Systemic risk assessment
through high order clustering coefficient

Roy Cerqueti[♭], Gian Paolo Clemente[♮], Rosanna Grassi[♯*]

[♭] University of Macerata, Department of Economics and Law.

Via Crescimbeni 20, 62100, Macerata, Italy.

Tel.: +39 0733 2583246; fax: +39 0733 2583205. Email: roy.cerqueti@unimc.it

[♮] Università Cattolica del Sacro Cuore, Milano, Department of Mathematics, Finance and Econometrics

Email: gianpaolo.clemente@unicatt.it

[♯] University of Milano - Bicocca, Department of Statistics and Quantitative Methods

Email: rosanna.grassi@unimib.it

Abstract

In this article we propose a novel measure of systemic risk in the context of financial networks. To this aim, we provide a definition of systemic risk which is based on the structure, developed at different levels, of clustered neighbours around the nodes of the network. The proposed measure incorporates the generalized concept of clustering coefficient of order l of a node i introduced in [10]. Its properties are also explored in terms of systemic risk assessment.

Empirical experiments on the time-varying global banking network show the effectiveness of the presented systemic risk measure and provide insights on how systemic risk has changed over the last years, also in the light of the recent financial crisis and the subsequent more stringent regulation for globally systemically important banks.

1   Introduction

The recent financial distress and its spread over the world economic realities have pointed the attention of practitioners and academics to the conceptualization and the management of systemic

---

*Corresponding author.

risk. Indeed, even if the starting point of the crisis is well localized in time and space – 2008 in the US, with the failure of Lehman Brothers – the effects of this negative event have been and are still pervasive quite everywhere.

The concept of systemic risk can be defined in a number of different ways, also on the basis of the identification of the context under investigation (see e.g. [14], [22], [30]). Under a very general point of view, systemic risk is the possibility that a negative occurrence at a local level might generate a collapse at a global level. The premise for the introduction of systemic risk is the definition of a system, which is nothing but a unified structure composed by distinct interconnected entities.

One of the most intuitive ways for modelling a systemic risk framework is through complex networks (for a recent survey about networks and systemic risk, we can refer to [9] or [28]). In fact, a network is a system composed by units – the so-called nodes – along with their interconnections – arcs or edges. Thus, systemic risk is the possibility that an exogenous shock in one of the nodes triggers the collapse of the entire network.

Therefore, it is not properly unexpected that several studies deal with systemic risk problems in the framework of complex networks (see e.g. [5], [13], [15], [23], [24] [32], [35]).

The basis of systemic risk lies in the way in which shocks propagate among the nodes of the network. Such a propagation is clearly strongly dependent on the position and the density of the edges, i.e. on the topological structure of the graph associated to the network. Indeed, as intuition suggests, the presence of a large number of interconnections leads to a more probable diffusion of the local shocks, hence yielding a high level of systemic risk. In this respect, it is worth mentioning [4], [6] and [25].

In the context of the relationship between interconnectedness and systemic risk, a relevant role is played by the concept of community. With the term community we refer to a set of nodes whose mutual links are of particular strength ([17], [19]). Thus, the assessment of the community structure of a network let understand how powerful the mutual interconnections among the nodes are, hence providing useful insights on the systemic risk. This argument suggests that the measurement of the entity of the communities strength might represent a crucial step for the exploration of systemic risk. In this respect, the clustering coefficient of a network is of peculiar relevance.

The clustering coefficient of a given node is a relative measure of the triangles including the considered node as a vertex with respect to the hypothetical ones. Triangles are the easiest geometric visualizations of the communities, since they offer the image of a non-exclusive interaction among different agents. Such a measure has been developed in all the cases of weighted, unweighted, directed and undirected networks (see e.g. [1], [11], [16], [29], [33] and [34]). The extension of this community measure to the overall network is obtained by simply taking the average of all the clustering coefficients of the nodes.

In line with [7], [27] and [31], this paper deals with the systemic risk assessment through the

2

analysis of the clustering coefficient of a network.We move from a way to formalize a concept of community more general and informative than the standard clustering coefficient. In fact, communities might be appreciated not only at the level of adjacent nodes, but also by considering nodes which are at the periphery of the network. In so doing, we are adding a stratification of the communities at different levels, in order to capture the presence of community effects at different distances from the considered nodes.

To this aim, we employ the high order clustering coefficient, recently introduced by [10]. Furthermore, we adapt it to the systemic risk context in order to assess how each node is embedded in the entire system. In this way, we are providing specific indices that captures how a network is clustered at different levels. Additionally, a global systemic risk index is defined with the purpose of catching both the community structure around the node and the level of mutual interconnections of nodes that are at a specific geodesic distance. In other words, it indicates the presence of high (or low) clustered areas, revealing parts of the network where the risk diffusion could spread easily. Furthermore, being this index defined as a weighted average of high order clustering coefficients measured at different levels, we are able to modulate the effects of both adjacent and peripheral nodes through the weights distribution. Indeed, we can take into major consideration either the interactions of a node with its neighbours or the community structures generated by nodes at a wide distance from the considered one.

Our theoretical proposal is validated over the paradigmatic example of Global Interbank Network, which is particularly suitable for our purpose. Indeed, the network structure of national interbank markets has been studied, at the global level, using the Bank of International Settlements (BIS) data set. Systemic risk is mainly related to the interbank context and the considered empirical data allow to provide meaningful insights also to this type of literature (see, for instance, [7], [18], [21], [20], [26], [27]).

We study the time-varying behaviour of the community structure of the global banking network over the sample period that goes from the first quarter of 2005 to the end of 2017. Data naturally induce a core-periphery network and then we focus on the behaviour of core countries, i.e. countries whose banking systems report data to the BIS. In particular, to disentangle the role of core countries that host global systemically important banks (GSIBs), as defined in the list provided by the Financial Stability Board since November 2001, we will separately analyse the behaviour of countries in which at least a GSIB is present. In other words, we start by measuring the community structures at different levels on the basis of full global banking network's topology over each quarter of the year. Then, we compute high order clustering coefficients at a global level by focusing on two different subsets: the set of countries in which at least a GSIB is present and the set containing other core countries. In this way, we provide two alternative systemic risk indices that allow to assess the state of risk and to describe the pattern over time for both groups.

3

We observe that, in the entire period, countries, where a GSIB is present, show a clustering coefficient higher than other core countries, confirming the important systemic role of the banks headquartered in these countries. Our analysis confirms evidence of a reduction of global banking connectedness as an effect of the cutback in cross-border lending, triggered by the subprime crisis and the subsequent sovereign debt crisis in the Euro area. Furthermore, results suggest a different pattern since 2011 between the two clusters. On one hand, other core countries show a tendency to diversify relationships as the average number of transactions increases and the average volume remains quite stable. On the other hand, lending countries, where a GSIB is present, are cutting their exposures in terms of numbers and volumes. We interpret it as an effect of the "systemic risk score" introduced by the Basel Committee on Banking Supervision at the end of 2011 and the effectiveness of GSIBs regulation in inducing these banks to contain their systemic nature.

Furthermore, by providing a separate analysis between coefficients that considers only either in-flows or out-flows, we describe the different behaviour of countries in terms of risk-driver or risk-taker. It is worth mentioning how other core countries are more affected as risk taker. As regard to the "out"coefficient we observe a greater value for countries where a GSIB is present in the period 2009-2011. It is also noticeable the reduction of GSIB systemic impact over the last years.

To sum up, the novelties of the present study are the following: first, we construct a new systemic risk measure by using the concept of stratified communities. In so doing, we include in the systemic risk assessment also core-periphery effects in the analysis of shocks propagation; second, we allow the tuning of the systemic risk measure on specific levels of the communities, in order to give credit to the action of peculiar parts of the considered network in assessing systemic risk. In this, we are basically introducing a family of systemic risk measures with a wide set of meanings and interpretations; third, we provide a deep analysis of the Global Interbanking System and infer aspects of the related systemic risk which seem to remain unexplored in the standard frameworks.

The rest of the paper is structured as follows. Preliminaries and notations are reported in Section 2. We summarize general notation in Subsection 2.1, while in Subsection 2.2 it is described the definition of the high order clustering coefficient introduced in [10], which is the basis of the systemic risk measure defined below. In Section 3, we provide a new indicator of systemic risk based on an extended version of the clustering coefficient. Some remarks on the systemic risk measure are in Subsection 3.1. By means of a small example, Section 3.2 stresses the potential of our proposal in respect to the classical weighted clustering coefficient defined in the literature. In Section 4, we provide a deep analysis of the interbank system. Conclusions follow. An Appendix contains the definition of the clustering coefficients in the directed case, for the convenience of the reader.

4

2 Preliminaries and notation

We here present the mathematical definitions supporting the structure of the paper, for the convenience of the reader.

2.1 General notation

We denote by $G = (V, E)$ a graph, being V the set of N vertices and E the set of m arcs (or edges), which are unordered pairs of vertices. Vertices i and j are said to be adjacent when $(i, j) \in E$. The degree d_i of i is the number of the edges incident upon i. A path connecting vertices i and j is a sequence of distinct vertices and edges between i and j. If there exists a path between i and j, then i and j are said to be connected. The graph G is connected if all pairs of vertices of G are connected.

The distance $d(i, j)$ is the length of any shortest path connecting i and j. Such a shortest path is said to be a geodesic between i and j. All the geodesics between i and j have, of course, the same length $d(i, j) = l$. We define the set $\mathcal{G}_{ij}(l)$ as the one collecting all the geodesics connecting the vertices i and j; the generic element of $\mathcal{G}_{ij}(l)$ is $g(l) = g_{ij}(l)$, and it represents the generic geodesic connecting i and j whose length is l. Clearly, if such a geodesic exists, it is not unique in general. Thus, when $\mathcal{G}_{ij}(l)$ is not empty, its cardinality is not necessarily equals to one, for each $i, j \in V$ and length $l > 0$.

By conventional agreement, we assume that $d(i, j) = \infty$ when i and j are not connected. The diameter of G, denoted by $diam(G)$, is an integer given by the length of any longest path of G, and can be properly defined once G is a connected graph.

For a connected graph, we define the set:

$$N_i(l) = \{j \in V | d(i, j) = l\},$$

with $l = 1, \ldots, diam(G)$, and use the notation $|N_i(l)| = d_i(l)$ to represent its cardinality.

If any edge $(i, j) \in E$ is associated to a positive real number w_{ij}, then both the edges and the graph are weighted. Once we set that $w_{ij} = 0$ if and only if $(i, j) \notin E$, then we can describe completely the edges of the graph through the real N-square matrix $\mathbf{W}$ with entries w_{ij}, which is the weighted adjacency matrix. In particular, if $w_{ij} = 1$ for all edges $(i, j) \in E$, then $\mathbf{W}$ is simply the adjacency matrix $\mathbf{A}$ and the graph is unweighted. We will collapse this case into the more general weighted one.

The strength of vertex i is the sum of the weights of the arcs incident upon i. We denote it by s_i. Clearly, $s_i = d_i$ in the unweighted case.

A weighted network is a graph with its weighted adjacency matrix.

The weight of a geodesic $g(l)$ between i and j is given by the sum of the weights of its edges, and will be denoted by $w_{ij}(l, g)$ hereafter. From this concept, we introduce the l-th order strength

of the node i as

$$s_i(l) = \sum_{j \in N_i(l)} w_{ij}(l),$$

with $w_{ij}(l) = \min\limits_{g(l) \in \mathcal{G}_{ij}(l)} \{w_{ij}(l, g)\}$.

When a direction is assigned to the edges of a graph G, then we obtain a directed graph $D = (V, E)$, and G represents the underlying graph of D. The directed edges of D are said arcs.

A directed path from i to j is a path whose arcs have the same direction, which is the one going out from i and going in j. The existence of a directed path from i to j implies that j is reachable from i. Such a directed path is said to be an out-path of i. In an intuitive way, one can say that the geodesic distance $\overrightarrow{d}(i, j)$ from i to j is the length of a geodesic out-path (or out-geodesic) connecting i and j, and it is set to $\overrightarrow{d}(i, j) = \infty$ when such an out-geodesic does not exist.

By reverting the argument above, one has that the out-path of i can be defined as an in-path of j, and we denote by $\overleftarrow{d}(i, j)$ the length of any geodesic in-path (or in-geodesic), with the usual agreement that $\overleftarrow{d}(i, j) = \infty$ when such an in-path does not exist.

The directed graph D is said to be strongly connected when all the pairs of two vertices are mutually reachable. D is said to be weakly connected if the underlying graph G is connected.

By replacing d with $\overrightarrow{d}$ and $\overleftarrow{d}$, one can rewrite the definitions of $N_i(l)$, $d_i(l)$, $\mathcal{G}_{ij}(l)$, $g(l) = g_{ij}(l)$, w_{ij}, $w_{ij}(l, g)$, $s_i(l)$, $w_{ij}(l)$ by $\overrightarrow{N}_i(l)$, $\overrightarrow{d}_i(l)$, $\overrightarrow{\mathcal{G}}_{ij}(l)$, $\overrightarrow{g}(l) = \overrightarrow{g}_{ij}(l)$, $\overrightarrow{w}_{ij}$, $\overrightarrow{w}_{ij}(l, g)$, $\overrightarrow{s}_i(l)$, $\overrightarrow{w}_{ij}(l)$ and $\overleftarrow{N}_i(l)$, $\overleftarrow{d}_i(l)$, $\overleftarrow{\mathcal{G}}_{ij}(l)$, $\overleftarrow{g}(l) = \overleftarrow{g}_{ij}(l)$, $\overleftarrow{w}_{ij}$, $\overleftarrow{w}_{ij}(l, g)$, $\overleftarrow{s}_i(l)$, $\overleftarrow{w}_{ij}(l)$, respectively.

2.2 The high order clustering coefficient

We now report the definition of the high order clustering coefficient introduced in [10], which is the basis of the systemic risk measure defined below.

In case of a weighted, undirected and connected graph G, we initially define a matrix $\mathbf{P}(l) = [p_{ij}(l)]_{i,j \in V}$ for $l = 1, \ldots, diam(G)$, whose entries are

$$p_{ij}(l) = \begin{cases} \frac{w_{ij}(l)}{s_i(l)} & \text{if } j \in N_i(l) \text{ and } N_i(l) \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

If $l = 0$, we define $\mathbf{P}(l) = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

The local clustering coefficient of order l is $\mathbf{c}(l) = [c_i(l)]_{i \in V}$, obtained as

$$\mathbf{c}(l) = \mathbf{P}(l)\mathbf{c} \tag{2}$$

where $\mathbf{c} = [c_i]_{i \in V}$ is the vector whose element c_i is the weighted local clustering coefficient of the node i (see [1]).

If the graph is directed, weighted and weakly connected, matrix $\mathbf{P}(l)$ in (1) becomes $\bar{\mathbf{P}}(l)$ with

entries:

$$\bar{p}_{ij}(l) = \begin{cases} \frac{\bar{w}_{ij}(l)}{\bar{s}_i(l)} & \text{if } j \in \bar{N}_i(l) \text{ and } \bar{N}_i(l) \neq \emptyset, \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

where:

(a) $\bar{N}_i(l) = \overrightarrow{N}_i(l)$, $\bar{w}_{i,j}(l) = \overrightarrow{w}_{ij}(l)$ and $\bar{s}_i(l) = \overrightarrow{s}_i(l)$ in case only out-paths of node i are taken into account, and $\bar{\mathbf{P}}(l) = \overrightarrow{\mathbf{P}}(l)$;

(b) $\bar{N}_i(l) = \overleftarrow{N}_i(l)$, $\bar{w}_{i,j}(l) = \overleftarrow{w}_{ij}(l)$ and $\bar{s}_i(l) = \overleftarrow{s}_i(l)$, considering only in-paths of i, and $\bar{\mathbf{P}}(l) = \overleftarrow{\mathbf{P}}(l)$;

(c) $\bar{N}_i(l) = N_i(l)$, $\bar{w}_{i,j}(l) = w_{ij}(l)$ and $\bar{s}_i(l) = s_i(l)$ when all directions are considered, and $\bar{\mathbf{P}}(l) = \mathbf{P}(l)$.

In all the previous cases, we assume $\bar{\mathbf{P}}(0) = \mathbf{I}$.

The local clustering coefficient of order l in (2) can be defined, respectively, for cases (a), (b) and (c), as follows:

$$\mathbf{c}^{in}(l) = \overleftarrow{\mathbf{P}}(l)\mathbf{c}^{in}. \tag{4}$$

$$\mathbf{c}^{out}(l) = \overrightarrow{\mathbf{P}}(l)\mathbf{c}^{out}. \tag{5}$$

$$\mathbf{c}^{all}(l) = \mathbf{P}(l)\mathbf{c}^{all}. \tag{6}$$

where $\mathbf{c}^{in} = [c_i^{in}]_{i \in V}$ and $\mathbf{c}^{out} = [c_i^{out}]_{i \in V}$ are vectors with entries c_i^{in} and c_i^{out}, respectively, that represent the in and out weighted local clustering coefficients of node i, while $\mathbf{c}^{all}$ is the vector of local clustering coefficient for the graph D. We report the definition of such clustering coefficients in the Appendix, for the convenience, and they have been introduced in [11].

3 Systemic risk measure

We here propose a new indicator of systemic risk based on an extended version of the clustering coefficient described in the previous Section and which seems to be particularly effective for our purpose.

If the case of an undirected graph, we introduce the vector $\mathbf{h} = [h(l)]_{l=0,\cdots,diam(G)}$ such that

$$h(l) = \frac{1}{N} \sum_{i \in V} c_i(l). \tag{7}$$

Observe that $h(l)$ provides a feedback on how the nodes of the network are clustered together at a specific level l, being the mean of the clustering coefficients of order l. Then, the vector $\mathbf{h}$ collects the measures of all clusters. As a consequence, the distribution of its elements could give insights on the systemic risk of the network at each level l, as we will better be explained in the next Subsection.

This suggests, as a quite natural further step, to use the elements of **h** to define a new measure of systemic risk. We define the index $h^\star$, based on the h's, by taking their weighted mean as follows:

$$h^\star = \sum_{l=0}^{diam(G)} x_l h(l), \tag{8}$$

where $x_l \in [0, 1]$ such that $\sum_{l=0}^{diam(G)} x_l = 1$.

Notice that by means of $h^\star$ we are providing a specific systemic risk index that considers how each node is embedded in the network. In particular, the measure takes into account either the whole community structure around each node as well as the level of mutual interconnection of the nodes at a geodesic distance $l \geq 1$. The presence of x_l allows to introduce some flexibility in the computation of $h^\star$. Some remarks about possible distribution of x_l are reported in Subsection 3.1. Moreover, the effect of different weights x_l will be tested in the numerical Section.

Also, observe that proposing this measure we are in line with the scientific debate on how clustering coefficient of a network might be viewed as a measure of systemic risk (see e.g. [27] and [31]).

Analogously to formula (7), in the case of directed graph D, we can define $\mathbf{h}^{in} = [h^{in}(l)]_{l=0,\cdots,diam(G)}$, $\mathbf{h}^{out} = [h^{out}(l)]_{l=0,\cdots,diam(G)}$ and $\mathbf{h}^{all} = [h^{all}(l)]_{l=0,\cdots,diam(G)}$ where

$$h^{in}(l) = \frac{1}{N} \sum_{i \in V} c_i^{in}(l), \qquad h^{out}(l) = \frac{1}{N} \sum_{i \in V} c_i^{out}(l), \qquad h^{all}(l) = \frac{1}{N} \sum_{i \in V} c_i^{all}(l), \tag{9}$$

and, in this case, the global clustering coefficients are:

$$h^{in,\star} = \sum_{l=0}^{diam(G)} x_l h^{in}(l), \qquad h^{out,\star} = \sum_{l=0}^{diam(G)} x_l h^{out}(l), \qquad h^{all,\star} = \sum_{l=0}^{diam(G)} x_l h^{all}(l). \tag{10}$$

According to the literature which states that a high value of clustering coefficient is associated to a high level of systemic risk, we assume that the highest the value of $h^\star$'s in (8) and (10), the highest the systemic risk of the network.

3.1 Some remarks on the systemic risk measure

Notice that the coefficient x_l in (8) and (10) represents the "weight" to be assigned to $h(l)$ in the analysis of the entire community structure related to the nodes of the network.As already pointed out, $h(l)$ brings information on the community structure at level l. In this respect, the selection of a specific distribution for the x's leads to different ways to include the periphery of the graph in the systemic risk measurement.

A high concentration at a specific level l could indicate a particular attention to a more clustered area, revealing the intuition of the policy maker to focus on a part of the network where the risk diffusion could spread easily. For instance, if the mass of the weights x's is concentrated over the small values of l, then the considered systemic risk measure will take into major consideration the community structures close to the nodes of the graph. Differently, the case of concentration of the

x's over large values of l is associated to a systemic risk more sensitive to communities far from the nodes.

The corner case $x_l = 1$ reduces the vectors **h**'s to the clustering coefficients of order l. In this particular situation, systemic risk measures $h^\star$'s take into consideration communities at geodesic distance l from the nodes. This analysis might be relevant for the assessment of the stratified community structure of the network. Indeed, a complete analysis of the corner cases $x_l = 1$ with $l = 0, 1, \ldots, diam(G)$ allows to characterize the vulnerability of the network and its ability of absorbing shocks by assessing the presence of core-periphery communities. Notice that the very special case of $x_l = 1$ for $l = 0$ is the standard conceptualization of the clustering coefficient.

Moreover, the distinction between directed and undirected graphs leads to remarkable differences in the definition of the systemic risk measure of the network. When dealing with undirected graph, strong communities are associated to high values of the weights between the unordered couples of nodes forming an edge (see formula (2)), and this property is clearly included in the definition of the systemic risk measure, defined through formula (8). In the directed case, weights should be intended with a direction and the couples of nodes become ordered. Thus, a node can be associated to a strong community in terms of in-paths and a weak one when considering the out-paths (see (4) and (5)). Such a characteristic of the nodes represents the basis of a concept of "directed" systemic risk measure (see $h^{in,\star}$ and $h^{out,\star}$ in (10)), and this contributes to the understanding of the vulnerability of the network. Indeed, a shock occurring at a node $i \in V$ with a high level of h_i^{out} is expected to be rapidly propagated to the other nodes of the network, while a high value of h_i^{in} is associated to a probable infection of node i when the shock comes from outside it.

3.2 Simulated example

To show the effectiveness of the proposed high order clustering coefficients in capturing how a network is clustered at different levels – hence leading to a a powerful definition of a systemic risk measure, as presented in the previous section – we provide a simple weighted and directed graph D of 12 nodes (see Figure 1). This allows to easily stress the potential of the index $h(l)$ in respect to the classical weighted clustering coefficient defined in the literature.

We recall that the clustering coefficient should reflect the combined effect of the weights and the presence of triangles. Notice that, we focus on a directed graph since the vector of high order clustering coefficients $\mathbf{c}(l)$ of the underlying graph G can be obtained by simply rescaling the coefficients of D (see [10] for details).

Since the graph is directed, we can refer to different kind of geodesics (depending on whether the direction of the arcs is considered or not). We first compute the vector $\mathbf{c}^{all}$ of coefficients as provided in [11]. We remind that each coefficient c_i^{all} includes all kind of triangles which a node
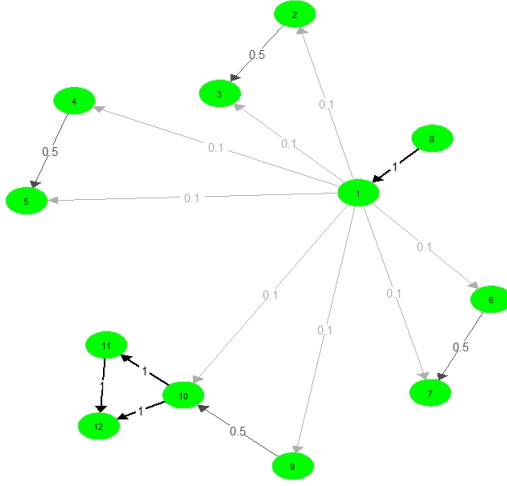
Figure 1: The simple weighted and directed graph D. Edges' opacity is proportional to weights.

belongs to.

We then evaluate clustering coefficients of order l by means of formula (6). Values are reported in Table 1.

Node	$\mathbf{c}^{all}(0)$	$\mathbf{c}^{all}(1)$	$\mathbf{c}^{all}(2)$	$\mathbf{c}^{all}(3)$
1	0.028	0.204	0.5	0
2	0.5	0.421	0.232	0.5
3	0.5	0.421	0.232	0.5
4	0.5	0.421	0.232	0.5
5	0.5	0.421	0.232	0.5
6	0.5	0.421	0.232	0.5
7	0.5	0.421	0.232	0.5
8	0	0.028	0.458	0.5
9	0.5	0.144	0.396	0
10	0.167	0.482	0.261	0
11	0.5	0.333	0.300	0.387
12	0.5	0.333	0.300	0.387
$h^{all}(l)$	0.391	0.338	0.300	0.356

Table 1: Clustering coefficients and systemic risk indicators of order l for the graph D when all paths are considered (neglecting directions).

x_l	$h^{all,\star}$
Decreasing Weights	0.360
Uniform Weights	0.346
Increasing Weights	0.340

Table 2: Values of the global systemic risk measure $h^{all,\star}$ for different weights' distribution

First, some remarks on the values of high order coefficients should be done. Looking at the clustering $\mathbf{c}^{all}(0)$, the node 1 has the lowest positive value (0.028), being part of triangles with low weights. This seems to suggest that node 1 is not remarkable in the context of communities. However, this value does not reflect the fact that the node 1 is adjacent to nodes with maximum clustering. In other words, the interconnection with high clustered nodes is softened by the presence of low weights. In addition, from node 1 we reach nodes 11 and 12 in two steps, having clustering equal to 0.5. The coefficients of order 1 and 2 ($c_1^{all}(1) = 0.204$, $c_1^{all}(2) = 0.5$) are increasing, showing that an higher order of clustering is able to capture the intensity of the communities around the node at distances greater than 1.

Node 10 has the second lowest value ($c_{10}^{all}(0) = 0.167$), but the value of $c_{10}^{all}(1)$ is almost three times. This suggests that the node forms communities with adjacent nodes with maximum clustering. It is worth noting the case of the node 8. This node does not contribute to form triangles, then its clustering coefficient is equal to zero. However, it has only one adjacent node and its direct connection with node 1 implies that the coefficient at level 1 completely absorbs the clustering value of the node 1. Also, clustering coefficients of order 2 and 3 are extremely high, due to connections to high clustered nodes through geodesics of length 2 and 3. These aspects allow to interpret the position of this node in a completely different perspective, especially in spreading risk.

The value of $h^{all,*}$ synthesizes the overall community structure of the network, thus providing a measure of the systemic risk associated to it. The choice of weights x_l can modulate the intensity of the measure $h(l)$ in formula (10), giving to this global network indicator a high degree of flexibility.

Here three possible scenarios for the weights x's are considered:

- Decreasing weights $x_l = \frac{(l+1)^{-1}}{\sum_{l=0}^{diam(G)} (l+1)^{-1}} = \frac{(l+1)^{-1}}{H_G}$ where H_G is the harmonic number of order $diam(G) + 1$.

- Uniform weights $x_l = \frac{1}{diam(G)+1}$

- Increasing weights $x_l = \frac{(l+1)}{\sum_{l=0}^{diam(G)} (l+1)} = \frac{2(l+1)}{(diam(G)+1)(diam(G)+2)}$

For instance, assuming that weights x_l are decreasing, we are reducing the impact of $h(l)$ respect the whole system when the distance l increases.

In this example, values of $h(l)$ state, on average, the presence of a strong community structure around the single node, as well as mutual interconnections at the maximal geodesic distance. As a consequence, both the closest and peripheral nodes have, on average, a similar influence on the node. In this regard, the distribution of weights is not very informative in this case. Indeed, different weights lead to very close values of $h^{all,\star}$ (see Table 2).

Moving to the analysis of the directed case, we can separately investigate patterns of in or out-clustering by using formulae (4) and (5).

Values are reported in Table 2. According to in-clustering, the vector $\mathbf{c}^{in}(0)$ considers in-triangles which a node belongs to. Referring to the 1st order, only two nodes (11 and 12) are reachable from a node with a positive in-clustering coefficient (i.e. the node 10). Notice that $c_{12}^{in}(1)$ is one half of $c_{11}^{in}(1)$ because the node is also reachable from 11.

Respect to the out-clustering, the node 8 is of interest; indeed, although its clustering coefficient is equal to zero, this node has positive coefficients of order 1 and 2, due to its connections via out-paths to nodes involved in out-triangles.

From the evaluation of the in- and out-patterns, a different behaviour in terms of systemic risk emerges, reflected by measures $h^{in,\star}$ and $h^{out,\star}$. The graph has a structure that seems more sensitive in receiving than spreading risk, and this characteristic is persistent also with different weights distributions (Table 4). In particular, the community structure at level 0 favours the receiving of the risk (captured by the highest $h^{in}(0)$). There is the presence of few nodes that spread risk also to peripheral nodes.

Node	$\mathbf{c}^{in}(0)$	$\mathbf{c}^{in}(1)$	$\mathbf{c}^{in}(2)$	$\mathbf{c}^{in}(3)$	$\mathbf{c}^{out}(0)$	$\mathbf{c}^{out}(1)$	$\mathbf{c}^{out}(2)$	$\mathbf{c}^{out}(3)$
1	0	0	0	0	0.071	0.063	0	0
2	0	0	0	0	0	0	0	0
3	0.50	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0.50	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0.50	0	0	0	0	0	0	0
8	0	0	0	0	0	0.071	0.063	0
9	0	0	0	0	0	0.500	0	0
10	0.50	0	0	0	0.500	0	0	0
11	0	0.50	0	0	0	0	0	0
12	0.50	0.25	0	0	0	0	0	0
Average	0.208	0.063	0	0	0.048	0.053	0.005	0

Table 3: Clustering coefficients and systemic risk indicators of order l for the graph D considering either in-paths or out-paths.

x_l	$h^{in,\star}$	$h^{out,\star}$
Decreasing Weights	0.115	0.036
Uniform Weights	0.068	0.026
Increasing Weights	0.035	0.015

Table 4: Values of the global systemic risk measure $h^{in,\star}$ and $h^{all,\star}$ for different weights' distribution

4 Numerical Analysis

As in [7], [21] and [27] we designed a global banking network using the Bank for International Settlements (BIS) consolidated statistics, which measure bank exposures to different countries. These statistics capture worldwide-consolidated claims of internationally active banks headquartered in BIS reporting countries. In particular, we consider international claims by a reporting country toward banks in counterparty countries. In this way, we focus on the lending activity of international banks. Here, nodes are countries and weighted arcs represent positive cross-border exposures.

We model each quarter of the year over the sample period (from the first quarter of 2005 to the end of 2017) through a single network, each one referred to links between banks of approximately 200 countries[1]. Figure 2 depicts the network at four different time periods. Density is equal to 0.046 at the 4th quarter of 2005, then the network is sparse. It is noticeable that networks become slowly denser over time (in terms of the number of transactions): density is indeed equal to 0.055 at the end of the time-period (4Q-2017). The number of arcs moves from 1540 to 2513 over the sample period, while the number of countries remains stable. The majority of countries is separated by at most two steps. Only very few countries are reachable with three steps.

As also shown in Figure 2, data are designed so that the resulting networks are characterized by core - periphery structures[2]. The core group consists of countries whose banking systems report data to the BIS in the analysed time-period, whereas other countries belong to the periphery group. According to the BIS data, periphery countries are analysed only as borrowers because for their banking systems only information on inflows is available. Thus, the selection of the core is strictly dependent on the data structure provided by the BIS. Table 5 shows the list of 24 core countries which reported to the BIS data about incoming and outgoing exposures of international financial claims.

A specific analysis will regard also a subset of the core countries that host global systemically

---

[1]The number of countries varies according to different time-periods. Indeed, few isolated nodes are present at specific times.

[2]In network theory, a core-periphery structure identifies a well-designed network model such that some nodes are densely connected, whereas others are sparsely connected, in a peripheral position (see [8]).
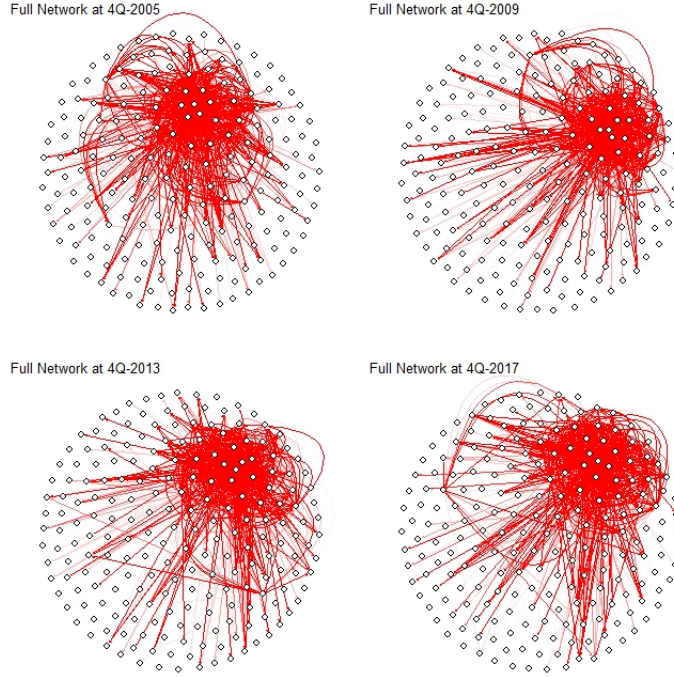
Figure 2: Cross-border global banking networks as on the end of 2005, 2009, 2013 and 2017. Red arrows represent arcs. Arcs opacity is proportional to weights (i.e., intensity of the exposures).

important banks (GSIBs)[3]. Since November 2011, the Financial Stability Board (FSB) releases the list of G-SIBs each year, based on the Basel Committee on Banking Supervision (BCBS) score of systemic risk (see [2] and [3]). These banks are asked to hold more capital (on top of Basel 3) and are subject to regulations that are more stringent. In particular, we will define GSI Countries, those countries in which at least a G-SIB is present (see Table 5 for a list of GSIB countries).

In order to understand the role of G-SIB, we will separate countries in three different clusters. In particular, we focus on GSI countries and other core countries (i.e. core countries that are not classified as GSI). A third residual group considers periphery countries. Figure 3 gives an idea of the role of different countries in the network at two different time periods. It is remarkable that specific core countries become more integrated in the dense part of the network by increasing, in particular, their out-flows.

According to this classification, we are able to compute global indicators based on a specific subset of countries. Initially, we compute classical local clustering coefficients $c_i^{all}(0)$ for the weighted and directed network (see [11]) and we aggregate these local coefficients separately for GSI and other core countries. Hence, we obtain two different estimates of $h^{all}(0)$ according to two different subsets

---

[3]See the FSB website for the updated list of G-SIBS banks (http://www.fsb.org/2017/11/fsb-publishes-2017-g-sib-list/) and the BCBS website for more information on how to assess systemically important banks (https://www.bis.org/bcbs/gsib/)

Australia (AU)	Core	
Austria (AT)	Core	
Belgium (BE)	Core	GSI
Canada (CA)	Core	
Chile (CL)	Core	
China (CN)		GSI
China Taipei (TW)	Core	
Finland (FI)	Core	
France (FR)	Core	GSI
Germany (DE)	Core	GSI
Greece (GR)	Core	
India (IN)	Core	
Ireland (IE)	Core	
Italy (IT)	Core	GSI
Japan (JP)	Core	GSI
Netherlands (NL)	Core	GSI
Norway (NO)	Core	
Portugal (PT)	Core	
Singapore (SG)	Core	
Spain (ES)	Core	GSI
Sweden (SE)	Core	GSI
Switzerland (CH)	Core	GSI
Turkey (TR)	Core	
United Kingdom (GB)	Core	GSI
United States (US)	Core	GSI

Table 5: List of core countries and countries where at least one GSIB has its headquarters.

of nodes.

In the entire period, GSI countries show a clustering coefficient higher than other core countries, confirming the important systemic role of the banks headquartered in these countries (see Figure 4). Furthermore, results suggest a different pattern between the two clusters of countries. In particular, GSI countries show a fall in clustering, starting from year-end 2008. This reduction is in line with the general reduction of clustering for the full network in 2009-2010, provided in [27] and ascribed to the perturbation in financial markets triggered by the Lehman failure.

It is interesting to notice that financial communities tends to weaken after the year 2011 for GSI countries. This phenomenon is also more intense starting from 2013, the year when the BCBS revised its methodology to assess GSIBs and the higher loss absorbency requirement, to better comply with the purpose of reducing the extent of failure of these banks.

This behaviour is partially explained by the pattern of out-degree and out-strength reported in Figure 5. There is a tendency in the network, particularly with other core countries, to diversify relationships as the average number of transactions increases and the average volume remains quite
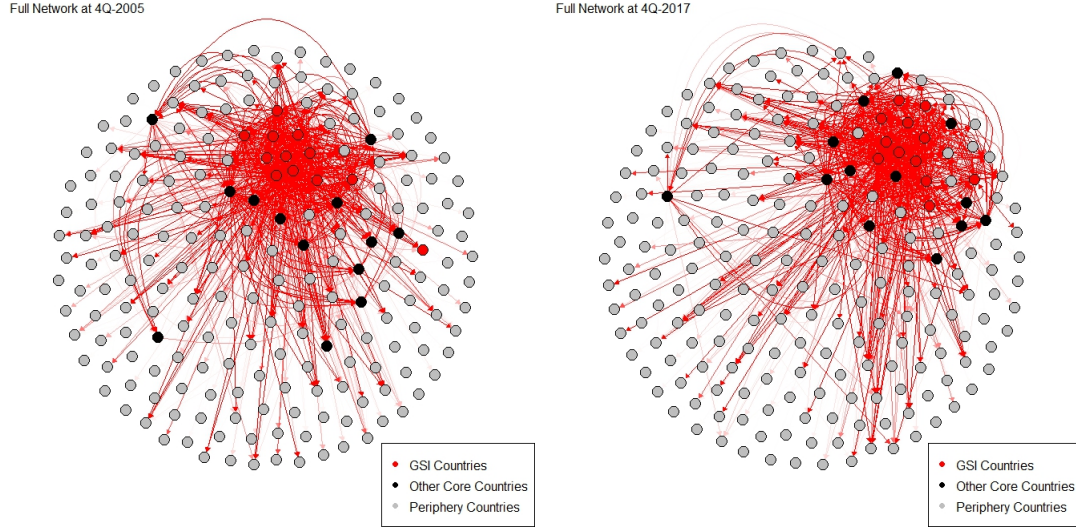
Figure 3: Cross-border global banking networks as on the end of 2005 and 2017 with the evidence of GSI, other core and periphery countries.
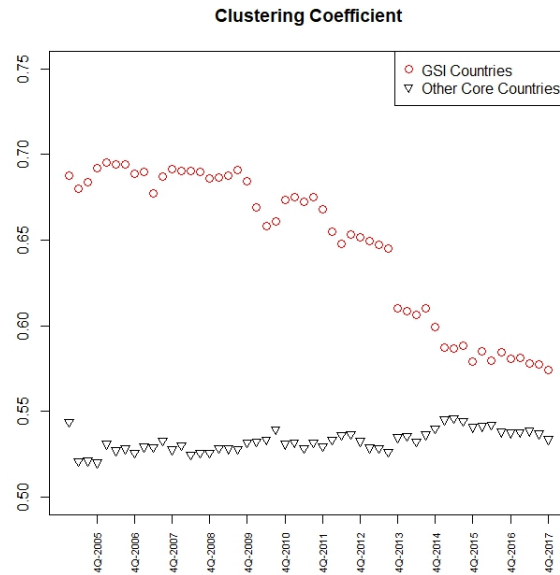

Figure 4: Global clustering Coefficients $h^{all}(0)$ computed by averaging local coefficients $c_i^{all}(0)$ at two different levels considering either only GSI countries or only other core countries respectively.

stable. The same pattern is not observed for GSI countries. In particular, considering the volume of transactions, evidence suggests that GSI countries are cutting their exposures to almost all partners. This behaviour provides an evidence of the effectiveness of the regulation, dampening the GSIB systemic impact.

In order to test the behaviour of clustering coefficients of order l, we compute $\mathbf{c}^{all}(1)$ by means
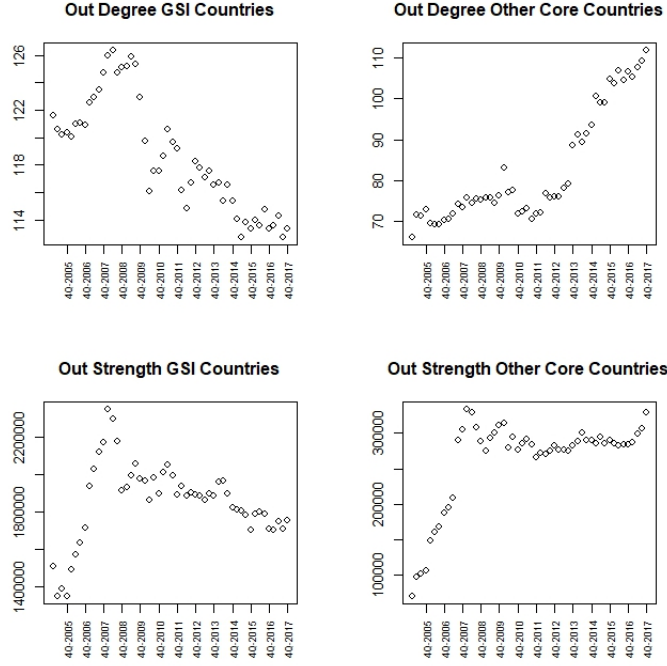
16

Figure 5: Out-degree and out-strength for GSI and other core countries.

of formula (6). In order to obtain two values of the synthetic indicator, one for GSI subset and one for other core countries subset, we average local coefficients of order 1 of countries belonging to the same subset. Results are reported in Figure 6.

We examine, in this way, the systemic risk at different observation scales. While classical clustering coefficient c_i^{all} takes into account how the node i form communities in the context of the overall system, $c_i^{all}(1)$ measures how much the neighbours of node i form triangles and communities. Patterns of Figure 6 show that GSI countries are also connected to well-established communities confirming the central role of these countries in term of systemic risk.

It is worth mentioning that GSI countries, against the reduction of average number and volume of transactions, maintain a high-level of communities, captured by $h^{all}(1)$, due to their relations with neighbours that belong to many triangles. This effect is partially induced by the contraction of lending activity towards periphery countries, leading to an increase of clustering coefficient of order 1.

On the left side, Figure 7 depicts global clustering coefficients of order 2 ($h^{all}(2)$) computed for the two subsets of countries. A similar pattern between GSI and other core countries is observed, specially over the last period. This behaviour can be partially explained by observing the ratio of geodesics of length 1 to the total number of geodesics for each cluster of countries (Figure 7, right side). Indeed, as already seen in Figure 5, the average number of borrowers for the two subsets of
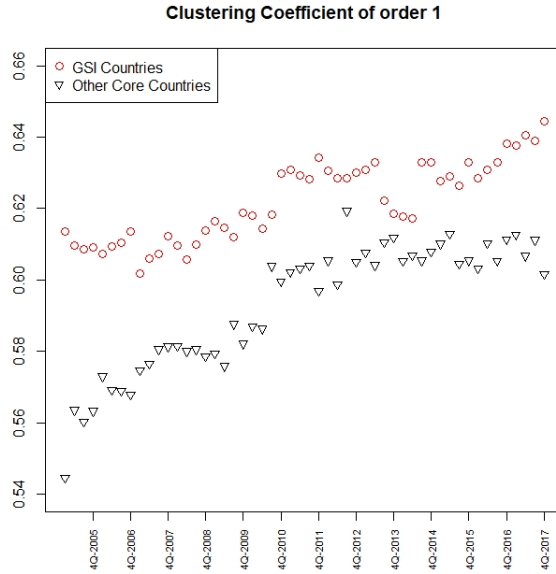
17

Figure 6: Global clustering coefficients of order 1 $h^{all}(1)$ computed by averaging local coefficients $c_i^{all}(1)$ at two different levels considering either only GSI countries or only other core countries respectively.
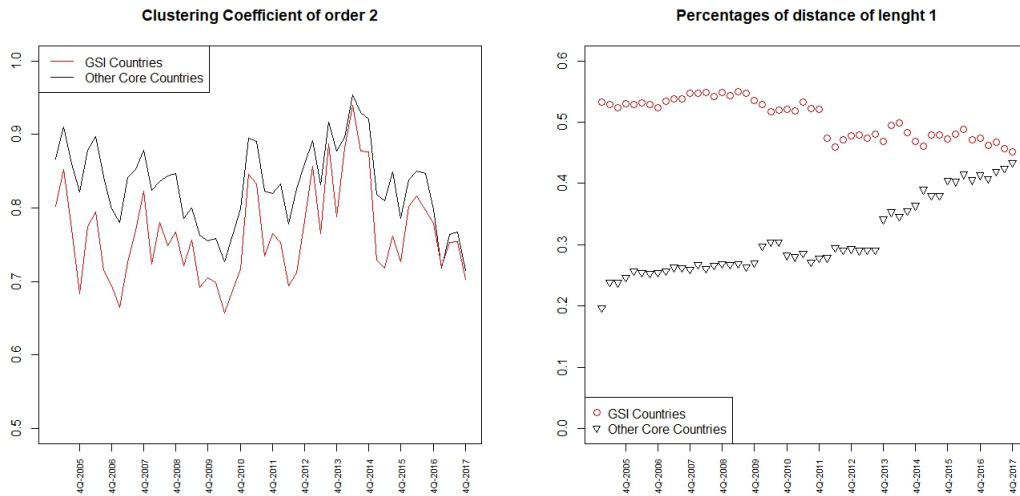

Figure 7: On the left side, Figure shows global clustering coefficients of order 2 $h^{all}(2)$ computed by averaging local coefficients $c_i^{all}(2)$ at two different levels considering either only GSI countries or only other core countries respectively. On the right side, Figure displays the percentage of geodesic of length 1 with respect to the total for each cluster (GSI and other core countries)

lenders behaves in a opposite manner since 2008-2009. On one hand, GSI countries reduced the average number of transactions, on the other hand, other core countries increase it. At the end of

the period, very close out-degrees are observed for both subsets, so that the percentage of nodes at distance 1 becomes very similar and clustering coefficient of order 2 is obviously affected by this fact. Additionally, the similarity of the 2-order clustering coefficients displays that both subsets are connected in two steps to comparable communities.

Figure 8 shows the high order clustering $\mathbf{h}^{all,*}$ computed by considering different choices of weights, for both subsets of countries. Specifically, we test here the same distributions of weights described in Section 3.2, reporting the results in Figure 8. A different weights' concentration produces distinct patterns, in terms of comparison between GSI and other countries. Focusing, for instance, on a decreasing distribution, it is confirmed the prominent role of GSI countries in spreading and receiving risk to (and from) their neighbours. On the contrary, assigning more weight to the relationships between nodes at higher distance, other core countries tend to have a pattern in line with GSI countries. Hence, we provide here a different view of systemic risk: unless GSI countries surely play a key role in spreading and receiving risk, by looking beyond the adjacent nodes, we deduce that also banks of other core countries can significantly contribute to risk diffusion. Furthermore, independently from the weights' distribution, in all cases, we have that the behaviour of both subsets tends to be aligned over the last years.

Figure 8: Figure reports the high order clustering coefficient for different choices of weights, where $\mathbf{h}^{all,*}$ is computed by averaging local coefficients for the two subsets of countries.

The numerical analysis has been also extended by considering in a separate way only either in-paths or out-paths. In this way, we catch only the effect of community structures in spreading or receiving risk. To this aim, we report in Figure 9 values of $h^{in}(0)$ and $h^{out}(0)$ for both subsets of countries. It is worth mentioning the different pattern between other core and GSI countries. The former ones tend to have higher connections of the in-type to their neighbours. GSI countries have instead, on average, an higher role in spreading risk towards their adjacent nodes. In particular, it is noticeable the specific pattern $h^{out}(0)$ for GSI countries since the end of 2011. Indeed, the structure of financial communities of out-type tends to weaken, with a significant decrease from 2013, probably due to the reaction of banks in these countries to the systemic risk regulation.

Figure 9: In and out clustering coefficients $h^{in}(0)$ and $h^{out}(0)$. Both coefficients are computed by averaging local coefficients considering either only GSI countries or only other core countries respectively.

As regard to level 1, we have a very different picture of the network (see Figure 10). GSI countries tend to show an higher level of connections of the in-type with strong community structures. On other hand, GSI countries spread, on average, risk toward countries that have lower out-clustering coefficients $h^{out}(0)$. So it is interesting to note that, in this case, there is only a low further propagation of risk. On other hand, we observe an increasing pattern of $h^{out}(1)$ for core countries over time showing that banks of these countries are more and more connected to banks of risk-giver countries.

Values of $h^{in}(2)$ and $h^{out}(2)$ have been reported in Figure 11. In these networks, the maximal length of in-geodesics is equal to 2. However, we have a very low proportion of countries (see Figure 12, left side) that are reachable in two steps via an in-path[4]. So, differences observed for the $h^{in}(2)$ clustering between the two subsets are mainly motivated by the behaviour of specific countries. On average, other core countries, when act as borrowers, are more connected, via in-paths of length 2, to strong community structure. Concerning the out-clustering of level 2, all core countries show very low connections. Although, over time, the percentages of out-distances of length 2 is reducing for GSI countries and increasing for other core countries, all countries are, on average, mostly connected at distance 2 with peripheral countries characterized by low risk.

Now, we focus on the systemic risk measures $h^{in,\star}$ and $h^{out,\star}$ regarding in and out-flows respectively (see Figure 13). As stressed in [31], an higher clustering coefficient of the in-type may reflect higher systemic risk because failure of the borrowing node in an "in" triangle can trigger

---

[4]The ratio of in-geodesics of length 1 to the total number of in-geodesics is equal to 96% for GSI countries and around 90% for other core countries.
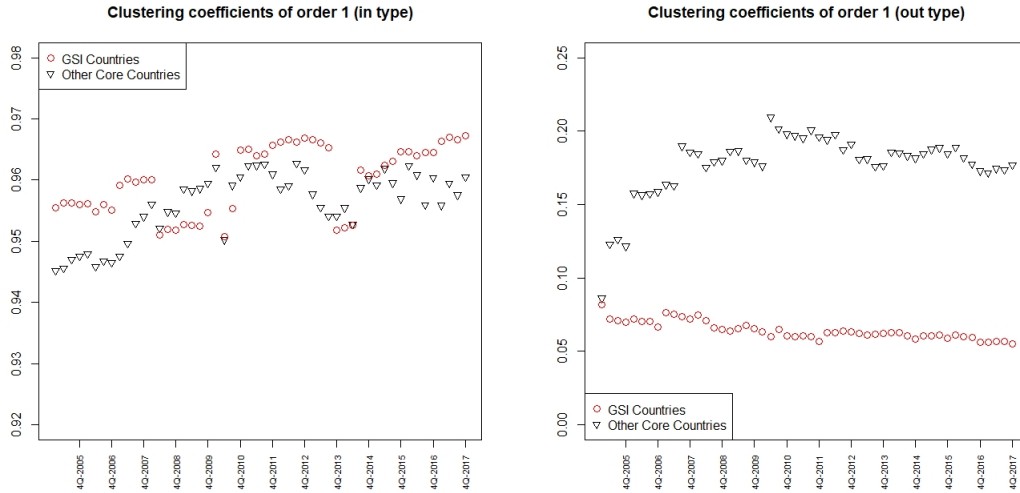
22

Figure 10: In and out clustering coefficients of order 1, $h^{in}(1)$ and $h^{out}(1)$. Both coefficients are computed by averaging local coefficients considering either only GSI countries or only other core countries respectively.

Figure 11: In and Out Clustering of order 2, $h^{in}(2)$ and $h^{out}(2)$. Both coefficients are computed by averaging local coefficients considering either only GSI countries or only other core countries respectively.

simultaneous non-repayments to the lending nodes, and this can make them unable to honour their own obligations. We show that both the in-clustering and the $h^{in,\star}$ assess the high state of stress in the network. It is worth mentioning how other core countries are more affected as risk taker than GSI countries. This effect is more evident when weights are more concentrated on adjacent nodes
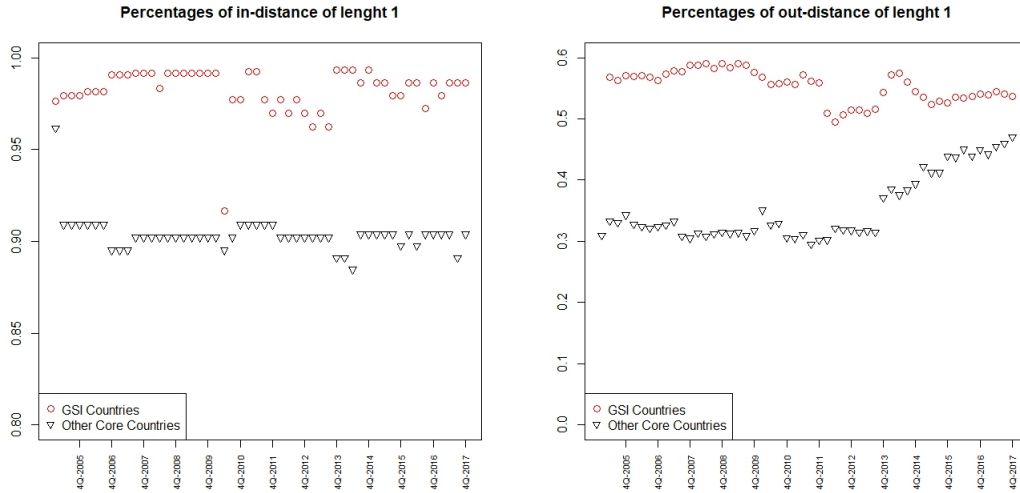
23

Figure 12: Percentages of directed (in and out respectively) geodesic of length 1 with respect to the total for each cluster (GSI and other core countries)

because of the high level of interaction of these countries as borrowers. It is instead more noticeable the effect of weights' distribution on GSI countries. We have indeed that values of $h^{in,\star}$ for this subset are decreasing approximatively from 0.87 to 0.7 when increasing weights are chosen. These countries are less affected, when they act as borrowers, by the effect of countries that are at higher distances. Differently focusing on $h^{out,\star}$, we observe a greater value for GSI countries in the period 2009-2011. It is also confirmed the reduction of GSIB systemic impact over the last years, mainly characterized by the reduction of the exposure of GSI countries.

Figure 13: In and out high order clustering coefficients $h^{\star,in}$ and $h^{\star,out}$. Coefficients are computed for both subsets of countries and with different weights.

5 Conclusions

Systemic risk in finance is a concept not easy to be formalized through a quantitative measure and a huge and fast growing literature is interested in this issue. A quite natural approach is based on the use of complex networks.

In a financial system, the interconnectedness among entities plays a fundamental role in situations of distress. Moving from this fact, we exploit the concept of community, usually relevant in understanding the relationship between interconnectedness and systemic risk. In particular, we consider a generalization of the concept of clustering coefficient in order to catch both the presence of clustered areas around a node and/or high levels of mutual interconnections at different distances from the

node itself. We provide a new systemic risk measure computed as the weighted average of high order clustering coefficients at different levels. On one hand, this proposal leads to a synthetic indicator able to assess the general state of stress of the financial system. On the other hand, the distribution of weights allows to introduce a degree of flexibility, in order to modulate the effects of both adjacent nodes and peripheral nodes.

An empirical application to time-varying global banking network is developed. Results show the effectiveness of these measures in reflecting how systemic risk has changed over the last years, also in the light of the recent financial crisis. Furthermore, we emphasize a different pattern of behaviour between countries where a GSIB is headquartered and other core countries, more noticeable since 2013. This effect, that could be interpreted as a reaction to the specific regulation inducing banks to contain their "systemic" nature, is in line with the recent report by the Committee on the Global Financial System [12] that shows that GSIBs become more selective and have also repositioned themselves toward less complex activities, as a response to the regulatory reforms process that is under way.

References

[1] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, 101(11):3747–3752, 2004.

[2] Basel Committee on Banking Supervision. Finalising post-crisis reforms. Technical report, 2017.

[3] Basel Committee on Banking Supervision. Global systemically important banks: revised assessment methodology and the higher loss absorbency requirement. Technical report, 2018.

[4] S. Battiston, D. Delli Gatti, M. Gallegati, B. Greenwald, and J. E. Stiglitz. Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *Journal of economic dynamics and control*, 36(8):1121–1141, 2012.

[5] S. Battiston and S. Martinez-Jaramillo. Financial networks and stress testing: Challenges and new research avenues for systemic risk analysis and nancial stability implications. *Journal of Financial Stability*, 35(6-16), 2018.

[6] M. Billio, M. Getmansky, A. W. Lo, and L. Pelizzon. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of financial economics*, 104(3):535–559, 2012.

[7] P. Bongini, G.P. Clemente, and R. Grassi. Interconnectedness, G-SIBs and network dynamics of global banking. *Finance Research Letters*, 2018.

[8] S. P. Borgatti and M. G. Everett. Models of core/periphery structures. *Social networks*, 21(4):375–395, 2000.

[9] F. Caccioli, P. Barucca, and T. Kobayashi. Network models of financial systemic risk: A review. *Journal of Computational Social Science*, 1(1):81–114, January 2018.

[10] R. Cerqueti, G.P. Clemente, and R. Grassi. Stratified communities in complex business networks. *Available on ArXiv*, 2018.

[11] G.P. Clemente and R. Grassi. Directed clustering in weighted networks: a new perspective. *Chaos, Solitons & Fractals*, 107(26-38), 2018.

[12] Committee on the Global Financial System. Structural changes in banking after the crisis,. Technical Report 60, 2018.

[13] R. Cont and A. Minca. Credit default swaps and systemic risk. *Annals of Operations Research*, 247(2):523–547, 2016.

[14] O. De Bandt and P. Hartmann. Systemic risk: a survey. Technical Report 35, European Central Bank, 2000.

[15] D. Di Gangi, F. Lillo, and D. Pirino. Assessing systemic risk due to fire sales spillover through maximum entropy network reconstruction. *Journal of Economic Dynamics and Control*, 94:117–141, 2018.

[16] G. Fagiolo. Clustering in complex directed networks. *Physical Review E*, 76(2), aug 2007.

[17] S. Fortunato. Community detection in graphs. *Physics reports*, 486(3-5):75–174, 2010.

[18] R. Garratt, L. Mahadeva, and K. Svirydzenka. Mapping systemic risk in the international banking network. Technical Report 43, Bank of England, 2011.

[19] M. Girvan and M.E.J. Newman. Community structure in social and biological networks. *PNAS*, 99(12):7821–7826, 2002.

[20] P. Giudici, P. Sarlin, and A. Spelta. The interconnected nature of financial systems: direct and common exposures. *Journal of Banking & Finance*, 2017.

[21] P. Giudici and A. Spelta. Graphical network models for international financial flows. *Journal of Business & Economic Statistics*, 34(1):128–138, 2016.

[22] A.G. Haldane and R.M. May. Systemic risk in banking ecosystems. *Nature*, 469(7330):351, 2011.

[23] A. Helfgott. Operationalising systemic resilience. *European Journal of Operational Research*, 268(3):852–864, 2018.

[24] A. Hübsch and U. Walther. The impact of network inhomogeneities on contagion and system stability. *Annals of Operations Research*, 254(1-2):61–87, 2017.

[25] S. Markose, S. Giansante, and A. R. Shaghaghi. 'Too interconnected to fail' Financial network of US CDS market: Topological fragility and systemic risk. *Journal of Economic Behavior & Organization*, 83(3):627–646, 2012.

[26] P. McGuire and N.A. Tarashev. Tracking international bank flows. Technical report, BIS Quarterly Review, 2006.

[27] C. Minoiu and J.A. Reyes. A network analysis of global banking: 1978-2010. *Journal of Financial Stability*, 9(2):168–184, July 2013.

[28] A.R. Neveu. A survey of network-based analysis and systemic risk measurement. *Journal of Economic Interaction and Coordination*, 13(2):241–281, July 2018.

[29] J.P. Onnela, J. Saramäki, J. Kertész, and K. Kaski. Intensity and coherence of motifs in weighted complex networks. *Physical Review E*, 71(6), 2005.

[30] J.C. Rochet and J. Tirole. Interbank lending and systemic risk. *Journal of Money, Credit and Banking*, 28(4):733–762, 1996.

[31] B.M. Tabak, M. Takamib, J.M.C. Rochac, D.O. Cajueirod, and S.R.S. Souzae. Directed clustering coefficient as a measure of systemic risk in complex banking networks. *Physica A: Statistical Mechanics and its Applications*, 394:211–216, January 2014.

[32] G. Torri, R. Giacometti, and Paterlini S. Robust and sparse banking network estimation. *European Journal of Operational Research*, 270(1):51–65, 2018.

[33] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications.* Cambridge University Press, New York, NY., July 1994.

[34] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, jun 1998.

[35] Y. Zhu, F. Yang, and W. Ye. Financial contagion behavior analysis based on complex network approach. *Annals of Operations Research*, 268(1-2):93–111, 2018.

Appendix

We here report the definition of the clustering coefficients in the directed case (see [11] for details). The overall clustering coefficient for weighted and directed network is defined as:

$$c_i^{all} = \frac{\frac{1}{2}[(\mathbf{W} + \mathbf{W}^T)(\mathbf{A} + \mathbf{A}^T)^2]_{ii}}{s_i(d_i - 1) - 2s_i^{\leftrightarrow}}$$

where $s_i^{\leftrightarrow}$ is the strength related to bilateral arcs between the node i and its adjacent nodes, defined as:

$$s_i^{\leftrightarrow} = \sum_{j \neq i} a_{ij} a_{ji} \frac{(w_{ij} + w_{ji})}{2}.$$

The numerator of the overall coefficient takes into account all directed triangles that a node i actually forms with its neighbours, weighted with the average weight of the links connecting i to its adjacent nodes. Then, it is divided by all possible (appropriately weighted) directed triangles that i could form. Observe that $2s_i^{\leftrightarrow}$ represents the number of "false" triangles, being formed by i and by a pair of directed arcs pointing to the same node, e.g., $i \to j$ and $j \to i$. Indeed, being the network directed, i can form up to two triangles with each pair of adjacent nodes, including two "false" potential triangles for each bilateral link. To this reason, $2s_i^{\leftrightarrow}$ does not contribute to the number of possible directed triangles and has to be removed by the denominator.

Notice that, if the network is undirected, $\mathbf{c}^{all} = \mathbf{c}$.

Accordingly to the previous definition, in and out weighted local clustering coefficients of node i are defined, respectively, as:

$$c_i^{in} = \frac{\frac{1}{2}[\mathbf{W}^T(\mathbf{A} + \mathbf{A}^T)\mathbf{A}]_{ii}}{\overleftarrow{s}_i\left(\overleftarrow{d}_i - 1\right)}$$

and

$$c_i^{out} = \frac{\frac{1}{2}[\mathbf{W}(\mathbf{A} + \mathbf{A}^T)\mathbf{A}^T]_{ii}}{\overrightarrow{s}_i\left(\overrightarrow{d}_i - 1\right)}.$$