \quad (2)$$

where Y^* is a continuous random vector, X represents an $n \times k$ matrix of explanatory variables with related coefficient vector β , ϵ is the error term and W is the spatial lag weight matrix with ρ the associated coefficient, which in our application to defaults will be interpreted as a contagion parameter.

The model implies heteroskedastic errors e as follows:

$$Y^* = (I - \rho W)^{-1}(X\beta + \epsilon) = (I - \rho W)^{-1}X\beta + e, \quad (3)$$

where

$$e = (I - \rho W)^{-1}\epsilon \quad (4)$$

and

$$\text{var}(e) = \text{var}[(I - \rho W)^{-1}\epsilon] = \sigma_\epsilon^2 [(I - \rho W)'(I - \rho W)]^{-1}. \quad (5)$$

The defined model has been used by Calabrese et al. (2017) to study default interdependence in the European banking sector. Relative to the estimation, Calabrese and Elkink (2014) have provided a review of the main methodologies for model (3) in the literature. Among the various approaches, we focus on the Generalized Method of Moments (GMM) proposed by Pinkse and Slade (1998). They derive the Generalized Method of Moments (GMM) moment equations from the likelihood function of a spatial error probit model, for which Klier and McMillen (2008) provide the extension to logit models. The GMM approach does not rely on a potentially inaccurate assumption of normally distributed errors and is therefore more robust than maximum likelihood methods.

In general, a GMM estimator is defined by:

$$\hat{\theta} \equiv \arg \min_{\theta} m_n(\theta) \Omega_n m_n(\theta)', \quad (6)$$

where $m_n(\theta)$ are the moment conditions and Ω_n is a weighting matrix to be determined.

In our case, we have:

$$\theta = [\rho, \beta]$$

To construct the moments, following Pinkse and Slade (1998) we use the generalized residuals

$$u_i = y_i - p_i, \quad (7)$$

where:

$$p_i = \Pr[y_i = 1] = \frac{\exp^{(I - \hat{\rho}W)^{-1}X\hat{\beta}}}{1 + \exp^{(I - \hat{\rho}W)^{-1}X\hat{\beta}}}$$

It follows from specification (3) that the elements of the spatially lagged dependent vector WY^* are correlated with those of the error vector, hence the need for instrumental variables. Following Kelijian and Prucha (1998), who suggest to choose the instruments as a subset of the linearly independent columns of:

$$H = \{X, WX, W^2X, W^3X, \dots\}$$

we define the instrument matrix¹

$$Z = \{X, WX\}$$

Thus, generating the moment conditions via the identity:

$$E[Z'u] = 0$$

$\hat{\theta}$ can be estimated by the following

$$\hat{\theta} = \arg \min_{\theta} u'Z\Omega Z'u \quad (8)$$

The estimation algorithm used in our application is explained in detail in section 2.3.

¹As explained in Kelijian and Prucha (2010), H proxies the expected value of WY^* using its projection on X .

2.2. The Network

The spatial regression model we propose is based on an exogenously defined network, where the nodes correspond to individual companies and the ties express the volume of trade between any pair of companies, i.e., the trade flow from company i to company j , for each i and each j . This information is generally not available, so we must approximate it using data on aggregate input-output trade between sectors.

The World Input Output Trade (WIOT) statistics provide information on the aggregate trade volumes of 52 economic sectors in each country with all sectors in all countries.

For a given country, define A as the sector of company i , B as the sector of company j , and let f_{AB} be the trade flow from sector A to sector B , while f_{BA} is the trade flow from sector B to sector A .

Replacing the individual flows with the aggregate ones, the entries of the approximate trade matrix F are then obtained as:

$$f_{ij} = f_{AB} = \sum_{l \in A} \sum_{m \in B} f_{lm}$$

To use these data for proxying the individual companies' flows, we need to calculate the proportion of each company in terms of size over its sector using a suitable measure, such as turnover or the value of trade receivables (for inflows) and payables (for outflows).

Consider, for example, the case of determining the trade flows from company i , belonging to sector A , to company j , belonging to sector B , knowing the individual trade payables and receivables.

We first calculate the ratio between company i trade payables \tilde{x}_i and the sum of sector A trade payables:

$$x_i = \frac{\tilde{x}_i}{\sum_{l \in A} \tilde{x}_l}$$

Then we calculate the ratio between company j trade receivables \tilde{y}_j and the sum of sector B trade receivables:

$$y_j = \frac{\tilde{y}_j}{\sum_{m \in B} \tilde{y}_m}$$

The product $x_i y_j$ is a proxy of the proportion of flows from company i to company j on the total flows from sector A to sector B .

Repeating this calculation for all companies, we get the matrix:

$$R = \langle x, y \rangle = \begin{pmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \ddots & \ddots & \vdots \\ x_n y_1 & x_n y_2 & \cdots & x_n y_n \end{pmatrix}$$

Finally, by calculating the entrywise product of R and the trade matrix F , we get the following matrix:

$$W = R \circ F = \begin{pmatrix} x_1 y_1 F_{1,1} & x_1 y_2 F_{1,2} & \cdots & x_1 y_n F_{1,n} \\ x_2 y_1 F_{2,1} & x_2 y_2 F_{2,2} & \cdots & x_2 y_n F_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ x_n y_1 F_{n,1} & x_n y_2 F_{n,2} & \cdots & x_n y_n F_{n,n} \end{pmatrix}$$

Note that the ij element can be interpreted as the proxy of the trade flow from company i to company j . Conversely, the ji element can be interpreted as the proxy of the trade flow from company j to company i . The estimated flows define the magnitude of intercompany connections. To use W as a spatial weighting matrix in our application, we need to set the entries on the diagonal to 0 and normalize the rows so as to sum to 1.

2.3. Estimation Procedure

To estimate the SAR model parameters, we use a two-step estimation procedure:

- (i) minimize Equation (8), letting $\Omega = I_{2k-1}$, to obtain parameter estimates $\hat{\theta}$ and calculate the optimal weighting matrix by computing the covariance of the moments:

$$\hat{S} = \frac{1}{n} Z' u u' Z$$

where the residual vector u is calculated as in (7).

- (ii) recompute the parameter estimates $\hat{\theta}$ by substituting the identity matrix with the optimal weight matrix:

$$\hat{\Omega} = \hat{S}^{-1}$$

Note that this procedure requires inversion and multiplication of large matrices, so the computation time can be very long when working with large datasets. Possible solutions should be based on suitable simplifications to the connectivity matrix W to make it more sparse, such as fixing a threshold for the relevance of trade flows. However, with our sample size ($n = 1,185$) the computational time for the two-step algorithm is more than acceptable. We remark that the employed data is available as **Supplementary Material**.

3. DATA AND RESULTS

In this section we empirically verify whether the predictive performance of P2P credit scoring models can be improved using correlation network models. In particular, we are interested in assessing significance and magnitude of the contagion parameter ρ . The more the contagion parameter is close to 1, the more the networking information can support credit risk evaluation. To achieve this goal, we have collected data from a European Credit Assessment Institution (ECAI), that supplies credit scorings to P2P platforms specialized in business lending. We use data relative to 1,185 borrowing Italian SMEs, in 2015–2016. The proportion of observed defaults in our sample is nearly 11%, which is large, in line with the observed impact of the recent financial crisis in Southern European countries. The available data include the status of the companies, classified as $[1 =$

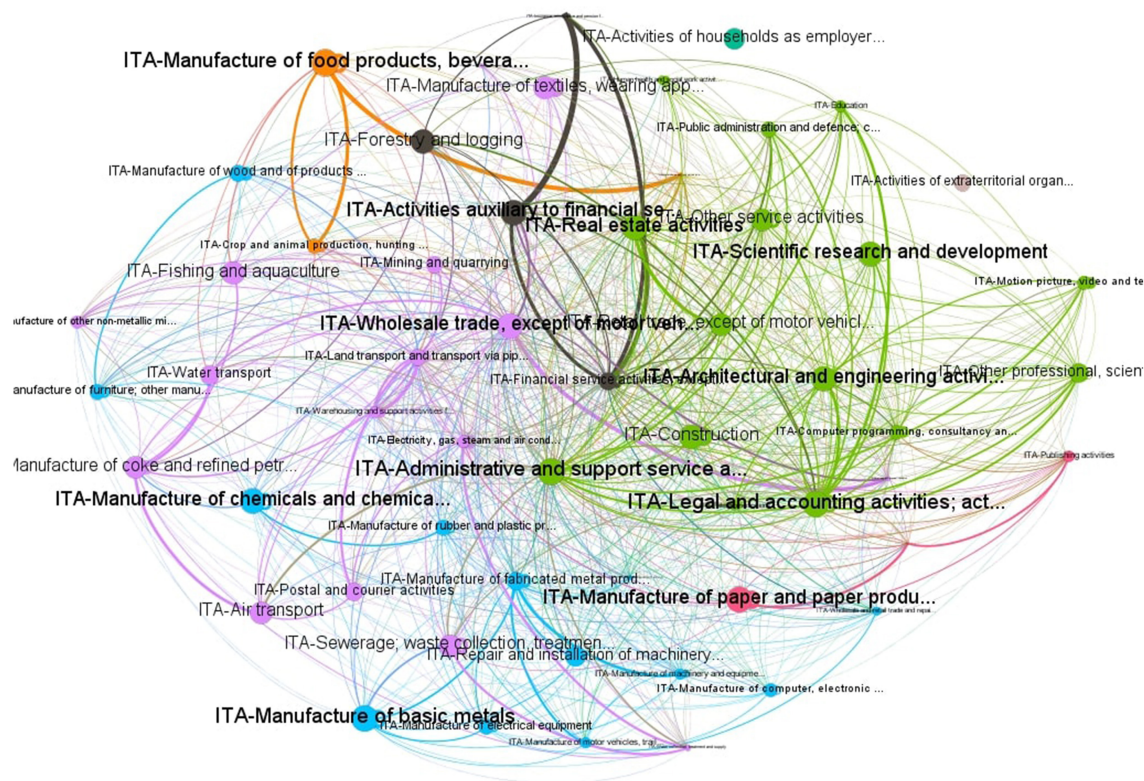


FIGURE 1 | Network of P2P Italian SMEs.

TABLE 1 | Results of estimation of non-spatial logit model.

	Estimate	Std. error	Pr(> z)
Intercept	-2.11	0.16	2.97e-38
β_1 (RATIO012)	-0.69	0.10	6.35e-11
β_2 (RATIO018)	0.02	0.10	0.84
β_3 (RATIO027)	-0.01	0.00	9.10e-04

Defaulted] and [0 = Active], in 2016 as well as some main financial information, for year 2015. From the available data, we select three financial ratios reflecting the three most important aspects related to default probability: operational performance, business sustainability and financial sustainability. Specifically, we consider:

- the return on equity ratio (RATIO012)
- the activity ratio, expressed as the ratio between sales and total assets (RATIO018);
- the solvency ratio, expressed as the ratio between the net income and the total debt (RATIO027)

The spatial weight matrix W has been built from the WIOT database, as described in section 2.2 and using turnover as a company size measure. **Figure 1** shows the network based on the estimated connections.

Table 1 shows the parameter estimates obtained using a simple logit model, without the spatial component.

TABLE 2 | Results of estimation of SAR model.

	Estimate	Std. error	Pr(> z)
ρ	0.78	0.23	5.44e-04
Intercept	0.44	0.46	0.35
β_1 (RATIO012)	-0.53	0.15	2.24e-04
β_2 (RATIO018)	0.05	0.13	0.69
β_3 (RATIO027)	-0.03	0.01	0.03

Then we estimate the SAR model (3) through the algorithm presented in section 2.3. The obtained results are reported in **Table 2**.

We first note from **Table 2** that the contagion parameter is significant and its value is high (0.78). The effect of financial ratios is stable, supporting the SAR specification including both a spatial and an exogenous component. Thus, considering a measure of connectivity between companies significantly explains the credit risk arising from P2P lending, improving the traditional analysis based on individual financial indicators.

Including the spatial component also improves model accuracy, as shown in **Figure 2** plotting the ROC curves of the simple logit and the spatial logit model. The AUC (Area Under the ROC Curve) values are 0.798 and 0.806, respectively. It is worth noting that the difference in the AUC values is modest and could turn out to be non-significant in an out-of-sample exercise.

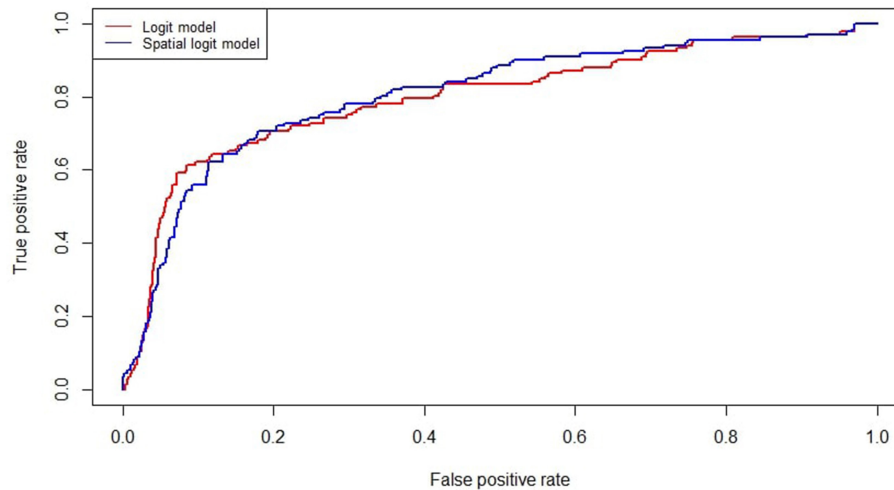


FIGURE 2 | ROC curves of baseline and spatial logit models.

However, the proposed specification defines a contagion model which can support the analysis of interconnectedness between the agents' default risk, even when this does not improve the predictive performance in a crucial way. Future research may concern dealing with unbalanced samples (as in Calabrese and Giudici, 2015) and/or with multiple data sources (as in Fagini and Giudici, 2011).

4. CONCLUSIONS

This paper provides a method, based on binary spatial regression models, to improve default prediction by estimating the interdependence between companies due to trade ties.

We have applied the methodology to a sample of Italian companies, finding evidence of a high level of spatial autocorrelation, interpretable as a credit contagion parameter.

The proposed model provides both a description of contagion (through the spatial component) and a predictive capability, differently from most existing contagion models, which provide either of the two. The model can be easily implemented, as a modification of a classical logistic regression that includes interconnectedness. We believe that the findings which can be derived from spatial autoregressive models may be useful, especially for P2P lenders who can use it to improve credit risk assessment.

From a methodological viewpoint, further research may involve employing a different generalized linear model, such as the generalized extreme value regression models discussed

in Calabrese and Elkink (2016). Moreover, the dependence structure could be extended to the dynamic case (Arakelian and Dellaportas, 2012).

DATA AVAILABILITY

The datasets for this manuscript are not publicly available because the data were provided by a private company. Requests to access the datasets should be directed to paolo.giudici@unipv.it.

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).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frai.2019.00006/full#supplementary-material>

REFERENCES

- Ahelegbey, D. F., Billio, M., and Casarin, R. (2016). Bayesian graphical models for structural vector autoregressive processes. *J. Appl. Econometr.* 31, 357–386. doi: 10.1002/jae.2443
- Arakelian, V., and Dellaportas, P. (2012). Contagion determination via copula and volatility threshold models. *Quant. Finan.* 12, 295–310. doi: 10.1080/14697680903410023
- Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and

- insurance sectors. *J. Finan. Econ.* 104, 535–559. doi: 10.1016/j.jfineco.2011.12.010
- Calabrese, R., and Elkink, J. (2014). Estimators of binary spatial autoregressive models: a Monte Carlo study. *J. Reg. Sci.* 54, 664–687. doi: 10.1111/jors.12116
- Calabrese, R., and Elkink, J. (2016). Estimating binary spatial autoregressive models for rare events. *Adv. Econometr.* 37, 147–168. doi: 10.1108/S0731-905320160000037012
- Calabrese, R., Elkink, J., and Giudici, P. (2017). Measuring bank contagion in Europe using binary spatial regression models. *J. Operat. Res. Soc.* 68, 1503–1511. doi: 10.1057/s41274-017-0189-4
- Calabrese, R., and Giudici, P. (2015). Estimating bank default with generalised extreme value regression models. *J. Operat. Res. Soc.* 66, 1783–1792.
- Claessens, S., Frost, J., Turner, G., and Zhu F. (2018). Fintech credit markets around the world: size, drivers and policy issues. *BIS Q. Rev.* 29–49.
- Diebold, F. X., and Yilmaz, K. (2014). On the network topology of variance decompositions: measuring the connectedness of financial firms. *J. Econometr.* 182, 119–134. doi: 10.1016/j.jeconom.2014.04.012
- Figini, S., and Giudici, P. (2011). Statistical merging of rating models. *J. Operat. Res. Soc.* 62, 1067–1074.
- Giudici, P., and Parisi, L. (2018). CoRisk: credit risk contagion with correlation network models. *Risks* 6:95. doi: 10.3390/risks6030095
- Giudici, P., and Spelta, A. (2016). Graphical network models for international financial flows. *J. Bus. Econ. Stat.* 34, 126–138. doi: 10.1080/07350015.2015.1017643
- Hautsch, N., Schaumburg, J., and Schienle, M. (2015). Financial network systemic risk contributions. *Rev. Finan.* 19, 685–738. doi: 10.1093/rof/rfu010
- Kelijian, H. H., and Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *J. Real Estate Finan. Econ.* 17, 99–121. doi: 10.1023/A:1007707430416
- Kelijian, H. H., and Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *J. Econometr.* 157, 53–67. doi: 10.1016/j.jeconom.2009.10.025
- Klier, T., and McMillen, D. (2008). Clustering of auto supplier plants in the United States: generalized method of moments spatial logit for large samples. *J. Bus. Econ. Stat.* 26, 460–471. doi: 10.1198/073500107000000188
- Pinkse, J., and Slade, M. E. (1998). Contracting in space: an application of spatial statistics to discrete-choice models. *J. Econometr.* 85, 125–154. doi: 10.1016/S0304-4076(97)00097-3

Conflict of Interest Statement: 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 Agosto, Giudici and Leach. 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.



Applied Machine Learning for Stochastic Local Volatility Calibration

Jürgen Hakala*

Quantitative Modeling Department, Leonteq Securities AG, Zurich, Switzerland

Stochastic volatility models are a popular choice to price and risk-manage financial derivatives on equity and foreign exchange. For the calibration of stochastic local volatility models a crucial step is the estimation of the expected variance conditional on the realized spot. The spot is given by the model dynamics. Here we suggest to use methods from machine learning to improve the estimation process. We show examples from foreign exchange.

Keywords: radial basis functions, machine learning, local stochastic volatility, derivatives pricing, finance

OPEN ACCESS

Edited by:

Peter Schwendner,
Zurich University of Applied Sciences,
Switzerland

Reviewed by:

Natalie Packham,
Hochschule für Wirtschaft und Recht
Berlin, Germany
Francesco Caravelli,
Los Alamos National Laboratory
(DOE), United States

*Correspondence:

Jürgen Hakala
juergen.hakala@leonteq.com

Specialty section:

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

Received: 17 January 2019

Accepted: 25 April 2019

Published: 17 May 2019

Citation:

Hakala J (2019) Applied Machine
Learning for Stochastic Local Volatility
Calibration. *Front. Artif. Intell.* 2:4.
doi: 10.3389/frai.2019.00004

1. INTRODUCTION

For derivatives pricing a major breakthrough was achieved with the risk-neutral valuation principle (Black and Scholes, 1973). Initially the models assumed a deterministic, state-independent volatility of the underlying price process. For many classes of underlying this assumption is too restrictive as it does not allow for an implied volatility that depends on strike as it is observed in the market, at least since the Black Friday in 1987, see for a review and attempted explanation (Benzoni et al., 2011).

Hence the most natural extension of the existing models was to postulate either a state-dependent volatility often duped as local volatility (Derman and Kani, 1994; Dupire, 1994) or to postulate an additional process for the volatility (e.g., Hull and White, 1987; Heston, 1993) which are labeled as stochastic volatility models.

Looking at the properties of these two model classes it was found (Hagan et al., 2002) that local volatility is postulating a dynamics which is not found in real markets. In foreign exchange options markets stochastic volatility models tend to exaggerate the effect of volatility convexity and at the same time these models are unable to match the short-dated volatility smile observed in market-prices. As a practical workaround, models that mix the local volatility and stochastic volatility were developed (Said, 1999; Blacher, 2001). It was observed that the calibration of SLV models is a hard problem which requires either a specific parametrization to derive fast pricing of vanilla options or quite time-consuming numerical optimization procedures (Guyon and Henry-Labordere, 2011). See as well Homescu (2014) for a great summary and best practice of local stochastic volatility models.

A shortcut to derive manageable calibration times was developed by Guyon and Henry-Labordere (2011) and Van der Stoep et al. (2014) using a Monte Carlo procedure to derive the required estimation of the conditional variance. In this paper we suggest to use methods from machine learning, in particular radial basis functions and variations thereof to derive fast and efficient estimators.

2. LOCAL STOCHASTIC VOLATILITY CALIBRATION

The LSV model in general is of the form:

$$\begin{aligned} dS_t &= \mu(t)S_t dt + \sigma(S_t, t)f(V_t)S_t dW_t \\ dV_t &= \mu_V(V_t)dt + \xi\chi(V_t)dX_t \\ < dW_t, dX_t > &= \rho dt \end{aligned}$$

with spot S_t , variance V_t , drift μ , (state-dependent) drift for the variance μ_V , vol of variance $\xi\chi(V_t)$, and correlation ρ . The LSV calibration is the process to determine the leverage function σ given the local volatility function σ_{Dupire}^2 and all the other parameters of the model. There is a fundamental relationship of the leverage function and the local volatility function (Dupire, 1996) where the expectation $E^{P(S_t, V_t, \sigma)}$ of the conditional variance V_t is taken with respect to the risk-neutral measure induced by the model. The notation indicates that $P(S_t, V_t, \sigma)$ is the joint probability of spot process S_t , variance process V_t and the solution for σ which depends on the probability distribution of (S_t, V_t) .

$$\sigma_{Dupire}^2(S_t, t) = E^{P(S_t, V_t, \sigma)}(V_t | S_t = S_t) \sigma^2(S_t, t) \quad (1)$$

Plugging the solution into the model equation makes this a McKean SDE where the expectation depends on the probability of the process itself.

To solve this equation Monte Carlo simulation can be used. The equations are discretized and the forward propagation of the spot S_t and variance V_t is interleaved with the estimation of conditional expectation using the realized paths of S_t and V_t . Contrary to standard Monte Carlo where all paths develop independently we need to bring all simulated paths to the estimation procedure. Using Euler discretization

$$\begin{aligned} \Delta \ln(S_t) &= \mu(t)\Delta t - \frac{1}{2}\sigma^2(S_t, t)f^2(V_t)\Delta t \\ &\quad + \sigma(S_t, t)f(V_t)\left(\sqrt{1-\rho^2}\Delta W + \rho\Delta X\right) \\ \Delta V_t &= \mu_V(V_t)\Delta t + \xi\chi(V_t)\Delta X \end{aligned}$$

with $\Delta W, \Delta X$ independent increments. The estimation of the conditional expectation can be seen as finding the function $R(S) = E(V_t | S_t = S)$ based on the samples as observed pairs

$$(S_t^1, V_t^1), \dots, (S_t^n, V_t^n)$$

where the spot S_t^i and variance V_t^i are the time t realizations of spot and variance on path i . Originally it was proposed to estimate the function R using kernel regression (Guyon and Henry-Labordere, 2011):

$$R((S^1, V^1), \dots, (S^N, V^N))(S) = \frac{\sum_{i=1}^N V_i K_h(S - S_i)}{\sum_{i=1}^N K_h(S - S_i)} \quad (2)$$

with Kernel functions K_h , where we dropped the t index as it is clear from the context. Alternatively (Van der Stoep et al., 2014) proposed to use binning techniques or sets of polynomials.

Subsequently we will evaluate alternative regression techniques to estimate the conditional expectation based on the realized paths. This can be rephrased as a supervised learning problem where each path is a (noisy) example.

3. REGRESSION AS A SUPERVISED LEARNING PROBLEM

The task to find a relationship between some input variables and an output from examples is one of the problems tackled by machine learning and is well studied as supervised learning. There are many classes of supervised learning algorithms and setups and we would like to demonstrate guidelines to which specific choices are suitable for the problem at hand. The basic problem is, given a set of examples x_i, y_i to find a function $f(x)$ such that an error functional is minimized. The task is to find a function such that there is low error on unseen examples, this is called generalization. There is a balance to strike between the error on the examples used for training and the error on the validation set of examples which are examples not used during training, for a fundamental analysis of the learning theory and the relation between capacity and the generalization (see e.g., Vapnik, 2013), in particular chapter 4.

3.1. Kernel Regression

The approach taken in Guyon and Henry-Labordere (2011) as stated above is Nadarajan-Watson kernel regression which is one of the so called non-parametric methods. The method is identical to Equation (2). The estimator is given as:

$$R((x_1, y_1), \dots, (x_N, y_N))(X) = \frac{\sum_{i=1}^N y_i K_h(x - x_i)}{\sum_{i=1}^N K_h(x - x_i)}$$

In this approach a Kernel function $K(x)$ is used, which satisfies:

$$\begin{aligned} K(x) &\geq 0 \\ K(x) &= K(-x) \\ \int_{-\infty}^{\infty} K(x)dx &= 1 \\ K_h(x) &= \frac{1}{h}K\left(\frac{x}{h}\right) \end{aligned}$$

There is a variety of Kernel functions well studied in the literature (Härdle, 1990):

- Gaussian $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$
- Quartic $\frac{15}{16}(1-x^2)^2$
- Epanechnikov $\frac{3}{4}(1-x^2)$
- Sigmoid $\frac{2}{\pi} \frac{1}{e^x + e^{-x}}$

Often the Kernel function used is Gaussian hence the support of the function is infinite or it will be the Epanechnikov Kernel which has bounded support.

The crucial choice is the bandwidth of the Kernel functions. There is a rule-of-thumb derived from normal distribution

assumptions, (Silverman, 1986):

$$h = \left(\frac{4\sigma^5}{3n} \right)^{\frac{1}{5}}$$

for the standard deviation σ of the data and n data points.

Alternatively cross validation, particular “leave-one-out cross validation” can be used to determine an optimal Kernel width. Cross validation is quite costly computationally and hence can only be used to cross check *ad-hoc* choices.

Local Linear Kernel Regression is a variation of Kernel regression which employs local linear terms and which is given by the solution of

$$R((x_1, y_1), \dots, (x_N, y_N))(x) = \min_{\alpha, \beta} \sum_{i=1}^N (y_i - \alpha - (x - x_i)\beta)^2 K_h(x - x_i)$$

The minimum is found by solving a 2×2 linear system .

In general the Kernel approaches suffers from some systematic shortcomings, mainly the fact that all examples are used, no compression happens, secondly a bias is introduced close to the boundary and the difficult choice of suitable bandwidth, where practically sound theoretical methods as cross validation cannot be used for computation time reasons.

3.2. Radial Basis Functions

Radial Basis Functions (RBF) and Partition of Unity Radial Basis Functions (PURBF) respectively take the form

$$RBF(x) = \sum_{i=1}^C w_i K_{h_i}(x - c_i)$$

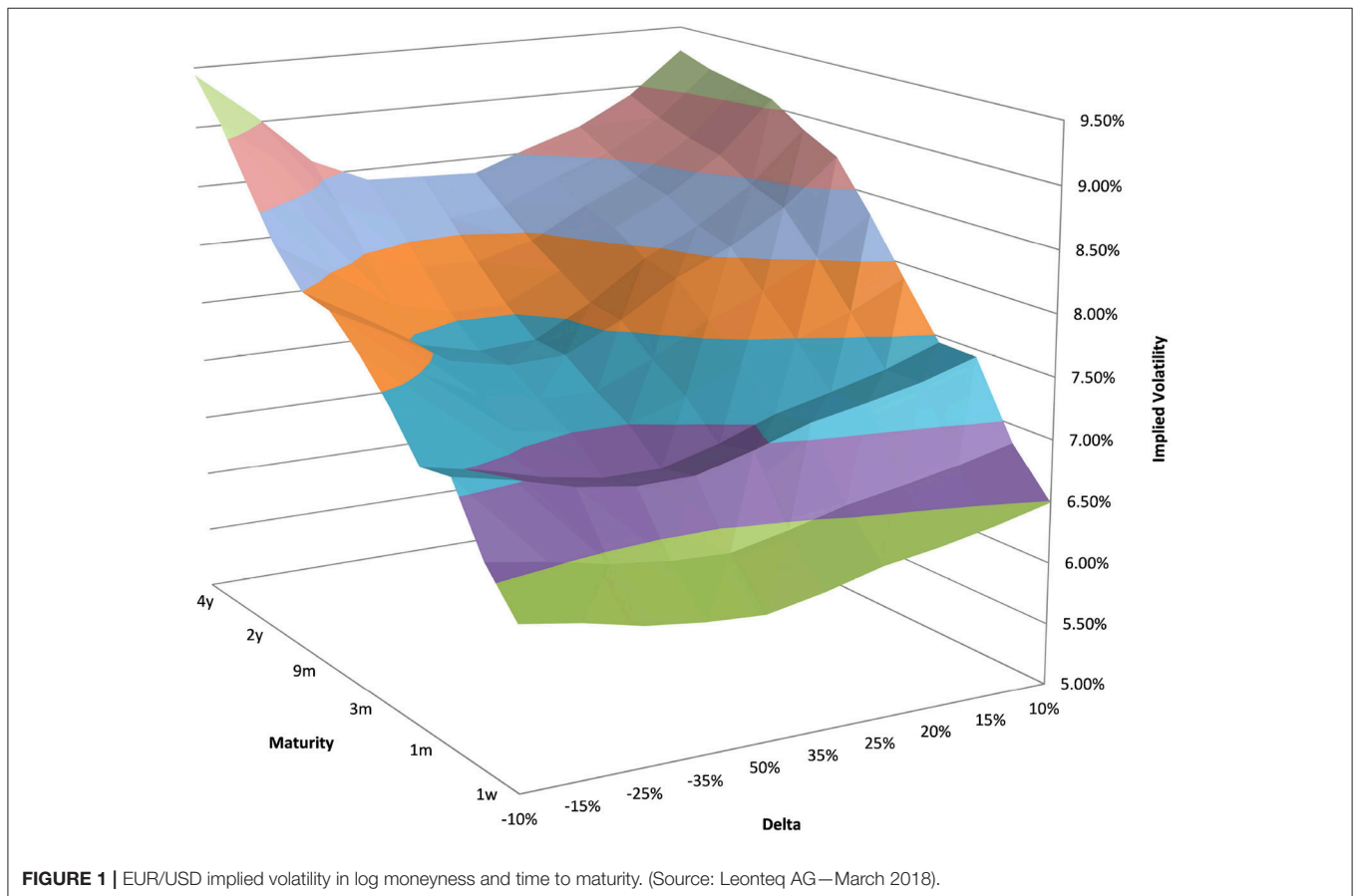
$$PURBF(x) = \frac{\sum_{i=1}^C w_i K_{h_i}(x - c_i)}{\sum_{i=1}^C K_{h_i}(x - c_i)}$$

PURBF are quite similar in functional form to Kernel regression. The main difference is that the number of basis functions is much smaller than the number of examples. It was proven that RBF and PURBF are universal function approximators (Hakala et al., 1994) which makes them suitable to approximate our estimation problem. If the L_2 norm is used the weights are optimized by solving the normal equation

$$LS = \frac{1}{2N} \sum_{i=1}^N (y_i - RBF(x_i))^2.$$

The solution s given by the weights w_i , which satisfy

$$w_i = (A^T A)_{ik}^{-1} A_{kj} y_j$$



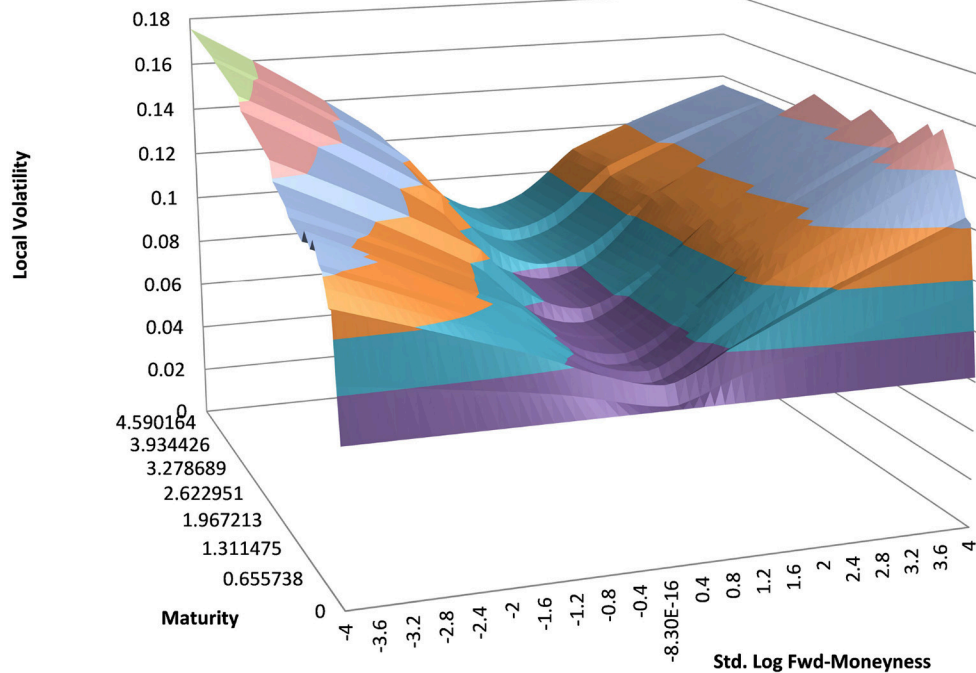


FIGURE 2 | EUR/USD local volatility in log moneyness and time to maturity. (Source: Leonteq AG—March 2018).

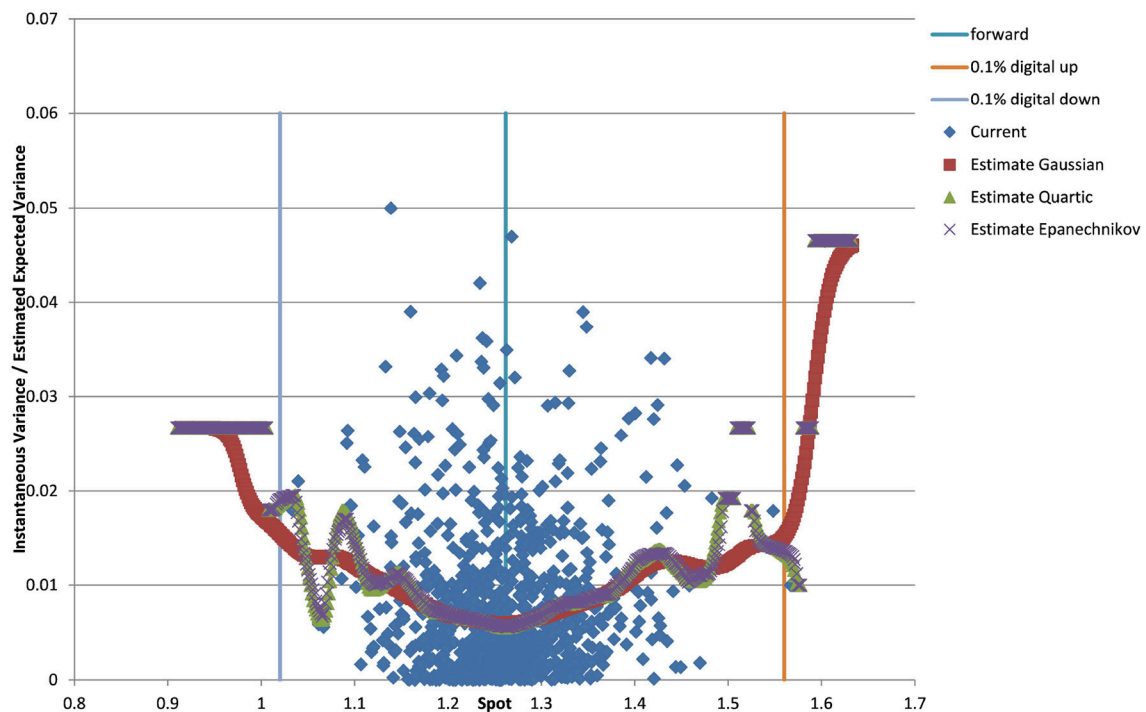


FIGURE 3 | EUR/USD 6M estimated variance conditional on realized spot and realized paths for various kernel functions as well as forward and digital levels. (Source: Leonteq AG—March 2018).

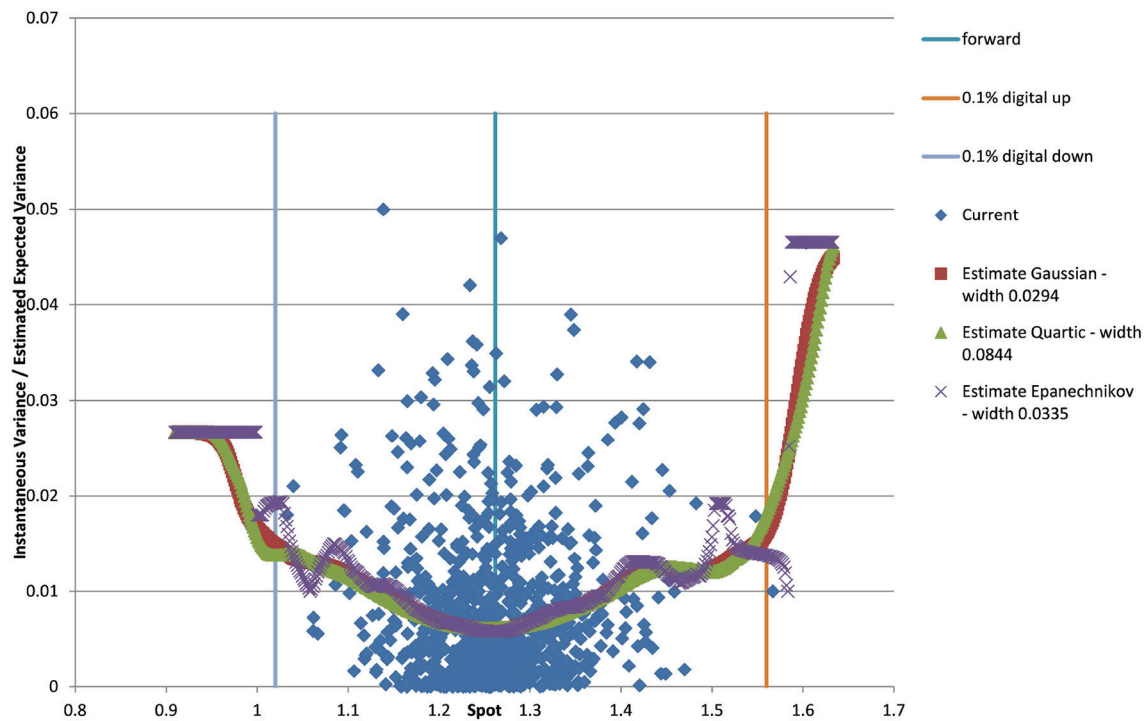


FIGURE 4 | EUR/USD 6M estimated variance conditional on realized spot and realized paths for various kernel functions and cross validated width as well as forward and digital levels. (Source: Leonteq AG—March 2018).

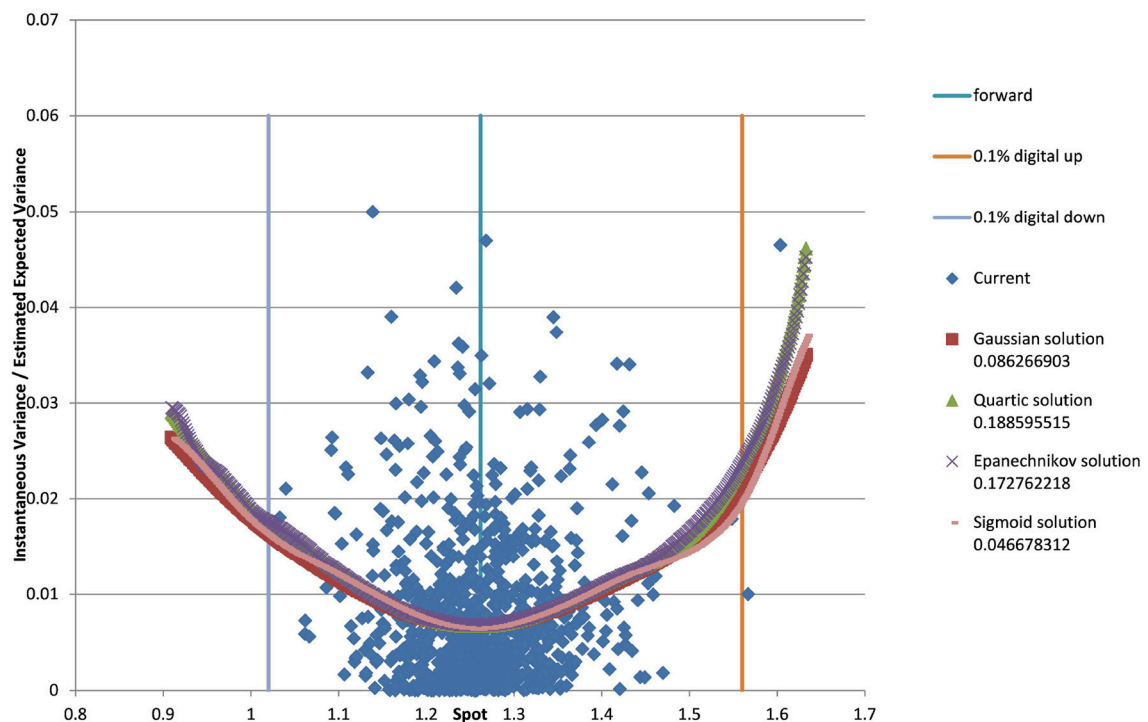


FIGURE 5 | EUR/USD 6M estimated variance conditional on realized spot and realized paths for various local linear kernel functions and cross validated width as well as forward and digital levels. (Source: Leonteq AG—March 2018).

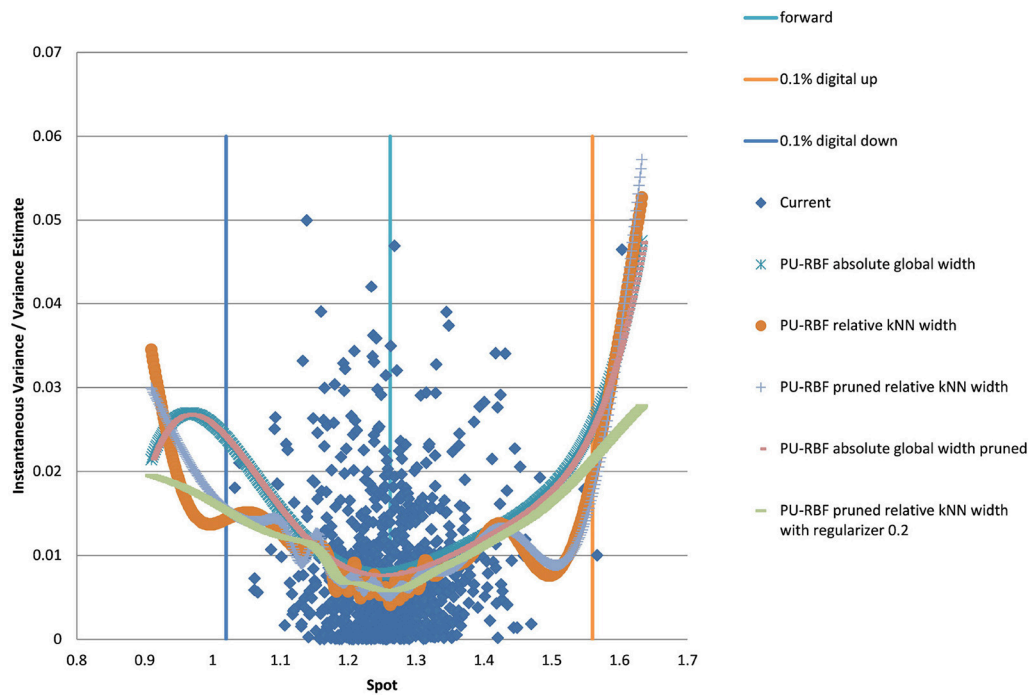


FIGURE 6 | EUR/USD 6M estimated variance conditional on realized spot and realized paths using PURBF as well as forward and digital levels. (Source: Leonteq AG—March 2018).

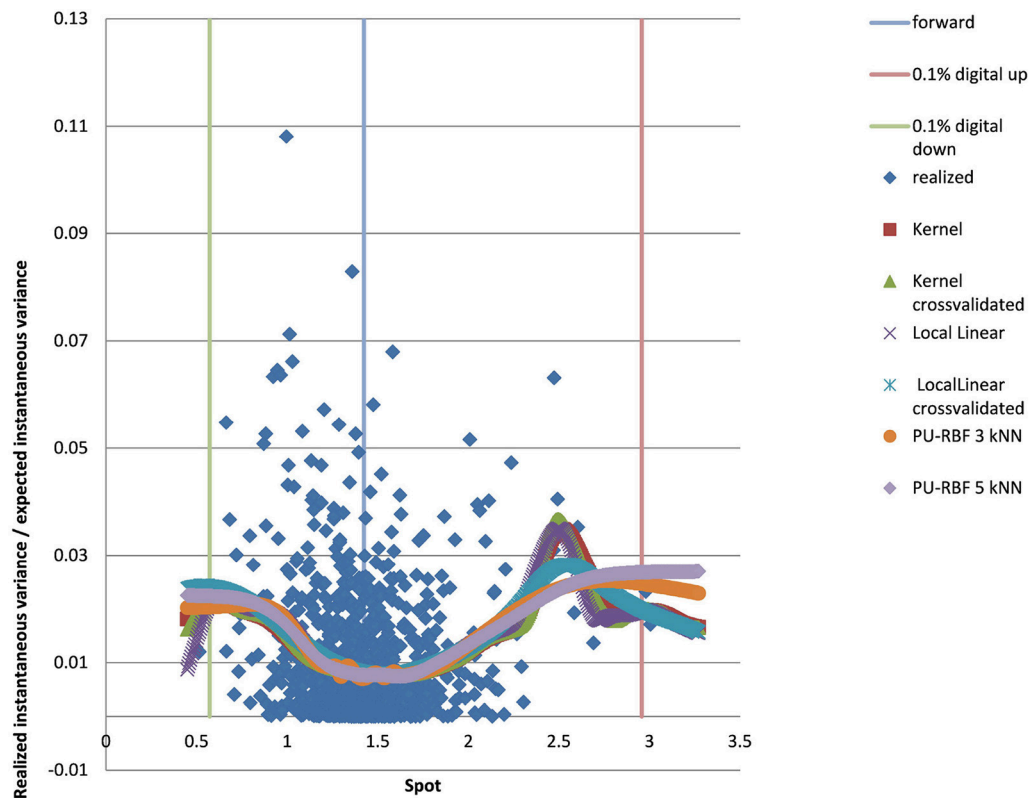


FIGURE 7 | EUR/USD 5Y estimated variance conditional on realized spot and realized paths using kernel regression, local linear kernel regression, and PURBF as well as forward and digital levels. (Source: Leonteq AG—March 2018).

with the matrix A_{ij} given as

$$A_{ij} = K_{h_j}(x_i - c_j). \quad (3)$$

The remaining parameters are determined heuristically:

- centers c_i - in 1D we chose $\min(x_i), \max(x_i)$, and a random subset of the remaining x_i . Centers which are too close (in their local width) are pruned—the criterion for unit j is $\min_i \left(\frac{|c_i - c_j|}{h_j} \right) \leq \Theta$ with a suitable global pruning constant Θ . There are resource allocating approaches as well (e.g., Fritzke, 1994) which seem less suitable here.
- width h_i - are chosen individually as a multiple of the k nearest neighbors or *ad-hoc*.

3.2.1. Regularization

Often the solution of the normal equation will be ill-conditioned. To counteract the bad conditioning of the problem and to get a better generalization we will use a regularizer on the L_2 norm of the weights (e.g., Goodfellow et al., 2016, Chapter 7.1).

$$LSR = \frac{1}{2N} \sum_{i=1}^N (y_i - RBF(x_i))^2 + \lambda \sum_{j=1}^C w_j^2$$

The corresponding solution is given as

$$w_i = (A^T A - \lambda \text{ id})_{ij}^{-1} A_{jk} y_k$$

with identity matrix id . The same solution applies for the PURBF function instead of the RBF one.

3.3. Computational Efficiency

For standard Kernel Regression computational effort is mainly due to sorting $O(n \log(n))$ of the spot observations to enable an efficient lookup of relevant spot observations during the retrieval phase. Optimal determination of width (cross-validation) requires the evolution of all kernels at all points several times which is very costly compared to the lookup. Local linear Kernel Regression requires an additional inversion of a 2×2 matrix which is negligible. For RBF and PURBF the solution to a small linear system is required. In particular the size is much smaller than the number of samples. Sorted examples can be used to optimize the training as the required matrix is determined by sums over the samples. Width and pruning computations require local computation of the order of the number of kernel functions. Overall the computational effort for RBF/PURBF is comparable and might be smaller in the retrieval phase than for Kernel Regression itself.

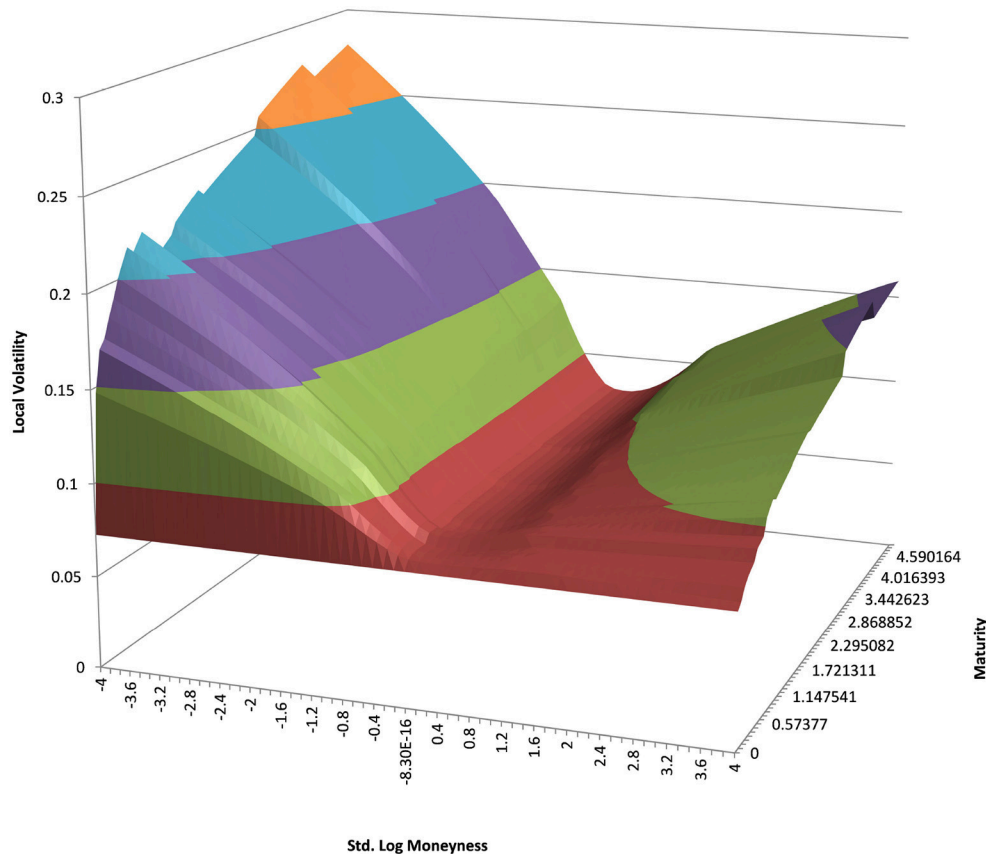


FIGURE 8 | USD/JPY local volatility in log moneyess and time to maturity. (Source: Leonteq AG—March 2018).

3.4. Alternative Architectures

In the last couple of years popularity of multilayer perceptron (MLP) and deep versions thereof grew enormously. For our application we rule out these architectures as the training is much more involved in the MLP case with a many remaining questions about a suitable number of hidden units, number of layers, type of activation functions. We could envision to use a pretrained MLP to get the solution without training. We postpone this approach for potential future use.

4. APPLICATION TO LSV MODEL IN FOREIGN EXCHANGE

The model we will study is of Heston type

$$\begin{aligned} dS_t &= \mu(t)S_t dt + \sigma(S_t, t)\sqrt{V_t}S_t dW_t \\ dV_t &= \kappa(\bar{V} - V_t)dt + \xi\sqrt{V_t}dX_t \\ \langle dW_t, dV_t \rangle &= \rho dt \end{aligned}$$

with mean reversion speed κ , mean reversion level \bar{V} .

The advantage is that we have a semi-closed form solution for vanilla call- and put options in the Heston model without the leverage function hence the first step is to calibrate the Heston model and then apply a scaling to the vol of vol parameter to reduce the SV impact and to let the local volatility compensate to match the vanilla option market. In this study we will use a volatility mixing of 66% which means that we scale the vol of variance by this factor before calibrating of the leverage function.

To compare the performance of the various regression algorithms on this model we will show for a specific slice the realized spot/variance and the corresponding results of the regression functions.

4.1. Example EUR/USD 6M

You can see the volatility surface in **Figure 1** and the corresponding local volatility surface in **Figure 2**. The snapshot of data, including spot, volatility, and interest rates was taken in March 2018. We show the results of different kernel estimators, using Silverman's rule of thumb for the width, including the samples indicated as *Current*, as well as the forward and the level of a 0.1% digital on the upside and downside **Figure 3**.

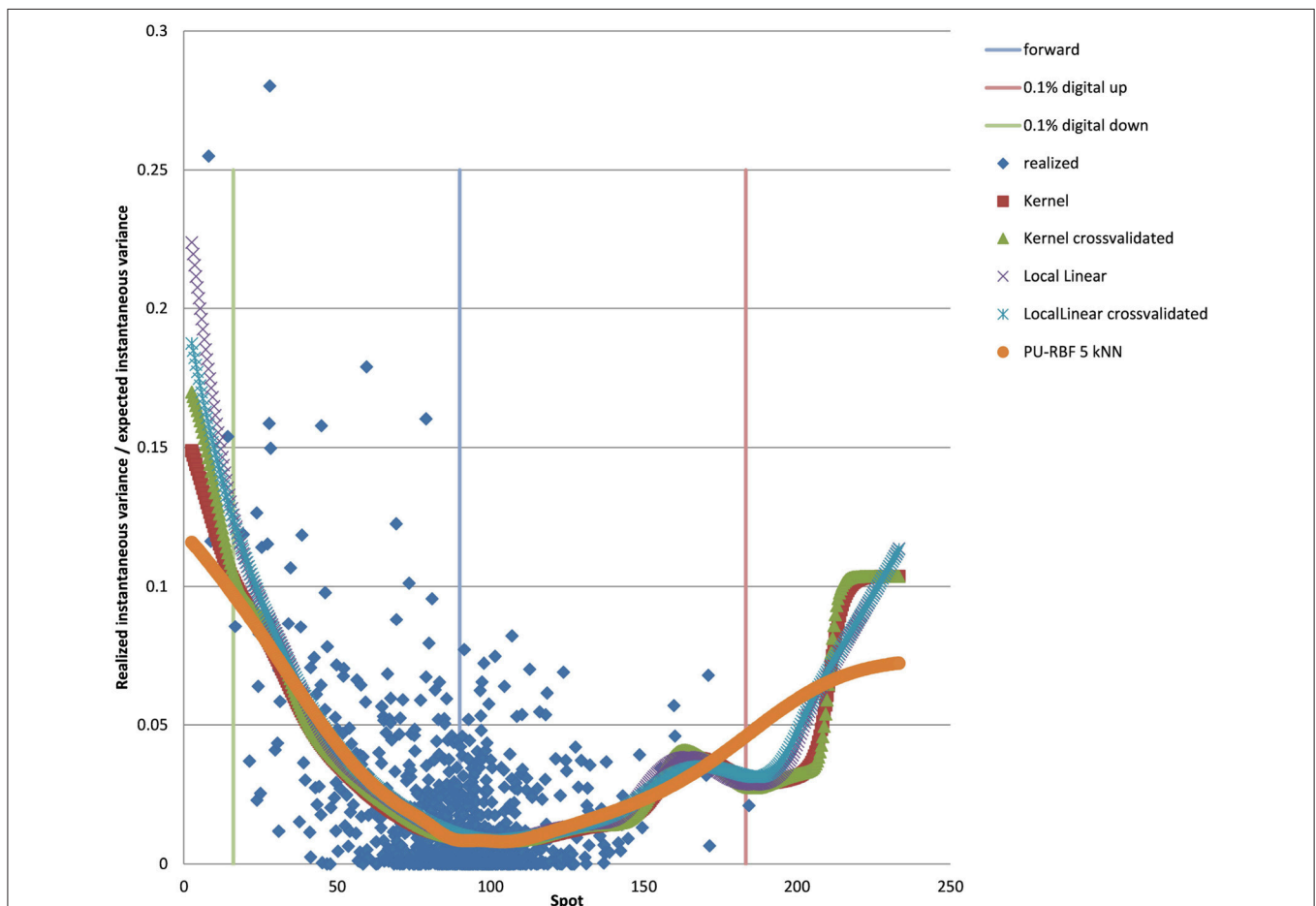


FIGURE 9 | USD/JPY 5Y estimated variance conditional on realized spot and realized paths using kernel regression, local linear kernel regression, and PURBF as well as forward and digital levels. (Source: Leonteq AG—March 2018).

In **Figure 4** you see the results of Kernel Regression for Silverman's rule and additionally the cross-validated width for the same kernel functions. Notice that the optimal width varies between different kernels. We show results for local linear kernel regression in **Figure 5**. It can be seen that the bias at the boundaries is reduced in comparison to the kernel regression. Again we show results for Silverman's rule and additionally the cross-validated width.

For the PURBF we show results in **Figure 6** using a global width, relative knn width, pruned and relative knn width, pruned with global width, and pruned, knn width and regularizer ($\lambda = 0.2$). We use 40 units in all cases as this number seems sufficiently versatile for the number of particles we want to use (2,048).

The last version with regularizer, pruning, and local width is the preferred version as it shows a smooth behavior without a bias in the boundaries and matches the part with many data-points in the middle without oscillations.

4.2. Example EUR/USD 5Y

We show the results for 5Y maturity and the same volatility surface in **Figure 7**. Among the tested approaches the PURBFwith 5 nearest neighbors performs best.

4.3. Example USD/JPY 5Y

We show the results for USD/JPY, see the local volatility surface in **Figure 8**. The estimation across the spot range is

shown in **Figure 9**. Again the PURBFwith 5 nearest neighbors performs best.

4.4. Example EUR/BRL 3Y

We show the results for EUR/BRL, which is a highly skewed and highly drifting underlying. See the local volatility surface in **Figure 10**. The estimation across the spot range is shown in **Figure 11**. Note that in this case the range of spot realizations is quite skewed as is expected from the skewed volatility surface. Nevertheless the PURBFwith 5 nearest neighbors puts a relatively smooth estimator through the samples and performs better than other methods.

4.5. Pricing Examples

To see the impact on exotics pricing we look at one-touch options. A one-touch option pays one unit of the counter currency at the maturity date if the spot trades at or beyond the touch-level at any time during the life of the option. We show the impact as a function of the Black-Scholes price (TV), similar to (Clark, 2011). The TV of a one-touch can be between 0% and the discount factor to maturity, which is in the range of 100%. For fixed market parameters like spot, volatility and the risk-neutral drift TV is a function of the touch-level only, hence makes a unique scale to show the model impact. The deviation of the LSV model price from the TV is the desired effect of an alternative model, which incorporates volatility risk management

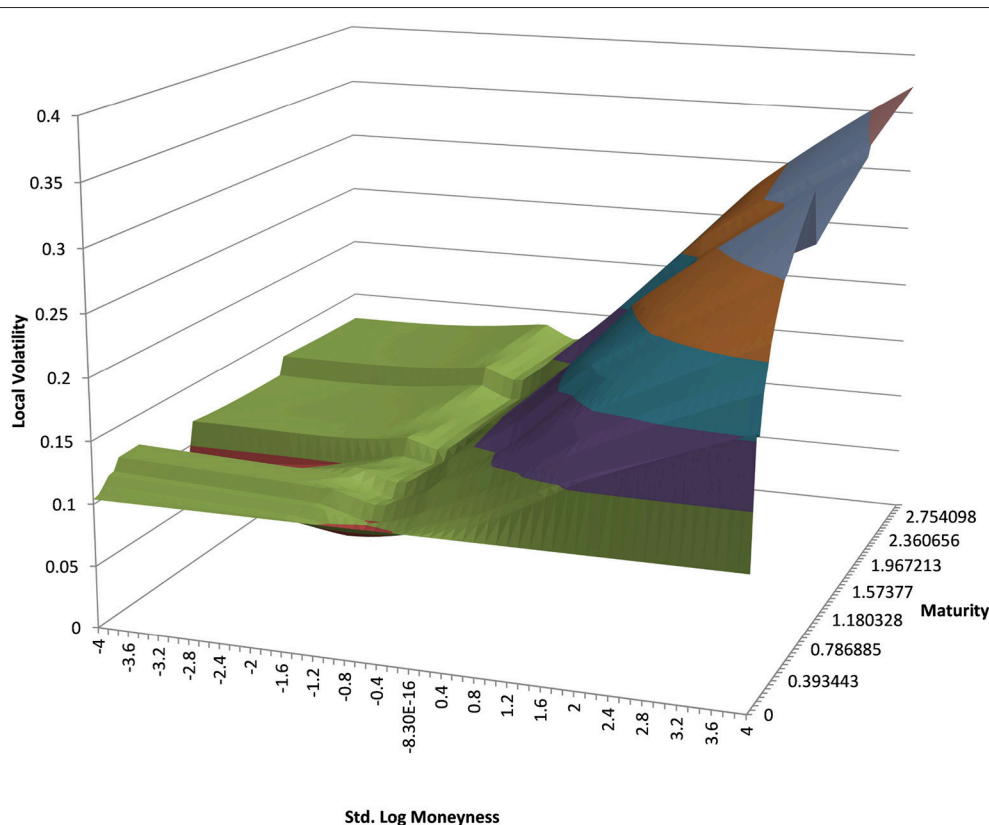


FIGURE 10 | EUR/BRL local volatility in log moneyness and time to maturity. (Source: Leonteq AG—March 2018).

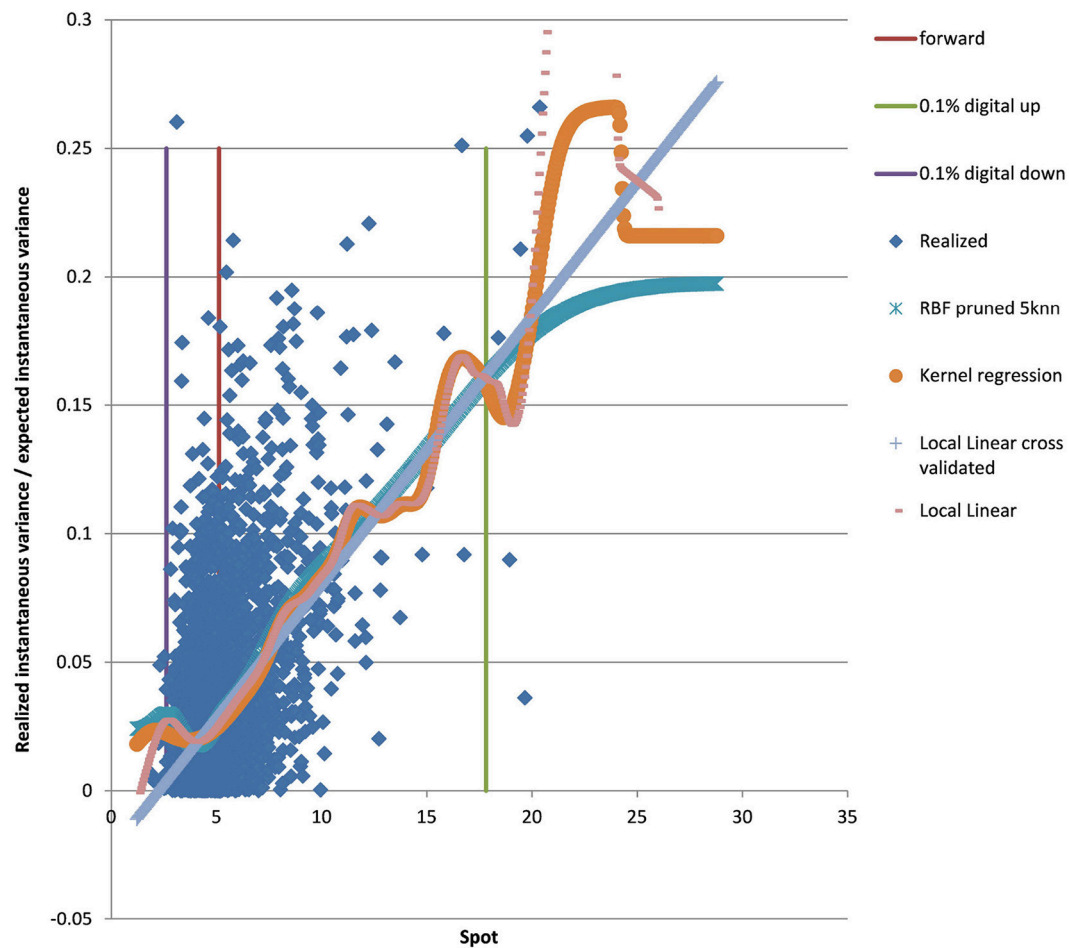


FIGURE 11 | EUR/BRL 3Y 5Y estimated variance conditional on realized spot and realized paths using kernel regression, local linear kernel regression, and PURBF as well as forward and digital levels. (Source: Leonteq AG—March 2018).

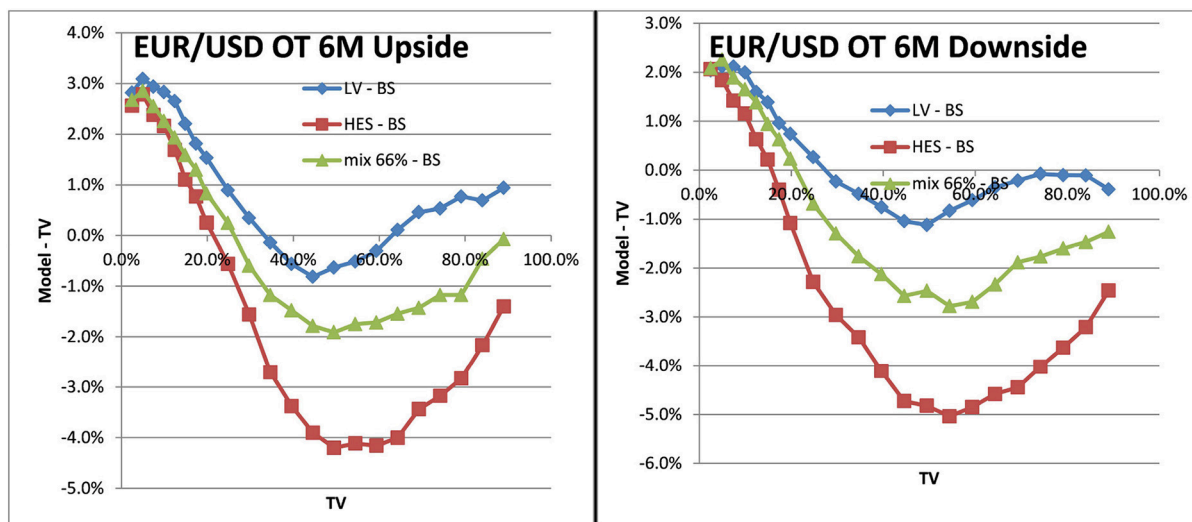


FIGURE 12 | One-touch prices EUR/USD 6M LV vs. TV, SV vs. TV, LSV vs. TV. Upside one-touches on the left, downside on the right. (Source: Leonteq AG—March 2018).

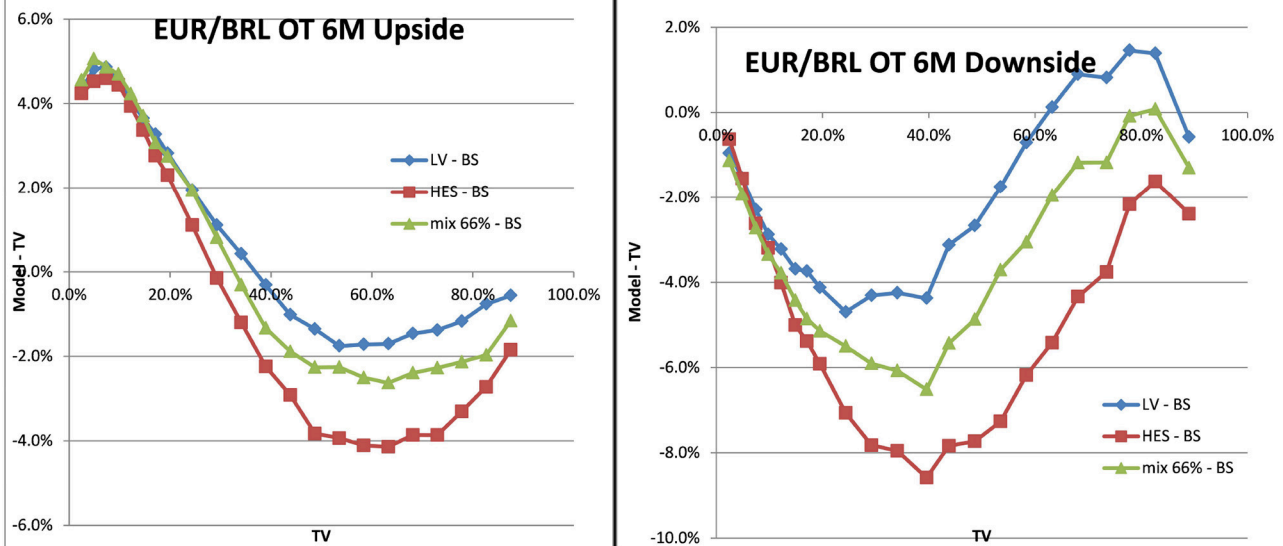


FIGURE 13 | One-touch prices EUR/BRL 6M LV vs. TV, SV vs. TV, LSV vs. TV. Upside one-touches on the left, downside on the right. (Source: Leonteq AG—March 2018).

and hedging, compared to the Black–Scholes model. The form of the deviation is not obvious and would require a rather complicated hedging argument of volatility risk and cross spot–volatility risk.

With the mixed local stochastic volatility model and mixing rate of 66% we expected the LSV price to be within the bounds of stochastic and local volatility price. We use Monte Carlo pricing with a fixed number of 32,000 paths (antithetic) and Quasi Random Numbers, a time discretization of 5 days and fixed 2048 particles. We denote the Black–Scholes prices as BS or TV (theoretical value) in the graphs and use LV as abbreviation for prices in local volatility and HES for the Heston model without local volatility component. The prices can be seen in **Figure 12** for EUR/USD 6M and in **Figure 13** for EUR/BRL 6M. We observe the expected behavior in all cases, the mixed local stochastic volatility prices are within the range of local and stochastic volatility prices and the mixing parameter can be used to adjust the behavior to observed exotics prices (e.g., one touches) in the market. Usually this mixing parameter is quite stable across longer periods, often weeks or even months.

5. CONCLUSION

We apply machine learning principles to improve the calibration process of the local stochastic volatility models. The suggested meta parameters and heuristics seem to apply to a wide variety of

underlyings in FX, liquid pairs like EUR/USD as well as emerging markets as EUR/BRL. The computational efficiency is at about the same level as for the formerly suggested Kernel Regression based approach. The results given by the PURBF function with pruning, regularization, and local width determined by 5 nearest neighbor performed significantly better than the Kernel based approaches, hence we would suggest to consider this approach in the calibration process.

Further work will be dedicated to improve the computational speed and to establish better measures of the quality. In particular in situations where vol surfaces are almost arbitrageable we will need the method to continue to provide numerically stable results.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

I would like to thank the Quant team of Leonteq to provide such a stimulating atmosphere, providing an environment to engage on a wide variety of practical and research oriented aspects of Quantitative Finance. In particular thanks to Nadzeya Bedziuk and Dmitry Davydov to critically crosscheck the document, still all errors are due to me.

REFERENCES

Benzoni, L., Pierre, C.-D., and Goldstein, R. S. (2011). Explaining asset pricing puzzles associated with the 1987 market crash. *J. Financ. Econ.* 101, 552–573. doi: 10.1016/j.jfineco.2011.01.008

Blacher, G. (2001). “A new approach for designing and calibrating stochastic volatility models for optimal delta-vega hedging of exotics,” in *ICBI Global Derivatives, Conference Presentation* (Juan-Les-Pins).

Black, F., and Scholes, M. (1973). The pricing of options and corporate liabilities. *J. Polit. Econ.* 81, 637–654.

- Clark, I. J. (2011). *Foreign Exchange Option Pricing: A Practitioner's Guide*. John Wiley & Sons.
- Derman, E., and Kani, I. (1994). Riding on a smile. *Risk* 7, 32–39.
- Dupire, B. (1994). Pricing with a smile. *Risk* 7, 18–20.
- Dupire, B. (1996). “A unified theory of volatility,” in *Derivatives Pricing: The Classic Collection*, ed P. Carr (Risk Books), 185–196.
- Fritzke, B. (1994). Fast learning with incremental RBF networks. *Neural Process. Lett.* 1, 2–5. doi: 10.1007/BF02312392
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. Available online at: <https://www.deeplearningbook.org/>
- Guyon, J., and Henry-Labordere, P. (2011). The smile calibration problem solved. *SSRN Electron. J.* doi: 10.2139/ssrn.1885032
- Hagan, P. S., Kumar, D., Lesniewski, A. S., and Woodward, D. E. (2002). “Managing smile risk,” in *The Best of Wilmott*, ed P. Wilmott (Hoboken, NJ: Wiley), 249–296.
- Hakala, J., Koslowski, C., and Eckmiller, R. (1994). “Partition of unity rbf networks are universal function approximators,” in *ICANN 94* (Berlin; Heidelberg: Springer), 459–462.
- Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press.
- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.* 6, 327–343. doi: 10.1093/rfs/6.2.327
- Homescu, C. (2014). Local stochastic volatility models: calibration and pricing. *SSRN Electron. J.* doi: 10.2139/ssrn.2448098
- Hull, J., and White, A. (1987). The pricing of options on assets with stochastic volatilities. *J. Finan.* 42, 281–300. doi: 10.1111/j.1540-6261.1987.tb02568.x
- Said, K. (1999). Pricing exotics under the smile. *Risk* 12, 72–75.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. New York, NY: Chapman Hall.
- Van der Stoep, A. W., Grzelak, L. A., and Oosterlee, C. W. (2014). The heston stochastic-local volatility model: efficient Monte Carlo simulation. *Int. J. Theor. Appl. Finan.* 17:1450045. doi: 10.2139/ssrn.2278122
- Vapnik, V. (2013). *The Nature of Statistical Learning Theory*. Springer Science & Business Media.

Legal Disclaimer: This publication serves only for information purposes and is not research; it constitutes neither a recommendation for the purchase or sale of any financial products nor an offer or invitation for an offer. No representation or warranty, either express or implied is provided in relation to the accuracy, completeness, or reliability of the information herein. Before investing in financial products, investors are highly recommended to contact their financial advisor for advice specifically focused on the investor's individual situation; the information contained in this document does not substitute such advice.

Conflict of Interest Statement: The 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 © 2019 Hakala. 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.



Network Based Scoring Models to Improve Credit Risk Management in Peer to Peer Lending Platforms

Paolo Giudici^{1*}, Branka Hadji-Misheva² and Alessandro Spelta¹

¹ Department of Economics and Management, Fintech Laboratory, University of Pavia, Pavia, Italy, ² School of Engineering, Zurich University of Applied Sciences (ZHAW), Winterthur, Switzerland

OPEN ACCESS

Edited by:

Ronald Hochreiter,
Vienna University of Economics and
Business, Austria

Reviewed by:

Simone Righi,
University College London,
United Kingdom
Francesco Caravelli,
Los Alamos National Laboratory
(DOE), United States

*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: 27 February 2019

Accepted: 23 April 2019

Published: 24 May 2019

Citation:

Giudici P, Hadji-Misheva B and
Spelta A (2019) Network Based
Scoring Models to Improve Credit Risk
Management in Peer to Peer Lending
Platforms. *Front. Artif. Intell.* 2:3.
doi: 10.3389/frai.2019.00003

Financial intermediation has changed extensively over the course of the last two decades. One of the most significant change has been the emergence of FinTech. In the context of credit services, fintech peer to peer lenders have introduced many opportunities, among which improved speed, better customer experience, and reduced costs. However, peer-to-peer lending platforms lead to higher risks, among which higher credit risk: not owned by the lenders, and systemic risks: due to the high interconnectedness among borrowers generated by the platform. This calls for new and more accurate credit risk models to protect consumers and preserve financial stability. In this paper we propose to enhance credit risk accuracy of peer-to-peer platforms by leveraging topological information embedded into similarity networks, derived from borrowers' financial information. Topological coefficients describing borrowers' importance and community structures are employed as additional explanatory variables, leading to an improved predictive performance of credit scoring models.

Keywords: contagion, credit risk, credit scoring, network models, peer to peer lending

1. INTRODUCTION

Financial intermediation has changed extensively over the course of the last two decades mostly due to technological advancement. One of the most significant change has been the emergence of FinTech that is nowadays altering many financial products, services, production processes, and organizational structure. In the context of commercial credit, FinTech solutions have introduced many opportunities for both lenders and borrowers thus redefining the role of traditional intermediaries. Peer-to-peer lending platforms, often abbreviated P2P lending, allow private individuals to directly run small and, in most cases, unsecured loans to private borrowers or small and medium enterprises (SME). The recent advances in information technology have enabled these online platforms to provide an alternative to traditional financial intermediaries, by delivering more cost efficient, consumer friendly and transparent lending services, improving the overall value for customers (for a review see e.g., Claessens et al., 2018; Giudici and Misheva, 2018).

The literature identifies many factors which explain the increasing role of P2P lending platforms in the global world of finance (see e.g., Serrano-Cinca and Gutiérrez-Nieto, 2016). For instance, P2P platforms are not required to respect bank capital requirements nor to pay fees associated with state deposit insurance practices, and this allows them to operate with lower costs. Thus, borrowers benefit because they are able to receive credits at lower interest rates, and in some cases with little or no collateral, whereas lenders can receive higher rates of return on investment, due to reduced transaction costs (see Emekter et al., 2015). Second, advancements in information

technology have also been a key force driving the exponential growth of P2P platforms (see Guegan and Hassani, 2017). In this context, many P2P platforms rely not only on "hard" but also on "soft" i.e., social network activity information for the purpose of evaluating a candidate's creditworthiness, a practice not typically employed by traditional banks. The third factor explaining the rapid growth of P2P platforms is related with regulatory aspects. With the new revised Payment Service Directive (PSD2), that came in effect in 2018, the monopoly which banks have on their clients account information and payment transactions becomes weaker as this information can be disclosed through application payment interfaces. From a different viewpoint, the rapid growth of the importance of P2P lending platforms can pose significant risks to financial stability. This because P2P lenders typically produce inadequate measures of credit risk. In comparison with traditional banks, P2P platforms are less able to eliminate asymmetric information, thus increasing the risk of bad debt accumulation because they have no access to detailed information on borrowers past financial transaction.

Moreover, P2P lending activity is built on the basis of a "many-to-many" approach, in which the financial intermediary empowers each lender to decide to whom borrower to lend and for what amount. This leads to a strong interdependence between the borrowers and the lenders, which may generate high levels of contagion and systemic risk.

Even more importantly, P2P lenders allow for direct matching between borrowers and lenders, without the loans being held on the intermediary's balance-sheet; in other words, in a P2P platform, the risk is fully born by the lender. From a risk-return perspective, while in classical banking a financial institution chooses its optimal trade-off between risks and returns (subject to regulation constraints), in P2P lending, the platform maximizes its returns without taking care of the risks which are borne by the lenders.

The misaligned incentives, asymmetric information, differences in the business model and in the risk ownership may lead to the platform not being able to correctly distinguish between different risk classes which in turn can impact the overall stability of the financial system. In this paper we propose to exploit topological information embedded into similarity networks to increase the predictive performance of some credit scoring models.

Understanding the structure of a similarity network (see Mantegna and Stanley, 1999) is indeed instrumental for understand the origin of companies failures and to inform policymakers on how to prepare for, and recover from, adverse shocks hitting the network. Similarity patterns between companies' features can be extracted from a distance matrix and they can reveal how credit risk is related to the topology of the network. To account for such topological information we rely on centrality measures and community structure detection (see e.g., Newman, 2018). We show that the inclusion of these variables into credit scoring models does improve their predictive utility. Results confirm the validity of this approach in discriminating between defaulted and sound institutions, thus, the proposed methodology can constitute a new instrument in both policy-makers and practitioners toolboxes. We remark that our work is

related to two main other recent research streams. First, some authors have carried out investigations on the accuracy of credit scoring models of P2P platforms (Serrano-Cinca et al., 2016). We improve these contributions by extending the methodology to also account for the interconnections that emerge between economic agents. Second, our network approach relates to a recent and fast expanding line of research which focuses on the application of network analysis tools, for the purpose of understanding flows in financial markets, as in the papers of Allen and Gale (2000), Leitner (2005), and Giudici and Spelta (2016). We improve these contributions, extending them to the P2P context and linking network models, that are often merely descriptive, with statistical and machine learning models, thus providing a predictive framework. The rest of the paper is organized as follows: section 2 introduces the data set we employ in the analysis together with the description of the credit scoring models and of the performance measures. In this section we also present the metric used for extracting distances between the borrowing companies and the methods employed for building the networks and for extracting topological information. Section 3 is devoted to show the results of the analysis and the comparison between the performances of the credit scoring models with and without the topological information. Section 4 concludes.

2. DATA AND METHODOLOGY

In this section we first describe the data set employed in our analysis and the necessary pre-processing stage. Subsequently we introduce the families of credit scoring models and the non-parametric measures used for testing the performance of such models. Then we focus on showing how one can extract relevant patterns of similarities to build up meaningful networks from balance-sheet features of borrowing companies.

We consider data supplied by the European External Credit Assessment Institution (ECAI) that specializes in credit scoring for P2P platforms focused on SME commercial lending. Specifically, the analysis relies on a data set, that is composed of official financial information (financial ratios constructed on the basis balance sheet and income statement information) on 4514 Italian SMEs which represent the target of P2P lending platforms. **Appendix A** provides a table encompassing formulas to compute such ratios. **Table 2**, instead, provides the summary statistics of the variables included in this data set and information concerning their mean value aggregated by the status of the companies (active and defaulted). It is important to note that none of the variables included in data set contains missing values and the proportion of defaulted companies is 11%.

What is noticeable from **Table 1**, is that, as in most real-world data sets (and particularly those reflecting the operations of start-ups and small and medium enterprises), for most variables, there is a noticeable presence of unusually large or small values when compared to the mean. The literature recognizes many methods for dealing with outliers however in most cases the correct application of these methods is based on very strong assumptions concerning the size and distribution of the data set as well as the

TABLE 1 | Summary statistics of variables included in the dataset.

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max	Active	Default
ratio001	4,514	8.885	19.155	−64.430	1.303	9.680	206.550	8.85	9.15
ratio002	4,514	1.264	3.333	−10	0	1.2	33	1.25	1.35
ratio003	4,514	1.444	0.761	0.170	1.070	1.520	8.270	1.49	1.09
ratio004	4,514	1.536	1.201	0.010	0.970	1.720	13.710	1.6	1.04
ratio005	4,514	1.190	1.024	0.000	0.610	1.407	10.880	1.24	0.76
ratio006	4,514	7.726	23.277	−33.140	0.940	4.890	297.020	7.93	6.09
ratio008	4,514	23.068	70.271	−285.860	1.240	16.317	566.960	26.22	−2.33
ratio011	4,514	0.028	0.147	−1	0.01	0.1	0	0.05	−0.13
ratio012	4,514	−0.069	0.790	−8.540	0.000	0.210	1.080	0.01	−0.69
ratio017	4,514	1.372	1.068	0.010	0.680	1.740	8.420	1.38	1.30
ratio018	4,514	1.335	1.064	0.010	0.640	1.700	8.420	1.34	1.29
ratio019	4,514	0.194	0.498	−3.320	0.010	0.390	3.950	0.21	0.05
ratio027	4,514	36.513	92.893	−191.630	2.470	27.608	747.010	40.18	6.96
ratio029	4,514	0.062	0.196	−2	0.02	0.1	1	0.08	−0.12
ratio030	4,514	0.068	0.216	−2	0.02	0.1	1	0.09	−0.12
DIO	4,514	105.228	355.807	0	1	80	5.569	100.61	142.47
DPO	4,514	75.934	111.651	0	0	99.8	1.467	67.35	145.18
DSO	4,514	95.732	128.370	0	0	136	1.465	91.07	133.32
turnover	4,514	3,344.479	7,580.559	6	594	2,761.8	76.403	3,542.27	1,749.41

For each measure we report the average (Mean) along with the standard deviation (St. Dev.), the minimum (Min), the 25-th and 75-th percentiles (Pctl), the maximum (Max), mean value of the variable for active companies (Active), mean value of the variable for defaulted companies (Defaulted).

randomness of the outliers. In this context, we do not substitute or cancel outliers because we believe they can provide important insights concerning the companies included in the sample. All data and code employed is available as **Supplementary Material**.

2.1. Credit Risk Models

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 (see Khandani et al., 2010; Yu et al., 2010; Khashman, 2011; Lessmann et al., 2015; Abellán and Castellano, 2017 to cite few).

Consider N firms having observation regarding T different variables (usually balance-sheet measures or financial ratios). For each institution n define a variable γ_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. In a nutshell, credit risk models develop relationships between the explanatory variables embedded in T and the dependent variable γ .

Against this background, we employ logistic regression, discriminant analysis, classification and regression trees and support vector machine (Anderson, 2007). The following paragraphs briefly summarize the characteristics of the models we use for the present analysis.

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 v.s. 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 α is the model intercept while β_t is the t -th regression coefficient. It follows that the probability of default can be found as:

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

Discriminant analysis assumes that different classes generate data based on different Gaussian distributions. Linear discriminant analysis (LDA) approaches the problem by assuming that the conditional probability density functions $p(\mathbf{x}|\gamma = 0)$ and $p(\mathbf{x}|\gamma = 1)$ are both normally distributed with mean and covariance parameters (μ_0, \mathbf{V}_0) and (μ_1, \mathbf{V}_1) respectively. In this context, the decision rule is based on the Linear Score Function, a function of the population means for each of the populations, i , as well as the pooled variance-covariance matrix.

Classification and regression trees (CART) is another widely used statistical technique in which a dependent variable is associated with a set of input factors through a recursive sequence of simple binary relations. Put simply, it is a step-by-step process which results in a decision tree which is constructed either by splitting or not splitting each node into daughter nodes. The splitting strategy follows a node impurity function meaning that at each stage of the recursive partitioning, all possible splits are considered and the one which leads to the greatest increase in node purity is chosen.

Support vector machine (SVM) classifies data by detecting the best hyperplane that separates all data points of one class from those of the other class. Given a data set of N institutions of the

form $(\mathbf{x}_1, \gamma_1), \dots, (\mathbf{x}_N, \gamma_N)$ where the γ_n indicates the class to which the point \mathbf{x}_n belongs. Each \mathbf{x}_n is a T -dimensional real vector. SVM finds the “maximum-margin hyperplane” that separates data points \mathbf{x}_n for which $\gamma = 1$ from the data points for which $\gamma = 0$, which is defined so that the distance between the hyperplane and the nearest point \mathbf{x}_n from either group is maximized. In formula:

$$\max_{\mathbf{w} \in \mathbb{R}^T, b \in \mathbb{R}} \min_{\mathbf{x} \in A \cup B} \frac{|\mathbf{w}'\mathbf{x}_i + b|}{\|\mathbf{w}\|} \quad (3)$$

where A and B are disjoint subsets and $\mathbf{w}\mathbf{x} - b = 0$ represents a hyperplane.

2.2. Assessing Model Performance

For evaluating the performance of each model, we employ, as a reference measure, the indicator $\gamma \in \{0, 1\}$ that is a binary variable which takes value one whenever the institutions has defaulted and value zero otherwise. For detecting default events represented in γ , 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 γ 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) and Gini coefficient. The ROC curve plots the false positive rate (FPR) against the true positive rate (TPR). To be more explicit:

$$FPR = \frac{FP}{FP + TN} \quad (4)$$

$$TPR = \frac{TP}{TP + FN} \quad (5)$$

Moreover, we also compute other measures for assessing models performance such as the accuracy and the KS statistic. The overall accuracy of each model can be computed as:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

and it characterizes the proportion of true results (both true positives and true negatives) among the total number of cases under examination. In this context a key issue is setting the threshold at which a company is classified as belonging to one class rather than another.

Additional to this, another often-used characteristic in describing the quality of the model (or the scoring function) is the Kolmogorov-Smirnov statistic (KS). This metric too seeks to jointly consider specificity and sensitivity and it corresponds to the maximum value of their sum as the threshold is varied. Put differently, it represents the maximum difference between the cumulative distribution of active and defaulted companies. Consequently, the KS statistics is defined as:

$$KS = \max_j |F_{Active}(x_j) - F_{Defaulted}(x_j)|$$

For back-testing, while assessing the performance of each model, available information must be exploited in a realistic manner. To this end, we perform repeated sub-sampling validation approach. Specifically, we randomly split the data set in 10 training and validation data sets. For each such split, the model is fitted on the training data set and predictive utility is assessed on the corresponding testing data. The results concerning the model accuracy (area under the ROC curve, KS statistic, Gini index) are then averaged over the splits.

2.3. The Distance Metric

In the present study we exploit information derived from financial statements of borrowing companies collected in a vector \mathbf{x}_n representing the financial composition of the balance-sheet of institution n . We define a metric that provides the relative distance between companies by applying the standardized Euclidean distance between each pair $(\mathbf{x}_i, \mathbf{x}_j)$ of institutions feature vectors. More formally, we define the pairwise distance d_{ij} as:

$$d_{ij} = (\mathbf{x}_i - \mathbf{x}_j) \Delta^{-1} (\mathbf{x}_i - \mathbf{x}_j)' \quad (7)$$

where Δ is a diagonal matrix whose i -th diagonal element represent the standard deviation of the series. Namely, each coordinate difference between pairs of vectors $(\mathbf{x}_i - \mathbf{x}_j)$ is scaled by dividing by the corresponding element of the standard deviation. The distances can be embedded into a $N \times N$ dissimilarity matrix \mathbf{D} such that the closer the companies i, j features are in the Euclidean space, the lower the entry d_{ij} .

Although \mathbf{D} can be informative about the distribution of the distances between the companies, the fully-connected nature of this set does not help to find out whether there exist dominant patterns of similarities between institutions. Therefore, to extract such patterns we derive the Minimal Spanning Tree (MST) representation of borrowing companies' balance-sheet similarities (see Mantegna and Stanley, 1999; Bonanno et al., 2003; Spelta and Araújo, 2012).

2.4. The Minimal Spanning Tree

To find out the MST representation of the system, we perform hierarchical clustering by applying the nearest neighbor method. At the initial step, we consider N clusters corresponding to the N institutions. Then, at each subsequent step, two clusters l_i and l_j are merged into a single cluster if:

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

with the distance between clusters being defined as:

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

with $r \in l_i$ and $q \in l_j$. These operations are repeated until a single cluster emerges. This clustering process is also known as the single link method since one obtains the MST of a network. Given a connected graph, the corresponding MST is a tree of $N - 1$ edges that provides the minimum value of the sum of

the edge distances. More specifically, the hierarchical clustering procedure takes $N - 1$ steps to be completed when the graph is composed by N nodes, and it exploits, at each step, a particular distance $d_{ij} \in \mathbf{D}$ to merge two clusters into a single one.

In order to extract relevant information from the topology of the network for discriminating between borrowing companies, we compute different measures from complex network theory. In particular, the research in network theory has dedicated a huge effort to developing measures of interconnectedness, related to the detection of the most important player in a network. Moreover, beside investigating the importance each institution has in the network, we are also interested in assessing whether the network is characterized by a community structure and to exploit such feature. This topological characteristic indicates the presence of sets of companies usually defined as very dense sub-graphs, with few connections between them.

2.5. Network Measures

Various measures of centrality have been proposed in network theory such as the count of neighbors of a node has, i.e., the degree centrality, or measures based on the spectral properties of the graph (see Perra and Fortunato, 2008). These measures are feedback, also know as global, centrality measures and provide information on the position of each node relative to all other nodes. For our purposes we employ both families of centrality measures. In particular, for each node we compute the degree and strength centrality. The degree k_i of a vertex i with $(i = 1, \dots, N)$ is the number of edges incident to it. More formally, let the binary representation of the network be $\hat{\mathbf{D}}$ such that:

$$\hat{\mathbf{D}}_{ij} = \begin{cases} 1 & \text{if } d_{ij} > 0 \\ 0 & \text{otherwise} \end{cases}$$

then, the degree a vertex i is:

$$k_i = \sum_{j=1}^N \hat{\mathbf{D}}_{ij}. \quad (8)$$

Similarly, the strength centrality measures the average distance of a node with respect to its neighbors. Formally the strength of vertex i is:

$$s_i = \sum_{j=1}^N d_{ij}. \quad (9)$$

Moreover, since several studies have found the presence of sets of very dense sub-graphs, with few connections between them, as a result of similar patterns at the micro-level (see Pecora et al., 2016; Spelta et al., 2018), we also apply the Louvain Method to extract the community structure of the network (see Blondel et al., 2008). The identified communities maximize system's modularity, a measure that quantifies the strength of the division of the system into communities of densely interconnected nodes that are only sparsely connected with the rest of the system (see Newman, 2006). The modularity of our system is:

$$Q = \frac{1}{2m} \sum_{i,j} [D_{ij} - \frac{s_i s_j}{2m}] \delta(c_i, c_j) \quad (10)$$

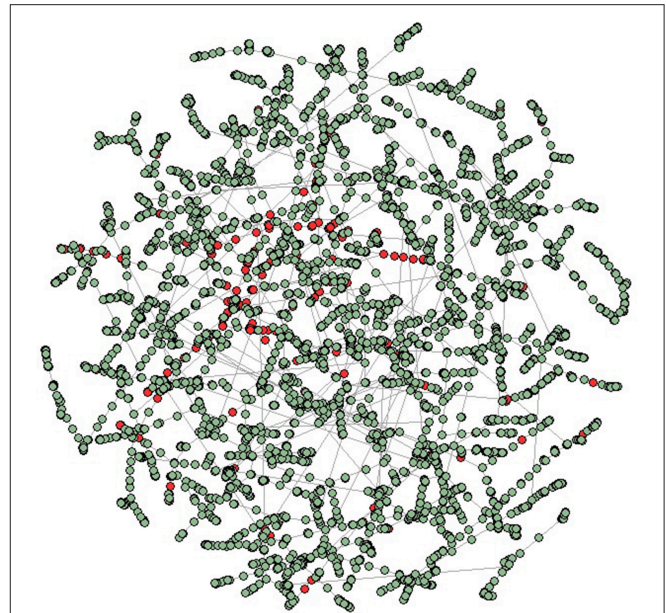


FIGURE 1 | Minimal spanning tree representation of the borrowing companies networks. The tree has been obtained by using the standardized Euclidean distance between institutions features and the Kruskal algorithm. In the panel, nodes are colored according to their financial soundness, red nodes represent defaulted institutions while green nodes are associated with active companies.

where d_{ij} is the weight of the edge between nodes i and j , s_i is the sum of the weights of the edges attached to node i , c_i is the community to which node i belongs, $\delta(u, v)$ is equal to 1 when $u = v$ and zero otherwise, and $m = \frac{1}{2} \sum_{i,j} D_{ij}$. The final step of our model specification is to embed the obtained centrality measures as well as information on the community structure of the network, into a predictive model. We propose to extend Chinazzi and Reyes, who incorporate network measures in a linear regression model, to the credit scoring context (i.e., logistic regression, linear discriminant analysis, CART, and SVM).

3. RESULTS

This section is devoted to show the results of the analysis. First, we report the MST representation of the similarity network obtained from companies' feature distances. We show nodes colored according to their financial soundness, red nodes represent defaulted institutions while green nodes represent sound and active companies, see **Figure 1**. Notice how, defaulted institutions occupy precise portion of the network, namely, such companies belong to the leafs of the tree and form clusters. This, in other words, suggests those companies form communities.

Information concerning the community structure of the networks and the centrality measures are used to provide synthetic topological variables at the node level. Such variables are embedded into the credit scoring models to assess whether they contain relevant information useful for forecasting institutions default.

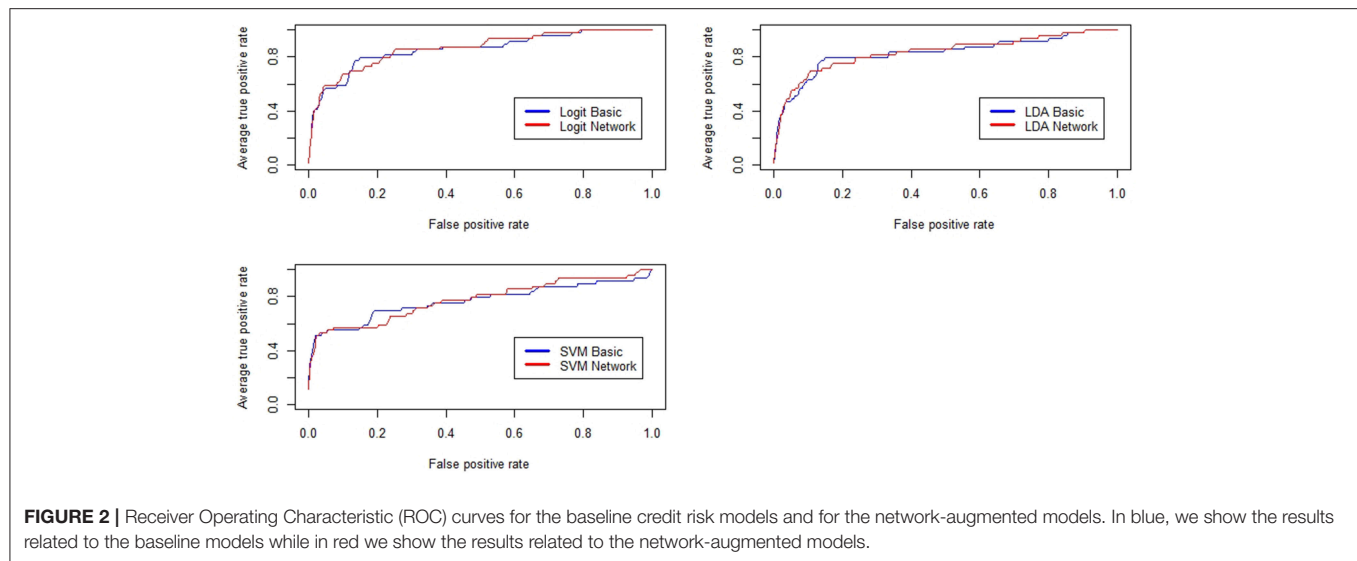


TABLE 2 | Summary Statistics of non-parametric analysis.

	AUC		KS		Gini		Accuracy	
	Basic	Network	Basic	Network	Basic	Network	Basic	Network
Logit	79.631	80.793	52	52	59.262	61.586	90.193	90.09661
LDA	77.759	79.16	51	52.8	55.518	58.32	90.122	89.98844
CART	67.973	67.973	35.5	35.946	35.946	35.5	90.832	90.82413
SVM	76.81	77.65	53.62	50	51	55.3	92.44444	92.22222

Summary statistics of the non-parametric analysis. From the left to the right: area under the ROC curve (AUC), KS Statistic (KS), Gini Index (Gini), Model accuracy (Accuracy), and area under the Precision Curve (AUCPR). For each measure and for all the tested models we report the results obtained by the baseline scenario and for the network-augmented configurations.

Figure 2 reports the results related to the performance of some of the models tested in the paper. Basically, the upper left panel shows the results from the logistic regression, the upper right panel encompasses the same information from the discriminant analysis while the bottom panel refers to the performance curves of the SVM classifier.

For sake of comparison, we have reported several measures of predictive utility so to show that, overall, the inclusion of topological information regarding similarity patterns among companies feature, increases the forecasting performance of various credit scoring models even when the data sets are imbalanced between the two classes (defaulted vs. active). Notice how, for most of the cases, red lines representing the performance of the models feeded with network measures lie above the blue lines representing baseline classifiers. Considering that graphically the improvements might not be fully visible, performance improvements for all the tested models are also reported in **Table 2**. The table summarizes the values of the measures employed to assess the predictive gain of the network-augmented credit scoring models. We report, the area under the ROC curve (AUC), the KS statistic, the Gini Index and the overall model accuracy (ACC).

From the results collected in **Table 2**, it is clear that the inclusion of topological variables describing institutions centrality in the similarity networks and the community structure composing such networks increases the predictive performance

of the methods used for credit scoring even if the forecasting gain obtained differ from model to model. In particular, we observe an increase of the predictive utility values for the logistic regression, the linear discriminate analysis and the SVM classifier once network parameters are added to the specification. Concerning the overall models accuracy, the ACC measure is less sensitive to the inclusion of topological variables with values between the baseline and network-augmented methods remaining quite similar across all models. Even though the increases in predictive utility across models are not very large, it might make significant difference for P2P lending platforms. Furthermore, we also notice that the predictive utility of the CART model does not change with the inclusion of the community and network parameters in the models specification. Future research may concern dealing with unbalanced samples (as in Calabrese and Giudici, 2015) and/or with multiple data sources (as in Figini and Giudici, 2011).

4. CONCLUSION

FinTech services, such as peer-to-peer lending platforms, are becoming part of the everyday life. Such new technologies can increase financial inclusion, but they can bring the cost of an increase credit risks. To cope with such risk, fintech risk management becomes a central point of interest for regulators and supervisors, to protect consumers and preserve financial

stability. In this work we have shown that topological information embedded into similarity networks can be exploited to increase the predictive performance of credit scoring models usually applied by P2P lending companies. Topological information are summarized computing centrality measures and community detection. The forecasting gain obtained by the inclusion of these variables has been then measured by employing non-parametric statistics. Standard performance measures such as ROC, precision recall and accuracy reveal the usefulness of the proposed methodology to build an early-warning signal suitable for both policy makers and supervisors as well as for practitioners.

DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the **Supplementary Files**.

AUTHOR CONTRIBUTIONS

It is the result of a joint work between the three authors in which, however, PG supervised the work and provided

the necessary research framework. BH-M wrote sections Introduction, Credit Risk Models, Assessing Model Performance and Results. AS wrote sections The Distance Metric, The Minimal Spanning Tree, Network Measures, and Conclusion.

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). 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.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frai.2019.00003/full#supplementary-material>

REFERENCES

- Abellán, J., and Castellano, J. G. (2017). A comparative study on base classifiers in ensemble methods for credit scoring. *Expert Sys. Appl.* 73, 1–10. doi: 10.1016/j.eswa.2016.12.020
- Allen, F., and Gale, D. (2000). Financial contagion. *J. Political Econ.* 108, 1–33. doi: 10.1086/262109
- Anderson, R. (2007). *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation*. Oxford: Oxford University Press.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech. Theor. Exp.* 2008:P10008. doi: 10.1088/1742-5468/2008/10/P10008
- 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
- Calabrese, R., and Giudici, P. (2015). Estimating bank default with generalised extreme value regression models. *J. Operat. Res. Soc.* 66, 1783–1792. doi: 10.1057/jors.2014.106
- Claessens, S., Frost, J., Turner, G., and Zhu, F. (2018). Fintech credit markets around the world: size, drivers and policy issues. *BIS Q. Rev.* 29–49.
- Emekter, R., Tu, Y., Jirasakuldech, B., and Lu, M. (2015). Evaluating credit risk and loan performance in online peer-to-peer (p2p) lending. *Appl. Econ.* 47, 54–70. doi: 10.1080/00036846.2014.962222
- Figini, S., and Giudici, P. (2011). Statistical merging of rating models. *J. Operat. Res. Soc.* 62, 1067–1074. doi: 10.1057/jors.2010.41
- Giudici, P., and Misheva, B. H. (2018). P2p lending scoring models: do they predict default? *J. Digit. Bank.* 2, 353–368.
- Giudici, P., and Spelta, A. (2016). Graphical network models for international financial flows. *J. Business Econ. Stat.* 34, 128–138. doi: 10.1080/07350015.2015.1017643
- Guegan, D., and Hassani, B. (2017). *Regulatory Learning: Credit Scoring Application of Machine Learning*. Technical report, HAL.
- Khandani, A. E., Kim, A. J., and Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *J. Bank. Finance* 34, 2767–2787. doi: 10.1016/j.jbankfin.2010.06.001
- Khashman, A. (2011). Credit risk evaluation using neural networks: Emotional versus conventional models. *Appl. Soft Comput.* 11, 5477–5484. doi: 10.1016/j.asoc.2011.05.011
- Leitner, Y. (2005). Financial networks: contagion, commitment, and private sector bailouts. *J. Finance* 60, 2925–2953. doi: 10.1111/j.1540-6261.2005.00821.x
- Lessmann, S., Baesens, B., Seow, H.-V., and Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *Eur. J. Oper. Res.* 247, 124–136. doi: 10.1016/j.ejor.2015.05.030
- Mantegna, R. N., and Stanley, H. E. (1999). *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge: Cambridge University Press.
- Newman, M. (2018). *Networks*. Oxford: Oxford University Press.
- Newman, M. E. (2006). Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8577–8582. doi: 10.1073/pnas.0601602103
- Pecora, N., Kaltwasser, P. R., and Spelta, A. (2016). Discovering sifs in interbank communities. *PLoS ONE* 11:e0167781. doi: 10.1371/journal.pone.0167781
- Perra, N., and Fortunato, S. (2008). Spectral centrality measures in complex networks. *Phys. Rev. E* 78:036107. doi: 10.1103/PhysRevE.78.036107
- Serrano-Cinca, C., and Gutiérrez-Nieto, B. (2016). The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (p2p) lending. *Decis. Support Syst.* 89, 113–122. doi: 10.1016/j.dss.2016.06.014
- 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
- Spelta, A., Flori, A., and Pammolli, F. (2018). Investment communities: behavioral attitudes and economic dynamics. *Soc. Netw.* 55, 170–188. doi: 10.1016/j.socnet.2018.07.004
- Yu, L., Yue, W., Wang, S., and Lai, K. K. (2010). Support vector machine based multiagent ensemble learning for credit risk evaluation. *Expert Syst. Appl.* 37, 1351–1360. doi: 10.1016/j.eswa.2009.06.083

Conflict of Interest Statement: 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 Giudici, Hadji-Misheva and Spelta. 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.

APPENDIX

A. FINANCIAL RATIOS

Since the data set is composed of ratios between financial and balance-sheet statements here we report the formulas employed to compute such ratios.

TABLE A1 | Description of variables included in the dataset.

ID	FORMULA	Type	ID	FORMULA	Type
RATIO001	(Total assets - Shareholders Funds)/Shareholders Funds	Continuous	RATIO019	Interest paid/(Profit before taxes + Interest paid)	Continuous
RATIO002	(Long term debt + Loans)/Shareholders Funds	Continuous	RATIO027	EBITDA/interest paid	Continuous
RATIO003	Total assets/Total liabilities	Continuous	RATIO029	EBITDA/Operating revenues	Continuous
RATIO004	Current assets/Current liabilities	Continuous	RATIO030	EBITDA/Sales	Continuous
RATIO005	(Current assets - Current assets: stocks)/Current liabilities	Continuous	RATIO036	Constraint EBIT	Dichotomous
RATIO006	(Shareholders Funds + Non current liabilities)/Fixed assets	Continuous	RATIO037	Constraint PL before tax	Dichotomous
RATIO008	EBIT/interest paid	Continuous	RATIO039	Constraint Financial PL	Dichotomous
RATIO011	(Profit (loss) before tax + Interest paid)/Total assets	Continuous	RATIO040	Constraint P/L for period th EUR	Dichotomous
RATIO012	P/L after tax/Shareholders Funds	Continuous	DPO	Trade Payables/Operating revenues	Continuous
RATIO013	GROSS PROFIT/Operating revenues	Continuous	DSO	Trade Receivables/Operating revenues	Continuous
RATIO017	Operating revenues/Total assets	Continuous	DIO	Inventories/Operating revenues	Continuous
RATIO018	Sales/Total assets	Continuous	NACE	Industry classification on NACE code, 4 digits precision	Dichotomous



An Artificial Intelligence Approach to Regulating Systemic Risk

Sharyn O'Halloran^{1*} and Nikolai Nowaczyk^{2†}

¹ Columbia University, New York, NY, United States, ² School of International and Public Affairs, Department of Political Science, Quaternon Risk Management, Dublin, Ireland

OPEN ACCESS

Edited by:

Peter Schwendner,
Zurich University of Applied Sciences,
Switzerland

Reviewed by:

Arianna Agosto,
University of Pavia, Italy
Paolo Pagnottoni,
University of Pavia, Italy

*Correspondence:

Sharyn O'Halloran
so33@columbia.edu

[†]These authors have contributed
equally to this work

Specialty section:

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

Received: 06 February 2019

Accepted: 03 May 2019

Published: 29 May 2019

Citation:

O'Halloran S and Nowaczyk N (2019)
An Artificial Intelligence Approach to
Regulating Systemic Risk.
Front. Artif. Intell. 2:7.
doi: 10.3389/frai.2019.00007

We apply an artificial intelligence approach to simulate the impact of financial market regulations on systemic risk—a topic vigorously discussed since the financial crash of 2007–09. Experts often disagree on the efficacy of these regulations to avert another market collapse, such as the collateralization of interbank (counterparty) derivatives trades to mitigate systemic risk. A limiting factor is the availability of proprietary bank trading data. Even if this hurdle could be overcome, however, analyses would still be hampered by segmented financial markets where banks trade under different regulatory systems. We therefore adapt a simulation technology, combining advances in graph theoretic models and machine learning to randomly generate entire financial systems derived from realistic distributions of bank trading data. We then compute counterparty credit risk under various scenarios to evaluate and predict the impact of financial regulations at all levels—from a single trade to individual banks to systemic risk. We find that under various stress testing scenarios collateralization reduces the costs of resolving a financial system, yet it does not change the distribution of those costs and can have adverse effects on individual participants in extreme situations. Moreover, the concentration of credit risk does not necessarily correlate monotonically with systemic risk. While the analysis focuses on counterparty credit risk, the method generalizes to other risks and metrics in a straightforward manner.

Keywords: artificial intelligence, graph theoretic models, data science, machine learning, stochastic Linear Gauss-Markov model, financial risk analytics, systemic risk, financial regulation

1. FRONTIERS OF ARTIFICIAL INTELLIGENCE

Predicting the next financial crisis is like forecasting the weather, a plethora of variables must converge at just the right moment in just the right way, invariably, leading experts to arrive at wildly conflicting prognostications. Advances in artificial intelligence (AI) methodologies have enhanced the robustness of such predictive models by introducing schemes based on skeletonization that extract vertices and edges from an initial graph and algorithms that prune unlikely outcomes by sifting through hundreds of thousands of factors to match shapes to known prototypes¹.

Artificial intelligence, which incorporates machine learning and data science, places data within a context through pattern recognition and iterative learning. What is new about the latest incarnation of the AI framework is that it draws on many disciplines, such as statistics and computer science, but also biology, psychology, and game theory, and employs a myriad of techniques, including:

¹See Kamani et al. (2018) for an application of this technique to forecast severe climate events.

1. Rules-based systems that set parameters and conditions to enable scenario testing;
2. Machine learning that applies algorithms to decipher patterns and linkages in the data by continuously updating “learning” through an iterative process;
3. Neural networks that identify interconnected nodes through multi-layered data to derive meaning;
4. Deep learning that leverages pools of high-dimensional data to identify patterns of patterns; and
5. Pattern recognition that uses tools, such as natural language processing to classify and interpret data.

What does this methodology tell us about predicting financial disasters or, even more importantly, how to avoid them? The turmoil following the 2008 collapse of Lehman Brothers, gave rise to a lively debate on how to regulate financial markets. Governments have imposed a number of regulations to reduce systemic risk or the possibility that an adverse event at a single financial institution could trigger severe instability or the collapse of an entire industry or economy. To mitigate effects of cascading defaults, for instance, regulators introduced the collateralization of derivative trades and incentivized dealers to clear trades on centralized exchanges as opposed to over-the-counter.

The financial crisis not only called into question the soundness of such regulations, but also the process to evaluate the efficacy of new regulations being put into place. Although a decade has passed, regulators and industry participants alike failed to arrive at a consensus on: (1) Have the regulations implemented post-crisis reduced systemic risk? (2) How can we predict the impact of a financial regulation before it is implemented? and (3) How can we evaluate which regulation is best to avert yet another “Financial Katrina?” As many governments once again face pressure to rollback far reaching financial legislation, it is necessary to know which regulations promote safety and soundness of the financial system and which add undue burdens on markets.

In this paper, we analyze credit exposures created by contracts among financial institutions that arise when one party defaults or fails to repay the contracted amount, or counterparty credit risk. We develop a graph model that characterizes a financial system as a network, similar to skeletal representations in meteorology, where the nodes of the graph represent a bank and the vertices represent credit relations, each with various weights. We introduce an analytic tool that simulates a financial system based on real case trade data. Through an iterative process, we evaluate, predict and optimize the amount of collateralization required to mitigate counterparty credit risk at the trade, bank and systemic level.

The analysis shows that collateralization reduces the costs of resolving risk in a financial system, yet it does not change the distribution of those costs among banks and can have adverse effects on individual participants in extreme situations. Consistent with the work of Battiston et al. (2012a,b) we also find that diversification is not sufficient to ward against systemic financial failures; indeed, it may exacerbate it. The analysis measures the impact of collateralization on counterparty credit risk exposure in the derivatives market, but the method

generalizes to other types of risks and metrics in a straightforward manner. The approach developed enables regulators and industry participants alike to conduct iterative scenario testing and thereby provides a unique opportunity to make informed decisions about the impact of public policy before the next crisis strikes.

2. MODELS IN CRISIS: A NEW APPROACH

The 2008 financial crisis was the perfect storm of failures: Wall Street, regulators, hedge funds, all played a part. Government's response has been to introduce a number of new regulations to improve the safety and soundness of the banking system as well as mitigate systemic risk. These include: capital buffers, leverage requirements and restrictions on derivatives. This has taken place at both domestic and global levels.

The question is, given all these regulations are we better off now than before? In particular, is the financial system more transparent and accountable than prior to the crisis? After all it was the oblique, complex derivatives that exasperated the mortgage crisis and almost brought down the international system in the first place.

The financial industry's response to these regulations has been to build black box risk models developed, for the most part, in institutional silos. The implication is that financial firms currently conduct risk exposure analysis absent shared standard models to use as benchmarks and validate results.

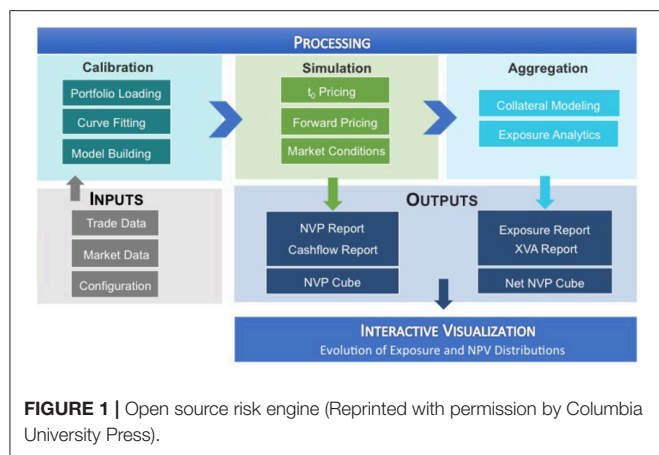
Yet, regulations require transparency and flexibility, and these requirements cannot be met by traditional silo-ed approaches. In response, collaborative efforts among academia, industry, and government have formed. Even the banks have come together in a previously unheard of data consortium, AcadiaSoft.

This reorganization has been accompanied by paradigm shifts from proprietary, homegrown software to open source. Even in financial risk management open source solutions, such as ORE, see Open Source Risk Engine (2016), have emerged. This trend has facilitated the use of AI technologies in the solution space, including: machine learning, natural language processing, AI and neural networks, provide powerful tools to augment risk analysis. In addition, these technologies provide new ways of developing models.

2.1. Open Source Risk Engine (ORE)

ORE computes the risks in a derivative portfolio from the perspective of a single bank. Schematically, it works as follows, see also **Figure 1**: It consumes trade data, market data and some configuration files as inputs, identifies all risk factors of the trade portfolio and performs a MonteCarlo simulation. This allows the computation of risk analytics at portfolio, asset class, and counterparty levels. See Lichters et al. (2015); Open Source Risk Engine User Guide (2017).

These analytics provide a benchmark that can be shared by regulators and industry participants to calibrate models around risk tolerance. As the assumptions are commonly known, it enables conversations around why and how various models deviate from the standard benchmarks.



2.2. A Systemic Risk Engine

One can aggregate firm specific risk metrics produced by the ORE into a systemic risk engine to assess the impact that regulations have on the financial system as a whole. This requires that the analysis takes into account not only the impact that financial transactions have on a financial institution but also the impact that each institution has on the system. Netting these input and output effects provide a more realistic picture of the impact of a regulation on the risks in the financial system. Moreover, adopting graph modeling enables visualization, calculation and testing of the robustness of various hypotheses under alternative parameter assumptions. More technical details on the technology stack used in the simulation can be found in Anfuso et al. (2017); O'Halloran et al. (2017b).

2.3. Columbia Data Science Institute FinTech Lab

The Columbia FinTech Lab housed in the Data Science Institute provides an easily accessible demonstration of how these tools can produce risk analytic measures. The Fintech Lab website, see Columbia University Fintech Lab (2018), provides a graphic display and interface that demonstrates how such analysis can be conducted.

3. USE CASE: SYSTEMIC FINANCIAL RISK

ORE has been built to compute the risks in a derivatives portfolio from the perspective of a single bank with purposes of serving as a bank risk management system or validating such a system. Its applications have an interesting pivot, however. Because, the computations of those risks from the perspective of one bank requires the above mentioned inputs, market data, trade data, netting agreements and other simulation parameters, one can use ORE to compute systemic risk, by running the computation from the perspective of all banks in a system.

The results include all risks of all banks in a financial system. As the same models are used for each bank, the resulting risk metrics are consistent and comparable across all banks. Those metrics can be computed under different regulatory regimes, allowing a consistent evaluation of the impact of financial regulation on systemic risk.

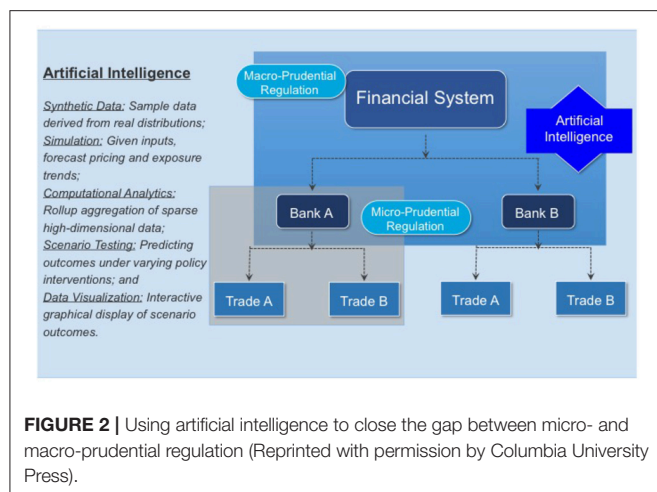
In practice, performing such a computation is difficult as one crucial input, the trade data of all the banks in the system, is proprietary and thus inaccessible. However, if the purpose of such a computation is to evaluate the impact of a financial regulation in general or to guide regulatory decision making bodies, it is, in fact, undesirable for the outcome to depend overly on current trade data. Trading activity in the global financial system is significant. Millions of transactions change the trade portfolios of the market participants every day, even every second. Changes in financial regulation, however, happen over a period of decades. The regulations around Initial Margin, for instance, a direct reaction to the financial crisis in 2007–2008, are still not fully implemented and will not be implemented fully before the early 2020s. Given the different time scales for changes in trade portfolios and changes in financial regulation, it would be an undesirable feature of financial regulation if its impact strongly depended on current trade data as this would signal overfitting of regulation to the current market.

Ideally, financial regulation should have the desired impact and that impact should be largely invariant under trading activity. Consequently, the evaluation of a regulation should be largely independent of changes in trade data. The precise trade data of the current financial system, therefore, should not be needed to evaluate the impact of a regulation. What is needed to study the impact of a regulation on a financial system is simply trade data, preferably as realistic as possible, but not necessarily the live deals of the current dealer banks. Our approach is to use a simulation technology. We randomly generate entire financial systems, including trade data, and calibrate those random generators to realistic distributions. The result is a representative sample of possible financial systems, which is transparent and completely accessible on all levels, from a single trade to the entire system.

3.1. Literature Review of Systemic Risk Metrics

This simulation approach has the advantage of bridging the gap that traditionally separates micro- and macro-prudential regulation, see Figure 2. The micro-prudential side considers a single bank in all its complexity and is primarily interested in the risks this bank is exposed to as a result of the trades in its portfolio. The metrics in which those risks are measured are standardized and their use is enforced globally by regulators. Examples include Value-at-Risk (VaR) to measure market risk, Effectivized Expected Positive Exposure (EEPE) for credit risk, Liquidity Coverage Ratio (LCR) for liquidity risk or a Basel-II traffic light test for model risk. Even though the concrete value of a metric like EEPE can differ between two banks that use internal models, the regulatory framework around internal models is designed to minimize those differences and the method, at least, is consistent. The only drawback of the micro-prudential view is that it considers only one bank in isolation making it difficult to study systemic risk.

In contrast, macro-prudential regulation considers an entire financial system with all its banks, but evaluates each and every bank from a high level perspective only. From a macro-prudential view, the amount of risk a bank is exposed to is of less interest than the amount of risk a bank induces into the financial system.



In particular, the question on whether or not a bank default could result in the default of the system is of particular importance (“too big to fail”). An excellent overview is provided by Battiston and Martinez-Jaramillo (2018) of the relationship between micro-prudential policies, which focus on individual exposures and leverage and capital ratios, and macro-prudential network-based policies.

In sharp contrast to the micro-prudential risk metrics, there is no clear definition on what systemic risk precisely means nor how it should be measured. In Bisias et al. (2012), the U.S. Office for Financial Research discusses 31 different metrics of systemic risk². A closer look at these metrics, however, reveals that these are not simply different mathematical functions measuring the same quantity, but different underlying notions of systemic financial risk. Most of these metrics focus on the analysis of market data, such as housing prices or government bonds and their correlations. For instance, Billio et al. (2012) use Principal Components Analysis (PCA) and Granger Causality to study the correlations between the returns of banks, asset managers and insurance. Unfortunately, most of those macro-prudential metrics are unsuited to guiding decision making bodies or regulatory interventions—precisely because their micro-prudential nature remains unclear (with CoVaR, which relies on a quantile of correlated asset losses, being a notable exception; see Adrian and Brunnermeier, 2016).

More recently, Sedunov (2016) compares the performance of three institution-level systemic risk exposures to forecast the financial crisis, including Exposure CoVaR, Granger causality, and Systemic Expected Shortfall. Using data from the 25 largest U.S. banks, insurers, and brokers, the analysis shows that CoVaR is the measure that best forecasts the within-crisis performance of financial institutions over multiple crisis periods. By contrast, neither Granger causality nor expected shortfall metrics predict within crisis performance. A key indicator in forecasting crisis exposures is the size of the financial institution.

²Similarly, in a meta analysis of the literature on systemic financial risk, Silva et al. (2017) find that from a sample of 266 articles published from 1990 to 2016, 134 articles directly addressed measures or indices of systemic risk.

3.2. AI: Bridging the Gap Between Micro- and Macro-prudential Regulation

As Figure 2 demonstrates micro-prudential regulation is directed toward the safety and soundness of an individual bank. Financial crises, however, result from the external actions of a bank, which may or may not be correlated with its compliance with regulatory standards. A lessons of the 2007–09 crisis is that macro-prudential regulation focused only on the risks taken by individual banks is insufficient to prevent crises.

An AI framework provides a way to bridge this gap. First, synthetic data of a financial system can be derived by sampling data from real market, portfolio and bank trades. Second, given these inputs, simulations can be constructed to forecast pricing and exposure trends. Computational analytics provide models for prediction and accuracy testing of sparse, high dimensional data. Scenario testing enables comparisons of different policy interventions on market outcomes. Finally, graphical visualization based on pattern recognition facilitates classifying outcomes.

3.3. Weighted Degree Metrics

This 2-fold divergence in metrics—the gap between micro- and macro- prudential regulation and the different notions of systemic risk—is unfortunate from a methodological point of view. The various notions of systemic risk are a consequence of the fact that this is a relatively new field and that the financial system and hence systemic risks are very complex and have many different facets. The gap between micro- and macro-prudential regulation has historic origins: The obvious approach of studying the macro-prudential impact of a regulation on an entire financial system as an aggregation of all its micro-prudential impacts has failed in the past due to the complexities of both levels.

In recent years there have been tremendous technological advances in handling big and highly complex data sets. Therefore, our approach is to use the standardized micro-prudential risk metrics and aggregate them in a graph model of systemic risk.

The advantages of this methodology are manifold. First, of the 266 papers reviewed by Silva et al. (2017), the analysis shows that only 20 articles used a combination of computational, simulation, and mathematical modeling. AI techniques enable iterative hypothesis testing to decipher patterns and linkages in the data, thereby providing more robust models and estimates of systemic risk. Second, Battiston and Martinez-Jaramillo (2018) note that existing research addresses systemic risk from either a micro-prudential or a macro-prudential level, absent any analysis of how link the two. By contrast, we derive a systemic risk metric from the ground up. The total risk exposure in the financial system is an aggregate estimate of individual firms’ credit risk exposure, thereby providing an indicator of how much risk a firm generates and how much it absorbs. And third, as documented by Silva et al. (2017), network analysis (Battiston et al., 2012a,b), cascade models (Capponi and Chen, 2015), and even examinations of the topological structure of inter-bank networks (Caccioli et al., 2015) are readily adopted constructs to evaluate contagion effects among financial institutions. Here, we

employ the mathematics of graph models to analyze the credit risk in financial systems.

4. GRAPH MODEL OF SYSTEMIC RISK

The trade data in a financial system is naturally organized in an undirected *trade relation graph* $G = (B, T)$: The nodes B represent the banks and the links T represent the trade relations. The graph is undirected because a trade relation is symmetric—a deal is only a done deal if both sides sign it. For formal details on graph models, see Erdős and Rényi (1959, 1960); Bales and Johnson (2006). An example of a trade relation graph is shown in **Figure 3**, where six banks (labeled A–F here) are trading bilaterally with each other in five trade relations. Any additional data on the trade portfolios can be attached to the links, for instance as a list of trade ids. The details of the trades are then stored in a database. This model serves both as a representation of a financial system and as a data format for the random generation of financial system, c.f. section 5.2. Optionally, one can also attach more information on the nodes in that graph, for instance a bank's core capital ratio.

Each trade in a trade relation imposes various types of risks (as well as rewards) on potentially both banks and these risks can be computed in various metrics by means of mathematical finance. By computing a fixed set of risk metrics for all trade relations in a trade relation graph, we obtain a risk graph that captures the risks between all the various banks in the system, see **Figure 4** for the example. Formally, the risk graph $RG = (B, A, w)$ is computed out of the trade relation graph as follows: The risk graph has the exact same nodes B as the trade relation graph, but each undirected trade relation $t \in T$ is replaced by two directed arrows $a_1, a_2 \in A$ representing the risks the bank at the tail induces onto the bank on the head and vice versa as a consequence of their trade relation. Finally, we attach a (possibly multivariate) weight function $w(a)$ onto the arrows $a \in A$ that quantify the risks. An example we will use later is EEPE (Effectivized Expected Positive

Exposure) to measure credit risk³. Another example could be the PFE (Potential Future Exposure) over a certain time horizon at a fixed quantile (analogous to US stress testing). Notice that the amount of risk that is induced by a bank b_1 onto a bank b_2 may or may not be the same as the amount of risk induced from b_2 onto b_1 even though both are in the same trade relation. For example, the loss an issuer of an FX option might suffer as a result of the buyer defaulting is at most zero, while the buyer can in theory suffer a unlimited losses. Notice that this use of a directed graph to model exposures in a financial system is consistent with (Detering et al., 2016), who use this to study default contagion.

The weight functions, that is, the risk metrics, can be computed using ORE. The resulting data produces a weight $w(a)$ for each arrow $a \in A$ in a risk graph. This provides a complete picture of risk in the financial system modeled by the trade relation graph in established micro-prudential risk metrics. We then aggregate this data by a purely graph theoretic construction from the arrows of the risk graph to the nodes and then further to a systemic level as follows: For each bank $b \in B$, we compute the weighted in/out-degree

$$w^-(b) := \sum_{\substack{a \in A \\ a \text{ ends at } b}} w(a), \quad w^+(b) := \sum_{\substack{a \in A \\ a \text{ starts at } b}} w(a). \quad (1)$$

The in-degree $w^-(b)$ represents the total amount of risk the bank b is exposed to from the system and thus corresponds to the micro-prudential view of b . The out-degree $w^+(b)$ represents the total amount of risk the bank b induces into the system and thus corresponds to the macro-prudential view of b . Therefore, this graph theoretic construction bridges the gap between the micro- and the macro-prudential by providing a coherent metric of both in the same model. In the example shown in **Figure 4**,

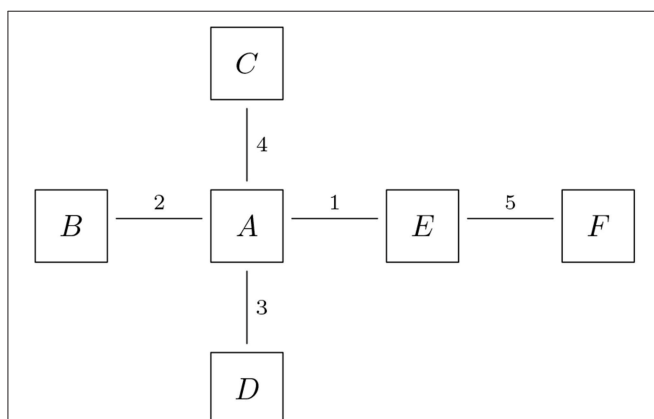


FIGURE 3 | Trade relations: the nodes represent the banks, the links represent the trade relations and the labels on the links represent the trade or portfolio IDs (Reprinted with permission by Columbia University Press).

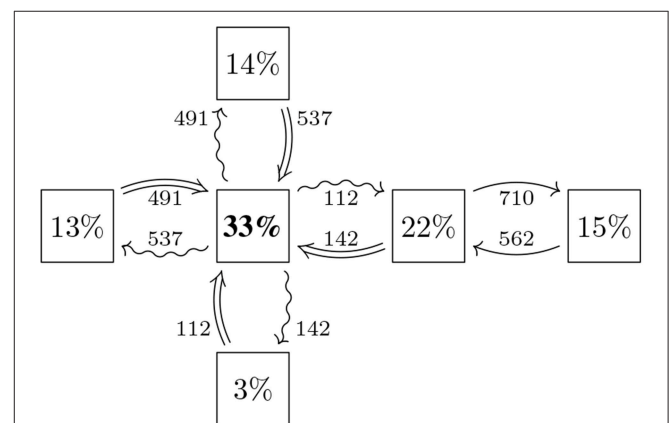


FIGURE 4 | Exposures: the nodes represent the same banks as in **Figure 3**, the arrows represent that risk is induced from the bank on the tail onto the bank on the head, the weights on the arrows quantify that risk and the percentages in the nodes represent the share of risk induced by that bank (Reprinted with permission by Columbia University Press).

³This is a regulatory standard metric to measure exposure. Notice that the exposure is a key ingredient in the calculation of capital requirements. Thus, a reduction in exposure automatically causes a reduction in capital requirements.

the in-degree of the big bank A in the middle is $w^-(A) = 537 + 142 + 112 + 491 = 1282$ and the relevant arrows going into A are highlighted as \Rightarrow . The out-degree is $w^+(A) = 491 + 112 + 142 + 537 = 1282$ and the outgoing arrows are highlighted as \sim .

In a second step, we aggregate the risk metrics to a system wide level by computing $w(G) := \sum_{a \in A} w(a)$ the total weight in the system. It is instructive to express the weighted in- and out-degree as a percentage of that total, i.e., to compute

$$\rho^-(b) := \frac{w^-(b)}{w(G)}, \quad \rho^+(b) := \frac{w^+(b)}{w(G)}, \quad (2)$$

a relative version of the weighted in/out-degree. In the example shown in **Figure 4**, the total amount of risk in the system in $w(G) = 3,836$ and e.g., counterparty A has $\rho^+(A) = w^+(A)/w(G) = 1,282/3,836 = 33\%$. Any of the quantities

$$w(G), \quad \max_{b \in B} w^+(b), \quad \max_{b \in B} \rho^+(b) \quad (3)$$

are (possibly \mathbb{R}^k valued) metrics that capture the total amount of weight in the graph and its concentration. These metrics serve as weighted degree metrics of systemic risk.

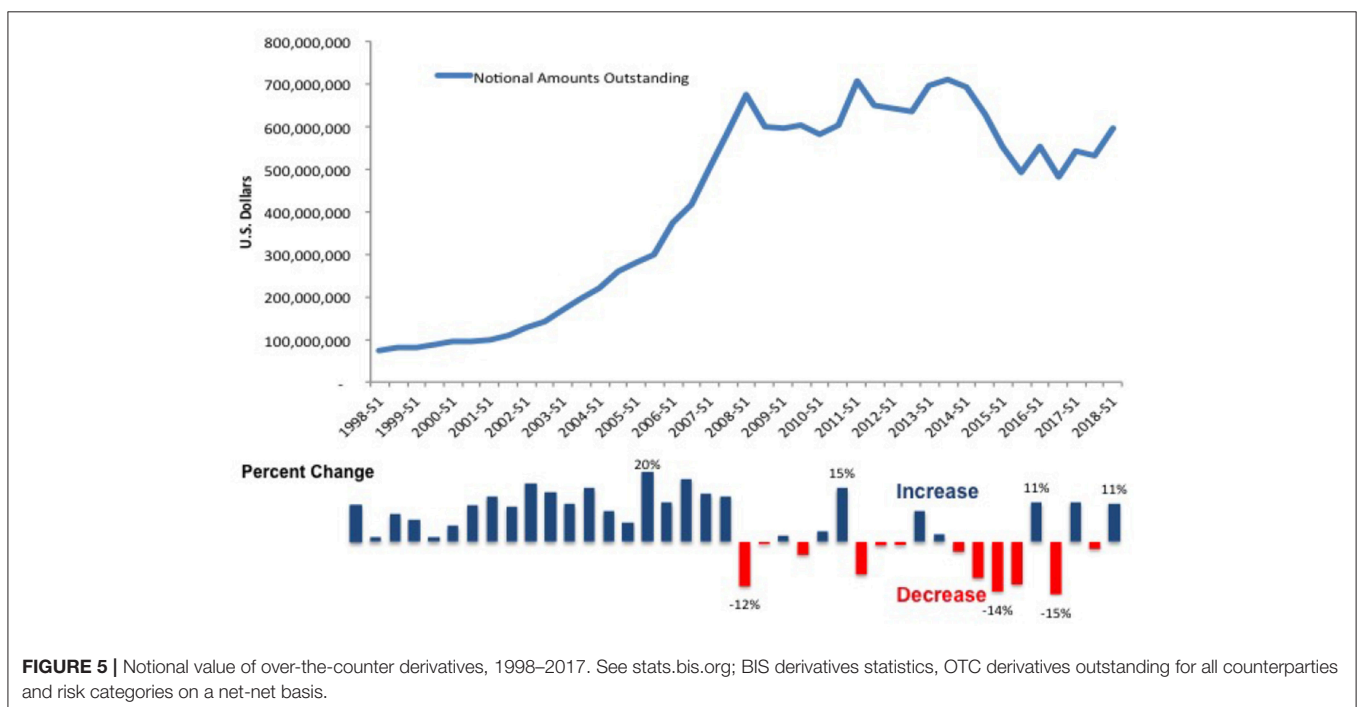
5. COLLATERALIZATION

The financial crisis exposed vividly the credit risk component in derivative contracts. Any two banks that enter into a derivative contract fix the terms and conditions of the contract at inception and both commit to payments according to the contract until it matures. While the rules on how to compute the payment amounts are fixed at inceptions, the payment amounts themselves are not as they depend on future market conditions.

In particular in the interest rate derivatives market that has an estimated total aggregated notional in the hundreds of trillions, the maturities of these contracts can be several decades. This exposes the two trading counterparties to each others credit risk: A payment in 10 years would simply not happen if one of the counterparties defaults in 9 years. As a derivative contract with a defaulted counterparty is worth zero, a default induces a significant shock to the value of a derivatives book of a bank.

Figure 5 shows the magnitude of the over-the-counter derivative market. The top part of the chart displays the notional amounts of outstanding derivatives in millions of U.S. dollars from 1998 to 2018. The data covers all derivative types, e.g., currency and interest rate swaps, for all risk types and all countries. The graph illustrates a steeply rising trend that peaks during the financial crisis, 2007–2009. The bottom half of the chart shows the increases and decreases in the trend line. The onset of the liquidity crisis in the U.S. and the sovereign debt crisis in Europe led to decreases in derivative trading activity. The subsequent introduction of new regulatory standards to force dealers to trade derivatives through central counterparties (CCPs) or exchanges precipitated sharp declines in notional amounts. By the end-June 2018, however, the notional value outstanding had once again reached 595 trillion USD, close to pre-crisis levels. The resumption of an upward trend suggests that despite new regulations to push more dealers onto central clearing platforms, banks continue to use non-standard derivative contracts.

Figure 6 compares OTC derivative gross market values and gross credit exposure from 1998 to 2018. The solid line shows the gross values, which measure a bank's total exposure to financial markets or the investment amount at risk. Once again, the trend peaks before the crisis and declines afterwards. This



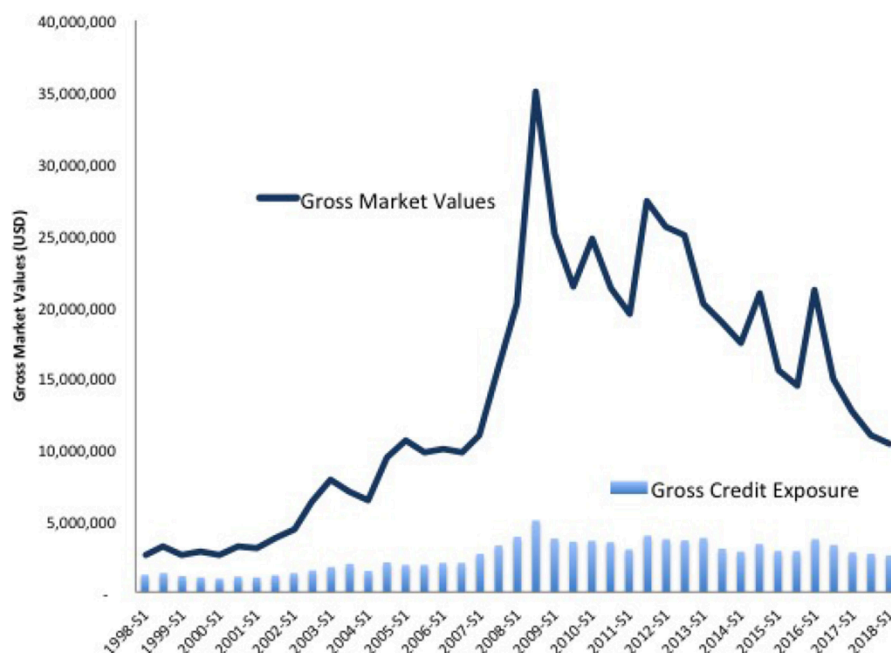


FIGURE 6 | Gross credit exposure and gross market values, 1998–2018. See stats.bis.org; BIS derivatives statistics, OTC derivatives outstanding for all counterparties and risk categories on a net-net basis.

time, however, the line continues its decent. For regulators, this indicates the success of stringent clearing and collateral requirements. By contrast, gross credit exposures, shown on the bottom of **Figure 6** by the light blue bar chart, tell a different story. Credit exposure is the total amount of credit made available to a borrower by a lender and calculates the extent to which a lender is exposed to the risk of loss in the event of the borrower's default. The chart shows that while market values have decreased, credit exposures have remained unchanged. In short, the credit risk resulting from a failure has not altered even as the total amount of market risk has declined. Moreover, the proportion of outstanding OTC derivatives that dealers cleared through CCPs held steady, at around 76 percent for interest rate derivatives and 54 percent for credit default swaps (CDS)⁴.

These data highlight that regulatory interventions may have unintended consequences. Adopting an AI framework—e.g., generating synthetic data from real bank distribution, simulating a financial system, and conduct scenario testing by introducing policy interventions and compare outcomes, may help avert implementing poorly tailored policies.

For example, a standard financial regulation to mitigate credit risk exposures is collateralization. That means that the two counterparties exchange collateral (typically in cash or liquid bonds) with each other during the lifetime of the trade. In a first step, counterparties exchange variation margin (VM) to cover the current exposure to daily changes in the value of a derivatives portfolio, sometimes subject to thresholds and minimum transfer amounts. This regulation is already fully phased in. In a second

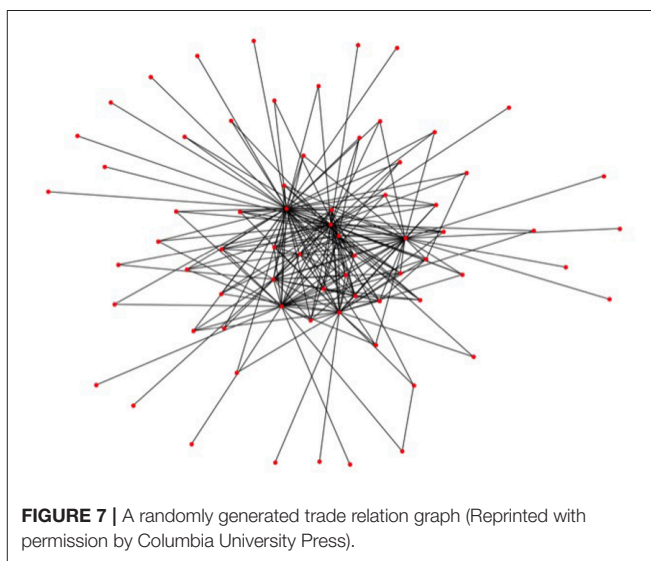


FIGURE 7 | A randomly generated trade relation graph (Reprinted with permission by Columbia University Press).

step, one can post initial margin (IM) to each other to cover for the potential exposure to close out risk after a default would occur. A more detailed description of these regulations can be found in (O'Halloran et al., 2017a, section 4); see also the Basel Committee on Banking Supervision (2015); Andersen et al. (2016, 2017); ISDA (2016); Caspers et al. (2017).

5.1. Collateralization Regimes

These collateralization regulations lead to four different regulatory regimes:

1. All derivative trades are uncollateralized.

⁴See Bank of International Settlements, Statistical release: OTC derivatives statistics at end- June 2018.

2. All derivative trades are VM collateralized, but some may only be partially collateralized due to thresholds and minimum transfer amounts.
3. All derivative trades are fully VM collateralized.
4. All derivative trades are fully VM collateralized and also fully IM collateralized.

For reasons of clarity, we exclude regime (2) from the present discussion. It is obvious that collateralization mitigates the exposure to credit risk on a micro-prudential level from the perspective of each counterparty⁵. We now test the hypothesis that collateralization also reduces systemic risk using the graph model from section 4 and simulated financial systems.

We consider regime (1) as our baseline scenario and will compute all relative impacts with reference to (1).

5.2. Simulation Technology

We use a systemic risk engine, see O'Halloran et al. (2017b), to compare the collateralization regimes described above. The engine generates trade relation graphs using the Python libraries `numpy.random` and `networkx` and then computes the risk metrics associated to all trades in all trade relations using an open source risk engine, see Open Source Risk Engine (2016). The resulting risk data is then aggregated using `pandas`. This process is repeated for each of the collateralization regimes such that their effect on the computed risk metrics can be systematically studied.

⁵In the language of section 4 this means that collateralization reduces the $w^{-(b)}$, i.e., the amount of risk bank is exposed to, where w is a credit risk metric (EEPE in our case).

5.3. Synthetic Data

The first step in the generation of the data is the generation of financial systems like **Figure 3**, where we want to calibrate the distributions of our random generator to realistic data. A statistical analysis of the macro exposures in the Brazilian banking system carried out in Cont et al. (2013) (based on central bank data) has shown that the degrees of the nodes in the trade relation graph, i.e., the number of links attached to each node, follow approximately a Pareto distribution. Therefore, we randomly generate Pareto distributed sequences and then compute a graph, which realizes that sequence. While the first step is straightforward, the second is a hard problem in discrete mathematics, which is still under active research. For the purposes of this paper, we use the so called erased configuration model as implemented in the Python library `networkx` and described in Newman (2003). Further details can also be found in Britton et al. (2006), Bayati et al. (2010). The resulting graphs look like **Figure 7**. We can see that the Pareto distributed node degree yields to graphs which have a few nodes with many links representing a few big banks, and many nodes with only one or a few links representing a large number of smaller firms in the system.

The trades in the trade relations are interest rate swaps (fixed vs. floating) and FX forwards in EUR and USD. Technically, these are implemented as boilerplate ORE XMLs and the trade parameters are chosen at random. For the FX forwards we use uniformly distributed maturities of up to 5Y, uniformly distributed notionals of between 100k and 100m and log-normally distributed strikes. For the interest rate swaps we use the same distributions for the notionals and the fixed rates are uniformly distributed between 0.01 and 5%. A coin flip decides whether or not a generated trade is an FX forward or an interest rate swap and the same applies to the long/short flag.

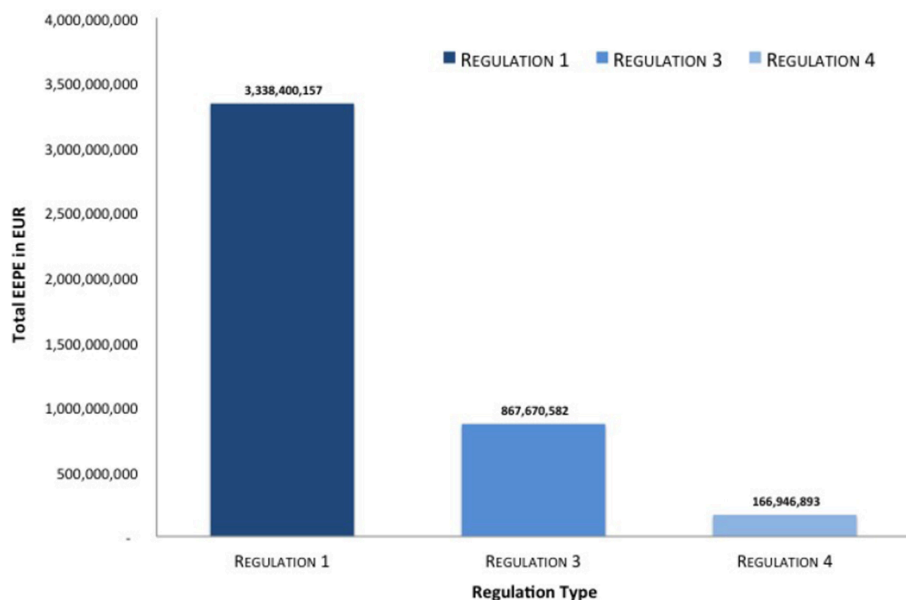


FIGURE 8 | Total reduction of EEPE (Reprinted with permission by Columbia University Press).

We run this simulation with parameters, which can be summarized as follows:

- Risk Type: Counterparty Credit Risk
- Risk Metric (choice of weight function w): EEPE (Effectivized Expected Positive Exposure)
- Asset classes: IR/FX Derivatives
- Number of financial systems: 10
- Number of banks in each system: ≤ 50
- Number of trades: 2,360
- Number of netting sets: 1,378
- Number of Monte Carlo paths: 500

5.4. Results

In **Figure 8** we see a highly aggregated overview of the results of the simulation. We can see that measured in average total levels of credit risk [i.e., $w(G)$] measured in $w = \text{EEPE}$ collateralization reduces this risk. The relative reduction between regime (1), that is the uncollateralized business, and regime (3), that is the fully VM collateralized business, is 74% and the relative reduction between regime (1) and (4), that is the fully VM and IM collateralized business is even 95%. Notice that this level of aggregation is even higher than in macro-prudential regulation as we aggregate across multiple financial systems representing possible future states of the world.

As all data is created during the simulation and thus completely accessible, we can now drill down to the macro-prudential view and study the impact of those regulations on an example system. In **Figures 9–11** we see the risk graph of a financial system under the three regulatory regimes. The size of the node indicates the amount of risk the bank at that node induces into the system, that is the $w^-(b)$. We see that collateralization significantly reduces risk in the entire system.

This optical impression can be confirmed by drilling down further to the micro-prudential view. In **Figure 12** we plot the $\text{EEPE}^+(b)$ for every bank b in the system. We can confirm that

the impact of collateralization on every bank is qualitatively the same as on the average, that is it reduces individual risk, but the amount of reduction can vary among the banks. It is interesting to note that the concentration of those risks, see **Figure 13**, i.e., the $\rho^+(b)$ stays mostly the same across the regulations and for banks, where it does change, it is not necessarily smaller. We conclude that collateralization has the desired effect of reducing total levels of risk of each counterparty, but is inadequate to address concentration risks.

We can now drill down even further than the micro-prudential level. As a byproduct of the simulation, we obtain exposure data of 1,378 netting sets, which we can mine to gain insight into all the micro impacts of the various regulations. In **Figure 14** we see the distribution of relative reductions in EEPE of the various netting sets when comparing REG_1 (uncollateralized) with REG_3 (VM collateralized). While most of the netting sets show a significant relative reduction in

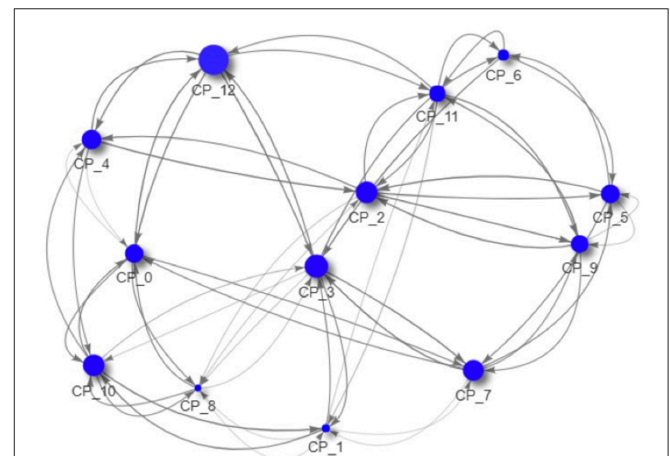


FIGURE 10 | Example of a financial system (VM collateralized) (Reprinted with permission by Columbia University Press).

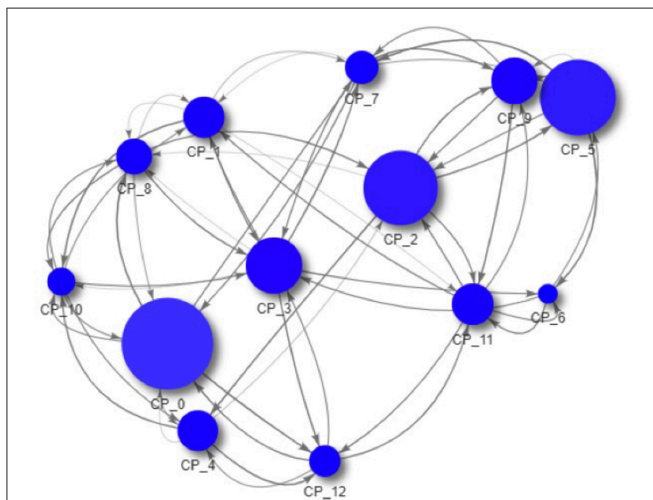


FIGURE 9 | Example of a financial system (uncollateralized) (Reprinted with permission by Columbia University Press).

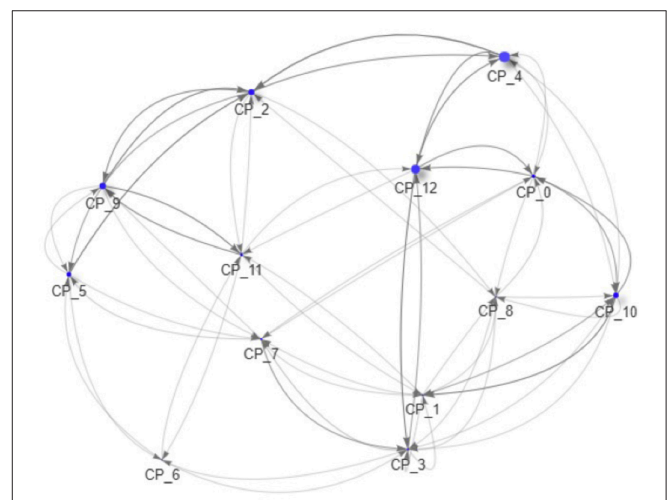


FIGURE 11 | Example of a financial system (VM & IM collateralized) (Reprinted with permission by Columbia University Press).

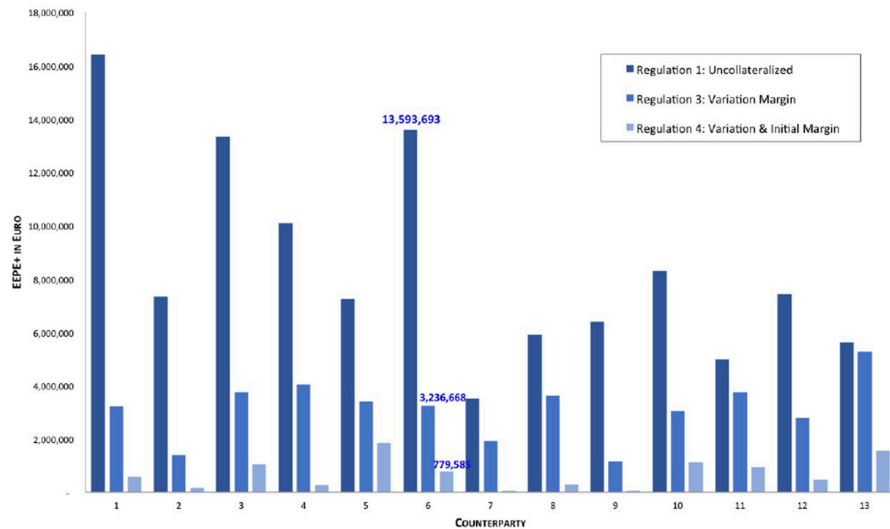


FIGURE 12 | Impact of collateralization on individual banks (EEPE+) (Reprinted with permission by Columbia University Press).

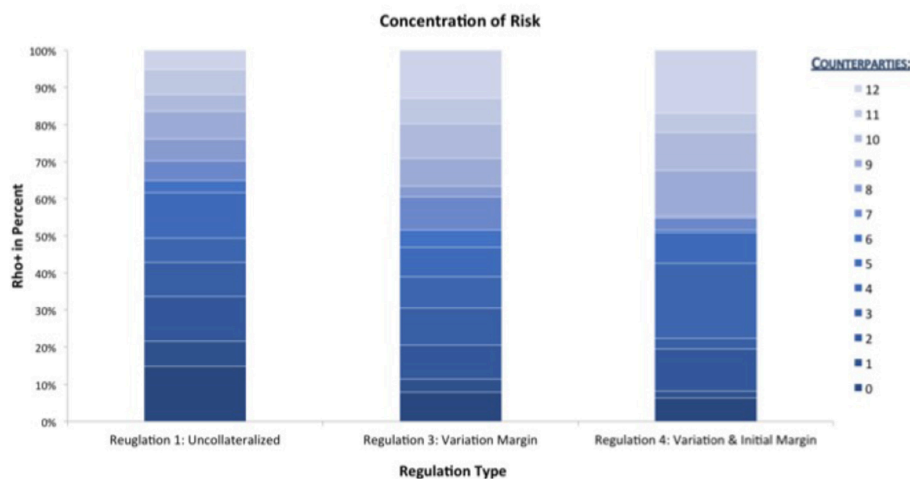


FIGURE 13 | Impact of collateralization on individual banks (ρ +) (Reprinted with permission by Columbia University Press).

exposure, we can see that some of them also show a significant relative increase in exposure. The explanation for this is as follows: Assume bank *A* has trades in a netting set with bank *B*. These trades are deeply out of the money for bank *A*, meaning the markets have moved into bank *B*'s favor. Then the uncollateralized exposure for bank *A* is very low⁶. Under VM collateralization however, as the trades are deeply in the money for bank *B*, bank *B* will call bank *A* for variation margin. Bank *A* will then pay the variation margin to bank *B*, where it is exposed to the default risk of *B*, because *B* might rehypothecate⁷ this variation margin. In some situations

this results in higher exposure under VM collateralization than under no collateralization. We see that on a micro level, VM collateralization can have an adverse effect in rare cases of netting sets, which are deeply out of the money.

Initial Margin cannot be rehypothecated and, therefore, posted Initial Margin is not treated as being at risk⁸. In **Figure 15** we see the relative reductions in EEPE of the various netting sets when comparing REG_3 (VM collateralization) vs. REG_4 (VM & IM collateralization). Here, we can see that the effect of the

⁶Due to the finite number of MonteCarlo paths, it is sometimes even numerically zero in the simulation.

⁷i.e., posting margin received from one counterparty to another.

⁸It should be highlighted that in our simulation we model the bilateral trading between various banks, where Initial Margin is posted into segregated accounts. Derivatives that are cleared through a central counterparty (CCP) or exchange traded derivatives (ETDs) are not in scope of this simulation.

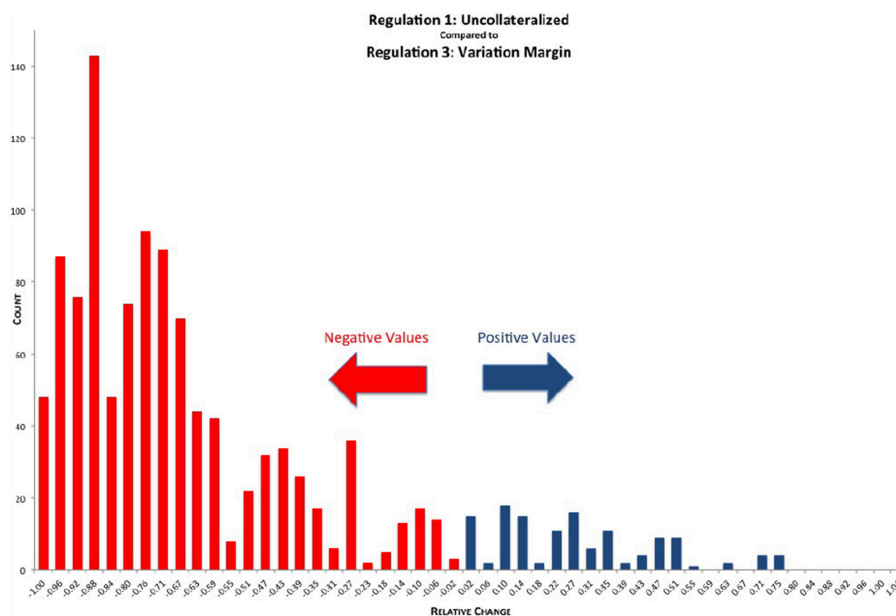


FIGURE 14 | Histogram of relative reduction in EEPE over all netting sets (197 out of 1,378 have more than 150% increase and are not shown, 29 of those have zero uncollateralized EEPE). Mean: -57.42% , SD: 38.33% (Reprinted with permission by Columbia University Press).

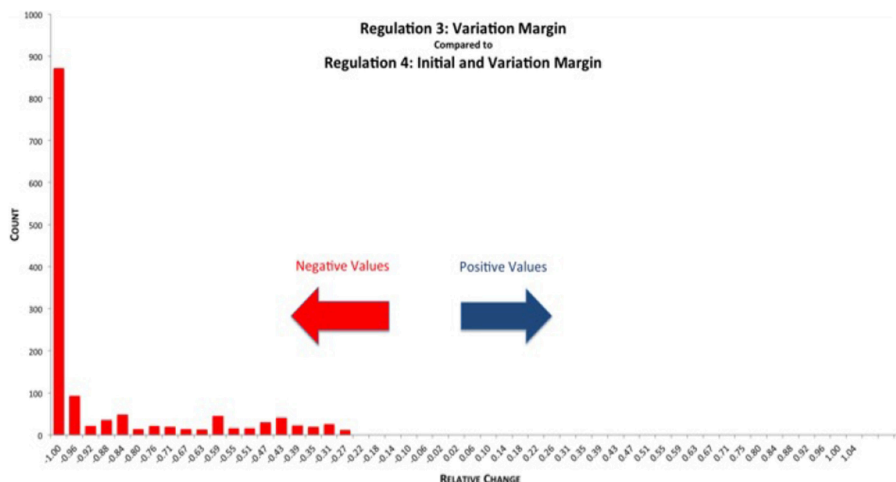


FIGURE 15 | Histogram of relative reduction in EEPE over all netting sets (0 out of 1,378 have more than 150% increase and are not shown, 0 of those have zero VM collateralized EEPE). Mean: -85.78% , SD: 20.76% (Reprinted with permission by Columbia University Press).

additional IM overcollateralization unambiguously reduces the exposure further.

When comparing REG_1 (uncollateralized) vs. REG_4 (VM & IM collateralization) directly, we can see in **Figure 16** that the reduction in exposure is larger and distributed more narrowly compared with just the VM collateralization, see **Figure 14**. There are still some netting sets left, which show an increase due to posted variation margin. However, this increase is smaller than under REG_3, as it is partially mitigated by the additional IM collateral.

It should be noted that while the increases in exposure we see in **Figures 14, 16** are large in relative terms, they are actually

quite small in absolute terms. In **Figure 17** we compute the total increases and decreases in EEPE of all the netting sets separately.

5.5. Summary

The directed weighted graph metrics provide a useful comparative statistics to evaluate the impact of various regulatory regimes on systemic risk. Applied to our hypothesis testing we arrive at the following conclusions:

- Collateralization reduces systemic credit risk significantly (measured in EEPE, i.e., the cost of resolving a failed system).

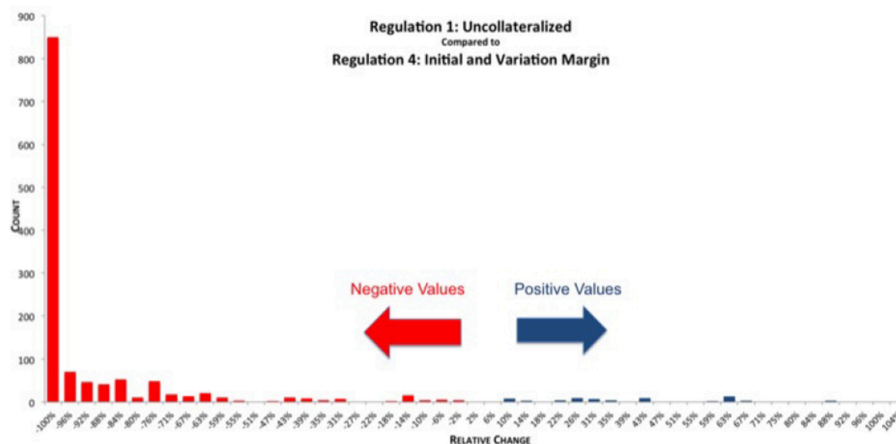


FIGURE 16 | Histogram of relative reduction in EEPE over all netting sets (69 out of 1,378 have more than 150% increase and are not shown, 29 of those have zero uncollateralized EEPE). Mean: -83.83% , SD: 34.44% (Reprinted with permission by Columbia University Press).

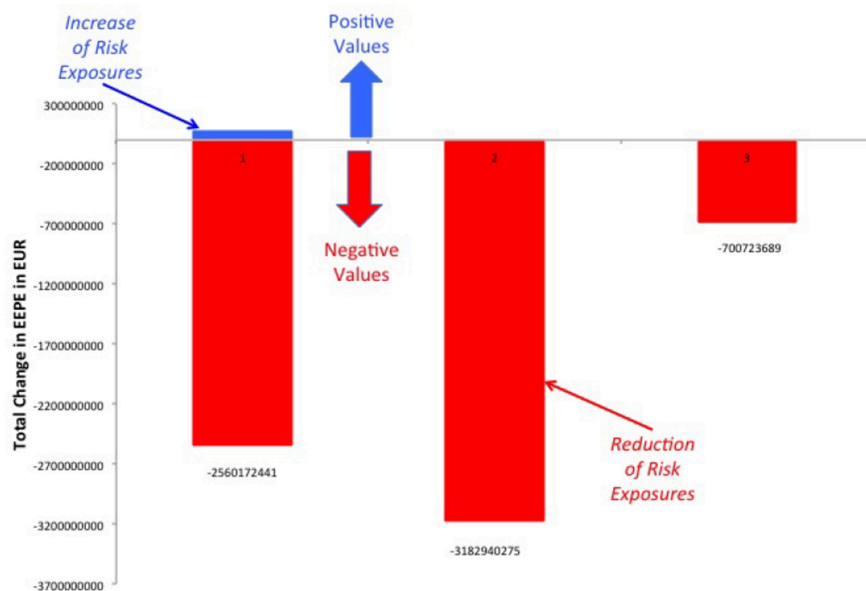


FIGURE 17 | Total increases and decreases in EEPE of all netting sets between the various regimes (Reprinted with permission by Columbia University Press).

- Collateralization does not materially change the concentration of credit risk in a financial system.
- In corner cases (deeply out of the money portfolios), VM collateralization can increase credit risk.

Notice that these results are an interplay of the aggregated macro-exposures and a systematic analysis of all micro-exposures, which would not be possible outside of the present framework.

6. CONCLUSION

Over the past two decades, the interconnected nature of global financial markets has increased dramatically, exacerbating threats to the financial system through the domino effect, the fire-sale

effect, and oversized role certain firms. Just like predicting the weather, financial service firms are now more interconnected and inherently more complex than ever before. The financial crisis highlighted the dangers of relying too heavily on proprietary models developed in silos. The open source paradigm introduced provides a means to benchmark models and to have common standards across the industry. The analytic approach adopted merges the structural and predictive properties of graph model and AI techniques to generate a financial system from real distributions of bank trading data.

Our analysis advances the literature in three ways:

1. Provides a simulation environment that enables iterative stress testing to decipher patterns and linkages in the data,

thereby providing more robust models and estimates of systemic risk;

2. Generates entire financial systems from a single trade to the entire trade relation graph, resulting in a graph model of systemic risk derived from random distributions of bank data; and
3. Aggregates standardized micro-prudential risk exposures into a macro-systemic risk metric that captures both the amount of risk a firm generates and the amount it absorbs.

We will expand the substantive analysis and methodological approach developed here in a number of directions:

Large scale simulation: The simulation performed to obtain the results in section V-D ran on a standard desktop computer. We plan to deploy the systemic risk engine in a cloud environment and run a large scale simulation to achieve an even higher statistical robustness.

Dependence on distributions of the trades: During the random generation of the trade relations, various distributional choices have to be made. It is interesting to study the dependence of the results on those choices. We expect them to be fairly stable under distributions.

Agent based creation of trade relation graphs: The current model assumes a Pareto distribution of the trade relations. It would be interesting to enhance the nodes representing the banks in the graph model with dynamic agent based rules of trading and study under what conditions the resulting trade relations are Pareto distributed.

Joint modeling of all major risk classes: In the present article we study the impact of collateralization on credit risk. However, regulation can affect all types of risk and the metrics used to measure it. We plan to conduct a joint analysis of market risk, credit risk, liquidity risk, operational risk and model risk. In fact, there is significant interplay between the various risks. **Figures 5, 6** highlight a paradox: the notional amount of OTC derivatives increased simultaneously as market risk decreased and credit risk

remained changed. One explanation is that collateralization may decrease market risk at the expense of increased liquidity risk. We can test this possibility with the AI framework detailed above⁹.

Initial Margin and Funding Costs: We believe that the key to understanding the interplay between credit risk and liquidity risk, in particular when studying the impact of collateralization, is its effect on funding costs and other value adjustments of derivative trades, the so called XVAs. These quantify the price of the reduction in risk.

Derivatives Market vs. Money Market: It is to be expected that collateralization will not only impact the derivatives market, but also the money market. As initial margin cannot be rehypothecated, its impact could be large. It is therefore interesting to study the interplay between those markets, both in case studies and simulations.

Central Clearing: The current analysis focuses on the study of the impact of collateralization on systemic risk as this was one of the major regulatory responses to the crisis. Another response was the incentivization of central clearing, which can be studied in a similar fashion. Notice that the graph model presented in section 4 is already able to capture the effect of this regulation: Any bilateral trade relation of a bank *A* with a bank *B* has to be replaced by two trade relations—one of bank *A* with the clearing house and another one for bank *B* with the clearing house. We expect to obtain results quantifying how much safer a clearing house needs to be in order to reduce systemic risk compared to bilateral trading.

AUTHOR CONTRIBUTIONS

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

⁹Another way the divergence between market and credit risk can occur is if banks resort to means outside the regulatory matrix. For example, banks can transact derivative contracts in tax havens, use instruments currently not covered by regulations or decide not to insure against the possibility of default altogether.

REFERENCES

- Adrian, T., and Brunnermeier, M. K. (2016). CoVaR. *Am. Econ. Rev.* 106, 1705–1741. doi: 10.1257/aer.20120555
- Andersen, L., Pykhtin, M., and Sokol, A. (2016). *Credit Exposure in the Presence of Initial Margin*. SSRN pre-print. Available online at: <https://ssrn.com/abstract=2806156>
- Andersen, L., Pykhtin, M., and Sokol, A. (2017). *Does Initial Margin Eliminate Counterparty Risk?* Risk Magazine.
- Anfuso, F., Aziz, D., Giltinan, P., and Loukopoulos, K. (2017). *A Sound Modelling and Backtesting Framework for Forecasting Initial Margin Requirements*. 86–91.
- Bales, M., and Johnson, S. (2006). Graph theoretic modeling of large-scale semantic networks. *J. Biomed. Informatics* 39, 451–464. doi: 10.1016/j.jbi.2005.10.007
- Basel Committee on Banking Supervision. (2015). *Margin Requirements for Non-centrally Cleared Derivatives*. BCBS 317. Available online at: <https://www.bis.org/bcbs/publ/d317.htm> (accessed February 3, 2019).
- Battiston, S., Gatti, D., Gallegati, M., Greenwald, B., and Stiglitz, J. (2012b). Default cascades: When does risk diversification increase stability? *J. Financ. Stabil.* 8, 138–149. doi: 10.1016/j.jfs.2012.01.002
- Battiston, S., Gatti, D. D., Gallegati, M., Greenwald, B., and Stiglitz, J. (2012a). Liaisons dangereuses: increasing connectivity, risk sharing, and systemic risk. *J. Dyn. Control* 36, 1121–1141. doi: 10.1016/j.jedc.2012.04.001
- Battiston, S., and Martinez-Jaramillo, S. (2018). Financial networks and stress testing: Challenges and new research avenues for systemic risk analysis and financial stability implications. *J. Financ. Stabil.* 35, 6–16. doi: 10.1016/j.jfs.2018.03.010

- Bayati, M., Kim, J., and Saberi, A. (2010). A sequential algorithm for generating random graphs. *Algorithmica* 58, 860–910. doi: 10.1007/s00453-009-9340-1
- Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *J. Financ. Econ.* 104, 535–559. doi: 10.1016/j.jfineco.2011.12.010
- Bisias, D., Flood, M., Lo, A. W., and Stavros, V. (2012). *A Survey of Systemic Risk Analytics*. Washington, DC: Office of Financial Research, Working Paper #0001.
- Britton, T., Deijfen, M., and Martin-Löf, A. (2006). Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.* 124, 1377–1397. doi: 10.1007/s10955-006-9168-x
- Caccioli, F., Farmer, J. D., Foti, N., and Rockmore, D. (2015). Overlapping portfolios, contagion, and financial stability. *J. Econ. Dyn. Control* 51, 50–63. doi: 10.1016/j.jedc.2014.09.041
- Capponi, A., and Chen, P. (2015). Systemic risk mitigation in financial networks. *J. Dyn. Control* 58, 152–166. doi: 10.1016/j.jedc.2015.06.008
- Caspers, P., Giltinan, P., Lichters, R., and Nowaczyk, N. (2017). Forecasting initial margin requirements—a model evaluation. *J. Risk Manage. Financ. Inst.* 10, 365–394. doi: 10.2139/ssrn.2911167
- Columbia University FinTech Lab. (2018). Available online at: <http://fintech.datascience.columbia.edu/>
- Cont, R., Moussa, A., and Santos, E. (2013). “Network structure and systemic risk in banking systems,” in *Handbook on Systemic Risk*, eds J. Fouque, and J. Langsam (Cambridge: Cambridge University Press), 327–368.
- Detering, N., Meyer-Brandis, T., Panagiotou, K., and Ritter, D. (2016). Managing default contagion in inhomogeneous financial networks. *arXiv[Preprint]. arXiv:1610.09542*.
- Erdős, P., and Rényi, A. (1959). On random graphs. *Publ. Math.* 6, 290–297.
- Erdős, P., and Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* 5, 17–61.
- ISDA. (2016). *ISDA SIMM (TM) Methodology*. Version R1.0, September 2016.
- Kamani, M. M., Farhat, F., Wistar, S., and Wang, J. Z. (2018). Skeleton matching with applications in severe weather detection. *Appl. Soft Comput.* 70, 1154–1166. doi: 10.1016/j.asoc.2017.05.037
- Lichters, R., Stamm, R., and Gallagher, D. (2015). *Modern Derivatives Pricing and Credit Exposure Analysis: Theory and Practice of CVA and XVA Pricing, Exposure Simulation and Backtesting (Applied Quantitative Finance)*. London: Palgrave Macmillan.
- Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Rev.* 45, 167–256. doi: 10.1137/S003614450342480
- O'Halloran, S., Nowaczyk, N., and Gallagher, D. (2017a). “Big data and graph theoretic models: simulating the impact of collateralization on a financial system,” in *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017, ASONAM '17* (New York, NY: ACM), 1056–1064. doi: 10.1145/3110025.3120989
- O'Halloran, S., Nowaczyk, N., and Gallagher, D. (2017b). “A data science approach to predict the impact of collateralization on systemic risk,” in *After the Crash: Financial Crises and Regulatory Responses*, eds S. O'Halloran, and T. Groll (New York, NY: Columbia University Press).
- Open Source Risk Engine User Guide. (2017). Available online at: <http://www.opensourcerisk.org/documentation/>
- Open Source Risk Engine. (2016). Available online at: www.opensourcerisk.org. First release in October.
- Sedunov, J. (2016). What is the systemic risk exposure of financial institutions? *J. Financ. Stabil.* 24, 71–87. doi: 10.1016/j.jfs.2016.04.005
- Silva, W., Kimura, H., and Sobreiro, V. (2017). An analysis of the literature on systemic financial risk: a survey. *J. Financ. Stabil.* 28, 91–114. doi: 10.1016/j.jfs.2016.12.004
- Conflict of Interest Statement:** 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 O'Halloran and Nowaczyk. 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.



Factorial Network Models to Improve P2P Credit Risk Management

Daniel Felix Ahelegbey^{1*}, Paolo Giudici² and Branka Hadji-Misheva³

¹ Department of Mathematics and Statistics, Boston University, Boston, MA, United States, ² Department of Economics and Management, University of Pavia, Pavia, Italy, ³ Zurich University of Applied Sciences (ZHAW), Zurich, Switzerland

This paper investigates how to improve statistical-based credit scoring of SMEs involved in P2P lending. The methodology discussed in the paper is a factor network-based segmentation for credit score modeling. The approach first constructs a network of SMEs where links emerge from comovement of latent factors, which allows us to segment the heterogeneous population into clusters. We then build a credit score model for each cluster via lasso-type regularization logistic regression. We compare our approach with the conventional logistic model by analyzing the credit score of over 1,5000 SMEs engaged in P2P lending services across Europe. The result reveals that credit risk modeling using our network-based segmentation achieves higher predictive performance than the conventional model.

OPEN ACCESS

Edited by:

Dror Y. Kenett,
Johns Hopkins University,
United States

Reviewed by:

J. D. Opdyke,
Allstate Insurance Company,
United States
Apama Gupta,
Rensselaer Polytechnic Institute,
United States

*Correspondence:

Daniel Felix Ahelegbey
dfkahey@bu.edu

Specialty section:

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

Received: 26 February 2019

Accepted: 13 May 2019

Published: 04 June 2019

Citation:

Ahelegbey DF, Giudici P and
Hadji-Misheva B (2019) Factorial
Network Models to Improve P2P
Credit Risk Management.
Front. Artif. Intell. 2:8.
doi: 10.3389/frai.2019.00008

Keywords: credit risk, factor models, FinTech, peer-to-peer lending, credit scoring, lasso, segmentation

1. INTRODUCTION

Issuance of loans by traditional financial institutions, such as banks, to other firms and individuals, is often associated with major risks. The failure of loan recipients to honor their obligation at the time of maturity leaves the banks vulnerable and affects their operations. The risk associated with such transactions is referred to as credit risk. It is well known that some percentage of these non-performing loans are eventually imputed to economic losses. To minimize such risk exposures, various methods have been extensively discussed in the credit risk literature to enable credit-issuing institutions to undertake a thorough assessment to classify loan applicants into risky and non-risky customers. Some of these methods range from logistic and linear probability models to decision trees, neural networks and support vector machines. A conventional individual-level reduced-form approach is the credit scoring model which attributes a score of credit-worthiness to each loan applicant based on the available history of their financial characteristics. See Altman (1968) for some pioneer works on corporate bankruptcy prediction models using accounting-based measures as variables. For a comprehensive review on credit scoring models, see Alam et al. (2010).

Recent advancements gradually transforming the traditional economic and financial system is the emergence of digital-based systems. Such systems present a paradigm shift from traditional infrastructural systems to technological (digital) systems. Financial technological (“FinTech”) companies are gradually gaining ground in major developed economies across the world. The emergence of Peer-to-Peer (P2P) platforms is a typical example of a FinTech system. The P2P platform aims at facilitating credit services by connecting individual lenders with individual borrowers without the interference of traditional banks as intermediaries. Such platform serves as a digital financial market and an alternative to the traditional physical financial market. P2P platforms significantly improve the customer experience and the speed of the service and reduce costs to both individual borrowers and lenders as well as small business owners. Despite

the various advantages, P2P systems inherit some of the challenges of traditional credit risk management. In addition, they are characterized by the asymmetry of information and by a strong interconnectedness among their users (see e.g., Giudici et al., 2019) that makes distinguishing healthy and risky credit applicants difficult, thus affecting credit issuers. There is, therefore, a need to explore methods that can help improve credit scoring of individual or companies that engage in P2P credit services.

This paper investigates how factor-network-based segmentation can be employed to improve the statistical-based credit score for small and medium enterprises (SMEs) involved in P2P lending. The approach is to first constructs a network of SMEs where links emerge from comovement of the latent factors that drive the observed financial characteristics. The network structure then allows us to segment the heterogeneous population into two sub-groups of connected and non-connected clusters. We then build a credit score model for each sub-population via lasso-type regularization logistic regression.

The contribution to the literature of this paper is manifold. Firstly, we extend the ideas contained in the factor network-based classification of Ahelegbey et al. (2019) to a more realistic setting, characterized by a large number of observations which, when links between them are the main object of analysis, becomes extremely challenging.

Secondly, we extend the network-based scoring model proposed in Giudici et al. (2019) to a setting characterized by a large number of explanatory variables. The variables are selected via lasso-type regularization (Tibshirani, 1996; Hastie et al., 2009) and, then, summarized by factor scores. Thus, we contribute to network-based models for credit risk quantification. Network models have been shown to be effective in gauging the vulnerabilities among financial institutions for risk transmission (see Battiston et al., 2012; Billio et al., 2012; Diebold and Yilmaz, 2014; Ahelegbey et al., 2016a), and a scheme to complement micro-prudential supervision with macro-prudential surveillance to ensure financial stability (see IMF, 2011; Moghadam and Viñals, 2010; Viñals et al., 2012). Recent application of networks have been shown to improve loan default predictions and capturing information that reflects underlying common features (see Letizia and Lillo, 2018; Ahelegbey et al., 2019).

Thirdly, our empirical application contributes to modeling credit risk in SMEs particularly engaged in P2P lending. For related works on P2P lending via logistic regression (see Andreeva et al., 2007; Barrios et al., 2014; Emekter et al., 2015; Serrano-Cinca and Gutiérrez-Nieto, 2016). We model the credit score of over 15,000 SMEs engaged in P2P credit services across Southern Europe. We compare the performance of our network-based segmentation credit score model (NS-CSM) with the conventional single credit score model (CSM). We show via our empirical results that our network-based segmentation presents a more efficient scheme that achieves higher performance than the conventional approach.

The paper is organized as follows. Section 2 presents the factor network segmentation methodology and the lasso-type regularization for credit scoring. Section 3 discusses the

empirical application of our segmentation approach against the conventional single model.

2. METHODOLOGY

We present the formulation and inference of a latent factor network to improve credit scoring and model estimation. Our objective is to analyze the characteristics of the borrowers to build a model that predicts the likelihood of their default.

2.1. Logistic Model

Let Y be a vector of independent observations of the loan status of n firms, such that $Y_i = 1$ if firm- i has defaulted on its loan obligation, and zero otherwise. Furthermore, let $X = \{X_{ij}\}$, $i = 1, \dots, n$, $j = 1, \dots, p$, be a matrix of n observations with p financial characteristic variables or predictors. The conventional parameterization of the conditional distribution of Y given X is the logistic model with log-odds ratio given by

$$\log\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + X_i\beta \quad (1)$$

where $\pi_i = P(Y_i = 1|X_i)$, β_0 is a constant term, $\beta = (\beta_1, \dots, \beta_p)'$ is a $p \times 1$ vector of coefficients and X_i is the i -th row of X .

2.2. Decomposition of Data Matrix by Factors

The dataset X can be considered as points of n -institutions in a p -dimensional space. It can also be interpreted at observed outcomes driven by some underlying firm characteristics. More specifically, X can be expressed as a factor model given by

$$X = FW + \varepsilon \quad (2)$$

where F is $n \times k$ matrix of latent factors, W is $p \times k$ matrix of factor loadings, ε is $n \times p$ matrix of errors uncorrelated with F . The error term ε is typically assumed to be multivariate normal but F in general case need not be multivariate normal (see Tabachnick et al., 2007). Lastly, $k < p$ is the number of factors required to summarize the pattern of correlations in the observed data matrix X . In the context of our application, we set k to be the number of factors that account for approximately 95% of the variation in X .

2.3. Factor Network-Based Segmentation

We present the construction of network structure for the segmentation of the population. Following the literature on graphical models (see Carvalho and West, 2007; Eichler, 2007; Ahelegbey et al., 2016a,b), we represent the network structure as an undirected binary matrix, $G \in \{0, 1\}^{n \times n}$, where G_{ij} represents the presence or absence of a link between nodes i and j . We construct G via similarity of the latent firm characteristics, such that $G_{ij} = 1$ if the latent coordinates of firm- i are strongly related to firm- j , and zero otherwise.

Given the latent factors matrix, F , we construct a network where the marginal probability of a link between nodes- i and j by

$$\gamma_{ij} = P(G_{ij} = 1|F) = \Phi[\theta + (FF')_{ij}] \quad (3)$$

where $\gamma_{ij} \in (0, 1)$, Φ is the standard normal cumulative density function, $\theta \in \mathbb{R}$ is a network density parameter, and $(FF')_{ij}$ is the i -th row and the j -th column of FF' . Under the assumption that G is undirected, it follows that $\gamma_{ij} = P(G_{ij} = 1|F) = P(G_{ji} = 1|F) = \gamma_{ji}$. We validate the link between nodes- i and j in G by

$$G_{ij} = \mathbf{1}(\gamma_{ij} > \gamma) \quad (4)$$

where $\mathbf{1}(\gamma_{ij} > \gamma)$ is the indicator function, i.e., unity if $\gamma_{ij} > \gamma$ and zero otherwise, and $\gamma \in (0, 1)$ is a threshold parameter. By definition, the parameters θ and γ control the density of G . Following Ahelegbey et al. (2019), we set $\theta = \Phi^{-1}(\frac{2}{n-1})$. To broaden the robustness of the results, we compare $\gamma = \{0.05, 0.1\}$ to capture a sparse but closely connected community.

2.4. Estimating High-Dimensional Logistic Models

When estimating high-dimensional logistic models with a relatively large number of predictors, there is the tendency to have redundant explanatory variables. Thus, to construct a predictable model, there is the need to select the subset of predictors that explains a large variation in the probability of defaults. Several variable selection methods have been discussed and applied for various regression models. In this paper, we consider variants of the lasso regularization for logistic regressions (Hastie et al., 2009).

2.4.1. Lasso

The lasso estimator (Tibshirani, 1996) solves a penalized log-likelihood function given by

$$\arg \min_{\beta} \sum_{i=1}^n \left[Y_i(\beta_0 + X_i\beta) - \log(1 + \exp(\beta_0 + X_i\beta)) \right] - \lambda \sum_{j=0}^p |\beta_j| \quad (5)$$

where n is the number of observations, p the number of predictors, and λ is the penalty term, such that large values of λ shrinks a large number of the coefficients toward zero.

2.4.2. Adaptive Lasso

The adaptive lasso estimator (Zou, 2006) is an extension of the lasso that solves

$$\arg \min_{\beta} \sum_{i=1}^n \left[Y_i(\beta_0 + X_i\beta) - \log(1 + \exp(\beta_0 + X_i\beta)) \right] - \lambda \sum_{j=0}^p w_j |\beta_j| \quad (6)$$

where w_j is a weight penalty such that $w_j = 1/|\hat{\beta}_j|^v$, with $\hat{\beta}_j$ as the ordinary least squares (or ridge regression) estimate and $v > 0$.

2.4.3. Elastic-Net

The elastic-net estimator (Zou and Hastie, 2005) solves the following

$$\arg \min_{\beta} \sum_{i=1}^n \left[Y_i(\beta_0 + X_i\beta) - \log(1 + \exp(\beta_0 + X_i\beta)) \right] - \lambda \sum_{j=0}^p (\alpha |\beta_j| + (1 - \alpha) \beta_j^2) \quad (7)$$

where $\alpha \in (0, 1)$ is an additional penalty such that when $\alpha = 1$ we a lasso estimator (L_1 penalty), and when $\alpha = 0$ a ridge estimator (L_2 penalty). For the elastic-net estimator, we set $\alpha = 0.5$ giving equal weight to the L_1 and L_2 regularization.

2.4.4. Adaptive Elastic-Net

The adaptive elastic-net estimator (Zou and Zhang, 2009) combines the additional penalties of the adaptive lasso and the

TABLE 1 | Description of the financial ratios with summary of mean statistics according to default status.

Var	Formula (description)	Active (mean)	Defaulted (mean)
V1	(Total Assets - Shareholders Funds)/Shareholders Funds	8.87	9.08
V2	(Longterm debt + Loans)/Shareholders Funds	1.25	1.32
V3	Total Assets/Total Liabilities	1.51	1.07
V4	Current Assets/Current Liabilities	1.6	1.06
V5	(Current Assets - Current assets: stocks)/Current Liabilities	1.24	0.79
V6	(Shareholders Funds + Non current liabilities)/Fixed Assets	8.07	5.99
V7	EBIT/Interest paid	26.39	-2.75
V8	(Profit (loss) before tax + Interest paid)/Total Assets	0.05	-0.13
V9	P/L after tax/Shareholders Funds	0.02	-0.73
V10	Operating Revenues/Total Assets	1.38	1.27
V11	Sales/Total Assets	1.34	1.25
V12	Interest Paid/(Profit before taxes + Interest Paid)	0.21	0.08
V13	EBITDA/Interest Paid	40.91	5.71
V14	EBITDA/Operating Revenues	0.08	-0.12
V15	EBITDA/Sales	0.09	-0.12
V16	Constraint EBIT	0.13	0.56
V17	Constraint PL before tax	0.16	0.61
V18	Constraint Financial PL	0.93	0.98
V19	Constraint P/L for period	0.19	0.64
V20	Trade Payables/Operating Revenues	100.3	139.30
V21	Trade Receivables/Operating Revenues	67.59	147.12
V22	Inventories/Operating Revenues	90.99	134.93
V23	Total Revenue	3557	2083
V24	Industry Classification on NACE code	4566	4624
Total number of institutions (%)		13413(89.15%)	1632(10.85%)

elastic-net to solve the following

$$\arg \min_{\beta} \sum_{i=1}^n \left[Y_i(\beta_0 + X_i\beta) - \log(1 + \exp(\beta_0 + X_i\beta)) \right] - \lambda \sum_{j=0}^p (\alpha w_j |\beta_j| + (1 - \alpha) \beta_j^2) \quad (8)$$

In the empirical work, we focus on estimating the credit score using the four lasso-type regularization methods. We select the regularization parameter using 10-fold cross-validation on a grid of λ values for the penalized logistic regression problem. Two λ 's are widely considered in the literature, i.e., λ_{\min} and λ_{1se} . The former is the value of the λ that minimizes the mean square cross-validated errors, while the latter is the λ value that corresponds to one standard error from the minimum mean square cross-validated errors. Our preliminary analysis shows that λ_{1se} produces a larger penalty that is too restrictive in the sense that we lose almost all the regressors. Although our goal is to encourage a sparse credit scoring model for the purpose of interpretability, we do not want to impose too much sparsity that renders the majority of the features insignificant. Thus, we rather choose λ_{\min} over λ_{1se} . For the additional penalty terms, we set $\alpha = 0.5$, $\nu = 2$, and $\hat{\beta}_j$ as the ridge regression estimate.

TABLE 2 | The eigenvalues of the singular value decomposition to determine the factors to retain.

No.	Eigenvalue	Variance explained (%)	Cumulative (%)
1	5.18	21.60	21.60
2	2.58	10.73	32.33
3	2.50	10.41	42.74
4	1.60	6.69	49.42
5	1.42	5.92	55.34
6	1.30	5.40	60.74
7	1.16	4.82	65.55
8	1.09	4.56	70.11
9	0.99	4.11	74.22
10	0.93	3.88	78.10
11	0.80	3.35	81.45
12	0.79	3.31	84.76
13	0.75	3.11	87.87
14	0.56	2.35	90.22
15	0.53	2.21	92.43
16	0.51	2.12	94.55
17	0.43	1.80	96.35
18	0.37	1.54	97.89
19	0.17	0.69	98.58
20	0.11	0.47	99.05
21	0.09	0.36	99.41
22	0.07	0.27	99.68
23	0.06	0.26	99.94
24	0.01	0.06	100.00

3. APPLICATION

3.1. Data: Description and Summary Statistics

To illustrate the effectiveness of the application of factor network methodology in credit scoring analysis, we obtained data from the European External Credit Assessment Institution (ECAI) on 15045 small-medium enterprises engaged in Peer-to-Peer lending on digital platforms across Southern Europe.

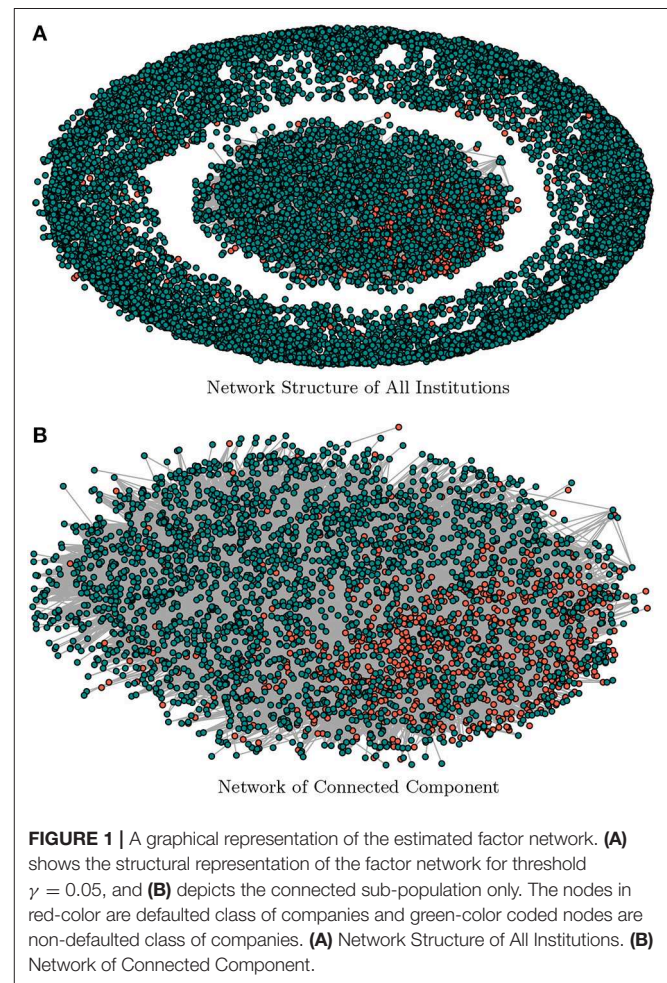


FIGURE 1 | A graphical representation of the estimated factor network. **(A)** shows the structural representation of the factor network for threshold $\gamma = 0.05$, and **(B)** depicts the connected sub-population only. The nodes in red-color are defaulted class of companies and green-color coded nodes are non-defaulted class of companies. **(A)** Network Structure of All Institutions. **(B)** Network of Connected Component.

TABLE 3 | Summary statistic of connected and non-connected sub-population obtained from the factor network-based segmentation for threshold values of $\gamma = \{0.05, 0.1\}$.

Threshold	Status	Conn-sub		Non-conn-sub	
$\gamma = 0.05$	Default	964	22.4%	668	6.2%
	Non-Default	3,341	77.6%	10,072	93.8%
	Total	4,305	28.6%	10,740	71.4%
$\gamma = 0.1$	Default	816	24%	816	7%
	Non-Default	2,580	76%	10,833	93%
	Total	3,396	22.6%	11,649	77.6%

TABLE 4 | Estimated coefficients from lasso (top left), adaptive lasso (top right), elastic-net (bottom left) and adaptive elastic-net (bottom right).

	CSM	NS-CSM(C)	NS-CSM(NC)	CSM	NS-CSM(C)	NS-CSM(NC)
	lasso			Adaptive lasso		
V1	0.0535	.	0.0375	.	.	.
V2	.	0.0332
V3	−0.4468	−0.2818	−1.0148	−0.5298	−0.3539	−1.1990
V4	−0.3549	−0.1294	−0.5556	−0.2928	−0.1368	−0.5137
V5
V6	0.0774	.	0.1460	0.0440	.	0.0213
V7	0.2818	.	.	0.2116	.	.
V8	−0.3933	−0.3408	0.1185	−0.4356	−0.3463	.
V9	−0.0360	0.0365	−0.4690	.	.	−0.5577
V10	−0.0701	0.0287
V11	0.1291	.	0.0550	.	.	.
V12	0.0265	0.0222	0.0204	.	.	.
V13	−0.2419	.	.	−0.1759	.	.
V14	−0.0399	−0.0776	.	.	−0.113	.
V15	−0.0751	−0.0396	0.0128	−0.0520	.	.
V16	0.0520	0.2851	.	.	0.2245	.
V17	0.2213	0.1650	0.1761	0.2529	0.2092	.
V18	0.0396	0.0661	0.0143	.	0.0484	.
V19	0.2540	0.0291	0.2096	0.2755	.	0.2151
V20	0.0412	.	0.2429	.	.	0.1950
V21	0.2212	0.1620	0.2969	0.2410	0.1721	0.3185
V22	0.0930	.	0.1470	0.0541	.	0.0219
V23	−0.2262	−0.0649	−0.3452	−0.2213	−0.0650	−0.3826
V24	−0.0062	−0.0641	0.0343	.	−0.0645	.
	Elastic-net			Adaptive elastic-net		
V1	0.0548	.	0.0568	.	.	.
V2	1.0e−04	0.0372
V3	−0.4472	−0.2692	−1.0132	−0.5293	−0.3538	−1.2208
V4	−0.3628	−0.1286	−0.6051	−0.2900	−0.1350	−0.6034
V5	0.0048	−0.0123
V6	0.0780	−0.0028	0.1862	0.0422	.	0.1528
V7	0.3003	.	.	0.1925	.	.
V8	−0.3926	−0.3310	0.2054	−0.4363	−0.3474	0.1672
V9	−0.0356	0.0435	−0.4884	.	.	−0.5195
V10	−0.1419	0.0315
V11	0.2016	0.0112	0.1025	.	.	.
V12	0.0299	0.0299	0.0545	.	.	.
V13	−0.2595	.	.	−0.1571	.	.
V14	−0.0374	−0.0785	.	.	−0.1112	.
V15	−0.0777	−0.0468	0.0597	−0.0499	.	.
V16	0.0600	0.2902	0.0669	.	0.2256	.
V17	0.2173	0.1588	0.1701	0.2527	0.2097	0.1147
V18	0.0417	0.0769	0.0439	.	0.0459	.
V19	0.2538	0.0502	0.2042	0.2747	.	0.2151
V20	0.0425	.	0.3139	.	.	0.2571
V21	0.2210	0.1634	0.3113	0.2409	0.1721	0.3036
V22	0.0933	0.0012	0.1727	0.0533	.	0.1047
V23	−0.2286	−0.0728	−0.3754	−0.2185	−0.0616	−0.4114
V24	−0.0077	−0.0724	0.0464	.	−0.0619	.

CSM is the benchmark credit score model, NS-CSM(C) is the network segmented connected sub-population credit score model, and NS-CSM(NC) is the network segmented non-connected sub-population credit score model, estimated for threshold value $\gamma = 0.1$.

The observation on each institution is composed of 24 financial characteristic ratios constructed from official financial information recorded in 2015. **Table 1** presents a description of the financial ratios with summary of mean statistics of the institutions grouped according to their default status. In all, the data consists of 1,632 (10.85%) defaulted institutions and 13,413 (89.15%) non-defaulted companies.

3.2. Decomposition of the Observed Data Matrix by Factors

To estimate the underlying factors that drive the observed data matrix, we decompose the matrix of observed financial characteristics via a singular value decomposition given by,

$$X = UDV = FW + \varepsilon \quad (9)$$

where U and V are orthonormal, and $D = \Lambda^{1/2}$ is a diagonal matrix of non-negative and decreasing singular values, with Λ as the diagonal matrix of the non-zero eigenvalues of $X'X$ and

XX' . U is $n \times p$, D is $p \times p$ and V is $p \times p$. Following the error approximation criteria, we obtain the factor matrix by, $F = U_{n,k} D_{k,k}$ and $W = V_{k,p}$, where $U_{n,k}$ is $n \times k$ matrix composed of the first k columns of U , $k < p$, $D_{k,k}$ is $k \times k$ matrix comprising the first k columns and rows of D , and $V_{k,p}$ is $k \times p$ matrix of factor loadings. The matrix F can therefore be interpreted as a projection of X onto the eigenspace spanned by $U_{n,k}$. We determine k by observing the number of eigenvalues associated with the largest variance matrix. **Table 2** shows the eigenvalues of the singular value decomposition to determine the factors to retain. The eigenvalues reported are the normalized squared diagonal terms of D . From the table, we set $k = 17$ since the first 17 eigenvalues explain about 95% of the total variation in X .

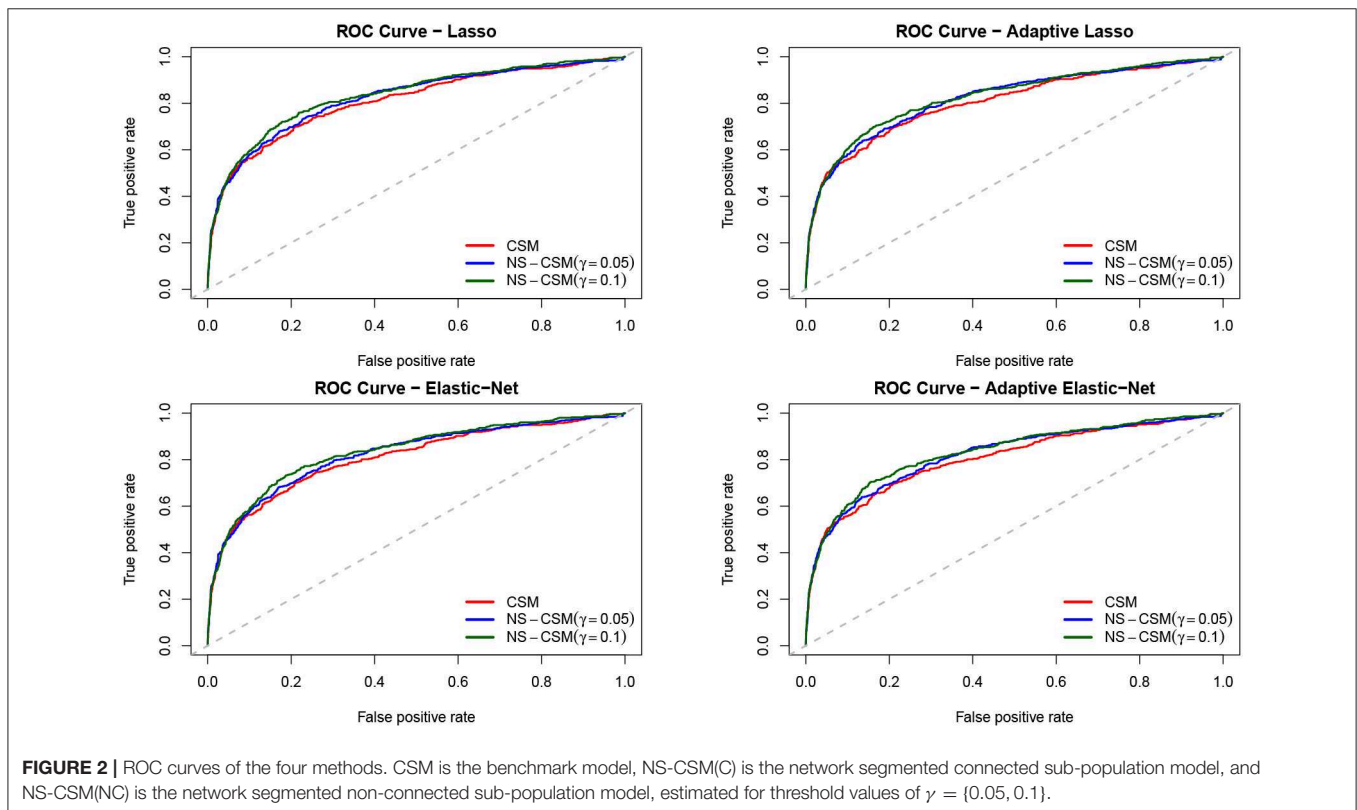
3.3. Factor Network Analysis

We use the estimated factor matrix, F , to construct the network for the segmentation of the companies. For purposes of graphical representations and to keep the companies name anonymous, we report the estimated network by representing the group of institutions with color-codes. The defaulted companies are represented in a red color code, and non-defaulted companies in the green color code (see **Figure 1**). **Table 3** reports the summary statistics of the estimated network in terms of the default-status composition of the SMEs. For robustness purposes, we compare the results obtained with a threshold value $\gamma = 0.05$ against $\gamma = 0.10$.

The result for the threshold $\gamma = 0.05$ of **Table 3** shows that the connected sub-population is composed of 4,305 companies

TABLE 5 | Number of selected variables of the four methods.

	Lasso	Adaptive lasso	Elastic-net	Adaptive elastic-net
CSM	22	12	24	12
NS-CSM(C)	16	10	20	10
NS-CSM(NC)	17	9	18	11



which constitute 28.6% of the full sample. The non-connected sub-population is composed of 10,740 (71.4%). The percentage of the defaulted class of companies are 22.4 and 6.2% among the connected- and non-connected sub-population, respectively. We notice that higher threshold values (say $\gamma = 0.1$) decrease (increase) the total number of connected (non-connected) sub-population and vice versa. Such higher threshold values also lead to a lower (higher) number of defaulted class of connected (non-connected) SMEs but (and) constituting a higher percentage of the defaulted population. **Figure 1** presents the graphical representation of the estimated factor network with the sub-population of defaulted and non-defaulted companies color coded as red and green, respectively. **Figure 1A** shows the structural representation of both connected and non-connected sub-population while **Figure 1B** depicts the structure of connected sub-population only.

3.4. Credit Score Modeling

We compare the lasso, adaptive lasso, elastic-net, and adaptive elastic-net variable selection methods to model the credit score of the listed companies in our dataset. To estimate the models, we standardized each series to a zero mean and unit variance. **Table 4** reports the variable selection and estimated coefficients of the four methods. The column CSM represents the benchmark credit scoring model, NS-CSM(C) - the network segmented connected sub-population credit scoring model, and NS-CSM(NC) for the network segmented non-connected sub-population credit scoring model. The top left panel represents the lasso method, the adaptive lasso is on the top right panel, elastic-net at the bottom left and adaptive elastic-net at the bottom right.

Table 5 reports the number of variables selected by each of the four competing methods for the credit score model estimation. From the table, the elastic-net is the least parsimonious, followed by the lasso, and lastly, the adaptive elastic-net and adaptive lasso are the most parsimonious. From **Tables 4, 5**, we observed a significant difference in the number of selected explanatory variables for the benchmark model and the network segmented

models. More precisely, the former model the credit score of a given company by using more variables while the latter on the other hand uses a significantly lower number of variables. The similar results across the four variable selection methods, given their similarities, is not terribly surprising. But they do indicate that the general approach appears to be robust in this setting, which was the main purpose of the testing. The network-based segmentation framework is therefore more parsimonious than the benchmark full population credit score model, and this helps in interpretability.

3.5. Comparing Default Predicting Accuracy

We analyzed the performance of the models by splitting the sample into 70% training and 30% testing sample. We now compare the default prediction accuracy of the models in terms of the standard area under the curve (AUC) derived from the receiver operator characteristic (ROC) curve. The AUC depicts the true positive rate (TPR) against the false positive rate (FPR) depending on some threshold. TPR is the number of correct positive predictions divided by the total number of positives. FPR is the ratio of false positives predictions overall negatives. See **Figure 2** for the plot of the ROC curve for the competing methods.

The comparison of the ROC curves from the competing methods shows that the CSM (in red) lies below the rest. Clearly, the curves of NS-CSM ($\gamma = 0.1$) depicted in green seems to dominate the others. The summary of the area under the ROC curve reported in **Table 6** shows that NS-CSM ($\gamma = 0.1$) is ranked first, followed by NS-CSM ($\gamma = 0.05$), and the lowest AUC is obtained by the CSM. Overall, in terms of default predictive accuracy, the result of the AUC shows the NS-CSM outperforms the CSM, on average by two percentage points. This is an advantage that can be further increased considering as the cut-off the observed default percentages, which are different in the two samples.

We investigate whether the AUC of the network segmented model is significantly different from the benchmark model for the four methods. We applied the DeLong test (DeLong et al., 1988) to investigate the pairwise comparison of the AUC of the benchmark model (i.e., CSM) and that of the NS-CSM for $\gamma = \{0.05, 0.1\}$. We perform these tests under the null-hypotheses that $H_0: \text{AUC (CSM)} \geq \text{AUC (NS-CSM)}$ and the alternative hypotheses, $H_1: \text{AUC (CSM)} < \text{AUC (NS-CSM)}$. **Table 7** reports the one-sided statistical test of the AUC of the

TABLE 6 | Comparing area under the ROC curve (AUC) of the four methods.

	Lasso	Adaptive lasso	Elastic-net	Adaptive elastic-net
CSM	0.8089	0.8061	0.8090	0.8061
NS-CSM($\gamma = 0.05$)	0.8214	0.8204	0.8225	0.8207
NS-CSM($\gamma = 0.1$)	0.8330	0.8277	0.8342	0.8312

TABLE 7 | AUC of the benchmark model relative to the network segmented models under the four methods.

		Statistic	P-value	Significance			Statistic	P-value	Significance
CSM	NS-CSM($\gamma = 0.05$)		-0.7639	0.2225		Adaptive lasso	-0.8598	0.1950	
	NS-CSM($\gamma = 0.1$)		-1.4972	0.0672			-1.3129	0.0946	
CSM	NS-CSM($\gamma = 0.05$)		-0.8241	0.2050		Adaptive elastic-net	-0.8728	0.1914	
	NS-CSM($\gamma = 0.1$)		-1.5770	0.0574			-1.5327	0.0627	

benchmark model relative to the network segmented models. The result of the De Long test shows that while the ROC of CSM is not statistically different from that of NS-CSM ($\gamma = 0.05$), the difference between the ROC of NS-CSM ($\gamma = 0.1$) and the benchmark (CSM) is statistically significant at 90% confidence level for all four methods.

In conclusion, our proposed factor network approach to credit score modeling presents an efficient framework to analyze the interconnections among the borrowers of a peer to peer platform and provides a way to segment a heterogeneous population into clusters with more homogeneous characteristics. The results show that the lasso logistic model for credit scoring leads to better identification of the significant set of relevant financial characteristic variables, thereby producing a more interpretable model, especially when combined with the segmentation of the population via the factor network-based approach. These empirical results are promising, but certainly not definitive. More research is required to determine whether the observed 'lift' truly is significant rather than just an artifact of random chance or spurious correlation, especially given the fact that these p -values are not calibrated in any way (e.g., Sellke et al., 2001) and Calabrese and Giudici (2015). Further research may include a Bayesian approach, as in Figini and Giudici (2011) and Giudici (2001). We therefore find evidence of a modest improvement in the default predictive performance of our model compared to the conventional approach.

4. CONCLUSION

This paper improves credit risk management of SMEs engaged in P2P credit services by proposing a factor network-based approach to segment a heterogeneous population into a cluster of homogeneous sub-populations and estimating a credit score model on the clusters using a lasso-type regularization logistic model.

REFERENCES

- Ahelegbey, D., Giudici, P., and Hadji-Misheva, B. (2019). Latent factor models for credit scoring in P2P systems. *Physica A* 522, 112–121. doi: 10.1016/j.physa.2019.01.130
- Ahelegbey, D. F., Billio, M., and Casarin, R. (2016a). Bayesian graphical models for structural vector autoregressive processes. *J. Appl. Econom.* 31, 357–386. doi: 10.1002/jae.2443
- Ahelegbey, D. F., Billio, M., and Casarin, R. (2016b). Sparse graphical vector autoregression: a bayesian approach. *Ann. Econ. Stat.* 123/124, 333–361. doi: 10.15609/annaeconstat2009.123-124.0333
- Alam, M., Hao, C., and Carling, K. (2010). Review of the literature on credit risk modeling: development of the past 10 years. *Banks Bank Syst.* 5, 43–60.
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *J. Finance* 23, 589–609. doi: 10.1111/j.1540-6261.1968.tb00843.x
- Andreeva, G., Ansell, J., and Crook, J. (2007). Modelling profitability using survival combination scores. *Eur. J. Oper. Res.* 183, 1537–1549. doi: 10.1016/j.ejor.2006.10.064
- Barrios, L. J. S., Andreeva, G., and Ansell, J. (2014). Monetary and relative scorecards to assess profits in consumer revolving credit. *J. Oper. Res. Soc.* 65, 443–453. doi: 10.1057/jors.2013.66
- Battiston, S., Gatti, D. D., Gallegati, M., Greenwald, B., and Stiglitz, J. E. (2012). Liaisons dangereuses: increasing connectivity, risk sharing, and systemic risk. *J. Econ. Dyn. Control* 36, 1121–1141. doi: 10.1016/j.jedc.2012.04.001
- Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *J. Financ. Econ.* 104, 535–559. doi: 10.1016/j.jfineco.2011.12.010
- Calabrese, R., and Giudici, P. (2015). Estimating bank default with generalized extreme value regression models. *J. Oper. Res. Soc.* 66, 1783–1792. doi: 10.1057/jors.2014.106
- Carvalho, C. M., and West, M. (2007). Dynamic matrix-variate graphical models. *Bayesian Anal.* 2, 69–98. doi: 10.1214/07-BA204
- DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44, 837–845. doi: 10.2307/2531595
- Diebold, F., and Yilmaz, K. (2014). On the network topology of variance decompositions: measuring the connectedness of financial firms. *J. Econom.* 182, 119–134. doi: 10.1016/j.jeconom.2014.04.012
- Eichler, M. (2007). Granger causality and path diagrams for multivariate time series. *J. Econom.* 137, 334–353. doi: 10.1016/j.jeconom.2005.06.032

We demonstrate the effectiveness of our approach through empirical applications analyzing the probability of default of over 15,000 SMEs involved in P2P lending across Europe. We compare the results from our model with the one obtained with standard single credit score methods. We find evidence that our factor network approach helps to obtain sub-population clusters such that the resulting models associated with these clusters are more parsimonious than the conventional full population approach, leading to better interpretability and to a modest improved default predictive performance.

DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the supplementary files.

AUTHOR CONTRIBUTIONS

In this manuscript, all the authors investigated how to improve the credit scoring of SMEs involved in P2P lending via a factor network-based segmentation method. The contribution of this work is manifold. DA extended a recently proposed concept of factor network-based classification to a more realistic setting. PG contributed to network-based models for credit risk quantification using a lasso logistic regression. BH-M presented an application of our approach to model the credit score of over 15,000 SMEs engaged in P2P credit services across Southern Europe.

FUNDING

Funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 825215 (Topic: ICT-35-2018 Type of action: CSA).

- Emekter, R., Tu, Y., Jirasakuldech, B., and Lu, M. (2015). Evaluating credit risk and loan performance in online peer-to-peer (P2P) lending. *Appl. Econ.* 47, 54–70. doi: 10.1080/00036846.2014.962222
- Figini, S., and Giudici, P. (2011). Statistical merging of rating models. *J. Oper. Res. Soc.* 62, 1067–1074. doi: 10.1057/jors.2010.41
- Giudici, P. (2001). Bayesian data mining, with applications to benchmarking and credit scoring. *Appl. Stochastic Models Bus. Ind.* 17, 69–81. doi: 10.1002/asmb.425
- Giudici, P., Hadji-Misheva, B., and Spelta, A. (2019). Network-based scoring models to improve credit risk management in peer to peer lending platforms. *Front. Artif. Intell.* 2:3. doi: 10.3389/frai.2019.00003
- Hastie, T., Tibshirani, R., and Friedman, M. F. (2009). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Berlin: Springer.
- IMF (2011). *Global Financial Stability Report: Grappling with Crisis Legacies*. Technical report, World Economic and Financial Services.
- Letizia, E., and Lillo, F. (2018). *Corporate Payments Networks and Credit Risk Rating*. arXiv:1711.07677. Available online at: <https://arxiv.org/abs/1711.07677>
- Moghadam, R., and Viñals, J. (2010). *Understanding Financial Interconnectedness*. Mimeo, Washington, DC: International Monetary Fund.
- Sellke, T., Bayarri, M., and Berger, J. O. (2001). Calibration of p values for testing precise null hypotheses. *Am. Stat.* 55, 62–71. doi: 10.1198/000313001300339950
- Serrano-Cinca, C., and Gutiérrez-Nieto, B. (2016). The use of profit scoring as an alternative to credit scoring systems in peer-to-peer lending. *Decis. Support Syst.* 89, 113–122. doi: 10.1016/j.dss.2016.06.014
- Tabachnick, B. G., Fidell, L. S., and Ullman, J. B. (2007). *Using Multivariate Statistics*, Vol. 5. Boston, MA: Pearson Publisher.
- Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *J. R. Stat. Soc. B* 58, 267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x
- Viñals, J., Tiwari, S., and Blanchard, O. (2012). *The IMF's Financial Surveillance Strategy*. Washington, DC: International Monetary Fund.
- Zou, H. (2006). The adaptive lasso and its oracle properties. *J. Am. stat. Assoc.* 101, 1418–1429. doi: 10.1198/016214506000000735
- Zou, H., and Hastie, T. (2005). Regularization and Variable Selection via the Elastic Net. *J. R. Stat. Soc. B* 67, 301–320. doi: 10.1111/j.1467-9868.2005.00503.x
- Zou, H., and Zhang, H. H. (2009). On the adaptive elastic net with a diverging number of parameters. *Ann. stat.* 37, 1733. doi: 10.1214/08-AOS625

Conflict of Interest Statement: 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 Ahelegbey, Giudici and Hadji-Misheva. 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.



Validation of PARX Models for Default Count Prediction

Arianna Agosto^{1*} and Emanuela Raffinetti²

¹ Department of Economics and Management, University of Pavia, Pavia, Italy, ² Department of Economics, Management and Quantitative Methods, University of Milan, Milan, Italy

OPEN ACCESS

Edited by:

Joerg Osterrieder,
Zurich University of Applied Sciences,
Switzerland

Reviewed by:

Apama Gupta,
Rensselaer Polytechnic Institute,
United States
Fabian Woebeking,
Goethe-Universität Frankfurt am Main,
Germany

Branka Hadji Misheva,
Zurich University of Applied Sciences
(ZHAW), Switzerland

*Correspondence:

Arianna Agosto
arianna.agosto@unipv.it

Specialty section:

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

Received: 27 February 2019

Accepted: 21 May 2019

Published: 12 June 2019

Citation:

Agosto A and Raffinetti E (2019)
Validation of PARX Models for Default
Count Prediction.
Front. Artif. Intell. 2:9.
doi: 10.3389/frai.2019.00009

The growing importance of financial technology platforms, based on interconnectedness, makes necessary the development of credit risk measurement models that properly take contagion into account. Evaluating the predictive accuracy of these models is achieving increasing importance to safeguard investors and maintain financial stability. The aim of this paper is two-fold. On the one hand, we provide an application of Poisson autoregressive stochastic processes to default data with the aim of investigating credit contagion; on the other hand, focusing on the validation aspects, we assess the performance of these models in terms of predictive accuracy using both the standard metrics and a recently developed criterion, whose main advantage is being not dependent on the type of predicted variable. This new criterion, already validated on continuous and binary data, is extended also to the case of discrete data providing results which are coherent to those obtained with the classical predictive accuracy measures. To shed light on the usefulness of our approach, we apply Poisson autoregressive models with exogenous covariates (PARX) to the quarterly count of defaulted loans among Italian real estate and construction companies, comparing the performance of several specifications. We find that adding a contagion component leads to a decisive improvement in model accuracy with respect to the only autoregressive specification.

Keywords: credit risk, systemic risk, contagion, PARX models, validation measures

1. INTRODUCTION

The credit market is experiencing a large growth of innovative financial technologies (fintechs). In particular, peer-to-peer lending platforms propose a business model that disintermediates the links between borrowers and lenders and is based on a stronger interconnectedness between the agents with respect to the traditional banking system. Furthermore, peer-to-peer lenders often do not have access to individual borrowers' data usually employed in banks' credit scoring models, such as financial ratios and credit bureau information. In this context, models analyzing correlation in the default dynamics of different agents or sectors can effectively support credit risk assessment.

More generally, interconnectedness, already known as a trigger of the great financial crisis in 2008–2009, is recognized as a source of *systemic risk*, i.e., according to the European Central Bank, “the risk of experiencing a strong systemic event, which adversely affects a number of systemically important intermediaries or markets.” The impact that an event experienced by an economic agent or sector can have on other institutions in the market is often referred to as *contagion*. From an econometric viewpoint, statistical methods able to properly measure the systemic risk that arises from interconnectedness are necessary to safeguard both traditional intermediaries and peer-to-peer lending investors, therefore maintaining financial stability.

The first systemic risk measures have been proposed for the financial sector, in particular by Adrian and Brunnermeier (2016) and Acharya et al. (2012). These works consider financial market data, calculating the estimated loss probability distribution of a financial institution, conditional on an extreme event in the financial market. Being applied to market prices, these models are based on Gaussian processes.

Financial market data have also been used in another recent approach to systemic risk, based on correlation network models, where contagion effects are estimated from the dependence structure among market prices. The first contributions in this framework are Billio et al. (2012) and Diebold and Yilmaz (2014), who derived contagion measures based on Granger-causality tests and variance decompositions. Ahelegbey et al. (2016) and Giudici and Spelta (2016) have extended the methodology introducing stochastic correlation networks.

Networks represent a relevant modeling approach in peer-to-peer platforms, where continuous credit demand and lending activity makes available large amounts of transaction data. Network models have been recently applied to peer-to-peer lending platforms data by Ahelegbey et al. (2019) and Giudici et al. (2019).

Another possible approach to analyze contagion is to build discrete data models for the counts of default events. Including exogenous covariates in such models allows to test whether the failure of a given firm increases the probability that other failures occur conditional on a set of risk factors. For example, Lando and Nielsen (2010) model default times by Poisson processes with macroeconomic and firm-specific covariates entering the default intensities. Their methodology does not directly include a contagion component, but investigates possible contagion effects by testing whether the Poisson model is misspecified. Default counts are also modeled by Koopman et al. (2012) and, recently, by Azizpour et al. (2018), who use a binomial specification where the probability of default is a time-varying function of underlying factors, also including unobserved components.

Among the approaches to default counts modeling we focus on PARX models developed by Agosto et al. (2016), including autoregressive and exogenous effects in a time-varying Poisson intensity specification. A recent extension by Agosto and Giudici (Submitted) makes PARX models suitable to investigate default contagion. In this paper, PARX models are applied to default counts data in the Italian real estate sector.

Validation is a critical issue in credit risk modeling, because of the interest in selecting indicators able to predict the default peaks, and achieves further importance in artificial intelligence systems, where the traditional accuracy measures based on probabilistic assumptions cannot always be implemented.

In the specific case of contagion analysis, such as the one presented in this paper, model selection also assumes an explanatory role: the comparison of alternative specifications, including contagion components or not and considering different exogenous risk factors, provides a deeper insight into default correlation.

In our empirical application we validate the models applied to default counts using several measures, including the Rank Graduation index *RG*, recently developed by Giudici and

Raffinetti (Submitted). In Giudici and Raffinetti (Submitted), the purpose was to propose an index that is objective and not endogenous to the system itself. The Rank Graduation index (*RG*) was originally developed to deal with two real machine learning applications characterized, respectively, by a binary and a continuous response variable. It is based on the calculation of the cumulative values of the response variable, re-ordered according to the ranks of the values predicted by the considered model. Giudici and Raffinetti (Submitted) showed that the *RG* metric is more effective than the *AUROC* (typically used for models with binary response variables) and the *RMSE* (typically used for models with continuous response variables). Specifically, in the binary case, it appears as an objective predictive accuracy diagnostic, since built on re-ordering the response variable values according to the predicted values themselves, and, in the continuous case, it is not affected by the presence of outliers. Here, the application of the Rank Graduation index is extended to the case of default count data and the related results are compared to those obtained with traditional measures, such as the likelihood-based criteria and *RMSE*. Given its attractive features and properties, both regulators and supervisors may be interested in the *RG* employment in artificial intelligence applications, in order to better understand and manage the business models and avoid decisions based upon wrong outputs which may lead to losses or reputational risks.

The paper is organized as follows. Section 2 describes PARX models and how they can be used to study the default count dynamics and investigate possible contagion effects. Section 3 provides an overview of the main validation criteria and the basic elements characterizing the Rank Graduation measure. Section 4 presents the empirical findings derived from the application and validation of PARX models for default counts. Section 5 concludes.

2. PARX MODELS

The approach to default counts modeling applied in this work is based on PARX models (Agosto et al., 2016). PARX models assume that a count time series y_t , conditional on its past, follows a Poisson distribution with a time-varying intensity $\lambda_t > 0$, whose formulation includes an autoregressive part and a d -dimensional vector of exogenous covariates $x_t: = (x_{1t}, x_{2t}, \dots, x_{dt})' \in \mathbb{R}^d$:

$$y_t | \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t) \Leftrightarrow P(y_t = y | \mathcal{F}_{t-1}) = \frac{\lambda_t^y \exp(-\lambda_t)}{y!} \quad (1)$$

$$\lambda_t = \omega + \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{i=1}^q \beta_i \lambda_{t-i} + \sum_{i=1}^d \gamma_i f(x_i)$$

with \mathcal{F}_{t-1} denoting the σ -field $\sigma\{y_0, \dots, y_{t-1}, \lambda_0, \dots, \lambda_{t-1}, x_0, \dots, x_{t-1}\}$, $\omega > 0$, $\alpha_i \geq 0$ ($i = 1, 2, \dots, p$) and $\beta_i \geq 0$ ($i = 1, 2, \dots, q$).

When the vector of unknown parameters $\gamma: = (\gamma_1, \dots, \gamma_d)$ is null, the model reduces to Poisson Autoregression (PAR) developed by Fokianos et al. (2009), who showed how including past values of the intensity λ_t allows for parsimonious modeling

of long memory effects. Note that exogenous covariates are included through a non-negative link function to guarantee that intensity is positive.

The presence of both dynamic and exogenous effects makes PARX models suitable for describing count time series of events that cluster in time, as defaults are known to do. Furthermore, it can be shown that including an autoregressive component as well as covariates in a Poisson process generates overdispersion, that is unconditional variance larger than the mean, a typical feature of default count time series.

Agosto et al. (2016) applied model (1) to Moody's rated US corporate default counts, with the aim of distinguishing between the impact of past defaults on current default intensity—possibly due to contagion effects—and the impact of macroeconomic and financial variables acting as common risk factors. Recently, Agosto and Giudici (Submitted) proposed to extend PARX models to accomplish investigation of default contagion effects. Differently from model (1) and following Fokianos and Tjøstheim (2011), they use a log-linear intensity specification. This allows to consider negative dependence on exogenous covariates, which can be useful in credit risk applications.

Letting y_{jt} the number of defaults in economic sector (or, more generally, group of borrowers) j at time t and y_{kt} the number of defaults in sector k , they define the following model:

$$y_{jt} | \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_{jt}) \quad (2)$$

$$\log(\lambda_{jt}) = \omega_j + \sum_{i=1}^p \alpha_{ji} \log(1 + y_{jt-i}) + \sum_{i=1}^q \beta_{ji} \log(\lambda_{jt-i})$$

$$+ \sum_{i=1}^r \gamma_{ji} x_{t-i} + \sum_{i=1}^s \zeta_{ji} \log(1 + y_{kt-i})$$

with $\omega_j, \alpha_{ji}, \beta_{ji}, \gamma_{ji}, \zeta_{ji} \in \mathbb{R}$ and $x_{t-i} := (x_{1t-i}, x_{2t-i}, \dots, x_{dt-i})' \in \mathbb{R}^d$ being a vector of lagged exogenous covariates. In model (2), that the authors call Contagion PARX, ζ_j measures the effect of the covariate default count process on the response default counts, which can be interpreted as a contagion effect. Taking the $\log(\cdot) + 1$ of counts allows to deal with possible zero values. This specification can easily be extended to the case where the default counts of a set of different sectors, rather than only one covariate default series, are included among the regressors.

3. MODEL VALIDATION

A basic issue of the artificial intelligence systems is the validation process for the model prediction quality assessment. In this paper, we consider the available literature for validation procedures and illustrate a new practice for the validation.

In literature, several metrics aimed at comparing and improving the models are available, depending on the nature of data. As mentioned above, one of the focus of this paper is on the use of the Poisson autoregressive models for modeling default counts. The presence of a discrete response variable suggests the choice of the Root Mean Squared Error (RMSE) and the criteria based on likelihood, such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), as

the most widely employed measures for the model predictive accuracy evaluation.

It is worth noting that in the model validation research field, the lack of a standard metric, working regardless of the nature of the response variable to be predicted, is still a crucial drawback to be faced. Recently, Giudici and Raffinetti (Submitted) have worked out one possible solution by proposing a new measure, the *RG Rank Graduation index*, which is based on the calculation of the cumulative values of the response variable, according to the ranks of the values predicted by a given model. The main features of the *RG* criterion together with a brief description of the conventional validation measures are provided in the following subsections.

3.1. Conventional Model Validation Measures

The RMSE, AIC, and BIC criteria, intended as some of the most broadly used metrics for the model validation, are defined as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \quad (3)$$

where the y_i 's and \hat{y}_i 's represent the response variable observed and predicted values (with $i = 1, \dots, n$), respectively,

$$AIC = -2\log L(\hat{\theta} | x_1, \dots, x_n) + 2k \quad (4)$$

and

$$BIC = -2\log L(\hat{\theta} | x_1, \dots, x_n) + k\log(n), \quad (5)$$

where θ is the set of model parameters, $\log L(\hat{\theta} | x_1, \dots, x_n)$ is the log-likelihood of the model given the data x_1, \dots, x_n when evaluated at the maximum log-likelihood estimate of θ ($\hat{\theta}$), k is the number of the estimated parameters in the model and n is the number of observations.

The best model, in terms of predictive accuracy, is the one that provides the minimum RMSE, AIC and BIC (for more details, see e.g., Kuha, 2004; Hyndman and Koehler, 2006).

3.2. The *RG* as an Additional Model Validation Criterion

Besides the conventional model validation criteria, the *RG* measure deserves a wider discussion, especially because it appears as a more general predictive accuracy criterion which does not depend on the type of data to be analysed. As mentioned above, in Giudici and Raffinetti (Submitted), the *RG* was proposed as a unique metric to assess the model predictive accuracy in presence of both binary and continuous response variables. Moreover, due to its features and construction it fulfills some attractive properties: (1) it appears as an objective criterion compared with the *AUROC* metric, which depends on the arbitrary choice of the cut-off points; (2) it is a robust criterion since non-sensitive to the presence of outliers. Given the topic of this paper, related to the employment of discrete data models for default counts, it is therefore worth to extend the frontiers of the *RG* application

areas to the context of discrete response variables.

The interest in applying the *RG* index to default count data is also linked to some typical features shown by the time series of defaults. The common presence of peaks and outliers makes indeed preferable to evaluate predictive accuracy of default count models through concordance measures rather than error measures that are known to be sensitive to outliers.

In order to better highlight the main strengths of our validation approach, a brief overview on the *RG* construction seems to be basic. The proposal is based on the so-called *C* concordance curve, which is obtained by ordering the normalized *Y* response variable observed values according to the ranks of the predicted \hat{Y} values provided by the model.

Let *Y* be a discrete response variable and let X_1, \dots, X_p be a set of *p* explanatory variables. Suppose to apply a model such that $\hat{y} = f(X)$. The model predictive accuracy is assessed by measuring the distance between the set of the *C* concordance curve points, whose coordinates are denoted with $(i/n, (1/(n\bar{y})) \sum_{j=1}^i y_{\hat{r}_j})$, where $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ and $y_{\hat{r}_j}$ represents the *j*-th response variable value ordered by the rank of the corresponding predicted value \hat{y}_j (with $j = 1, \dots, i$ and $i = 1, \dots, n$), and the set of the bisector curve points of coordinates $(i/n, i/n)$. As an example, the graphical representation of the *C* concordance (in red) and bisector (in black) curves is displayed in **Figure 1**. **Figure 1** reports also two other curves: the response variable L_Y Lorenz curve (in blue), which is defined by the normalized *Y* values ordered in non-decreasing sense, and the response variable L'_Y dual Lorenz curve (in green), which is defined by the normalized *Y* values ordered in non-increasing sense.

Both the response variable Lorenz and dual Lorenz curves take a remarkable role in the *RG* measure construction, especially the response variable L_Y Lorenz curve. Indeed, since the model predictive accuracy degree depends on the distance between the bisector and the *C* concordance curves, it follows that the more the *C* concordance curve moves away from the bisector curve, the more the model predictive accuracy improves. This because the bisector curve detects a model without predictive capability. Indeed, if $\hat{y}_i = \bar{y}$, for any $i = 1, \dots, n$, through some manipulations, the coordinates of the *C* concordance curve becomes $(i/n, i/n)$, which perfectly corresponds to the coordinates of points characterizing the bisector curve. Analogously, if the *C* concordance curve perfectly overlaps with the L_Y Lorenz curve, then the model is perfect because it preserves the ordering between the observed response variable *Y* values and the corresponding \hat{Y} estimated values. In such a case, the coordinates of the *C* concordance curve become $(i/n, (1/(n\bar{y})) \sum_{j=1}^i y_{(j)})$, where $y_{(j)}$'s, with $j = 1, \dots, i$ and $i = 1, \dots, n$, are the response variable values ordered in non-decreasing sense.

Based on the above considerations, the *RG* measure takes the following expression:

$$RG = \sum_{i=1}^n \frac{\left\{ (1/(n\bar{y})) \sum_{j=1}^i y_{\hat{r}_j} - i/n \right\}^2}{i/n} = \sum_{i=1}^n \frac{\{C(y_{\hat{r}_i}) - i/n\}^2}{i/n}, \quad (6)$$

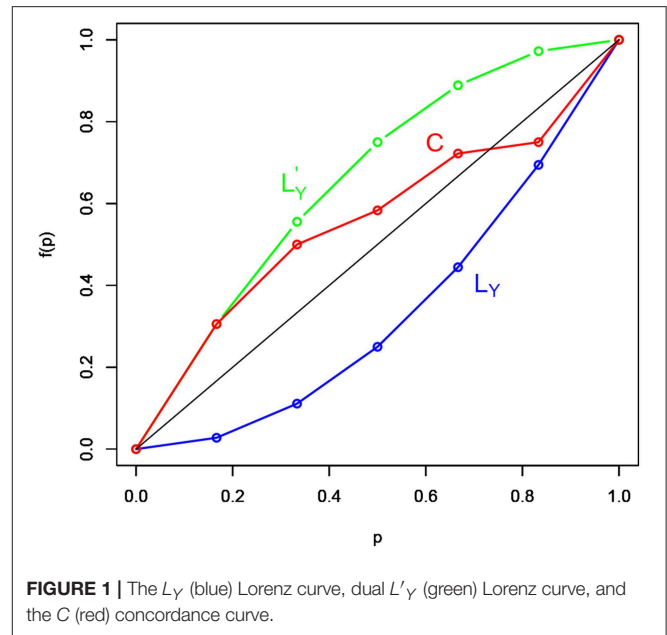


FIGURE 1 | The L_Y (blue) Lorenz curve, dual L'_Y (green) Lorenz curve, and the *C* (red) concordance curve.

where $C(y_{\hat{r}_j}) = \frac{\sum_{j=1}^i y_{\hat{r}_j}}{\sum_{i=1}^n y_i}$ represents the cumulative values of the (normalized) response variable *Y*. The *RG* measure in (6) appears as an absolute metric, since it takes values in the close range $[0, RG_{max}]$, where RG_{max} is the maximum value that can be achieved. Trivially, the maximum *RG* value can be reached if the model perfectly explains the response variable, meaning that the *C* concordance curve indifferently overlaps with the response variable Lorenz or dual Lorenz curves. Indeed, the distance between the *Y* Lorenz or dual Lorenz curves and the bisector curve is the same, being the two curves symmetric around the bisector curve. A normalized *RG* measure is then defined as the ratio between the absolute *RG* measure and its maximum value RG_{max} .

Finally, we remark that when some of the \hat{Y} values are equal to each other, we take into account the adjustment suggested by Ferrari and Raffinetti (2015) in order to solve the re-ordering problem. Specifically, the original *Y* values associated with the equal \hat{Y} values are substituted by their mean.

4. APPLICATION

In this section we provide the application of PARX models to Italian corporate default counts data in the real estate sector and their evaluation through different validation measures. Bank of Italy's Credit Register collects the quarterly number of transitions to *bad loans* in major economic sectors. Bad loans are exposures to insolvent debtors that cannot be recovered and that the bank must report as balance sheet losses. Being an absorbent state, the number of loans turned out to be *bad* in a given period can be used as a proxy of the default count at that time. The data are quarterly and divided by economic sector. Among the sectors included in the database we focus on the Real Estate and Commercial ones, using data covering the period March

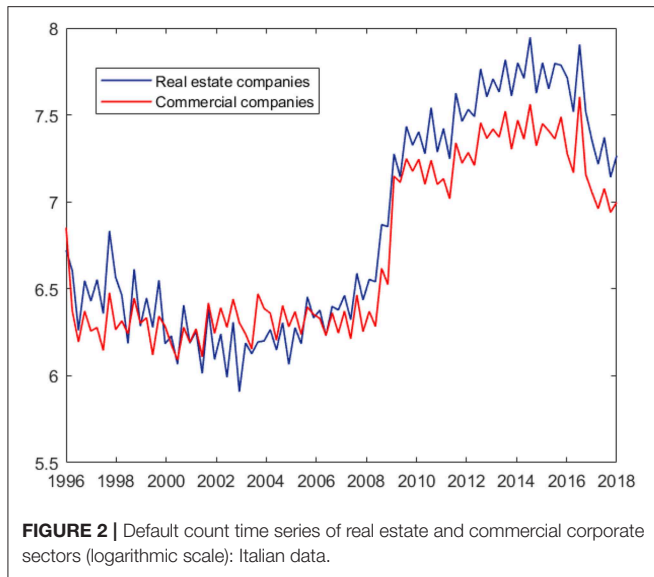


FIGURE 2 | Default count time series of real estate and commercial corporate sectors (logarithmic scale): Italian data.

TABLE 1 | Summary statistics for the real estate sector default counts: Italian data.

Mean	Std. Dev.	Min	Max
1132.9	711.7	368	2825

1996–June 2018 (90 observations). The real estate sector includes both real estate and construction companies and was one of the most hit by the recent financial crisis. Our choice is motivated by the economic interest in verifying the impact that the default dynamics of commercial firms, highly influenced by the changes in consumption behavior, may have on the real estate sector. Possible contagion from the commercial to the real estate sector is mainly due to the decrease of both business and private investments by the owners of commercial activities, causing a reduction in the demand of new buildings and real estate services.

Figure 2 shows the default count time series of the two economic sectors considered. Both series exhibit clustering and a possible structural break in 2009, with an increase in both level and variability. **Table 1** reports the main summary statistics for the response variable of our exercise, that is the default counts among real estate Italian firms, while **Figure 3** shows the autocorrelation function of the series. Both the presence of overdispersion (the empirical variance is 506468.7 and the empirical average 1132.9) and the slowly decaying autocorrelation encourage the use of PARX to model the data.

To investigate credit contagion effects between the two sectors and show our validation procedure, we consider the model regressing real estate sector default counts on their past values and on past commercial sector default counts.

An important robustness and validation step when applying PARX models is assessing the effects of including exogenous covariates summarizing the macroeconomic context, such as the business cycle. The aim is to verify to what extent the macroeconomic stress affecting all the economic agents and sectors explains the default and contagion dynamics.

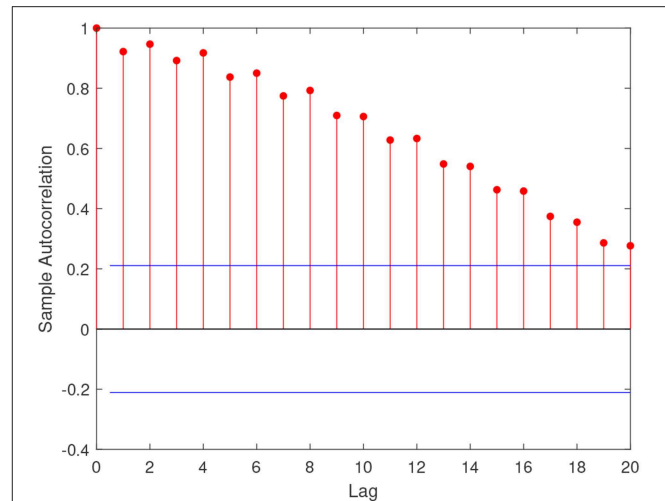


FIGURE 3 | Sample autocorrelation function of real estate default count time series: Italian data.

TABLE 2 | Parameter estimates for real estate sector default counts.

Variable	Estimate	Standard error	t-stat
Constant	−0.1339	0.3285	0.4075
Real estate sector bad loans in t-1	0.6062	0.1591	3.8103***
Commercial sector bad loans in t-1	−0.2886	0.2689	−1.0732
Commercial sector bad loans in t-2	0.7161	0.1299	5.5129***
GDP growth rate in t-1	−0.0341	0.0284	−1.2009
GDP growth rate in t-2	−0.0705	0.0274	−2.5732**

*** $p < 0.001$; ** $p < 0.01$.

Thus, we first estimate a model (Full Contagion PARX) that, according to specification (2), includes both a contagion component and the exogenous covariate GDP in a log-linear intensity specification¹:

$$\log(\lambda_t) = \omega + \alpha \log(1 + y_{t-1}) + \gamma_1 GDP_{t-1} + \gamma_2 GDP_{t-2} + \zeta_1 \log(1 + y_{Ct-1}) + \zeta_2 \log(1 + y_{Ct-2}) \quad (7)$$

where GDP_t is the Italian GDP growth rate and y_{Ct} is the number of defaults among commercial sector companies at time t .

From **Table 2**, reporting the parameter estimates for the model above, note that the effect of GDP variation on the real estate sector default risk is significant at the second lag, suggesting a delayed effect of the business cycle on the corporate solvency dynamics which is reasonable from an economic point of view. Also the impact of commercial sector default counts turns out to be significant with a two quarters lag.

In order to highlight the contribution of the different components—autoregressive, contagion, and exogenous—and validate the model we then consider two alternative specifications.

¹The number of lags has been determined through preliminary model selection based on likelihood ratio and BIC criterion.

TABLE 3 | Validation measures for the considered models.

Model	RMSE	AIC	BIC	RG
Full contagion PARX model	207.68	−1,256,019	−1,256,004	6.098
Contagion PARX model	222.02	−1,255,643	−1,255,633	6.114
PAR model	272.06	−1,254,332	−1,254,327	5.796

We first estimate a PARX model that, following specification (1), includes an autoregressive and an exogenous component in a linear intensity specification:

$$\lambda_t = \omega + \alpha y_{t-1} + \gamma_1 GDP_{t-1}^- + \gamma_2 GDP_{t-2}^- \quad (8)$$

where $GDP^- := \mathbb{I}_{GDP < 0} |GDP|$, that is the absolute value of the negative part of GDP growth rate. This ensures that default intensity is positive, as needed in the linear specification. Fitting the model above, we do not find significant effects of GDP decrease on the real estate sector. Thus, the model reduces to an only autoregressive Poisson model as the previously cited PAR. According to this result, while negative correlation with the business cycle taken into account by the log-linear model significantly explains the default dynamics, the positive association between the GDP decrease and the default counts is not significant in our exercise. This highlights the advantage of using specifications that allow to consider negative dependence.

The last competing model is a Contagion PARX without other covariates than commercial sector default counts [γ parameters equal to 0 in specification (2)]:

$$\log(\lambda_t) = \omega + \alpha y_{t-1} + \zeta_1 \log(1 + y_{Ct-1}) + \zeta_2 \log(1 + y_{Ct-2}) \quad (9)$$

We now compare the in-sample performances of the three models above: PAR model, Contagion PARX model, and Full Contagion PARX model by using the RMSE, AIC, BIC and RG validation measures. The results are illustrated in **Table 3**.

First note that the Full Contagion PARX model is the most performing according to RMSE, AIC, and BIC criteria. In particular, moving from the PAR to the Contagion PARX specification leads to a decrease of nearly 24% in the RMSE. The model ordering changes when considering the RG index. The model showing the higher RG index is indeed the Contagion PARX one, with a value of 6.114. The Full Contagion PARX model shows a slightly lower value (6.098), while the RG index of the PAR model is 5.796. As $RG_{max} = 6.709$, it follows that the PAR model explains the 86.4% of the variable ordering, compared with the 90.9% of the Full Contagion PARX Model and the 91.1% of the Contagion PARX Model.

According to all the considered measures, adding the contagion component leads to a decisive increase in model performance with respect to the only autoregressive specification,

with a decrease of 18% in RMSE and an increase of nearly 3.5% in accuracy. Considering the negative association between the macroeconomic stress and default risk considerably reduces the error measure—the decrease in RMSE with respect to the Contagion PARX model is around 7% - but does not improve model performance in terms of accuracy, measured through the RG index. In such a case, the choice of the preferable specification depends on the objective of model comparison. If the aim, as in our contagion analysis, is validating a model that well explains the empirical distribution of the data even with a limited number of parameters, rather than getting a point forecast of the response variable, decisions based on a concordance measure are more appropriate.

5. CONCLUSION

In this paper, we have illustrated an application of PARX models, which investigate contagion through Poisson autoregressive stochastic processes, and we have evaluated the predictive accuracy of different specifications. While previous works focused on the theory development and extension of PARX, we concentrate on the issue of validating these models and measuring the contribution of contagion and exogenous components to their predictive performance. For doing so, we resorted to a novel metric, called RG index, which is independent on the involved response variable nature. Specifically, the RG measure, originally considered in the cases of binary and continuous data, was here extended with the aim of covering also the case of discrete data.

Fitting several PARX-type specification to the quarterly count of defaulted loans in the Italian real estate sector, we find evidence of a significant effect of commercial sector defaults on real estate default risk. We also find that considering the effect of the business cycle improves model performance according to likelihood-based criteria and traditional error measures, but it does not increase predictive accuracy according to the new concordance metric.

DATA AVAILABILITY

Publicly available datasets were analyzed in this study. This data can be found here: <https://www.bancaditalia.it/statistiche/basi-dati/bds/index.html>.

AUTHOR CONTRIBUTIONS

AA is a post-doctoral research fellow at University of Pavia, Department of Economics and Management. She has a research experience in Econometrics and Quantitative Finance with application to risk management, especially credit risk and contagion models. ER is Assistant Professor in Statistics at University of Milan, Department of Economics, Management, and Quantitative Methods. She has a research experience in dependence analysis and model validation criteria.

REFERENCES

- Acharya, V., Engle, R., and Richardson, M. (2012). Capital shortfall: a new approach to ranking and regulating systemic risks. *Am. Econ. Rev.* 102, 59–64. doi: 10.1257/aer.102.3.59
- Adrian, T., and Brunnermeier, M. K. (2016). CoVaR. *Am. Econ. Rev.* 106, 1705–1741. doi: 10.1257/aer.20120555
- Agosto, A., Cavaliere, G., Kristensen, D., and Rahbek, A. (2016). Modeling corporate defaults: poisson autoregressions with exogenous covariates (PARX). *J. Empir. Finance* 38, 640–663. doi: 10.1016/j.jempfin.2016.02.007
- Ahelegbey, D., Billio, M., and Casarin, R. (2016). Bayesian graphical models for structural vector autoregressive processes. *J. Appl. Econometr.* 31, 357–386. doi: 10.1002/jae.2443
- 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
- Azizpour, S., Giesecke, K., and Schwenkler, G. (2018). Exploring the sourcing of default clustering. *J. Financ. Econ.* 129, 154–183. doi: 10.1016/j.jfineco.2018.04.008
- Billio, M., Getmansky, M. W., Lo, A., and Pelizzon, L. (2012). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *J. Financ. Econ.* 104, 535–559. doi: 10.1016/j.jfineco.2011.12.010
- Diebold, F., and Yilmaz, K. (2014). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *J. Econometr.* 182, 119–134. doi: 10.1016/j.jeconom.2014.04.012
- Ferrari, P. A., and Raffinetti, E. (2015). A different approach to dependence analysis. *Multivar. Behav. Res.* 50, 248–264. doi: 10.1080/00273171.2014.973099
- Fokianos, K., Rahbek, A., and Tjøstheim, D. (2009). Poisson autoregression. *J. Am. Stat. Assoc.* 104, 1430–1439. doi: 10.1198/jasa.2009.tm08270
- Fokianos, K., and Tjøstheim, D. (2011). Log-linear Poisson autoregression. *J. Multivar. Anal.* 102, 563–578. doi: 10.1016/j.jmva.2010.11.002
- Giudici, P., Hadji-Misheva, B., and Spelta, A. (2019). Network based scoring models to improve credit risk management in peer to peer lending platforms. *Front. Artif. Intell.* 2:3. doi: 10.3389/frai.2019.00003
- Giudici, P., and Spelta, A. (2016). Graphical network models for international financial flows. *J. Bus. Econ. Stat.* 34, 128–138. doi: 10.1080/07350015.2015.1017643
- Hyndman, R. J., and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *Int. J. Forecast.* 22, 679–688. doi: 10.1016/j.ijforecast.2006.03.001
- Koopman, S. J., Lucas, A., and Schwaab, B. (2012). Dynamic factor models with macro, frailty, and industry effects for U.S. default counts: the credit crisis of 2008. *J. Bus. Econ. Stat.* 30, 521–532. doi: 10.1080/07350015.2012.700859
- Kuha, J. (2004). AIC and BIC: comparisons of assumptions and performance. *Sociol. Methods Res.* 33, 188–229. doi: 10.1177/0049124103262065
- Lando, D., and Nielsen, M. (2010). Correlation in corporate defaults: contagion or conditional independence? *J. Financ. Intermed.* 19, 355–372. doi: 10.1016/j.jfi.2010.03.002

Conflict of Interest Statement: 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 Agosto and Raffinetti. 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.



Neural Network Models for Bitcoin Option Pricing

Paolo Pagnottoni*

Department of Economics and Management, University of Pavia, Pavia, Italy

Despite the current growing interest in Bitcoins—and cryptocurrencies in general—financial instruments, as well as studies related to them, are quite underdeveloped. Therefore, this article aims to provide a suitable pricing model for options written on this peculiar underlying. This is done through an artificial neural network approach, where classical pricing models—namely the trinomial tree, Monte Carlo simulation, and explicit finite difference method—are used as input layers. Results show that options written on Bitcoin turn out to be systematically overpriced when considering classical methods, whereas a noticeable improvement in price predictions is achieved by means of the proposed neural network model.

Keywords: cryptocurrencies, bitcoin, option pricing, neural network, alternative option pricing methods

OPEN ACCESS

Edited by:

Jochen Papenbrock,
Independent Researcher, Frankfurt,
Germany

Reviewed by:

Francesco Caravelli,
Los Alamos National Laboratory
(DOE), United States
Emanuela Raffinetti,
University of Milan, Italy

*Correspondence:

Paolo Pagnottoni
paolo.pagnottoni01
@universitadipavia.it

Specialty section:

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

Received: 31 October 2018

Accepted: 29 April 2019

Published: 03 July 2019

Citation:

Pagnottoni P (2019) Neural Network
Models for Bitcoin Option Pricing.
Front. Artif. Intell. 2:5.
doi: 10.3389/frai.2019.00005

1. INTRODUCTION

Stock options are a category of financial derivatives which became widely employed by investors and speculators during the last few decades. Nevertheless, investors may ineffectively manage their portfolios if they are not able to value options in a proper way. For this reason, a reliable methodology capable to yield an option's current price or forecast is fundamental for investors in order to produce a rigorous decision making. This is particularly true when considering non-mature and volatile markets like the cryptocurrency one.

The theory of option pricing is broad and involves various types of pricing techniques, largely parametric ones. The most widely known option pricing method is arguably the one defined by Black and Scholes (1973). Although this technique has been widely employed by practitioners, its strict set of assumptions, as well as subjectivity with respect to the parameter choices, often yields to unreliable results to some extent. To illustrate, the leptokurtic behavior of return distributions and the volatility smiles and skews are features that cannot be captured by such a simplistic technique.

Besides the Black-Scholes model and its modifications, other parametric models have been developed and became widely used, among which the (binomial and trinomial) tree models, the finite difference method and the Monte Carlo simulation. While tree models converge to the Black-Scholes one in case the time occurring between steps is small enough, other methodologies take into consideration pricing aspects that these two models do not. Indeed, the Monte Carlo simulation allows for random shocks other than those provided by the volatility and the movement probabilities of the tree models, whereas the finite difference method relies on a different simulation scheme. This is the reason why in this paper examines and includes tree models, the Monte Carlo simulation, and the finite difference method as pricing methodologies.

Alongside the category of classical derivative and option pricing models, non-parametric models, such as neural networks gradually emerged, mainly thanks to their improved predictive performance with respect to the former techniques. Yao et al. (2000) predicts prices related to the Nikkei 225 index futures using back-propagation neural networks. Their results show that, despite the Black-Scholes model is still good for pricing at-the-money options, the neural network

outperforms it, in particular when considering volatile markets. Another research conducted by Liang et al. (2009) motivates this paper's approach, as the authors use classical models (binomial tree, finite difference method, and Monte Carlo simulation) in a first stage to forecast the option price and refine those forecasts through neural networks and support vector machines in a second stage. This technique allows to notably reduce forecast error, i.e., substantially improves price forecasts in their Hong Kong option market framework. Nonetheless, there are many other examples on neural network models for derivative securities pricing which found that neural networks outperform classical models—see, for instance, Hutchinson et al. (1994), Malliaris and Salchenberger (1996), Amilon (2003), Binner et al. (2005), and Lin and Yeh (2005).

Research related to the cryptocurrency market, as the phenomenon itself, is relatively new. Despite that, there is a massive interest of the academic community in investigating this new market and its peculiar features from all points of view, with a particular focus on Bitcoin. Indeed, since Nakamoto (2008) introduced the concept of Bitcoin as a purely peer-to-peer version of electronic cash, researches developed following different and multidisciplinary fields. Some researchers provide a general descriptive analysis of the cryptocurrency framework. To illustrate, in Dwyer (2015) we may find a detailed overview on technical issues of Bitcoin and the cryptocurrency market in general. Also White (2015) goes through the key concepts of cryptocurrencies, while focusing on the so called “Altcoins”¹. A further study by Kroll et al. (2013) examines the Bitcoin mining process thoroughly. Another stream of the literature, with studies conducted by Brandvold et al. (2015) and Pagnottoni and Dimpfl (2019), finds the leader and follower Bitcoin exchanges of the price discovery process through an econometric analysis of its price across different exchange.

Despite the quite wide set of studies in the cryptocurrency area, to the best of our knowledge there is not yet any research trying to address option pricing related to Bitcoin (or cryptocurrency) derivatives. The aim of this study is to propose a pricing methodology that is feasible to price cryptocurrency options. Without loss of generality, the paper focuses on European style Bitcoin put and call options which became recently available on the market. To this end, the study makes use of a two stage approach. The first stage consists of option pricing through parametric approaches, such as tree models, finite difference method, and Monte Carlo simulation. In the second stage, artificial neural networks are employed in order to combine the parametric option pricing approaches and capture the residual errors by learning schemes in the current status of the option market. Their performance is then compared to the conventional option pricing techniques obtained in the first stage. Results point to the predominance of the neural network models with respect to the conventional methods in pricing Bitcoin options and, therefore, in capturing their real price dynamics. As a robustness check, an out-of-sample analysis confirm the previous result, as well as a cross validation analysis through

random sub-sampling reveals that—despite there is still some room for improvement—results are arguably stable and the neural network is a suitable model in order to price options written on Bitcoin.

The remainder of the paper proceeds as follows. Section 2 outlines the methodology employed. Section 3 describes and analyzes the data. Section 4 presents the results. Section 5 illustrates the robustness analysis conducted. Section 6 concludes.

2. METHODOLOGY

This section briefly introduces the classical parametric option pricing techniques used in this paper: specifically, tree models, finite difference method, and Monte Carlo simulation. After that, I discuss the neural network model and the comprehensive approach for option pricing.

The following notation will be used. S represents the underlying asset price, C is the option price, K is the options' exercise price, σ denotes the asset price volatility, r represents the risk-free interest rate, Δt is the time interval (i.e., the time period length), and T is the time to maturity.

2.1. Tree Models

Tree models are widely used not only to price European style options, but also closed-form American options, as they can account for the early exercise feature. Milestone references for binomial trees are the ones of Cox et al. (1979) and Rendleman and Barter (1979). Further extensions are proposed by Boyle (1977), Nelson and Ramaswamy (1990), and Hull and White (1990a).

In the binomial tree setup, the underlying asset price $S_{t,i}$ with $t = 0, 1, 2, \dots, n - 1$ may either experience an up movement to $S_{t+1,i}$ or a down movement to $S_{t+1,i+1}$, with $t = 1, 2, \dots, n$. This happens according to an upward rate u and a downward rate d , which Cox et al. (1979) define as:

$$u = e^{\sigma\sqrt{\Delta t}}, \quad d = e^{-\sigma\sqrt{\Delta t}} \quad (1)$$

where $\Delta t = \frac{T}{n}$ denotes the time step from t to $t + 1$ and n the total number of time steps in the binomial tree.

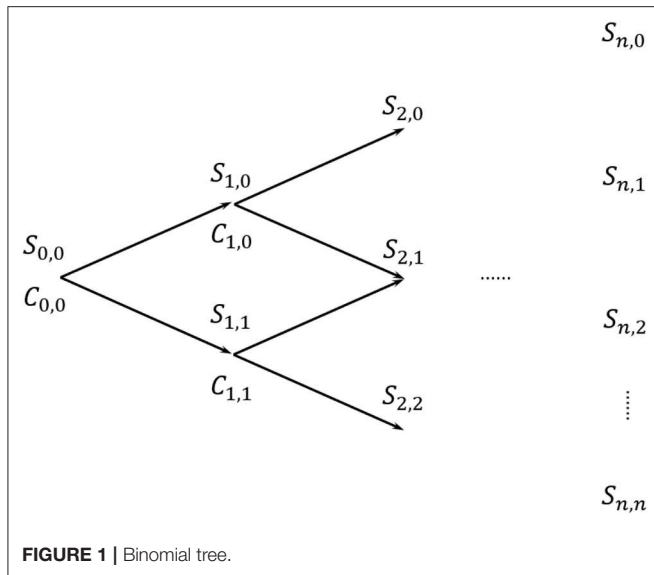
A graphical representation of a n -step binomial tree is illustrated in Figure 1. Arrows constitute possible paths for the price dynamics, whereas nodes represent the underlying price $S_{t,i}$ from which the option price $C_{t,i}$ is computed. Option prices are then recursively computed from the last ones to the first one, going backwards, according to the following:

$$C_{t-\Delta t,i} = e^{-r\Delta t}(pC_{t,i+1} + (1-p)C_{t,i}) \quad (2)$$

where r is the risk-free rate, and the probabilities of up (p) and down (p_d) movements are defined as

$$p = \frac{e^{r\Delta t} - d}{u - d}, \quad p_d = 1 - p \quad (3)$$

¹“Altcoin” stands for “alternative coin.” The term is used to indicate all cryptocurrencies except for Bitcoin.



The trinomial tree (**Figure 2**) works in a similar way. However, in this setup, the underlying asset price $S_{t,i}$ with $t = 0, 1, 2, \dots, n-1$ may either experience an up movement to $S_{t+1,i+1}$, a middle movement to $S_{t+1,i}$ or a down movement to $S_{t+1,i-1}$, with $t = 1, 2, \dots, n$. This happens according to an upward rate u , downward rate d and middle rate m defined as:

$$u = e^{\sigma\sqrt{2\Delta t}}, \quad d = e^{-\sigma\sqrt{2\Delta t}}, \quad m = 1 \quad (4)$$

In this case, the probabilities of up (p), down (p_d) and middle (p_m) movements are defined as:

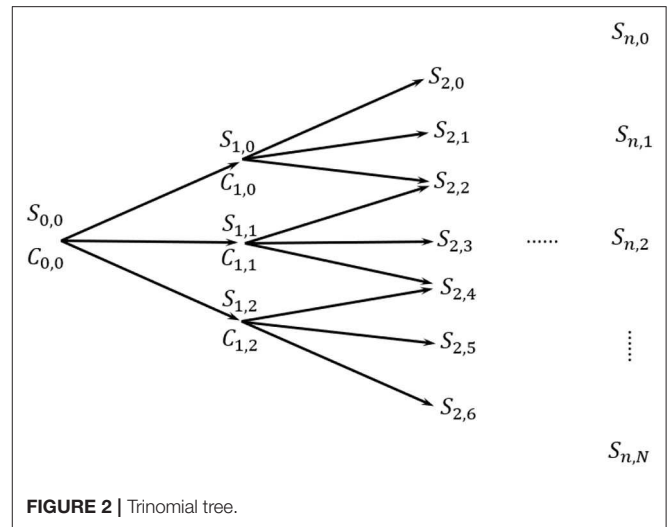
$$p = \left(\frac{e^{(r)\frac{\Delta t}{2}} - e^{-\sigma\sqrt{\frac{\Delta t}{2}}}}{e^{\sigma\sqrt{\frac{\Delta t}{2}}} - e^{-\sigma\sqrt{\frac{\Delta t}{2}}}} \right)^2, \quad p_d = \left(\frac{e^{\sigma\sqrt{\frac{\Delta t}{2}}} - e^{(r)\frac{\Delta t}{2}}}{e^{\sigma\sqrt{\frac{\Delta t}{2}}} - e^{-\sigma\sqrt{\frac{\Delta t}{2}}}} \right)^2, \quad p_m = 1 - (p + p_d) \quad (5)$$

Among the advantages of using the trinomial trees, computational efficiency as well as precision are of our interest. Indeed, the trinomial tree should yield to more precise prices with less time steps if compared to the binomial counterpart.

2.2. Finite Difference Method

As extensively described in Brennan and Schwartz (1977), the finite difference method allows to price options through the solution of some differential equations with respect to the option prices. These equations are transformed into difference equations, whose solutions are iteratively solved by CPUs.

According to the finite difference method, the time to maturity T is segmented into p equally sized time periods Δt , whereas the asset price is segmented into q steps of length ΔS , ranging from a minimum of 0 to a maximum of S_{max} . This can be represented as



a grid in which the horizontal line is the number of periods and the vertical one the asset prices.

In the present case, the application uses the so called explicit finite difference method, which solves the differential equations in a forward way, as elucidated by Hull and White (1990b). The reason behind our choice is that the explicit finite difference method is arguably more efficient than the implicit one, which in contrast solves the differential equations backwards. In particular, the equation to be solved is the well-known partial differential equation of Black-Scholes, i.e.,

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS \frac{\partial C}{\partial S} = rC \quad (6)$$

Where $i = 1, 2, \dots, p$ and $j = 1, 2, \dots, q$. The discrete version of Equation (6) is:

$$-\frac{C_{ij} - C_{i-1,j}}{\Delta t} = \frac{1}{2}\sigma^2 \frac{C_{i,j+1} - 2C_{ij} + C_{i,j-1}}{\Delta S^2} + rS \frac{C_{ij} - 2C_{i,j-1} + C_{i,j-2}}{2\Delta S} - rC_{i+1,j}. \quad (7)$$

The option price can then be derived as:

$$C_{ij} = \frac{1}{1 + r\Delta t} (pC_{i+1,j+1} + p_m C_{i+1,j} + p_d C_{i+1,j-1}) \quad (8)$$

where the probabilities associated with an up, middle or down movement are respectively:

$$p = S_j r \frac{\Delta t}{2\Delta S} + \frac{1}{2} S_j^2 \sigma^2 \frac{\Delta t}{\Delta S^2} \quad (9)$$

$$p_m = 1 - S_j^2 \sigma^2 \frac{\Delta t}{\Delta S^2} \quad (10)$$

$$p_d = -\frac{S_j r \Delta t}{2 \Delta S} + \frac{1}{2} S_j^2 \sigma^2 \frac{\Delta t}{\Delta S^2} \quad (11)$$

For a detailed explanation of the finite difference method, refer to Brennan and Schwartz (1977) and Hull and White (1990b).

2.3. Monte Carlo Simulation

The Monte Carlo simulation is used to obtain the underlying asset price at the option maturity by means of averaging a sufficiently high number of stochastic asset price paths, obtained by assuming that the underlying price follows a log-normal distribution, that is simulating L scenarios for the underlying price evolution as:

$$S_T = S_t e^{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{T-t}\Delta W_t} \quad (12)$$

where W_t denotes a standard Wiener process at time t .

After that, option prices are found by discounting that average result backwards. In other words, given the payoffs at maturity T of call and put options, respectively as:

$$C_T = \max(0, S_T - K), \quad P_T = \max(0, K - S_T) \quad (13)$$

the resulting call and put prices are obtained as an average of the L simulated scenarios, i.e.,

$$C_t = \frac{1}{L} \sum_{l=1}^L C_l, \quad P_t = \frac{1}{L} \sum_{l=1}^L P_l \quad (14)$$

where $l = 1, 2, \dots, L$.

2.4. Neural Networks to Improve Precision

Option prices dynamics depend on several variables as well as on an economic environment and rules that continuously change. Despite parametric methods mimic the behavior of real option prices, it may be argued that they do not fully reflect the actual market evolution of option prices.

To cope with that, similarly to Liang et al. (2009), this paper defines a two-step procedure in order to consistently evaluate option prices. The first step consists of pricing options according to the three parametric methods described above, i.e., tree models, finite difference method, and Monte Carlo simulation. The prices obtained in the first step are then used as input training vector of a neural network model in the second step. As a consequence, once the main information regarding an option's price are captured through the parametric methods in the first step, the machine learning neural network can concentrate its modeling power to approximate the non-linear features of the option pricing errors. A graphical representation of the model can be found in Figure 3.

It is well-known that the option market is a complex system with non-linear characteristics. This further motivates our approach, since the use of a particular kind of neural network

model, the multilayer perceptron one, allows to account for these features. Indeed, through the multilayer perceptron neural network one is able to include hidden layers and non-linear activation functions that may capture the non-linearity of the option market. An organic description of multilayer perceptron neural networks can be found, for example, in Haykin et al. (2009).

2.5. Performance Assessment

In this subsection the assessment criteria used to evaluate our models are presented. Performances of our pricing methods are judged according to three widely employed measures, i.e., the mean absolute error (MAE), mean squared error (MSE), and the mean absolute percentage error (MAPE). These criteria are defined by

$$MAE = \frac{1}{N} \sum_{n=1}^N |A_{t,n} - F_{t,n}| \quad (15)$$

$$MAPE = \frac{1}{N} \sum_{n=1}^N \left| \frac{A_{t,n} - F_{t,n}}{A_{t,n}} \right| \quad (16)$$

$$MSE = \frac{1}{N} \sum_{n=1}^N (A_{t,n} - F_{t,n})^2 \quad (17)$$

where A is the actual option value and F is the fitted value obtained by the corresponding pricing model, being t the specific time at which the option is evaluated and N the number of observations.

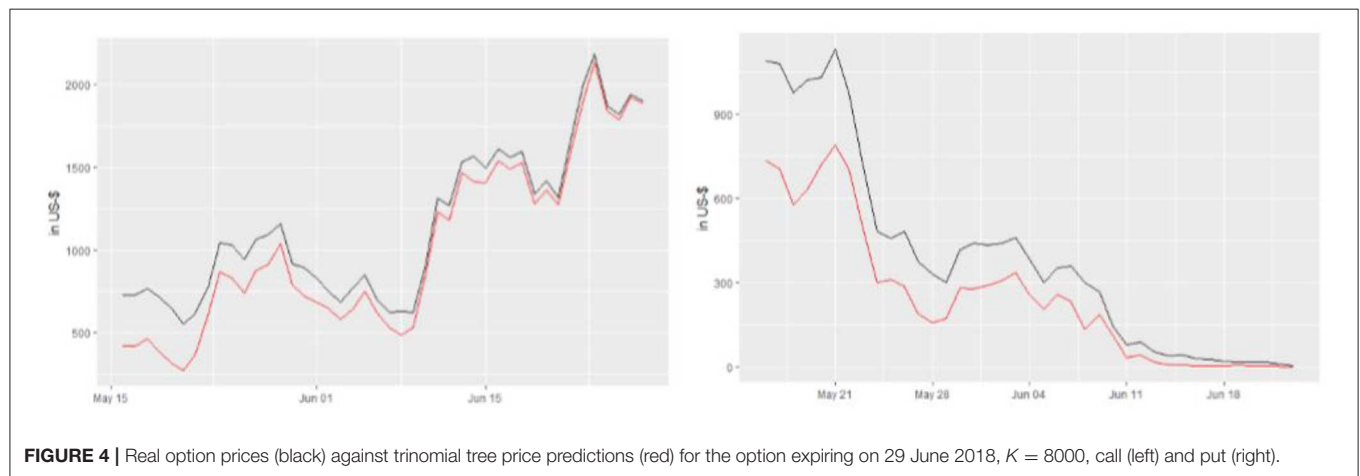
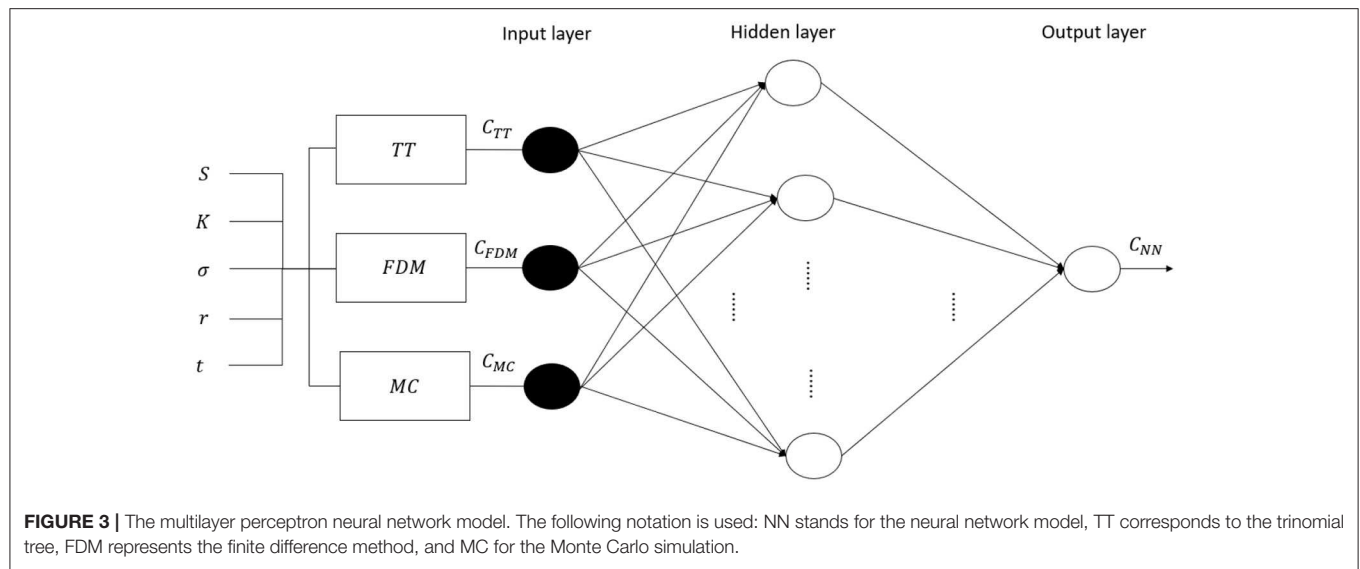
3. DATA

An option market for cryptocurrencies—and Bitcoin—is gradually emerging. I analyze data from deribit.com, a platform offering trading of futures and European style options written on Bitcoin. In particular, the corresponding underlying on which the options are written consists of the deribit BTC index².

Data are collected from 16 May 2018 to 15 July 2018, on a daily basis, every day at the same time (11:00 UTC). To be precise, the retrieved data are the deribit BTC index and all available option prices related to that day (European calls and puts).

Following Liang et al. (2009) the analysis is restricted to options having a time to maturity comprised between 5 and 20 days, as well as to in-the-money options having a spread which is lower than 50%. In this way it is possible to overcome price fluctuations related to the expiration effect and liquidity problems linked to the long term time to maturity options, as well as to eliminate outliers reflecting expectations which are somehow not rational and may heavily affect results. Furthermore, the choice of such a maturity

²Detailed information regarding the deribit BTC index can be found on www.deribit.com



range is in line with the peculiar short term feature of cryptocurrency options, whose maturities are generally smaller than the ones related to traditional option markets. To illustrate, the majority of options in our full dataset were issued only 8 days before maturity.

Given the set of restrictions adopted above, the dataset ends up with a total number of 281 call and 695 put prices. In the current analysis, the first 10 weeks will be used for the estimation purposes, while the last 2 weeks will be used for out-of-sample performance assessment.

As far as the parameter specifications, a 15-days historical volatility for the deribit BTC index and the 2-months Libor interest rate as risk-free rate are used. Moreover, the finite difference method has a grid of size $3T$ and the Monte Carlo simulation involves 10,000 repetitions.

The neural network involves several specifications, too. Firstly, the study relies on the widely spread backpropagation algorithm for the parameter estimation. Secondly, the most widely employed activation functions are tested in order to

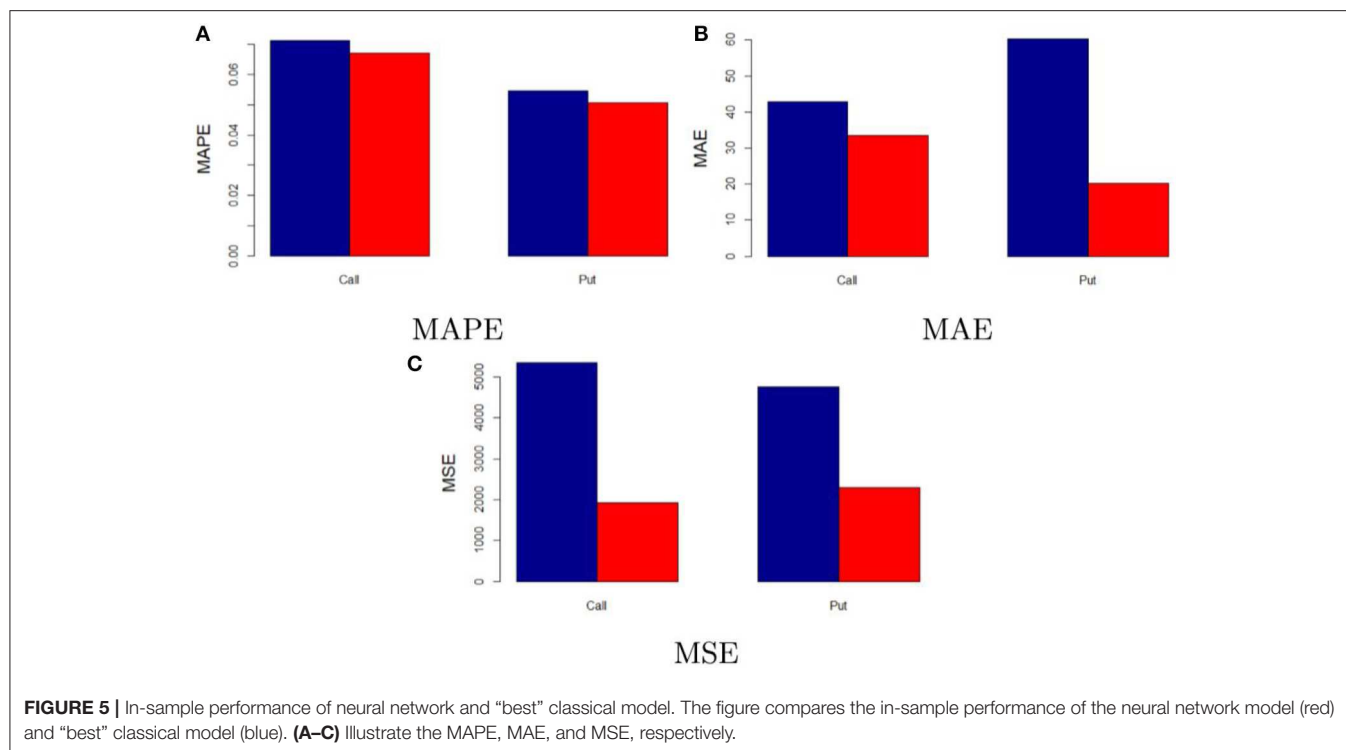
TABLE 1 | In-sample performance of neural network and classical models.

	TT	FDM	MC	NN
CALL				
MAPE	0.0713	0.0713	0.0716	0.0670
MAE	42.78	42.79	43.2	33.55
MSE	5,362.41	5,362.65	5,401.13	1,926.66
PUT				
MAPE	0.0546	0.0547	0.0546	0.0506
MAE	56.00	56.05	56.08	33.63
MSE	4,764.71	4,764.81	4,765.29	2,299.11

The following notation is used: NN represents the neural network model, TT corresponds to the trinomial tree, FDM stands for finite difference method, and MC for the Monte Carlo simulation.

choose the one ensuring the best performance in terms of fitting³. Results indicate that the sigmoid function is the one ensuring

³In particular, the following activation functions are tested: sigmoid, Taylor, identity, tanh, softplus, gauss.



the smallest sizes of prediction error. Thirdly, an analysis of the optimal number of hidden layers and neurons in the network is conducted, following the iterative procedure described in Stathakis (2009). Results suggest a model having two neurons and one hidden layer.

4. EMPIRICAL FINDINGS

4.1. Experimental Results on Selected Options

In this section results are presented distinguishing between call and put options.

Without loss of generality, a plot of a representative option price evolution against one of the parametric methods (the trinomial tree) prediction is shown in **Figure 4**. Overall, classical parametric option pricing methods (i.e., trinomial tree, finite difference method and Monte Carlo simulation) lead to price predictions which are consistently lower than the actual option prices, both in the put and the call cases. Consequently, it may be argued that options written on Bitcoin are systematically overpriced by the platform when considering the parametric methods in question. Notwithstanding this, theoretical prices yielded by parametric methods converge to the real option prices as the time to maturity becomes smaller. This is in line with the behavior of the traditional markets for option exchanges, where a small time to maturity leads to a convergence of theoretical and real option prices.

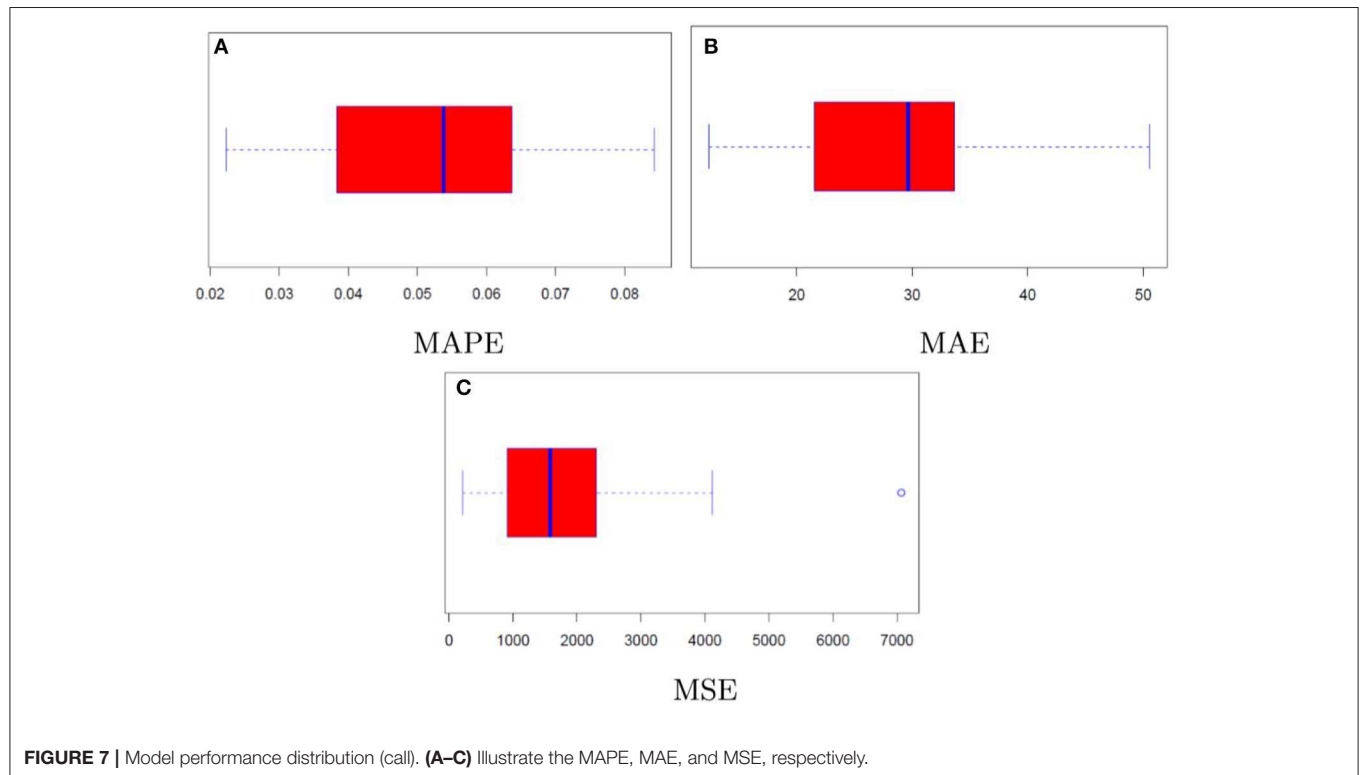
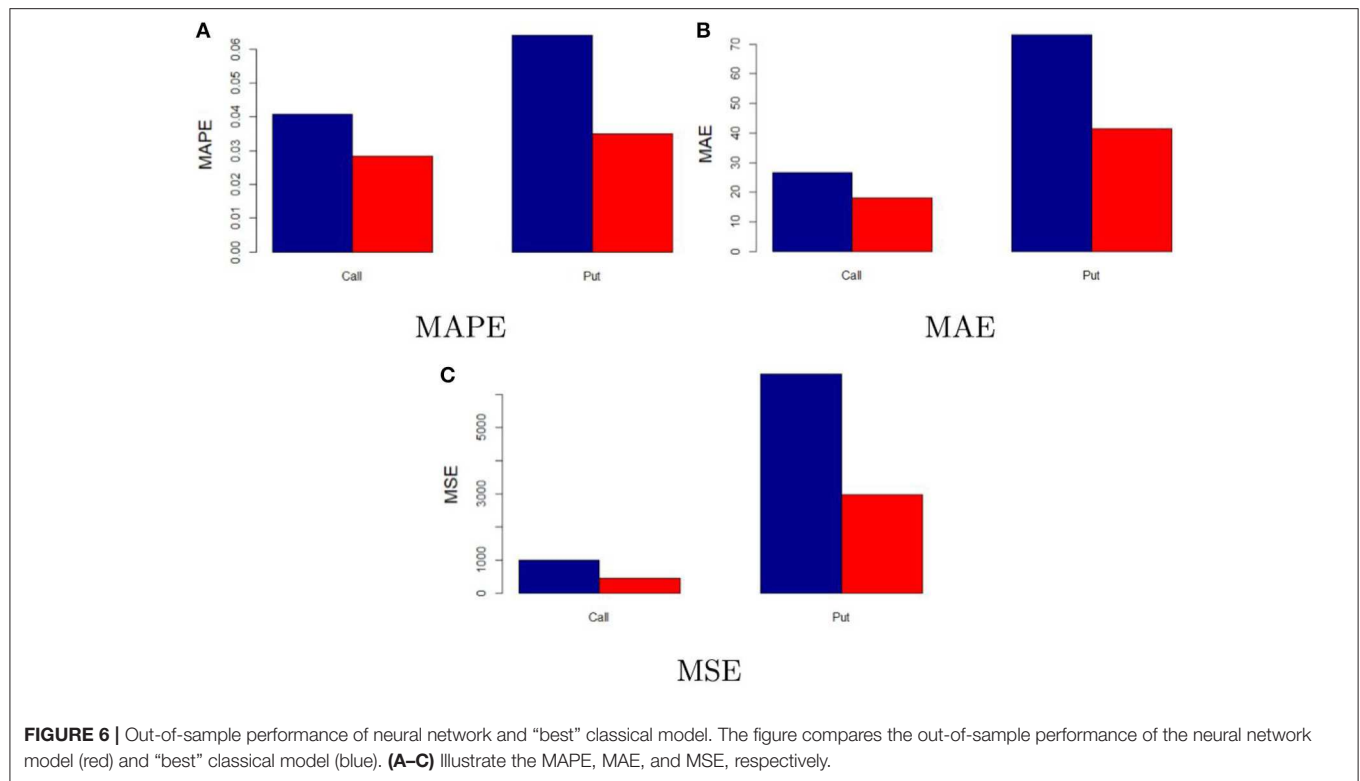
Prediction errors associated with each category of options are illustrated in **Table 1**. Absolute and relative model performance measures are quite comparable across the considered classical parametric methods. Besides that, it is clear that the neural

TABLE 2 | Out-of-sample performance of neural network and classical models.

	TT	FDM	MC	NN
CALL				
MAPE	0.0429	0.0429	0.0425	0.0283
MAE	26.64	26.65	26.77	17.93
MSE	1,016.11	1,016.28	1,026.79	441.94
PUT				
MAPE	0.0642	0.0643	0.0642	0.035
MAE	73.4	73.4	73.23	41.45
MSE	6,668.17	6,667.56	6,646.12	2,978.26

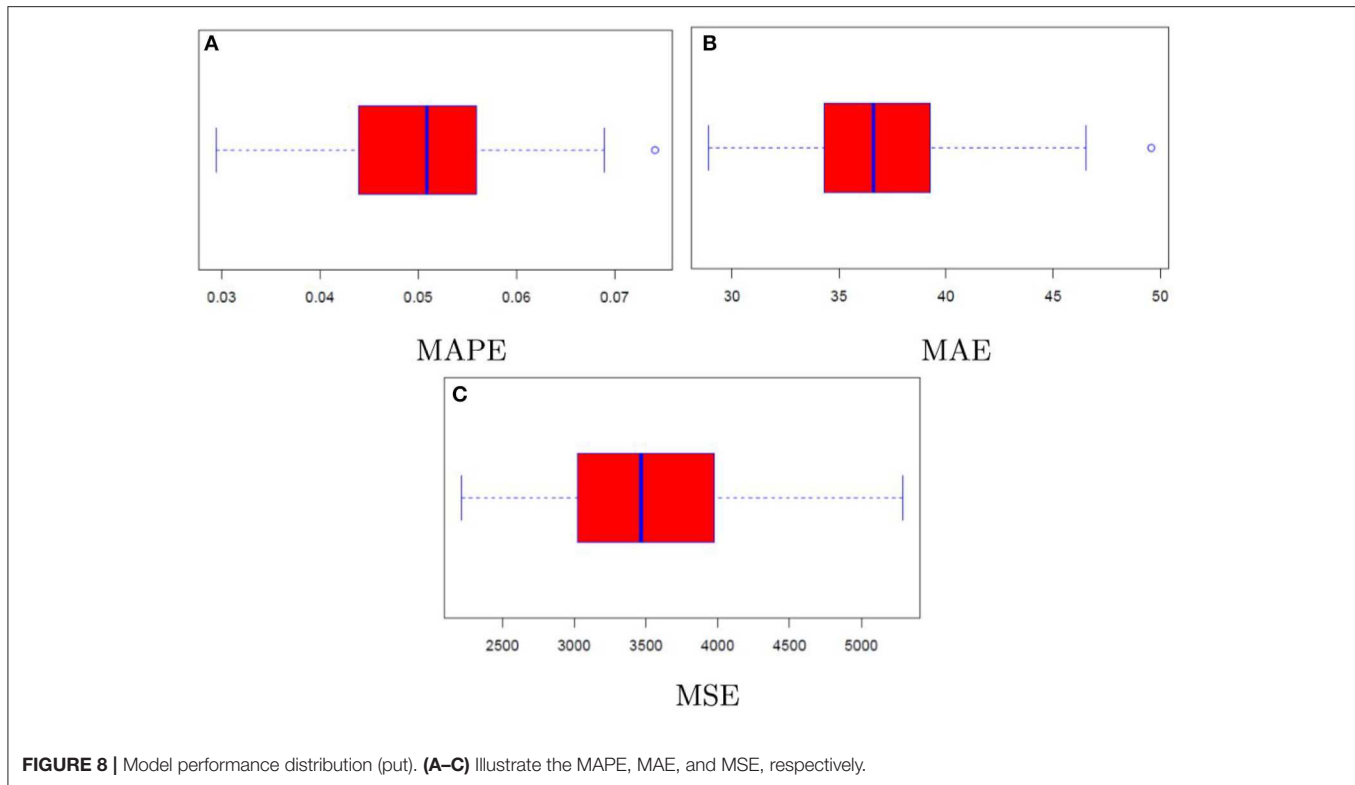
The following notation is used: NN represents the neural network model, TT corresponds to the trinomial tree, FDM stands for finite difference method, and MC for the Monte Carlo simulation.

network outperforms them in terms of prediction accuracy. This is also graphically represented in **Figure 5**, which shows the model performance metrics of the neural network against those of the “best” classical model, meaning the parametric model among the ones used in this study showing the lowest prediction error. To illustrate, when comparing the neural network and the “best” classical model performances the MAPE lowers by 6% in the call case and 7.33% in the put one, the MAE by 21.58% (call) and 0.4% (put) as well as the MSE by 64.07% (call) and 51.75% (put). This is mainly due to the fact that the multilayer perceptron neural network can deal with the complexity and non-linearity of the option market and the cryptocurrency market. Indeed, price predictions yielded in the first step by the conventional approach are then refined into the second step by the neural network, which focuses on lowering the errors existing between the real option prices and the predicted ones.



The obtained results are in accord with the existing literature on option pricing through non-parametric methods and, particularly, neural networks—see Hutchinson et al. (1994), Malliaris and Salchenberger (1996), Amilon (2003), Binner et al.

(2005), and Lin and Yeh (2005). Indeed, all these studies point to an overall predominance of neural network based models in pricing options with respect to conventional methodologies. It may be argued that this holds true also for particular markets



like the cryptocurrency one, whose particular features are well-captured by non-parametric models, such as the neural network.

5. ROBUSTNESS ANALYSIS

With the aim of testing the robustness of our model, this section provides an out-of-sample performance analysis as well as a cross-validation analysis through repeated random sub-sampling.

5.1. Out-of-Sample Performance

The out-of-sample performance is tested on the options available on the deribit platform between 1 August 2018 and 15 August 2018. Options are selected according to the same criteria described in section 3. The final out-of-sample dataset consists of 29 call and 47 put option prices.

Results of the out-of-sample performance of the investigated models are illustrated in **Table 2**. At a first glance, one may notice that results linked to both absolute and relative performances change quite consistently. This is mainly due to the different structure of the out-of-sample dataset, in particular to the different maturities and market expectations.

As also depicted in **Figure 6**, it is clear that the neural network model proposed still outperforms the considered parametric methods. In addition, the difference in performance is even higher than the in-sample one. When comparing the performance of the neural network and the “best” classical model, the MAPE lowers by 33.41% in the call case and 45.48% in the put one, the MAE by 32.7% (call) and 43.4% (put) as well as the MSE by 55.23% (call) and 55.06% (put). This provides further support

to the fact that the neural network is a feasible model to price Bitcoin options.

5.2. Cross-Validation

To further assess the robustness of our proposed model, the approach of repeated random sub-sampling for cross-validation purposes is adopted. In other words, the dataset is randomly split into training and validation set for 50 times and then the methodology and procedures described in this study are repeated. In this way one is able to determine whether the neural network performance achieved in the results section are stable, as well as to evaluate the model’s relative performance after random sub-sampling with respect to the conventional option pricing methods.

Results linked to the random sub-sampling procedure are illustrated through the boxplots contained in **Figure 7** (call case) and **Figure 8** (put case). Overall, outcomes are satisfactory provided that performance variability lies in ranges which are arguably not too wide. To illustrate, the interquartile ranges for MAPE and MAE are respectively <3% and below 10 USD in the call case, whereas in the put case they amount to roughly 1% and 5 USD.

Furthermore, comparing the distributions of the assessment criteria with the results in **Table 1**, it may be noticed that even in the context of resampling the neural network achieves again satisfactory results in terms of precision. Indeed, despite the MAPE results coming from the repeated random sub-sampling are partly worse than those of classical option pricing methods, the absolute assessment criteria still point to a substantial improvement when considering the neural network model rather than the conventional option pricing methods.

To conclude, there may be room for improvement in the modeling strategy, as well as this needs to be adapted to the specific case of interest. As an example, it can be argued that the neural network performances would benefit from increasing the number of observations and, specifically, by using high frequency data. In addition, as the market is highly volatile and the option market follows fast changing rules and patterns, different choices of the neural network specifications—different input layers, structure of the layers, activation functions, etc.—may result more feasible in other contexts. Nevertheless, it may be claimed that the multilayer perceptron neural network model proposed is suitable for pricing options written on Bitcoin. Moreover, it may be argued that its application can be extended to the whole cryptocurrency framework, as well as to traditional markets.

6. CONCLUSION

This paper proposes an approach that relies on artificial neural network models for the purpose of Bitcoin option pricing. The methodology involves a first step in which options are priced according to some of the most widely employed parametric methodologies, i.e., tree models, Monte Carlo simulation, and finite difference method. The option prices obtained in this way are then used as input layers in a second step by the neural network, which is capable to refine the price predictions delivered by the parametric models in the first step. We believe

that the proposed model can be extended, without loss of generality, to other cryptocurrency derivatives, as well as to traditional ones.

Empirical results show that the investigated conventional pricing methodologies yield to the conclusion that Bitcoin options are extensively overpriced. In contrast, by applying the proposed neural network model one is able to better represent the real market dynamics of Bitcoin option prices. Indeed, prediction errors consistently reduce when comparing the neural network pricing model to the classical parametric ones.

Further studies may benefit and improve prediction precision by using high frequency data as well as different model specifications. As an example, improvements could be achieved by the use of different models, such as stochastic volatility models, as input layers in the proposed neural network framework.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

I heartily thank Matthias for the time he spent on the insightful discussions we had toward this common interesting topic. I also thank Dina for her great help in getting the dataset.

REFERENCES

- Amilon, H. (2003). A neural network versus black-scholes: a comparison of pricing and hedging performances. *J. Forecast.* 22, 317–335. doi: 10.1002/for.867
- Binner, J. M., Bissoondeal, R. K., Elger, T., Gazely, A. M., and Mullineux, A. W. (2005). A comparison of linear forecasting models and neural networks: an application to Euro inflation and Euro Divisia. *Appl. Econ.* 37, 665–680. doi: 10.1080/0003684052000343679
- Black, F., and Scholes, M. (1973). The pricing of options and corporate liabilities. *J. Political Econ.* 81, 637–654.
- Boyle, P. P. (1977). Options: a Monte Carlo approach. *J. Financ. Econ.* 4, 323–338.
- Brandvold, M., Molnár, P., Vagstad, K., and Andreas Valstad, O. C. (2015). Price discovery on bitcoin exchanges. *J. Int. Financ. Market. Inst. Money* 36, 18–35. doi: 10.1016/j.intfin.2015.02.010
- Brennan, M. J., and Schwartz, E. S. (1977). Finite difference methods and jump processes arising in the pricing of contingent claims: a synthesis. *J. Financ. Quant. Anal.* 12:659.
- Cox, J. C., Ross, S., and Rubinstein, M. (1979). Option pricing: a simplified approach. *J. Financ. Econ.* 7, 229–263.
- Dwyer, G. P. (2015). The economics of bitcoin and similar private digital currencies. *J. Financ. Stabil.* 17, 81–91. doi: 10.1016/j.jfs.2014.11.006
- Haykin, S. S., Haykin, S. S., Haykin, S. S., Elektroingenieur, K., and Haykin, S. S. (2009). *Neural Networks and Learning Machines*, vol. 3. Upper Saddle River, NJ: Pearson.
- Hull, J., and White, A. (1990a). Pricing interest-rate-derivative securities. *Rev. Financ. Stud.* 3, 573–592.
- Hull, J., and White, A. (1990b). Valuing derivative securities using the explicit finite difference method. *J. Financ. Quant. Anal.* 25:87.
- Hutchinson, J. M., Lo, A. W., and Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *J. Finance* 49, 851–889.
- Kroll, J. A., Davey, I. C., and Felten, E. W. (2013). “The economics of bitcoin mining, or bitcoin in the presence of adversaries,” in *Proceedings of WEIS*.
- Liang, X., Zhang, H., Xiao, J., and Chen, Y. (2009). Improving option price forecasts with neural networks and support vector regressions. *Neurocomputing* 72, 3055–3065.
- Lin, C.-T., and Yeh, H.-Y. (2005). The valuation of taiwan stock index option price: comparison of performances between black-scholes and neural network model. *J. Stat. Manag. Syst.* 8, 355–367. doi: 10.1080/09720510.2005.10701164
- Malliaris, M., and Salchenberger, L. (1996). Using neural networks to forecast the s&p 100 implied volatility. *Neurocomputing* 10, 183–195.
- Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Available online at: www.bitcoin.org
- Nelson, D. B., and Ramaswamy, K. (1990). Simple binomial processes as diffusion approximations in financial models. *Rev. Financ. Stud.* 3, 393–430.
- Pagnottoni, P., and Dimpfl, T. (2019). Price discovery on bitcoin markets. *Dig. Financ.* 1, 1–23. doi: 10.1007/s42521-019-00006-x
- Rendleman, R. J., and Bartter, B. J. (1979). Two-state option pricing. *J. Finance* 34, 1093–1110.
- Stathakis, D. (2009). How many hidden layers and nodes? *Int. J. Remote Sens.* 30, 2133–2147. doi: 10.1080/01431160802549278
- White, L. H. (2015). The market for cryptocurrencies. *Cato J.* 35, 383–402. doi: 10.2139/ssrn.2538290
- Yao, J., Li, Y., and Tan, C. L. (2000). Option price forecasting using neural networks. *Omega* 28, 455–466. doi: 10.1016/S0305-0483(99)00066-3

Conflict of Interest Statement: The 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 © 2019 Pagnottoni. 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.



Artificial Intelligence (AI) in the Financial Sector—Potential and Public Strategies

Stephan Bredt*

Ministry of Economics, Energy, Transport and Housing of the State of Hessen, Wiesbaden, Germany

OPEN ACCESS

Edited by:

Paolo Giudici,
University of Pavia, Italy

Reviewed by:

Shatha Qamhieh Hashem,
An-Najah National University, Palestine
Bertrand Kian Hassani,
University College London,
United Kingdom

*Correspondence:

Stephan Bredt
stephanbredt@gmx.de

Specialty section:

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

Received: 04 March 2019

Accepted: 19 August 2019

Published: 04 October 2019

Citation:

Bredt S (2019) Artificial Intelligence (AI)
in the Financial Sector—Potential and
Public Strategies.
Front. Artif. Intell. 2:16.
doi: 10.3389/frai.2019.00016

AI is providing a significant basis for future technological innovation. The financial sector will be transformed by AI, offering the opportunity for better and more tailor-made services, cost reduction, and the development of new business models. The Federal and the Hessen governments recently published roadmaps for the further development of AI in Germany and Hessen, respectively. The Federal Government will invest three billion euros over the next 5 years in a variety of research and business sectors whereas the State of Hessen will set up a new AI-oriented institute of applied research and business development and spend one billion euros over the next 5 years on digitalization development. The public strategies for building AI hubs are still extremely diverse. However, the focus is on stronger application of research results in business activities, on increasing networks and ecosystems and predominantly on building on existing centers of excellence. The Frankfurt Rhein Main Region, already a strong hub for fintech, cyber security, and AI, will especially benefit from these programs. The financial center Frankfurt offers a vivid and fast growing tech and start up community as well as an academic and data infrastructure unprecedented in Europe: the largest data and cloud service hub in continental Europe, the world's largest internet knot, universities, research institutes with global quality research in AI, as well as companies and consultancies specialized in AI and neighboring areas such as fintech and cyber security.

Keywords: artificial intelligence, public strategies, financial sector, national strategy for artificial intelligence, artificial intelligence and market surveillance, artificial intelligence ecosystem frankfurt rhein main, start up ecosystem frankfurt rhein main

INCREASING IMPORTANCE OF AI FOR SOCIETIES AND THE ECONOMY

AI is recognized as a combination of new technologies, processes, and methods with an increasing importance for the current and future development of our societies and economies. AI is applied today in various diverse sectors such as medical diagnostics, optical character recognition, automotive autonomous driving, and financial services. Already today large corporates and small and medium enterprises in Germany use AI technologies. AI has become a part of daily life for millions of consumers. The application of AI is seen as a potential driver of disruptive technological development and innovation.

That is why recently a flood of studies and reports by research and public institutions as well as consultancies has been published to clarify the opportunities, challenges, and steps ahead (Schwab K., World Economic Forum, 2016, The fourth industrial revolution: what it means, how to respond, 2016; Bafin, Big Data trifft auf künstliche Intelligenz, Herausforderungen und Implikationen für Aufsicht und Regulierung von Finanzdienstleistungen, 2018; BMBF, Die Revolution der Künstlichen Intelligenz, 2018).

Estimates calculate that productivity in Germany could be increased by 29% until 2035 (Accenture, Why AI is the Future of Growth, 2016). Investment in AI is increasing sharply, in start ups as well in existing companies. It is calculated that in 2017, about 17 billion euros were invested globally in AI technology (Accenture, Weg ohne Ziel? Wie Deutschland ein Spitzenstandort für Künstliche Intelligenz werden kann, 2018). Additionally AI will have an increasing impact on financial services. It is estimated that by 2035, banks could improve their productivity annually by 4.3%¹. In financial services business, AI could transform the financial sector in the following three aspects.

First, AI could improve the quality of products and service for clients due to a broader and deeper analytical basis and information. Second, AI could lead to higher efficiency and lower costs, e.g., in the area of compliance and fraud detection or anti-money laundering measures. Furthermore, public institutions like Financial Market or Tax Supervisory authorities could benefit from AI technology in that sense (Bafin, Big Data trifft auf künstliche Intelligenz, Herausforderungen und Implikationen für Aufsicht und Regulierung von Finanzdienstleistungen, 2018). Third, AI could become a central innovation driver. Although it is not quite clear yet what the financial service provider of tomorrow will look like, it seems probable that AI will transform financial service providers into data- and AI-based businesses (Accenture, Hessens Ambitionen für Künstliche Intelligenz, 2018, p. 11s.).

In Germany, the Federal Government and the State of Hessen government have set up strategic road maps to develop AI made in Germany and AI made in Hessen. These roadmaps are first instruments to shape the development of AI in the financial sector and beyond. They must be embedded in broader public strategies such as development of research and the creation of an innovative environment for start ups and incumbents.

STRENGTHENING THE TECH AND AI ECOSYSTEM

The Hessen Government has analyzed the situation of the Frankfurt Metropolitan Region for its status and prospects as a Tech and AI Hub in recent months and years and developed a road map, and it has taken steps to develop it along this roadmap. After the set up of the TechQuartier in 2016, a Masterplan for the

region as a Start Up hub was developed in 2017 (Techquartier, Masterplan Start-Up Region Rhein Main, 2018). Subsequently strengths and opportunities of the region were analyzed by the Frankfurt ecosystem report by Startup Genome in 2018. Finally, an analysis by Accenture on the status and opportunities of the region in the AI sector with concrete proposals was presented in 2018. These reports are the basis of the following section.

The Environment: The Tech Ecosystem in Frankfurt Metropolitan Region

Alongside the public universities with more than 100,000 students, more than 20 private universities and other institutions of higher research, Frankfurt boasts world-class research organizations such as the Fraunhofer Institutes, Leibniz Association, and Max Planck Institutes, among others. The region's 22 research institutions are responsible for breakthrough research, innovative products, and new processes. Academic research in Frankfurt is complemented by corporate research and development (R&D) in production, life sciences, robotics, and artificial intelligence. Corporate R&D spending reached 5.5 billion euros in 2017 (Startup Genome, Frankfurt Startup Ecosystem Report, 2018, p. 16).

The Frankfurt Metropolitan Tech ecosystem can be characterized by three sub-sectors for which the region has potential to build global competitiveness and economic value. These are closely integrated, with alignment across talent and the types of problems that startups are addressing: 1. Fintech, 2. AI, Big Data & Analytics, and 3. Cybersecurity.

Whereas Fintech is characterized in the region by its dynamic start up environment, the exceptional role of the region in cyber security is shaped by research institutions. The Center for Research in Security and Privacy (CRISP) in Darmstadt alone is host to more than 450 researchers in this sector and is complemented by research at the universities and the Fraunhofer Institute. Recently, Chancellor Merkel announced that these institutions will be supported by the Federal Government and be developed as the national cyber security hub.

The creation of the TechQuartier by the State of Hessen with the support of numerous financial service providers in 2016 provided a fast growing platform an light house and most important ecosystem for start ups, businesses, and researchers. The TechQuartier (TQ) is the central platform for the startup community in the Frankfurt Metropolitan Region, enriching the vibrant startup ecosystem with its unique community of more than 100 start ups and 30 corporate partners and academic institutions. Nearly 400 tech start ups are now active in the Frankfurt region, and recent exits and funding rounds have helped drive total ecosystem value to \$1.8 billion. These start ups enjoy a supportive environment that includes 32 incubators, 24 coworking spaces, and 10 accelerators.

In 2018, the so called Masterplan Start-Up Region Frankfurt Rhein-Main, developed by TQ and partners was presented. It endorses an ambitious strategy to develop Frankfurt Metropolitan Region as the leading FinTech-Hub in continental Europe and home to 1,000 start ups in 2022.

¹Accenture, How AI Boosts Industry Profits an Innovation, 2017. Available online at: https://www.accenture.com/fr-fr/_acnmedia/36dc7f76eb444cab6a7f44017cc3997.pdf

Frankfurt's multinational corporations operate several programs to support start ups and, increasingly, these companies are investing more into early-stage companies. This has helped Frankfurt achieve one of the fastest growth rates in early-stage funding in the world in recent years. As a center of global finance—home to the European Central Bank and several international banking headquarters—this corporate support has led to one of the world's strongest clusters of Fintech startups. Frankfurt's leading Fintech sub-sector has been catalyzed by financial support (over half of venture capital investment in the region has gone into Fintech startups in the last 5 years) and the \$800 million acquisition of 360T in 2015 by Deutsche Börse (See Startup Genome, Frankfurt Startup Ecosystem Report, 2018, Startup Genome, p. 5). However, there are still challenges to the Frankfurt Start Up Ecosystem:

- Over the past few years, Frankfurt has enjoyed one of the world's fastest growth rates in early-stage funding. Yet as startup output rises, that growth needs to continue and even increase, meaning that Frankfurt will need millions of additional dollars per year in early-stage funding (European Parliament, 2018, p. 6).
- Frankfurt has several areas of latent entrepreneurial talent that need to be better activated (European Parliament, 2018, p. 6).
- Frankfurt currently faces gaps in scaleup creation. The exception is the extent to which founders in Frankfurt report the ways they build Global Connectedness (such as meeting founders from international ecosystems locally). In January 2019, TQ, Goethe University and Yi Shi Foundation published a report that addresses this problem and indicates which instruments and strategies could help to overcome this problem (TechQuartier, Goethe University and Yi Shi Foundation for innovations, Scale-Ups In Europe, an untapped potential, 2019).
- The gaps in Frankfurt's startup ecosystem relate to the global orientation of the ecosystem and experience levels (Ecosystem Report p. 39).

The AI Ecosystem

The AI research in the Frankfurt Metropolitan region has been classified as competitive on a global level (Accenture, Hessens Ambitionen für Künstliche Intelligenz, 2018, p. 15).

The Technical University Darmstadt belongs to the most important universities globally in informatics. Research at TU Darmstadt covers the full scale of AI research (Machine Learning, Reinforcement Learning, Deep Learning, Supervised und Unsupervised Learning, Computer Vision, NLP, Robotik, Predicative Systems). The Technical University Darmstadt runs an Autonomous Systems Lab for Machine Learning for Intelligent Systems and Robotics with research centered around the goal of bringing advanced motor skills to robotics using techniques from machine learning and control (European Parliament, 2018, p. 22).

Additionally the Frankfurt Big Data Lab Start-up Program at Goethe University offers general training courses for data computation and analytics by startups. Frankfurt School of Finance & Management runs a Center for Human and Machine

Intelligence (HMI), which conducts basic and applied research at the intersection of artificial intelligence and machine learning, decision and social science, and finance and management. Just recently, January 2019, the Frankfurt School opened their AI Lab as a place for testing, learning, and developing AI-based ideas and strategies.

Also on the business and start up side, there are growing activities: Some 8.5% of all startups in Frankfurt are in the Artificial Intelligence or Big Data & Analytics sub-sector and, over the past 5 years, the sub-sector captured 13% of all local VC investment (European Parliament, 2018 p. 22).

Still, there are challenges that need to be addressed: more AI talents need to be educated. There are not enough and additionally they do rarely move to finance professionally (Accenture, Hessens Ambitionen für Künstliche Intelligenz, 2018). The number of students in informatics in Hessen increased in the last 10 years by ca. 73% to 1,897 in 2016/2017. However, still only 0.8% of all students in Hessen study informatics (without mixed curricula such as "Wirtschaftsinformatik"). The top employers of graduates from German universities in informatics are Google (25.2%), BMW Group (10.6%), Microsoft (10.5%), Apple (9.9%), and SAP (9.7%). Deutsche Bank is the first financial service provider in that list and is ranked 53 (1.3%).

PUBLIC STRATEGIES FOR SUPPORTING AI IN GERMANY AND HESSEN

The National AI Strategic Report

Germany's national government published its AI Strategic Report in December 2018. The strategy is broad in both focus of industries and technologies as well as in instruments to strengthen AI in Germany (Bundesregierung, Strategie Künstliche Intelligenz der Bundesregierung, 2018). The financial sector is included beside many other business areas (Bundesregierung, p. 25). The instruments mentioned are strengthening the startup environment in AI, building on existing instruments like the Digital Hub Initiative of the Federal Government, the creation of new institutions like the Agency for Disruptive Technologies, 12 centers and "application hubs," the expansion of venture capital offering, extended research (100 professor positions), and the creation of academic networks (Bundesregierung, p. 6). Industry-supported or -led initiatives can also be eligible for support.

The program will support institutions nationwide that already are focused on AI technology such as the Deutsches Zentrum für Künstliche Intelligenz or Fraunhofer Institutes and specialized universities. The federal government will cooperate closely with the federal states for an effective execution of the program.

The State of Hessen and especially the Frankfurt Metropolitan Region will potentially benefit from that national program. The TechQuartier and the Digital Hub in Darmstadt are both included in the Federal Digital Hub Strategy and are therefore potentially eligible for financial assistance. Also, the abovementioned inclusion of the financial sector and its supervisory institutions offers widespread opportunities for

projects to be supported by the national program. The Darmstadt Digital Hub, also part of the Federal Digital Hub Strategy and focused on cyber security, could also benefit from the program. Besides, there are points of contact with national programs that are already enrolled in the Frankfurt Metropolitan Region such as cyber security and so called “Mittelstandsförderung für Digitalization.”

It needs to be decided in the future where exactly some of the three billion euros to be spent by the federal government until 2015 will be invested. The idea exposed in the national program is that this support will be leveraged by investment of academic, commercial, or other public institutions such as the federal states (Bundesregierung, p. 6).

The program has been welcomed and criticized. Welcomed, because it offers for the first time a more comprehensive view and action plan to this important topic, and because it is including concrete approaches and resources. It was criticized because it is missing a kind of focal point but is comprising potentially too many institutions and topics and therefore losing traction and visibility. However, with its decentralized approach, it can be stated that the strategy paper corresponds very well with the German academic and economic national institutional set up. It is left to the competition and efficiency of several academic and economic institutions and players where lighthouses of AI will develop. Still, it would be an advantage and should be a target to develop a globally renowned top institution for AI. For the moment the German Center for AI in Saarbrücken and Kaiserslautern is seen as such an institute, which will also benefit from the national strategy. The national strategy could potentially and adequately strengthen this and other players.

The State of Hessen AI and Tech Strategy

Parallel to the creation of the national strategy for AI Made in Germany, the Hessen State government decided to build up an AI hub in the Frankfurt Metropolitan Region in August 2018. This decision was supported by an analysis of the AI capacities of the Frankfurt Metropolitan Region by Accenture for the State of Hessen (Accenture, *Hessens Ambitionen für Künstliche Intelligenz*, 2018). Before, an analysis from Startup Genome had already encouraged increasing the support for the development of AI start ups in the Frankfurt Metropolitan Region (European Parliament, 2018). The core finding and then proposal of the Accenture report was that there are already high quality research and business activities in the region but that sufficient interconnectedness was missing. Also, a finding was that research results should be more effectively introduced into business activities.

These findings and corresponding proposals and measures were then inscribed in the Coalition Treaty 2018 in Hessen. A specific focus in the coalition treaty is on the development of AI (Koalitionsvertrag zwischen CDU Hessen und Bündnis90/Die Grünen Hessen für die 20. Legislaturperiode, 2018 p. 178s). It includes the creation of a tech campus with 20 professor positions and is supposed to overcome the shortcomings found by the Accenture analysis concerning interconnectedness in the AI ecosystem in the Frankfurt Metropolitan Region.

The tech campus is supposed to strengthen applied research in AI and deliver a growing number of coders and IT specialists for a

growing AI economy. It seems open to decision currently, which kind of institution the TechCampus will develop into. There exist several successful tech campuses in Germany which could serve as a role model: the CDTM in Munich, the Code University in Berlin, and the Hasso Plattner Institute at Potsdam University. Other federal states and cities in Germany have also published plans to develop such campuses, e.g., the states of Hamburg and North Rhine Westphalia.

Already now, AI activities have been intensified by activities of TechQuartier and industry partners concerning the ecosystem, comprising accelerator programs and seminars. The next step under preparation by the Ministry of Economics together with TQ and industry partner is to make use of the outstanding data infrastructure the region provides: national and Europe-wide data are available with the federal statistics office in Wiesbaden, the Bundesbank research center in Frankfurt, the Goethe University participating in a Europe wide data project on financial data going back to the nineteenth century, the ECB, and EIOPA und Bafin collecting financial data broadly and in depth. Besides, Frankfurt is home to the continent's largest offers in cloud services and data centers, of commercial financial data providers such as Deutsche Börse and Schufa. It will be an opportunity and challenge to make use of these data pools for AI purposes. The idea is to set up or open platforms as far as possible for start ups and new technologies and applied research as the provision of sufficient data is understood as the most relevant basis of AI applications. There will be public and private interest in projects to be developed on these platforms: financial market supervision instruments for the supervisory mentioned above and based in Frankfurt or AI-based tools for business processes in diverse industries. Moreover, university labs could offer access for students to such data pools.

Legal restrictions in the EU and Germany are simultaneously both chance and impediment: chance, because a safe and reliable legal environment attracts data providers and companies outsourcing data, and impediment because a broad use of these data is still often prohibited; some may only be used for academic research, others may not be combined with other specific data. In general, it will be a challenge to define broad limits for the outsourcing for companies' data. The security of data stored in cloud service could be improved and enhanced for this purpose. An international cloud provider could potentially offer security standards to be fully controlled by the outsourcing company.

The coalition treaty includes also the decision to strengthen the access to venture capital for AI start ups (Koalitionsvertrag, p. 175). It was decided to set up a specialized fund with a volume with up to 200 million euros contributed to publicly and privately equally. It was also decided to invest generally 1 billion euros for “digitalization” measures and programs, concerning public institutions, infrastructure and business development. Besides, the existing structures such as the TQ are being focused more on AI-related projects and startup programs.

CONCLUSION

In 2018, public strategies and programs for the development of AI have leaped forward significantly (Basel Committee on

Banking Supervision (BCBS), 2018; European Commission, 2018; European Parliament, 2018). Germany and the State of Hessen are investing significant resources to strengthen their already highly competitive AI ecosystems, research, and technology. Other federal states have already set up dedicated technology innovation hubs or are currently planning to do so. The federal AI program will strengthen cooperation of national and state programs and hub development, building on existing centers of excellence. Several analyses have found that the startup ecosystem in Frankfurt Metropolitan Region

is fast developing as an early stage ecosystem, and is offering a high potential for development in AI. After the path has been laid with the national and the Hessen AI strategies, the years to come require now efficient execution of these plans and programs.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

REFERENCES

- Accenture, *Hessens Ambitionen für Künstliche Intelligenz, Ein Beitrag zur nationalen KI-Strategie am Beispiel des Finanzsektors* (2018). Available online at: https://wirtschaft.hessen.de/sites/default/files/media/hmwvl/20180925_ki_studie_hessen_report_final_im_auftrag_von_0.pdf
- Accenture, *Weg ohne Ziel? Wie Deutschland ein Spitzenstandort für Künstliche Intelligenz werden kann* (2018). Available online at: https://www.accenture.com/_acnmedia/pdf-90/accenture-weg-ohne-ziel-studie.pdf
- Accenture, *Why AI is the Future of Growth* (2016). Available online at: https://www.accenture.com/t20170524t055435_w_/ca-en/_acnmedia/pdf-52/accenture-why-ai-is-the-future-of-growth.pdf
- Bafin, *Big Data trifft auf künstliche Intelligenz, Herausforderungen und Implikationen für Aufsicht und Regulierung von Finanzdienstleistungen* (2018). Available online at: https://www.bafin.de/SharedDocs/Downloads/DE/dl_bdai_studie.pdf?__blob=publicationFile&v=3
- BMBF, *Die Revolution der Künstlichen Intelligenz* (2018). Available online at: <https://www.bmbf.de/de/die-revolution-der-kuenstlichen-intelligenz-6355.html>
- Bundesregierung, *Strategie Künstliche Intelligenz der Bundesregierung* (2018). Available online at: https://www.bmbf.de/files/Nationale_KI-Strategie.pdf
- Koalitionsvertrag zwischen CDU Hessen und Bündnis90/Die Grünen Hessen für die 20. Legislaturperiode (2018). Available online at: https://www.hessen.de/sites/default/files/media/staatskanzlei/koalitionsvertrag_20_wahlperiode.pdf
- Schwab K., *World Economic Forum, 2016, The fourth industrial revolution: what it means, how to respond* (2016). Available online at: <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>
- Startup Genome, *Frankfurt Startup Ecosystem Report, Driving Growth and Scaleups with Universities and Corporations* (2018). Available online at: <https://creativehubfrankfurt.de/wp-content/uploads/2018/06/Frankfurt-Startup-Ecosystem-Report-2018.pdf>; <https://startupgenome.com/reports>
- TechQuartier, *Goethe University and Yi Shi Foundation for innovations, Scale-Ups In Europe, an untapped potential* (2019). Available online at: <https://scaleupsineurope.techquartier.com/requestaccess>; <https://techquartier.com/scaleupsineurope-pressrelease/>
- Techquartier, *Masterplan Start-Up Region Rhein Main* (2018). Available online at: https://techquartier.com/wp-content/uploads/2018/02/masterplan_public.pdf
- Basel Committee on Banking Supervision (BCBS) (2018). *Sound Practices: Implications of Fintech Developments for Banks and Bank Supervisors*. Bank for International Settlements. Available online at: <https://www.bis.org/bcbs/publ/d431.pdf>
- European Commission, Joint Research Center (2018). *Artificial Intelligence: A European Perspective*. Available online at: <https://ec.europa.eu/jrc/en/publication/eurscientific-and-technical-research-reports/artificial-intelligence-european-perspective>
- European Parliament, Think Tank (2018). *European Artificial Intelligence (AI) Leadership, the Path for an Integrated Vision*. Available online at: [http://www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_STU\(2018\)626074](http://www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_STU(2018)626074)

Conflict of Interest: The 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 © 2019 Bredt. 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^{1*}, Martin Schüle² and Martin Hillebrand³

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

OPEN ACCESS

Edited by:

Dror Y. Kenett,
Johns Hopkins University,
United States

Reviewed by:

Apama 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.

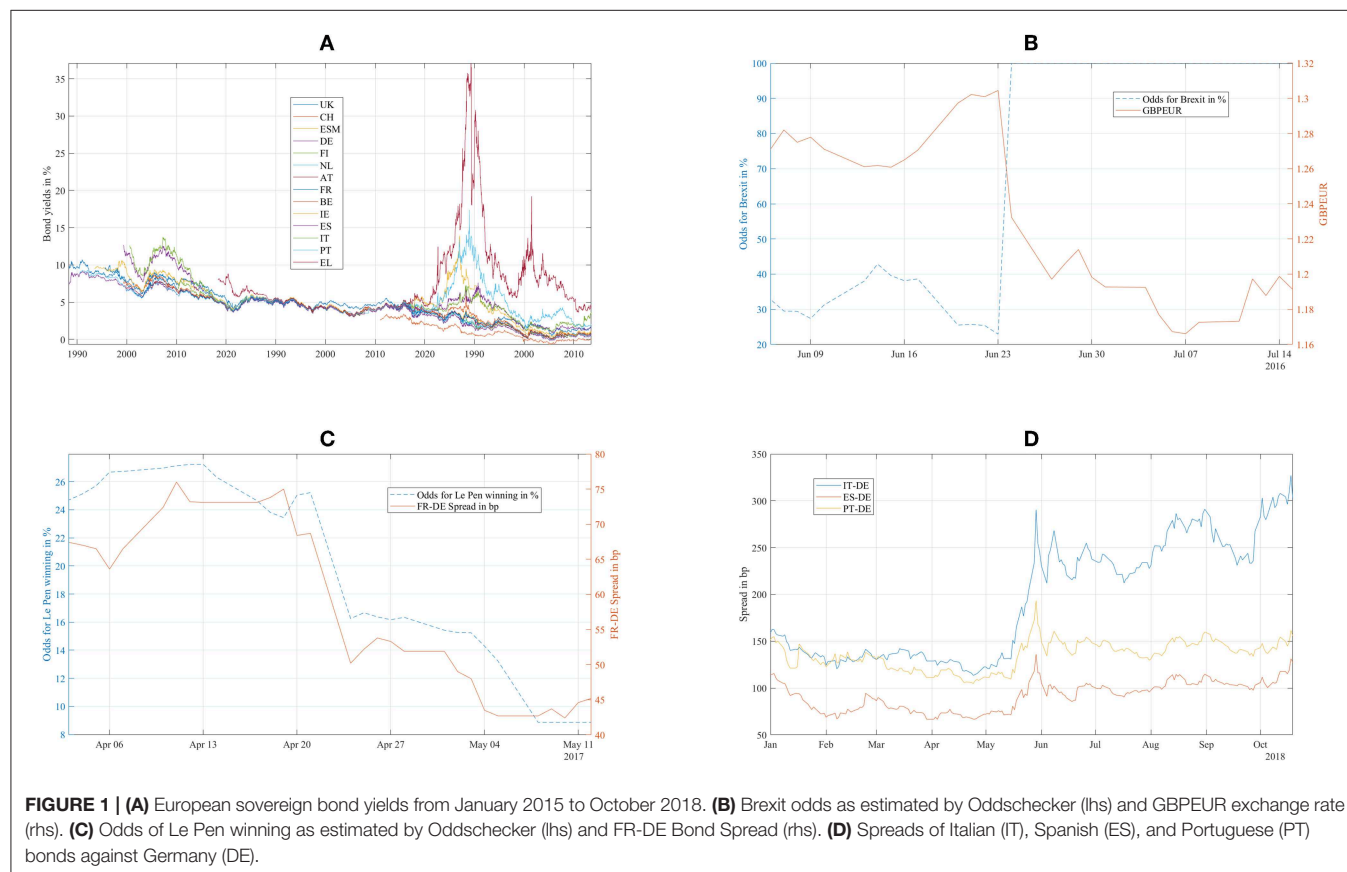
The spread increase in Euro area bonds from 2010 to 2012 has been discussed thoroughly by academia as well as by central bank research and related European institutions, for example by Beirne and Fratscher (2013) and Tola and Waelti (2015). D’Agostino and Ehrmann (2014) pointed to an overreaction of 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,



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

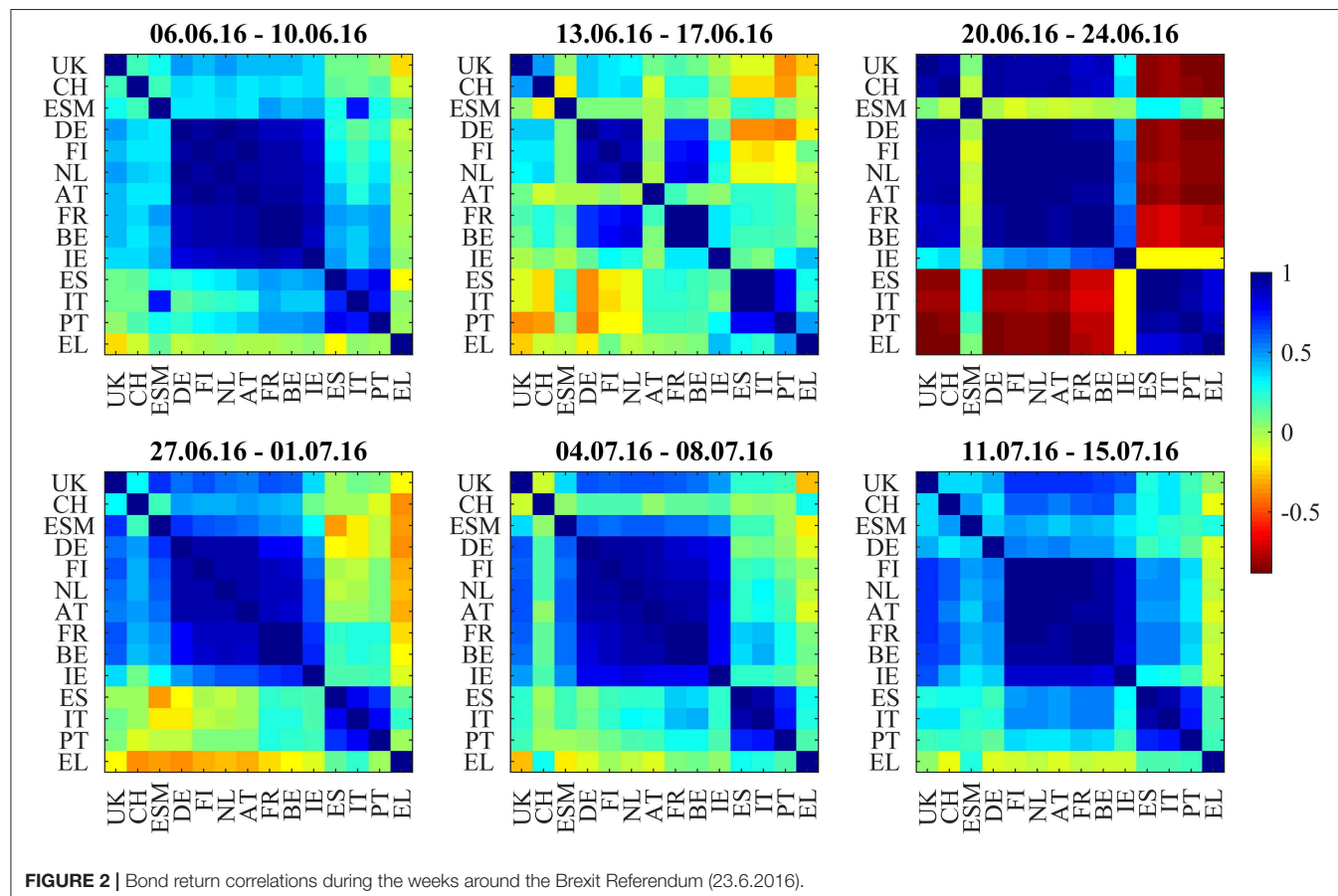
2017, the most important topics were the relationship toward the EU and immigration. The spread between French and German bonds closely followed the odds of the right-wing Marie Le Pen winning in the second round (Bird and Sindreu, 2017; Macintosh, 2017). After Emmanuel Macron won the second round on 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).

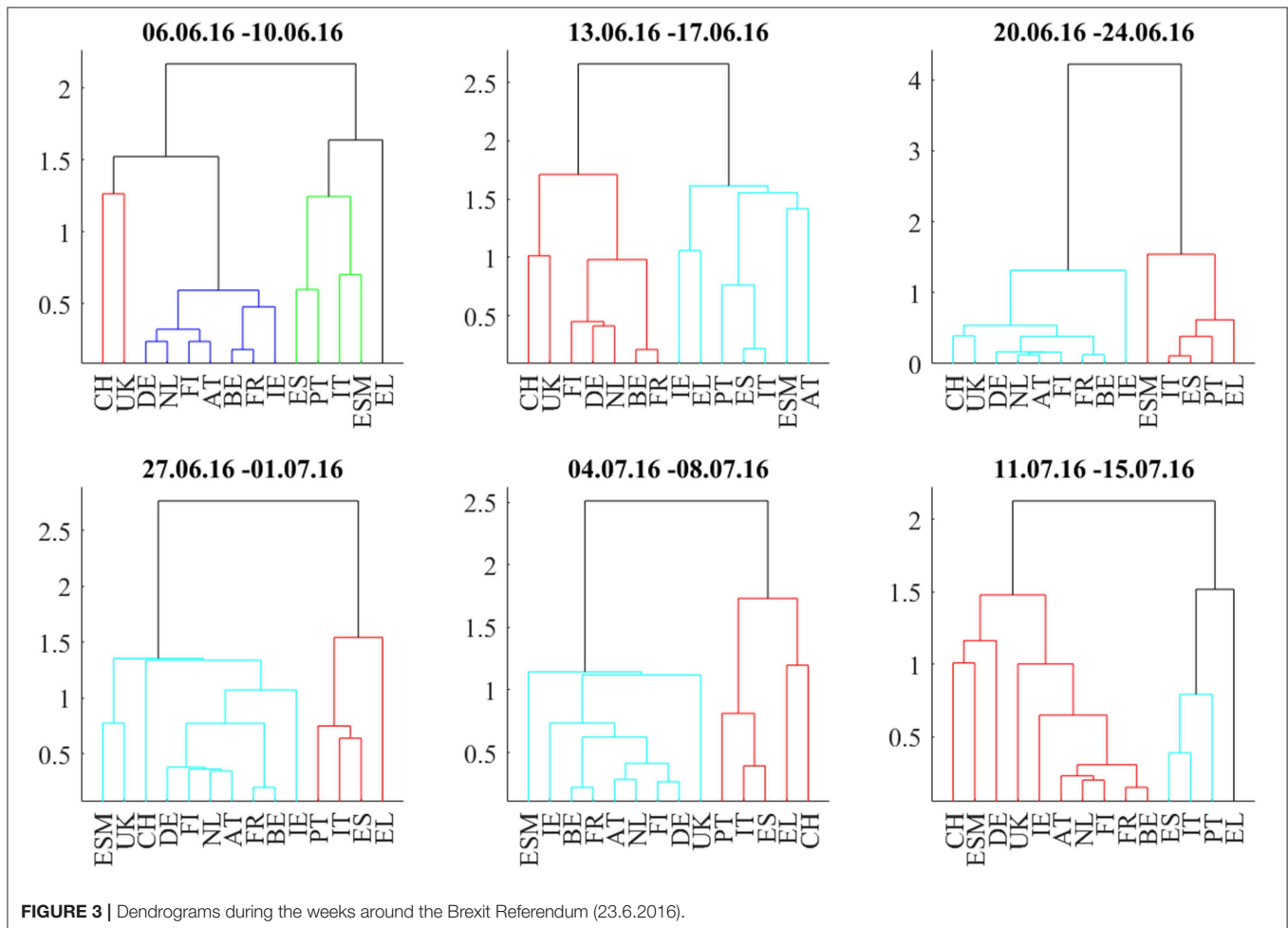
DATA AND METHODS

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



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 .

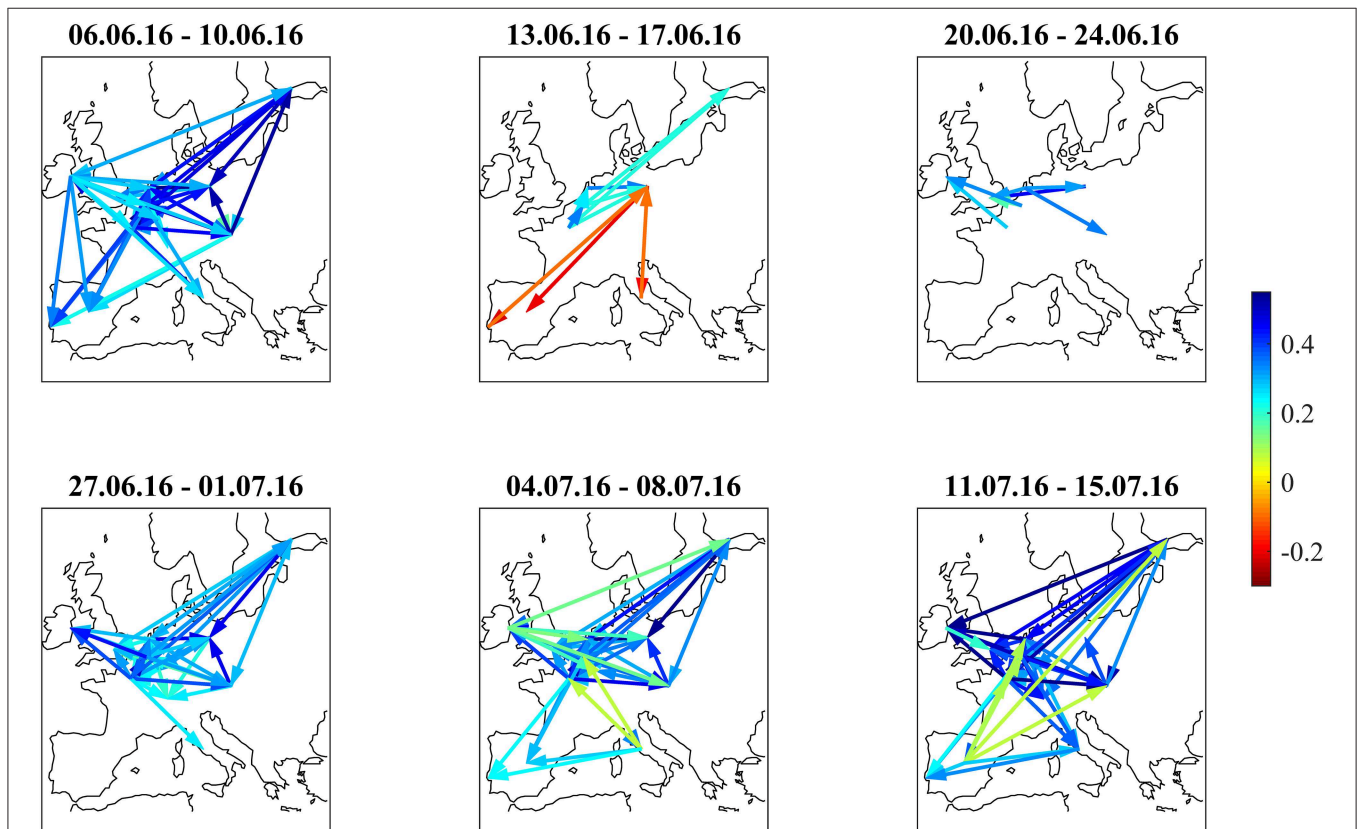


FIGURE 4 | Filtered correlation influence networks during the weeks around the Brexit Referendum (23.6.2016).

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 Popoyan (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

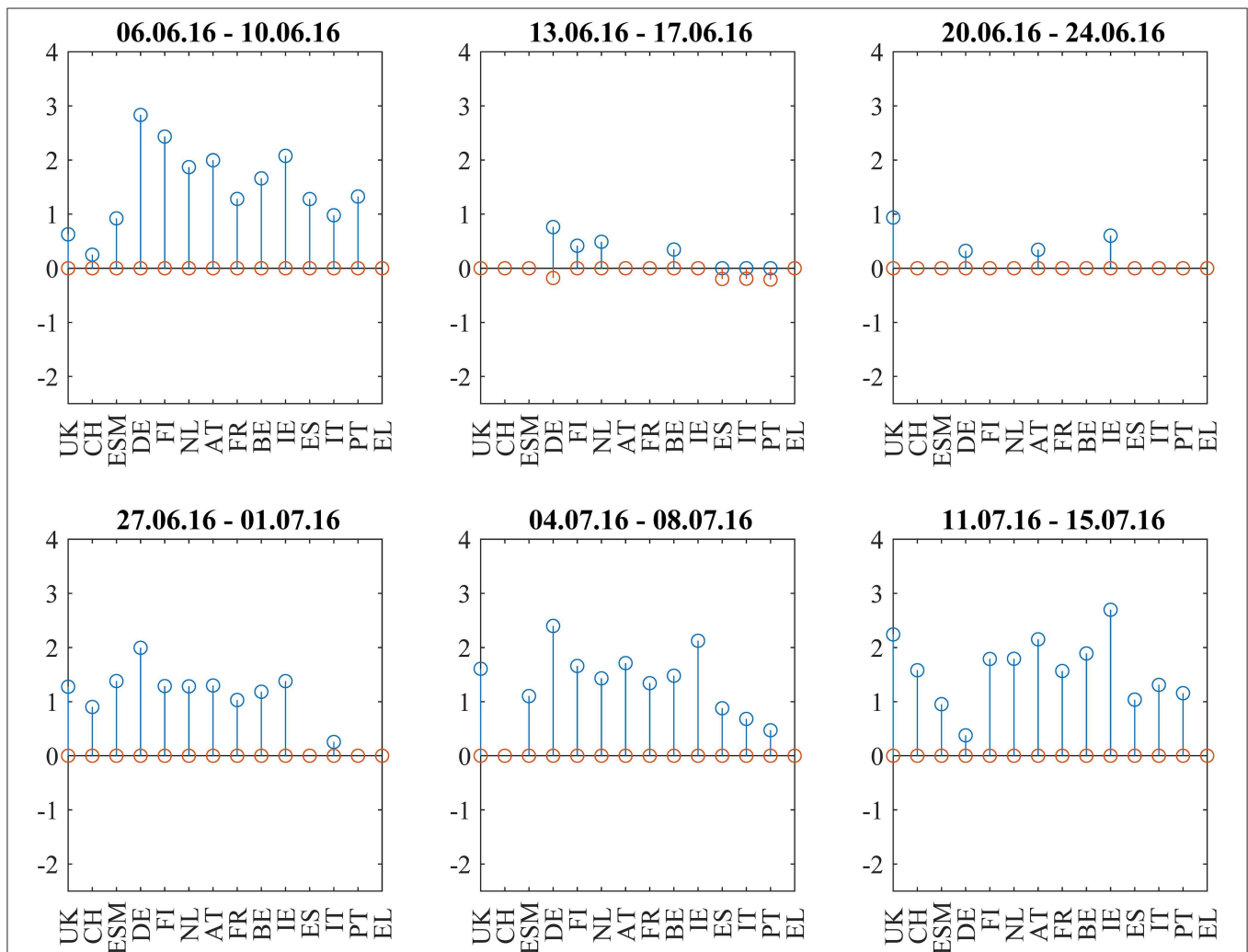


FIGURE 5 | Cumulative positive (blue) and negative (red) incoming filtered correlation influences per market during the weeks around the Brexit Referendum (23.6.2016).

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

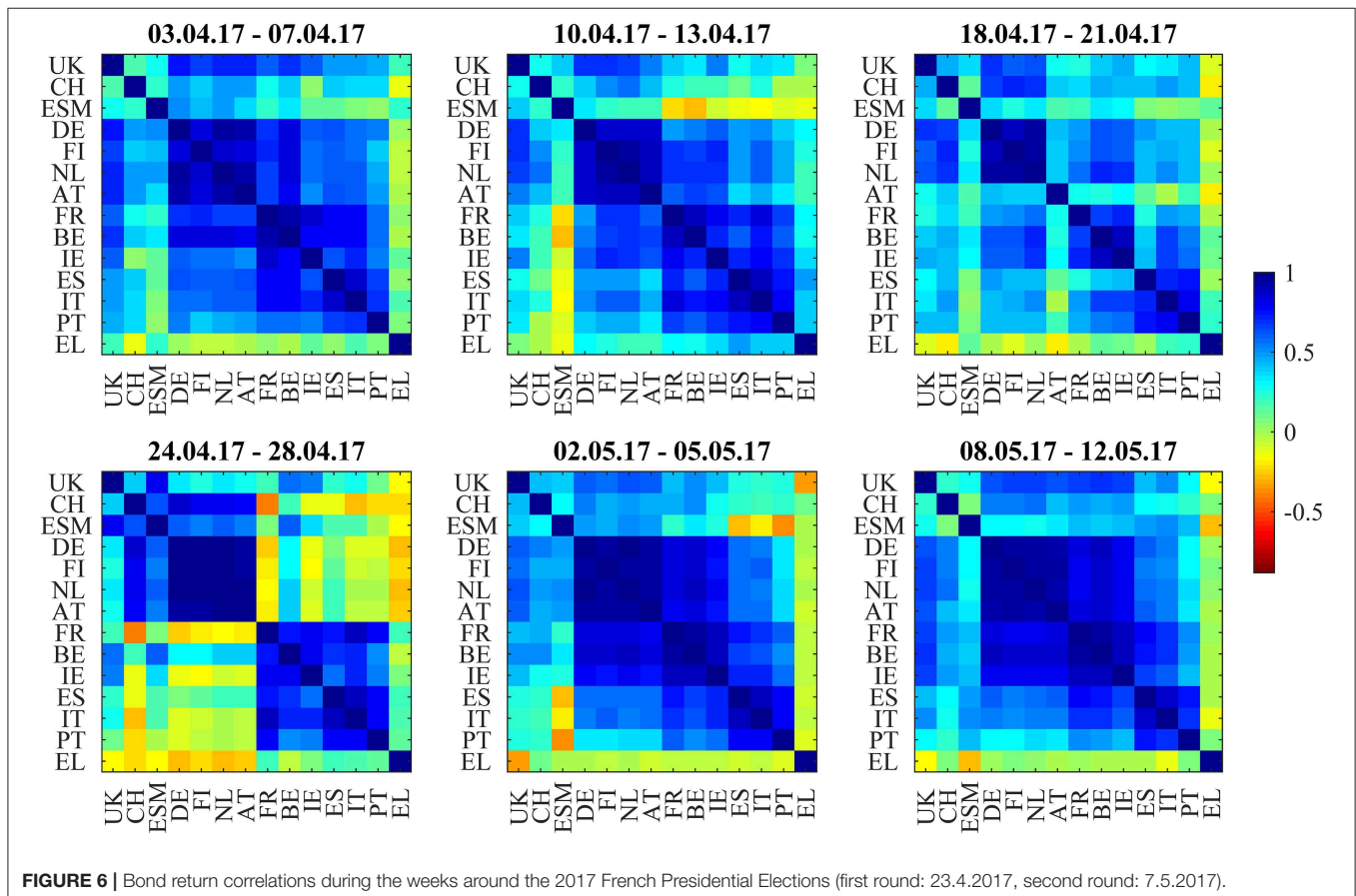
In the Discussion section, we discuss the bond return correlation matrices, hierarchical clusters and filtered correlation influence networks for the three political situations “Brexit referendum,” “French presidential elections,” and “Italian budget negotiations” as main results. A supplementary spreadsheet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

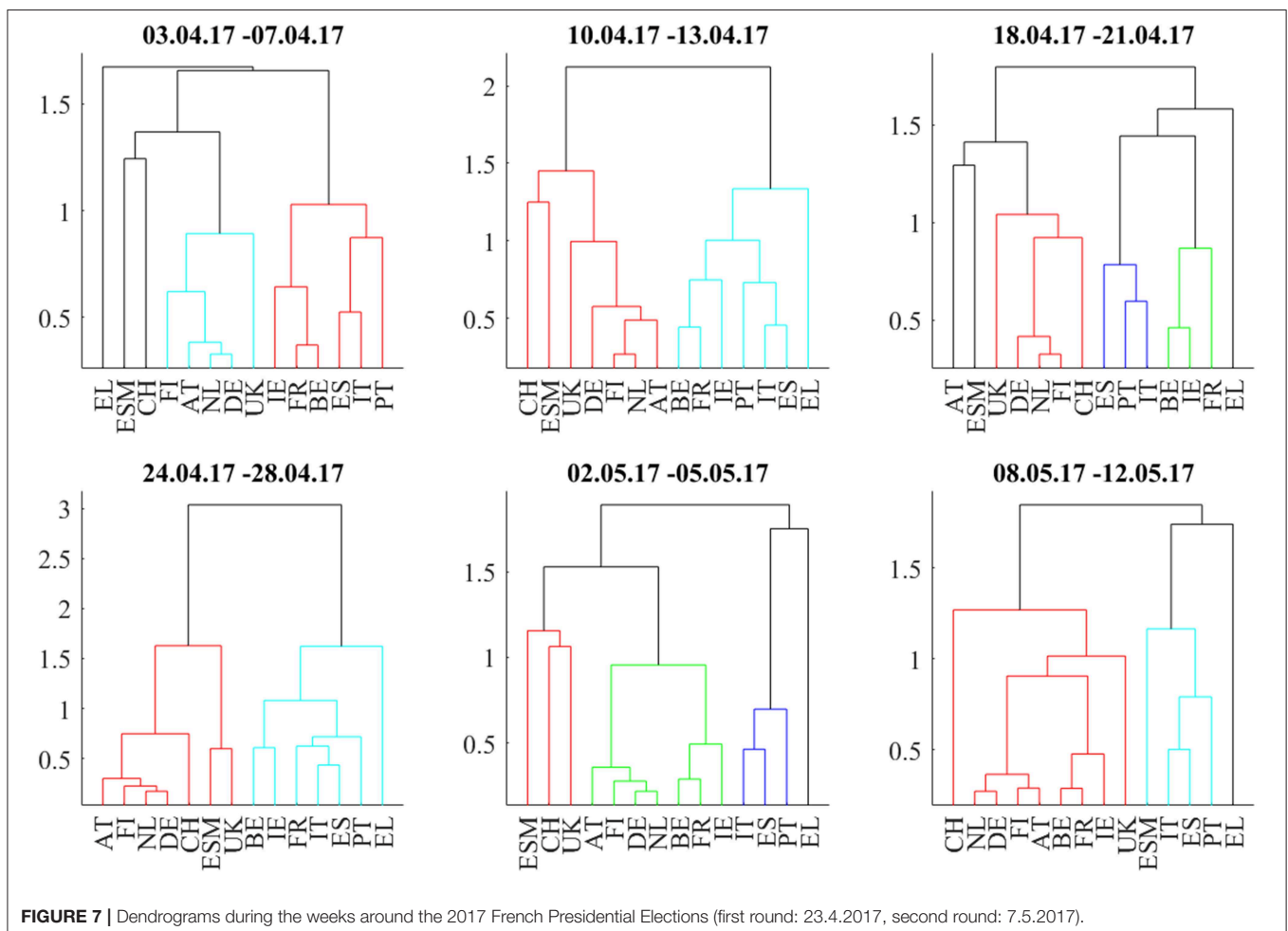
Supplementary Figure 3 shows the cumulative outgoing filtered correlation influences per market.

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: ODCHEAV 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 dendrograms. **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 dendrograms, 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

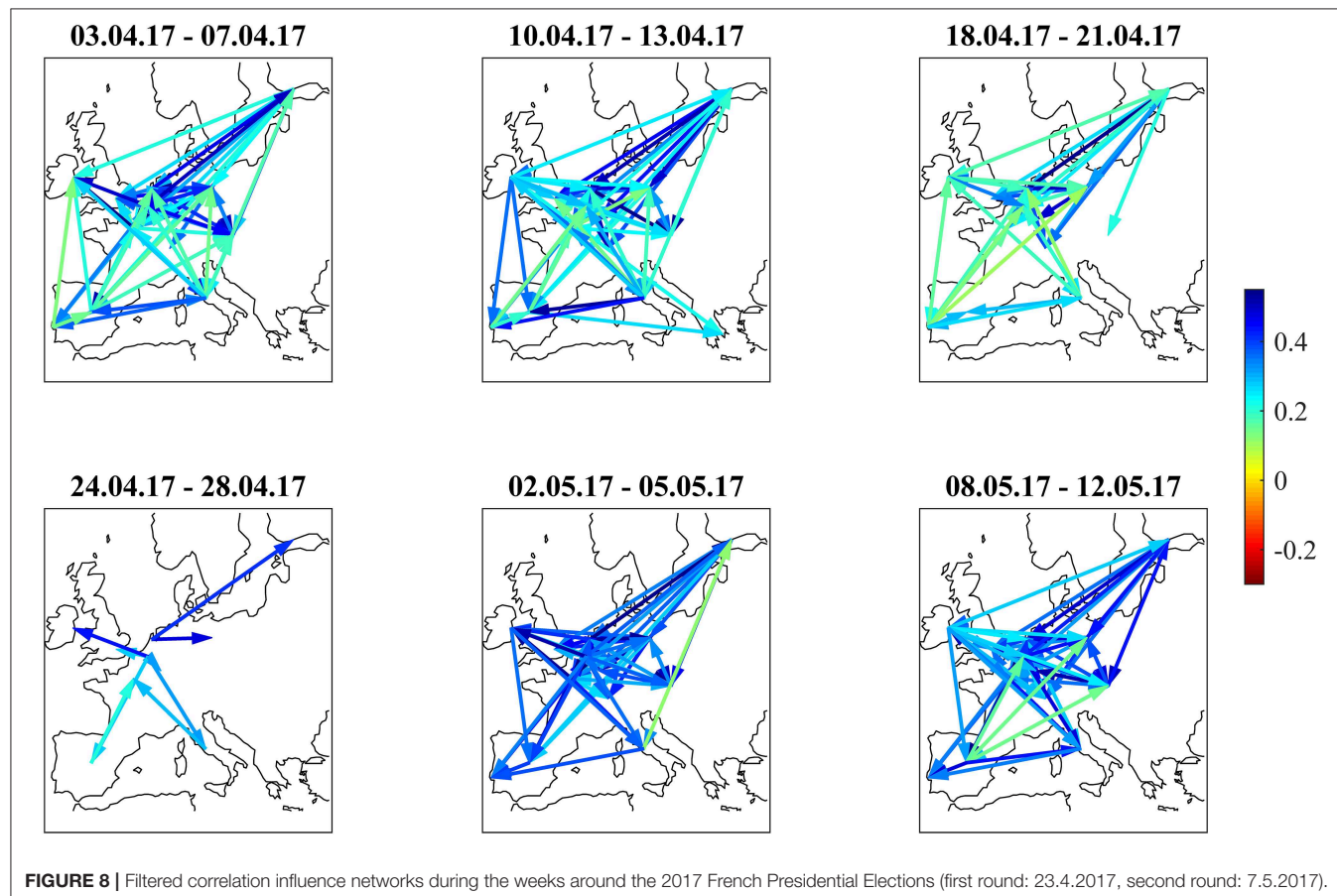
economy. British bonds even gained in market value, consistent with the core Euro area bonds. The dendrogram of this week confirms the strong core-periphery segmentation. The network shows only a few connections that pass the noise filter.

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

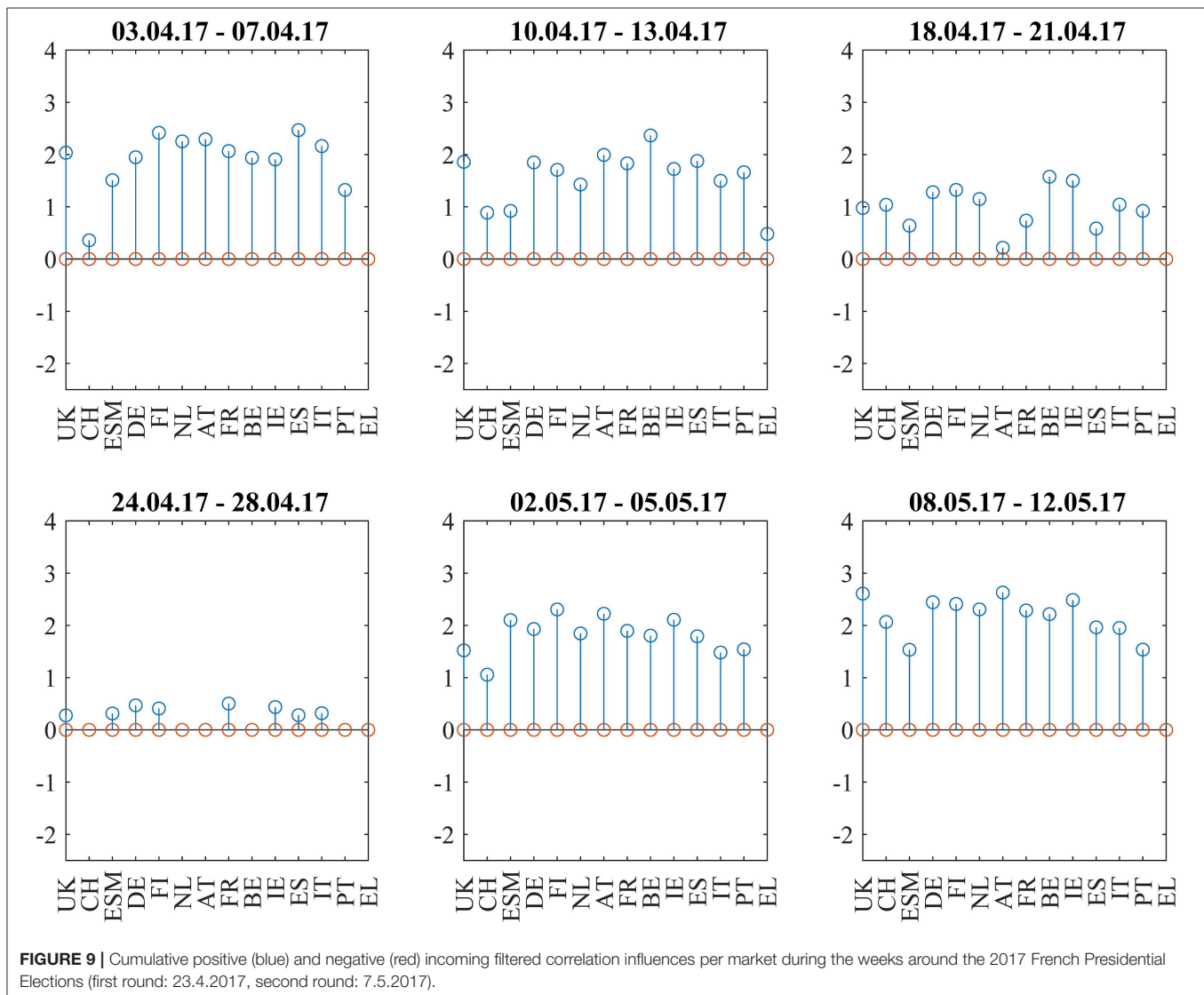


and Macron) and then further until 43 bp at the second round (May 7, resulting in the victory of Macron).

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 dendrograms 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 dendrograms 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.

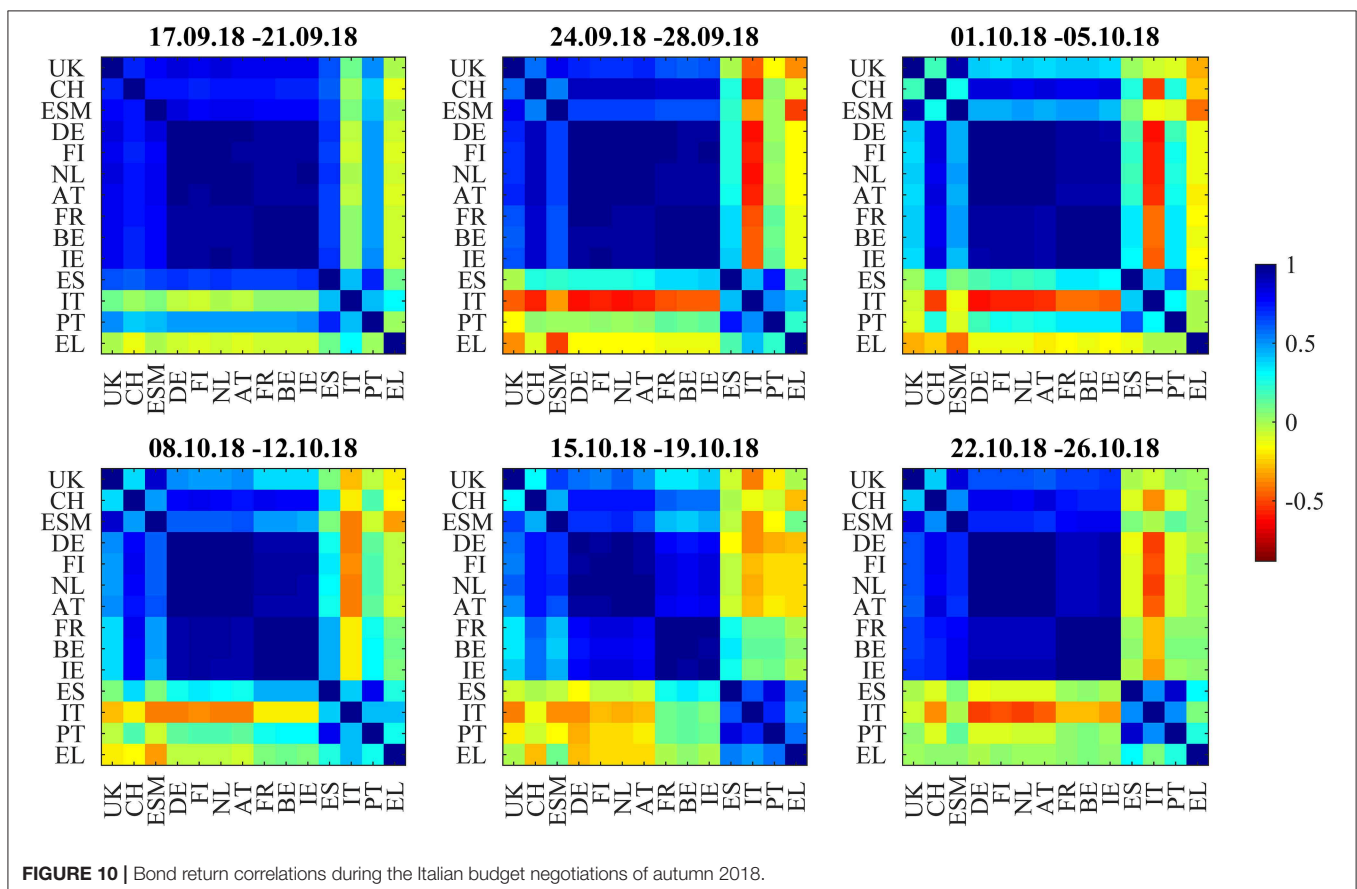
Italian Budget Negotiations

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



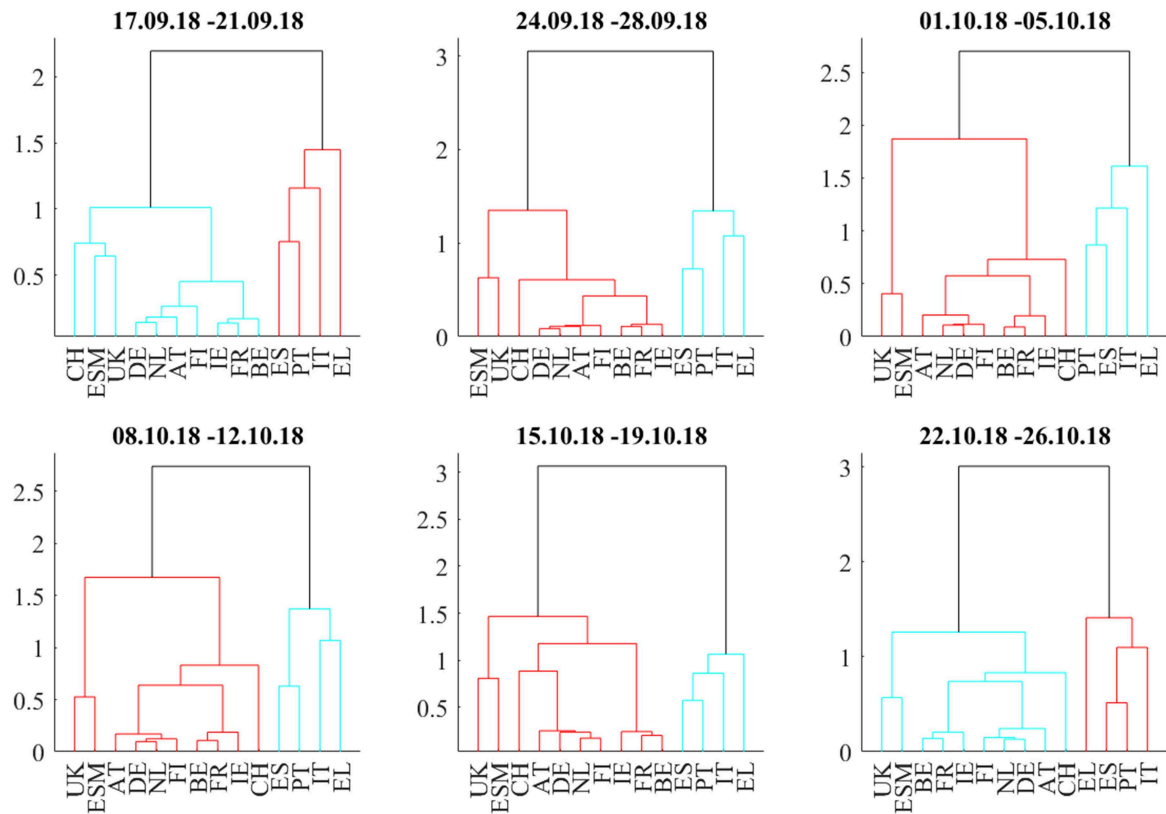


FIGURE 11 | Dendrograms during the Italian budget negotiations of autumn 2018.

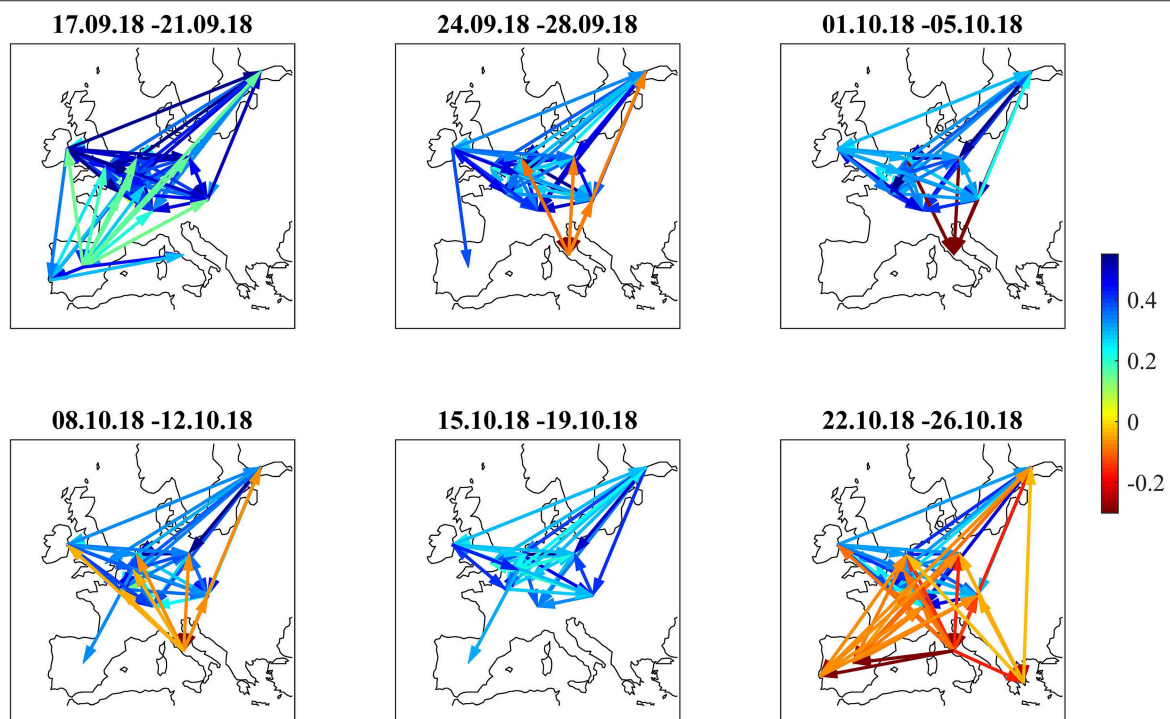


FIGURE 12 | Filtered correlation influence networks during the Italian budget negotiations of autumn 2018.

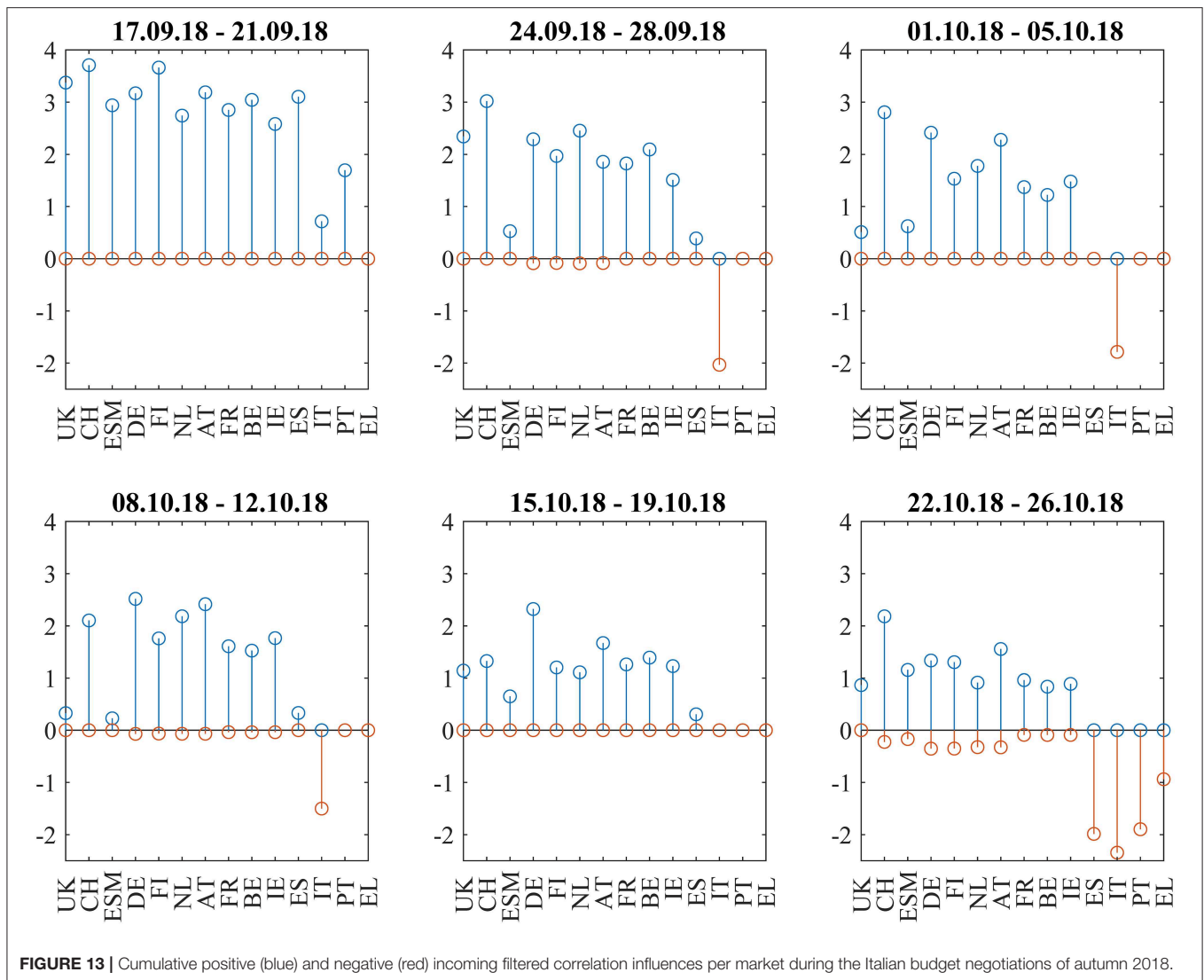


FIGURE 13 | Cumulative positive (blue) and negative (red) incoming filtered correlation influences per market during the Italian budget negotiations of autumn 2018.

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.jimonfin.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.jimonfin.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/europes-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.jimonfin.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., Euan, 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., Gudín, 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
- Shoemith, 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.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership