Luu, Duc Thi; Lux, Thomas; Yanovski, Boyan

Working Paper

Structural correlations in the Italian overnight money market: An analysis based on network configuration models

Economics Working Paper, No. 2017-02

Provided in Cooperation with:
Christian-Albrechts-University of Kiel, Department of Economics

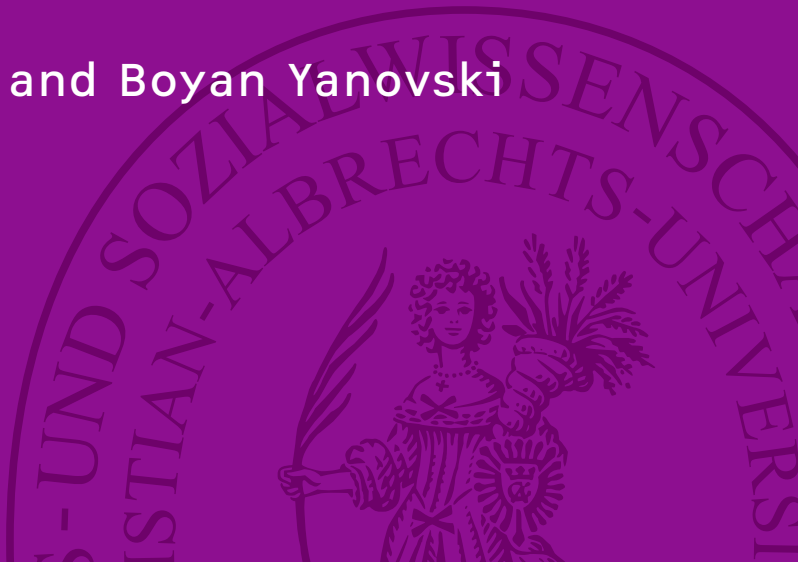
Suggested Citation: Luu, Duc Thi; Lux, Thomas; Yanovski, Boyan (2017) : Structural correlations in the Italian overnight money market: An analysis based on network configuration models, Economics Working Paper, No. 2017-02, Kiel University, Department of Economics, Kiel

This Version is available at:
https://hdl.handle.net/10419/156399

C A U
Christian-Albrechts-Universität zu Kiel
Department of Economics
Economics Working Paper
No 2017-02

Structural correlations in the Italian overnight money market: An analysis based on network configuration models

by Duc Thi Luu, Thomas Lux and Boyan Yanovski



issn 2193-2476

Structural correlations in the Italian overnight money market: An analysis based on network configuration models

Duc Thi Luu[1,2], Thomas Lux[1], Boyan Yanovski[1]

(1) Department of Economics, University of Kiel, Germany

(2) OFCE-SciencesPo, Nice, France

Ver. March 28, 2017

Abstract

We study the structural correlations in the Italian overnight money market over the period 1999-2010. We show that the structural correlations vary across different versions of the network. Moreover, we employ different configuration models and examine whether higher-level characteristics of the observed network can be statistically reconstructed by maximizing the entropy of a randomized ensemble of networks restricted only by the lower-order features of the observed network. We find that often many of the high order correlations in the observed network can be considered emergent from the information embedded in the degree sequence in the binary version and in both the degree and strength sequences in the weighted version. However, this information is not enough to allow the models to account for all the patterns in the observed higher order structural correlations. In particular, one of the main features of the observed network that remains unexplained is the abnormally high level of weighted clustering in the years preceding the crisis, i.e. the huge increase in various indirect exposures generated via more intensive interbank credit links.

JEL classification: G21; G01; E42.

Keywords: Interbank Network; Structural Correlations; Clustering Coefficients; Configuration Models; Network Reconstruction.

1 Introduction

Understanding the topological structure of complex systems is crucial in many areas, e.g. in ecology, physics, neuroscience, epidemiology, economics, and finance. Statistics pertaining to properties related to single nodes, linked node pairs and linked node triplets are often referred to as structural correlations of the first, second and third order respectively. The study of these structural correlations is one of the most common approaches for examining the properties of a network. The degree and strength sequences are examples of first order structural correlations. Statistics pertaining to properties related to linked node pairs reveal information about the type of mixing (assortative vs disassortative) that takes place in the network, while those related to linked node triplets are indicative of the clustering behavior.

In terms of second order correlations, a network would exhibit assortative mixing if its nodes are predominantly connected to other nodes having similar degrees or strengths. In contrast, disassortative mixing occurs when the connected nodes are dissimilar (see, for example, Newman, 2002; Newman, 2003a). This concept can be extended to directed networks yielding four mixing categories, i.e. in-in, in-out, out-in, and in-in mixing as illustrated in Figure (1) (see, for example, Foster et al., 2010; Piraveenan et al., 2012; van der Hoorn and Litvak, 2015). It should be emphasized that, in many real world networks, the mixing behavior of the directed version can differ a lot from the one observed in the undirected version. Furthermore, the same directed network can have assortative and disassortative aspects related to the mixing categories mentioned above (see, for example, Foster et al., 2010).

At the level of a single node, in the binary case, second order structural correlations can be expressed in terms of a relationship describing the average degree of the nearest neighbors (ANND) of a node as a function of that node's own degree. If the ANND is an increasing function of degree, this can be considered evidence in favor of assortative mixing. In contrast, a decreasing function would signal disassortativity. For the whole network, the Pearson correlation coefficient between the degrees of pairs of linked nodes is often used to assess whether a network displays disassortative or assortative mixing (Newman, 2002; Newman, 2003a).

In addition, we can decompose the overall assortativity coefficient into the contributions of each node, i.e. we can measure the local assortativity associated with each node. Such a decomposition can reveal which nodes contribute to the overall observed mixing nature of the network and which are associated with the opposite type of mixing (see, for example, Piraveenan at al., 2012). For instance, a globally assortative network may be locally

disassortative and vice versa. It is worth noting that two networks with the same degree distribution and the same global level of assortativity may display different patterns of local assortativity.

The analysis of the second order structural correlations in the binary case can be straightforwardly extended to weighted networks by employing a measure that takes the average strength of the nearest neighbors (ANNS) of a node or by computing the Pearson correlation coefficient between the strengths of pairs of linked nodes.

As is common in the literature, we use clustering coefficients as measures of the third order structural correlations in the network. A clustering coefficient measures the tendency of two neighbors of a particular node to also be connected to each other (e.g. Newman, 2003b). If we define a node triplet as three nodes connected by at least 2 edges, then, considering a network as a whole, the transitivity ratio (T) is equal to the number of triplets in which all three nodes are directly connected (forming a triangle) as a fraction of all node triplets (e.g. Newman, 2003b). An alternative measure is proposed by Watts and Strogatz (1998), which can capture the observed local clustering. The average of these local clustering coefficients can be used as an alternative measure of clustering for the whole network. The difference between the transitivity ratio and the average clustering coefficient is that, while in the former we calculate the ratio of the means, in the latter we take the mean of the ratios. In addition, for the directed version of a network, it is useful to differentiate between different relationship types depending on the direction of the edges in a triangle, i.e. inward, outward, cyclic, and middleman relationships, since as shown in Figure (2), the different relationships have different implications in terms of the risk exposure the individual banks are facing and in terms of systemic risk (see, for example, Fagio, 2007; Tabak et al., 2014). In weighted networks, the weighted clustering coefficients can be formulated in several ways, depending on how we take into account the roles of the strengths and weights of the nodes in each triangle (see, for example, Barrat et al., 2004; Onnela et al., 2005; Zhang and Horvath, 2005; Holme et al., 2007 [1]).

To assess whether the observed higher order structural correlations in a network are typical of a network with the observed lower order structural correlations, we can employ a randomization procedure based on the observed lower order patterns in the attempt to arrive at a suitable null model to test against non-random patterns. Such null models create a whole ensemble of networks out of a subset of the information necessary to completely define the observed network. This is why this technique can also be used for filling in

[1] We refer the readers to Saramäki et al. (2007) for a comparison between different methods for calculating the local weighted clustering coefficients.

unavailable information. The most basic null models are the random graph models (RGM), which specify only global constraints such as the node degree average in the binary case or the node strength average in the weighted case. Since in these models, all nodes are treated homogeneously, there is no difference between the expected topological properties across nodes, which does not happen often in real world networks. In order to capture the intrinsic heterogeneity in the capacity of the individual nodes, a popular approach is to generate the microcanonical ensemble of networks having exactly the same degree sequence (or the same strength sequence in weighted networks) as the one in the observed network (see, for example, Maslov and Sneppen, 2002; Maslov et al., 2004; Zlatic et al., 2011). However, this "hard" approach suffers from various limitations [2]. Based on the maximum-entropy and maximum-likelihood methods, recent advances in the specification of configuration models propose a "soft" approach that enforces the constraints on average over an ensemble of randomized networks (e.g. Garleschelli and Loffredo, 2008; Squartini and Garlaschelli, 2011; Squartini et al., 2011a; Squartini et al., 2011b; Mastrandrea et al., 2014; Squartini et al., 2015). This approach allows us to sample network ensembles more efficiently and in an unbiased manner (Squartini et al., 2015).

In this paper we analyze the structural correlations in a particular financial system, i.e. the Italian electronic market for interbank deposits (e-MID). While some of the network properties of the e-MID market have been previously studied (see, for example, De Masi et al., 2006; Fricke, 2012; Fricke et al., 2013; Finger et al., 2013; Fricke and Lux, 2015a; Fricke and Lux, 2015b; Squartini et al., 2015; Cimini et al., 2015a), what is novel in our paper is that: (i) we provide a more comprehensive analysis of the structural correlations in all versions of the network, and employ both local as well as global measures for analyzing such patterns; (ii) we employ configuration models to investigate whether the intrinsic node heterogeneity represented by the degree sequence (in the binary network) and/or strength sequence (in the weighted network) can explain higher order structural correlations observed in the system; (iii) we utilize the so called Directed Enhanced Configuration Model as a null model for the directed weighted version of the network, which makes use of the available information about the direction of the edges in the network.

We use quarterly data for the e-MID network over the period 1999-2010 and restrict our analysis to the Italian banks participating in this market, because foreign banks are not frequently active in the market. Particularly, from the onset of the financial crisis in 2008 onward, non-Italian banks have basically withdrawn from this electronic market (e.g. see

[2]See Squartini and Garlaschelli (2011) or Squartini et al. (2015) for a deeper discussion.

Fricke et al., 2013) [3].

The remainder of this paper is structured as follows. In Sec. 2 we provide a general framework for analyzing the structural correlations in different versions of the observed network as well as the algorithm for generating an ensemble of randomized networks from given constraints. In Sec. 3, we analyze the structural correlations in the undirected and directed binary versions of the e-MID network, and then compare the results to those obtained from the associated null models. In Sec. 4, we provide a similar analysis of the undirected as well as directed weighted versions of the network. Sec. 5 contains a discussion of the results as well as directions for future research. At the end of this paper, the Appendix provides additional details concerning the measures of structural correlations.

2 Structural correlations in complex networks

2.1 Undirected networks

2.1.1 General notation

In the undirected version, suppose we have a network (of size n) characterized by a symmetric adjacency matrix $A = \{a_{ij}\}_{nxn}$ and a symmetric weighted matrix $W = \{w_{ij}\}_{nxn}$ ($a_{ii} = w_{ii} = 0$). The degree and strength sequences for each node i are respectively defined as

$$k_i^{un} = \sum_{j=1}^{n} a_{ij}, \tag{1}$$

and

$$s_i^{un} = \sum_{j=1}^{n} w_{ij}. \tag{2}$$

The total degree and total strength over all nodes in the network are given by

$$m = \frac{1}{2} \sum_{i=1}^{n} k_i^{un}, \tag{3}$$

[3]The transactions between banks are aggregated into quarterly data, since at the higher frequencies the matrix of the trades between banks is very sparse. From the network perspective, we are more interested in existing long-term relationships (credit lines) rather than single transactions. Since such credit lines will typically not be activated on each day, a sufficiently long horizon is necessary to extract such information from the data. For a more detailed description of the e-MID dataset, we refer the readers to the studies of Fricke et al. (2013) and Finger et al. (2013), or to the e-MID website `http://www.e-mid.it/`.

and

$$w_{tol} = \frac{1}{2}\sum_{i=1}^{n} s_i^{un}. \tag{4}$$

2.1.2 Structural correlations in undirected networks

Assortativity Analysis

Regarding assortativity, we use two measures, i.e. the average degree (strength in the weighted case) of the nearest neighbors as well as the Pearson correlation coefficient (hereafter: Pearson coefficient) between degrees (strengths in the weighted case).

The average degree and strength of the nearest neighbors

The average degree of the nearest neighbors (ANND) of node i in the binary version of a network is given by

$$k_{nn,i}^{un} = \frac{\sum_{j=1}^{n} a_{ij} k_j^{un}}{k_i^{un}}. \tag{5}$$

For the weighted version, the average strength of the nearest neighbors (ANNS) of node i is defined as

$$s_{nn,i}^{un} = \frac{\sum_{j=1}^{n} a_{ij} s_j^{un}}{k_i^{un}}. \tag{6}$$

Treating k_{nn}^{un} as a function of k^{un}, an overall positive (negative) correlation between k_{nn}^{un} and k^{un} suggests assortative (disassortative) mixing in the binary version of the network. In the weighted case, a positive (negative) correlation between s_{nn}^{un} and s^{un} evidences assortative (disassortative) mixing.

We can also compute the averages of ANND and ANNS over the whole network respectively as

$$\bar{k}_{nn}^{un} = \frac{1}{n}\sum_{i=1}^{n} k_{nn,i}^{un} \tag{7}$$

and

$$\bar{s}_{nn}^{un} = \frac{1}{n}\sum_{i=1}^{n} s_{nn,i}^{un}. \tag{8}$$

Pearson correlation coefficient of the node degrees and of the node strengths

The second measure of mixing computes the Pearson's correlation between two degree sequences (see the Appendix for details). Practically, the main idea to measure such a correlation is that, from the adjacency matrix, first, we obtain a list of m edges, that is the list of pairs of nodes (i_e, j_e) where $a_{i_e j_e} = 1$, (for $e = 1, 2, ..., m$, $1 \leq i_e, j_e \leq n$). Next, for each e, we get two degrees $k_{i_e}^{un}$, $k_{j_e}^{un}$, and two strengths $s_{i_e}^{un}$, $s_{j_e}^{un}$ associated with the pair of nodes (i_e, j_e) and compute the correlation coefficients of the degrees and strengths at either

ends of an edge (e.g. Newman, 2002; Newman, 2003a). In the binary case, if the correlation coefficient of the degrees, r_{bin}^{un}, is negative, it signals the presence of disassortativity, while a positive value implies the opposite. In the weighted case, the same interpretation holds for the correlation coefficient of the strengths, r_{w}^{un}, but r_{bin}^{un} and r_{w}^{un} are not necessary equal.

Clustering coefficients

According to Watts and Strogatz (1998), the undirected binary clustering associated with node i is defined as

$$C_{bin,i}^{un} = \frac{\sum_{j\neq i}\sum_{k\neq i,j} a_{ij}a_{jk}a_{jk}}{\sum_{j\neq i}\sum_{k\neq i,j} a_{ij}a_{ik}}. \tag{9}$$

Following Onnela et al. (2005), we obtain the local weighted clustering associated with node i in undirected version of the network as

$$C_{w,i}^{un} = \frac{\sum_{j\neq i}\sum_{k\neq i,j} w_{ij}^{\frac{1}{3}} w_{jk}^{\frac{1}{3}} w_{jk}^{\frac{1}{3}}}{\sum_{j\neq i}\sum_{k\neq i,j} a_{ij}a_{ik}}. \tag{10}$$

Note that $C_{w,i}^{un}$ in Eq. (10) is invariant to weight permutation for each triangle and it takes into account the weights of all associated edges. In addition, it is easy to show that if $A = W$, we will have $c_{w,i}^{un} = c_{bin,i}^{un}$.

To analyze the evolution of the third order correlations over time, we define the average of $\{C_{bin,i}^{un}\}_{i=1}^{n}$ as

$$\bar{C}_{bin}^{un} = \frac{1}{n}\sum_{i=1}^{n} C_{bin,i}^{un}, \tag{11}$$

and the average of $\{C_{w,i}^{un}\}_{i=1}^{n}$ as

$$\bar{C}_{w}^{un} = \frac{1}{n}\sum_{i=1}^{n} C_{w,i}^{un}. \tag{12}$$

2.2 Directed networks

2.2.1 General definitions

In a directed network, the two matrices A and W may not be symmetric (i.e. $A \neq A^T$ and $W \neq W^T$). We then distinguish between in-degree and out-degree for every node i as

$$k_i^{in} = \sum_{j=1}^{n} a_{ji}, \tag{13}$$

and

$$k_i^{out} = \sum_{j=1}^{n} a_{ij}. \tag{14}$$

Similarly, instead of the total strength of the undirected version, we can distinguish between in-strength and out-strength for every node i.

Matters become even more complex for the higher-order structural correlations. Taking the directions of edges into account (as in Figure (1)), two types of nodes (giving and receiving) give rise to four types of relationships and four versions of ANND and ANNS for each node i: in-in, in-out, out-in, and out-out versions of ANND and ANNS, e.g.

$$k_{nn,i}^{in-in} = \frac{\sum_{j=1}^{n} a_{ji} k_j^{in}}{k_i^{in}}, \tag{15}$$

and

$$s_{nn,i}^{in-in} = \frac{\sum_{j=1}^{n} a_{ji} s_j^{in}}{k_i^{in}}, \tag{16}$$

with the other three types of statistics being obtained by replacing in-degrees and in-strengths by the pertinent entities.

In each version, the interpretation of the relationship between the ANND and node degree and between the ANNS and node strength is similar to the one for the measures discussed in the undirected case. That is, a negative (positive) relationship signals disassortativity (assortativity) in the respective class of relationships.

Similarly, the four possible combinations between giving and receiving nodes are associated with four global assortativity coefficients, i.e. r_{bin}^{in-in}, r_{bin}^{in-out}, r_{bin}^{out-in}, and $r_{bin}^{out-out}$ (see the Appendix for further details). Their weighted counterparts are r_{w}^{in-in}, r_{w}^{in-out}, r_{w}^{out-in}, and $r_{w}^{out-out}$. The algorithm for calculating these binary (weighted) coefficients is still similar to the one used for r_{bin}^{un} (or r_{w}^{un} for the weighted case), except for the requirement that the directions of edges (see Figure (1)) must be taken into account.

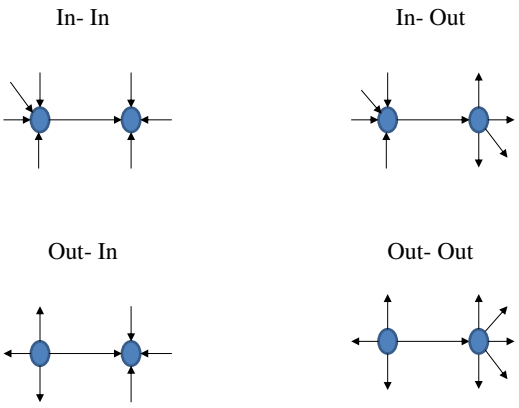


Figure 1: Degree-degree dependencies in the directed version.

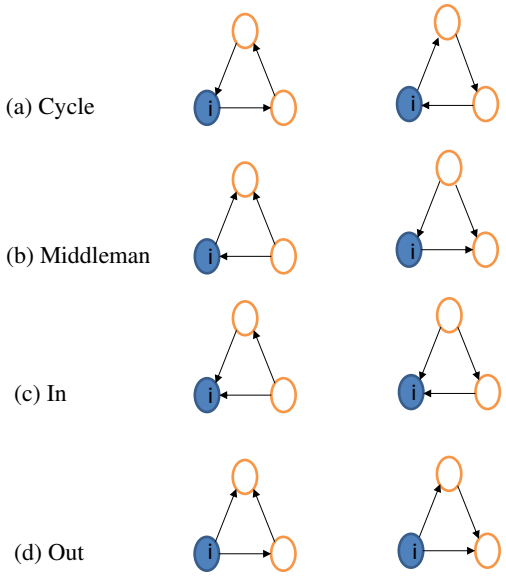


Figure 2: Directed triangles and the corresponding (binary) clusterings associated with a node i. (a) Cycle clustering, (b) Middleman clustering, (c) In clustering, (d) Out clustering.

Clustering coefficients also come in four different versions in the directed network. The different types are exhibited in Figure (2). The local binary clustering coefficient in a directed network associated with cyclical clustering is, for example, defined as:

$$C_{bin,i}^{cyc} = \frac{\sum_{j\neq i}\sum_{k\neq i,j} a_{ij}a_{jk}a_{ki}}{(\sum_{j\neq i} a_{ij}\sum_{j\neq i} a_{ji}) - (\sum_{j\neq i} a_{ij}a_{ji})}. \tag{17}$$

Note that, in the binary case we have $a_{ij}^2 = a_{ij}$ $(\forall i, j)$, so that

$$C_{bin,i}^{cyc} = \frac{\sum_{j\neq i}\sum_{k\neq i,j} a_{ij}a_{jk}a_{ki}}{k_i^{in}k_i^{out} - k_i^{\leftrightarrow}}, \tag{18}$$

and other types are obtained by appropriate variation of the indices, where $k_i^{\leftrightarrow}$ is the number of nodes j in the neighborhood of the node i such that $a_{ij} = a_{ji} = 1$.

The local clustering coefficient associated with cyclical combination in the directed weighted version is defined as

$$C_{w,i}^{cyc} = \frac{\sum_{j\neq i}\sum_{k\neq i,j} w_{ij}^{\frac{1}{3}}w_{jk}^{\frac{1}{3}}w_{ki}^{\frac{1}{3}}}{(\sum_{j\neq i} a_{ij}\sum_{j\neq i} a_{ji}) - (\sum_{j\neq i} a_{ij}a_{ji})} = \frac{\sum_{j\neq i}\sum_{k\neq i,j} w_{ij}^{\frac{1}{3}}w_{jk}^{\frac{1}{3}}w_{ki}^{\frac{1}{3}}}{k_i^{in}k_i^{out} - k_i^{\leftrightarrow}}. \tag{19}$$

2.3 Configuration models

In this subsection we will summarize the main ideas behind the algorithm involved in the extraction of hidden (latent) variables from an observed network and their role in the network randomization process (see, for example, Squartini and Garlaschelli, 2011; Squartini et al., 2011a; Squartini et al., 2011b; Mastrandrea et al., 2014; Squartini et al., 2015). For a more detailed explanation of the derivation of the family of Exponential Random Graph Model based on the maximum-entropy method, as well as on how to use the maximum-likelihood method to solve for the hidden variables under given constraints, we refer readers to the studies by Park and Newman (2004), Squartini and Garlaschelli (2011), and Squartini et al. (2015).

Undirected Binary Configuration model (UBCM)

In the UBCM, briefly, the entropy of a randomized ensemble of networks is maximized under the constraint that the node degrees in the observed network $\{k_i^{un}\}_{i=1}^n$ should match the averages of node degrees in the randomized ensemble. Mathematically, we need to solve the following system of n equations to obtain the non-negative hidden variables $\{x_i^*\}_{i=1}^n$ that carry the information from the constraints and allow us to perform an efficient unbiased

sampling of the ensemble

$$\sum_{j\neq i}\frac{x_i^*x_j^*}{1+x_i^*x_j^*}=k_i^{un},\forall i=1,2,...,n. \tag{20}$$

Once obtained, the hidden variables can be used to compute the probability p_{ij} of a link between any two nodes i and j, which in turn allows us to easily sample the ensemble associated with the above constraints

$$p_{ij}=\langle a_{ij}\rangle=\frac{x_i^*x_j^*}{1+x_i^*x_j^*}, \tag{21}$$

where $\langle a_{ij}\rangle$ is the notation for the expectation of a_{ij} over the ensemble.

Directed Binary Configuration model (DBCM)

In the DBCM, the constraints are the observed out-degree and in-degree sequences $\{k_i^{out}\}_{i=1}^n$ and $\{k_i^{in}\}_{i=1}^n$. We need to solve the following system of 2n equations to obtain the associated non-negative hidden variables $\{x_i^*\}_{i=1}^n$ and $\{y_i^*\}_{i=1}^n$

$$\begin{cases}\sum_{j\neq i}\frac{x_i^*y_j^*}{1+x_i^*y_j^*}=k_i^{out},\forall i=1,2,...,n,\\ \sum_{j\neq i}\frac{x_j^*y_i^*}{1+x_j^*y_i^*}=k_i^{in},\forall i=1,2,...,n.\end{cases} \tag{22}$$

The probability of a link from node i to j is given by

$$p_{ij}=\langle a_{ij}\rangle=\frac{x_i^*y_j^*}{1+x_i^*y_j^*}, \tag{23}$$

and the probability of a link from node j to i is given by

$$p_{ji}=\langle a_{ji}\rangle=\frac{x_j^*y_i^*}{1+x_j^*y_i^*}. \tag{24}$$

Undirected Weighted Configuration model (UWCM)

Similarly, suppose that in an undirected weighted network we want to extract n hidden variables $\{x_i^*\}_{i=1}^n$ associated with the observed strength sequence $\{s_i^{un}\}_{i=1}^n$, ($\{x_i^*\}_{i=1}^n\in[0,1)$). The maximum likelihood method involves solving the following system of n equations for the hidden variables

$$\sum_{j\neq i}\frac{x_i^*x_j^*}{1-x_i^*x_j^*}=s_i^{un},\forall i=1,2,...,n. \tag{25}$$

The expected link weight between node i and node j is given by

$$\langle w_{ij} \rangle = \frac{x_i^* x_j^*}{1 - x_i^* x_j^*}. \tag{26}$$

The probability of a link weight w_{ij} between node i and node j in the UWCM is

$$q(w_{ij}) = (p_{ij})^{w_{ij}}(1 - p_{ij}), \tag{27}$$

for $w_{ij} > 0$, where $p_{ij} = \langle a_{ij} \rangle$ is the probability of a link between two nodes (i, j), which is given by

$$p_{ij} = \langle a_{ij} \rangle = x_i^* x_j^*. \tag{28}$$

Directed Weighted Configuration model (DWCM)

In the DWCM, the constraints are the observed out-strength and in-strength sequences (i.e. $\{s_i^{out}\}_{i=1}^n$ and $\{s_i^{in}\}_{i=1}^n$). Mathematically, we need to solve the following system of 2n equations to obtain the hidden variables $\{x_i^*\}_{i=1}^n$ and $\{y_i^*\}_{i=1}^n \in [0, 1)$, which are respectively associated with $\{s_i^{out}\}_{i=1}^n$ and $\{s_i^{in}\}_{i=1}^n$

$$\begin{cases} \sum_{j \neq i} \frac{x_i^* y_j^*}{1 - x_i^* y_j^*} = s_i^{out}, \forall i = 1, 2, ..., n, \\ \sum_{j \neq i} \frac{x_j^* y_i^*}{1 - x_j^* y_i^*} = s_i^{in}, \forall i = 1, 2, ..., n. \end{cases} \tag{29}$$

The expected link weights between node i and node j are given by

$$\langle w_{ij} \rangle = \frac{x_i^* y_j^*}{1 - x_i^* y_j^*}, \tag{30}$$

and the probability of a link weight w_{ij} from node i to node j in the DWCM is the same as in the equation (27).

Undirected Enhanced Configuration model (UECM)

In the UECM, we use both the degree sequence $\{k_i^{un}\}_{i=1}^n$ as well as the strength sequence $\{s_i^{un}\}_{i=1}^n$ as constraints. The associated non-negative hidden variables $\{x_i^*\}_{i=1}^n$ and $\{y_i^*\}_{i=1}^n$ ($\{y_i^*\}_{i=1}^n \in [0, 1)$) are then the solution to the following system of 2n equations

$$\begin{cases} \sum_{j \neq i} \frac{x_i^* x_j^* y_i^* y_j^*}{1 - y_i^* y_j^* + x_i^* x_j^* y_i^* y_j^*} = k_i^{un}, \forall i = 1, 2, ..., n, \\ \sum_{j \neq i} \frac{x_i^* x_j^* y_i^* y_j^*}{(1 - y_i^* y_j^*)(1 - y_i^* y_j^* + x_i^* x_j^* y_i^* y_j^*)} = s_i^{un}, \forall i = 1, 2, ..., n. \end{cases} \tag{31}$$

It should be noted that, in the UECM, the probability of a link (i.e. $\langle a_{ij} \rangle$) and the

expected weight (i.e. $\langle w_{ij}\rangle$) between node i and node j depend on the information encoded in the strengths as well as in the degrees. More specifically, they are given by

$$p_{ij} = \langle a_{ij}\rangle = \frac{x_i^* x_j^* y_i^* y_j^*}{1 - y_i^* y_j^* + x_i^* x_j^* y_i^* y_j^*}, \tag{32}$$

and

$$\langle w_{ij}\rangle = \frac{x_i^* x_j^* y_i^* y_j^*}{(1 - y_i^* y_j^*)(1 - y_i^* y_j^* + x_i^* x_j^* y_i^* y_j^*)}. \tag{33}$$

In this model the probability of a link weight w_{ij} between two nodes (i, j) is given by

$$q(w_{ij}) = \begin{cases} 1 - p_{ij}, & \text{if } w_{ij} = 0 \\ p_{ij}(r_{ij})^{w_{ij}-1}(1 - r_{ij}), & \text{if } w_{ij} > 0, \end{cases} \tag{34}$$

where $r_{ij} = y_i^* y_j^*$, and p_{ij} is defined by Eq. (32).

Directed Enhanced Configuration model (DECM)

In the DECM, the non-negative hidden variables $\{x_i^*\}_{i=1}^n$, $\{y_i^*\}_{i=1}^n$, $\{z_i^*\}_{i=1}^n$, $\{t_i^*\}_{i=1}^n$ ($\{z_i^*\}_{i=1}^n$, $\{t_i^*\}_{i=1}^n \in [0, 1)$) extracted from the four sequences of constraints $\{k_i^{out}\}_{i=1}^n$, $\{k_i^{in}\}_{i=1}^n$, $\{s_i^{out}\}_{i=1}^n$, and $\{s_i^{in}\}_{i=1}^n$ are the solution to the following system of 4n equations

$$\begin{cases} \sum_{j\neq i} \frac{x_i^* y_j^* z_i^* t_j^*}{1 - z_i^* t_j^* + x_i^* y_j^* z_i^* t_j^*} = k_i^{out}, \forall i = 1, 2, ..., n, \\ \sum_{j\neq i} \frac{x_j^* y_i^* z_j^* t_i^*}{1 - z_j^* t_i^* + x_j^* y_i^* z_j^* t_i^*} = k_i^{in}, \forall i = 1, 2, ..., n, \\ \sum_{j\neq i} \frac{x_i^* y_j^* z_i^* t_j^*}{(1 - z_i^* t_j^*)(1 - z_i^* t_j^* + x_i^* y_j^* z_i^* t_j^*)} = s_i^{out}, \forall i = 1, 2, ..., n, \\ \sum_{j\neq i} \frac{x_j^* y_i^* z_j^* t_i^*}{(1 - z_j^* t_i^*)(1 - z_j^* t_i^* + x_j^* y_i^* z_j^* t_i^*)}) = s_i^{in}, \forall i = 1, 2, ..., n. \end{cases} \tag{35}$$

Similar to the UECM, in the DECM, the probability of a link (i.e. $\langle a_{ij}\rangle$) and the expected weight (i.e. $\langle w_{ij}\rangle$) from node i to node j depend on information encoded in the two sequences of observed degrees as well as in the two sequences of observed strengths. More specifically, we have

$$p_{ij} = \langle a_{ij}\rangle = \frac{x_i^* y_j^* z_i^* t_j^*}{1 - z_i^* t_j^* + x_i^* y_j^* z_i^* t_j^*}, \tag{36}$$

$$\langle w_{ij}\rangle = \frac{x_i^* y_j^* z_i^* t_j^*}{(1 - z_i^* t_j^*)(1 - z_i^* t_j^* + x_i^* y_j^* z_i^* t_j^*)}, \tag{37}$$

The probability of a link weight w_{ij} from node i to node j is still defined as in equation (34).

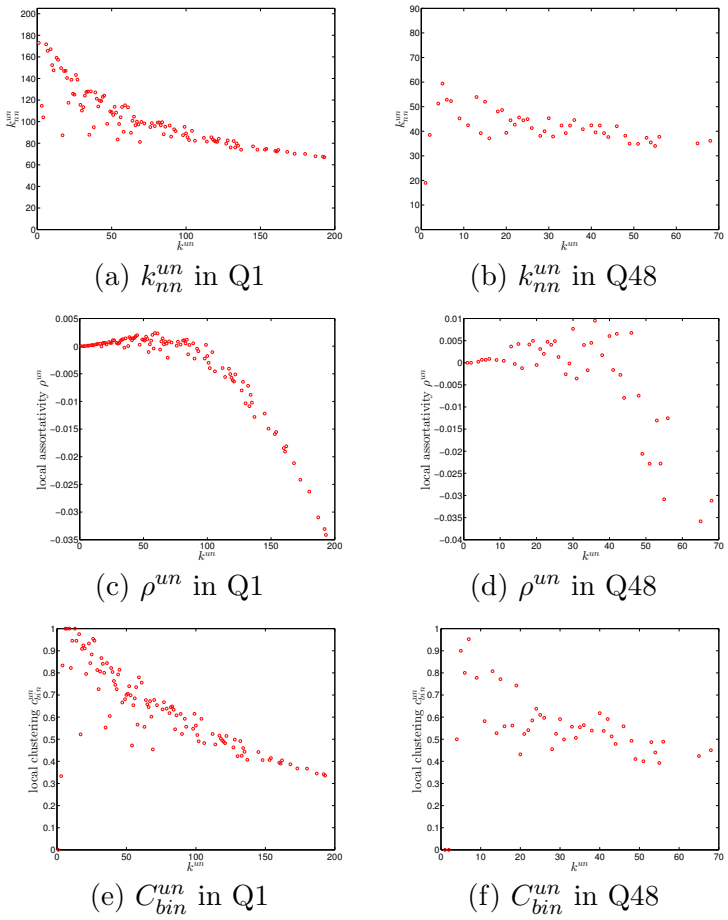
Note that, the expected values of the second and third structural correlations in the randomized networks can be analytically computed via the hidden variables extracted from

each configuration model or numerically computed by taking the average over a simulated ensemble. In our study, for each considered null model, we generate an ensemble of 1000 randomized networks, and then take the averages of the measures in question over the ensemble.

3 Findings for the binary network

3.1 Structural correlations in the undirected binary e-MID network

We first investigate the degree dependencies in the undirected binary e-MID network by examining the relationship between the node degree (k^{un}) and the average degree of its neighbors (k^{un}_{nn}). The overall disassortativity in this version of the network is evidenced by the negative relationship between these two quantities as shown in panels (a) and (b) of Figure (3), in which the measures for the networks from Q1 (the first quarter in our data set) and from Q48 (the last quarter in our data set) are plotted as an example. Note that the overall negative correlation between k^{un}_{nn} and k^{un} is also observed in all 48 quarters from 1999 to 2010. In addition, generally, we find that the absolute value of this correlation is declining over time.



(a) k_{nn}^{un} in Q1 (b) k_{nn}^{un} in Q48

(c) ρ^{un} in Q1 (d) ρ^{un} in Q48

(e) C_{bin}^{un} in Q1 (f) C_{bin}^{un} in Q48

Figure 3: ANND (panels a, b), local assortativity ρ^{un} (panels c, d), and local clustering coefficients C_{bin}^{un} (panels e, f) in the undirected binary e-MID network, in Q1 and Q48.

Next, we now turn to the Pearson correlation coefficient of degrees r_{bin}^{un} as an overall indicator of degree dependencies in the network. As shown in Figure (4), over time, overall, the network exhibits disassortativity as signaled by the negative coefficient. Consistent with what we discovered in our analysis of the measure ANND, the absolute value of r_{bin}^{un} is also declining from 1999 to 2010.

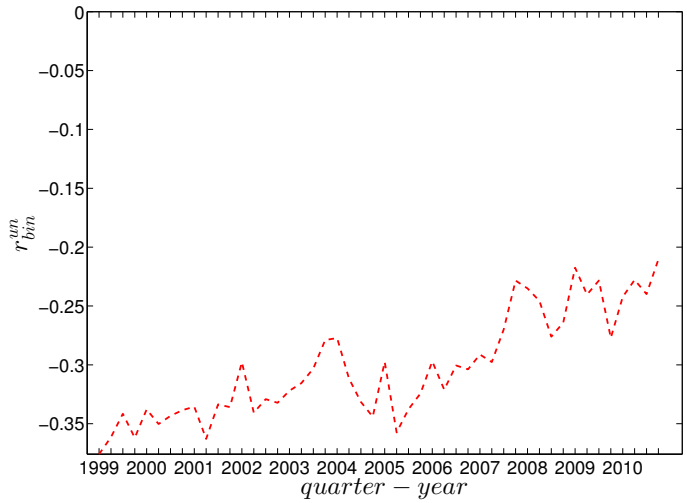


Figure 4: Evolution of the overall assortativity indicator r_{bin}^{un} in the undirected binary e-MID network.

For a more comprehensive assessment of the degree dependencies in the network, we employ the local assortativity coefficients ρ^{un} that indicate the contribution of each node to the global level of assortativity r_{bin}^{un} (see the Appendix for further details). The basic idea is that the numerator in the Pearson correlation coefficient proposed by Newman (2002, 2003a) can be reformulated based on the contribution of the individual nodes instead of in terms of the edges (see, for example, Piraveenan et al., 2010) leading to the decomposition

$$r_{bin}^{un} = \sum_{i=1}^{n} \rho_i^{un}. \tag{38}$$

In panels (c) and (d) of Figure (3) we plot ρ^{un} against k^{un} to investigate which nodes (in terms of their degrees) contribute most to r_{bin}^{un}. It appears that the hubs are the mainly contributors to the overall disassortativity of the network, while smaller degree nodes sometimes exhibit assortativity. This also reveals that adding or removing a hub from a network may have a large impact on its overall mixing nature.

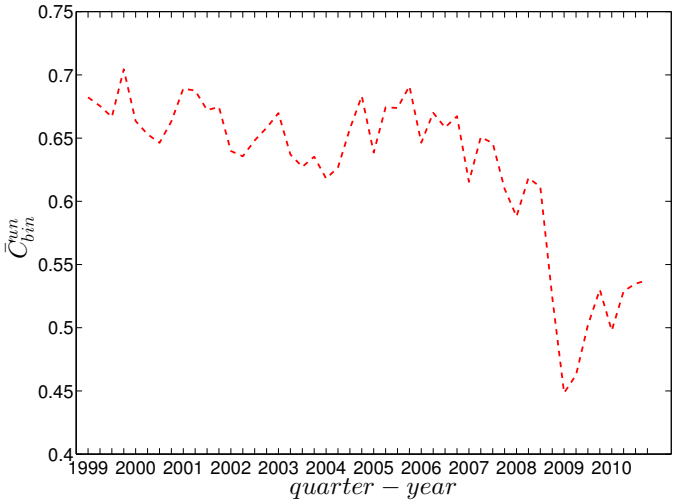


Figure 5: Evolution of the average of local clustering coefficients (i.e. $\bar{C}^{un}_{bin}$) in the undirected binary e-MID network.

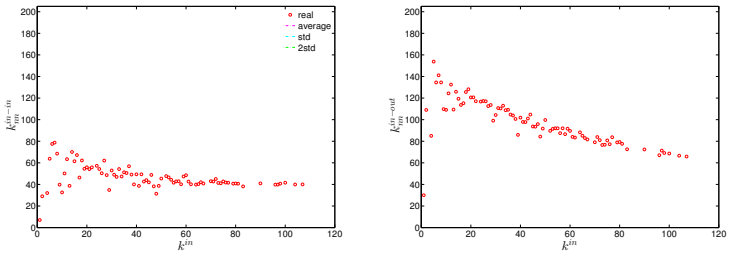
For the third order correlations, we employ the local clustering coefficient proposed by Watts and Strogatz (1998). In this simple version of the network (undirected binary case) clustering refers to the extent to which two connected nodes in the network have common neighbors. We observe that, overall, the undirected local clustering is a decreasing function of degree (panels (e) and (f) of Figure (3)), meaning that the neighbors of highly (poorly) connected banks are poorly (highly) interconnected. In fact, this relationship is typically found in many real world networks exhibiting a high heterogeneity in the degrees and a disassortative mixing nature (e.g. see Newman, 2003b). In our network, the bank degrees are highly heterogeneous, and the small (large) degree banks seem to have larger (smaller) local clustering coefficients because they are mostly connected to large (small) degree banks.

The evolution of the average of the undirected local binary clustering coefficients over all nodes is shown in Figure (5), where we can see a significant reduction in $\bar{C}^{un}_{bin}$ around the time of the financial crisis.

3.2 Structural correlations in the directed binary e-MID network

We now extend our analysis to the directed version of the binary network. Figures (6) and (7) show the relationship between ANND and node degree for the four types of mixing, i.e. in-in, in-out, out-in, and out-out. In the same network some types of mixing can be assortative, while others disassortative. For instance, while in Q1, overall, ANND is a decreasing function of the associated degree in all four cases, this relationship breaks down for the in-in and out-out mixing in Q48. In contrast, the overall negative correlation between ANND and the associated degree for the in-out and out-in mixing is observed in almost all quarters.

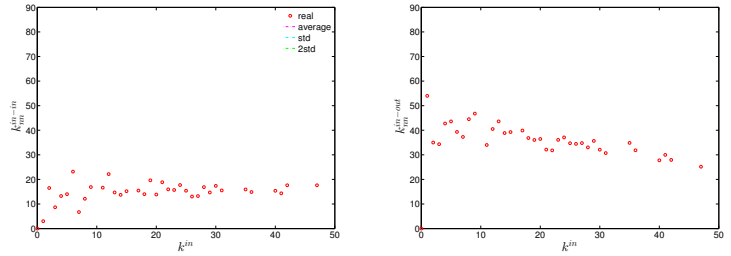
For a more general assessment of the overall mixing nature in the directed binary network, we calculate the Pearson correlation coefficient in each category of mixing and show its evolution over time (see Figure (8)). In comparison to the undirected version, the directed binary network displays more complicated degree dependencies. We can see that r_{bin}^{out-in} and r_{bin}^{in-out} display a different behavior than r_{bin}^{in-in} and $r_{bin}^{out-out}$. More specifically, while in the out-in and out-in categories we persistently observe disassortativity in all quarters, the other two categories switch between displaying assortativity and disassortativity over time. The interpretation of the mixing observed in the various categories is similar to the interpretation of mixing in an undirected binary network. For instance, a negative value of r_{bin}^{out-in}, signaling disassortativity in the out-in category, indicates that a high out-degree bank tends to have out-going links to low in-degree banks, and/or that a low out-degree bank tends to have out-going links to high in-degree banks. The mixing we observe in the out-in category (r_{bin}^{out-in}) comes closest to the one observed in the undirected network captured by r_{bin}^{un}. The similarity between these two quantities was mathematically proven by van der Hoorn and Litvak (2015). In addition, although the in-out mixing category exhibits disassortativity, in many quarters the coefficient r_{bin}^{in-out} is very close to zero.



(a) k_{nn}^{in-in} (b) k_{nn}^{in-out}

(c) k_{nn}^{out-in} (d) $k_{nn}^{out-out}$

Figure 6: ANND in the directed binary e-MID network, in Q1. k_{nn}^{in-in} (panel a), k_{nn}^{in-out} (panel b), k_{nn}^{out-in} (panel c), $k_{nn}^{out-out}$ (panel d).



(a) k_{nn}^{in-in} (b) k_{nn}^{in-out}

(c) k_{nn}^{out-in} (d) $k_{nn}^{out-out}$

Figure 7: ANND in the directed binary e-MID network, in Q48. k_{nn}^{in-in} (panel a), k_{nn}^{in-out} (panel b), k_{nn}^{out-in} (panel c), $k_{nn}^{out-out}$ (panel d).

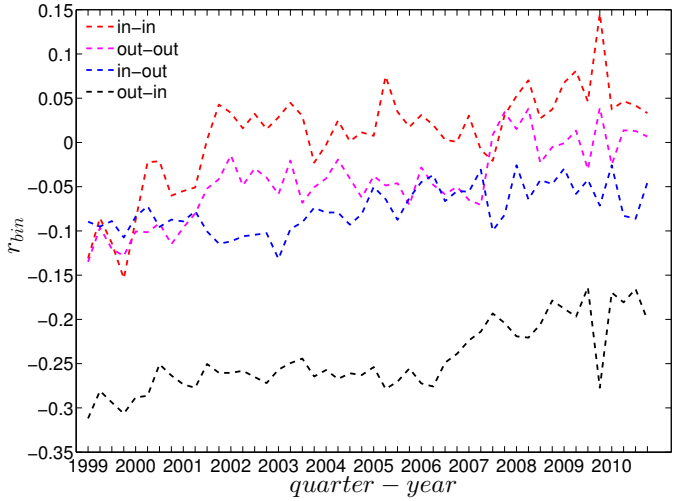
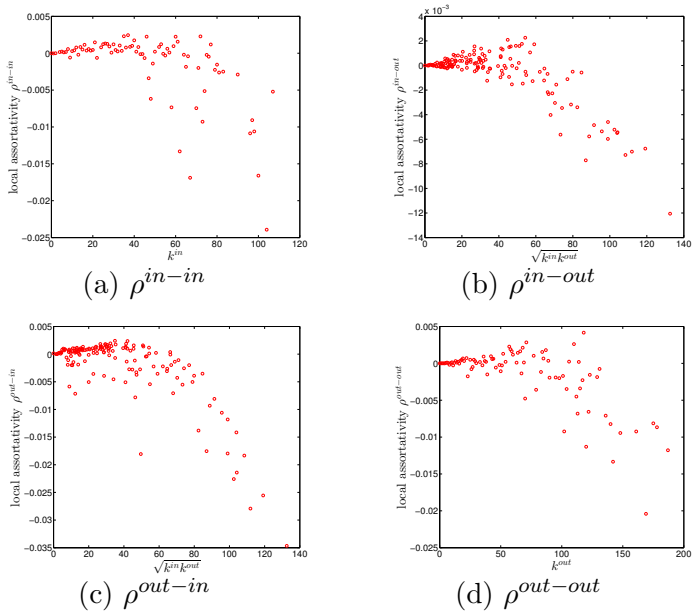


Figure 8: Evolution of the overall assortativity indicators in the directed binary e-MID network.

Similarly to the undirected case, we define the local assortativity measures for a given node i as ρ_i^{in-in}, ρ_i^{in-out}, ρ_i^{out-in}, and $\rho_i^{out-out}$ corresponding to the four mixing categories in the directed version of the network. Note that again equalities of the following form must hold:

$$r_{bin}^{in-in} = \sum_{i=1}^{n} \rho_i^{in-in}, \tag{39}$$

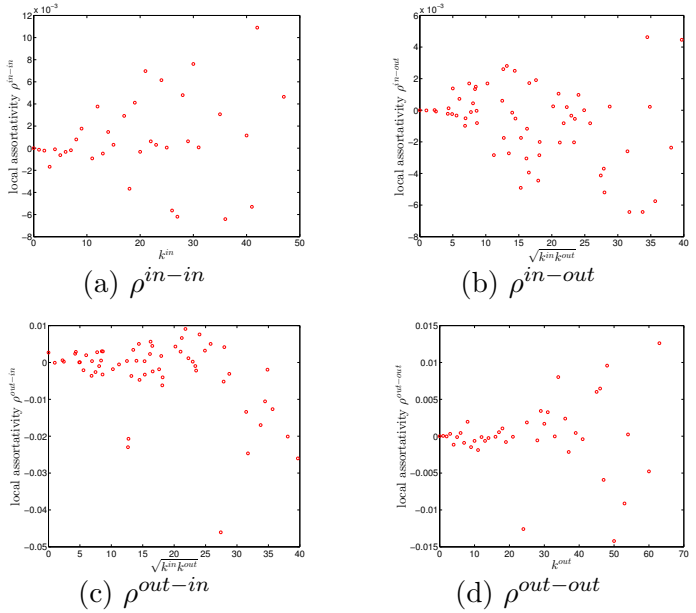
and analogously for the other measures. The measures ρ_i^{in-in}, ρ_i^{in-out}, ρ_i^{out-in}, and $\rho_i^{out-out}$ give us useful information about the contribution of each node to the respective overall assortativity indicators.



(a) ρ^{in-in} (b) ρ^{in-out}

(c) ρ^{out-in} (d) $\rho^{out-out}$

Figure 9: Local assortativity in the directed binary e-MID network, in Q1. ρ^{in-in} (panel a), ρ^{in-out} (panel b), ρ^{out-in} (panel c), $\rho^{out-out}$ (panel d).



(a) ρ^{in-in} (b) ρ^{in-out}

(c) ρ^{out-in} (d) $\rho^{out-out}$

Figure 10: Local assortativity in the directed binary e-MID network, in Q48. ρ^{in-in} (panel a), ρ^{in-out} (panel b), ρ^{out-in} (panel c), $\rho^{out-out}$ (panel d).

The local assortativity indicators in the two quarters Q1 and Q48 are respectively shown in Figures (9) and (10). In these figures, each local assortativity indicator is shown as a function of the corresponding degree [4]. The results indicate that, first, given an overall level

[4]In the cases of ρ^{in-out} and ρ^{out-in}, we plot them against $k^{in-out} = \sqrt{k^{in}k^{out}}$, since each of them

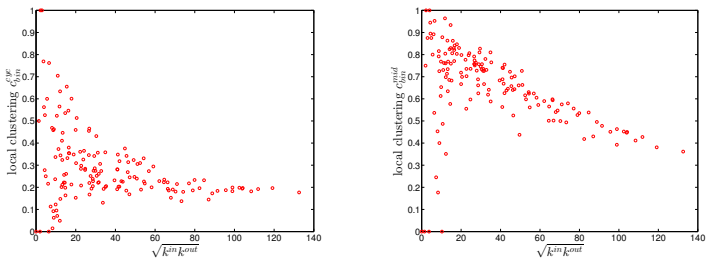
of assortativity in a particular category, the contribution of nodes of different degrees varies across the four mixing categories. In the out-in mixing category, we observe that, on the one hand, the hubs contribute most to the overall level of assortativity; on the other hand, small degree nodes are characterized by small values of assortativity or disassortativity. In addition, the contributions of medium degree nodes are more volatile than those of the small degree nodes. This is very similar to what we found in the undirected version of the network. However, the behavior of the local assortativity indicators becomes more complicated for the other mixing categories. For example, the contributions of hubs and medium degree nodes can fluctuate a lot, so that it becomes difficult to classify which type of nodes plays an important role for the overall level of assortativity.

Next, we turn to the third order correlations between banks in the directed binary network. We focus on investigating local clustering as a function of degree for the four cases shown in Figure (2) (see, for example, Fagiolo, 2007). In the following discussion we will be referring to nodes i, j, k as an example of three vertices in a network building a triangle. It is clear that the directions of the edges now matter for the clustering analysis. The measures C^{mid}, C^{cyc}, C^{out} and C^{in} summarize the prevalence of a particular type of relationship that a node has with its neighbors. For instance, larger values of C^{mid} (see panel (b) of Figure (2)) may represent a higher systemic risk associated with that node, since bank i can be a source of risk as well as be exposed to risk from other banks. Clustering relationships of the type shown in panel (c) of Figure (2) are also conducive to systemic risk since a default of bank i would affect both its partners. Larger values of C^{in} indicate a higher systemic risk dependence on the same funding sources in the interbank market. This is, however, not the case for cyclical clustering relationships (captured by C^{cyc}) since in this type of relationships exposures can cancel each other out (see panel (a) of Figure (2)). Finally, large values of C^{out} associated with bank i indicate risk exposure of bank i itself, since both banks j and k can affect bank i in case either of them would default (see panel (d) of Figure (2).

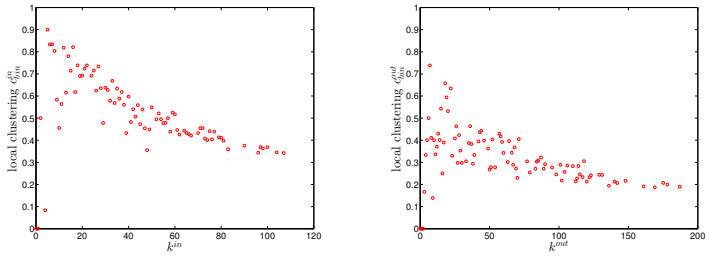
For each type of clustering relationship, we first consider the local clustering coefficient as the function of the corresponding degree [5]. Typically, in each case, a general negative relationship is observed in the first quarters, but for later quarters this relationship becomes flatter (see, for example, Figures (11) and (12)).

depends on both k^{in} and k^{out} (see the Appendix for more detailed derivations).

[5]In the cases of C^{cyc}_{bin} and C^{mid}_{bin}, we plot them against $k^{in-out} = \sqrt{k^{in}k^{out}}$.

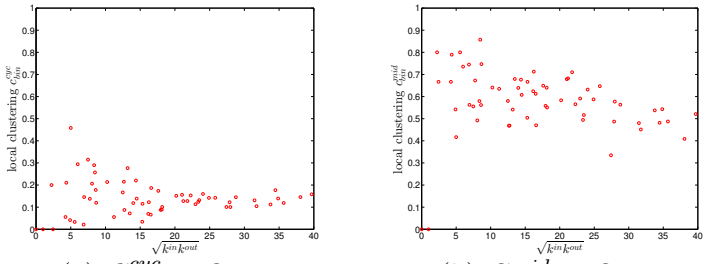


(a) C_{bin}^{cyc} in Q1 (b) C_{bin}^{mid} in Q1

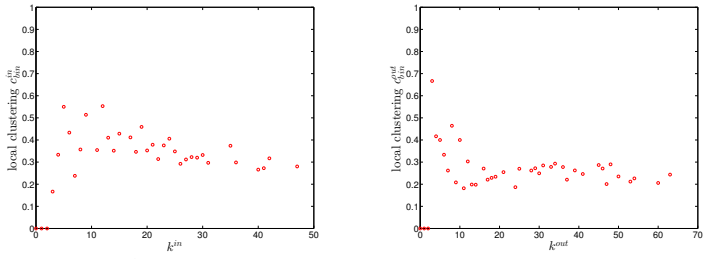


(c) C_{bin}^{in} in Q1 (d) C_{bin}^{out} in Q1

Figure 11: Local clustering coefficients C_{bin}^{cyc} (panel a), C_{bin}^{mid} (panel b), C_{bin}^{in} (panel c), C_{bin}^{out} (panel d) in the directed binary e-MID network, in Q1.



(a) C_{bin}^{cyc} in Q48 (b) C_{bin}^{mid} in Q48



(c) C_{bin}^{in} in Q48 (d) C_{bin}^{out} in Q48

Figure 12: Local clustering coefficients C_{bin}^{cyc} (panel a), C_{bin}^{mid} (panel b), C_{bin}^{in} (panel c), C_{bin}^{out} (panel d) in the directed binary e-MID network, in Q48.

We now take the averages of the local clustering coefficients across all nodes and then investigate their evolution over time. We observe that, first, for the most part, the averages $\bar{C}^{mid}$, $\bar{C}^{in}$, $\bar{C}^{out}$, and $\bar{C}^{cyc}$ are in descending order, with clustering relationships of the cyclical

and out-type being much less common than the other two. We consider this prevalence of the middleman and in-type clustering relationships as evidence of the presence of systemic risk in the network. Second, similarly to what we observed in the undirected network for $\bar{C}^{un}_{bin}$, the averages of the local clustering coefficients for all four clustering types dramatically decrease around the time of the financial crisis, evidencing structural change in the third order correlations between banks.

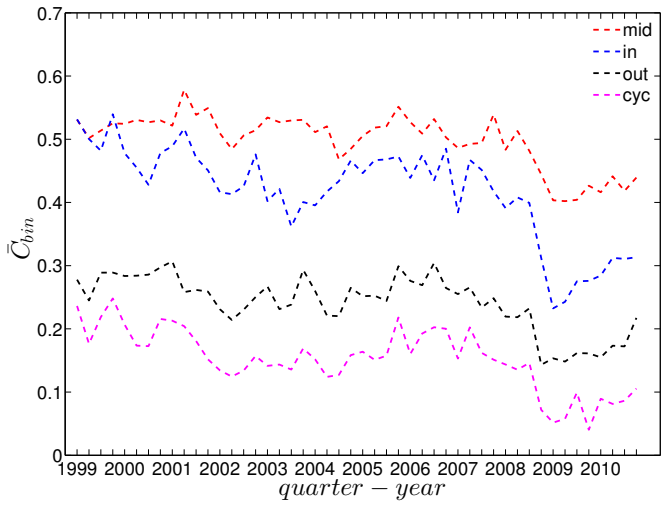


Figure 13: Evolution of the averages of local clustering coefficients (i.e. $\bar{C}^{mid}$, $\bar{C}^{in}$, $\bar{C}^{out}$, and $\bar{C}^{cyc}$) in the directed binary e-MID network.

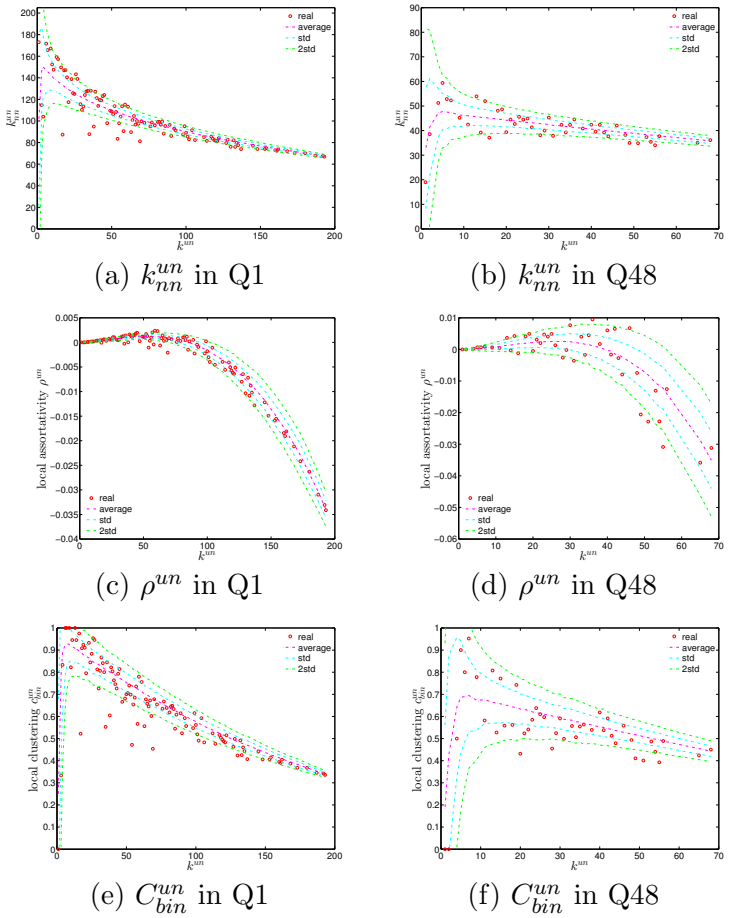
3.3 Comparisons to the configuration models

When comparing the phenomenological properties of the data to those of the configuration models, we basically ask the question whether the higher-level characteristics are the mere consequence of the observed features of lower order. Features that could not be accounted for by the configuration model would indicate facets of the data that need additional behavioral explanations.

Undirected Binary Network

We first employ the undirected binary configuration model (UBCM), which maintains the intrinsic heterogeneity in the degree sequence of the undirected binary version of the observed e-MID network. Figure (14) shows a comparison between various higher order structural correlations observed in the e-MID network and the same structural correlations observed in the randomized ensemble for the first and last quarters. Note that, in each panel of Figure (14), besides the observed and the expected values (over the randomized

ensemble), we also report the regions of ±1 standard deviation (std.) and ± 2 std. away from the expectations. In most cases, as shown in panels (a) to (f), the local behavior of the structural correlations is well replicated by the UBCM. As shown in panel (a) of Figure (15), the average of the ANNDs over all nodes ($\bar{k}^{un}_{nn}$) is also located inside the ±2 std. band when plotted over time. In contrast, in terms of our measure of global assortativity (r^{un}_{bin}), in almost all of the quarters, the observed values lie outside the ±2 std. band (see panel (b) of Figure (15)). A similar result is obtained for the evolution of the average of the local clustering coefficients ($\bar{C}^{un}_{bin}$) with many significant deviations, but the main trends of the observed and the expected values are similar (see panel (c) of Figure (15)).

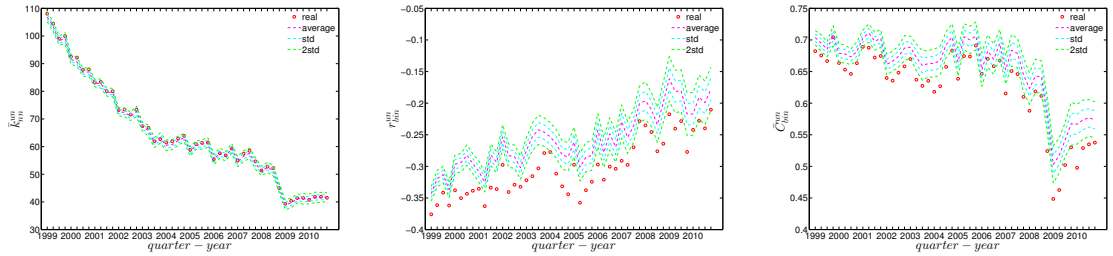


(a) k^{un}_{nn} in Q1 (b) k^{un}_{nn} in Q48

(c) ρ^{un} in Q1 (d) ρ^{un} in Q48

(e) C^{un}_{bin} in Q1 (f) C^{un}_{bin} in Q48

Figure 14: ANND (panels a, b), local assortativity ρ^{un} (panels c, d), local clustering coefficients C^{un}_{bin} (panels e, f) in the observed e-MID network and in the UBCM, in Q1 and Q48.



(a) $\bar{k}^{un}_{nn}$ (b) r^{un}_{bin} (c) $\bar{C}^{un}_{in}$

Figure 15: Evolution of $\bar{k}^{un}_{nn}$ (panel a), r^{un}_{bin} (panel b), and $\bar{C}^{un}_{in}$ (panel c) in the observed e-MID network and in the UBCM.

Directed Binary Network

Recalling that under the directed binary configuration model (DBCM), both out-going and in-coming degrees are enforced on average over the ensemble, we show the comparisons between the structural correlations of observed network and those obtained from that model in Figures (16) and (17) (for ANND), Figures (18) and (19) (for the local assortativity indicators), and Figures (20) and (21) (for the local clustering coefficients).

In addition, as for the undirected version, we also compare the evolution of the global indicators with the evolution of their expected values obtained from the DBCM. We show the results for the averages of the ANNDs (i.e. $\bar{k}^{in-in}_{nn}$, $\bar{k}^{in-out}_{nn}$, $\bar{k}^{out-in}_{nn}$, $\bar{k}^{out-out}_{nn}$) in Figure (22), for the global assortativity indicators (r^{in-in}_{bin}, r^{in-out}_{bin}, r^{out-in}_{bin}, $r^{out-out}_{bin}$) in Figure (23), and for the averages of the local clustering coefficients ($\bar{C}^{in-in}_{bin}$, $\bar{C}^{in-out}_{bin}$, $\bar{C}^{out-in}_{bin}$, $\bar{C}^{out-out}_{bin}$) in Figure (24).

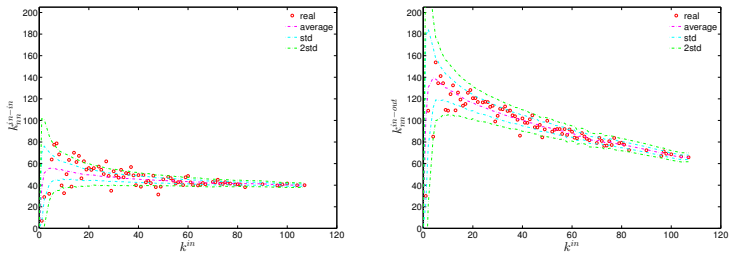
First, regarding the local indicators (see from Figure (16) to Figure (21)), in most cases, the observed ANNDs, local assortativity indicators, and local clustering coefficients are in agreement with those evaluated under the DBCM. Since the few observed points significantly deviating from the expected ones might not reveal any patterns (under the DBCM), they might be seen as the expected rejections one obtains for a large sequence of simultaneous tests.

Second, regarding the evolution of the averages of the ANNDs, Figure (22) shows that, $\bar{k}^{in-in}_{nn}$, $\bar{k}^{in-out}_{nn}$, and $\bar{k}^{out-in}_{nn}$ always lie within the ± 2 std. band, while $\bar{k}^{out-out}_{nn}$ is underestimated for most of the time.

Third, in terms of the global assortativity indicators, for the most part, r^{in-in}_{bin} and r^{out-in}_{bin} are located inside the $\pm$ 2 std. band, while r^{in-out}_{bin} and $r^{out-out}_{bin}$ are mostly being overestimated (see Figure (23)).

Finally, over time, the averages of the local clustering coefficients $\bar{C}^{in-in}_{bin}$, $\bar{C}^{in-out}_{bin}$, $\bar{C}^{out-in}_{bin}$,

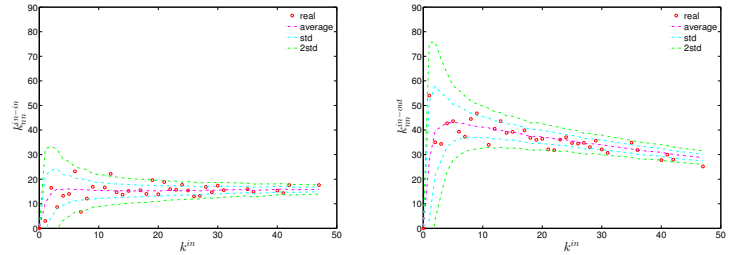
and $\bar{C}_{bin}^{out-out}$ are generally in agreement with their expected values from the DBCM, as shown in Figure (24).



(a) k_{nn}^{in-in} (b) k_{nn}^{in-out}

(c) k_{nn}^{out-in} (d) $k_{nn}^{out-out}$

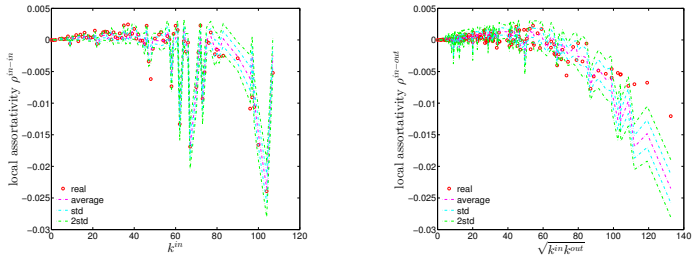
Figure 16: ANND in the observed e-MID network and in the DBCM, in Q1. k_{nn}^{in-in} (panel a), k_{nn}^{in-out} (panel b), k_{nn}^{out-in} (panel c), $k_{nn}^{out-out}$ (panel d).



(a) k_{nn}^{in-in} (b) k_{nn}^{in-out}

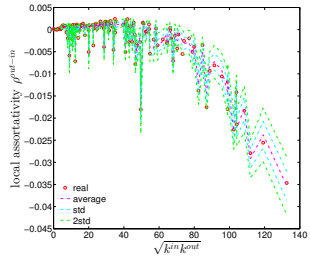
(c) k_{nn}^{out-in} (d) $k_{nn}^{out-out}$

Figure 17: ANND in the observed e-MID network and in the DBCM, in Q48. k_{nn}^{in-in} (panel a), k_{nn}^{in-out} (panel b), k_{nn}^{out-in} (panel c), $k_{nn}^{out-out}$ (panel d).

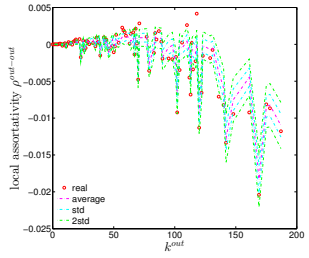


(a) ρ^{in-in}

(b) ρ^{in-out}

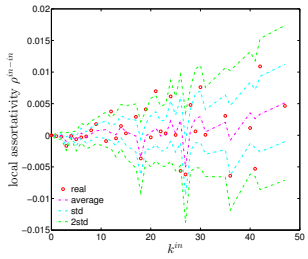


(c) ρ^{out-in}

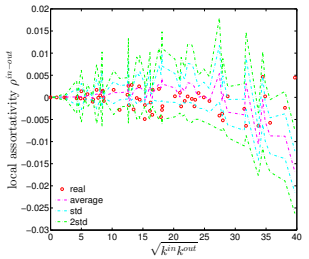


(d) $\rho^{out-out}$

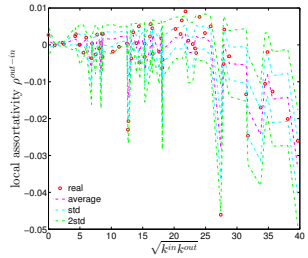
Figure 18: Local assortativity in the observed e-MID network and in the DBCM, in Q1. ρ^{in-in} (panel a), ρ^{in-out} (panel b), ρ^{out-in} (panel c), $\rho^{out-out}$ (panel d).



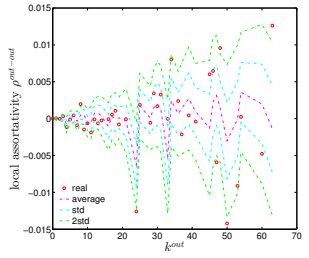
(a) ρ^{in-in}



(b) ρ^{in-out}

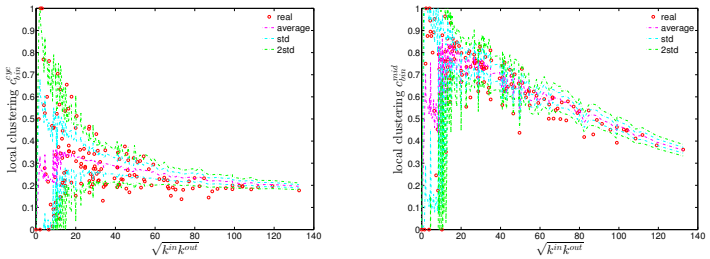


(c) ρ^{out-in}



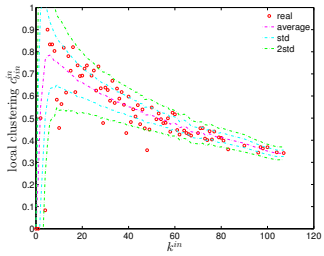
(d) $\rho^{out-out}$

Figure 19: Local assortativity in the observed e-MID network and in the DBCM, in Q48. ρ^{in-in} (panel a), ρ^{in-out} (panel b), ρ^{out-in} (panel c), $\rho^{out-out}$ (panel d).

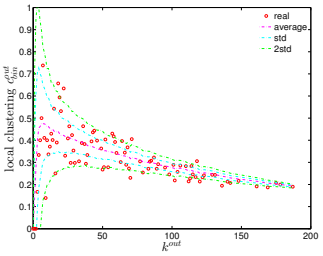


(a) C^{cyc}_{bin} in Q1

(b) C^{mid}_{bin} in Q1

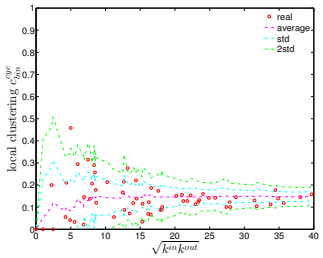


(c) C^{in}_{bin} in Q1

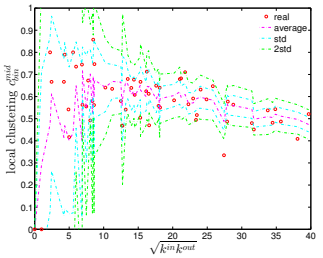


(d) C^{out}_{bin} in Q1

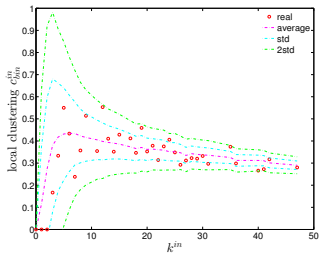
Figure 20: Local clustering coefficients C^{cyc}_{bin} (panel a), C^{mid}_{bin} (panel b), C^{in}_{bin} (panel c), C^{out}_{bin} (panel d) in the observed e-MID network and in DBCM, in Q1.



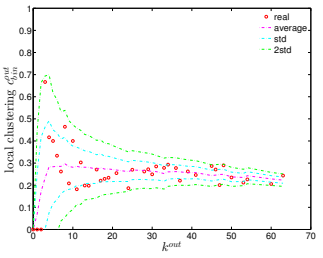
(a) C^{cyc}_{bin} in Q48



(b) C^{mid}_{bin} in Q48

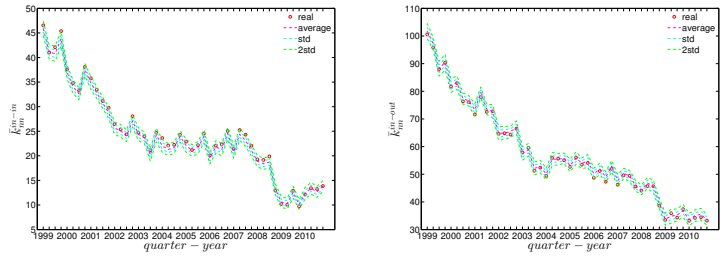


(c) C^{in}_{bin} in Q48

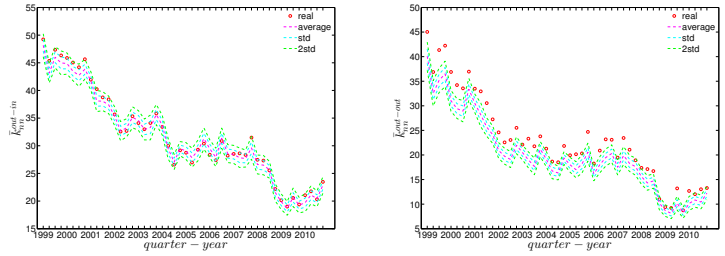


(d) C^{out}_{bin} in Q48

Figure 21: Local clustering coefficients C^{cyc}_{bin} (panel a), C^{mid}_{bin} (panel b), C^{in}_{bin} (panel c), C^{out}_{bin} (panel d) in the observed e-MID network and in DBCM, in Q48.

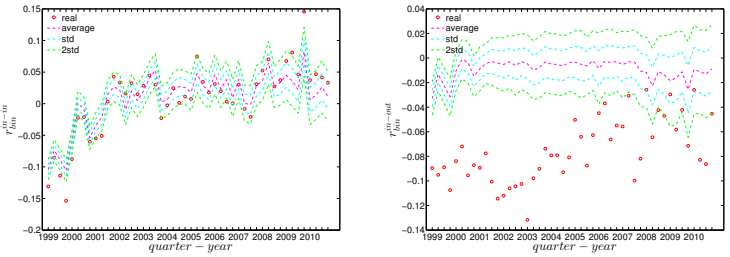


(a) $\bar{k}_{nn}^{in-in}$ (b) $\bar{k}_{nn}^{in-out}$

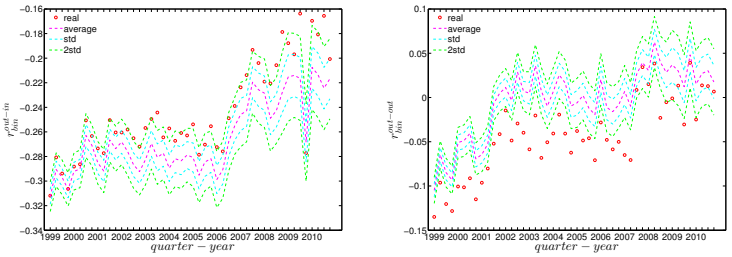


(c) $\bar{k}_{nn}^{out-in}$ (d) $\bar{k}_{nn}^{out-out}$

Figure 22: Evolution of the averages of ANNDs in the observed e-MID network and in the DBCM. $\bar{k}_{nn}^{in-in}$ (panel a), $\bar{k}_{nn}^{in-out}$ (panel b), $\bar{k}_{nn}^{out-in}$ (panel c), $\bar{k}_{nn}^{out-out}$ (panel d).

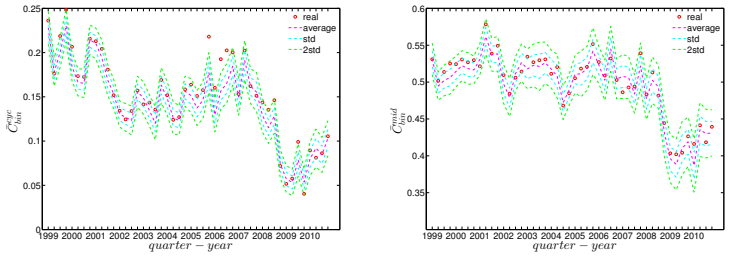


(a) r_{bin}^{in-in} (b) r_{bin}^{in-out}

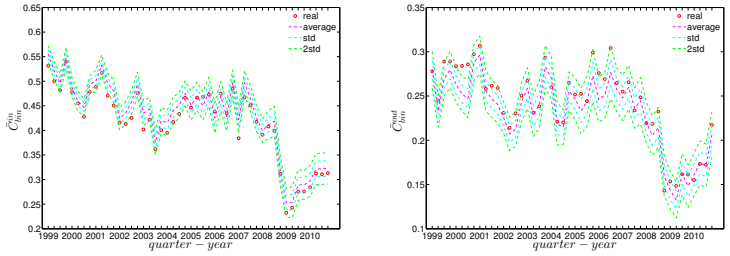


(c) r_{bin}^{out-in} (d) $r_{bin}^{out-out}$

Figure 23: Evolution of the global assortativity indicators in the observed e-MID network and in the DBCM. r_{bin}^{in-in} (panel a), r_{bin}^{in-out} (panel b), r_{bin}^{out-in} (panel c), $r_{bin}^{out-out}$ (panel d).



(a) $\bar{C}^{cyc}_{bin}$ (b) $\bar{C}^{mid}_{bin}$



(c) $\bar{C}^{in}_{bin}$ (d) $\bar{C}^{out}_{bin}$

Figure 24: Evolution of the averages of clustering coefficients in the observed e-MID network and in the DBCM. $\bar{C}^{cyc}_{bin}$ (panel a), $\bar{C}^{mid}_{bin}$ (panel b), $\bar{C}^{in}_{bin}$ (panel c), $\bar{C}^{out}_{bin}$ (panel d).

4 Findings for the weighted network

In the binary version of the observed network, we treat all edges as if they were homogeneous. However, in reality, the capacity and intensity of the relations between banks can be very heterogeneous, consequently the weighted version can have different properties compared to its binary counterpart. In this section, we investigate the structural correlations in the weighted e-MID network. For the sake of simplicity, we do not consider the local weighted assortativity in this section, since breaking down the overall weighted assortativity measure into the contributions of the individual nodes is much more complicated than in the binary case.

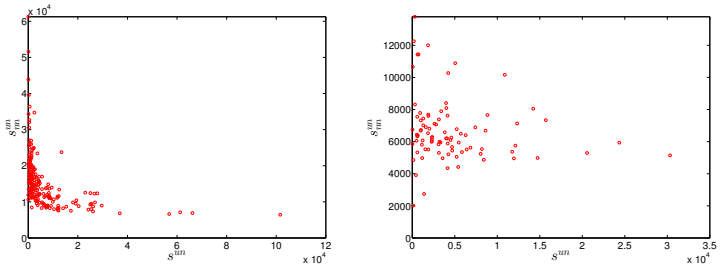
Regarding the null models, instead of preserving the observed degree sequence(s) as in the Binary Configure Models (i.e. UBCM, DBCM), first, we employ the weighted configuration model preserving the observed strength sequence(s) (i.e. the UWCM model in the undirected case and the DWCM model in the directed case) and examine whether the chosen null models can replicate the structural correlations in the observed weighted network. As a second step, we consider the enhanced configuration models which maintain both the observed degree as well as strength sequences (i.e. the UECM and the DECM respectively in the undirected and directed cases) and repeat the same exercise.

4.1 Structural correlations in the undirected weighted e-MID network

We report the strength dependencies in Figure (25) by considering the relationship between s_{nn}^{un} (ANNS) and s^{un} in the first and last quarters. We observe that s_{nn}^{un} is generally a declining function of s^{un}, although this feature becomes less pronounced towards the end of our time series. This relationship is confirmed by the negative value of the global weighted assortativity measure r_w^{un} (see Figure (26)). This signals that the prevalence of disassortative mixing in the undirected weighted e-MID network does not only apply to the degrees, but also to the strengths of nodes. Furthermore, it should be emphasized that, in comparison to the undirected binary version of the network, the undirected weighted network exhibits less disassortativity overall, since r_w^{un} is smaller than r_{bin}^{un} in absolute value.

In our analysis of the third order correlations, in contrast to what we discovered in the binary version, we find that, on average banks with higher strength also have higher local clustering coefficients (see Figure (27)). This is mainly because the heterogeneity in the transaction volumes across banks in every triangle is now taken into account and the

average transactions of banks with high strength are much larger than those of banks with low strength. Furthermore, we observe three very distinct phases in the evolution of the average of the local weighted clustering coefficients, i.e. before 2002, from 2002 to 2006, and from 2007 onward, which might reflect effects arising from the adoption of the euro as well as from the recent financial crisis (see Figure (28)). In particular, we find that $\bar{C}_w^{un}$ is much higher from 2002 to 2006 than in the years before and after. The same results still hold if we normalize all weights by the total weight average.



(a) s_{nn}^{un} in Q1 (b) s_{nn}^{un} in Q48

Figure 25: ANNS in the undirected weighted e-MID network, in Q1 and Q48.

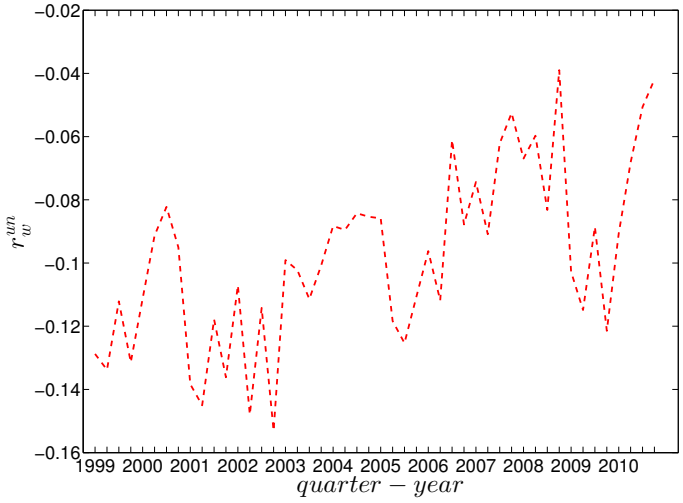
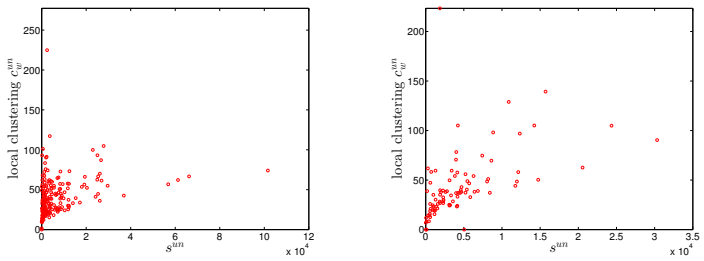


Figure 26: Evolution of global weighted assortativity r_w^{un} in the undirected weighted e-MID network.



(a) C_w^{un} in Q1 (b) C_w^{un} in Q48

Figure 27: Local clustering coefficients C_w^{un} in the undirected weighted e-MID network, in Q1 and Q48.

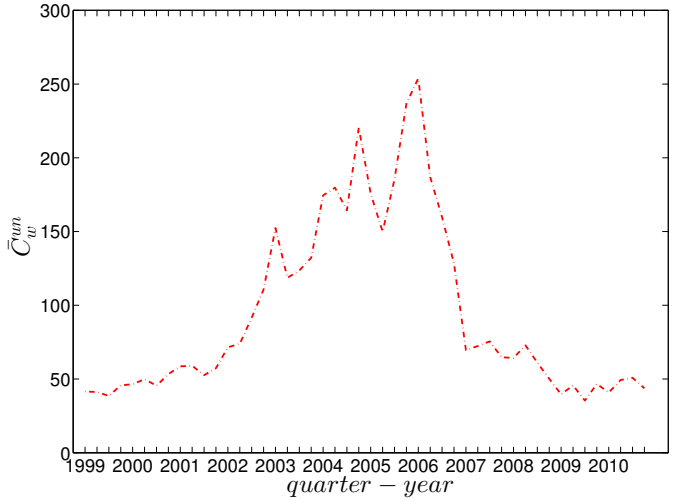


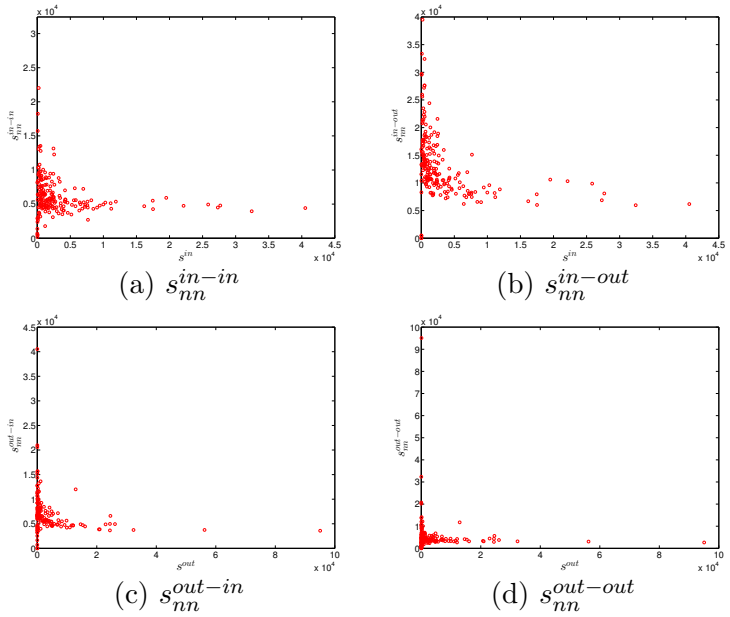
Figure 28: Evolution of the average of local weighted clustering coefficients (i.e. $\bar{C}_w^{un}$) in the undirected weighted e-MID network.

4.2 Structural correlations in the directed weighted e-MID network

In the directed weighted version of the e-MID network, to analyzing the structural correlations, we employ average nearest neighbor strength measures for the various mixing categories ($s_{nn,i}^{in-in}, s_{nn,i}^{in-out}, s_{nn,i}^{out-in}, s_{nn,i}^{out-out}$), global weighted assortativity indicators (r_w^{in-in}, r_w^{in-out}, r_w^{out-in}, $r_w^{out-out}$), and weighted clustering coefficients ($C_w^{cyc}, C_w^{mid}, C_w^{in}, C_w^{out}$).

First, Figures (29) and (30) show the relationship between the ANNSs and the associated strengths for all four mixing categories in Q1 and Q48. Over time, while in the first quarters, the ANNSs are a declining function of the associated strengths, in many later quarters this relationship again seems to break down, especially for the mixing categories in-in, in-out, and

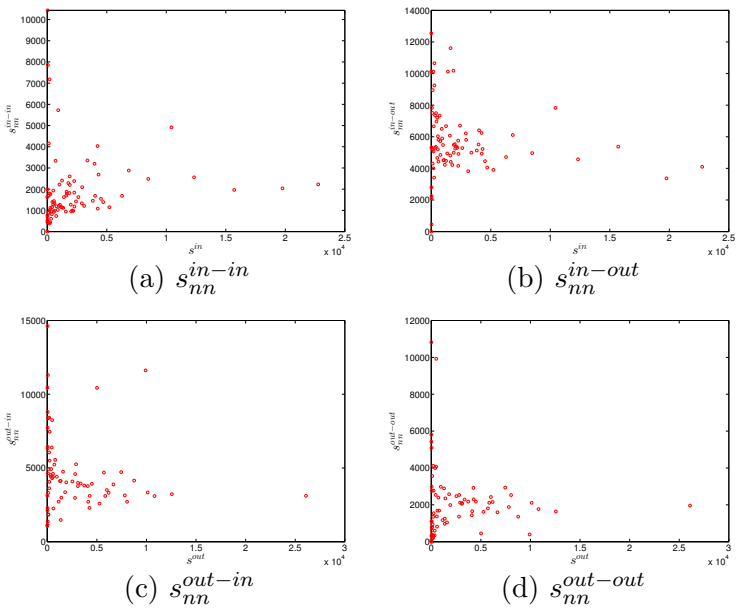
out-out. To obtain the overall level of strength dependency of bank interactions for each mixing category, we calculate the global assortativity indicators $r_w^{in-in}, r_w^{in-out}, r_w^{out-in}, r_w^{out-out}$, and show their evolution over time in Figure (31). The results indicate that, while the out-in mixing is disassortative for the most part, the other three categories do not seem to exhibit a distinct mixing nature. In comparison to the directed binary version, the absolute values of $r_w^{in-in}, r_w^{in-out}, r_w^{out-in}, r_w^{out-out}$ are often smaller than those of $r_{bin}^{in-in}, r_{bin}^{in-out}, r_{bin}^{out-in}, r_{bin}^{out-out}$. An interesting observation is that, among the four mixing categories, the weighted assortativity in the out-in category is closest to the undirected weighted assortativity, i.e. $r_w^{out-in} \sim r_w^{un}$. For the binary versions of the network, when comparing the mixing patterns in the directed and undirected case, we made the same observation.



(a) s_{nn}^{in-in} (b) s_{nn}^{in-out}

(c) s_{nn}^{out-in} (d) $s_{nn}^{out-out}$

Figure 29: ANNSs in the directed weighted e-MID network, in Q1. s_{nn}^{in-in} (panel a), s_{nn}^{in-out} (panel b), s_{nn}^{out-in} (panel c), $s_{nn}^{out-out}$ (panel d).



(a) s_{nn}^{in-in} (b) s_{nn}^{in-out}

(c) s_{nn}^{out-in} (d) $s_{nn}^{out-out}$

Figure 30: ANNSs in the directed weighted e-MID network, in Q48. s_{nn}^{in-in} (panel a), s_{nn}^{in-out} (panel b), s_{nn}^{out-in} (panel c), $s_{nn}^{out-out}$ (panel d).

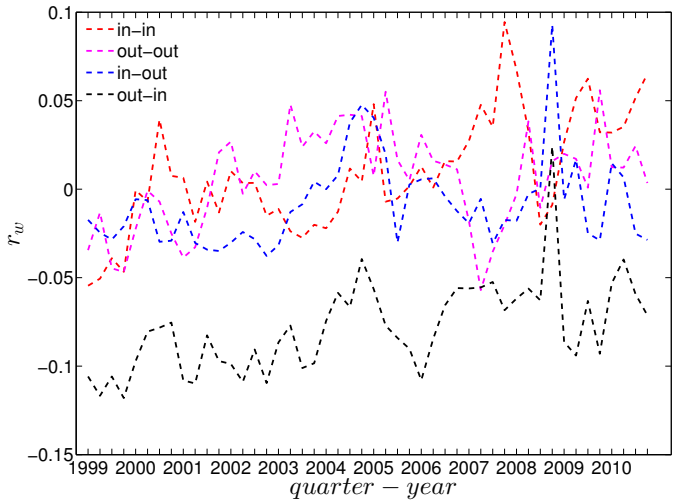
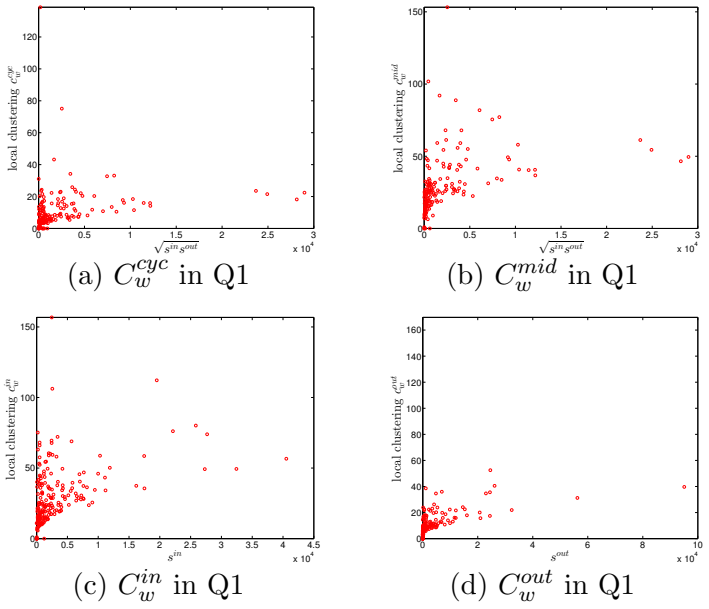


Figure 31: Evolution of the directed weighted assortativity indicators, i.e. r_w^{in-in}, r_w^{in-out}, r_w^{out-in}, and $r_w^{out-out}$ in the directed weighted e-MID network.

Second, the local weighted clustering coefficients for the four clustering types C_w^{cyc}, C_w^{mid}, C_w^{in}, C_w^{out} are plotted against the associated strengths in Figures (32) and (33) [6]. We observe that, generally, higher (lower) strengths correspond to higher (lower) local weighted clustering coefficients.

[6] In the cases of C_w^{cyc} and C_w^{mid}, we plot them against $s^{in-out} = \sqrt{s^{in} s^{out}}$.

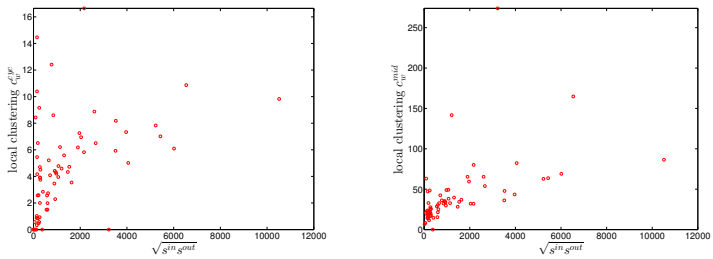
The evolution of the averages of the local weighted clustering coefficients also exhibits three different phases, i.e. before 2002, from 2002 to 2006, and from 2007 onward. For all types of clustering, the averages in the period from 2002 to 2006 are higher than those in the other two periods. Recall that, on average, larger values of $\bar{C}_w^{mid}$, $\bar{C}_w^{in}$ imply higher risk from concentrated funding lines, while larger values of $\bar{C}_w^{out}$ reveal the high exposure of the associated bank to risk from defaults of their borrowers. The order and magnitude of different combinations of $\bar{C}_w$ shown in Figure (34) thus reveal the importance of both types of risk in the period from 2002 to 2006 in the weighted version of the network. It should be emphasized that, even when all weights are normalized by the average weight over the whole network, we still observe a similar trend, signaling that the evolution of the averages of the directed local weighted clustering coefficients is not only driven by changes in the overall transaction volume (overall strength of the interactions) but also by changes in the frequency of aforementioned tripartite relations among banks.



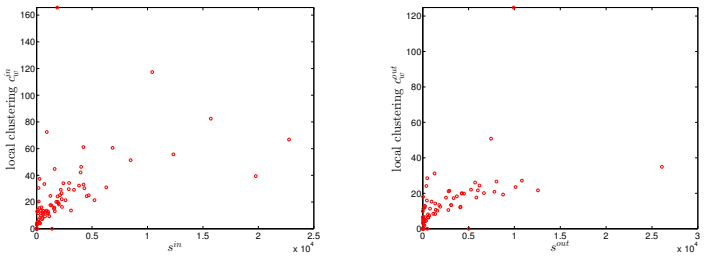
(a) C_w^{cyc} in Q1 (b) C_w^{mid} in Q1

(c) C_w^{in} in Q1 (d) C_w^{out} in Q1

Figure 32: Local weighted clustering coefficients in the directed weighted e-MID network, in Q1. C_w^{cyc} (panel a), C_w^{mid} (panel b), C_w^{in} (panel c), C_w^{out} (panel d).



(a) C_w^{cyc} in Q48 (b) C_w^{mid} in Q48



(c) C_w^{in} in Q48 (d) C_w^{out} in Q48

Figure 33: Local weighted clustering coefficients in the directed weighted e-MID network, in Q48. C_w^{cyc} (panel a), C_w^{mid} (panel b), C_w^{in} (panel c), C_w^{out} (panel d).

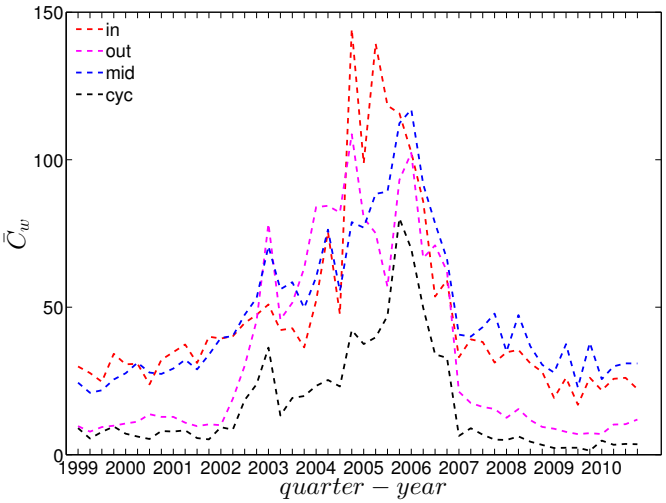


Figure 34: Evolution of the averages of local weighted clustering coefficients, i.e. $\bar{C}_w^{cyc}$, $\bar{C}_w^{mid}$, $\bar{C}_w^{in}$, and $\bar{C}_w^{out}$ in the directed weighted e-MID network.

4.3 Comparisons to the weighted configuration models

Undirected Weighted Network

To examine the role of the heterogeneity in the local constraints for the emergence of higher order structural correlations in the weighted version of the observed network, we employ the UWCM, which preserves the observed strength sequence, and the UECM, which enforces both the observed degree as well as strength sequences onto the randomized ensemble.

First, the observed values of the measure ANNS as well as of the local weighted clustering coefficients (as can be seen in Figures (35) and (36)) strongly deviate from their respective expectations under the UWCM. In contrast, we find that the UECM model is able to reproduce the main features of such measures (see Figures (37) and (38)).

For a more detailed comparison between the two models, we compare the z-scores of the measure ANNS as well as of the local weighted clustering coefficients evaluated under the UWCM with those for the same measures evaluated under the UECM (see subsection B of the Appendix for a more detailed explanation). More specifically, for every bank i, we define the z-scores

$$z_{\mathrm{ANNS}}^{\mathrm{UWCM}}(i) = \frac{\mathrm{ANNS}(i) - \langle \mathrm{ANNS}(i) \rangle_{\mathrm{UWCM}}}{\sigma[\mathrm{ANNS}(i)]_{\mathrm{UWCM}}}, \tag{40}$$

and

$$z_{\mathrm{ANNS}}^{\mathrm{UECM}}(i) = \frac{\mathrm{ANNS}(i) - \langle \mathrm{ANNS}(i) \rangle_{\mathrm{UECM}}}{\sigma[\mathrm{ANNS}(i)]_{\mathrm{UECM}}}, \tag{41}$$

where $\mathrm{ANNS}(i)\rangle_{\mathrm{UWCM}}$ and $\mathrm{ANNS}(i)\rangle_{\mathrm{UECM}}$ are respectively the expected values of the measure ANNS for bank i evaluated under the UWCM and the UECM; and $\sigma(\mathrm{ANNS}(i))_{\mathrm{UWCM}}$ and $\sigma(\mathrm{ANNS}(i))_{\mathrm{UECM}}$ are respectively the standard deviations of ANNS(i) evaluated under the UWCM and the UECM [7].

Similarly, the z-scores for the local weighted clustering coefficients for bank i evaluated under the UWCM and the UECM are defined as

$$z_{C_w}^{\mathrm{UWCM}}(i) = \frac{C_w^{un}(i) - \langle C_w^{un}(i) \rangle_{\mathrm{UWCM}}}{\sigma[C_w^{un}(i)]_{\mathrm{UWCM}}}, \tag{42}$$

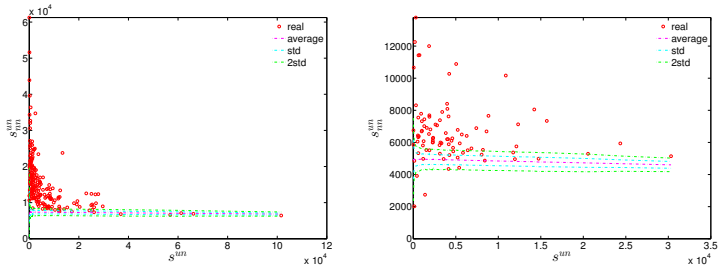
and

$$z_{C_w}^{\mathrm{UECM}}(i) = \frac{C_w^{un}(i) - \langle C_w^{un}(i) \rangle_{\mathrm{UECM}}}{\sigma[C_w^{un}(i)]_{\mathrm{UECM}}}. \tag{43}$$

We show the comparisons between the z-scores under the two considered configuration

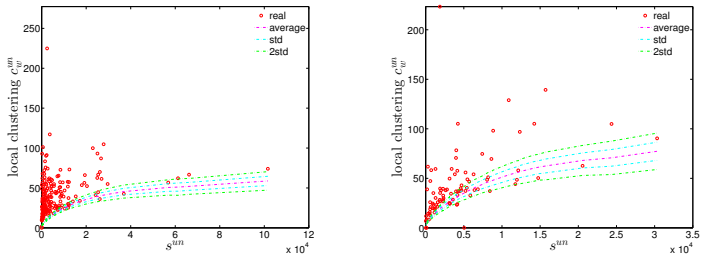
[7] Throughout this paper, the notation $\langle \mathrm{X} \rangle_{\text{null model}}$ and $\sigma[\mathrm{X}]_{\text{null model}}$ are respectively the expected value and standard deviation of X evaluated under the referenced null model.

models in Figure (39) and Figure (40). For almost all banks, we find that $|z^{\text{UECM}}_{\text{ANNS}}| < |z^{\text{UWCM}}_{\text{ANNS}}|$ and $|z^{\text{UECM}}_{C^{un}_w}| < |z^{\text{UWCM}}_{C^{un}_w}|$.



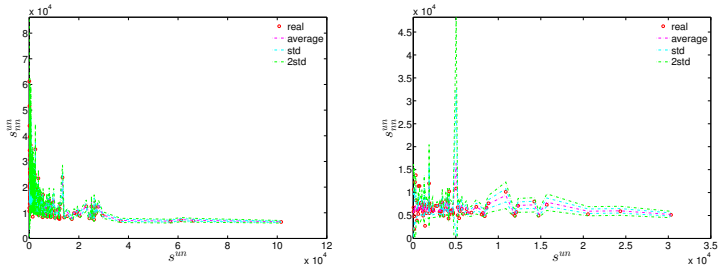
(a) s^{un}_{nn} in Q1 (b) s^{un}_{nn} in Q48

Figure 35: ANNS in the observed e-MID network and in the UWCM, in Q1 and Q48.



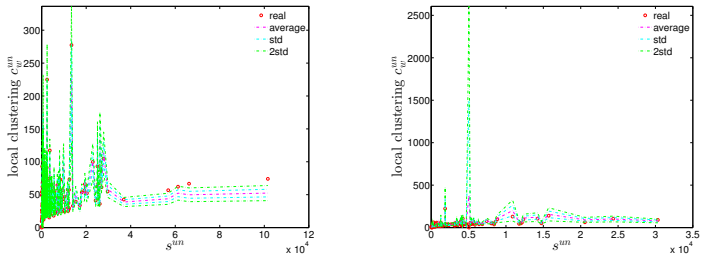
(a) C^{un}_w in Q1 (b) C^{un}_w in Q48

Figure 36: Local weighted clustering coefficients C^{un}_w in the observed e-MID network and in the UWCM, in Q1 and Q48.



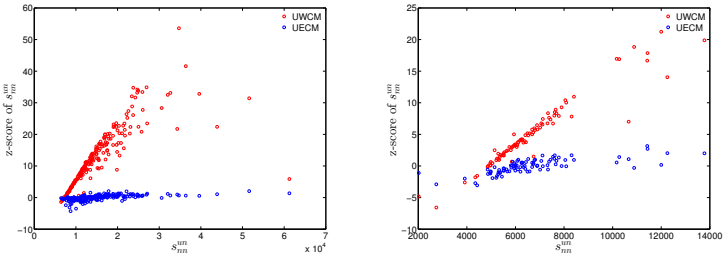
(a) s^{un}_{nn} in Q1 (b) s^{un}_{nn} in Q48

Figure 37: ANNS in the observed e-MID network and in the UECM, in Q1 and Q48.



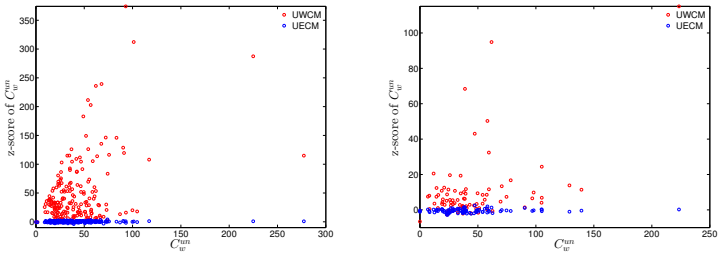
(a) C_w^{un} in Q1 (b) C_w^{un} in Q48

Figure 38: Local weighted clustering coefficients C_w^{un} in the observed e-MID network and in the UECM, in Q1 and Q48.



(a) z-scores of s_{nn}^{un} in Q1 (b) z-scores of s_{nn}^{un} in Q48

Figure 39: z-scores of s_{nn}^{un} vs. s_{nn}^{un} in the UWCM and the UECM, in Q1 and in Q48. Panel (a) for z-scores of s_{nn}^{un} in Q1, panel (b) for z-scores of s_{nn}^{un} in Q48.

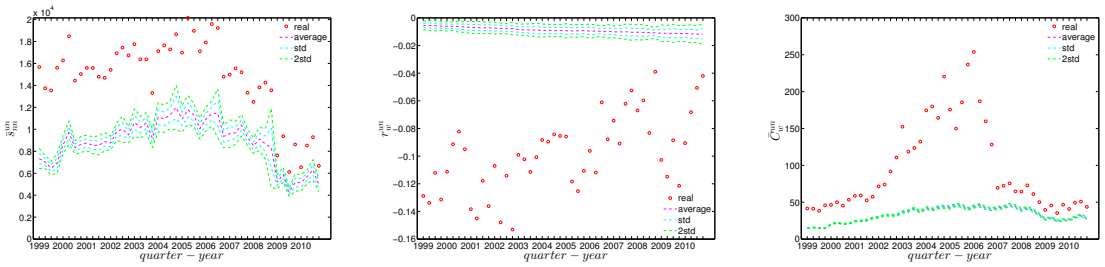


(a) z-scores of C_w^{un} in Q1 (b) z-scores of C_w^{un} in Q48

Figure 40: z-scores of C_w^{un} vs. C_w^{un} in the UWCM and the UECM. Panel (a) for z-scores of C_w^{un} in Q1, panel (b) for z-scores of C_w^{un} in Q48.

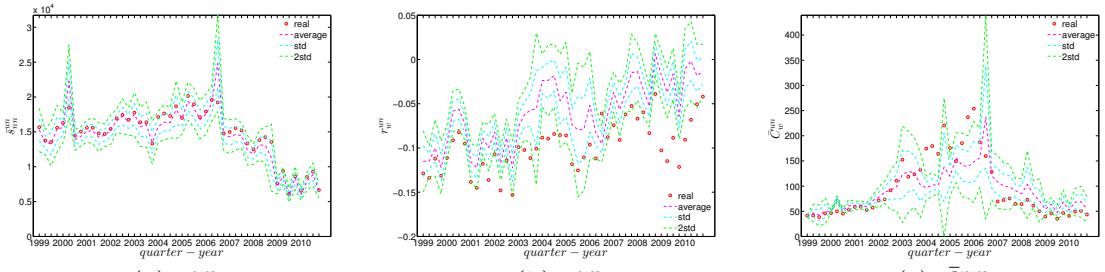
We now compare the evolution of $\bar{s}_{nn}^{un}$, r_w^{un}, and $\bar{C}_w^{un}$ for the observed network with the one obtained for these measures under the UWCM and the UECM. In Figure (41), we see that for most of the time, the observed values of $\bar{s}_{nn}^{un}$, r_w^{un}, and $\bar{C}_w^{un}$ lie outside the $\pm$ 2 bands associated with the UWCM. In contrast, in Figure (42), we see that the evolution of these measures is well captured by the ECM. The observed values of $\bar{s}_{nn}^{un}$ and $\bar{C}_w^{un}$ and the expected ones obtained from the ECM are in very close agreement. Even in the case of r_w^{un}, for which

several significant deviations are found, the main features of its evolution are well reproduced by the ECM.



(a) $\bar{s}_{nn}^{un}$ (b) r_w^{un} (c) $\bar{C}_w^{un}$

Figure 41: Evolution of $\bar{s}_{nn}^{un}$ (panel a), r_w^{un} (panel b), and $\bar{C}_w^{un}$ (panel c) in the observed e-MID network and in the UWCM.



(a) $\bar{s}_{nn}^{un}$ (b) r_w^{un} (c) $\bar{C}_w^{un}$

Figure 42: Evolution of $\bar{s}_{nn}^{un}$ (panel a), r_w^{un} (panel b), and $\bar{C}_w^{un}$ (panel c) in the observed e-MID network and in the UECM.

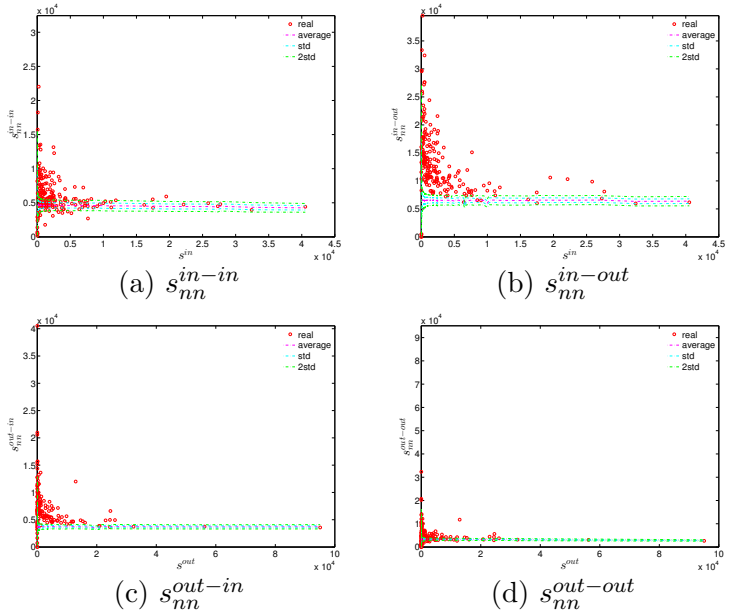
Directed Weighted Network

We now extend our comparison between the observed network and the configuration models to the directed weighted version. For this purpose, again the two relevant null models are employed, i.e. the DWCM and the DECM.

First, regarding the directed versions of the measure ANNS, we compare s_{nn}^{in-in}, s_{nn}^{in-out}, s_{nn}^{out-in}, and $s_{nn}^{out-out}$ of the observed network in the two chosen quarters with those obtained from the DWCM in Figures (43) and (44), and with those obtained from the DECM in Figures (45) and (46). Similar to the undirected weighted case, the z-scores of the directed weighted versions of the measure ANNS evaluated under these two models are also reported in Figures (47) and (48). Overall, we see that the main features of the measure ANNS are replicated much better by the DECM than by the DWCM. Furthermore, typically for almost all banks, we find that $|z_{\text{ANNS}}^{\text{DECM}}| < |z_{\text{ANNS}}^{\text{DWCM}}|$.

In terms of the third order structural correlations, the DECM again outperforms the DWCM in terms of reproducing the main features of local weighted clustering coefficients.

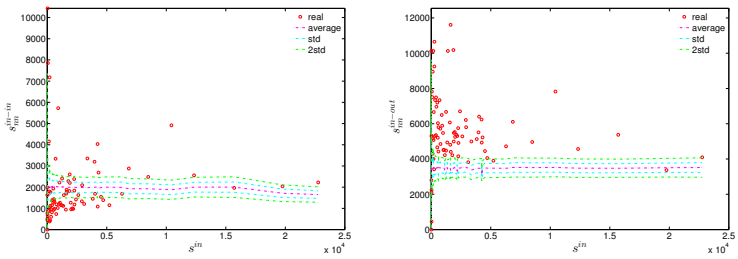
This is visualized in Figures (49), (50), (51), and (52). In addition, for each type of local weighted clustering coefficients we also calculate the z-scores evaluated under the DWCM and the DECM. As shown in Figures (53) and (54), we observe that on average $|z^{\mathrm{DECM}}_{C^{cyc}_w}| < |z^{\mathrm{DWCM}}_{C^{cyc}_w}|$, $|z^{\mathrm{DECM}}_{C^{mid}_w}| < |z^{\mathrm{DWCM}}_{C^{mid}_w}|$, $|z^{\mathrm{DECM}}_{C^{in}_w}| < |z^{\mathrm{DWCM}}_{C^{in}_w}|$, and $|z^{\mathrm{DECM}}_{C^{out}_w}| < |z^{\mathrm{DWCM}}_{C^{out}_w}|$, although there is often a certain deviation from this general tendency at the lower end of the spectrum of degrees and strengths.



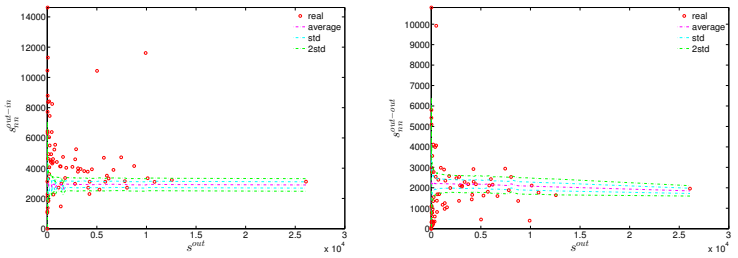
(a) s^{in-in}_{nn} (b) s^{in-out}_{nn}

(c) s^{out-in}_{nn} (d) $s^{out-out}_{nn}$

Figure 43: ANNSs in the observed e-MID network and in the DWCM, in Q1. s^{in-in}_{nn} (panel a), s^{in-out}_{nn} (panel b), s^{out-in}_{nn} (panel c), $s^{out-out}_{nn}$ (panel d).

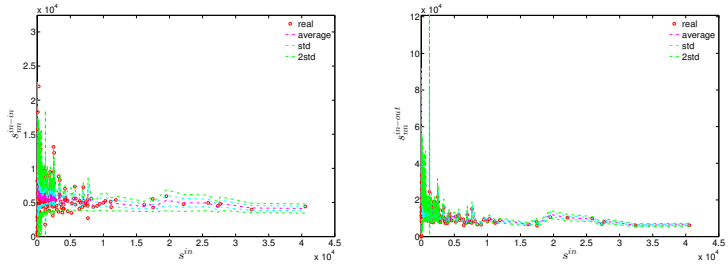


(a) s_{nn}^{in-in} (b) s_{nn}^{in-out}

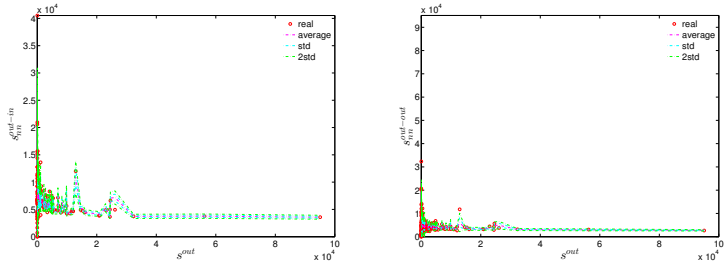


(c) s_{nn}^{out-in} (d) $s_{nn}^{out-out}$

Figure 44: ANNSs in the observed e-MID network and in the DWCM, in Q48. s_{nn}^{in-in} (panel a), s_{nn}^{in-out} (panel b), s_{nn}^{out-in} (panel c), $s_{nn}^{out-out}$ (panel d).

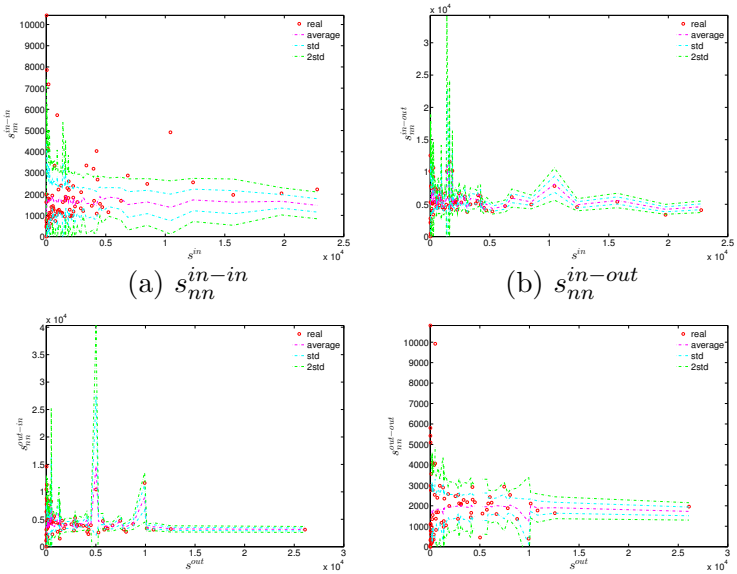


(a) s_{nn}^{in-in} (b) s_{nn}^{in-out}



(c) s_{nn}^{out-in} (d) $s_{nn}^{out-out}$

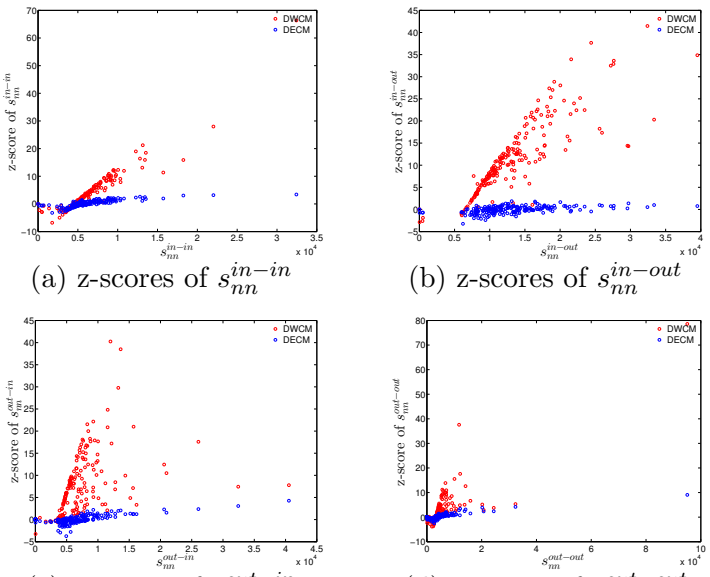
Figure 45: ANNSs in the observed e-MID network and in the DECM, in Q1. s_{nn}^{in-in} (panel a), s_{nn}^{in-out} (panel b), s_{nn}^{out-in} (panel c), $s_{nn}^{out-out}$ (panel d).



(a) s_{nn}^{in-in} (b) s_{nn}^{in-out}

(c) s_{nn}^{out-in} (d) $s_{nn}^{out-out}$

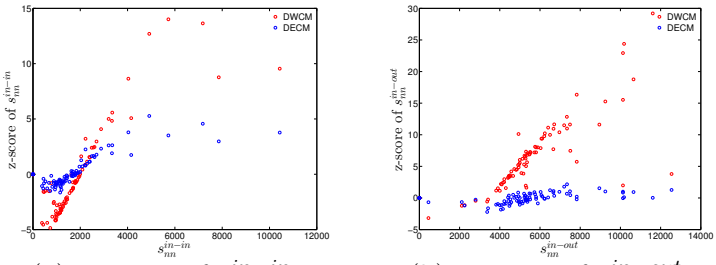
Figure 46: ANNSs in the observed e-MID network and in the DECM, in Q48. s_{nn}^{in-in} (panel a), s_{nn}^{in-out} (panel b), s_{nn}^{out-in} (panel c), $s_{nn}^{out-out}$ (panel d).



(a) z-scores of s_{nn}^{in-in} (b) z-scores of s_{nn}^{in-out}

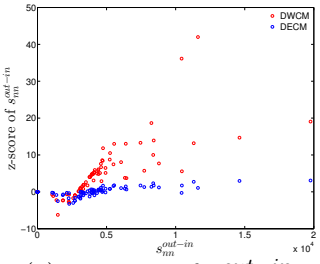
(c) z-scores of s_{nn}^{out-in} (d) z-scores of $s_{nn}^{out-out}$

Figure 47: z-scores of ANNSs vs. ANNSs, in the DWCM and DECM models, in Q1. Panels (a) for s_{nn}^{in-in}, (b) for s_{nn}^{in-out}, (c) for s_{nn}^{out-in}, (d) for $s_{nn}^{out-out}$.

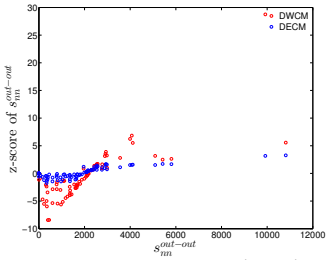


(a) z-scores of s_{nn}^{in-in}

(b) z-scores of s_{nn}^{in-out}

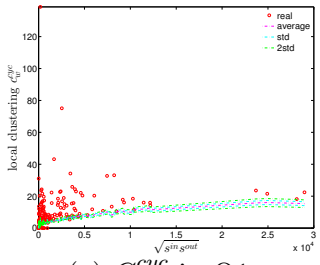


(c) z-scores of s_{nn}^{out-in}

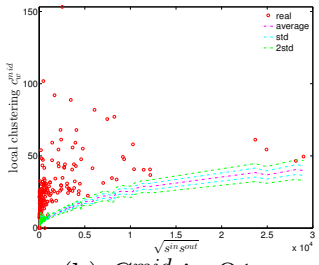


(d) z-scores of $s_{nn}^{out-out}$

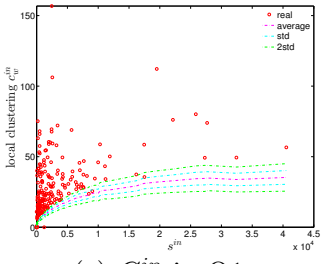
Figure 48: z-scores of ANNSs vs. ANNSs, in the DWCM and DECM models, in Q48. Panels (a) for s_{nn}^{in-in}, (b) for s_{nn}^{in-out}, (c) for s_{nn}^{out-in}, (d) for $s_{nn}^{out-out}$.



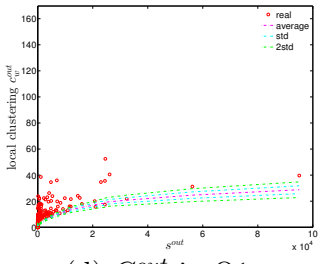
(a) C_w^{cyc} in Q1



(b) C_w^{mid} in Q1

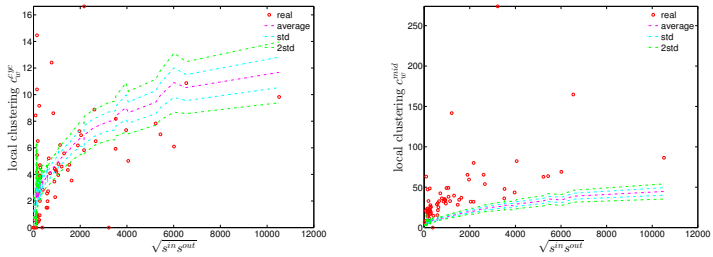


(c) C_w^{in} in Q1



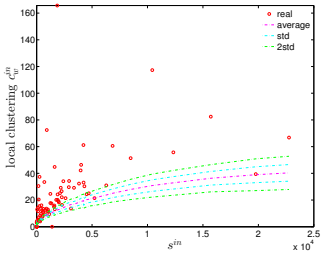
(d) C_w^{out} in Q1

Figure 49: Local weighted clustering coefficients in the observed e-MID network and in the DWCM, in Q1. C_w^{cyc} (panel a), C_w^{mid} (panel b), C_w^{in} (panel c), C_w^{out} (panel d).

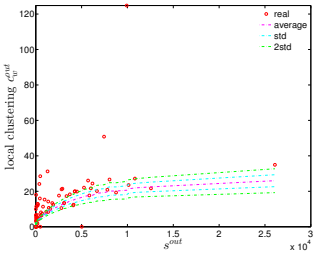


(a) C_w^{cyc} in Q48

(b) C_w^{mid} in Q48

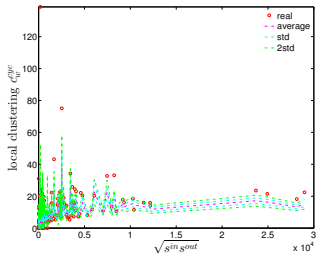


(c) C_w^{in} in Q48

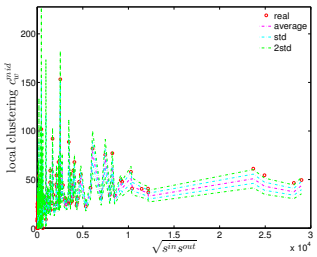


(d) C_w^{out} in Q48

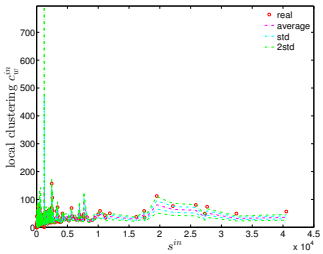
Figure 50: Local weighted clustering coefficients in the observed e-MID network and in the DWCM, in Q48. C_w^{cyc} (panel a), C_w^{mid} (panel b), C_w^{in} (panel c), C_w^{out} (panel d).



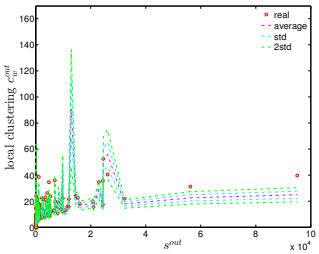
(a) C_w^{cyc} in Q1



(b) C_w^{mid} in Q1

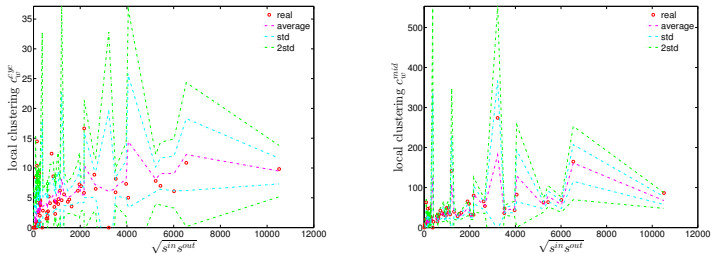


(c) C_w^{in} in Q1



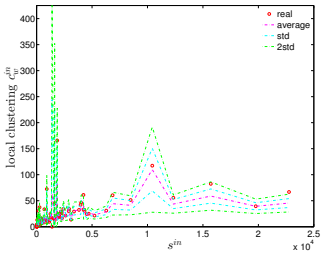
(d) C_w^{out} in Q1

Figure 51: Local weighted clustering coefficients in the observed e-MID network and in the DECM, in Q1. C_w^{cyc} (panel a), C_w^{mid} (panel b), C_w^{in} (panel c), C_w^{out} (panel d).

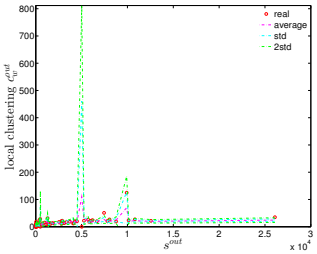


(a) C_w^{cyc} in Q48

(b) C_w^{mid} in Q48

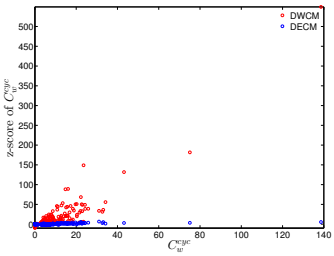


(c) C_w^{in} in Q48

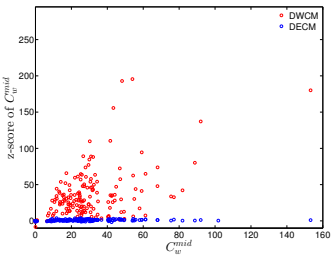


(d) C_w^{out} in Q48

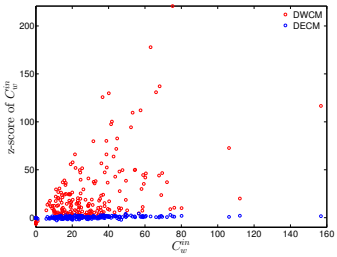
Figure 52: Local weighted clustering coefficients in the observed e-MID network and in the DECM, in Q48. C_w^{cyc} (panel a), C_w^{mid} (panel b), C_w^{in} (panel c), C_w^{out} (panel d).



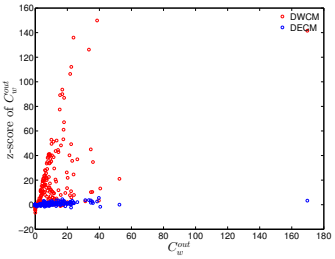
(a) z-scores of C_w^{cyc}



(b) z-scores of C_w^{mid}

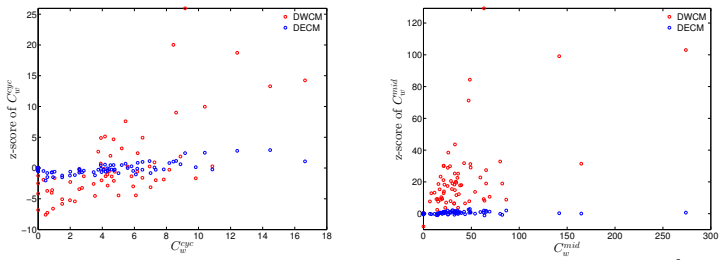


(c) z-scores of C_w^{in}

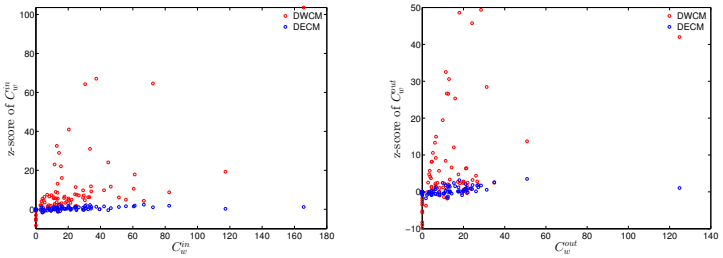


(d) z-scores of C_w^{out}

Figure 53: z-scores of C_w vs. C_w, evaluated under the DWCM and DECM models, in Q1. Panel (a) for C_w^{cyc}, panel (b) for C_w^{mid}, panel (c) for C_w^{in}, panel (d) for C_w^{out} .



(a) z-scores of C_w^{cyc} (b) z-scores of C_w^{mid}



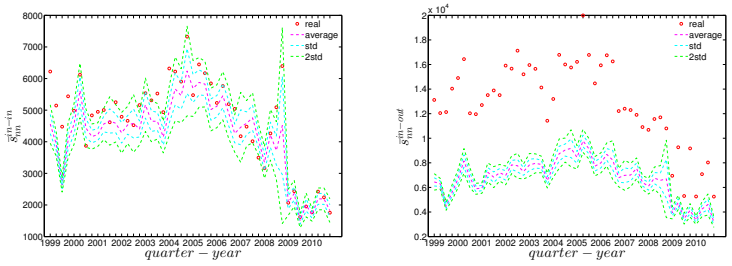
(c) z-scores of C_w^{in} (d) z-scores of C_w^{out}

Figure 54: z-scores of C_w vs. C_w, evaluated under the DWCM and DECM models, in Q48. Panel (a) for C_w^{cyc}, panel (b) for C_w^{mid}, panel (c) for C_w^{in}, panel (d) for C_w^{out} .

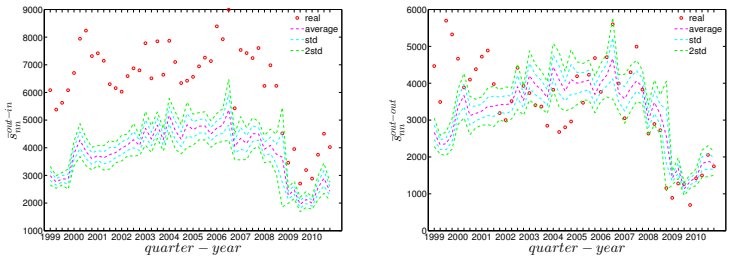
Finally, we analyze the predictive power of the two considered null models in terms of the evolution of the averages of the various versions of the measure ANNSs (i.e. $\bar{s}_{nn}^{in-in}$, $\bar{s}_{nn}^{in-out}$, $\bar{s}_{nn}^{out-in}$ and $\bar{s}_{nn}^{out-out}$), the global weighted assortativity indicators (i.e. r_w^{in-in}, r_w^{in-out}, r_w^{out-in} and $r_w^{out-out}$), and the averages of the local weighted clustering coefficients (i.e. $\bar{C}_w^{cyc}$, $\bar{C}_w^{mid}$, $\bar{C}_w^{in}$ and $\bar{C}_w^{out}$) (see also the next subsection for a further comparison).

Figures (55), (56), and (57) show significant deviations of the observed network from the DWCM over time. A comparison between the DECM and the observed network in terms of the aforementioned measures is shown in Figures (58), (59), and (60). Overall, we observe that, on the one hand, the DWCM is clearly dominated by the DECM, on the other hand, significant deviations from the DECM are still present in several quarters, regarding such as the average of the measure ANNS in the mixing category out-out ($\bar{s}_{nn}^{out-out}$), the global weighted assortativity indicators in the in-in and in-out categories, the average of the local weighted clustering coefficients $\bar{C}_w^{cyc}$ and the average of the local weighted clustering coefficients $\bar{C}_w^{out}$.

We emphasize that one of the main features not explained by the sequences of degrees and strengths of the network nodes themselves is the high level of clustering in the years preceding the crisis, i.e. the huge increase in various indirect exposures generated via more intensive interbank credit links.

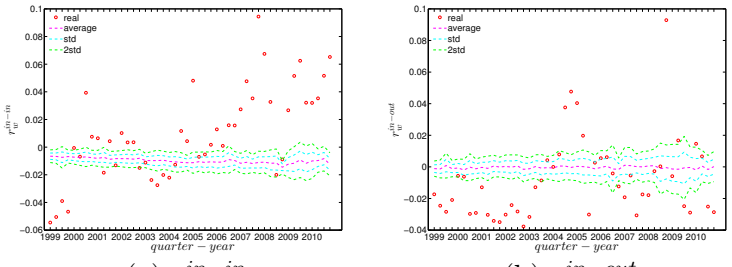


(a) $\bar{s}_{nn}^{in-in}$ (b) $\bar{s}_{nn}^{in-out}$

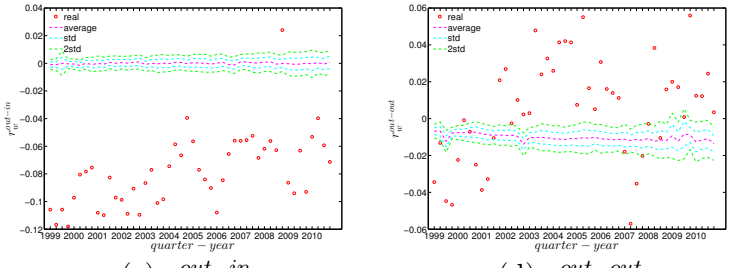


(c) $\bar{s}_{nn}^{out-in}$ (d) $\bar{s}_{nn}^{out-out}$

Figure 55: Evolution of the averages of ANNSs in the observed e-MID network and in the DWCM. $\bar{s}_{nn}^{in-in}$ (panel a), $\bar{s}_{nn}^{in-out}$ (panel b), $\bar{s}_{nn}^{out-in}$ (panel c), $\bar{s}_{nn}^{out-out}$ (panel d).

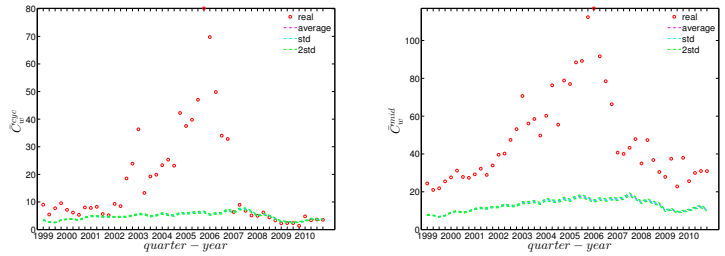


(a) r_w^{in-in} (b) r_w^{in-out}

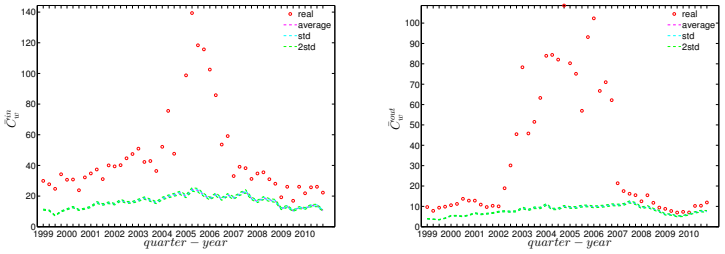


(c) r_w^{out-in} (d) $r_w^{out-out}$

Figure 56: Evolution of the global weighted assortativity indicators in the observed e-MID network and in the DWCM. r_w^{in-in} (panel a), r_w^{in-out} (panel b), r_w^{out-in} (panel c), $r_w^{out-out}$ (panel d).

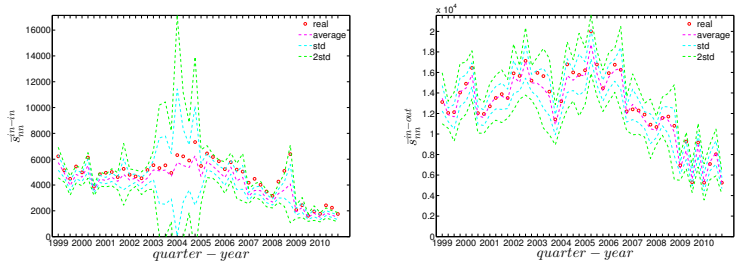


(a) $\bar{C}_w^{cyc}$ (b) $\bar{C}_w^{mid}$

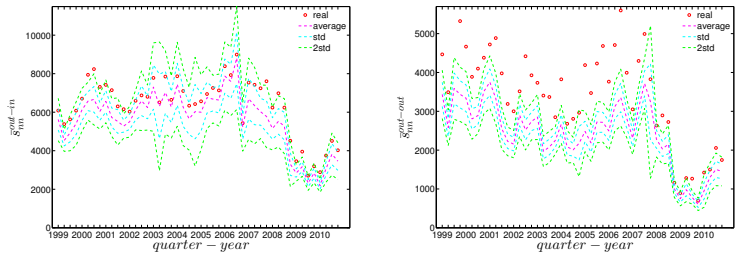


(c) $\bar{C}_w^{in}$ (d) $\bar{C}_w^{out}$

Figure 57: Evolution of the averages of local weighted clustering coefficients in the observed e-MID network and in the DWCM. $\bar{C}_w^{cyc}$ (panel a), $\bar{C}_w^{mid}$ (panel b), $\bar{C}_w^{in}$ (panel c), $\bar{C}_w^{out}$ (panel d).

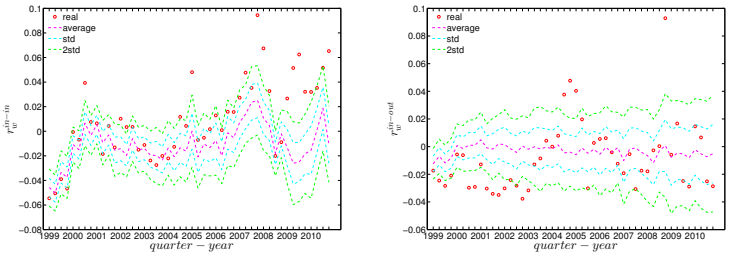


(a) $\bar{s}_{nn}^{in-in}$ (b) $\bar{s}_{nn}^{in-out}$

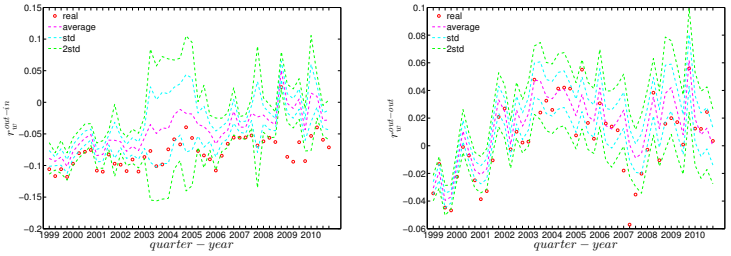


(c) $\bar{s}_{nn}^{out-in}$ (d) $\bar{s}_{nn}^{out-out}$

Figure 58: Evolution of the averages of ANNSs in the observed e-MID network and in the DECM. $\bar{s}_{nn}^{in-in}$ (panel a), $\bar{s}_{nn}^{in-out}$ (panel b), $\bar{s}_{nn}^{out-in}$ (panel c), $\bar{s}_{nn}^{out-out}$ (panel d).

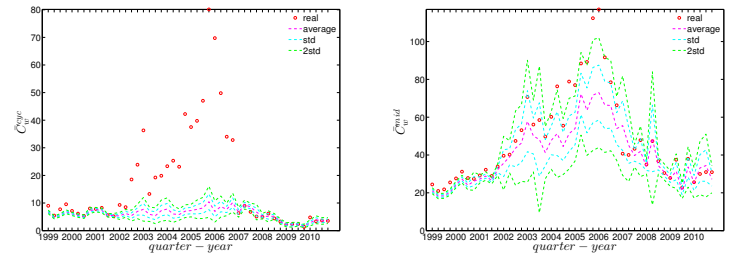


(a) r_w^{in-in} (b) r_w^{in-out}

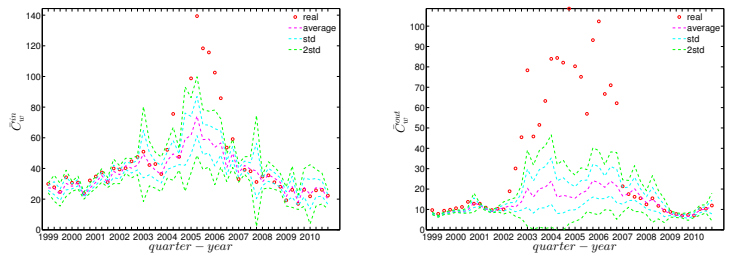


(c) r_w^{out-in} (d) $r_w^{out-out}$

Figure 59: Evolution of the global weighted assortativity indicators in the observed e-MID network and in the DECM. r_w^{in-in} (panel a), r_w^{in-out} (panel b), r_w^{out-in} (panel c), $r_w^{out-out}$ (panel d).



(a) $\bar{C}_w^{cyc}$ (b) $\bar{C}_w^{mid}$



(c) $\bar{C}_w^{in}$ (d) $\bar{C}_w^{out}$

Figure 60: Evolution of the averages of local weighted clustering coefficients in the observed e-MID network and in the DECM. $\bar{C}_w^{cyc}$ (panel a), $\bar{C}_w^{mid}$ (panel b), $\bar{C}_w^{in}$ (panel c), $\bar{C}_w^{out}$ (panel d).

Note that, although, in general, we find that the family of Enhanced Configuration Models outperforms the family of Weighted Configuration Models in terms of replicating the

main features of the structural correlations in the weighted version of the observed network, solving the system (35) to extract the hidden variables in the DECM (or system (31) in the UECM for the undirected version of the network) is much more computationally demanding than solving the system (29) for the DWCM (or sys. (25) for the UWCM for the undirected version of the network) [8,9].

4.4 z-scores analysis revealing structural changes in the weighted system

To analyze the evolution of the discrepancies between the referenced models and the observed network, we define z-scores for the global indicators, i.e. for $\bar{s}_{nn}^{un}, r_w^{un}, \bar{C}_{un}^w$ in the undirected weighted network (evaluated under the UWCM and the UECM) and for $\bar{s}_{nn}^{in-in}, \bar{s}_{nn}^{in-out}, \bar{s}_{nn}^{out-in}, \bar{s}_{nn}^{out-out}$, $r_w^{in-in}, r_w^{in-out}, r_w^{out-in}, r_w^{out-out}$, $\bar{C}_w^{cyc}, \bar{C}_w^{mid}, \bar{C}_w^{in}$, and $\bar{C}_w^{out}$ in the directed weighted network (evaluated under the DWCM and the DECM).

Before going into details, it should be noted that from Figure (61) to Figure (64), when comparing the UECM with the UWCM in the undirected version or the DECM with the DWCM in the directed version, some of the high z-scores under the UECM (or under the DECM) are blurred because of the presence of much larger z-scores under the UWCM (or under the DWCM).

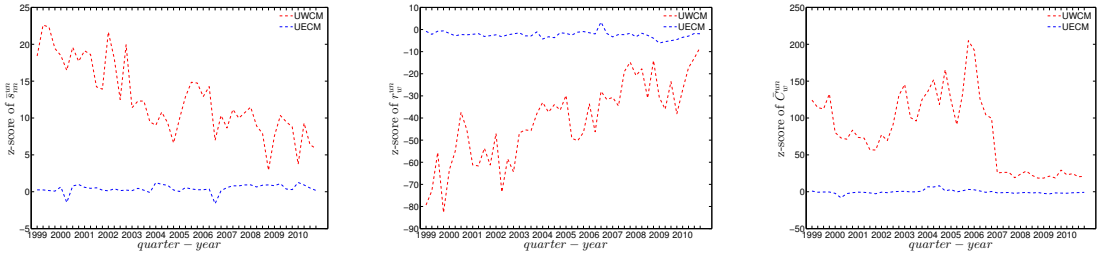
In the undirected weighted case, two important findings are obtained. First, overall, the z-scores are mostly smaller in absolute value under the UECM than under the UWCM (see Figure (61)). This is consistent with what we found for the local indicators and re-emphasizes the finding that the UECM out-performs the UWCM. Second, interestingly, in panel (c) of Figure (61) we see that the distance between the z-scores for $\bar{C}_w^{un}$ evaluated under the UWCM and the UECM increases over the period from 2002 to 2006, and then decreases sharply after the financial crisis. This suggests that the importance of particular basic features of a network (like its degree sequence or its strength sequence) for the emergence of higher order correlations structures can vary over time.

In the directed weighted case, similarly, we find that the z-scores under the DECM

[8] According to Squartini et al. (2015), solving system (35) for the DECM and solving the system (31) for the UECM may be very time consuming if the strength distribution contains big outliers and the degree distribution is narrow. This also happens in our study, and in fact our data set shows that the strength distribution is much wider than the degree distribution.

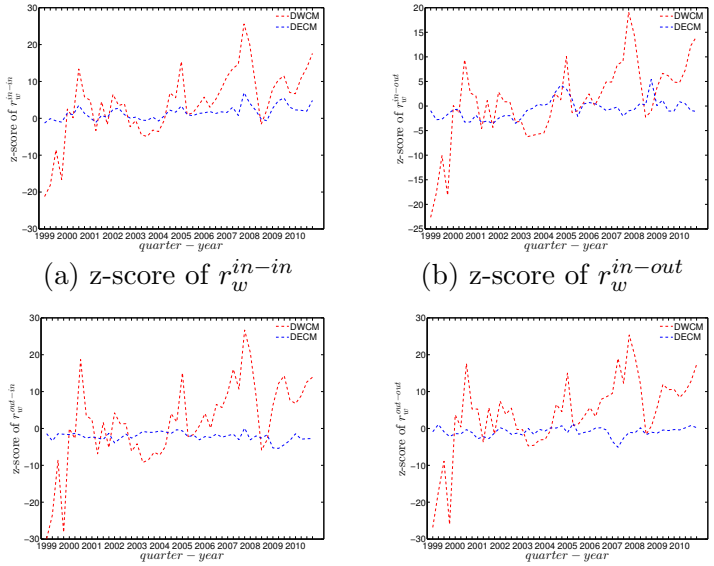
[9] Following Mastrandrea et al. (2014) and Squartini et al. (2015), in order to speed up the process of solving system (31) for the UECM and system (35) for the DECM, we have used the iteration method, which uses the output of the previous iteration as the initial value for the current one. However, it remains a very time consuming process to obtain an acceptable solution for the hidden variables.

are much smaller in absolute value than those evaluated under the DWCM (see Figures (62), (63), and (64)). As we can see in Figure (64), similar to the undirected version, the distance between the z-scores for each of the $\bar{C}_w$ evaluated under the DWCM and the DECM continuously increases during the period 2002 to 2006, and then decreases dramatically after the financial crisis.



(a) z-score of $\bar{s}^{un}_{nn}$ (b) z-score of r^{un}_{w} (c) z-score of $\bar{C}^{un}_{w}$

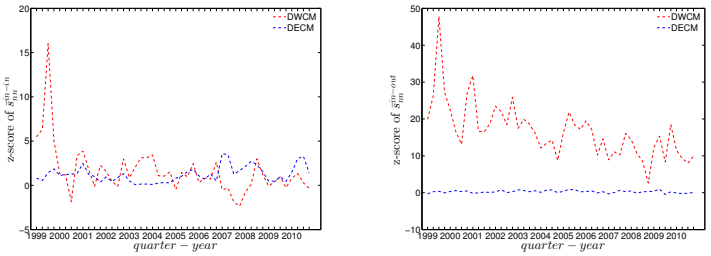
Figure 61: Evolution of z-scores for $\bar{s}^{un}_{nn}$ (panel a), r^{un}_{w} (panel b), and $\bar{C}^{un}_{w}$ (panel c) evaluated under the UWCM (red dashed lines) and the UECM (blue dashed lines).



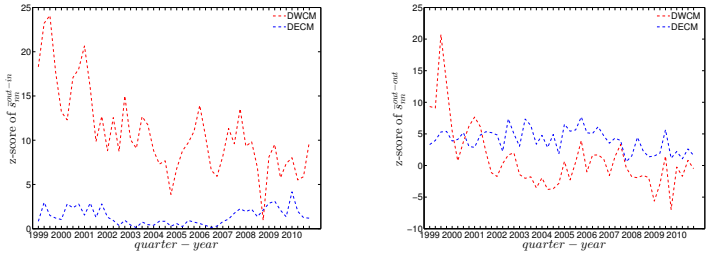
(a) z-score of r^{in-in}_{w} (b) z-score of r^{in-out}_{w}

(c) z-score of r^{out-in}_{w} (d) z-score of $r^{out-out}_{w}$

Figure 62: Evolution of z-scores for r^{in-in}_{w} (panel a), r^{in-out}_{w} (panel b), r^{out-in}_{w} (panel c), and $r^{out-out}_{w}$ evaluated under the DWCM (red dashed lines) and the DECM (blue dashed lines).

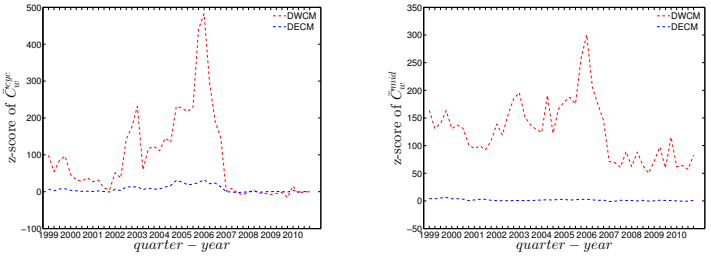


(a) z-score of $\bar{s}_{nn}^{in-in}$ (b) z-score of $\bar{s}_{nn}^{in-out}$

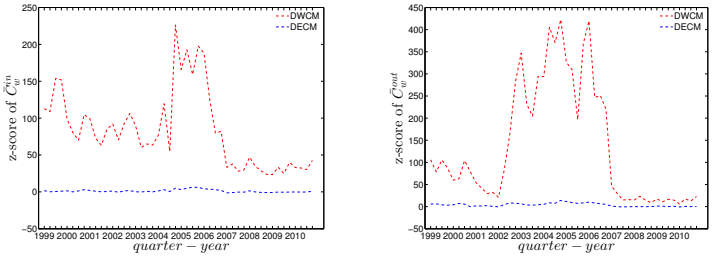


(c) z-score of $\bar{s}_{nn}^{out-in}$ (d) z-score of $\bar{s}_{nn}^{out-out}$

Figure 63: Evolution of z-scores for $\bar{s}_{nn}^{in-in}$ (panel a), $\bar{s}_{nn}^{in-out}$ (panel b), $\bar{s}_{nn}^{out-in}$ (panel c), and $\bar{s}_{nn}^{out-out}$ evaluated under the DWCM (red dashed lines) and the DECM (blue dashed lines).



(a) z-score of $\bar{C}_w^{cyc}$ (b) z-score of $\bar{C}_w^{mid}$



(c) z-score of $\bar{C}_w^{in}$ (d) z-score of $\bar{C}_w^{out}$

Figure 64: Evolution of z-scores for $\bar{C}_w^{cyc}$ (panel a), $\bar{C}_w^{mid}$ (panel b), $\bar{C}_w^{in}$ (panel c), and $\bar{C}_w^{out}$ evaluated under the DWCM (red dashed lines) and the DECM (blue dashed lines).

5 Conclusions

In this study, we investigated the structural correlations in the e-MID network. We find that the observed structural correlations can vary across different versions of the network (binary vs weighted and undirected vs directed). In the undirected version of the network, the mixing is disassortative in both the binary and the weighted case. In addition, when the directions of the edges are taken into account, we find that among the four mixing categories (i.e. in-in, in-out, out-in, and out-out), the global assortativity in the out-in category comes closest to the mixing observed in the undirected network. The similarity between these two quantities is suggested in the study by van der Hoorn and Litvak (2015). Due to the fact that only in the out-in mixing category the considered edge (see the out-in category in Figure (1)) contributes to the node degrees on both of its sides, this mixing category can be considered a generalization of the mixing in undirected networks. During our analysis of the evolution of the third order correlations among banks over time, we detected dramatic changes in the network structure surrounding the recent financial crisis in 2007. More specifically, in the weighted network, the averages of the local weighted clustering coefficients appear elevated from the adoption of the Euro up until 2006, and then decrease dramatically around the time of the financial crisis. We also report strong indications of elevated shared risk in the network, evidenced by the prevalence of the "middleman" and "inward" types of clustering in the network.

Moreover, by employing the various configuration models, we examined whether the information encoded in the local constraints (like the observed degree sequence and/or the strength sequence of a network) can explain higher order structural correlations. We find that, in the binary case, the degree sequence is informative in terms of explaining the main features of the structural correlations in the e-MID network. However, under closer scrutiny, the binary e-MID network does display some patterns that cannot simply be explained by the degree sequence in conjunction with the configuration model.

In the weighted version of the network, for the most part, the structural correlations in the observed e-MID network are deviating strongly from their respective expectations evaluated under the Weighted Configuration Models, which capture only the heterogeneity in the strength sequence(s) (i.e. the UWCM in the undirected version and the DWCM in the directed version). One possible explanation is that while all measures of structural correlations used in the weighted network depend on the elements of both the adjacency as well as the weighting matrices, neither the UWCM or DWCM utilize information about the node degrees (degree sequence), which is, in fact, found to be more important than the

strength sequence in reproducing the topological properties of real world networks (see, for example, Squartini et al., 2011a, Squartini et al., 2011b; Squartini et al., 2015).

Due to the failures of the UWCM and the DWCM, we consider the family of Enhanced Configuration Models, which constrains the degree as well as the strength sequences of the randomized ensemble to match those of the observed network on average (i.e. UECM in the undirected case and the DECM in the directed case). Our findings indicate that the randomized ensembles produced by the Enhanced Configuration Model have a much greater predictive power. This is in line with what was found in previous studies such as Mastrandrea et al. (2014) and Squartini et al. (2015), and is not very surprising since the Enhanced Configuration Models utilize more information when replicating the structural correlations of the observed network. The results obtained from the analysis of the DECM confirms the role that the distribution of the in-coming and out-going degrees together with their strengths (volumes) in directed weighted networks plays for the emergence of higher order structural correlations.

Still, a detailed comparison between the observed network and the Enhanced Configuration Models reveals that even this family of Configuration Models is not able to produce accurate estimates for all the measures of structural correlations we used, meaning that some of the patterns can be considered non-random or unexplained by the models. For instance, in the undirected network, we find that even when using the UECM, the weighted assortativity deviates significantly from the respective expected value in a couple of times. In the directed weighted network, the global weighted assortativity in the in-in as well as in the in-out mixing categories and the average of the local weighted clustering coefficients of "inward","outward", and "cyclical" clustering also display patterns in several quarters, mainly from 2002 to 2006, that deviate from the expectations based on the configuration model. The high degree of clustering in this episode is the one characteristic that can not be explained satisfactorily via the influence of lower-order characteristics like the degree and strength sequences. Hence, this finding points to a behavioral change in the formation of the credit network: A deliberate increase of indirect exposure through multiple credit relations. Interestingly, with the crisis year 2007, we find an abrupt reduction of all clustering coefficients to their "normal" levels implied by the degree and strength sequences.

The Enhanced Configuration Models also fail to reproduce the local behavior of certain banks captured by the local indicators of structural correlations. Unfortunately, because of the lack of more detailed information about the banks in the system, we can not identify the factors for the formations of such deviating patterns.

Interestingly, similar to the study of Squartini et al. (2013) [10], we also observe the evidence for structural changes when comparing the weighted version of the e-MID network with the weighted configuration models. More specifically, the distance between the predictions of the Weighted Configuration Models and of the Enhanced Configuration Models for the averages of local weighted clustering coefficients continuously increases from the adoption of the Euro up until the financial crisis in 2007 and then sharply decreases after that. This result can be interpreted as an indication of structural changes in the network associated with these two critical events. It also suggests that the importance of particular basic features of a network (like its degree sequence or its strength sequence) for the emergence of higher order correlation structures can vary over time.

Due to issues of confidentiality, in many cases, the biggest challenge in the analysis of complex real financial systems lies in the utilization of the limited available information. Our results can be understood as an evaluation of the potential of configuration models to reconstruct higher order topological properties of a network from limited information (e.g. see Mastrandrea et al., 2014; Cimini et al., 2015a). Meaningful systemic risk evaluation can be conducted on reconstructed networks only to the extent to which the reconstruction is reliable (see, for example, Cimini et al., 2015b).

In addition, the configuration models translate the local constraints in the observed network into hidden variables associated with the individual banks. It would be interesting to investigate whether some individual node characteristics (i.e. non-topological properties) correlate with the extracted hidden variables (see, for example, Garlaschelli and Loffredo, 2004; Garlaschelli et al., 2007; Garlaschelli and Loffredo, 2008; Almog et al., 2015), however, such additional information is unfortunately not available in our data set. This can be a fruitful direction for future research into financial networks.

Moreover, since the Exponential Random Graph Model is generic and flexible enough, one may want to investigate the extent to which it can be useful to use other statistics of the observed network as ensemble constraints. For instance, the average degree of the nearest neighbors or the local clustering coefficients might also prove informative in explaining particular topological properties of the observed network (see, for example, Park and Newman (2004) and Bianconi (2009) for employing different constraints). In addition, since the second and third order structural correlations are the main focus of this study, we suggest that the role of various constraints for the emergence of higher order correlations (or motifs) and for the meso-scale network structures such as the core-periphery and community structures

[10] Squartini et al. (2013) focus on the analysis of the binary version of the network of interbank exposures among Dutch banks over the period 1998-2008.

should be studied further.

6 References

Almog A., Squartini T., Garlaschelli D. 2015. A GDP-driven model for the binary and weighted structure of the international trade network. New Journal of Physics 17.

Barrat A., Barthélemy M., Pastor-Satorras R., Vespignani A. 2004. The architecture of complex weighted networks. Proc. Natl. Acad. Sci. 101 (11), pp. 3747-3752.

Bianconi G. 2009. Entropy of network ensembles. Physical Review E 79 (3).

Cimini G., Squartini T., Garlaschelli D., Gabrielli A. 2015a. Estimating topological properties of weighted networks from limited information. Physical Review E 92 (4).

Cimini G., Squartini T., Garlaschelli D., Gabrielli A. 2015b. Systemic risk analysis on reconstructed economic and financial networks. Scientific Reports 5.

De Masi G., Iori G., Caldarelli G. 2006. Fitness model for the Italian interbank money market. Physical Review E 74 (6).

Fagiolo G. 2007. Clustering in complex directed networks. Physical Review E 76 (2).

Finger K., Fricke D., Lux T. 2013. Network analysis of the e-MID overnight money market: the informational value of different aggregation levels for intrinsic dynamic processes. Computational Management Science 10 (2), pp. 187-211.

Foster J. G., Foster D. V., Grassberger P., Paczuski M. 2010. Edge direction and the structure of networks. Proc. Natl. Acad. Sci. 107 (24), pp. 10815-10820.

Fricke D. 2012. Trading strategies in the overnight money market: Correlations and clustering on the e-MID trading platform. Physica A: Statistical Mechanics and its Applications, 391 (24), pp. 6528-6542.

Fricke D., Finger K., Lux T. 2013. On assortative and disassortative mixing scale-free networks: The case of interbank credit networks. Kiel Working Papers 1830, Kiel Institute for the World Economy. Available at:
`https://www.ifw-members.ifw-kiel.de/publications/on-assortative-and-disassortative-mixing-scale-free-networks-the-case-of-interbank-credit-networks/1830_KWP.pdf`.

Fricke D., Lux T. 2015a. On the distribution of links in the interbank network: evidence from the e-MID overnight money market. Empirical Economics 49 (4), pp. 1463-1495.

Fricke D., Lux T. 2015b. Core-periphery structure in the overnight money market: Evidence from the e-MID trading platform. Computational Economics. 45 (3), pp. 359-395.

Garlaschelli D., Loffredo M. I. 2004. Fitness-dependent topological properties of the world trade web. Physical Review Letters 93 (18).

Garlaschelli D., Di Matteo T., Aste T., Caldarelli G., Loffredo M. 2007. Interplay between topology and dynamics in the world trade web. The European Physical Journal B 57 (2), pp. 159-164.

Garlaschelli D., Loffredo M. I. 2008. Maximum likelihood: Extracting unbiased information from complex networks. Physical Review E 78 (1).

Holme P., Park S. M., Kim B. J., Edling C. R. 2007. Korean university life in a network perspective: Dynamics of a large affiliation network. Physica A: Statistical Mechanics and its Applications 373, pp. 821-830.

Maslov S., Sneppen K. 2002. Specificity and stability in topology of protein networks. Science 296 (5569), pp. 910-913.

Maslov S., Sneppen K., Zaliznyak A. 2004. Detection of topological patterns in complex networks: Correlation profile of the internet. Physica A: Statistical Mechanics and its Applications 333, pp. 529-540.

Mastrandrea R., Squartini T., Fagiolo G., Garlaschelli D. 2014. Enhanced reconstruction of weighted networks from strengths and degrees. New Journal of Physics 16.

Newman M. E. J. 2002. Assortative mixing in networks. Physical Review Letters 89 (20).

Newman M. E. J. 2003a. Mixing patterns in networks. Physical Review E 67 (2).

Newman M. E. J. 2003b. The structure and function of complex networks. Society for Industrial and Applied Mathematics Review 45 (2), pp. 167-256.

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

Park J., Newman M. E. J. 2003. Origin of degree correlations in the Internet and other networks. Physical Review E 68 (2).

Park J., Newman M. E. J. 2004. Statistical mechanics of networks. Physical Review E 70 (6).

Piraveenan M., Prokopenko M., Zomaya A. 2010. Classifying complex networks using unbiased local assortativity. In Fellermann H., Dörr M., Hanczyc M. M., Ladegaard Laursen L., Maurer S., Merkle D., Monnard P. -A., Stoy K., Rasmussen S. (Eds.), *Artificial Life XII, Proc. 12th Int'l Conf. Synthesis and Simulation of Living Systems* (pp. 329-336).

Piraveenan M., Prokopenko M., Zomaya A. 2012. Assortative mixing in directed biological networks. IEEE/ACM Transactions on Computational Biology and Bioinformatics 9

(1), pp. 66-78.

Saramäki J., Kivelä M., Onnela J., Kaski K., Kertész J. 2007. Generalizations of the clustering coefficient to weighted complex networks. Physical Review E 75 (2).

Squartini T., Fagiolo G., Garlaschelli D. 2011a. Randomizing world trade. I. A binary network analysis. Physical Review E 84 (4).

Squartini T., Fagiolo G., Garlaschelli D. 2011b. Randomizing world trade. II. A weighted network analysis. Physical Review E 84 (4).

Squartini T., Garlaschelli D. 2011. Analytical maximum-likelihood method to detect patterns in real networks. New Journal of Physics 13.

Squartini T., van Lelyveld I., Garlaschelli D. 2013. Early-warning signals of topological collapse in interbank networks. Scientific Reports 3.

Squartini T., Mastrandrea R., Garlaschelli D. 2015. Unbiased sampling of network ensembles. New Journal of Physics 17.

Tabak B. M., Takami M., Rocha J. M. C., Cajueiro D. O., Souza S. R. S. 2014. Directed clustering coefficient as a measure of systemic risk in complex banking networks. Physica A: Statistical Mechanics and its Applications 394, pp. 211-216.

van der Hoorn P., Litvak N. 2015. Degree-degree dependencies in directed networks with heavy-tailed degrees. Internet Mathematics 11 (2), pp. 155-179.

Watts D. J., Strogatz S. H. 1998. Collective dynamics of "small-world" networks. Nature 393, pp. 440-442.

Zlatic V., Bianconi G., Díaz-Guilera A., Garlaschelli D., Rao F., Caldarelli G. 2009. On the rich-club effect in dense and weighted networks. The European Physical Journal B 67 (3), pp. 271-275.

Zhang B., Horvath S. 2005. A general framework for weighted gene co-expression network analysis. Statistical Applications in Genetics and Molecular Biology 4 (1).

7 Appendix

Assortativity Coefficients

Overall Assortativity

In an undirected network, define the list of m edges $\{A_e B_e\}_{e=1}^{m}$, where for each index e, the two nodes A_e, B_e stand for the ends of an edge. Note that, the overall assortativity

indicator r_{bin}^{un} can be calculated via $\{k_i^{un}\}_{i=1}^n$ and $\{k_{nn,i}^{un}\}_{i=1}^n$ as

$$r_{bin}^{un} = \frac{\sum_{i=1}^n (k_i^{un})^2 k_{nn,i}^{un} - \frac{1}{2m}[\sum_{i=1}^n (k_i^{un})^2]^2}{\sum_{i=1}^n (k_i^{un})^3 - \frac{1}{2m}[\sum_{i=1}^n (k_i^{un})^2]^2}, \tag{44}$$

where $m = \frac{1}{2}\sum_{i=1}^n k_i^{un}$ (e.g. Park and Newman, 2003).

In a directed network, suppose that we have a list of M edges $\{A_e B_e\}_{e=1}^M$, where for each index e, the two nodes A_e, B_e respectively stand for the source and target nodes (note that $M = \sum_{i=1}^n k_i^{in} = \sum_{i=1}^n k_i^{out}$).

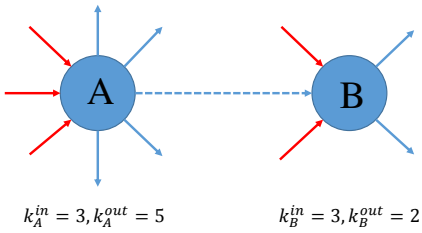


Figure 65: In-coming, out-going degrees to two vertices of an edge in directed networks.

Each node A_e or B_e has an in-coming degree and an out-going degree (see Figure (65)). Consequently, we have four combinations of degrees associated with each edge as mentioned in Figure (1). Therefore, regarding the degree dependencies, four separate indicators can be obtained, i.e. $r_{bin}^{in-in}, r_{bin}^{out-in}, r_{bin}^{in-out}, r_{bin}^{out-out}$. Similar to the undirected case, mathematically, these measures of overall assortativity actually depend on the degree sequences $\{k^{in}\}_{i=1}^n$, $\{k^{out}\}_{i=1}^n$ as well as the sequences of the average nearest neighbor degrees $k_{nn,i}^{in-in}, k_{nn,i}^{in-out}, k_{nn,i}^{out-in}, k_{nn,i}^{out-out}$ (e.g. Piraveenan et al., 2012; van der Hoorn and Litvak, 2015). More specifically, accordingly, they are given by

$$r_{bin}^{in-in} = \frac{\frac{1}{2}[\sum_{i=1}^n (k_i^{in})^2 k_{nn,i}^{in-in} + k_i^{in} k_i^{out} k_{nn,i}^{out-in}] - \frac{1}{M}[\sum_{i=1}^n (k_i^{in})^2 \sum_{i=1}^n (k_i^{in} k_i^{out})]}{\sqrt{\{\sum_{i=1}^n (k_i^{in})^3 - \frac{1}{M}[\sum_{i=1}^n (k_i^{in})^2]^2\}\{\sum_{i=1}^n (k_i^{in})^2 k_i^{out} - \frac{1}{M}[\sum_{i=1}^n (k_i^{in} k_i^{out})]^2\}}}, \tag{45}$$

$$r_{bin}^{in-out} = \frac{\frac{1}{2}\sum_{i=1}^{n} k_i^{in} k_i^{out}(k_{nn,i}^{in-in} + k_{nn,i}^{out-out}) - \frac{1}{M}[\sum_{i=1}^{n}(k_i^{in}k_i^{out})]^2}{\sqrt{\{\sum_{i=1}^{n}(k_i^{in})^2 k_i^{out} - \frac{1}{M}[\sum_{i=1}^{n}(k_i^{in}k_i^{out})]^2\}\{\sum_{i=1}^{n}(k_i^{out})^2 k_i^{in} - \frac{1}{M}[\sum_{i=1}^{n}(k_i^{out}k_i^{in})]^2\}}}, \tag{46}$$

$$r_{bin}^{out-in} = \frac{\frac{1}{2}\sum_{i=1}^{n}[(k_i^{out})^2 k_{nn,i}^{out-in} + (k_i^{in})^2 k_{nn,i}^{in-out}] - \frac{1}{M}[\sum_{i=1}^{n}(k_i^{in})^2 \sum_{i=1}^{n}(k_i^{out})^2]}{\sqrt{\{\sum_{i=1}^{n}(k_i^{in})^3 - \frac{1}{M}[\sum_{i=1}^{n}(k_i^{in})^2]^2\}\{\sum_{i=1}^{n}(k_i^{out})^3 - \frac{1}{M}[\sum_{i=1}^{n}(k_i^{out})^2]^2\}}}, \tag{47}$$

and

$$r_{bin}^{out-out} = \frac{\frac{1}{2}\sum_{i=1}^{n}[(k_i^{out})^2 k_{nn,i}^{out-out} + k_i^{out}k_i^{in}k_{nn,i}^{in-out}] - \frac{1}{M}[\sum_{i=1}^{n}(k_i^{out})^2 \sum_{i=1}^{n}(k_i^{in}k_i^{out})]}{\sqrt{\{\sum_{i=1}^{n}(k_i^{out})^3 - \frac{1}{M}[\sum_{i=1}^{n}(k_i^{out})^2]^2\}\{\sum_{i=1}^{n}(k_i^{out})^2 k_i^{in} - \frac{1}{M}[\sum_{i=1}^{n}(k_i^{in}k_i^{out})]^2\}}}. \tag{48}$$

Local Assortativity

The local assortativity statistics is obtained as the (unbiased) contribution of individual nodes to the overall (global) assortativity. The basic idea is that the numerator in the Pearson correlation coefficient proposed by Newman (2003) can be reformulated based on the contribution of individual nodes instead of edges (e.g. Piraveenan et al., 2010; Piraveenan et al., 2012).

It should be emphasized that, for the directed version of the measure of local assortativity introduced in Piraveenan et al. (2012), the two in-out and out-in degree dependencies are not differentiated, when in fact they exhibit totally different behaviors (as found in Foster et al. (2010) and in Sec. 3 of our study). In our study, the contributions to the in-out and out-in degree dependencies are distinguishable.

We denote the local assortativity measures for a given node i as ρ_i^{in-in}, ρ_i^{in-out}, ρ_i^{out-in}, and $\rho_i^{out-out}$ corresponding to the four mixing categories in the directed version and ρ_i^{un} is used for the undirected version. Note that the following equalities must hold:

$$r_{bin}^{un} = \sum_{i=1}^{n} \rho_i^{un}, \tag{49}$$

$$r_{bin}^{in-in} = \sum_{i=1}^{n} \rho_i^{in-in}, \tag{50}$$

$$r_{bin}^{in-out} = \sum_{i=1}^{n} \rho_i^{in-out}, \tag{51}$$

$$r_{bin}^{out-in} = \sum_{i=1}^{n} \rho_i^{out-in}, \tag{52}$$

$$r_{bin}^{out-out} = \sum_{i=1}^{n} \rho_i^{out-out}. \tag{53}$$

First, we define

$$\mu_{un} = \frac{1}{2m} \sum_{i=1}^{n} (k_i^{un})^2, \tag{54}$$

$$\mu_{in-in} = \frac{1}{M} \sum_{i=1}^{n} (k_i^{in})^2, \tag{55}$$

$$\mu_{out-out} = \frac{1}{M} \sum_{i=1}^{n} (k_i^{out})^2, \tag{56}$$

and

$$\mu_{in-out} = \mu_{out-in} = \frac{1}{M} \sum_{i=1}^{n} (k_i^{in} k_i^{out}). \tag{57}$$

Note that, in the undirected case, it can be shown that μ_{un} is equal to the average of the degrees of the target and source nodes in the edge list $\{A_e B_e\}_{e=1}^{m}$, i.e. $\mu_{un} = \frac{1}{2m}(\sum_{e=1}^{m} k_{A_e}^{un} + \sum_{e=1}^{m} k_{B_e}^{un})$. Similarly, in the directed case, given the edge list $\{A_e B_e\}_{e=1}^{M}$, it can be shown that μ_{in-in} and $\mu_{out-out}$ are respectively equal to the averages of the in-coming and out-going degrees from target and source nodes in the edge list. Mathematically, $\mu_{in} = \frac{1}{M} \sum_{e=1}^{M} k_{B_e}^{in}$ and $\mu_{out} = \frac{1}{M} \sum_{e=1}^{M} k_{A_e}^{out}$. In contrast, μ_{in-out} (μ_{out-in}) gives the average of out-going (in-coming) degrees of the target (source) nodes in the edge list. We have that $\mu_{in-out} = \frac{1}{M} \sum_{e=1}^{M} k_{A_e}^{in}$ and $\mu_{out-in} = \frac{1}{M} \sum_{e=1}^{M} k_{B_e}^{out}$.

Second, we define

$$\sigma_{un}^2 = \sum_{i=1}^{n} (k_i^{un})^3 - \frac{1}{2m} [\sum_{i} (k_i^{un})^2]^2, \tag{58}$$

$$\sigma_{in}^2 = \sum_{i=1}^{n} (k_i^{in})^3 - \frac{1}{M} [\sum_{i} (k_i^{in})^2]^2, \tag{59}$$

$$\sigma_{out}^2 = \sum_{i=1}^{n} (k_i^{out})^3 - \frac{1}{M} [\sum_{i=1}^{n} (k_i^{out})^2]^2, \tag{60}$$

$$\sigma_{in'}^2 = \sum_{i=1}^{n}(k_i^{in})^2 k_i^{out} - \frac{1}{M}[\sum_{i=1}^{n}(k_i^{in}k_i^{out}]^2, \tag{61}$$

and

$$\sigma_{out'}^2 = \sum_{i=1}^{n}(k_i^{out})^2 k_i^{in} - \frac{1}{M}[\sum_{i=1}^{n}(k_i^{in}k_i^{out}]^2. \tag{62}$$

The denominators in Eqs. (44), (45), (46), (47), (48) are respectively equal to σ_{un}^2, $\sigma_{in}\sigma_{in'}$, $\sigma_{in'}\sigma_{out'}$, $\sigma_{out}\sigma_{in}$, and $\sigma_{out}\sigma_{out'}$.

By decomposing the overall assortativity coefficient r_{bin}^{un} in Eq. (44), we obtain the local assortativity indicators. More specifically, the contribution of node i to r is

$$\rho_i^{un} = \frac{(k_i^{un})^2 k_{nn,i}^{un} - (k_i^{un})^2 \mu_{un}}{\sigma_{un}^2}. \tag{63}$$

Similarly, in the directed case, for each node i, we have four local assortativity indicators:

$$\rho_i^{in-in} = \frac{k_i^{in}[k_i^{in} * (k_{nn,i}^{in-in} - \mu_{in-out}) + k_i^{out}(k_{nn,i}^{out-in} - \mu_{in-in})]}{2\sigma_{in}\sigma_{in'}}, \tag{64}$$

$$\rho_i^{in-out} = \frac{[k_i^{in}k_i^{out}(k_{nn,i}^{out-out} + k_{nn,i}^{in-in}) - 2k_i^{in}k_i^{out} * \mu_{in-out}]}{2\sigma_{out'}\sigma_{in'}}, \tag{65}$$

$$\rho_i^{out-in} = \frac{[(k_i^{out})^2 * (k_{nn,i}^{out-in} - \mu_{in-in}) + (k_i^{in})^2(k_{nn,i}^{in-out} - \mu_{out-out})]}{2\sigma_{out}\sigma_{in}}, \tag{66}$$

$$\rho_i^{out-out} = \frac{k_i^{out}[k_i^{out} * (k_{nn,i}^{out-out} - \mu_{out-in}) + k_i^{in}(k_{nn,i}^{in-out} - \mu_{out-out})]}{2\sigma_{out}\sigma_{out'}}. \tag{67}$$