



Assessing the risk of default propagation in interconnected sectoral financial networks

Adrià Barja¹, Alejandro Martínez¹, Alex Arenas², Pablo Fleurquin¹, Jordi Nin³, José J. Ramasco^{4*}  and Elena Tomás¹

*Correspondence:

jramasco@ifisc.uib-csic.es

⁴Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Palma de Mallorca, Spain
Full list of author information is available at the end of the article

Abstract

Systemic risk of financial institutions and sectoral companies relies on their inter-dependencies. The inter-connectivity of the financial networks has proven to be crucial to understand the propagation of default, as it plays a central role to assess the impact of single default events in the full system. Here, we take advantage of complex network theory to shed light on the mechanisms behind default propagation. Using real data from the BBVA, the second largest bank in Spain, we extract a financial network from customer-supplier transactions among more than 140,000 companies, and their economic flows. Then, we introduce a computational model, inspired by the probabilities of default contagion, that allow us to obtain the main statistics of default diffusion given the network structure at individual and system levels. Our results show the exposure of different sectors to default cascades, therefore allowing for a quantification and ranking of sectors accordingly. This information is relevant to propose countermeasures to default propagation in specific scenarios.

Keywords: Financial networks; Default analysis; Financial sector analytics; SIS propagation models; Complex systems

1 Introduction

Interconnected financial networks are the fabric where economic agents from different sectors operate. One of the main challenges we face nowadays on financial networks is assessing systemic risk [1–3]. In the literature, systemic risk is defined as the probability of having large cascades of entangled economic events. Such cascades are triggered by causes that range from exogenous shocks to endogenous defaults. Besides, the succession of several defaults can jeopardize the full system because network financial inter-dependencies act as an economic sounding board. The interplay between the topology of the underlying interaction network and the easiness with which events propagate have proven to be essential to understand the proportion of the financial system affected by default avalanches and to assess the systemic risk [4].

Avalanches in financial systems are understood as dynamical processes that correlate individual economic states of the agents when a stress event materializes. This process resembles epidemic spreading in networks [5]. Under a simplifying assumption and to better explore the network potential for risk transmission, we model them in a similar way as epidemic spreading as was recently done in [6–8]. This is an oversimplification, the basic

mechanisms of default propagation and epidemic spreading have similarities as branching process and chain reactions. Still in economic systems, especially those involving large companies such as banks, there are other mechanisms that may delay or even prevent the final default. Given the simplicity of the epidemic models and the fact that our networks are mostly formed by small and medium enterprises, we have taken this approach in the hope of getting direct information on the multi-layer sectoral network interdependencies and on how the risk can pass from one to the other. The failure of one subject in the financial network generates a chain reaction through interconnections and causes shocks and therefore a default risk. This risk is understood as the incapability of one of the participants to perform their obligations, or at least to accomplish them properly, which leads to the interruption in the obligation payments of other participants.

One of the most commonly used contagion propagation models corresponds to the celebrated Susceptible–Infected–Susceptible (SIS). In a SIS model, individuals that are cured do not develop permanent immunity, but are again susceptible to the “disease”. Similarly, companies that manage to escape default by overcoming high economic stress can fall into trouble again later on. Additionally, SIS model provides valuable insights to understand how different situations may affect the outcome of the contagion process, *e.g.* what the most efficient technique is for isolating a limited number of companies in a given financial network to minimize the risk of observing an avalanche. Epidemic modeling is still the main application of SIS-like approaches, and the main driver behind the development and refinement of this framework through time. However, the contagion analogy has been applied in different contexts and in particular in those where it is important to consider the spatial and social structure of systems. Some examples are adoption of fads and innovations [9], propagation of news and rumors [10] and information diffusion [11]. These are phenomena for which the state of the agent is affected by the interaction with its neighbors. In the financial context there is a strong causal relation between the financial and economical state of a company’s clients and how this influences its economical well-being [12]. This dynamics resembles a Hawkes stochastic process [13], where one event, under certain circumstances, is able to generate a new set events allowing the diffusion of a given phenomena [14]. Under this hypothesis, epidemic modeling can shed light on how systemic risk propagates through financial networks. Besides the contagion analogy, there are other similarities between the transmission of diseases and the transmission of financial distress in financial networks. For example, both are branching processes where one event produces others. But we can also observe some differences. For instance, disease transmission is usually studied as a continuous phenomena whereas financial distress is studied in a discrete time scale. Also, there are different levels of homogeneity in both cases, usually financial networks are more heterogeneous than the population networks used in disease spreading research.

In this paper we provide a mechanistic model to assess the impact of a particular diffusion process of default on financial networks. To this end, we take as basis a probabilistic computational framework named microscopic Markov chain approach (MMCA) to compute the probability of the states of individual agents in contagion processes in complex networks [15–17], and adapt its formulation to the understanding of the default propagation in financial networks. Further, we analyze the behavior of our proposed model using real data from the anonymized database of BBVA from December 2015 to December 2016, covering around 140,000 public and private Spanish firms. We set default labels to 0 or 1

based on this data, depending on whether a given company was or not in default at the beginning of the considered period. By means of this data we have access to the real network of interactions and to the initial condition for the dynamics of the default endogenous propagation.

This paper is organized as follows. First, we review previous work on default propagation in financial networks. Next, in Sect. 3, we propose a contagion model adapting the well-known SIS Model. Section 4 provides a complete description of the data used, and its topological analysis. Section 5 includes a set of experiments to study the main characteristics of default propagation in each inter-connected sector. Later, we examine the implications of using our default propagation model in Sect. 6. Finally, Sect. 7 provides some conclusions and future work.

2 Related work

The use of networks in economy and finance has a long tradition. Initial works were conceptual, like [18], where the networks were proposed as tools to represent the interactions (as links) between economic agents playing the role of nodes. When data started to become available, the popularization of complex networks brought a change of paradigm, leading to several advances in the field.

For example, the properties of the economic interchange networks between countries were studied in [19]. Also, the network formed by companies holding shares of other companies was studied for the Milan, New York and NASDAQ stock exchange markets in [20]. Interestingly, these networks show a scale-free nature, which implies that investors having a large number of connections are not uncommon. Explanations for this have been searched in the network dynamics properties mixed with a “rich-gets-richer” effect by means of different approaches [21, 22]. More recent models have been also proposed in [23, 24], assuming different hypothesis. A complete review of empirical economic network models can be found at [25]. Beyond the distribution of connections, other characteristics such as the level of clustering have been studied [26]. Despite all these works, there are still numerous open challenges when it comes to fully understanding the structure and dynamics of economic and financial networks [1, 27, 28].

The reason why these networks attract so much attention is that, besides economic interchanges, financial risk also propagates through them [24, 29–32]. Their stability becomes thus an important question [33]. Furthermore, risk and economic distress, and even default in a second stage, can occur in cascades leading to serious systemic instabilities [32, 34]. Therefore, the resilience of the networks to contagion, as well as the circumstances under which it becomes systemic has been analyzed in many works [35–37]. Following this research line, a method called debtrank was introduced to find nodes in financial networks that can induce large cascades when perturbed [3]. This method allows to search for measures to mitigate risk propagation [38]. In the special case of networks where the nodes are banks and the links represent holding of different types of obligations, the complexity of the products traded such as derivatives [39] and the feedback-loops between solvency perception and stock and obligation values [40] can play an important role in economic distress propagation. Many of these previous works have been focused on banking [41, 42], where the risk propagation is related to the stress tests performed by central banks. These kind of models resorts on ad-hoc mathematical models for financial institutions. However, as it has been seen in the last crisis, the risk can spill out of the banking system to enter other economic sectors. This is why it is of high relevance to consider risk

propagation in more general economic networks, including different sectors and different types of nodes, ranging from large holdings to small companies or even the final individual consumers. This is precisely the direction that we take in the present work where we use general contagion model to evaluate the default spreading in a highly heterogeneous network.

3 Default contagion model

Inspired by the Microscopic Markov Chain Approach (MMCA) designed for epidemic spreading, first we propose an adaptation of the framework for modeling the default cascades observed in the transactions between different companies in real financial networks. Then we introduce some measures to dynamically analyze the default contagion process and its functional relations with any sectoral financial network.

3.1 MMCA model for default contagion

The original MMCA model was designed to cope with the propagation of epidemics [15], where the states of the agents (nodes) forming the network of contacts were binary, namely, susceptible or infected. In well-mixed populations, the differential equations governing the number of susceptible (S) and infected (I) individuals are

$$\begin{aligned}\frac{dS}{dt} &= -\tilde{\beta}S\frac{I}{N} + \tilde{\mu}I, \\ \frac{dI}{dt} &= \tilde{\beta}S\frac{I}{N} - \tilde{\mu}I,\end{aligned}\tag{1}$$

where $N = S(t) + I(t)$ is the (constant) size of the population. The term I/N accounts for the probability of contacting an infected individual in a well-mixed population of size N , $\tilde{\beta}$ is the infectivity rate (probability per unit time) for each contact, and $\tilde{\mu}$ is the rate at which one infected individual recovers. Their corresponding differential equations are

$$\begin{aligned}S(t + \Delta t) &= S(t) \left(1 - \tilde{\beta} \Delta t \frac{I(t)}{N} \right) + \tilde{\mu} \Delta t I(t), \\ I(t + \Delta t) &= I(t) \left(1 - \tilde{\mu} \Delta t + \tilde{\beta} \Delta t \frac{I(t)}{N} S(t) \right),\end{aligned}\tag{2}$$

or equivalently

$$I(t + \Delta t) = I(t) - \tilde{\mu} \Delta t I(t) + \tilde{\beta} \Delta t \frac{I(t)}{N} [N - I(t)].\tag{3}$$

Defining $\rho(t) = I(t)/N$ as the fraction of infected individuals in the population, Eq. (3) is written as

$$\rho(t + \Delta t) = \rho(t) - \tilde{\mu} \Delta t \rho(t) + \tilde{\beta} \Delta t \rho(t) [1 - \rho(t)].\tag{4}$$

Note that this discrete equation can be mapped to networks, and in a microscopic approximation, the density of infected individuals will correspond to the individual probability of being infected.

In the case of our economic networks, the state of the different agents corresponds to their liquidity in time according to their initial state and economical activity. To keep

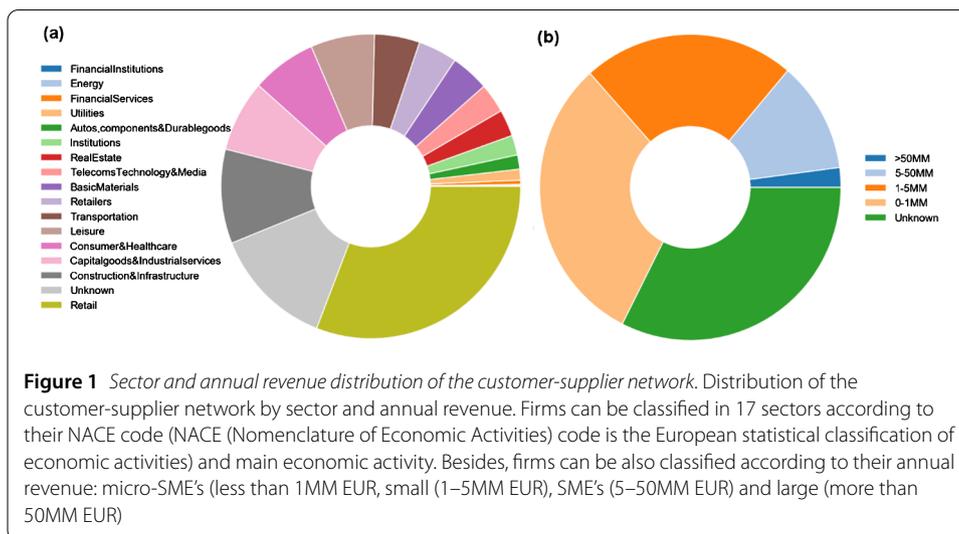
things as simple and realistic as possible, we assume a discrete-time version of the SIS model in networks, and adopt its terminology, defining $\mu = \tilde{\mu} \Delta t$ as the probability of default recovery, and $\beta = \tilde{\beta} \Delta t$ as the probability of default contagion. In this setting, μ reflects the average recovery time a company needs to overcome a default process, whereas β reveals the average number of interactions with an infected company required to be defaulted. Finally, the network of contacts, in this case, is represented as a weighted matrix, with matrix elements w_{ij} accounting for the total sum of money transfers from node i to node j in 12 months (from a fixed date). The equation governing the default system is then described as:

$$p_i(t + 1) = (1 - q_i(t))(1 - p_i(t)) + (1 - \mu)p_i(t) + \mu(1 - q_i(t))p_i(t), \tag{5}$$

where $p_i(t)$ is the probability of a node to be in default; $q_i = \prod_{j=1}^N (1 - \beta w_{ji} p_j(t))$ the probability that a given node i is not infected by any of its neighbors. The right hand side of Eq. (5) is explained as follows: $(1 - q_i(t))(1 - p_i(t))$ is the probability that a given node i is susceptible of entering into default $(1 - p_i(t))$ and it is infected, $(1 - q_i(t))$, by at least one of its neighbors in default. The term $(1 - \mu)p_i(t)$ is the probability that a node i in default does not recover $(1 - \mu)$. Finally, the term $\mu(1 - q_i(t))p_i(t)$ corresponds to the probability that a given node i recovers from default but is reinfected by at least one of its neighbors already in default $(1 - q_i(t))$.

According to the European Central Bank definition for risk classification [43], the susceptible state would correspond to a company which is in step 3 (default). In this step, the credit quality of the company is considered equivalent to a probability of default of between 0.10% and 0.40% over a one-year horizon. Therefore, after a given period of time (12–18 months), which depends on its revenue, it can go through the step 2 (cure) and finally come back to step 1 (normal) if it proves to have a good payment behavior.

By using this model, we computationally analyze the behavior of default contagion processes in a real topology created by the interactions among different companies. Additionally, we want to understand the main properties of default propagation and the emergent clusters containing defaulted companies in the full system. Moreover, having a well-defined sector distribution and their annual revenue, see Fig. 1, we can elucidate if default



cascades depend on sectors and their economical states. Our objective here is not to build an accurate description of the default process, since for this purpose our model should be more complex and realistic, but to explore the role that the network structure plays on the default propagation process with varying economic scenarios.

3.2 Dynamical analysis of default contagion

To analyze the potential for default propagation in a financial network using the defined contagion model, we define a default probability density ρ as

$$\rho = \frac{\sum_{i=1}^N p_i^{\text{ss}}}{N}, \quad (6)$$

where N is the number of companies and p_i^{ss} stands for the default probability of company i at the model's stationary state. Note that since we are considering a SIS modeling framework, by construction, the dynamics will always reach a steady state. As shown in Sect. 3.1, the default probability density depends on the default infection rate β and the recovery rate μ . For obtaining ρ , we monitor $p_i(t)$ as a function of time in a discrete manner, until the contagion model reaches the steady state. Note that time represents the iterations of the MMCA recursive model. To do so, we set $p_i(0)$ to the real company default label. For understanding the system dynamics, without loss of generality, we can apply the classic MMCA framework where the parameter, β_i , the contagion infectivity per node, is constant and equal for all companies in the network.

However, to make this model more realistic we introduce a variation to this setting using default recovery probability μ dependent on the relative in-degree of each company i . A particular μ_i for a company i is therefore defined as

$$\mu_i = \frac{k_i}{\max_{j=1}^N(k_j)}, \quad (7)$$

where $\max_{j=1}^N(k_j)$ is the maximum in-degree in the network. The intuition behind this heterogeneity in the recovery parameter μ_i is that companies having large number of customers will recover faster than those whose market risk is concentrated in a few customers. Note that this is just one of the possible variations of our default contagion framework. For instance, the generalization of the heterogeneous μ to other company characteristics/features as balance sheet information or any other individual attribute is straightforward. Moreover, not only the recovery rate but also the infection probability β can be defined as company dependent, and/or even depend on group of companies such as economical sector (C). So, Eq. (5) can be generalized as

$$\begin{cases} p_i(t+1) = (1 - q_i(t))(1 - p_i(t)) + (1 - \mu_i)p_i(t) + \mu_i(1 - q_i(t))p_i(t), \\ q_i = \prod_{j=1}^N (1 - \beta)w_{ji}p_j(t). \end{cases} \quad (8)$$

Accordingly, the proposed modification, Eq. (7), to obtain a in-degree dependent μ is one of the simplest approaches, since it only varies a parameter using a topological characteristic such as the relative in-degree.

3.3 Dynamical properties of the network: onset slope and sensitivity

Naturally, the observed dynamical behavior is the result of an interplay between the MMCA framework and the network topology. To understand the reasons behind the different sectors' response to default propagation, we will characterize the dynamical behavior of each sector by the onset threshold \mathcal{R}_0 and the *sensitivity* S to the initial set of defaulted companies. Both metrics are descriptors of the expected behavior in the steady state regime. The onset slope is measured by estimating numerically the critical $\beta_c(s)$ at which the first default cases start to appear in sector C . Practically, fixing μ , the parameter β is increased until the number of default cases in the sector in the stationary state goes over 1% of the real data defaults found in the sector, marking $\beta_c(s)$. All these calculations are done in the stationary state of the system. When $\beta_c(s)$ is plotted versus μ , one finds a noisy linear increase and, therefore, we define \mathcal{R}_0 as the slope of the linear fit.

This property reveals the spreading capacity of the infectious process in each sector. Larger values imply that when the life times of defaults in the companies of the sector become shorter, one needs higher infectivity to overcome the threshold. Sectors with larger \mathcal{R}_0 should be more resilient to general default. From the moment they start to show significant default, other sectors with lower \mathcal{R}_0 may be in very bad shape already. Furthermore, given a certain set of parameters, an isolated default event in one of the sectors with larger \mathcal{R}_0 can trigger an avalanche of default on weaker sectors, for which the conditions are favorable for contagion. In this sense, the \mathcal{R}_0 value of a sector is also related to the capacity of the sector to spread default.

Regarding the sector sensitivity to default propagation, this dynamical property measures the rate of change of $\rho(\beta, \mu)$ at the transition point (which is normally known as β cut-off). Computing sensitivity involves fitting a linear regression to the model response and using its standardized regression coefficients as direct measures of sensitivity. Therefore this metric describes how susceptible a sector is to default, quite the opposite to \mathcal{R}_0 , which characterizes how a sector affects the system. The relationship between these two dynamical descriptors and the network structure will be explored next. As mentioned before, these are defined at the steady state, but it is also important to understand how dynamics evolve in the transient regime. This analysis will be carried out by synthetically concentrating defaulted companies (in specific proportions) in the different sectors and exploring pair-wise sectoral interactions at the initial steps of the simulation.

3.4 How do sectoral properties of the nodes affect network dynamics?

Since economic crisis often start in a given sector and later expand to others (see for example the 2000 energy crisis and the 2008 financial crisis), we are also interested in exploring the dependence of default propagation on attributes related to each sector, such as sectoral default probability density and sectoral inter-connectivity. In particular, to study the dependence on the sectoral default probability density we rewrite Eq. (6) for a each sector (C) as:

$$\rho_C = \frac{\sum_{i \in C} P_i^{ss}}{N_C}. \quad (9)$$

We also propose a metric of in-sectoral inter-connectivity I_{in} that measures how well, on average, a sector is connected to other sectors by incoming links:

$$I_{in} = \frac{\sum_{i,j} (w_{ij} - \frac{s_i^{in}}{17})^2}{N_C}, \quad (10)$$

where N_C is the number of companies a certain sector C has, $i \in C$, w_{ij} are incoming links of i ($j \rightarrow i$) and s_i^{in} is the company i in-strength. In short, the in-sector inter-connectivity measures the mean square error with respect to a hypothetical equally distributed situation where we have a node with incoming weights from all C sectors with equal probability ($\frac{s_i^{in}}{C}$). So, the larger the value, the more heterogeneously connected the sector is to other sectors by incoming links.

4 Topological analysis of the client-supplier network

Now we describe the sectoral financial network used in the experiments carried out in this work. To do so, we first provide all the details about network construction. Second, we report commonly used network statistical descriptors.

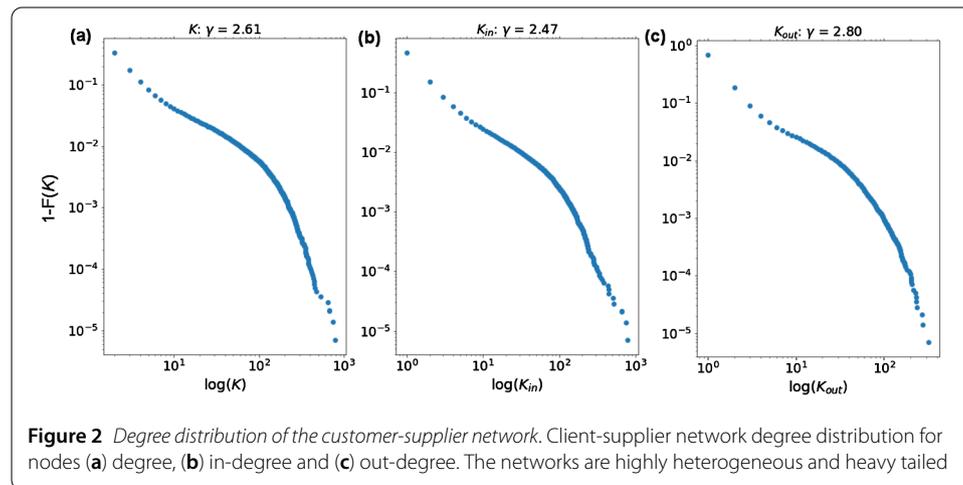
4.1 Network construction

Customer-supplier relationships highly depend on economical sectors and the financial situation of the companies involved. To properly model this situation with real data we gathered anonymized quarterly data from the official customer-supplier third party payment declarations collected by the BBVA risk management department. This declaration is used as a mechanism to avoid fraud in company VAT declarations. here, Spanish firms (our nodes) inform about their supplier payments and customer earnings. For each available company, we extracted its operating revenue and financial statement attributes: sector and default status. Collected data covers from December 2015 to December 2016. Default labels at the initial step were set to 0 or 1 depending on whether a company was in default or not at the beginning of this period. By using customer-supplier relationships, and after removing self-loops, a directed and weighted network with 142,477 nodes and 255,509 edges was obtained. Direction of edges follows the path of money injection (from the customer to the supplier). All edge weights (total money transferred) were aggregated annually and normalized by its source node out-strength. Note that, both BBVA customers and non-customers were included in a percentage of 63% and 37%, respectively. Therefore, the network contains an important percentage of missing values.

As illustrated in Fig. 1, most of the companies included in the network are micro-SME's and small companies, with an annual revenue smaller than 5 million euros (more than 80% of the informed values). Besides, the most common categories are retail (shops), followed by construction & industrial companies. It is important to mention that leisure and consumer & healthcare are also important sectors in the network. Although nodes belonging to energy and financial institutions are only a few, they account for 50% of the network's out-strength, whereas other sectors such as retail are relatively abundant in the network (30% of the nodes) but only account for 3% of the total out-strength. To model the probability of contagion, we normalized the edge weights by the out-strength of the source node (customer). As a result, Energy and Financial Institutions account for the 11% of the normalized out-strength in the network, and Retail for the 5% (20% and 7% of the corresponding global in-strength). See Table 1 for more details.

Table 1 Summary of network topological measures by sector where \bar{k}_{in} and \bar{k}_{out} stand for the average in-degree and out-degree, respectively

Sector	Size (%)	\bar{k}_{in}	\bar{k}_{out}	Default (%)	Rank hub	Rank auth
Financial institutions	0.046	39.613	45.529	3.650	17	1
Energy	0.083	12.844	8.666	1.111	14	2
Financial services	1.165	6.300	20.265	0.786	13	3
Utilities	1.529	5.589	5.903	1.264	11	4
Telecoms technology & media	3.299	5.960	5.194	1.776	2	5
Basic materials	2.745	5.789	5.350	2.782	6	6
Transportation	4.064	5.411	4.336	1.868	1	7
Retail	23.593	3.973	3.233	1.217	12	8
Retailers	4.273	5.001	3.613	1.885	5	9
Capital goods & industrial services	8.689	4.528	3.098	1.866	9	10
Autos, components & durable goods	1.470	4.454	2.991	1.786	10	11
Consumer & healthcare	7.055	3.259	3.770	1.539	8	12
Construction & infrastructure	8.907	3.067	3.270	3.071	3	13
Unknown	10.159	0.930	1.413	1.942	15	14
Real estate	6.843	1.517	1.844	3.603	7	15
Leisure	12.861	2.509	2.512	1.511	4	16
Institutions	3.219	5.547	10.764	0.535	16	17



4.2 Statistical descriptors

The first point we address is the degree distribution of the network. The degree distributions of the real network analyzed (total, in- and out-degree) presents a heavy-tail but does not fit to a power-law. This heavy-tail is relevant for the sake of the analysis given that it clearly indicates the existence of hubs. The complementary cumulative distribution functions are displayed in Fig. 2. Table 1 also sorts business sectors according to their average hub and authority score [44]. The main hub in the network is the transportation sector, followed by telecoms, Technology & media and construction & infrastructure sectors. The main authority is related to financial institutions, followed by energy and financial services. Hubs and authorities agree with the expected economic behavior. Hub sectors such as transportation, Construction & Infrastructure and telecoms, Technology & media, which require energy resources to produce goods, transport them, or even to run IT services, therefore those sectors have important connections to authorities such as energy. Note also sectors such as financial institutions and services (credit cards, insurances) arising as important authorities. This is a direct result of the financial needs that many the

companies have to operate in Spain. This need is mostly because the average number of days required to collect invoiced amounts from customers is quite high in Spain, around 90 days.

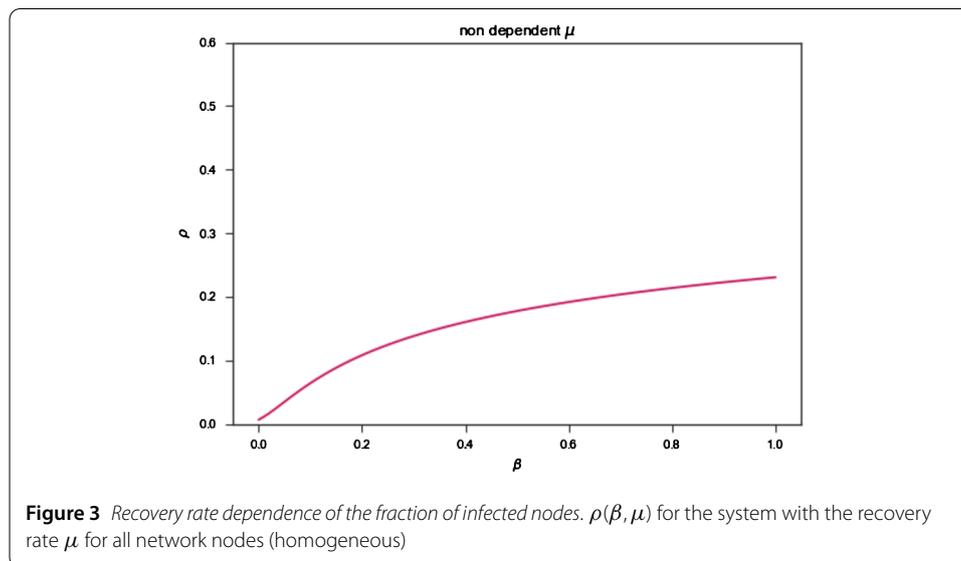
5 Experimental results

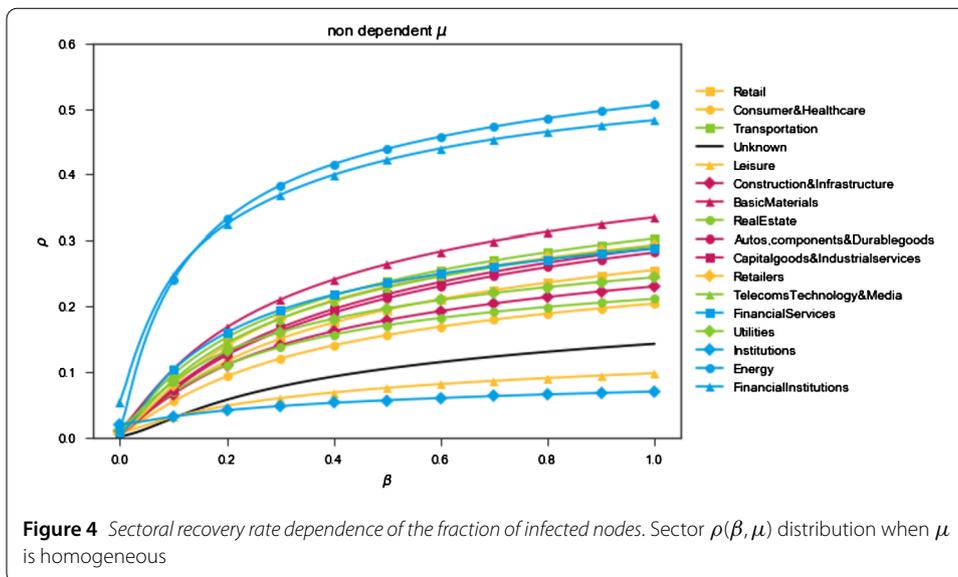
We will study next three different default propagation scenarios. The first one corresponds to the classical MMCA model where all nodes share the same recovery rate. In the second one we use the heterogeneous recovery rate measure introduced in Sect. 3.2. Besides, to increase our knowledge about the role of each sector in the default propagation process, we synthetically simulate default problems in each sector to analyze the different spreading speed in the transient state. Finally, we validate our findings comparing achieved results with a null model built by rewiring the edges of our network.

5.1 Default incidences for homogeneous recovery rate

We first explore the system response to the default dynamics using the baseline model. Figure 3 shows the default probability density at the steady state versus the default infection rate β for a default recovery rate of $\mu = 0.01$. For such a small recovery rate the system behaves almost like a Susceptible–Infected model, still there is a non-zero transition point from a low default rate to a system-wide default regime. As shown, for high infectious rates, the system default density is close to 0.20. This means that on average companies in the customer-supplier network, have a 0.20 probability of being in default for the set of parameters used.

Figure 4 shows $\rho(\beta, \mu)$ for each economical sector. Clearly, not all sectors behave in the same way regarding default dynamics. Broadly speaking, economical sectors can be grouped in three blocks given its response to default propagation. On one hand, (public) Institutions, Leisure and unknown sector reference show a low propensity to default propagation, where the sector default probability density range from approximately 0.04 to 0.10 for high infectious rate β . On the other hand, Financial Institutions and Energy evince a high propensity to default contagion, with ρ reaching almost 0.50. In other words, on average, each company of these two sectors has a 0.50 probability of being in default for





extreme parameter conditions. In-between, we find the rest of the sectors with density variations ranging from 0.15 to almost 0.30. Interestingly, the exposure of the economical sectors is quite different one to another. This result, may allow current risk assessment models (e.g. Generalized Linear Models, GLM) to include a quantification of sectoral risk and rank accordingly. In the following sections, we will assess the effect of the recovery rate variation and explore the reason behind the different sector response to default dynamics.

5.2 Impact of customer diversification on default incidence

In addition, we have compared the customer diversification variation model with a baseline model having constant default recovery rate μ . We simulate the latter with μ equal to the mean of μ_i for the whole network, specifically with $\mu = 0.005$. Consequently, the homogeneous μ (baseline model) used is 0.01. In our data, initial conditions for the number of companies in default at t_0 are more concentrated in Financial Institutions, 15% (relative to the sector) and the rest varies with a default rate between 6% and 1%. However, the MMCA modeling framework does not depend on the initial conditions when the steady state is reached.

Figure 5 depicts $\rho(\beta, \mu)$ distribution for each sector using a heterogeneous μ_i dependent of the companies customer diversification. Default density ρ increases when it is compared to the homogeneous μ scenario for all sectors. This means that the inclusion of recovery heterogeneity make sectors more prone to default. Nevertheless, this effect is likely related with the Spanish financial network structure, where medium and small companies with low diversification are predominant in the system (see Fig. 1). This fact produces that most nodes have small μ values (close to zero), whilst very few companies present μ values close to 1. However, there are some differences for a few sectors. We will explore the reason behind the observed dynamic behavior and its connection to the network structure in the following section.

5.2.1 Sector structure-function relationship

Table 2 summarizes metrics related to the network dynamics and topology for all sectors. The onset slope \mathcal{R}_0 and the rank according to the sensitivity value for the baseline

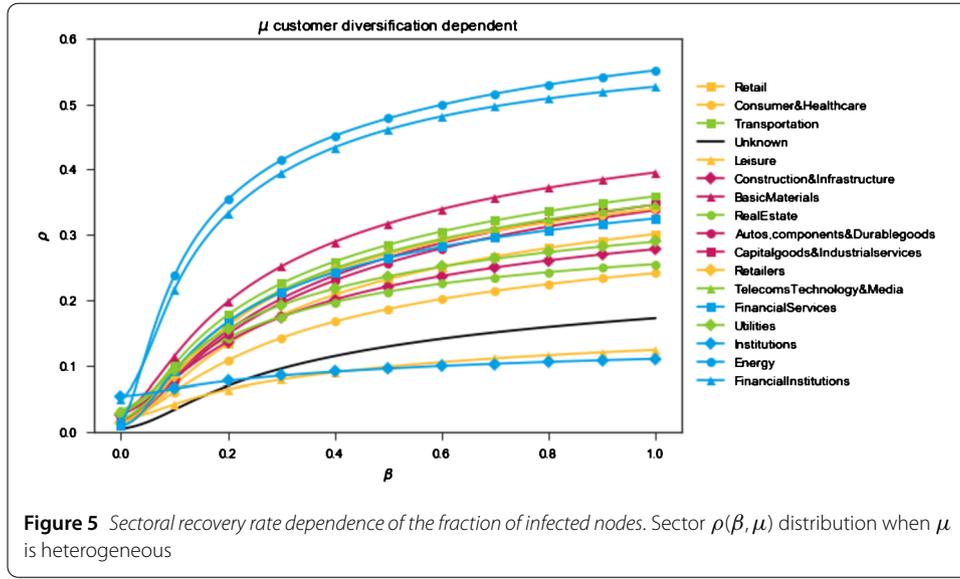


Table 2 Summary of network measures influencing dynamic default contagion by sector

Sector	γ	l_{in}	\mathcal{R}_0	S_{het}^{rank}	S_{hom}^{rank}
Energy	1.33	13.44	1.74	1	1
Financial institutions	1.27	2.39	3.81	2	2
Basic materials	1.58	41.07	2.89	3	5
Financial services	1.36	1.66	2.22	4	3
Transportation	1.44	16.27	2.33	5	4
Telecoms, technology & media	1.44	16.99	2.68	6	6
Retailers	1.54	28.18	2.70	7	8
Capital, goods & industrial services	1.47	9.51	2.93	8	9
Utilities	1.44	7.15	3.43	9	7
Autos, components & durable goods	1.52	21.24	2.82	10	10
Retail	1.58	29.80	2.73	11	11
Real estate	1.40	1.86	4.74	12	12
Construction & infrastructure	1.44	17.01	4.25	13	13
Consumer & health care	1.55	32.06	4.10	14	14
Unknown	1.59	5.16	3.69	15	15
Leisure	1.35	7.71	7.56	16	16
Institutions	1.54	5.65	-	17	17

model S_{hom}^{rank} and for the customer diversification variation S_{het}^{rank} are shown. Worth noting, we discarded the \mathcal{R}_0 value corresponding to Institutions (a sector including governmental institutions, religious organizations and others public institutions) because there were very few points, and as a consequence the regression slope is not statistically robust. In general, it is observed that there are slight variations in the sensitivity ranking but, as seen on Figs. 4 and 5, the customer diversification has not modified how most of the sectors are affected by the default dynamics. Also, there is quite a strong inverse relationship between the onset slope and the sensitivity, indicating that sectors which are susceptible to be affected are not those prone to propagate the default throughout the system. Main changes are observed in the basic materials and utility sectors. For example, default sensitivity of basic material companies which have a strong inter-sectoral inter-connectivity is reduced, showing that these companies have a very diverse customers, therefore making them more resistant to default spreading. Here it is important to highlight that the network only contains business to business relations, therefore sectors whose customers

are mostly individuals clients have lower probability to spread a default process. Besides, sectors having a larger number of authorities (nodes highly connected to hubs) exhibit a larger centrality making them more suitable to spread default problems.

We still lack an explanation of the observed behavior given the topological characteristics of each sector. To do so, we computed the slope coefficient of the inverse cumulative probability distribution of the companies in-strength for each sector (γ). The in-strength is defined as the sum of the incoming normalized weights for each company, meaning that the higher the value for a company the more probable that the default dynamics will affect it. As in most complex systems, this probability distribution is heavy-tailed signaling a Pareto like distribution where the slope coefficient can be computed. A smaller γ value signals that the sector is more probable to contain well connected companies (sector hubs). We observe that Energy and Financial Institutions are the most susceptible (S^{rank}) sectors, and coincide with lower values of γ (higher probability of sector hubs). The contrary happens to Institutions, and unknown sector reference. This highlights the fact that hub structures play an important role in the dynamics, and in the extent that a sector is affected by it. Clearly, leisure does not follow this explanation because it has a middle-range γ value. This could be due to the fact that it has 76% of its companies with zero in-strength. However, it is naive to think that only this structural property can explain the sector response to default dynamics. In Table 2, we can observe how less susceptible sectors are more equally inter-connected to other sectors (larger value of I_{in}). In practice, this causes the default spreading to be less likely to find a high probability path to these sectors because their incoming weights are less concentrated. As before, economical sectors can be grouped in three blocks given its capacity to interact and affect other sectors. On one hand, energy, financial institutions, financial services, transportation, telecommunications and retailers sectors are largely affected by the others due to their large sensitivity. Besides, these sectors are highly inter-connected with all the other sectors since their activity is traversal to all sectors and firms, therefore this high degree connectivity allows default to infect them easily. On the other hand, when consider leisure and unknown sectors, we observe that these are not affected by other sectors. For leisure sector, this is a consequence that most of its companies have zero in-strength. Similarly, companies belonging to the unknown sector are not BBVA clients, so their information is quite limited and most of their connections are not included in the network having, both, low in and out degree. In-between, there are sectors, such as (public) institutions and retail, that are stable independent of the perturbed sector. The main characteristic of these sectors is their low in-strength due to their customers not being companies or not having customers at all because they are public entities. In any case, default contagion does not reach these sectors and they kept healthy in all parameters setting.

Independently of the particular economic insights that may arise from these analysis, the methodology proposed in this paper has a greater advantage. It allows to perform experiments in a massive way. These large amount of data has enabled us to study how the time it takes the system to arrive to the steady state (convergence time) depends on a set of parameters such as the initial default rate or the infection rate β . This may have practical applications. For instance, if we can establish a relationship between simulations convergence time and real time, the risk departments could take advantage of this understanding to estimate the speed of default propagation among sectors.

5.3 Synthetic default assessment experiments

The methodology also allows for another kind of experiments; what we call synthetic default assessment. Previously, we have performed a sensitivity analysis using the real default distribution at the initial step. Now we are going to explore what happens when the initial default nodes are concentrated in a given sector and repeat the analysis for every sector. Previously, we have focused on the stationary state, where dynamic properties such as S and \mathcal{R}_0 were calculated without any dependence on the initial conditions. Now, we are interested in the transient phase, when the system has not reached the stationary state yet. In this transient phase, we can gain insights on the speed of default propagation among sectors.

To perform this experiment, we have randomly set different initial default values, specifically, 2%, 4%, 8% and 25% in each sector and no default in the remaining nodes (in the real data there is an initial default ratio equal to 1.703%, distributed across the sectors). Although, several experiments were performed, here we only include two of them, the ones described in Figs. 6 and 7. Simulations show that the moment when the system arrives to the stationary state depends on the initial default values, the infection rate β and the recovery rate μ . In particular, the larger the recovery rate μ , the faster the system arrives to the stationary state. Note that for initial default values of 25% (Fig. 7) there are no horizontal lines at iteration 20, indicating that the affected sector is still in the transient state, since it depends on the perturbed sector. We also observe that there are difference among sectors. Energy is the sector which first reaches the stationary for all values of the parameters.

We observe that most of the business sectors follow the tendency that the larger the number of sectors they affect, the fewer the number of sectors that at the same time are affecting them. An example can be found in Utilities; for a 25% perturbation all other sectors except one are affected, while Utilities itself is only affected by another sector (Unknown sector). Note that this is also the case for Transportation, and seems to be a constant throughout the other sectors. This is opposite to what happens to the sectors Energy and Financial Institutions, which is affected rapidly by all the other sectors. Knowing the dependence of the default speed contagion on the business sectors may allow risk assessment models to understand the conditions to react to a sudden perturbation of a sector, or to an event that may indicate the initial stages of a sector crisis.

5.4 Validation

To study if the topological properties of the customer-supplier network really affect default contagion, edges were rewired a million times in such a way that nodes kept their out-degree, in-degree and out-strength using the algorithm described in [45]. Doing this, 25 different networks with different topology were created. In Fig. 8, we present the differences of the trimmed mean of ρ as a function of β with regard to the original network by sector. Specifically, bloxplots were computed using the formula:

$$\Delta\rho_s = \frac{\left| \sum_{n=P_5}^{P_{95}} \frac{\rho_{r,n}}{N_{r,s}} - \sum_{n=P_5}^{P_{95}} \frac{\rho_{o,n}}{N_{o,s}} \right|}{\sum_{n=P_5}^{P_{95}} \frac{\rho_{o,n}}{N_{o,s}}}, \quad (11)$$

where P_i , ρ_r , ρ_o stand for the i th percentile, node ρ values of the rewired network and ρ values of the original network respectively. Besides, $N_{r,s}$ and $N_{o,s}$ are the number of nodes

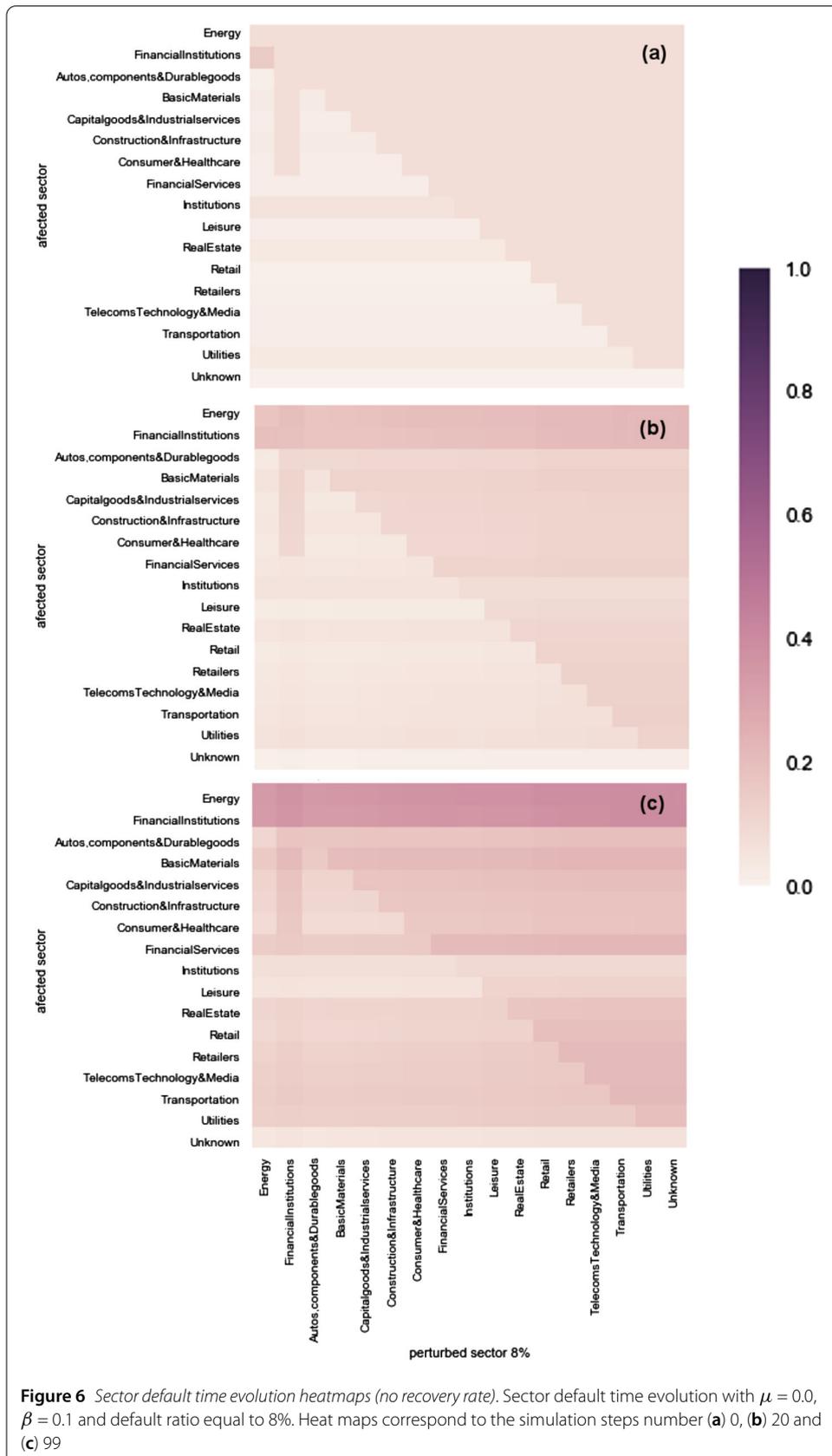
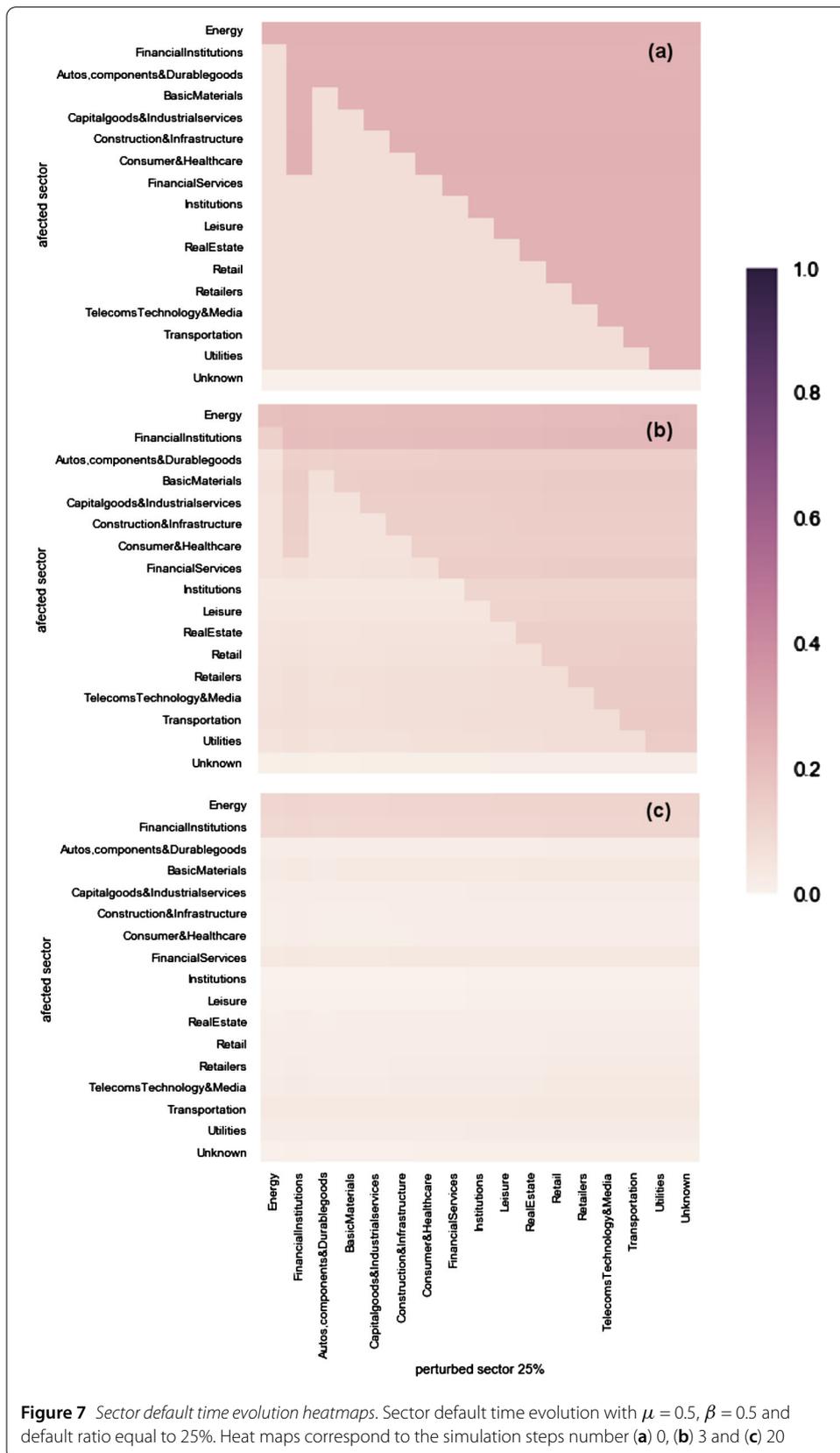
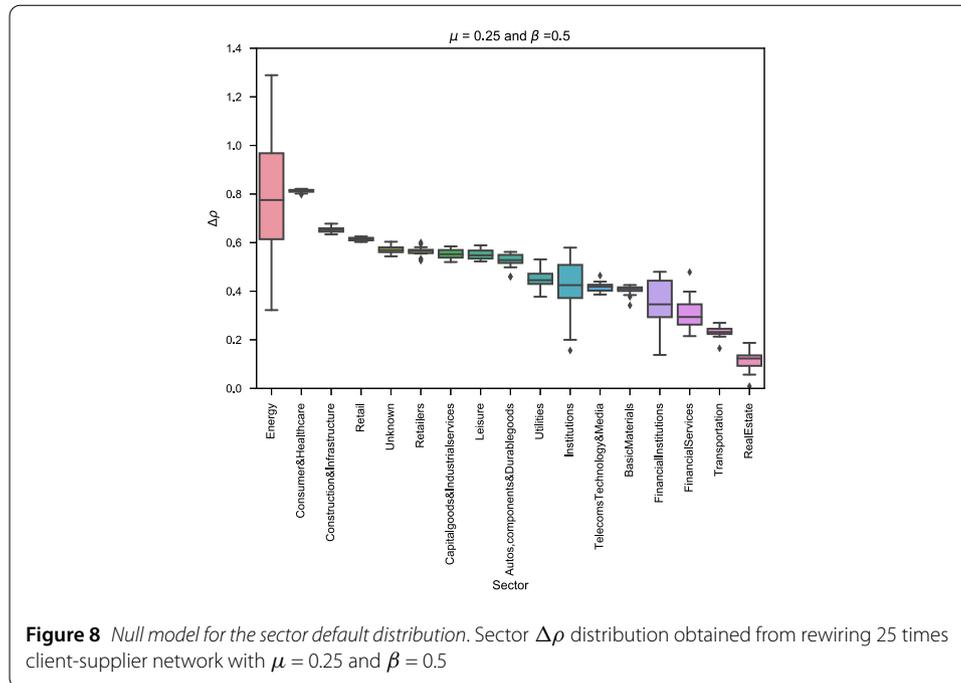


Figure 6 Sector default time evolution heatmaps (no recovery rate). Sector default time evolution with $\mu = 0.0$, $\beta = 0.1$ and default ratio equal to 8%. Heat maps correspond to the simulation steps number (a) 0, (b) 20 and (c) 99





within percentiles P_5 and P_{95} of the rewired and original network respectively. $\Delta\rho$ values of Fig. 8 show significant differences between customer-supplier network and the rewired one. Consequently, this confirms that the customer-supplier network structure plays an important role in all the aforementioned results.

6 Discussion

The methodology presented in this work applied to a real customer-supplier network of companies in Spain has allowed to gain new insights on financial data. We confirm quantitatively that business sector is a key component in default contagion, both in the strength and in the velocity of the contagion. Specifically we have found three differentiated blocks of behaviors depending on the business sectors. The ability to quantitatively estimate the size of this effect via topological analysis of the customer-supplier network demonstrated in this work may allow current risk assessment models to include them for future default prediction in a systematic way. Nowadays, this effect is only included from macro-financial data perspective. Therefore, there is a lack of relational information at micro-level. Customer-supplier network relations would cover this gap. We have proven that this dependence happens in at least two different scenarios: when recovery rate is homogeneous and when it is heterogeneous for every company. However, when the recovery rate depends on a topological property such as customer diversification, default contagion increases. This finding may be only a consequence of the structure of the customer-supplier network in Spain. Further studies could be carried out by using other country customer-supplier networks, and by using other dependence hypothesis for the recovery rate. This is where a major advantage of the methodology presented in this work comes: the ability to vary infection and recovery parameters at the micro-scale and study the effects on the dynamical properties of the network. For instance, we can base the model on more economically-wise hypotheses for the recovery rate (or infectious rate) that account for more realistic scenarios, such as the companies' revenue, and answer why some

sectors do not default although they should, provided the structure of the network. This allows risk experts to study many possible hypothesis. After carrying out a detailed analysis of how different parameters affects default contagion, we have seen that the number of hub companies a business sector includes is essential to estimate its sensitivity to default contagion. Although this is not the only variable to take into account, and results can not be generalized to other customer-supplier networks outside of Spain, we believe this is a first step towards the study of how the topological properties of the companies in the network affect default contagion. From a topological perspective, the results of our model applied to the real data reveal which sectors are more at risk in the propagation of default, which sectors are more resilient to the default avalanches, and what are the expectations for the cascades of default under different stochastic conditions. For that purpose, we have carried out two type of experiments:

- We have analyzed the dynamics of default contagion for different values of the recovery parameter μ , dependent and non-dependent on the node's features. Our methodology allows to tune parameters individually for every company and to carry out experiments for the simulation of future scenarios. In particular, we have studied the impact of the company's customer diversification on default propagation. A discussion on the connection between topological and dynamical properties is also included (Sect. 5.2.1);
- Our methodology also allows for another kind of experiment described in Sect. 5.3, where we have focused on the dynamics of default propagation at the transient state, and its dependence on the default initial conditions.

7 Conclusions

We have proposed a computational model, based on the probabilities of default contagion, to study the default diffusion at individual and aggregated levels. We have performed massive experiments based on this model by varying several parameters such as the initial default rate, the contagion rate β and the recovery rate μ . This methodology allows us to vary the parameters at the individual level to account for a more realistic scenarios. Our results show the relationship between dynamical and topological properties for more than 140,000 BBVA firms aggregated at a economic sector level, and also allow us to create a ranking of sectors by sensitivity to default, which can be used in potential applications. For future work, we would like to enrich the network adding different types of payments such as national transfers or direct debits, extending in this way our computational model to a multiplex network, finally we want to enrich model parameters considering for example companies' revenue.

Funding

JJR acknowledge partial funding from the Spanish Ministry of Science, Innovation and Universities, the National Agency for Research Funding AEI and FEDER (EU) under the grant PACSS (RTI2018-093732-B-C22) and the Maria de Maeztu program for Units of Excellence in R&D (MDM-2017-0711). AA acknowledges support by Spanish Ministry of Science, Innovation and Universities (grant PGC2018-094754-B-C21), Generalitat de Catalunya (grant 2017SGR896), Universitat Rovira i Virgili (grant 2017PFRURV-B2-41), ICREA Academia and the James S. McDonnell Foundation (grant 220020325). We acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abbreviations

BBVA, Banco Bilbao Vizcaya Argentaria, name of the second largest bank in Spain; GLM, Generalized Linear Models; MMCA, Microscopic Markov Chain Approach; SIS, Susceptible–Infected–Susceptible epidemic model; SME, Small and Medium-sized Enterprise; VAT, Value-Added Tax.

Availability of data and materials

The dataset is not publicly available. It was acquired by BBVA. Any of the authors based at BBVA may be contacted for further details about the dataset.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Conceptualization: JN, ET, PF, AA, JJR; methodology: AA; formal analysis: AB, JN, PF, ET; data curation: ET; writing (original draft preparation): PF, ET, JN; writing (review and editing): JN, PF, ET, AA, JJR; visualization: JN, AB, AM, ET; funding acquisition: JN. All authors read and approved the final manuscript.

Author details

¹BBVA Data & Analytics, Barcelona, Spain. ²Dept. Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Tarragona, Spain. ³ESADE, Universitat Ramon Llull, Barcelona, Spain. ⁴Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Palma de Mallorca, Spain.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 9 July 2018 Accepted: 17 October 2019 Published online: 04 November 2019

References

- Schweitzer F, Fagiolo G, Sornette D, Vega-Redondo F, Vespignani A, White DR (2009) Economic networks: the new challenges. *Science* 325:422–425. <https://doi.org/10.1126/science.1173644>
- Haldane AG, May RM (2011) Systemic risk in banking ecosystems. *Nature* 469:351–355. <https://doi.org/10.1038/nature09659>
- Battiston S, Puliga M, Kaushik R, Tasca P, Caldarelli G (2012) DebtRank: too central to fail? Financial networks, the FED and systemic risk. *Sci Rep* 2:541. <https://doi.org/10.1038/srep00541>
- Roukny T, Bersini H, Pirotte H, Caldarelli G, Battiston S (2013) Default cascades in complex networks: topology and systemic risk. *Sci Rep* 3:2759. <https://doi.org/10.1038/srep02759>
- Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A (2015) Epidemic processes in complex networks. *Rev Mod Phys* 87:925–979. <https://doi.org/10.1103/RevModPhys.87.925>
- Wu Y (2014) Reputation risk contagion and control of rural banks in China based on epidemic model. *Comput Model New Technol* 18:229–235
- Paltalidis N, Gounopoulos D, Kizys R, Koutelidakis Y (2015) Transmission channels of systemic risk and contagion in the European financial network. *J Bank Finance* 61:536–552. <https://doi.org/10.1016/j.jbankfin.2015.03.021>
- Kostylenko O, Rodrigues HS, Torres DFM (2019) The spread of a financial virus through Europe and beyond. *AIMS Math* 4:86. <https://doi.org/10.3934/Math.2019.1.86>
- Rogers EM (2010) *Diffusion of innovations*. Simon & Schuster, New York
- Jin F, Dougherty E, Saraf P, Cao Y, Ramakrishnan N (2013) Epidemiological modeling of news and rumors on Twitter. In: Proceedings of the 7th workshop on social network mining and analysis, p 8
- Khelil A, Becker C, Tian J, Rothermel K (2002) An epidemic model for information diffusion in MANETs. In: Proceedings of the 5th ACM international workshop on modeling analysis and simulation of wireless and mobile systems, Atlanta, GA, USA, pp 54–60
- Mizgier KJ, Wagner SM, Holyst JA (2012) Modeling defaults of companies in multi-stage supply chain networks. *Int J Prod Econ* 135:14–23
- Rizoiu M-A, Mishra S, Kong Q, Carman M, Xie L (2018) SIR-Hawkes: linking epidemic models and Hawkes processes to model diffusions in finite populations. In: Proc. of the World Wide Web conference (WWW), pp 419–428
- Glasserman P, Young HP (2016) Contagion in financial networks. *J Econ Lit* 54:779–831. <https://doi.org/10.1257/jel.2015.1228>
- Gómez S, Arenas A, Borge-Holthoefer J, Meloni S, Moreno Y (2010) Discrete-time Markov chain approach to contact-based disease spreading in complex networks. *Europhys Lett* 89:38009. <https://doi.org/10.1209/0295-5075/89/38009>
- Gómez S, Gómez-Gardeñes J, Moreno Y, Arenas A (2011) Nonperturbative heterogeneous mean-field approach to epidemic spreading in complex networks. *Phys Rev E* 84:036105. <https://doi.org/10.1103/PhysRevE.84.036105>
- Cai C-R, Wu Z-X, Guan J-Y (2014) Effective degree Markov-chain approach for discrete-time epidemic processes on uncorrelated networks. *Phys Rev E* 90:052803. <https://doi.org/10.1103/PhysRevE.90.052803>
- Callon M (1990) Techno-economic networks and irreversibility. *Sociol Rev* 38:132–161
- Serrano MA, Boguñá M (2003) Topology of the world trade web. *Phys Rev E* 68:015101. <https://doi.org/10.1103/PhysRevE.68.015101>
- Garlaschelli D, Battiston S, Castri M, Servedio VDP, Caldarelli G (2004) The scale-free topology of market investments. *Phys A, Stat Mech Appl* 350:491–499. <https://doi.org/10.1016/j.physa.2004.11.040>
- Jackson MO, Watts A (2002) The evolution of social and economic networks. *J Econ Theory* 106:265–295. <https://doi.org/10.1006/jeth.2001.2903>
- Coelho R, Neda Z, Ramasco JJ, Santos MA (2005) A family-network model for wealth distribution in societies. *Phys A, Stat Mech Appl* 353:515–528. <https://doi.org/10.1016/j.physa.2005.01.037>
- Gale DM, Kariv S (2007) Financial networks. *Am Econ Rev* 97:99–103
- Babus A (2016) The formation of financial networks. *Rand J Econ* 47:239–272. <https://doi.org/10.1111/1756-2171.12126>

25. Kosfeld M (2004) Economic networks in the laboratory: a survey. *Rev Network Econom* 3:1. <https://doi.org/10.2202/1446-9022.1041>
26. Onnela J-P, Kaski K, Kertész J (2004) Clustering and information in correlation based financial networks. *Eur Phys J B* 38:353–362. <https://doi.org/10.1140/epjb/e2004-00128-7>
27. Battiston S, Glattfelder JB, Garlaschelli D, Lillo F, Caldarelli G (2010) In: Estrada E, Fox M, Higham DJ, Oppo G-L (eds) *The structure of financial networks*. Springer, London, pp 131–163. https://doi.org/10.1007/978-1-84996-396-1_7
28. Nagurney A, Siokos S (2012) *Financial networks: statics and dynamics. Advances in spatial science*. Springer, Berlin
29. Caldarelli G, Battiston S, Garlaschelli D, Catanzaro M (2004) In: Ben-Naim E, Frauenfelder H, Toroczkai Z (eds) *Emergence of complexity in financial networks*. Springer, Berlin, pp 399–423. https://doi.org/10.1007/978-3-540-44485-5_18
30. Leitner Y (2005) Financial networks: contagion, commitment, and private sector bailouts. *J Finance* 60:2925–2953. <https://doi.org/10.1111/j.1540-6261.2005.00821.x>
31. Gai P, Kapadia S (2010) Contagion in financial networks. *R Soc Lond Proc, Ser A, Math Phys Eng Sci* 466:2401–2423. <https://doi.org/10.1098/rspa.2009.0410>
32. Battiston S, Caldarelli G (2013) Systemic risk in financial networks. *J Financ Manag Mark Inst*: 129–154. <https://doi.org/10.12831/75568>
33. Jackson MO, van den Nouweland A (2005) Strongly stable networks. *Games Econ Behav* 51:420–444. <https://doi.org/10.1016/j.geb.2004.08.004>
34. Lorenz J, Battiston S, Schweitzer F (2009) Systemic risk in a unifying framework for cascading processes on networks. *Eur Phys J B* 71:441. <https://doi.org/10.1140/epjb/e2009-00347-4>
35. Amini H, Cont R, Minca A (2013) Resilience to contagion in financial networks. *Math Finance* 26:329–365. <https://doi.org/10.1111/mafi.12051>
36. Glasserman P, Young HP (2015) How likely is contagion in financial networks? *J Bank Finance* 50:383–399. <https://doi.org/10.1016/j.jbankfin.2014.02.006>
37. Bardoscia M, Battiston S, Caccioli F, Caldarelli G (2017) Pathways towards instability in financial networks. *Nat Commun* 8:14416. <https://doi.org/10.1038/ncomms14416>
38. Thurner S, Poledna S (2013) Debrank-transparency: controlling systemic risk in financial networks. *Sci Rep* 3:1888. <https://doi.org/10.1038/srep01888>
39. Battiston S, Caldarelli G, Georg C-P, May RM, Stiglitz J (2013) Complex derivatives. *Nat Phys* 9:123–125. <https://doi.org/10.1038/nphys2575>
40. Battiston S, Caldarelli G, May RM, Roukny T, Stiglitz J (2016) The price of complexity in financial networks. *Proc Natl Acad Sci USA* 113:10031–10036. <https://doi.org/10.1073/pnas.1521573113>
41. Ait-Sahalia Y, Cacho-Diaz J, Laeven RJA (2015) Modeling financial contagion using mutually exciting jump processes. *J Financ Econ* 117:585–606. <https://doi.org/10.1016/j.jfineco.2015.03.002>
42. Feinstein Z (2017) Financial contagion and asset liquidation strategies. *Oper Res Lett* 45:109–114. <https://doi.org/10.1016/j.orl.2017.01.004>
43. European Central Bank (2005) *The financial risk management of the Eurosystem's monetary policy operations*. Eurosystem, Frankfurt
44. Kleinberg JM (1999) Authoritative sources in a hyperlinked environment. *J ACM* 46:604–632. <https://doi.org/10.1145/324133.324140>
45. Opsahl T, Colizza V, Panzarasa P, Ramasco JJ (2008) Prominence and control: the weighted rich-club effect. *Phys Rev Lett* 101:168702. <https://doi.org/10.1103/PhysRevLett.101.168702>

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
