# GEOMETRY OF DIFFUSION

## Structure of connectivity and systemic risk in financial networks

Master in Economics: Empirical Applications and Policies. Master Thesis.


### ABSTRACT

Ever since the 2008 global financial crises, both scholars and practitioners have argued that the interconnectivity of financial and banking networks is a key determinant of global systemic risk. Employing a simple diffusion model, we study the effect of short network cycles on the diffusion of idiosyncratic shocks on financial networks, an unexplored but important topic of research. We show the presence of trade-offs in the conditional effect of short cycles, and find preliminary theoretical evidence suggesting that cycles might at the same time stimulate and hinder diffusion. We subsequently undertake Monte Carlo analysis using core-periphery networks to analyze which direction prevails and under which conditions. We show that network cycles play a quantitatively important but complex role in the diffusion of financial shocks. In opposition to the notion of connectivity associated to short cycles, but in line with the trade-off, we detect non-monotonicities and significant interaction effects between triangles and squares that motivate further theoretical analysis.



***Keywords:*** complex networks, financial networks, systemic risk, diffusion, core-periphery, clustering, degree distribution, network cycles, short cycle distribution.



**Author:** Clavero Jover, Ion Cernin

**Director:** Kovářík, Jaromír





University of the Basque Country UPV/EHU

Bilbao, September 2021

# CONTENTS

# LIST OF ILLUSTRATIONS

## FIGURES



## TABLES

*"... **complexity and opaqueness** turned a housing price decline into a major financial crises, a development that few economists had anticipated. ... banks became reluctant to lend to each other for fear that the bank to which they lent might not be able to repay ..."*

Olivier Blanchard, *Macroeconomics*, 8th edition (2020)

*"The difficult task before market participants, policymakers, and regulators with systemic risk responsibilities such as the Federal Reserve is to find ways to preserve the benefits of **interconnectedness in financial markets** while managing the potentially harmful side effects."*

Janet Yellen (2013)

# 1 INTRODUCTION

Complexity and opaqueness of global financial markets, as the preceding quote by Olivier Blanchard points out, fed a spiraling economic meltdown in 2008. As the explosion of a housing prices bubble morphed into a banking crisis and, eventually, into a full-blown systemic crisis of secular dimensions, a strong call to disentangle the unknown risks arising from the opaque structure of interdependencies among financial intermediators arose. In the last years that call has evolved into an exciting field of study at the frontier of economic and finance: the analysis of economic and financial networks, which aims at measuring and understanding the systemic implications of network configurations.

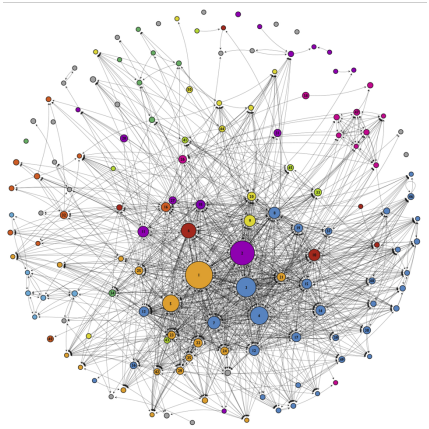


**Figure 1:** Top 150 global financial institutions by net worth in 2012 (Kopplin et al., 2012).

Financial institutions are exposed to each other due to a diverse form of interconnections, such as liabilities, cash lendings, over-the-counter transactions, or through correlated portfolia. However, such interconnectivity generates negative externalities. As a result, in the event of a idiosincratic shock, an agent in the network can fail, putting its dependent neighbors under stress, and potentially initiating a cascading sequence of defaults that compromise the whole system. Hence, the structure of the interdependencies, the *network*, is a source of financial and, by extension, global systemic risk.

The role of network structure on systemic risk has been pioneered by Allen and Gale (2000). Until 2008, though, it was not a central topic of research in economics. The prominence with which the *Financial Crisis Inquiry Report,* (United States Financial Crisis Inquiry Commission, 2011) featured the relationship of financial network structures with systemic risk, though, gives a measure of the weight that the topic has reached among regulators and governments in the aftermath of the global financial crisis. Eco-

nomic research in the field has, since, become dynamic and vibrant. This paper aims at contributing a small token to the broadening of our knowledge about networks.

# 2 MOTIVATION AND OBJECTIVES

The main goal of our research is to study the effect of short network cycles, namely triangles and squares, on shock diffusion in financial networks. While network connectivity and density as measures of systemic risk in financial systems have been abundantly analyzed in the literature (see Jackson and Pernoud (2021) and Cabrales, Gottardi, and Vega-Redondo (2017)) to the best of our knowledge no study on financial diffusion has established the role of cyclical structures (related to but not completely captured by degree distributions), although network cycles are ubiquitous in financial and banking networks (see Figures 1 and 2).

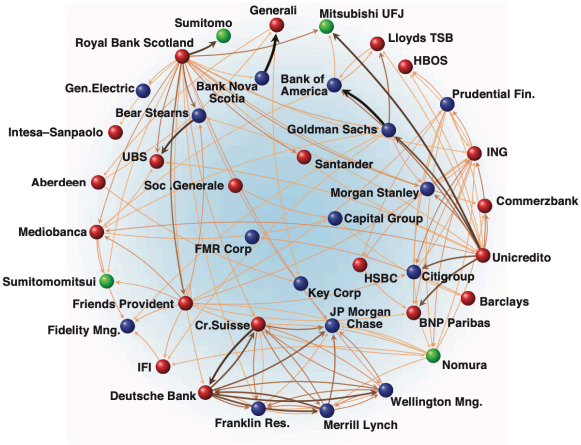



**Figure 2:** Global *core* financial institutions in 2009 (Schweitzer et al., 2009).

This paper introduces an initial approach to this sizable inquiry: we analyze a simple model of diffusion on networks, in which a bank is initially hit by an idiosyncratic shock and fails. As a result of this shock, other institutions that are connected to the initially hit bank fail with an exogeneous probability $p$. If they do fail, they further "infect" their partners, and so on. Although the model abstracts from certain specificities of financial diffusion, it allows us to provide a simple characterization of the diffusion process and the role of short network cycles[1].

In the following, we briefly summarize our findings. Theoretically, we show that short network cycles can exhibit both positive and negative impact on the diffusion of idiosyncratic financial shocks. To illustrate both directions, let us focus on the network

[1]The findings exposed in this work should be taken as preliminary analysis of a work in progress, with possible continuity in form of a PhD thesis.

in Figure 3: the distribution of connectivity and the link density (number of links) are the same in both networks in the figure. Therefore, the two networks allow us to make a comparison of the role of the triangle, conditional on the effect of connectivity and density traditionally considered in financial diffusion.

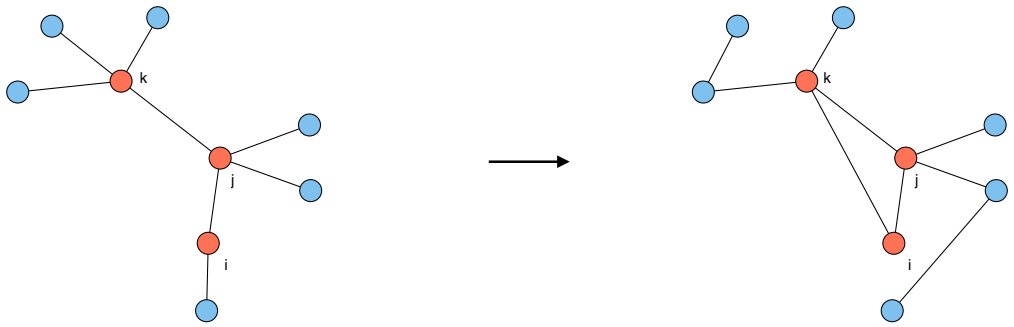



**Figure 3:** The addition of a triangle, a *ceteris paribus* illustration.

Lets imagine that node $i$ is hit by a shock and defaults. As a result of the shock, node $i$ infects node $j$ with probability $p$ in both networks. This is the only way in which $i$ can infect $j$ in the network on the left. In contrast, there is another infection path from $i$ to $j$ in the right network. If $i$ does not infect $j$, the perturbation can still flow from $i$ to $k$ and then from $k$ to $j$ due to the extra link the triangle provides in the right network. In this way, the triangle increases the likelihood of failure of node $j$. However, since the comparison holds the connectivity (degree and number of edges; definitions are provided on Section 3.1 on page 9) constant, after the previously described process, node $j$, once infected, cannot infect anybody else in the right, while it can still infect one additional node, $k$, in the left. Hence, in the *ceteris paribus* (keeping connectivity constant) comparison, the triangle decreases the number of connected agents. This effect contrasts the positive impact described above. In sum, the sign of the *net* effect of the triangle on diffusion is ambiguous.

Motivated by this trade-off, the second part of our analysis provides a Monte-Carlo analysis of the role of triangles on social networks, using core-periphery networks. As an identification strategy, we systematically manipulate the number of triangles and squares in the cores of the networks to be able to separate the effect of triangles (potentially squares) from that of connectivity and other relevant features. The simulation results confirm that the role of network cycles is economically important but complex. We first show that, in contrast to connectivity that increases the extent of financial contagion, the

(non-*ceteris paribus*) impact of triangles and squares is non-monotonic and interacts with the impact of other variables. In our regression analysis that exploits our identification strategy, we show that, *ceteris paribus*, triangles decrease the extent and number of periods of diffusion, while squares exert the contrary effect.

The implications of our findings can be far-reaching, particularly from a policy perspective, if they inform regulatory interventions regarding which network structure to promote or hinder and under which conditions. Furthermore, a robust knowledge of the consequences of specific network configurations can be instrumental in the design of economic mechanisms capable of internalizing the negative network externality in the agent's decision-making process.

There is abundant literature that deals with the determinants of diffusion: for example, Acemoglu et al. in their 2012 paper *The network origins of aggregate fluctuations* show the network determinants of shock diffusion. They argue that, in presence of intersectoral input–output linkages, microeconomic idiosyncratic shocks can lead to significant aggregate fluctuations. This crucial result contradicts the fundamental diversification argument, whereby, as Lucas (1977) and others argued, microeconomic shocks average out, and thus, only have negligible aggregate effects (Acemoglu et al., 2012). They do not, however, take into account the effect of short cycles in their analysis.

As for the role of short network cycles, Jackson and Pernoud (2021) discuss the role of short network cycles in the presence of multiple equilibria within financial systems, but they do not analyze their role in the diffusion of idiosyncratic shocks. Meanwhile, Espinosa, Kovářík and Ruiz-Palazuelos, in their 2021 paper *Are close-knit networks good for employment?* show that short social network cycles determine diffusion of job offers and, consequently, labor-market outcomes (Espinosa, Kovářík, & Ruiz-Palazuelos, 2021).

# 3 THEORY

In this section, we introduce the notation and define the key concepts for our analysis. Then, we introduce our core-periphery network. Lastly, we characterize the dynamics of diffusion in networks.

## 3.1 Definitions

The characterization of a network model requires fixing a number definitions. These provide the conceptual framework upon which we build and characterize both the core-periphery model and the diffusion dynamics.

**Definition 1: Network.** A network (or graph) $G = (N, E)$ is defined as a set of nodes $N = \{1, .., n\}$ and a set of edges $E$ linking pairs of nodes together.

**Definition 2: Induced subnetwork.** An induced subnetwork (or subgraph) $G[S_N]$ of $G$ is as a network formed by a subset $S_N \subset N$ of the nodes of $G$ and a subset $S_E \subset E$ of the edges of $G$ that have both endpoints in $S_N$. $G[S_N]$ is referred as the subnetwork induced in $G$ by $S_N$. We denote $ij \in E$ if $i$ and $j$ are connected in $G$. Alternatively, we will write $e_{ij} \in E$ in such a case.

**Definition 3: Network size.** The size of a network is its number of nodes, that is, the cardinality of the set $N = \{1, .., n\}$.

**Definition 4: Direct contact.** A direct contact, or first order contact, of a node $i \in N$ is any node $j \in N$ such that the edge $e_{ij}$ that links $i$ and $j$ is part of the set $E$ of edges of $G$, that is: $\forall j \in N : e_{ij} \in E$.

**Definition 5: First order degree.** The first order degree of node $i \in N$ is the cardinality of the set of nodes $D_i^1(G) = \{j \in N : e_{ij} \in E\}$, where $e_{ij}$ is the edge that links $i$ and $j$. That is, the first order degree of $i \in N$ is the number $d_i^1 = |D_i^1(G)|$ of nodes $j \in N$ that are direct contacts of $i$.

**Definition 6: Indirect contact.** An indirect contact, or second order contact, of a node $i \in N$ is any node $k \in N, k \neq i$ such that the edges $e_{ij}$ and $e_{jk}$ linking, respectively, node $i$ to node $j \in N$ and node $j$ to node $k$, are part of the set of edges of of $G$, that is: $\forall k \in N : k \neq i, e_{ij} \in E, e_{jk} \in E, j \in N$.

**Definition 7: Second order degree.** The second order degree of node $i \in N$ is the cardinality of the set of nodes $D_i^2(G) = \{k \in N : k \neq i, e_{ij} \in E, e_{jk} \in E, j \in N\}$, where $e_{ij}$ and $e_{jk}$ are the edges linking, respectively, node $i$ to node $j \in N$ and node $j$ to node $k \in N$. In other words, the second order degree of $i \in N$ is the number $d_i^2 = |D_i^2(G)|$ of nodes $k \in N$ that are indirect or second order contacts of $i$.

**Definition 8: Cycle.** A $K-$cycle in $G$ is a sequence of distinct nodes $i_1, i_2, \ldots, i_{K-1}, i_K$ such that $K > 2$, $e_{i_k i_{k+1}} \in E \ \forall k \in (1, \ldots, K-1)$, and $e_{i_1 i_K} \in E$. In other words, a $K-$cycle in $G$ is a sequence of $K$ linked nodes starting and ending in the same node.

**Definition 9: Triangle.** A triangle, or three-cycle, of a network $G$ is a cycle where $K = 3$, or a sequence of nodes $\{i, j, k\} \subset N$ such that $\{e_{ij}, e_{jk}, e_{ki}\} \subset E$.

**Definition 10: Square.** A square, or four-cycle, of a network $G$ is a cycle where $K = 4$, or a sequence of nodes $\{i, j, k, l\} \subset N$ such that $\{e_{ij}, e_{jk}, e_{kl}, e_{li}\} \subset E$.

**Definition 11: Clustering coefficient.** The clustering coefficient $C_i(G)$ of the node $i$ in network $G$ is the fraction of $i$'s first degree contacts that are contacts between themselves. Understanding it in cycle terms, it measures the number of triangles $i$ is involved in, over the number of all possible triangles linking $i$ with its contacts, were these to be connected among them. Formally:

$$C_i(G) = \frac{\sum_{j \neq i; k \neq j; k \neq i} g_{i,j} \cdot g_{i,k} \cdot g_{j,k}}{\sum_{j \neq i; k \neq j; k \neq i} g_{i,j} \cdot g_{i,k}},$$

where $g_{i,j} = 1$ if edge $e_{i,j} \in E$, and $g_{i,j} = 0$ otherwise (same for $g_{i,k}$ and $g_{j,k}$).

**Definition 12: Isomorphic networks.** Networks $G = (N, E)$ and $G' = (N', E')$ are isomorphic if there exists a bijection $f : N \to N'$ such that the edge $ij \in E$ if and only if the edge $f(i)f(j) \in E'$. That is, $f$ just relabels the nodes, and the networks are identical after said relabeling.

**Definition 13: Degree distribution.** The degree distribution of a network $G$, for a given order $k$, is a frequency distribution of the different degrees of nodes. That is, $P_k(d)$ is the fraction of nodes in $N$ that have $k-$order degree $d$ under a degree distribution $P$.

**Definition 14: Regular network.** A regular network of degree $d$, for a given order $k$, is one in which all nodes have the same $k-$order degree $d$, that is: $P_k(d) = 1$.

**Definition 15: Vertex-transitive network.** A vertex-transitive is a special case of

regular network, where every node is identical to every other in their first-order degree, second-order degree, and so on. That is, where $P_1(d_1) = 1, P_2(d_2) = 1, \ldots, P_k(d_k) = 1$.

## 3.2 Core-periphery networks

We propose a core-periphery model as an instrument to convey an important characteristic of financial networks: the distinction between central parts of the network, populated by big agents which are densely connected among them (the *core*) and a sparser, more diffuse outskirt, in which the size of the agents is smaller and the connections among its members scarce, with the majority of links directed towards the core (the *periphery*; see Figures 1 on page 4 and 2 on page 6). This duality synthetizes, hence, a clear morphological feature arising from the examination of financial networks, but, moreover, captures an important functional one: the flow of credit[2] follows a hierarchic order, much alike the network of arteries, vessels and capillaries in an organism distribute the blood from the heart to its last cell. Therefore, as a first approach to the specifics of network diffusion that we are focusing on (the effects of short cycles), it makes sense to initially abstract from the periphery, hierarchically subjected to the core, and studying first the dynamics among the systemic players, on which the rest of the network depends. Hence, after defining the core-periphery model here below, from then on *we will exclusively focus on the core, omitting the periphery.*

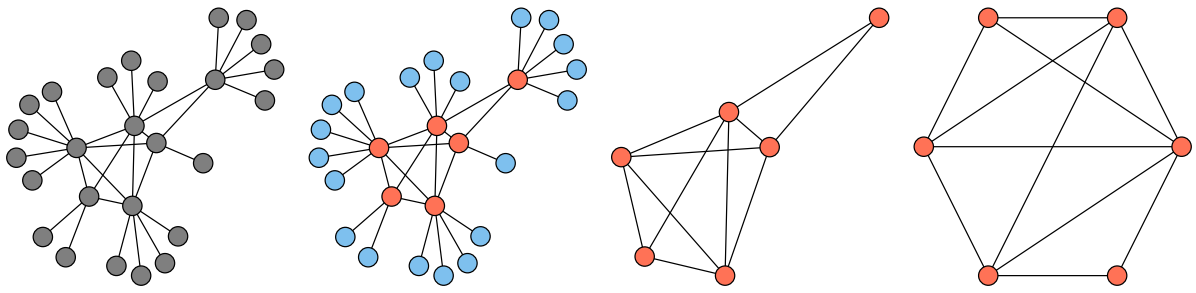


**Figure 4:** Core-periphery network: from left to right, the first figure features the whole network without distinction of core and periphery. On the second figure, the red nodes represent the core and the blue ones the periphery. The third picture shows the isolated core, as we will deal with from now on. Finally, the fourth picture shows the canonical representation of the core.

[2]Although a financial network does not necessarily represent creditor-debtor relationships, these follow closely mostly any other interrelation between financial institutions.

**Definition 16: Core-periphery network.** Let $G(N, E)$ be a core-periphery network (see Figure 4). The periphery $P$ is defined as a subset of the nodes in $G(N, E)$, $N_P \subset N$, such that their first degree order is 1, and as the subset of the edges in $G$, $E_P \subset E$, originating in $N_P$: $P(N_P, E_P)$. The core $C$ is hence defined as the sub-network composed by the subsets of nodes and edges in $G$ that do not intersect with those in $P$.

As previously argued, we abstract from the periphery in the current stage of our analysis. This simplification abides, as already mentioned, to the hierarchic structure of financial networks, but, given the definition of the model of diffusion that will follow, it makes also practical sense: the number of nodes infected in the periphery will be roughly proportional to the number of nodes in the core (see Section 3.3 on page 12). Once isolated from the periphery, the core is represented in its canonical form, as a ring conforming a polygon with as many sides as nodes, and with diagonals in its interior.

## 3.3 Dynamics of diffusion

We define diffusion as the process whereby a random, exogenous shock is capable of spreading, as a perturbation, from an infected node towards its contacts in the core. In our model, diffusion takes place from one node to its first order contacts, and unfolds dynamically according to the following assumptions:

- We define 2 categories of the status of nodes at every period: Infection status and susceptibility status. Nodes can be in the *infected* status in period $t$ ("new" infected), in which case they can transmit the perturbation to their contacts with probability $p \in (0, 1)$ in $t+1$; or they can be previously infected nodes (in $t-i$, $i \in \{1, 2, \ldots\}$; "old" infected) in which case, while still considered infected, they cannot transmit the perturbation.

  Regarding susceptibility, any node that has not been infected and is a direct neighbor of a node infected at $t$, is *susceptible* to infection at $t+1$. Any other node is *non-susceptible* to infection (including "old" infected nodes, non-infected nodes that are not direct neighbors of infected ones and, obviously, those currently infectious).

- At time $t = 1$, one of the nodes receives a shock, becoming *infected* and capable of transmitting the perturbation to its contacts in $t+1$. Every node in the core has the same probability of receiving the initial exogenous shock.

- The contacts of the shocked node become *susceptible to infection* in $t = 1$, and will become infected with probability $p$ in $t = 2$. The probability $p$ is, therefore, the basic probability assigned to the transmission of the perturbation among two linked nodes in one period.

- At $t = 2$, some of the previously susceptible nodes are infected, and capable of continuing transmission to their susceptible neighbors. The infectious node in $t = 1$, meanwhile, is no longer capable of transmitting the perturbation.

- The system goes on for $t > 2$ in this fashion, until no further contagion is possible.

Figure 5 offers a graphical example of the dynamics of diffusion in a core of size 6.

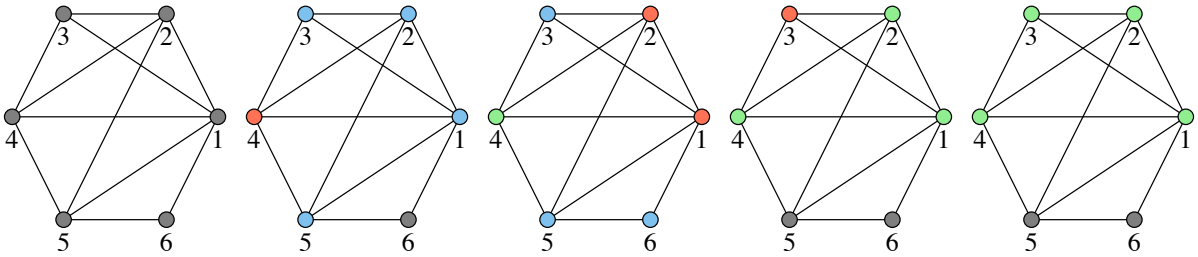



**Figure 5:** Example of diffusion dynamics in a core of size 6, from $t = 0$ on the left to $t = 4$ on the right. Red nodes are infected at time $t$, blue nodes are susceptible to being infected at time $t + 1$, gray nodes are not susceptible to being infected at time $t + 1$ and green nodes where infected in previous periods.

# 4 SIMULATION OF DIFFUSION MODEL

Once the core-periphery network model and the dynamics of diffusion are defined, using the package *igraph* in *R*, we construct a sample of cores in order to run simulations of the dynamic diffusion of a shock after randomly hitting one of the nodes.

## 4.1 Identification strategy

The cores, as defined in 3.2, are sub-networks of the core-periphery networks, with size $C$. Their basic configuration is a $C$ length cycle without shorter cycles (a ring; see Figure 7) which, in its canonical representation, is pictured as a regular polygon without diagonals.

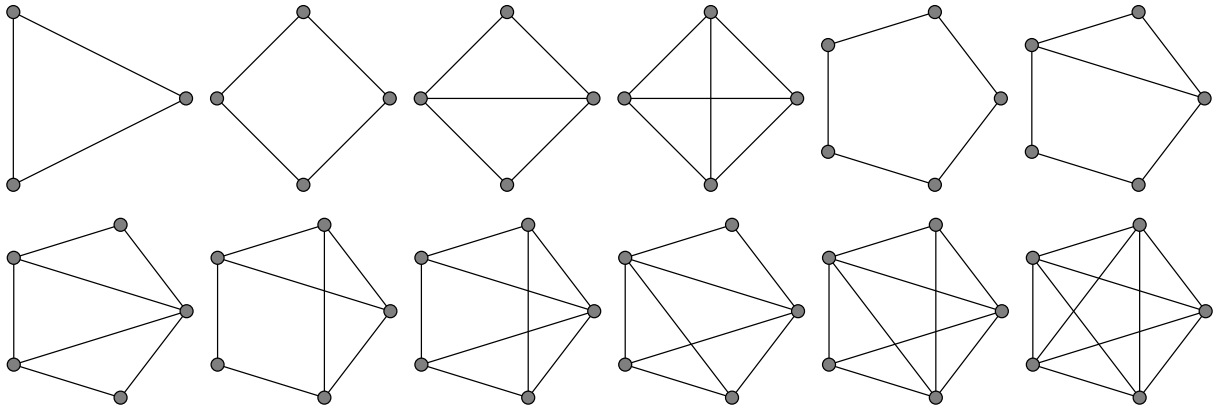
**Figure 6:** Basic configuration of cores as rings of 3, 4, 5 and 6 nodes.

The choice of rings as the most basic forms of cores lays in the fact that, for a given number of nodes, rings are the most simple (lowest number of edges) finite (finite number of nodes) vertex-transitive networks (see Definition 15) formed by a single component (a component is the biggest possible sub-network of a network such that all possible pairs of nodes in the sub-network are connected by a path of edges). They therefore provide an adequate foundation for the identification strategy. It should be noted that, as previously mentioned, there is no possibility of strict *ceteris paribus* when dealing with changes into one of the metrics characterizing a network. Rings, with the exceptions of those of size 3 and 4 (themselves a triangle and a square) provide a baseline in which, while all nodes are connected (a feature of what we understand as the *core* in a financial network), their clustering coefficient, number of triangles and number of squares, are all 0.

**Figure 7:** Identification strategy: comparison between two cores of size 6, with same number of diagonals, and just one difference in the ending node of one of the diagonals

From the basic rings, we proceed by adding diagonals linking nodes which are not direct contacts in the ring, until all distinct combinations of diagonals are exhausted when the number of diagonals reaches the value $D_C = C(C-3)/2$ (the maximum number of diagonals of a polygon; 0 in triangles, 2 in squares, 5 in pentagons, 9 in hexagons, etc). By distinct combinations of diagonals, we refer to those combinations of $0, 1, \ldots, D_C$ diagonals that result, for a given $C$ number of nodes, in a set of networks (cores) in which no pairs of cores are isomorphic (see Definition 12).

**(a)** Non-isomorphic cores of size 3, 4 and 5.



**(b)** Non-isomorphic cores of size 6.

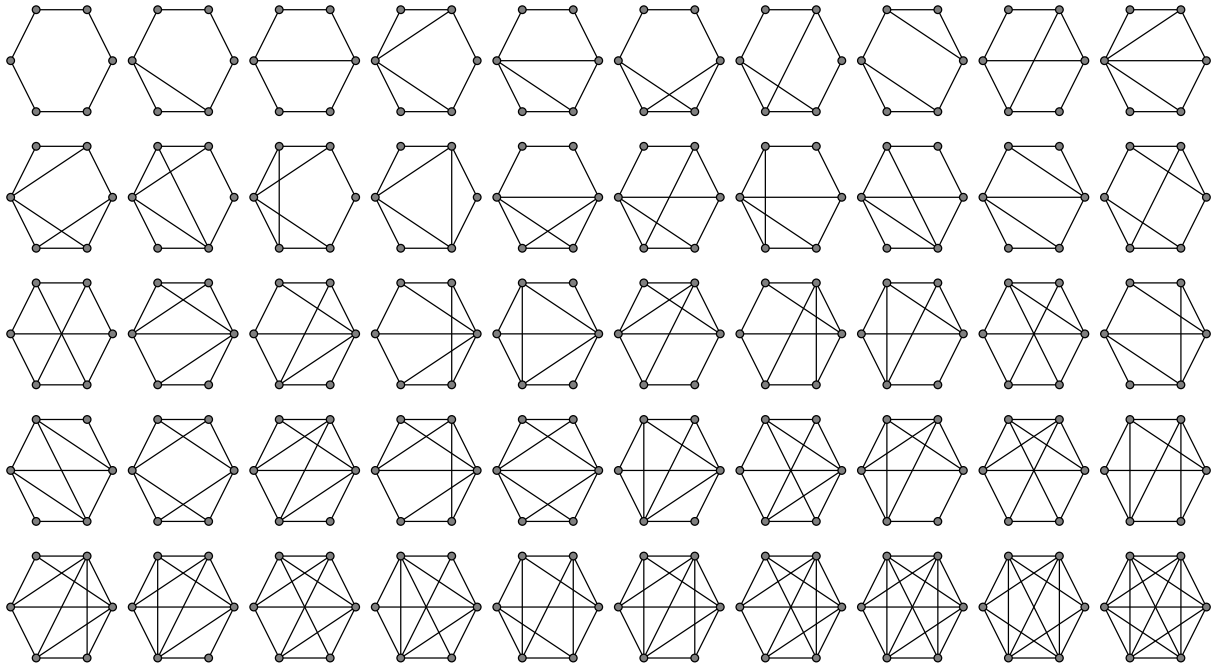


**Figure 8:** Full set of non-isomorphic cores, sizes 3, 4, 5 and 6.

The algorithm employed to identify *all* the distinct cores of size $N$, and *only* the distinct cores of size $N$ in the terms defined afore consists in, firstly, defining the rings of

size $N \in \{3, 4, 5, 6\}$, and, for each of the four rings, defining all the possible cores with all possible combinations of diagonals for $d \in \{1, 2, \cdots D_N\}$ total diagonals (for example: for cores of size 4, there would be 1 core with no diagonals, 2 cores with 1 diagonal, and 1 core with 2 diagonals). Secondly, comparing all the cores of the same size by pairs, to establish whether they are isomorphic or not, tending up with a set of non-isomorphic cores (following with the previous example, for the cores of 4 nodes and 1 diaghonal, only 1 of the 2 would be kept, as they are isomorphic). The *igraph* package in *R* is capable of establishing whether two graphs are isomorphic[3], but the comparison is always between the two elements of a set of cardinality 2. Hence, the approach we have followed, which becomes computationally intractable for cores beyond size 6, is, for the set of all possible cores of size $N$, evaluating all possible pairs for isomorphisms. Each time a pair of cores is isomorphic, one of them is dismissed. This evaluation of cores restarts for the remaining cores, until the evaluation of all possible pairs of cores outputs no isomorphisms.

Once the elimination of isomorphisms is finished, the set of cores ends up being formed by 1 core of size $N = 3$, 3 cores of size $N = 4$, 8 cores of size $N = 5$ and 50 cores of size $N = 6$ (see Figure 8).

## 4.2 Monte Carlo simulations

The simulation of the process described in Section 3.3 is implemented in the 62 cores described in Section 4.1. For each of the 62 cores, we simulate the process 500 times for each value of $p$. The vector of probabilities chosen is $p = (0.05, 0.10, 0.15, \ldots, 0.90, 0.95)$ (19 different basic transmission probability values). Hence, in total we run $62 \times 19 \times 500 = 589,000$ simulations.

[3] *igraph* implements the BLISS algorithm by Junttila and Kaski for undirected graphs. See *igraph* documentation and Junttila and Kaski (2007).

# 5 RESULTS

We focus on three variables in our empirical analysis. First of all, the *infection rate:* the number of nodes in the core that have ended up infected by the time the system stabilizes. Given the variability of core sizes, analyze the fraction of infected nodes over the core size. Furthermore, to avoid the biasing effect of counting the initial shocked node (more intense the smaller the core is), we compute the measure over the initially non-infected nodes, that is:

$$\text{Infection rate} = \frac{\#\,\text{Infected nodes}}{C-1}, \tag{1}$$

where $C$ is the size of the core.

The second outcome of interest is the number of periods that the system requires to stabilize: *periods to stabilization.* If, for example, after the shock hits in $t = 1$, no nodes become infected in $t = 2$, the system has stabilized in $t = 1$, and the number of periods required to stabilization is 0. Therefore, the number of periods to stabilization is equal to the value of the first period with identical state to the period before, minus 2.

Finally, we propose a variable measuring *intensity.* We compute it as the ratio of the infection rate over the periods to stabilization. It represents a (tentative) measure of the severity of the diffusion process, as it combines the *scope* (spatial dimension of the infection) with the *speed* (time dimension of the process). This approach can be useful in two ways: on the one hand, it allows us to grasp, at a glance, a combined measure of infection and speed. On the other hand, it can provide some preliminary evidence, thinking in terms of policy recommendation, of whether the number of the nodes infected or the speed at which they become infected is more relevant when confronting an event of diffusion in a financial network.

## 5.1 Descriptive analysis

In this section, we summarize the results of our analysis using simple descriptive statistics. A first look at the outcome of the simulations (see Table 1) shows that, for bigger cores, the average rate of infection is higher, with an overall infection rate of 61.5% of nodes (as defined in Equation 1). Regarding the number of periods it takes for the system to stabilize, according to intuition, it increases with increasing size, with an average of 3.48 periods overall.

| **Infection rate** | | | | **Periods to stabilization** | | | |
|---|---|---|---|---|---|---|---|
| Core size | Mean | Std. dev. | Freq. | Core size | Mean | Std. dev. | Freq. |
| 3 | 0.5843 | 0.4442 | 9,500 | 3 | 2.7001 | 0.8709 | 9,500 |
| 4 | 0.5922 | 0.4311 | 28,500 | 4 | 3.0599 | 1.0972 | 28,500 |
| 5 | 0.6068 | 0.4245 | 76,000 | 5 | 3.3140 | 1.2224 | 76,000 |
| 6 | 0.6182 | 0.4203 | 475,000 | 6 | 3.5445 | 1.3370 | 475,000 |
| Total | 0.6149 | 0.4219 | 589,000 | Total | 3.4777 | 1.3153 | 589,000 |

**Table 1:** Descriptive statistics of simulation results by core size. Infection rate (as percentage of nodes infected upon stabilization, measured over $C - 1$ nodes) and periods to stabilization.

Summarizing the results by the value of the basic probability of transmission ($p$) reveals, as expected, a monotonically increasing average number of infected nodes for increasing value of $p$ (see Figure 9 and Table 4 in Appendix A on page 29). This straightforward preliminary result of infection rates when conditioning only on $p$ contrasts with the outcome regarding the number of periods it takes the system to stabilize: there is clear preliminary evidence of non-monotonicity on the average number of periods. For increasing values of $p$, the system requires an increasing number of periods to stabilize up to a value around $p = 0.70$, when the number of periods decreases for increasing values of $p$. This is an important preliminary confirmation of the existence of trade-offs involving the value of the basic probability of transmission $p$.

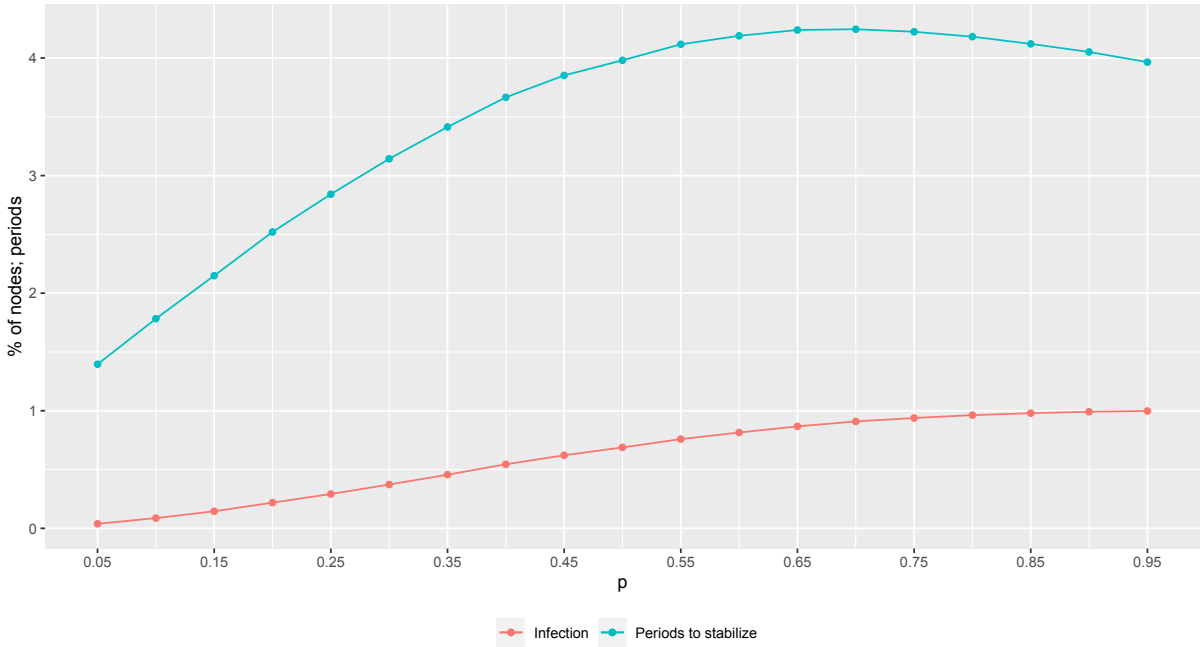



**Figure 9:** Average infection rate and periods to stabilization by $p$ (CI 95% too small to appreciate).

One way to approach the evaluation of the outcome variables (infection, periods to stabilization and intensity) controlling for short cycles (triangles and squares) and core size is through the inspection of Figures 10 (page 19), 11 (page 20) and 12 (21). They represent the average simulated outcomes only for size 6 cores (the most abundant ones) ordered in two different ways: the top figures represent the outcome variable for the cores ordered by average first order degree (and second order, in case of ties). The bottom figures, on the other hand, represent the same outcome variable, for the same cores, but ordered differently: starting for 0 squares, the cores are ordered by increasing number of triangles; then, for 1 square, again by increasing number of triangles, and so on. In this way, we end up having a number of "segments" representing the outcome variable for a given number of squares, ordered for increasing number of triangles. Furthermore, each outcome is represented for 3 different values of $p$: 0.25, 0.50 and 0.75.


**Figure 10:** Average infected nodes upon stabilization on 6 node cores, CI 95%. Ordered by ascending degree (top) and by increasing number of squares and triangles (bottom).

Figure 10 shows the absolute number of infected nodes when the diffusion stops, including the exogenously shocked node. When looking at the degree (top) the results are clear and straightforward: infection increases (roughly) monotonically for the average

degree of the core. This is an expected result, following the well known role of *unconditional* (of short cycle distribution) degree in diffusion in netwoks. When looking at short cycles (bottom) a prominent feature is apparent: an extra square seems to reduce the infection. However, this apparent result should not be taken as preliminary evidence: we do not know the value of the number of triangles before and after the jumps caused by the extra square. In any case, what seems more solid, is that extra triangles appear to increase infection unconditionally on degree.

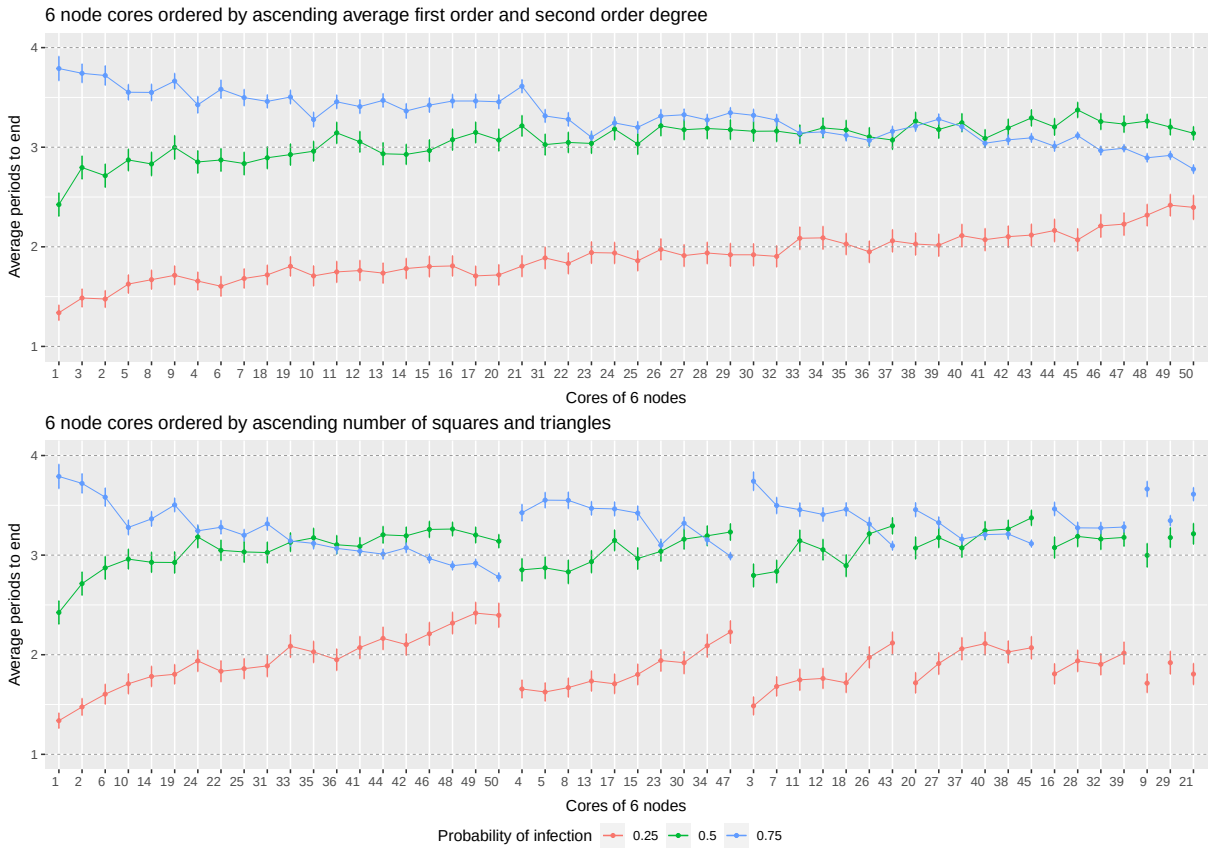



**Figure 11:** Average periods to stabilization on 6 node cores, CI 95%. Ordered by ascending degree (top) and by increasing number of squares and triangles (bottom).

Figure 11 presents the data for the average periods to stabilization, organized following the aforementioned procedure for infection in Figure 10. In this case, the preliminary results are very interesting: for $p = 0.25$ and $p = 0.50$, the time required for the system to stabilize (the number of periods) behaves as intuitively expected: higher degree, and higher number of triangles, increases the number of periods to stabilization. For higher values of probability, though (in this case, for $p = 0.75$) and in accordance with what Figure 9 on page 18 shows, increasing both the degree and the number of triangles (conditional on squares) reduces the infectivity: there is, again, *preliminary evidence of a*

*trade-off arising from the structure of the network, which accelerates or slows down the process given high or low values of p.* It should be noted, furthermore, that this is not an artifact originated in the fact that there are more or less nodes becoming infected (to less infected nodes, faster stabilization is expected): Figure 10 on page 19 tells that infection rises monotonically for every probability for both increasing degree and cycle distribution: hence, for high degrees and/or high density of triangles (which always appear to increase infection) low $p$ values slow down the process, and high $p$ values increase it (*less* periods).

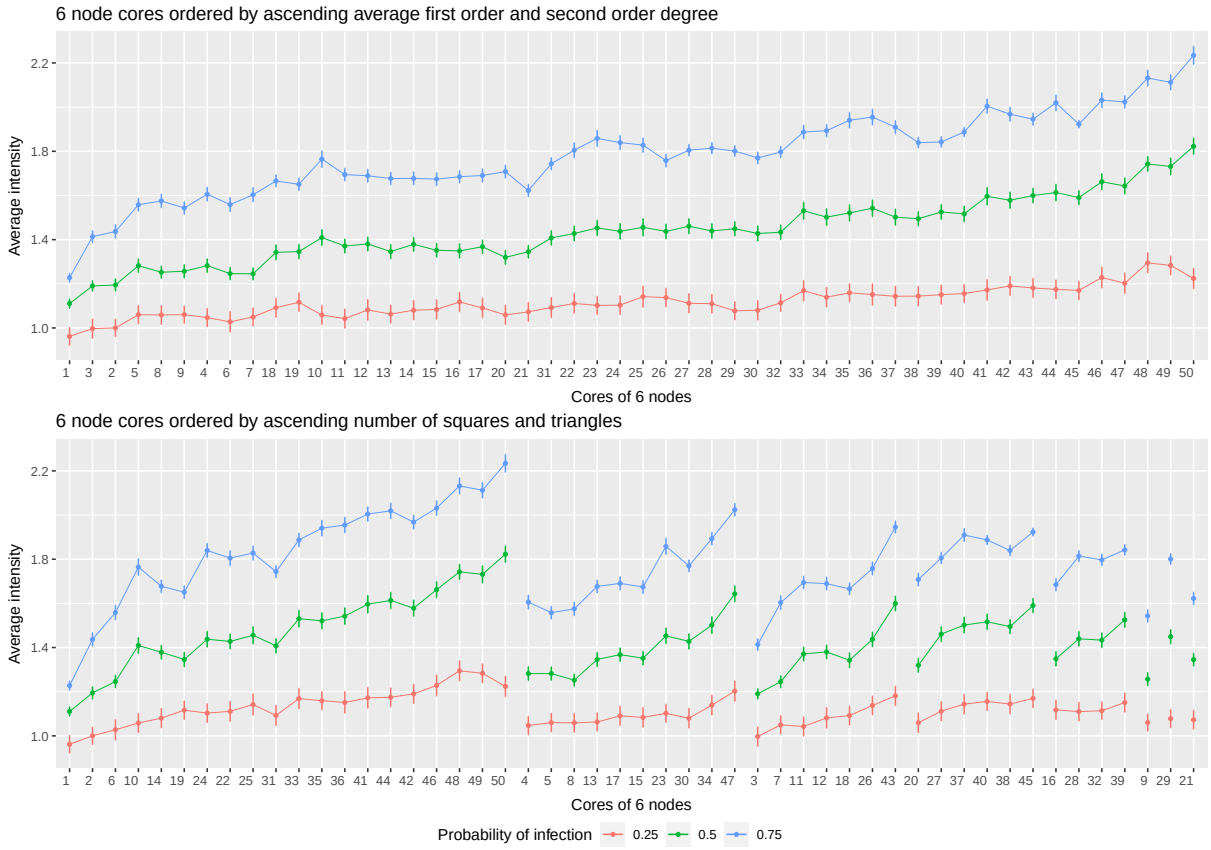



**Figure 12:** Average intensity on 6 node cores, CI 95%. Ordered by ascending degree (top) and by increasing number of squares and triangles (bottom).

Finally, Figure 12 plots the proposed measure of intensity in the same way as figures 10 and 11. In this case, we find again a (very roughly) monotonic increasing intensity for both degree and triangles. Two features should be noted, though: the first one is that, as this variable is the ratio of the number of infected nodes over the number of periods to stabilization, it represents both measures at the same time; however, as in this case we find no visual evidence of the trade-offs observed in the number of periods (Figure 11) we can preliminarily conclude that the magnitude of the effects on the number of infected cores is relatively more important than the effects on the periods to stabilization. The

second feature to note is that, for both degree and triangle ordering, the slope appears higher, the higher the $p$; intensity appears, then, to grow in a more pronounced way the higher the $p$, both when increasing the average degree or the number of triangles.

## 5.2 Estimations

The visual evidence provided by the figures in Section 5.1 has a fundamental limitation: it does not control for both degree and short cycles at the same time, and, hence, we cannot conclude anything firm from it, regardless of the interesting intuitions it provides on trade-offs. We therefore run regressions on the infection rate and periods to stabilization to adequately control for the different network characteristics when measuring the effects of short cycles.

In all the regressions of both outcomes we control for $p$ and $p^2$, to account for the non-linear effects of the basic probability of contagion. We also control for core size and for first and second order (average) degree of the core: specification (1). Specification (2) adds the number of triangles and squares as controls. Finally, specification (3) also controls for the interaction between squares and triangles.

Table 2 on page 23 shows the estimation results for the rate of infection. We run the three specifications with two econometric tools. On the one hand, we run a regular linear regression model. However, as the outcome is fractional (the percentage of nodes in the core that end up infected, excluding the originally shocked one) we also implement a fractional outcome model through a logit regression, which avoids obtaining predicted values out of the domain of definition of the outcome variable. We report the average marginal effects on the means. The estimations obtained with the two procedures are coherent, indicating the robustness of the results. We find strong statistical evidence of a *negative* average estimated effect of the addition of an extra triangle on the rate of infection, *ceteris paribus*. This result suggests that the prevalent effect of the presence of triangles is the *reduction* of infection, perhaps by the means described in Section 2. We also find evidence of a *positive* effect of squares, but of a much smaller magnitude than the effect of triangles. Finally, the interaction between squares and triangles is also significant, but its expected effect is of an even smaller magnitude than that of squares.

**Table 2:** Full-sample estimations of percentage of infected nodes

| | **Linear models** | | | **Fractional logit models**[2] | | |
|---|---|---|---|---|---|---|
| **Infected nodes**[1] **(%)** | (1) | (2) | (3) | (1) | (2) | (3) |
| $p$ | 2.056 | 2.056 | 2.056 | 0.993 | 0.990 | 0.990 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| $p^2$ | -0.882 | -0.882 | -0.882 | -0.015 | -0.012 | -0.012 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.052) | (0.051) |
| Core size | 0.0203 | 0.0364 | 0.0358 | 0.0213 | 0.0348 | 0 .0343 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| 1st order d. (avg. %) | 0.850 | 0.952 | 0.944 | 0.833 | 0.919 | 0.912 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| 2nd order d. (avg. %) | 0.306 | 0.231 | 0.229 | 0.285 | 0.218 | 0.216 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Triangles | | -0.00455 | -0.00444 | | -0.00391 | -0.00362 |
| | | (0.000) | (0.000) | | (0.000) | (0.000) |
| Squares | | 0.000933 | 0.000645 | | 0.001163 | 0.00161 |
| | | (0.001) | (0.030) | | (0.000) | (0.000) |
| Triangles × squares | | | 0.000155 | | | - |
| | | | (0.015) | | | - |
| $N$ | 589,000 | 589,000 | 589,000 | 589,000 | 589,000 | 589,000 |
| $R^2$ | 0.622 | 0.623 | 0.623 | - | - | - |

[1] As defined in Equation 1 (page 17)

[2] Average marginal effects computed on the means are reported

$p$-values in parentheses

Table 3 on page 25 shows the estimation results for the periods to stabilization. In this case the three specifications are estimated with a linear model and with a Poisson regression model, to account for the fact that the outcome is a counting variable. We report the incidence-rate ratios for the Poisson regressions (the exponentiated estimated coefficients). They are interpreted as the expected ratio between the baseline number of periods to stabilize and the new number of periods to stabilize upon the increase of the control in one unit. Again, the output of both procedures is highly significant and coherent: triangles are found to have a *negative* (increasing speed) and significant effect in the linear model, and in the Poisson regression, increasing the number of triangles in one unit *reduces* the expected number of periods to stabilization in less than 0.01%. Squares, on the other hand, show a significant and *positive* effect: an extra triangle slows down the process of stabilization. Finally, the effect of the interaction between triangle and square, while significant, is an order of magnitude smaller than the effect of the squares.

**Table 3:** Full-sample estimations of periods to stabilization

| | **Linear models** | | | **Poisson reg. models**[1] | | |
|---|---|---|---|---|---|---|
| **Periods to stabilize** | (1) | (2) | (3) | (1) | (2) | (3) |
| $p$ | 9.390 | 9.390 | 9.390 | 28.690 | 28.690 | 28.690 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| $p^2$ | -6.578 | -6.578 | -6.578 | 0.090 | 0.090 | 0.090 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Core size | 0.249 | 0.272 | 0.267 | 1.078 | 1.086 | 1.085 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| 1st order d. (avg. %) | 0.211 | 0.286 | 0.227 | 1.076 | 1.107 | 1.089 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| 2nd order d. (avg. %) | 0.151 | -0.0884 | -0.0985 | 1.0600 | 0.9863 | 0.9840 |
| | (0.005) | (0.110) | (0.075) | (0.000) | (0.391) | (0.316) |
| Triangles | | -0.00684 | -0.00597 | | 0.9977 | 0.9980 |
| | | (0.000) | (0.000) | | (0.000) | (0.000) |
| Squares | | 0.0124 | 0.0102 | | 1.0032 | 1.0027 |
| | | (0.000) | (0.000) | | (0.000) | (0.000) |
| Triangles × Squares | | | 0.00117 | | | 1.00029 |
| | | | (0.000) | | | (0.000) |
| $N$ | 589,000 | 589,000 | 589,000 | 589,000 | 589,000 | 589,000 |
| $R^2$ | 0.469 | 0.469 | 0.469 | - | - | - |

[1] Incidence-rate ratios are reported

$p$-values in parentheses

# 6 CONCLUSIONS

- The simulation results confirm that the role of network cycles is economically important but complex. We show that, in contrast to connectivity, that increases the extent of financial contagion, the impact of triangles and squares is non-monotonic and interacts with the impact of other variables.

- In our regression analysis, that exploits our identification strategy, we show that short cycle distribution, conditional on degree distributions and core sizes, is highly significant on both outcomes: proportion of infected nodes and speed of stabilization.

- Triangle distribution, conditional on degree distribution, shows a negative effect on infection. This might be, as we hypothesize, due to the fact that, for an extra triangle, *ceteris paribus*, the effect of node isolation (against infection) dominates over the conectivity effect (increased infection).

- Triangle distribution has an estimated conditional positive effect on speed (less periods to stabilization) in our regression analysis. However, we also find preliminary evidence of trade-offs between the value of the probability of node-to-node contagion $p$ and the number of triangles: triangles seem to have contradictory effects on the speed of stabilization, depending on the easiness with which shocks jump between linked nodes. This suggest non-monotonicity: higher number of triangles at high probabilities of contagion accelerate stabilization, while slowing it at low probabilities.

- Preliminary analysis of the intensity variable suggests that the effect of short cycles on fast and far-reaching episodes of diffusion might be relevant, with important implications for policy recommendation, as intensity combines the two dimensions most relevant to intervention: speed and scope.

# 7 DISCUSSION

From the initial stage in an ambitious project of research that this paper represents, we are aware of the areas in which we should be focusing our coming efforts:

- Our model is a simple model of diffusion in complex networks. It provides valuable insights on the effect of cycles conditional on degree distribution but it requires incorporating specific features of financial systems à la Jackson & Pernoud (2021).
- We need to incorporate the study of real-life networks: particularly those built on empirical data of financial (Jackson & Pernoud, 2021), banking and sectorial networks (Acemoglu et al., 2012) on which simulations can be run.
- We need to broaden the theoretical modeling of bigger core sizes and their metrics. To that end, we need to develop, or adopt, a more powerful and efficient algorithm.
- We have to undertake the theoretical characterization of the two opposing effects and the net effect of cycles on diffusion on core-periphery networks.

# REFERENCES

Allen, F., & Gale, D. (2000). Financial contagion. *Journal of Political Economy*, *108*(1), 1–33. Retrieved from http://www.jstor.org/stable/10.1086/262109

Cabrales, A., Gottardi, P., & Vega-Redondo, F. (2017). Risk-sharing and contagion in networks. *Review of Financial Studies*, *30*. Retrieved from https://www.researchgate.net/publication/268206043_Risk-sharing_and_contagion_in_networks

Espinosa, M. P., Kovářík, J., & Ruiz-Palazuelos, S. (2021). Are close-knit networks good for employment? [Working Papers]. (21.06). Retrieved from https://ideas.repec.org/p/pab/wpaper/21.06.html

Jackson, M. O., & Pernoud, A. (2021). Systemic risk in financial networks: A survey. *Annual Review of Economics*, *13*(1), 171-202. Retrieved from https://doi.org/10.1146/annurev-economics-083120-111540

Junttila, T., & Kaski, P. (2007). Engineering an efficient canonical labeling tool for large and sparse graphs. *Proceedings of the ninth Workshop on Algorithm Engineering and Experiments and the fourth Workshop on Analytic Algorithmics and Combinatorics*, 135-149. Retrieved from https://researchportal.helsinki.fi/en/publications/engineering-an-efficient-canonical-labeling-tool-for-large-and-sp (Society for Industrial and Applied Mathematics cop.; Workshop on Algorithm Engineering and Experiments.)

Kopplin, I., Hahn, & Zimmerman. (2012, 5). Wem gehört die welt? *Die Zeit*, *23*. Retrieved from https://www.zeit.de/zustimmung?url=https%3A%2F%2Fwww.zeit.de%2F2012%2F23%2FIG-Capitalist-Network

Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A., & White, D. (2009). Economic networks: The new challenges. *Science (New York, N.Y.)*, *325*, 422-5.

United States Financial Crisis Inquiry Commission. (2011). *The financial crisis inquiry report.* The Library of Congress. Retrieved from https://lccn.loc.gov/2011381760

# A FURTHER DESCRIPTIVE STATISTICS

**Table 4:** Descriptive statistics of simulation results by basic probability of transmission, $p$. Infection rate (as percentage of nodes infected upon stabilization, measured over $C-1$ nodes) and periods to stabilization.

| **Infection rate** | | | | **Periods to stabilization** | | | |
|---|---|---|---|---|---|---|---|
| Probability | Mean | Std. dev. | Freq. | Probability | Mean | Std. dev. | Freq. |
| 0.05 | 0.0387 | 0.1014 | 31,000 | 0.05 | 1.3955 | 0.6028 | 31,000 |
| 0.10 | 0.0871 | 0.1629 | 31,000 | 0.10 | 1.7825 | 0.8266 | 31,000 |
| 0.15 | 0.1451 | 0.2193 | 31,000 | 0.15 | 2.1491 | 0.9743 | 31,000 |
| 0.20 | 0.2184 | 0.2749 | 31,000 | 0.20 | 2.5206 | 1.1118 | 31,000 |
| 0.25 | 0.2921 | 0.3158 | 31,000 | 0.25 | 2.8419 | 1.1874 | 31,000 |
| 0.30 | 0.3723 | 0.3480 | 31,000 | 0.30 | 3.1427 | 1.2324 | 31,000 |
| 0.35 | 0.4561 | 0.3681 | 31,000 | 0.35 | 3.4134 | 1.2442 | 31,000 |
| 0.40 | 0.5445 | 0.3769 | 31,000 | 0.40 | 3.6658 | 1.2352 | 31,000 |
| 0.45 | 0.6213 | 0.3735 | 31,000 | 0.45 | 3.8523 | 1.1980 | 31,000 |
| 0.50 | 0.6883 | 0.3603 | 31,000 | 0.50 | 3.9807 | 1.1390 | 31,000 |
| 0.55 | 0.7588 | 0.3348 | 31,000 | 0.55 | 4.1167 | 1.0703 | 31,000 |
| 0.60 | 0.8150 | 0.3062 | 31,000 | 0.60 | 4.1889 | 1.0077 | 31,000 |
| 0.65 | 0.8671 | 0.2683 | 31,000 | 0.65 | 4.2384 | 0.9237 | 31,000 |
| 0.70 | 0.9091 | 0.2276 | 31,000 | 0.70 | 4.2451 | 0.8485 | 31,000 |
| 0.75 | 0.9382 | 0.1912 | 31,000 | 0.75 | 4.2234 | 0.7856 | 31,000 |
| 0.80 | 0.9627 | 0.1511 | 31,000 | 0.80 | 4.1815 | 0.7154 | 31,000 |
| 0.85 | 0.9795 | 0.1129 | 31,000 | 0.85 | 4.1203 | 0.6596 | 31,000 |
| 0.90 | 0.9915 | 0.0731 | 31,000 | 0.90 | 4.0516 | 0.5999 | 31,000 |
| 0.95 | 0.9980 | 0.0348 | 31,000 | 0.95 | 3.9652 | 0.5428 | 31,000 |
| Total | 0.6149 | 0.4219 | 589,000 | Total | 3.4777 | 1.3153 | 589,000 |