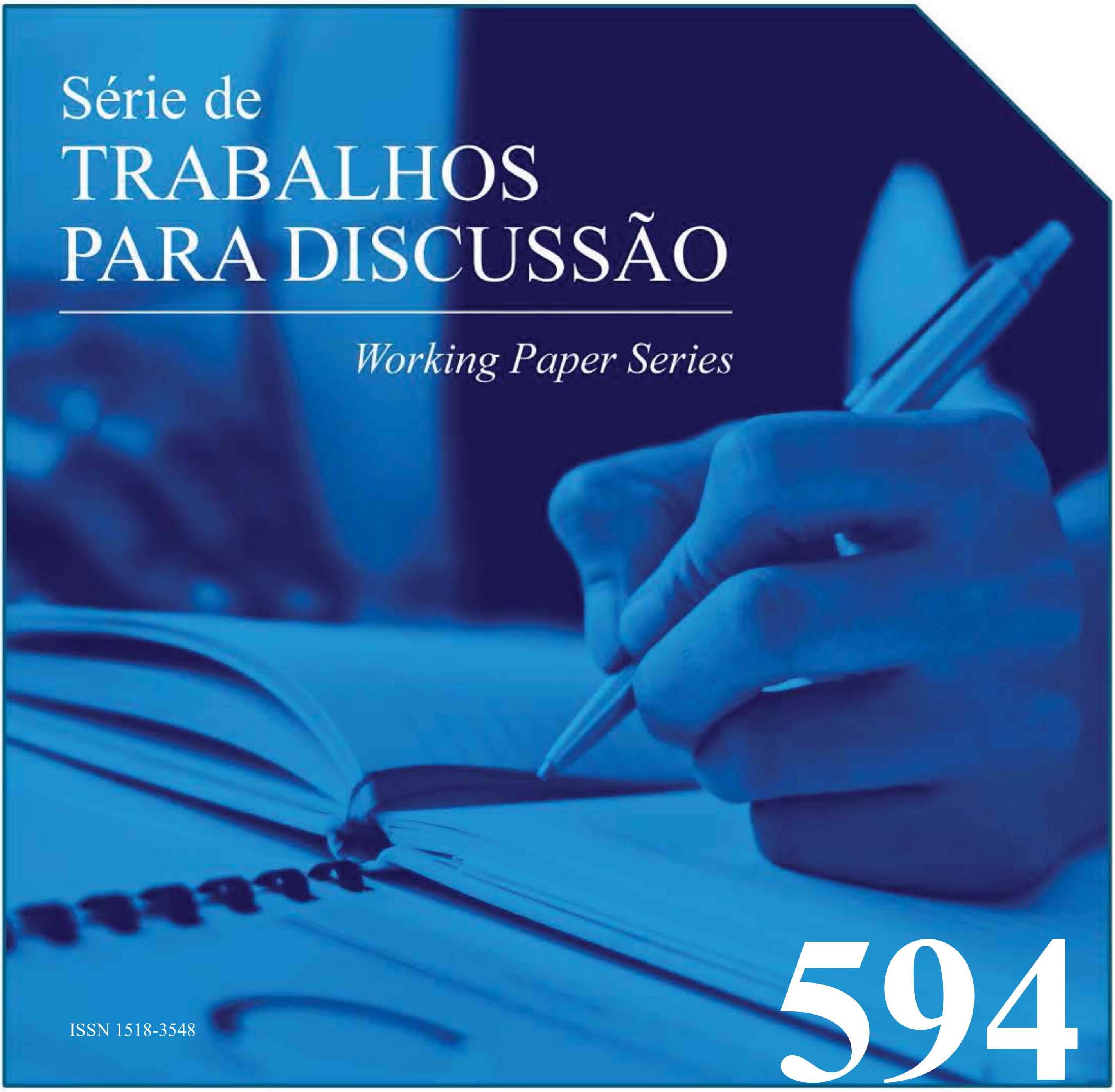


Série de TRABALHOS PARA DISCUSSÃO

Working Paper Series

ISSN 1518-3548

594

Agosto 2024

Critical Edges in Financial Networks

Michel Alexandre, Thiago Christiano Silva, Francisco Aparecido Rodrigues

BANCO CENTRAL DO BRASIL

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	no. 594	Agosto	2024	p. 3-25

Working Paper Series

Edited by the Research Department (Depep) – E-mail: workingpaper@bcb.gov.br

Editor: Rodrigo Barbone Gonzalez

Co-editor: Eurilton Alves Araujo Jr

Head of the Research Department: André Minella

Deputy Governor for Economic Policy: Diogo Abry Guillen

Citizen Service Division

Banco Central do Brasil

Deati/Diate

SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo

70074-900 Brasília – DF – Brazil

Toll Free: 0800 9792345

Fax: +55 (61) 3414-2553

Internet: http//www.bcb.gov.br/?CONTACTUS

Non-Technical Summary

A network is a collection of nodes linked through edges. For instance, in a financial network (say, a bank-firm credit network), the nodes are the banks and firms, and the edges, are the loans extended by the former to the latter. Due to these connections, a shock in one of the nodes can spread to the other nodes. For example, a firm hit by a negative shock will not fully honor its debt obligations against its creditor banks, imposing losses to these banks as well. Similarly, these banks will impose losses on their creditors in the interbank market, and so on. This additional loss engendered by this contagion process is called *systemic risk* (SR).

A key feature of a complex network – which is the case of real financial networks – is that its elements (nodes and edges) are heterogeneous regarding their importance to the whole network. Therefore, it is expected that each edge has a different impact in the SR of the network. Considering this, we propose a methodology to identify critical edges in financial networks – that is, those edges whose removal would cause a large impact on the SR of the network. This is done by comparing the SR of the financial network before and after the removal of the edge. The critical edges are the outliers considering this impact – that is, those whose impact is much larger than the average impact.

We apply this framework to a thorough Brazilian dataset to identify critical bank-firm edges. In our data set, banks and firms are connected through two financial networks: the interbank network and the bank-firm loan network. At least 18% of the edges are critical, but this fraction depends on the level of the initial shock. Employing machine learning (ML) techniques and a set of explanatory variables, we can predict the edges' critical status with a high level of accuracy (around 90%). Moreover, this analysis shows some important predictive features interact nonlinearly with the predicted critical status.

An interesting phenomenon emerges when we raise the value of the initial shock: the fraction of edges whose impact is not positive increases from virtually zero to almost 18%. We investigated this phenomenon using the same ML techniques and explanatory variables. We found the sign of the edge's impact is mainly related to the degree (i.e., the number of connections) of its origin node – that is, the lender. Edges whose origin node has a higher in-degree (that is, a high number of lending partners) have a higher probability of having a positive impact on the systemic risk of the whole network. On the other hand, edges whose origin node has a higher out-degree (number of borrowing partners) have a higher probability of having a non-positive impact on systemic risk.

Sumário Não Técnico

Uma rede é uma coleção de nós ligados por arestas. Por exemplo, em uma rede financeira — como uma rede de crédito banco-firma —, os nós são os bancos e as firmas e as arestas, os empréstimos concedidos pelos primeiros aos segundos. Devido a essas conexões, um choque em um dos nós pode se espalhar para os demais. Por exemplo, uma firma atingida por um impacto negativo não honrará integralmente as suas dívidas com os seus bancos credores, impondo perdas também para esses bancos. Da mesma forma, esses bancos imporão perdas aos seus credores no mercado interbancário, e assim por diante. Essa perda adicional gerada por este processo de contágio é denominada *risco sistêmico* (RS).

Uma característica fundamental de uma rede complexa -– que é o caso das redes financeiras reais — é que seus elementos (nós e arestas) são heterogêneos quanto à sua importância para o toda a rede. Portanto, espera-se que cada aresta tenha um impacto diferente no RS da rede. Considerando isso, propomos uma metodologia para identificar arestas críticas em redes financeiras – isto é, aquelas arestas cuja remoção causaria um grande impacto no RS da rede. Isto é feito comparando o RS da rede financeira antes e após a remoção da borda. As arestas críticas são os outliers considerando-se esse impacto — isto é, aquelas cujo impacto é muito maior que o impacto médio.

Aplicamos essa metodologia a um amplo conjunto de dados brasileiro para identificar arestas críticas entre bancos e empresas. No nosso conjunto de dados, os bancos e as firmas estão ligados através de duas redes financeiras: a rede interbancária e rede de empréstimos banco-firma. Pelo menos 18% das arestas são críticas, mas essa fracção depende do nível do choque inicial. Empregando técnicas de aprendizado de máquina (AM) e um conjunto de variáveis explicativas, podemos prever a criticalidade das arestas com um alto nível de precisão (cerca de 90%). Além disso, essa análise mostra que algumas variáveis preditivas importantes interagem de forma não linear com a criticalidade prevista das arestas.

Um fenômeno interessante surge quando aumentamos o valor do choque inicial: a fração de arestas cujo impacto não é positivo aumenta de praticamente zero para quase 18%. Nós investigamos esse fenômeno usando as mesmas técnicas de AM e variáveis explicativas. Descobrimos que o sinal do impacto da aresta está principalmente relacionado ao grau (ou seja, o número de conexões) do seu nó de origem – ou seja, o credor. Arestas cujo nó de origem possui um maior grau de entrada (ou seja, um maior número de concessores de empréstimo) têm uma maior probabilidade de ter um impacto positivo no risco sistêmico da rede. Por outro lado, arestas cujo nó de origem tem um grau de saída mais alto (maior número de tomadores de empréstimos) têm uma maior probabilidade de ter um impacto não positivo no risco sistêmico.

Critical Edges in Financial Networks

Michel Alexandre[*]

Thiago Christiano Silva[**]

Francisco Aparecido Rodrigues[***]

Abstract

In this study, we propose a method for the identification of influential edges in financial networks. In our approach, the critical edges are those whose removal would cause a large impact on the systemic risk of the financial network. We apply this framework to a thorough Brazilian data set to identify critical bank-firm edges. In our data set, banks and firms are connected through two financial networks: the interbank network and the bank-firm loan network. We found at least 18% of the edges are critical, in the sense they have a significant impact on the systemic risk of the network. We then employed machine learning (ML) techniques to predict the critical status and – for a large level of the initial shock – the sign of the impact of bank-firm edges on the systemic risk. The level of accuracy obtained in these prediction exercises was very high (above 90%). Posterior analysis through Shapley values shows: i) the PageRank of the edge's destination node (the firm) is the main driver of the critical status of the edges; and ii) the sign of the edges' impact depends on the degree of the edge's origin node (the bank).

JEL classification codes: C63, D85, G21, G23, G28

Keywords: critical edges, complex networks, financial networks, machine learning

[*]Research Department, Central Bank of Brazil, Brazilian Institute of Education, Development and Research (IDP), and University of São Paulo. SBS Quadra 3 Bloco B, 70074-900, Brasília, Brazil. E-mail: michel.alexandre@bcb.gov.br

[**]Research Department, Central Bank of Brazil and Universidade Católica de Brasília. SBS Quadra 3 Bloco B, 70074-900, Brasília, Brazil. E-mail: thiago.silva@bcb.gov.br

[***]Institute of Mathematics and Computer Science, University of São Paulo. Av. Trab. São Carlense 400, 13566-590, São Carlos, Brazil. E-mail: francisco@icmc.usp.br

1 Introduction

This paper deals with the impacts of interconnectedness on the stability of financial networks. A network can be represented as a collection of nodes (or vertices) connected through edges. Financial networks are ubiquitous, such as interbank markets (Allen and Gale, 2000; Bardoscia et al., 2016), bipartite bank-firm credit networks (De Masi and Gallegati, 2012; Lux, 2016), networks of correlated assets (Gabaix et al., 2003; Mantegna and Stanley, 1995), and so on. The type of node varies according to the network (e.g., banks in interbank networks, banks and firms in bank-firm credit networks), and so does the type of edge (e.g., loans in bipartite bank-firm credit networks, return correlations in networks of correlated assets).

The interconnections among financial agents configure a channel for both risk-sharing and shock propagation. It configures the so-called *robust-yet-fragile* nature of financial networks (Chinazzi and Fagiolo, 2015). Until recently, before the 2008 crisis, the conventional wisdom regarding the impact of interconnectedness in financial networks went hand in hand with the view expressed in Allen and Gale (2000) seminal paper. According to this paradigm, the benefits brought about by risk diversification overcome the contagion problems. Therefore, more interconnected financial networks would be more resilient. However, the 2008 financial turmoil defeated this view. The burst of the subprime mortgage bubble in the United States in 2007 led, through a network of complex financial linkages, to the collapse of big risk-taking banks, like that of Lehman Brothers in September 2008. Since then, interconnectedness in financial networks has been mostly associated with the propagation of losses between different financial actors (Acemoglu et al., 2015; Battiston and Martinez-Jaramillo, 2018; Gai and Kapadia, 2010; Glasserman and Young, 2015, 2016; Martinez-Jaramillo et al., 2019).

The propagation of shocks among different interconnected financial agents can lead to the collapse of a non-negligible part of the whole system. This risk is called *systemic risk* (SR). An interesting research question is how the tiniest variation in the interconnectedness of a financial network – namely, the inclusion/removal of an edge – affects SR. On the other hand, the most striking characteristic of complex networks – the case of real financial networks – regards the heterogeneity of their components (Albert and Barabási, 2002). Within this perspective, we can define *critical edges* as extraordinary edges that play a more significant role than other edges in the structure and function of the network (Yu et al., 2018). Specifically in the context of this study, the removal of a critical edge would have a much greater impact on the SR of the network than the removal of a non-critical edge.

In this paper, we propose a framework to the identification of critical edges in financial networks. Although node importance has been explored more often,[1] the study of the criticality of edges in networks has attracted much attention from researchers in many fields (Bröhl and Lehnertz, 2019; De la Cruz Cabrera et al., 2020; Hajarathaiah et al., 2024; Song, 2023; Yu et al., 2018; Zhao et al., 2020). Identifying critical edges helps in the design of procedures to keep or increase the resilience of a network (Song, 2023; Zhao et al., 2020). Moreover, the removal of edges as a strategy to protect the network is much less costly than the node-cutting strategy. For instance, the withdrawal of certain

[1] See, for instance, Alexandre et al. (2021), Ghanbari et al. (2018), Kuzubas et al. (2014), and Martinez-Jaramillo et al. (2014) specifically on the identification of important nodes in financial networks.

financial products is much more applicable than removing financial institutions to prevent financial crises (May et al., 2008). Many approaches have been proposed to identify the importance of edges in networks. Onnela et al. (2007) measure the importance of an edge based on the number of common neighbors between the nodes linked by the node. Yu et al. (2018) propose a measure of edge relevance based on their betweenness centrality and the number of cliques containing the edge. Betweenness centrality is also the basis of the metric of edge importance suggested by Kanwar et al. (2022). The metric proposed by Helander and McAllister (2018) exploits the properties of a k-shortest path algorithm. Zhao et al. (2020) create a second-order neighborhood index to quantify the relevance of an edge in a network.

As far as we know, the only study devoted specifically to the identification of critical edges in financial networks is that of Seabrook et al. (2021). The authors defined a structural importance metric, l_e, based on the change in the largest eigenvalues of the adjacency matrix of the network resulting from perturbations in it. The authors then propose a model of network evolution where this metric controls the probabilities of subsequent edge changes. Using synthetic data, they show how the parameters of the model are related to the capability of l_e to predict whether an edge will change. Assessing five real networks (four financial and one social), the authors showed l_e is slightly predictive of edge change in all cases, but only marginally so for two of the financial networks.

In this study, we propose a method for the identification of influential edges based on the impact of their removal on the stability of the financial network. First, we compute the impact of the edge by measuring the change in the systemic risk of the financial network caused by the removal of the edge. Next, we identify the critical edges as the outliers according to this impact, following the interquartile range (IQR) approach. Thus, the critical edges are those whose removal would cause a large impact on the systemic risk of the financial network. Finally, we employ machine learning (ML) techniques to classify an edge as critical using topological and financial features as predictive variables. Shapley values are used in the interpretation of the results, showing which features are more relevant in the definition of critical edges in financial networks.

We apply this framework to a thorough Brazilian dataset to identify critical bank-firm edges – that is, edges whose origin nodes are banks and destination nodes are firms. In other words, we are assessing the criticality of edges that represent loans extended by banks to firms. In our data set, the two types of agents – banks and firms – are connected through two financial networks: the interbank network and the bank-firm loan network. We found at least 18% of the edges are critical (depending on the level of the initial shock), in the sense they have a significant impact on the systemic risk of the network. Moreover, we observed that, for a larger level of the initial shock, the fraction of edges with non-positive impact rises significantly.

We then employ machine learning (ML) techniques and a set of predictive features to predict the critical status of the bank-firm edges, as well as the sign of the edge impact for a large level of the initial shock. The purpose is to identify which features (from the edges or the nodes) are the drivers of edge criticality and – in the case of large shocks – of positive impacts of the edge on the SR. The level of accuracy obtained in these prediction exercises was very high (above 90%). We use Shapley values for the sake of a better interpretability of our results. This analysis shows: i) the PageRank of the edge's destination – that is, the firm – is the main driver of the critical status of the edges; and ii) the sign of the edges' impact depends on the degree of the edge's origin node (the bank).

Our study is related to the literature on the relationship between interconnectedness and SR in financial networks. It is widely accepted that this relationship is nonlinear (Gai and Kapadia, 2010; Nier et al., 2007), and driven by other elements, such as the size of the initial shock (Acemoglu et al., 2015) and the loss distribution regime adopted by distressed debtors (Alexandre et al., 2023). In this study, we assess how the tiniest change in the interconnectedness – that is, removal/inclusion of edges – impacts the SR of the financial network. We show the edges have different degrees of importance to the resilience of the financial network, corroborating the heterogeneous nature of financial networks. Moreover, we identify which features – among a set of variables related to the edge and its nodes – drive the size and the sign of the impact of an edge on the stability of the financial network.

This paper proceeds as follows. The data set and methodological issues are discussed in Sections 2 and 3, respectively. Section 4 brings the results. Final considerations are presented in Section 5.

2 The data set

Our data set comprises two types of information: i) financial and supervisory (F & S) information on the agents, which can be financial institutions (FIs) or firms, and ii) information on the financial linkages between these agents. The FIs of our data set can be banks or credit unions. All these data are from December 2022.

We include in our data set FIs which are financial conglomerates or individual FIs. Concerning the firms, we include those belonging to the *Alexandria* database (Docha and Rodrigues, 2023) with positive net worth. The Alexandria database gathers information on more than 42,000 firms, of which more than 31,000 are non-financial companies (NFCs). Most of these firms are medium-sized and owned by non-residents. The firms belong to 18 different economic sectors, labeled from A to S (we are only considering NFCs, so financial firms are excluded). The F & S information on these agents includes net worth, financial indicators (e.g., ROE), type of control, economic sector, and so on.

Regarding the financial linkages, we consider two financial networks: the interbank (IB) and the bank-firm credit network. In the IB network, we consider the net financial exposures between Brazilian FIs in the IB market. This network comprises all types of unsecured financial instruments, such as credit, capital, foreign exchange operations, and money markets. These financial instruments are registered in different custodian institutions – Cetip, Central Bank of Brazil, and B3. Exposures between FIs belonging to the same financial conglomerate are excluded. The bank-firm credit network encompasses the corporate loans granted by FIs to non-financial firms. The source of this information is the Central Bank of Brazil's Credit Risk Bureau System (SCR). Besides the value of these exposures for both networks, we have more information on the bank-firm loans, such as the interest rate and maturity. Some topological metrics of both networks are presented in Table 1.

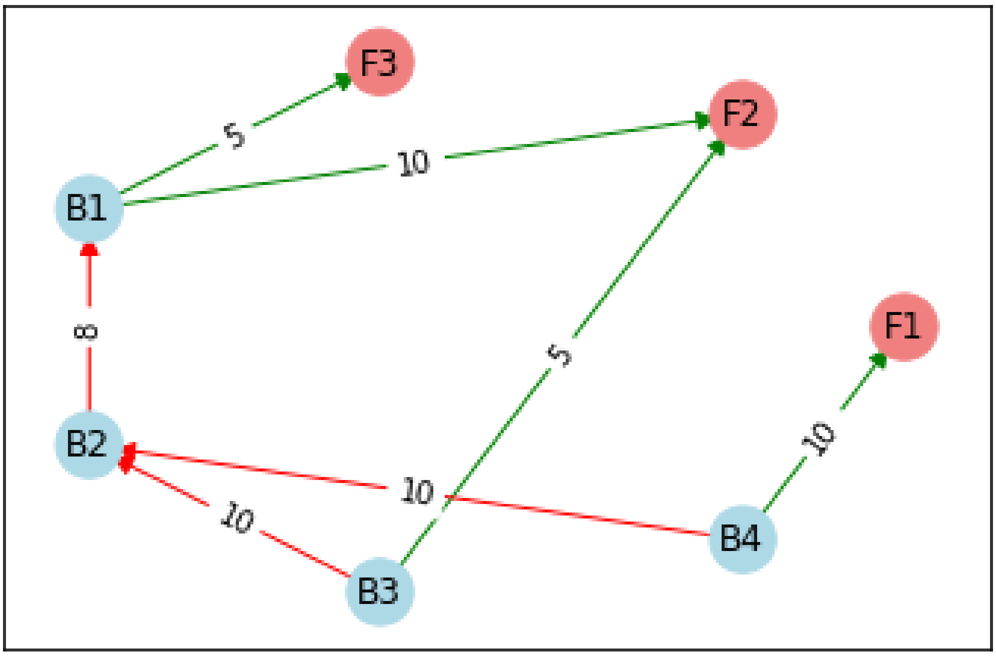
__Table 1:__ Some topological metrics of the networks

Type of network	Metric	Value
IB network	N. of nodes	803
	N. of edges	2971
	Assortativity	-0.3630
	Average degree	7.3998
	Average closeness centrality	0.1220
	Average eigenvector centrality	0.0174
Bank-firm network	N. of nodes	11,955
	N. of edges	28,157
	Assortativity	-0.3965
	Average degree	4.7105
	Average closeness centrality	0.0002
	Average eigenvector centrality	0.0063

3 Methodology

As stated in Section 2, we consider two types of financial linkages: the IB market and the bank-firm credit network. Figure 1 depicts a simple representation of our financial networks. The nodes in blue are the banks, labeled as B1,...,B4, and the nodes in coral are firms, labeled as F1,...,F3. The edges in red represented the IB connections and those in green, the bank-firm linkages. Edges are directed, showing the origin node and the destination node, and weighted. For instance, bank 3 (B3) has lent an amount of resources equal to 10 to bank 2 (B2) in the IB market and equal to 5 to firm 2 (F2).

__Figure 1:__ A stylized representation of a financial network



The first step is to compute the SR of the whole system depicted in Figure 1 by imposing losses in the form of net worth decrease to all firms at the same time. The net worth of both FIs and firms and the exposures among them in the two financial networks will be used at this step. The green edges are responsible for the *direct losses* this initial shock will cause on the banks. The red edges represent the IB connections that will amplify this initial shock – that is, they are responsible for the *contagion*.

We are interested in measuring the systemic relevance of bank-firm edges – that is, the green edges. This is done by removing one edge at a time and computing the proportional difference in the SR of the system. This metric is the edge's impact. An edge is deemed as critical if it is an outlier considering its impact, according to the IQR approach. Finally, we perform some prediction exercises. The topological variables built from the two networks, along with the F& S variables on the agents and the variables related to the bank-firm loan itself, will constitute the set of potential explanatory variables. These variables will be employed to predict the target variables – the edges' critical status and the sign of the edges' impact — through ML techniques. In the following subsections, we detail our methodological process.

3.1 Systemic risk

In what follows, we describe the *differential DebtRank* (DDR) approach (Bardoscia et al., 2015), used to compute the SR of the financial network. Let us define E_i as the equity of agent i and A_{ij}, the net exposure of agent i towards agent j. The agents correspond to banks and firms in Figure 1, and the exposures, to the weight of the edges between them. At period $t = 0$, we impose an exogenous shock on the network, in which some agents will lose a fraction ζ of their equity. At $t = 1$, the agents hit by the exogenous shock will transmit part of this loss to their creditors, by defaulting on part of their debt. In the next period, the agents that suffered some loss (that is, those with exposures towards the agents hit by the exogenous shock) will transmit part of this loss to their creditors. This process of loss transmission will continue in the subsequent periods. At a given period t, the accumulated loss transmitted by agent j to its creditor i is $L_{ij}(t)$. The aggregate loss suffered by i (considering all i's debtors) up to t is $L_i(t)$. The dynamics of these two variables are represented by the following equations:

$$\Delta L_{ij}(t) = min\left(A_{ij} - L_{ij}(t-1), A_{ij}\frac{[L_j(t-1) - L_j(t-2)]}{E_j}\right), \tag{1}$$

$$\Delta L_i(t) = min\left(E_i - L_i(t-1), \sum_j \Delta L_{ij}(t)\right), \tag{2}$$

in which $t \geq 0$. In Eqs. 1 and 2, $\Delta L_{ij}(t) = L_{ij}(t) - L_{ij}(t-1)$ is the new flow of loss transmitted by j to i at t, and $\Delta L_i(t) = L_i(t) - L_i(t-1)$ is the variation in the total loss transmitted to i by their debtors at t. Thus, the loss transmitted by a given debtor to a given creditor is the proportional loss of equity suffered by the debtor times the exposure of the creditor on it. Observe, from Eq. 1, that the loss imposed by j to i cannot be greater than i's exposures towards j. Moreover, an agent cannot suffer a loss greater than its equity (Eq. 2).

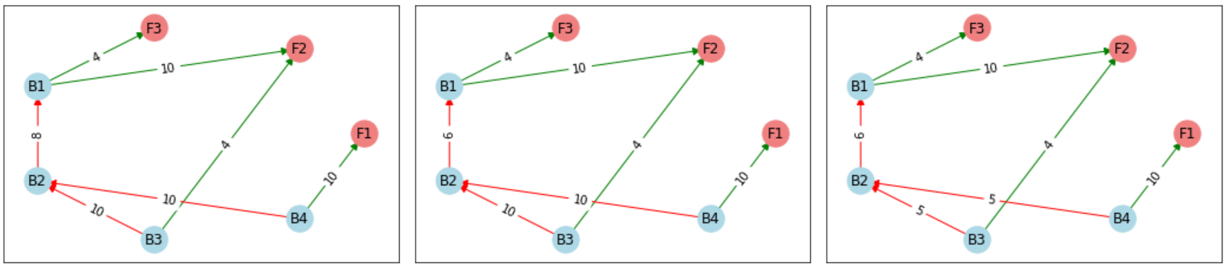
At $t = T \gg 0$, the system converges – i.e., when no more losses are transmitted. Then, we compute the *systemic risk* of the network according to the following equation:

$$S_\zeta = 100 \times \frac{\sum_i [L_i(T) - L_i(0)]}{\sum_i E_i}. \tag{3}$$

Therefore, the systemic risk S_ζ of the network is the percentage of the aggregate equity of the system which is lost after an exogenous shock of size ζ. Observe that we remove $L_i(0)$ from the computation, which is the loss resulting from the exogenous shock. Thus, S_ζ considers only the loss resulting from the contagion process.

Figure 2 gives a simple example of how SR is computed according to the DDR approach. Considering again Figure 1, suppose F3 loses 20% of its net worth. This firm will default the same fraction on the exposures of B1, its only creditor, towards it. For simplicity, assume all banks have a net worth equal to 4. The loss transmitted by F3 to B1 ($0.2 \times 5 = 1$) corresponds to 25% of B1's net worth. The exposure of B1 towards F3 reduces to 4 (Figure 2, panel (a)). In the next step, B1 defaults 25% on the exposures of B2 in the IB towards it, imposing to B2 a loss ($0.25 \times 8 = 2$) equal to 50% of its net worth (Figure 2, panel (b)). Finally, B2 defaults 50% on its creditors' exposures, B3 and B4 (Figure 2, panel (c)). However, as the loss suffered by an agent cannot be greater than its net worth (Equation 2), this loss is capped at 4 for both B3 and B4. The process stops, as B3 and B4 have no creditors. The SR engendered by an initial shock of 20% on F3's net worth is equal to the total loss excluding the initial shock ($2+4+4 = 10$) divided by the aggregate banks' net worth ($4 \times 4 = 16$) – that is, 62.5%.

Figure 2: *Example of SR computation through the DDR approach*



(a) *(b)* *(c)*

3.2 Classification of critical edges

The next step is to set the critical status of the edges. For each edge e of the network, we compute its impact in the systemic risk $I_{e,\zeta}$:

$$I_{e,\zeta} = \frac{S_\zeta - S_{-e,\zeta}}{S_{-e,\zeta}}, \tag{4}$$

where $S_{-e,\zeta}$ is the systemic risk of the network for an exogenous shock of size ζ computed after the removal of edge e. If $I_{e,\zeta} > 0$, it means the edge has a positive impact on the systemic risk of the network for that level of ζ.

Considering the distribution of $I_{e,\zeta}$, we assign each edge a label $C_{e,\zeta}$ according to

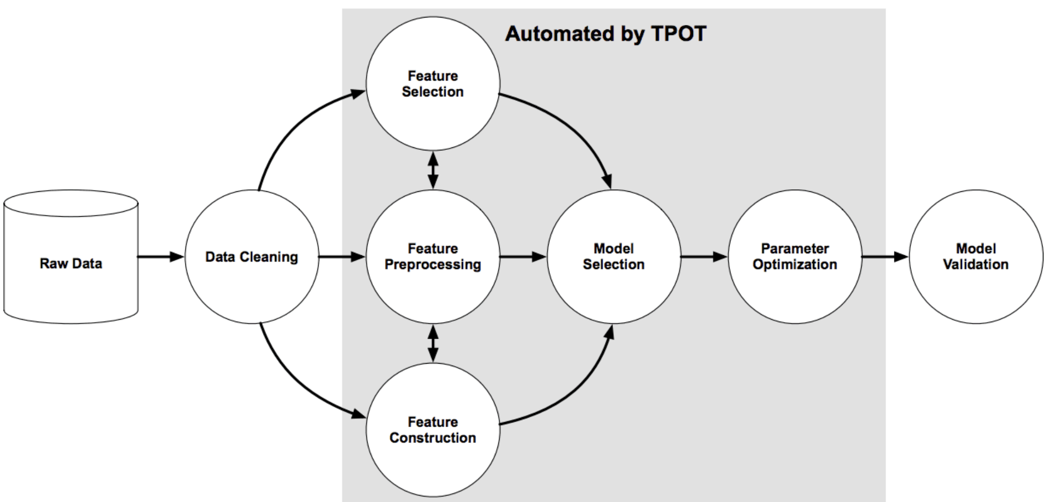
$$C_{e,\zeta} = \begin{cases} 1, & \text{if } Q1 - 1.5IQR > I_{e,\zeta} \mid Q3 + 1.5IQR < I_{e,\zeta} \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where Q1 is the first quartile of the distribution of $I_{e,\zeta}$, Q3 is the third quartile, and IQR is the interquartile range $Q3 - Q1$. Therefore, an edge will be classified as critical if it is an outlier considering the distribution of $I_{e,\zeta}$, according to the IQR criterion.

3.3 Machine learning techniques

The final step is the prediction of the target variables – the critical status of the edges and the sign of the edges' impact – through ML techniques. For this task, we rely on the *Tree-based Pipeline Optimization Tool* (TPOT), a programming-based automated machine learning (AutoML) system (Olson and Moore, 2016). TPOT uses a genetic programming (GP) stochastic global search procedure to efficiently discover the top-performing ML algorithm, as well as the optimal hyperparameters, for a given prediction problem. The search performed by TPOT encompasses all models included in the *Scikit-learn* library, including linear (e.g., linear regression, logistic regression, ElasticNet, etc.) and nonlinear (SVM, tree-based models, XGBoost, neural networks, etc.) models. The TPOT pipeline was used to perform the following tasks: data cleaning, feature selection, feature processing, feature construction, model selection, hyperparameter optimization, and model validation (Figure 3). The set of potential explanatory variables is presented in Table A1.

Figure 3: *Overview of the TPOT pipeline search. Source: Olson et al. (2016)*



For the interpretability of the predictions made by the ML model, we use the Shapley values approach, which originated from the coalition games theory (Shapley, 1953; Shoham and Leyton-Brown, 2008). This methodology informs on how important is a given feature in the prediction of the output, as well as whether this feature is positively or negatively correlated to the output. For the computation of Shapley values, we resort to the SHAP (SHapley Additive exPlanation) framework (Lundberg and Lee, 2017). Let g be an explainer model aiming at predicting an output. A set of M features will be used as inputs. The predicted value for a given data instance is given by

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z'_i, \tag{6}$$

where ϕ_0 is the mean output, ϕ_i is the SHAP value of feature i, and z' is a binary variable indicating whether feature i was included in the model or not. Therefore, the SHAP value ϕ_i indicates to what extent the inclusion of feature i in the model shifts (upwards or downwards) the predicted value from the mean output. If certain properties (local accuracy, missingness, and consistency) are met, ϕ_i corresponds to the original Shapley value (Lundberg and Lee, 2017). The SHAP value of feature i is defined by the following equation:

$$\phi_i = \sum_{S \subseteq M \setminus i} \frac{|S|!(|M| - |S| - 1)!}{M!} [F(S \cup \{i\}) - F(S)]. \tag{7}$$

Therefore, the SHAP value of feature i for a given data instance computes the difference between the predicted value of the instance using all features in S plus feature i, $F(S \cup \{i\})$, and the prediction excluding feature i, $F(S)$. This is weighted and summed over all possible feature vector combinations of all possible subsets S.

4 Results

4.1 General results

We computed the $I_{e,\zeta}$ (Eq. 4) for the bank-firm edges considering two values of ζ: 0.1 and 0.5. The great majority of observations (99.8% for $\zeta = 0.1$ and 82.8% for $\zeta = 0.5$) are positive. It means, in most of the cases, the inclusion of the edge increases the systemic risk. It corroborates the idea that financial networks are *robust-yet-fragile* (Haldane, 2013): if the financial network is sparse enough (what is the case of our network), shock propagation prevails over risk-sharing. Therefore, an increase in the interconnectedness of the network causes an increase in the systemic risk.

The majority of the values of $I_{e,\zeta}$ are very small, as well as the dispersion of these values. The maximum impact is around 2% for $\zeta = 0.1$ and 1% for $\zeta = 0.5$. According to the criteria set in Eq. 5, around 18.7% (22.0%) of the bank-firm edges are critical for $\zeta = 0.1$ ($\zeta = 0.5$). The average impact of critical edges is around 5 times higher than that of all edges. Moreover, the impact is higher for $\zeta = 0.1$. Some statistics of $I_{e,\zeta}$ are presented in Table 2.

Table 2: Statistics of $I_{e,\zeta}$

Statistic	All edges		Critical edges	
	$\zeta = 0.1$	$\zeta = 0.5$	$\zeta = 0.1$	$\zeta = 0.5$
Mean	3.16e-05	1.83e-05	1.60e-04	8.24e-05
Std. dev.	2.26e-04	1.51e-04	5.01e-04	3.14e-04
Min.	-1.92e-04	-1.92e-04	-1.92e-04	-1.92e-04
1^{st} quartile	2.37e-08	1.43e-14	3.00e-05	7.84e-06
Median	6.12e-07	3.48e-08	5.36e-05	1.86e-05
3^{rd} quartile	9.19e-06	2.24e-06	1.26e-04	5.41e-05
Max.	2.06e-02	1.06e-02	2.06e-02	1.06-02

4.2 Predicting edges' critical status

We then applied the TPOT approach in our set of predictive variables (Table A1) to predict the critical status of the bank-firm edges. We split our observations into in-sample (80%) and out-of-sample (20%) data sets. The in-sample data set was oversampled through the Synthetic Minority Oversampling Technique (SMOTE) methodology (Chawla et al., 2002) so the fraction of critical edges reached 50% before being used to train the model. We evaluate the model using accuracy as the performance metric through the *k-fold cross-validation* technique, with $k = 5$ and 10 repetitions.

The model selected by the TPOT pipeline was the *XGBoost* (Friedman et al., 2000) for $\zeta = 0.1$ and the *Random Forest* (Breiman, 2001) for $\zeta = 0.5$. The optimal hyperparameters are presented in Table 3. In the in-sample data set, the model has achieved an accuracy of over 90% (Figure 4), being slightly higher for $\zeta = 0.5$. Table 4 presents some performance metrics for the out-of-sample data set.

***Table 3:** Hyperparameters of the ML models for the prediction of the edges' critical status*

Model	Hyperparameter	Value
XGBoost	learning_rate	0.5
	max_depth	2
	min_child_weight	20
	n_estimators	100
	subsample	0.9
Random Forest	bootstrap	True
	criterion	Gini
	max_features	0.05
	min_samples_leaf	5
	min_samples_split	12
	n_estimators	100

Figure 4: *Histogram of accuracy – in-sample data set. Target variable: edges' critical status*

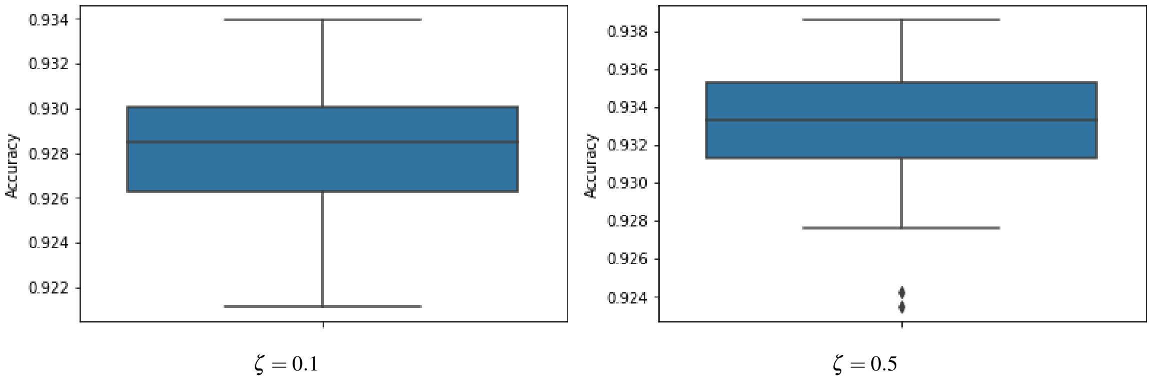


Table 4: *Performance metrics – out-of-sample data set. Target variable: edges' critical status*

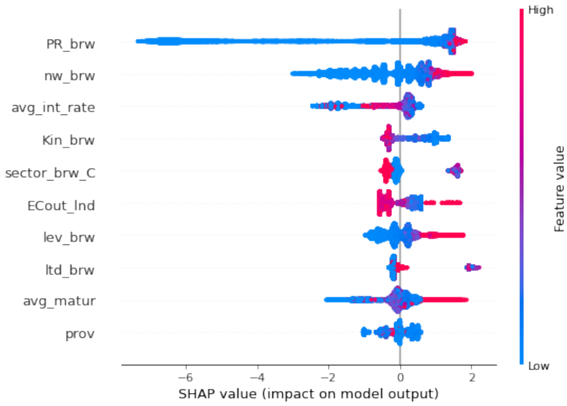
Metric	$\zeta = 0.1$	$\zeta = 0.5$
Precision	0.68	0.70
Recall	0.78	0.90
F1-score	0.72	0.79
Accuracy	0.88	0.89

The SHAP analysis (Figure 5) shows the main drivers of the critical status of the bank-firm edges. The figure can be interpreted as follows: for each feature, it is shown which observations have their predicted output increased (those at the right of the vertical axis) or decreased (those at the left of the vertical axis) after the inclusion of the feature in the model as a predictive variable. On the other hand, the observations closer to the red (blue) color spectrum have a higher (smaller) value of the feature. For instance, for $\zeta = 0.1$, edges with a smaller value of the borrower's PageRank (PR_brw) are concentrated at the left side of the vertical axis – that is, the probability of these edges being critical decreases when this feature is used as a predictive variable. Thus, there is a positive correlation between the probability of an edge being critical and the value of the PageRank of the firm which is the destination node of this edge.

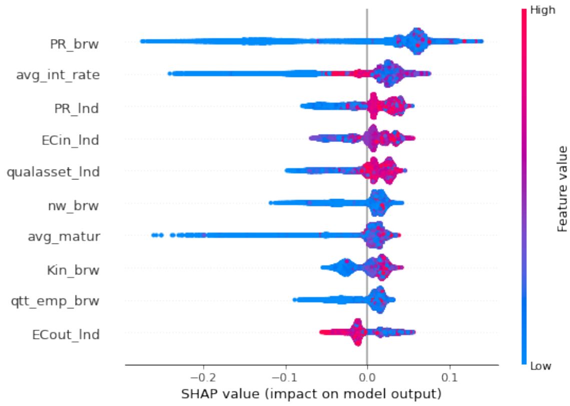
The features are sorted in descending order according to the mean absolute SHAP value. The relative importance of the features in predicting the critical status of the bank-firm edges varies according to the level of the initial shock. For $\zeta = 0.1$, the PageRank of the borrower (that is, the firm) is the main driver of the edge's critical status, followed by the borrower's net worth (nw_brw) and the average interest rate (avg_int_rate). The first two features are positively correlated to the probability of edges being critical. Thus, edges which are loans extended to firms with a large PageRank and a large net worth have a higher probability of being critical.

For $\zeta = 0.5$, the borrower's PageRank is also the main driver of the edge's criticality. The average interest rate appears as the second most important driver. The relationship between this feature and the target variable is nonlinear, as can be seen by the presence of edges with a high interest rate (red observations) in the middle of the horizontal axis. This nonlinear relationship between the interest rate and the edge's criticality is also observed for $\zeta = 0.1$. Comparing the two levels of ζ, the most striking difference is that the firm's net worth loses importance as a driver of the edge critical status. This result is expected, as firms with a not-so-large net worth can cause a large impact under a higher level of initial shock.

Figure 5: *SHAP values, 10 most important features. Target variable: edges' critical status*



$\zeta = 0.1$



$\zeta = 0.5$

Figures B1 and B2 depict the SHAP partial dependence plots of the most relevant continuous features for both values of ζ. These figures show how the expected value of the target variable varies according to the value of the feature. We can observe that some of these features present a nonlinear pattern of interaction with the target variable, corroborating the findings of Figure 5. For instance, in Figure B2, an increase in the interest rate initially leads to a rise in the predicted value of the target for small values of the feature. However, if this value surpasses a given threshold, further increases in the value of the feature lead to a decrease in the predicted value of the target variable.

4.3 Predicting the sign of the edges' impact

A striking difference related to the size of the initial shock regards the fraction of edges with non-positive impact. When $\zeta = 0.1$, virtually all edges (99.8%) have a positive impact. It means the inclusion of these edges leads to an increase in the systemic risk of the network. However, for $\zeta = 0.5$, a significant fraction of the edges has a non-positive impact. Around 14.7% of the edges have a null impact on the systemic risk and 3.1% have a negative impact. Thus, we decided to predict the sign of the edges' impact when $\zeta = 0.5$.

We created a dummy variable equal to one when the impact of the edge is null or negative, and zero otherwise. We followed the same methodological steps described in Section 4.2, but now to predict this dummy variable. The optimal hyperparameters of the model selected by the TPOT pipeline (XGBoost) are presented in Table 5.

Table 5: Optimal hyperparameters of the XGBoost model chosen by the TPOT pipeline to predict the sign of the edges' impact for $\zeta = 0.5$

Hyperparameter	Value
learning_rate	0.1
max_depth	8
min_child_weight	1
n_estimators	100
subsample	0.85

The accuracy in the in-sample data set is above 97% (Figure 6). The model also performed very well in the out-of-sample data set, as can be seen in Table 6.

Figure 6: Histogram of accuracy – in-sample data set. Target variable: sign of the edges' impact, $\zeta = 0.5$

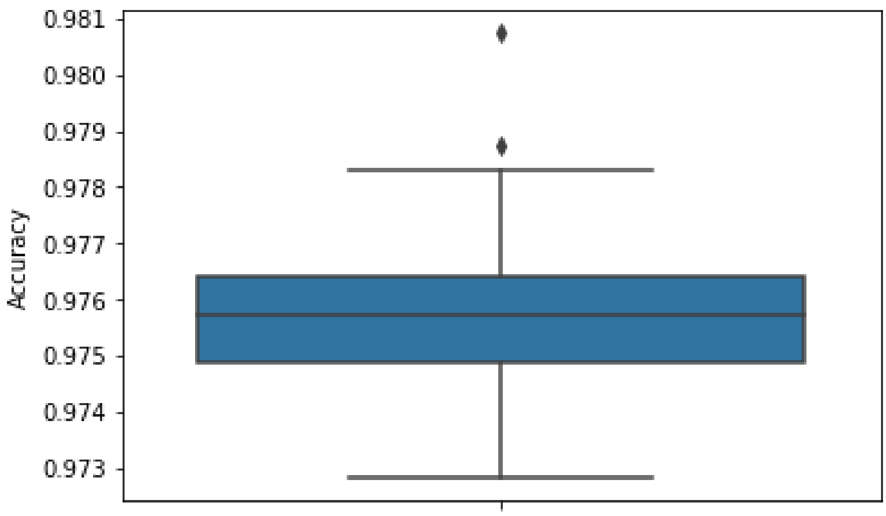


Table 6: Performance metrics – out-of-sample data set. Target variable: sign of the edges' impact, $\zeta = 0.5$

Metric	Value
Precision	0.86
Recall	0.93
F1-score	0.89
Accuracy	0.96

Figure 7 depicts the SHAP values. It shows the sign of the edges' systemic relevance is mostly related to the degree of the lender. The in-degree of the lender is negatively related to the probability of the sign of the edge's impact being non-positive. There is a critical value of the in-degree $k_C \approx 60$ above which this probability drops to zero (Figure 8, panel (a)). Thus, edges whose origin are lenders with a in-degree above this critical value have a greater chance of having a positive impact on the systemic risk. On the other hand, the probability of the edge having a non-positive impact increases with the out-degree of the lender, as can be seen in Figure 8, panel (b).

Figure 7: *SHAP values, 10 most important features. Target variable: sign of the edges' impact,* $\zeta = 0.5$

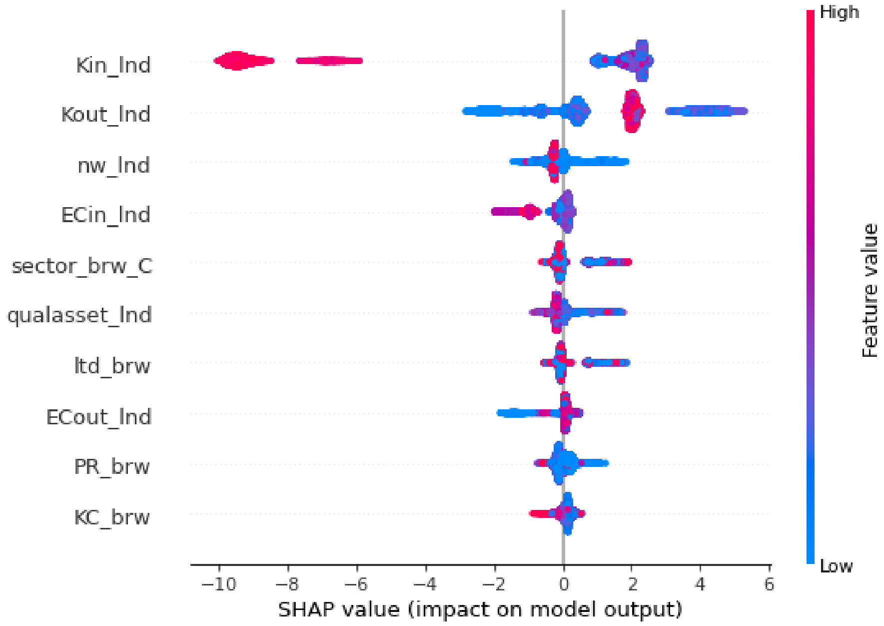
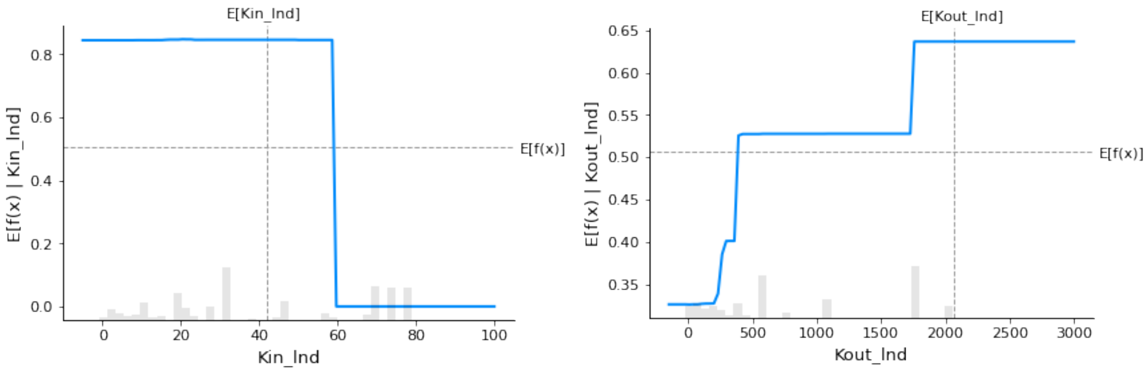


Figure 8: *SHAP partial dependence plots, lenders' in- and out-degree. Target variable: sign of the edges' impact,* $\zeta = 0.5$



(a) ***(b)***

5 Concluding remarks

In this paper, we presented a methodology to identify critical edges in financial networks. The advantage of our approach is that this is explicitly related to the impact of the edge on the systemic risk of the financial network. We first compute the systemic risk of the financial network. Then we assess how this measure of systemic risk is impacted by the removal of the edge. The edges deemed as critical are those which are outliers according to this impact. Thus, a critical edge is that whose removal would cause a large impact on the systemic risk of the network.

We applied this framework to a thorough Brazilian dataset to identify critical bank-firm edges. The majority of the edge impacts are positive. This finding goes in hand with the robust-yet-fragile nature of financial networks: in sparse networks (what is the case of real financial networks), an increase in the interconnectedness (represented by the addition of an edge) would have a positive impact on the systemic risk. The reason is that, for low levels of interconnectedness, shock propagation prevails over risk-sharing. For a smaller level of the initial shock ($\zeta = 0.1$), at least 18% of the edges are critical, but this percentage rises to 22% for $\zeta = 0.5$.

We applied ML-based models to predict the critical status of bank-firm edges, to identify which features drive the size, and – in the case of large shocks – the sign of the edges' impact on the SR. These models proved to have high predictive power for this specific task, showing a level of accuracy of around 90%. A posterior interpretative analysis carried out through Shapley values showed the main predictor of the critical status of bank-firm edges is the PageRank of the borrower — that is, the firm – for both values of ζ. Edges whose destination node has a large PageRank have a higher probability of being critical. The interest rate of the loan and the net worth of the firm are other important drivers of the edge's criticality. While the former feature interacts nonlinearly with the target variable, the latter loses importance for a higher value of the initial shock.

When the value of the initial shock ζ rises from 10% to 50%, an interesting phenomenon emerges: the fraction of edges whose impact is not positive increases from virtually zero to 17.8%. We investigated this phenomenon and found the sign of the edge's impact is mainly related to the degree of its origin node – that is, the lender. Edges whose origin node has a higher in-(out-)degree have a higher probability of having a positive (non-positive) impact on the systemic risk of the whole network.

This study contributes to the literature on the relationship between interconnectedness and systemic risk. Our results show that the impact of removing or adding edges – that is, changes in interconnectedness – on systemic risk is heterogeneous. Some edges are much more influential than others on the systemic risk of the financial network. In addition, we show which features – related to the origin node, the destination node, or the edge itself – identify an edge as critical. The relative importance of the features in predicting the systemic importance of the edge varies according to the level of the initial shock. Our results are also useful for policy-making purposes. They help financial regulators in the identification of financial transactions that will have the greatest impact on systemic risk and should therefore be closely monitored. Finally, by showing the firm's PageRank is the most important driver of the bank-firm edge's critical status, we corroborate the findings of other studies (Alexandre et al., 2021; Ghanbari et al., 2018; Kuzubas et al., 2014; Martinez-Jaramillo et al., 2014), according to which topological variables are at least as important as financial variables in driving systemic risk.

Appendix A Potential explanatory variables

Table A1: *List of potential explanatory variables*

Type	Description of the variable	Acronym
Lenders' F & S variables	Net worth	nw_lnd
	Return over equity	roe_lnd
	Leverage	lev_lnd
	Liquidity	liq_lnd
	Provisions-to-loans ratio	qualasset_lnd
	Dummy variable for private banks	private_lnd
	Dummy variable for foreign banks	foreign_lnd
	Dummy variable for credit unions	cred_union_lnd
Borrowers' F & S variables	Net worth	nw_brw
	Return over equity	roe_brw
	Leverage	lev_brw
	Bank debt-to-total debt ratio	bnkdebt_brw
	Market debt-to-total debt ratio	mktdebt_brw
	ROF debt-to-total debt ratio	rofdebt_brw
	Number of employees	qtt_emp_brw
	Average wage	avg_wage_brw
	Share of female employees	share_women_brw
	Age of the firm	age_brw
	Short-term external debt ratio	extdebtst_brw
	Dummy for Limited Society	ltd_brw
	Dummies for the economic sector	sector_brw_k^*
Topological variables**	In-degree	Kin
	Out-degree	Kout
	Core number	KC
	Closeness centrality (incoming links)	CCin
	Closeness centrality (outgoing links)	CCout
	Betweenness centrality	B
	Eigenvector centrality (incoming links)	ECin
	Eigenvector centrality (outgoing links)	ECout
	PageRank	PR
Transaction variables	Average interest rate	avg_int_rate
	Average maturity	avg_matur
	Provision-to-value ratio	prov

(*): k refers to the economic sector, represented by letters from A to S.

(**): The suffix "_lnd" (for lenders) or "_brw" (for borrowers) will be added to the acronym.

Appendix B SHAP partial dependence plots

Figure B1: *SHAP partial dependence plots, most important features. Target variable: edges' critical status,* $\zeta = 0.1$

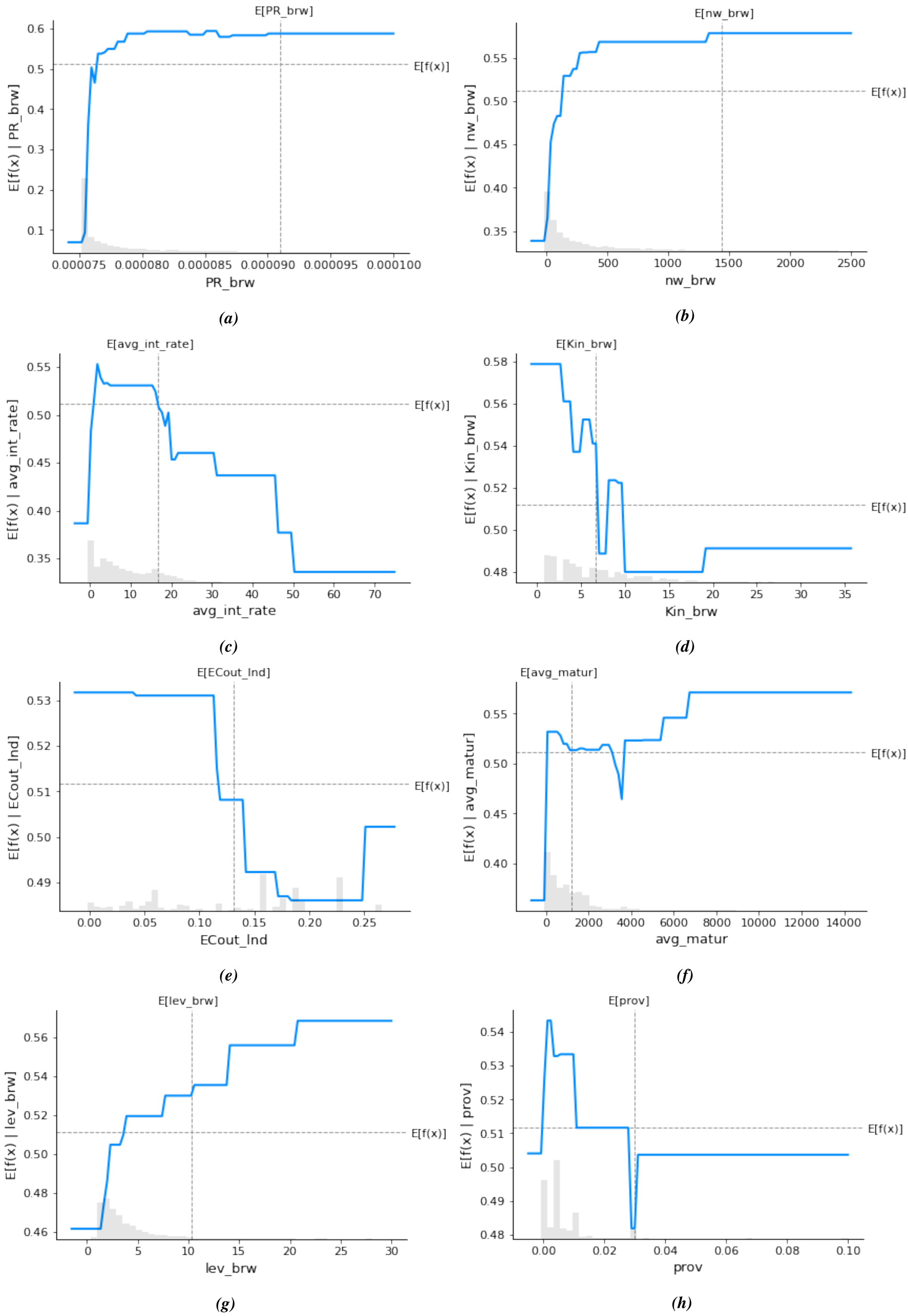
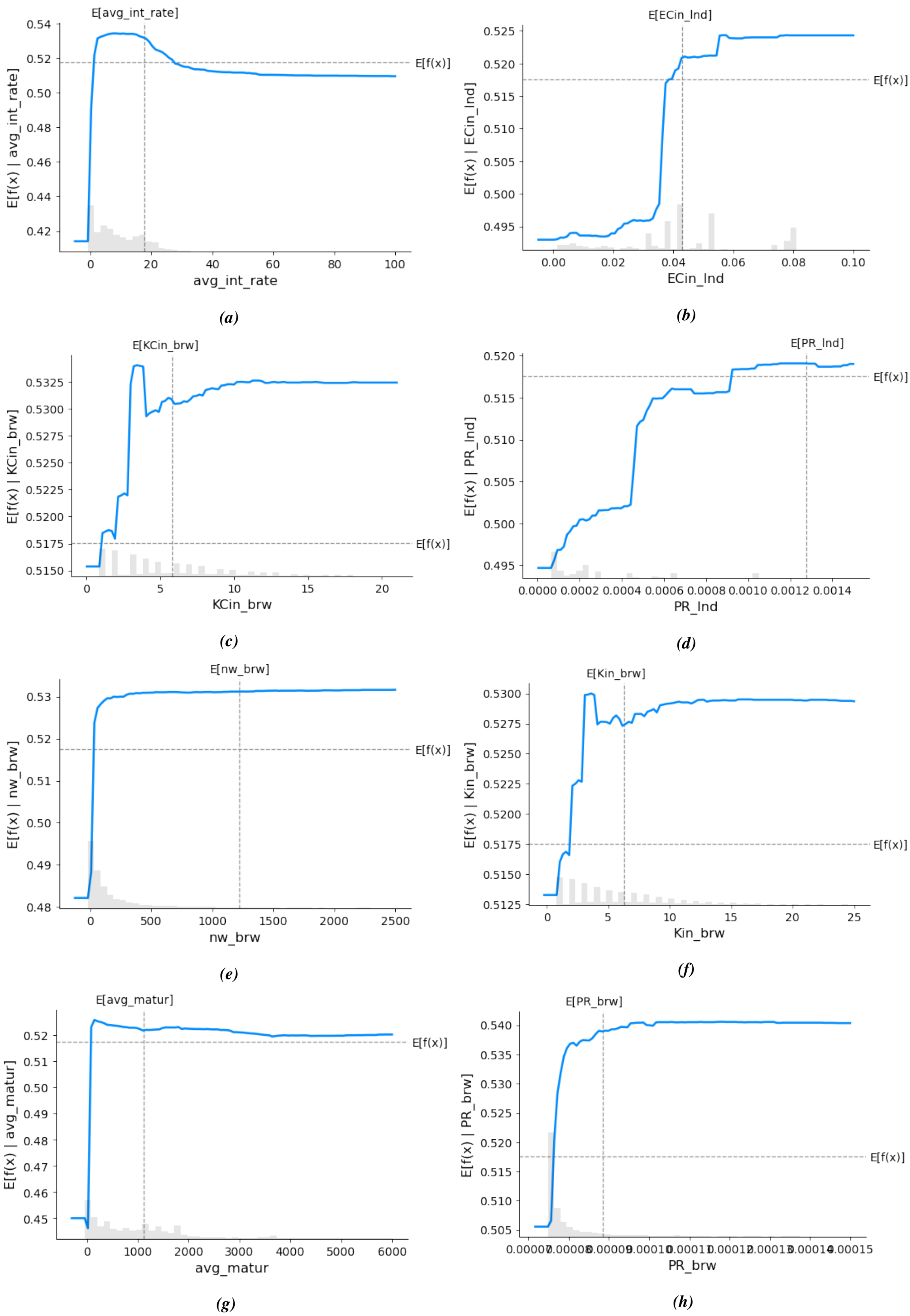


Figure B2: *SHAP partial dependence plots, most important features. Target variable: edges' critical status,* $\zeta = 0.5$



(a) *(b)* *(c)* *(d)* *(e)* *(f)* *(g)* *(h)*

References

Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608.

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47.

Alexandre, M., Silva, T. C., Connaughton, C., and Rodrigues, F. A. (2021). The drivers of systemic risk in financial networks: a data-driven machine learning analysis. *Chaos, Solitons & Fractals*, 153:111588.

Alexandre, M., Silva, T. C., Michalak, K., and Rodrigues, F. A. (2023). Does the default pecking order impact systemic risk? Evidence from Brazilian data. *European Journal of Operational Research*, 309(3):1379–1391.

Allen, F. and Gale, D. (2000). Financial contagion. *Journal of Political Economy*, 108(1):1–33.

Bardoscia, M., Battiston, S., Caccioli, F., and Caldarelli, G. (2015). Debtrank: A microscopic foundation for shock propagation. *PloS One*, 10(6):e0130406.

Bardoscia, M., Caccioli, F., Perotti, J. I., Vivaldo, G., and Caldarelli, G. (2016). Distress propagation in complex networks: the case of non-linear debtrank. *PloS One*, 11(10):e0163825.

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. *Journal of Financial Stability*, 35:6–16.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Bröhl, T. and Lehnertz, K. (2019). Centrality-based identification of important edges in complex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(3).

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16:321–357.

Chinazzi, M. and Fagiolo, G. (2015). Systemic risk, contagion, and financial networks: A survey. *Institute of Economics, Scuola Superiore Sant'Anna, Laboratory of Economics and Management (LEM) Working Paper Series No. 2013/08.*

De la Cruz Cabrera, O., Matar, M., and Reichel, L. (2020). Edge importance in a network via line graphs and the matrix exponential. *Numerical Algorithms*, 83:807–832.

De Masi, G. and Gallegati, M. (2012). Bank–firms topology in Italy. *Empirical Economics*, 43(2):851–866.

Docha, F. A. and Rodrigues, L. B. P. (2023). Industry and securities market influences on indebtedness: evidence from a large dataset from Brazil. *Central Bank of Brazil Working Paper Series No. 586*.

Friedman, J., Hastie, T., Tibshirani, R., et al. (2000). Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *Annals of statistics*, 28(2):337–407.

Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, H. E. (2003). A theory of power-law distributions in financial market fluctuations. *Nature*, 423(6937):267–270.

Gai, P. and Kapadia, S. (2010). Contagion in financial networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466(2120):2401–2423.

Ghanbari, R., Jalili, M., and Yu, X. (2018). Correlation of cascade failures and centrality measures in complex networks. *Future Generation Computer Systems*, 83:390–400.

Glasserman, P. and Young, H. P. (2015). How likely is contagion in financial networks? *Journal of Banking & Finance*, 50:383–399.

Glasserman, P. and Young, H. P. (2016). Contagion in financial networks. *Journal of Economic Literature*, 54(3):779–831.

Hajarathaiah, K., Enduri, M. K., Anamalamudi, S., Abdul, A., and Chen, J. (2024). Node significance analysis in complex networks using machine learning and centrality measures. *IEEE Access*, 12:10186–10201.

Haldane, A. G. (2013). Rethinking the financial network. In *Fragile stabilität–stabile fragilität*, pages 243–278. Springer.

Helander, M. E. and McAllister, S. (2018). The gravity of an edge. *Applied Network Science*, 3(1):1–35.

Kanwar, K., Kaushal, S., Kumar, H., Gupta, G., and Khari, M. (2022). BC_{DCN}: a new edge centrality measure to identify and rank critical edges pertaining to SIR diffusion in complex networks. *Social Network Analysis and Mining*, 12(1):49.

Kuzubas, T. U., Omercikoglu, I., and Saltoglu, B. (2014). Network centrality measures and systemic risk: An application to the Turkish financial crisis. *Physica A*, 405:203–215.

Lundberg, S. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. *arXiv preprint arXiv:1705.07874*.

Lux, T. (2016). A model of the topology of the bank–firm credit network and its role as channel of contagion. *Journal of Economic Dynamics and Control*, 66:36–53.

Mantegna, R. N. and Stanley, H. E. (1995). Scaling behaviour in the dynamics of an economic index. *Nature*, 376(6535):46–49.

Martinez-Jaramillo, S., Alexandrova-Kabadjova, B., Bravo-Benitez, B., and Solórzano-Margain, J. P. (2014). An empirical study of the Mexican banking system's network and its implications for systemic risk. *Journal of Economic Dynamics & Control*, 40:242–265.

Martinez-Jaramillo, S., Carmona, C. U., and Kenett, D. Y. (2019). Interconnectedness and financial stability. *Journal of Risk Management in Financial Institutions*, 12(2):168–183.

May, R., Levin, S., and Sugihara, G. (2008). Complex systems: Ecology for bankers. *Nature*, 451(7181):893–895.

Nier, E., Yang, J., Yorulmazer, T., and Alentorn, A. (2007). Network models and financial stability. *Journal of Economic Dynamics and Control*, 31(6):2033–2060.

Olson, R. S., Bartley, N., Urbanowicz, R. J., and Moore, J. H. (2016). Evaluation of a tree-based pipeline optimization tool for automating data science. In *Proceedings of the genetic and evolutionary computation conference*, pages 485–492.

Olson, R. S. and Moore, J. H. (2016). TPOT: A tree-based pipeline optimization tool for automating machine learning. In *Workshop on automatic machine learning*, pages 66–74.

Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J., and Barabási, A.-L. (2007). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, 104(18):7332–7336.

Seabrook, I. E., Barucca, P., and Caccioli, F. (2021). Evaluating structural edge importance in temporal networks. *EPJ Data Science*, 10(1):23.

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, 2(28):307–317.

Shoham, Y. and Leyton-Brown, K. (2008). *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press.

Song, J.-H. (2023). Important edge identification in complex networks based on local and global features. *Chinese Physics B*, (32):098901.

Yu, E.-Y., Chen, D.-B., and Zhao, J.-Y. (2018). Identifying critical edges in complex networks. *Scientific Reports*, 8(1):14469.

Zhao, N., Li, J., Wang, J., Li, T., Yu, Y., and Zhou, T. (2020). Identifying significant edges via neighborhood information. *Physica A: Statistical Mechanics and its Applications*, 548:123877.