Harmonic distances, centralities and systemic stability in heterogeneous interbank networks *

Gábor Fukker †

April 15, 2017

Abstract

This paper investigates the effects of contagion in interbank lending networks. I introduce a new systemic importance measure based on the harmonic distance of Acemoglu et al. (2015b) and, motivated by their theoretical results also for concentration centrality in Acemoglu et al. (2015a), compare them to well-known centrality measures already applied in the systemic risk literature which do not take into account the structure of a contagion mechanism. I derive an explicit formula of size-adjusted harmonic distances and extend it with the usage of liquid assets for a heterogeneous banking system. The simulation results on scale-free and complete networks do not confirm that these new measures would perform better than "off-the-shelf" centralities but their performance becomes similar to the best known measures in case of averaged networks which are applied in central banking analysis. Harmonic distances and concentration centrality are also capable of identifying systemically important institutions. Their time variation is also presented in an interbank network. I also test for the scale-free property of the Hungarian interbank lending network and besides, network measures as systemic risk indicators are analyzed on Hungarian data.

1 Introduction

Though the analysis of interbank contagion and the effect of institutional failures on a financial system dates back to 2001, when the seminal paper of Eisenberg and Noe (2001) was published, the interest in this topic flared up after the global financial crisis, when AIG was bailed out due to its global systemic importance. Possible asset side channels of contagion are numerous and can be categorized into two main sets: direct effects and indirect effects. The most well-known direct effect is interbank lending while indirect effects are channelled through asset prices. Possible liability side contagion channels are different types of bank runs (Upper, 2011).

In data one can scarcely encounter interbank contagions since institutions in distress are usually bailed out or are rescued by senior creditors. To overcome this fact, empirical researchers use simulated data or assume extreme stress events to real life data.

*Majority of this work has been published as part of the MNB Working Papers series.

†National Bank of Hungary, 1054 Szabadság tér 9., Budapest, Hungary, fukkerg@mnb.hu. The author is grateful to Péter Fáykiss for his valuable comments and ideas to the research topic, to Henrik Kucsera for the thorough reading and relevant comments on the paper and to Katalin Varga for her help in the writing of Subsection 8.4 and general discussions.

Regulatory response to system-wide stress events has been the intense development of macroprudential policy as microprudential interventions were unable to mitigate risks. To see this in the case of interbank contagion, it is enough to recognize that the interbank network is unknown from an individual banks balance sheet. From the point of view of individual institutions, they are susceptible to counterparty risk. They only quantify their partners' default but do not deal with their own role in the financial system. Systemic risk analysis tackles this flaw as its objective function is the stability of the whole financial system. Macroprudential tools aim to increase the resilience of financial institutions with liquidity regulations, improved capital requirements, and further targeted measures in order to achieve higher loss absorbing capacities in the financial system.

A financial institution is referred to as systemically important (SIFI) if its default could trigger a system-wide stress. Imposing higher capital requirements on institutions based on their interbank exposures and interconnectedness is assumed to improve financial stability and decrease social costs of a banking crisis. This idea is in the spirit of the assessment methodology for SIFIs proposed by the Basel Committee on Banking Supervision. Thus interbank contagion and interconnectedness is a typical focal point of SIFIs as confirmed in the following citation:

"The difficult task before market participants, policymakers, and regulators with systemic risk responsibilaties such as the Federal Reserve is to find ways to preserve the benefits of interconnectedness in financial markets while managing the potentially harmful side effects" Yellen (2013)

The risks of interconnectedness arise mostly in case of unsecured debt contracts: though financial institutions also conclude repo transactions, these are subject to a lesser degree of counterparty risk as they are secured by collateral. Therefore, the lack of repayment does not entail as high losses as an unsecured contract. One of the first papers that deal with bank interconnectedness through unsecured debt networks of financial institutions is the toy model of Allen and Gale (2000). They find, similarly to Freixas et al. (2000), that more diversified interbank liabilities lead to a more resilient system to the default of any bank. However, others argue that an increase in the number of interconnections leads to an increase in the probability of crisis (Vivier-Lirimont, 2006). These fundamental papers aim to study the stability of the whole network but it is also a key question how a large institution influences systemic stability. As Borgatti (2005) pointed out, when choosing and applying so-called centrality measures we need to identify and investigate the process taking place in the network and the role of institutions in the network. For the quantification of network systemic importance, Battiston et al. (2012) propose a measure deepening the idea of eigenvector centrality excluding walks from the network in which one or more nodes are repeated. Soramäki and Cook (2013) also propose an algorithm for identifying systemically important institutions. Alter et al. (2015) and Fink et al. (2015) empirically find that eigenvector and Bonacich centrality are key measures of systemic importance. This is theoretically confirmed in Acemoglu et al. (2015a) where a new centrality called concentration centrality is also introduced "which captures the concentration of an agent's influence on the rest of the agents", while the same authors (Acemoglu et al., 2015b) propose a new notion of distance called harmonic distance between banks which measures the possibility of contagion: "bank closest to all others according to our harmonic distance measure that may be too-interconnected-to-fail" and "systemic importance of a financial institution is captured via its harmonic distance to other banks, suggesting that this new notion of network distance should feature in theoretically-motivated policy analyses". Alter et al. (2015) investigate centrality based capital allocations in the German banking system. Fink et al. (2015) propose a framework to measure capital losses ($BSLoss$) to the banking system as the cost of interconnectedness and find high correlation between the costs and certain centrality measures. However, the so-called SIFI scores, the official measure for assessing systemically important financial institutions, have low correlation. This underpins the investigation of centrality measures.

The goal of this paper is to measure the performance and behaviour of harmonic distance and concentration centrality by quantifying systemic importance of institutions in case of individual defaults. A default of a liquidity providing institution may also give rise to liquidity shortage in the system as a whole. I do not intend to investigate this situation; on the other hand, I would like to deal with the case when a bank triggers a cascade only if it cannot meet its obligations.

The rest of this paper is structured as follows. In Section 2 I introduce the basic notations and centrality measures that have been used to measure systemic importance in networks. In Section 3 I refresh the underlying network equilibrium model of interbank payments and its generalized form. In Section 4 I summarize the main contributions of the two papers of Acemoglu et al. (2015a); Acemoglu et al. (2015b) that are related to the measurement of systemic importance of financial institutions and as an own result I propose an explicit formula for their measure harmonic distance. In Section 5 I extend the definition of harmonic distance with more flexible scaling of heterogeneous banks. In Section 6 I elaborate the problem of weakly connected networks and harmonic distances, introduce a generation process of artificial interbank networks and present the calculation of implied losses when a bank defaults. The co-movement of these implied losses and systemic importance measures are analyzed in Section 7, while Section 8 tests for the scale-free property of the Hungarian interbank lending network, demonstrates the behaviour of these measures in time and last but not least, shows an application for systemic stress indication. An explanation of results is presented in Section 9. Section 10 concludes.

2 Basic notations and centrality measures

A network can be represented by a graph of vertices and edges $G(V, E)$. Two vertices are neighbours if there exists an edge connecting them. $N(i)$ is the set of neighbours of vertex i. I will alternatively call vertices nodes. In case of financial networks nodes are financial institutions and edges represent liability connections among them. I will represent networks in the common way: the undirected adjacency matrix contains the undirected existence of connections among institutions as indicators, directed edges are represented in the directed adjacency matrix in the same way. In the presence of edge weights, indicators are substituted by real numbers, and I obtain the undirected weighted and directed weighted matrices of financial networks.

Centrality measures are used to quantify the importance of a node in a network. Several measures have been introduced in the literature, from natural ideas to more complicated ones. In the following I summarize some of the most frequently used centralities and the connections between them.

In the undirected adjacency network represented by matrix $\mathbf{A}$, for a given node i the *degree* is the number of nodes that are linked to i, i.e. the sum of rows of the adjecency matrix:

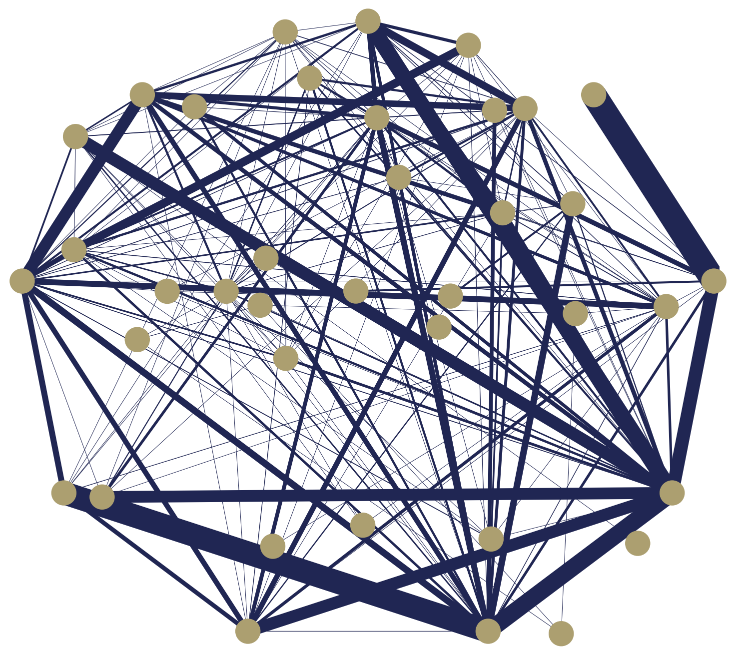

Figure 1: The aggregated Hungarian interbank lending network in 2015

$$d_i = \sum_{j=1}^{n} a_{i,j} = (\mathbf{A} \cdot \mathbf{1})_i,$$

where the multiplication by a vector of ones takes the sum of rows of the preceding matrix. In the undirected, weighted network $\mathbf{W}$ for a given node i the *weighted degree* is the sum of the weights of edges that are connected to node i,

$$w_i = \sum_{j=1}^{n} w_{i,j} = (\mathbf{W} \cdot \mathbf{1})_i.$$

Note that both degrees are the (weighted) number of steps from node i while a step is a special path of length 1. The number of steps of any length will come up at Bonacich centrality.

In the undirected network $\mathbf{W}$ the *closeness* of node i is given by the reciprocal of the maximal distance from all nodes in the network,

$$c_i = \frac{1}{\max_j d(i,j)},$$

where $d(i, j)$ denotes distance between node i and j, i.e. the minimum length of paths between them. Intuitively, the closeness of a node is high if its distance from other nodes is low, therefore it is close to all vertices in the network.

Betweenness is the number of shortest paths that contain a given node i. Paths of length 1 are excluded.

Let v_i denote the *eigenvector* centrality of node i which is implicitly defined by

$$v_i = \frac{1}{\lambda} \sum_{j \in N(i)} v_j,$$

that is, the importance of a node is proportional to the sum of importances of its neighbours. With the help of the adjacency matrix,

$$v_i = \frac{1}{\lambda} \sum_{j \in N(i)} v_j = \frac{1}{\lambda} \sum_j a_{i,j} \cdot v_j,$$

which in matrix notation becomes $\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$, thus $\mathbf{v}$ is an eigenvector of matrix $\mathbf{A}$. As one needs $\mathbf{v}$ to be elementwise positive, one chooses the eigenvector corresponding to the maximal eigenvalue. The Perron – Frobenius Theorem (see Appendix A for the exact statement) guarantees that the chosen vector is elementwise positive.

Bonacich centrality (Bonacich, 1987; Bonacich and Lloyd, 2001) is based on the idea of eigenvector centrality but has two flexibility parameters α and β,

$$b_i(\alpha, \beta) = \sum_j \alpha + \beta \cdot a_{i,j} \cdot b_j(\alpha, \beta),$$

which in matrix notation after rearranging turns into

$$\mathbf{b}(\alpha, \beta) = \alpha \cdot (\mathbf{I} - \beta \mathbf{A})^{-1} \cdot \mathbf{1}. \tag{1}$$

Note that for the special case $\alpha = 0, \beta = 1/\lambda$ in the definition Bonacich centrality is equal to eigenvector centrality and equation (1) is applicable only for $\alpha \neq 0$ and $\beta < 1$. Let $\mathbf{B} = (\mathbf{I} - \beta \mathbf{A})^{-1} = b_{i,j}$. This centrality has an important interpretation. By using the Neumann series representation of $\mathbf{B}$,

$$\mathbf{b}(\alpha, \beta) = \alpha \cdot \sum_{k=0}^{\infty} \beta^k \mathbf{A}^k \cdot \mathbf{1}.$$

If one looks at $\beta < 1$ as a probability, b_i is the expected number of paths from node i if every step has a probability β. From this point of view, Bonacich centrality is a closeness measure. Moreover, also a generalization of degree as being the expected number of paths of any length from the corresponding node. Choosing $\alpha = \beta$ is a special case also called *Katz centrality*.

Lemma 1 ((Bonacich and Lloyd, 2001)). *Eigenvector centrality is proportional to a limit of Bonacich centrality:*

$$\lim_{\beta \to \left(\frac{1}{\lambda_1}\right)-} (1 - \beta \lambda_1) \cdot \mathbf{b}(1, \beta) \propto \mathbf{v},$$

where λ_1 is the largest eigenvalue of $\mathbf{A}$.

For proof, see Appendix A.

All centralities are normed to range between 0 and 1. Degree centrality is divided by the complete network's number of edges, $n \cdot (n-1)/2$, weighted degree is divided by the sum of repayments in the whole network, closeness is between 0 and 1 by definition, betweenness is divided by the maximum number of shortest paths in a complete network, $(n-1) \cdot (n-2)/2$. Eigenvector centrality is divided by the 2-norm of the eigenvector.

At this point I refer to the paper of Alter et al. (2015) to emphasize the empirical usefulness of eigenvector centrality. They are investigating capital requirements in a way of capital reallocation in the German banking system with their main focus on interbank contagion and correlated credit exposures of financial institutions. In case of defaults, individual bankruptcy costs are proportional to total assets. The main goal is to compare two types of capital allocations. The first one focuses only on individual portfolio risk (VaR approach, benchmark), while the second one also takes into account the banks interconnectedness through interbank loans. In the latter case, the (VaR) benchmark capital requirement is reduced with a fixed fraction. The aggregate reduction is reallocated among banks according to their centrality (several centralities are tested) in the interbank network. Their main contribution is that they find *adjacency eigenvector* centrality the best measure as its expected bankruptcy cost is about 14 percent lower than in the benchmark case for the optimal reduction.

Fink et al. (2015) in their stress test setup, model the credit quality channel of interbank contagion. They mimic the capital management of a bank by regressing its Tier 1 capital ratio on the debtors PDs. Also, the change in the creditor banks capital ratio changes its own PD. The model is very flexible in defining the default event and balance sheet losses are sensitive to smaller shocks of a bank. Their related finding is that the measure of total balance sheet losses during a contagion mechanism highly correlates with eigenvector-like measure Bonacich centrality.

3 Model setup

In this section I introduce the theoretical framework of Acemoglu et al. (2015a); Acemoglu et al. (2015b), which are fundamental papers in the understanding of how time-invariant financial networks work. I start from a more general representation of the so called generalized economic networks, the special case of which is financial contagion in interbank networks.

3.1 Generalized economic networks

Let N be an economy of n agents $\{1, 2, \ldots, n\}$. An agent i has a state x_i ($x_i \in \mathbb{R}, i \in N$) which can be output, investment or liabilities. For an f continuous and increasing function let

$$x_i = f\left(\sum_{j=1}^{n} w_{i,j} \cdot x_j + \varepsilon_i\right), \tag{2}$$

which shows that states are interdependent due to strategic reasons, technology constraints or contractual obligations. They call f the interaction function which captures the interaction between the agents of the economy, ε_i are i.i.d. agent-level shocks with mean 0 and variance σ^2. $w_{i,j}$ is the sensitivity of agent i to the state of agent j. Let $\mathbf{W} = \{w_{i,j}\}_{i,j=1}^{n}$ denote the matrix of sensitivities. They naturally assume that rows sum up to 1, i.e. $\sum_{j=1}^{n} w_{i,j} = 1$.

The economy is said to be in an *equilibrium state* if for a given realization of shocks $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ equation (2) holds simultaneously for all i. They prove that equilibrium exists and is unique with the method of fixed point iteration.

The *macro state* of the economy is $y = g\left(h(x_1) + h(x_2) + \ldots + h(x_n)\right)$, where g is the aggregation function. $(g, h : \mathbb{R} \to \mathbb{R})$

3.2 Financial contagion

The phenomenon of financial contagion is a special case of the above general setup (next to several others in Acemoglu et al. (2015a)). Agents are banks and the connections are unsecured debt contracts. For the sake of simplicity each bank i's total claim size is ξ and the weight of it on bank j is $w_{i,j}$, therefore the claim on bank j is $w_{i,j} \cdot \xi$. Assume that $\mathbf{W}$ is doubly stochastic, that is $\sum_{j=1}^{n} w_{i,j} = \sum_{i=1}^{n} w_{i,j} = 1$. This ensures equal total claim size for all banks i.

After the realization of a shock ε_i, banks have to pay back their loans. A bank defaults if it can't meet its liabilities. If $y_{i,j}$ is the repayment of a loan from j to i, the cash flow of bank i is given by

$$c_i = e_i + \sum_{j=1}^{n} y_{i,j} + \varepsilon_i,$$

where e is outside assets. If $c_i \geq \xi$, bank i can meet its liabilities and therefore $y_{j,i} = w_{j,i} \cdot \xi$ for all j. Otherwise, if $0 < c_i < \xi$ the bank defaults and the creditors are repaid only c_i proportionally, $y_{j,i} = w_{j,i} \cdot c_i$. Thus an implicit equation can be set up:

$$y_{j,i} = \left[\min\left\{w_{j,i} \cdot \left(e_i + \sum_{k=1}^{n} y_{i,k} + \varepsilon_i\right), w_{j,i} \cdot \xi\right\}\right]^{+},$$

where $[\cdot]^{+} = \max\{\cdot, 0\}$. Summing up all over j and defining $y_i = \sum_{j=1}^{n} y_{j,i}$ as the total repayments of bank i one gets

$$y_i = \left[\min\left\{\left(e_i + \sum_{k=1}^{n} y_{i,k} + \varepsilon_i\right), \xi\right\}\right]^{+} = \left[\min\left\{\left(e_i + \sum_{k=1}^{n} w_{i,k} \cdot y_k + \varepsilon_i\right), \xi\right\}\right]^{+}.$$

From this equation system it is easy to see that for the interaction function $f(x) = [\min\{x+e, \xi\}]^{+}$, financial contagion is a special case of generalized economic networks.

4 Model implications

This section contains some main results of Acemoglu et al. (2015a); Acemoglu et al. (2015b) about the effect of shocks on agents and macro state, and a new measure called harmonic distance for financial contagion. I also present some additional theoretical results on a more general size-adjusted harmonic distance.

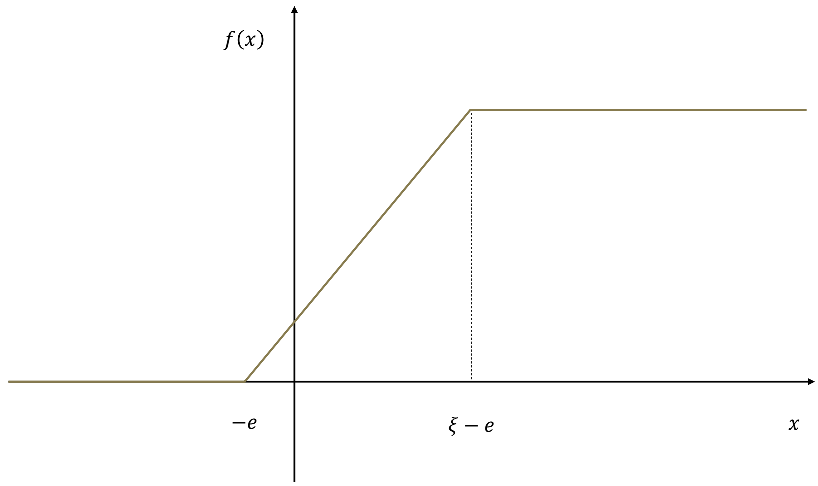

Figure 2: The graph of $f(x) = [\min\{x + e, \xi\}]^+$ for positive liquid assets e

4.1 Series expansion

It is natural question how a shock to agent p affects the state of agent i, therefore one is interested in the value of $\frac{\partial x_i}{\partial \varepsilon_p}$ The authors examine the case of small shocks, smooth interaction and aggregation functions then apply Taylor expansion around the point 0. By simply differentiating equation (2),

$$\frac{\partial x_i}{\partial \varepsilon_p} = f'\left(\sum_{j=1}^{n} w_{i,j} \cdot x_j + \varepsilon_i\right) \cdot \left(\sum_{j=1}^{n} w_{i,j} \cdot \frac{\partial x_i}{\partial \varepsilon_p} + \mathbf{1}\{p = i\}\right).$$

Evaluating at $\boldsymbol{\varepsilon}^T = (\varepsilon_1, \ldots, \varepsilon_n) = \mathbf{0}$ and rearranging the equation leads to the linear approximation

$$\left.\frac{\partial \mathbf{x}}{\partial \varepsilon_p}\right|_{\boldsymbol{\varepsilon}=\mathbf{0}} = f'(0) \cdot (\mathbf{I} - f'(0) \cdot \mathbf{W})^{-1} \cdot \mathbf{e}_p, \tag{3}$$

where $\mathbf{e_p}$ is the vector of zeros with one at the pth coordinate. Note that $(\mathbf{I} - f'(0) \cdot \mathbf{W})^{-1}$ is a very similar matrix to that in equation (1) with $\alpha = \beta = f'(0)$, $\mathbf{A} = \mathbf{W}$ and thus equation (3) is

$$\left.\frac{\partial \mathbf{x}}{\partial \varepsilon_p}\right|_{\boldsymbol{\varepsilon}=\mathbf{0}} = \alpha \cdot \mathbf{B} \cdot \mathbf{e}_p,$$

or equivalently

$$\left.\frac{\partial x_i}{\partial \varepsilon_p}\right|_{\boldsymbol{\varepsilon}=\mathbf{0}} = \alpha \cdot b_{i,p}.$$

Equation (3) is a simple differential equation system for vector $\mathbf{x}$. One can check that the solution is

$$x_i = \alpha \cdot \sum_{p=1}^{n} b_{i,p} \cdot \varepsilon_p \;\; \forall i$$

or $\mathbf{x} = \alpha \cdot \mathbf{B} \cdot \boldsymbol{\varepsilon} = f'(0) \cdot (\mathbf{I} - f'(0) \cdot \mathbf{W})^{-1} \cdot \boldsymbol{\varepsilon}$.

Further question is the macro state of the economy depending on a shock to agent p. By differentiating y one similarly gets

$$\frac{\partial y}{\partial \varepsilon_p} = g'(h(x_1) + \ldots + h(x_n)) \cdot \sum_{i=1}^{n} h'(x_i) \cdot \frac{\partial x_i}{\partial \varepsilon_p},$$

which after evaluating at $\boldsymbol{\varepsilon} = \mathbf{0}$ and using equation (3) yields

$$\left.\frac{\partial y}{\partial \varepsilon_p}\right|_{\boldsymbol{\varepsilon}=\mathbf{0}} = g'(0) \cdot h'(0) \sum_{i=1}^{n} \alpha \cdot b_{i,p}. \tag{4}$$

Note that $\sum_{i=1}^{n} b_{i,p}$ is the sum of the pth column of matrix $\mathbf{B}$. The solution of differential equations (4) gives the linear approximation of the macro state:

$$y = f'(0) \cdot g'(0) \cdot h'(0) \sum_{p=1}^{n} \sum_{i=1}^{n} b_{i,p} \cdot \varepsilon_p.$$

At this point one is able to shed light on the optimality of adjacency eigenvector centrality in the empirical work of Alter et al. (2015). Letting ($\mathbf{A}^* = \mathbf{A}^T$) and setting $f(x) = g(x) = h(x) = x$ (meaning $\alpha = \beta = 1$), the macro state of the economy is linearly approximated by

$$y = \sum_{p=1}^{n} \sum_{i=1}^{n} b_{i,p} \cdot \varepsilon_p.$$

Therefore the marginal effect on the macro state of a small shock to bank p is

$$\frac{\partial y}{\partial \varepsilon_p} = \sum_{i=1}^{n} b_{i,p} = (\mathbf{B}^T \cdot \mathbf{1})_p$$

Furthermore, if $\mathbf{A}^* = \mathbf{A}^T$ then $\mathbf{B}^* = \mathbf{B}^T$ and what one gets is exactly the Bonacich centrality vector $\mathbf{b}(\alpha, \beta)$ for $\alpha = \beta = 1$. Though $\mathbf{b}(\alpha, \beta)$ is not defined for $\beta = 1$, a limit of it exists as shown below with the help of Lemma 2.

Lemma 2. *The largest eigenvalue of a stochastic matrix is 1.*

For proof, see Appendix A.

This Lemma together with Lemma 1 leads to the fact that

$$\lim_{\beta \to 1-} (1 - \beta) \cdot \mathbf{b}(1, \beta) \propto \mathbf{v},$$

i.e. for a doubly stochastic matrix eigenvector is indeed a quasi-optimal measure of macro state (or in the case of capital allocation, aggregate loss) in the presence of small shocks and linear interaction function. Moreover, there is theoretical evidence (Acemoglu et al., 2015a) that for any interaction function, an agent with higher Bonacich centrality may propagate negative shocks more extensively which is closely related to the identification and assessment of systemically important institutions. This result holds only for the first order approximation.

If one takes the second order approximation of the macro state (similarly to the first order Taylor expansion above), another centrality measure can catch the extent of shock propagation ability, the standard deviation of the pth column of matrix $\mathbf{B}$, which they call *concentration centrality*:

$$con_p = \text{stdev}(b_{1,p} \dots, b_{n,p}).$$

Node i is said to be systemically more important than j if $y_{(i)} > y_{(j)}$, where $y_{(i)}$ denotes the macro state when i is hit with a negative shock. Precisely, in case of concave interaction function, institution i is systemically more important than j if $con_i > con_j$.

4.2 Size-adjusted harmonic distance

In fact, contagion interactions are not linear and a second order approximation may not be eligible (indeed, the special interaction function depicted on Figure 2 cannot be approximated by a quadratic function). One have to rely on the special interaction function defined in Subsection 3.2. Acemoglu et al. (2015b) deeply analyze stability in financial networks in case of identical institutions. They show that as long as the size of a negative shock is below a certain threshold, less fragility is obtained by a more equal distribution of interbank liabilities. This phenomenon changes if the size of shock is above this threshold. On the contrary, they show that a more equal distribution may increase the number of defaults. The authors present a new notion of distance called *harmonic distance* building on mean hitting time, an existing concept of the theory random walks on graphs, or more generally Markov chains. Mean hitting time of a random walk on a graph from vertex i to j is the expected number of steps of a random walk from i until it reaches j. It is easy to see that the recursive equation (5) for harmonic distance is exactly the same concept for financial networks. For further knowledge on random walks on graphs or Markov chains, see the survey of Lovász (1993) or the book of Levin et al. (2009). Note that harmonic distance is not a distance in a mathematical way because it is not symmetric.

The *size-adjusted harmonic distance* from bank i to bank j is given by

$$h_{i,j} = \theta_i + \sum_{k \neq j} \left(\frac{y_{i,k}}{y_k} \right) \cdot h_{k,j},$$

with the restriction that $h_{i,i} = 0$, where $y_{i,k}$ is again the liability of i to k, while y_i is total repayments of i, θ_i is the scaling factor of liabilities and liquid assets such that $\theta_i \cdot y = y_i$. By choosing $y = 1$, the definition becomes

$$h_{i,j} = y_i + \sum_{k \neq j} \left(\frac{y_{i,k}}{y_k} \right) \cdot h_{k,j}, \tag{5}$$

which representation will suggest an extension of this definition in the next section. The intuitive meaning of this definition is the following. In case of the default of bank i, losses of y_i are generated to its neighbours. One can think of $\left(\frac{y_{i,k}}{y_k} \right)$ as the probability of shock propagation towards node k, and from k to j the distance is recursively defined.

If I set $y_i = 1$ for all i, harmonic distance becomes exactly the same as the mean hitting time of a Markov chain. It is shown that if a bank j defaults all other banks also default if and only if their harmonic distance from j is smaller than a specific threshold (Proposition 8). The counterpart of this statement (Proposition 12) is that if $y_i \neq 1$, all other banks other than j also default if and only if their harmonic distance from j is smaller than the given threshold. *This implies that banks that are closer in harmonic distance to the defaulted bank are more vulnerable to distress.* Since the original definition is recursive, it is not useful for empirical investigations, the following proposition gives the explicit formula of harmonic distance of any two banks after the introduction of some technical notations.

Let $\mathbf{Y} = \{y_{i,j}\}_{i,j=1}^n = \{y_i\}_{i=1}^n$, i.e. the elements of the ith row are the total liabilities of i. Denote the probability of shock propagation matrix as $\mathbf{Q} = \{q_{i,k}\}_{i,k=1}^n = \left\{\frac{y_{i,k}}{y_k}\right\}_{i,k=1}^n$. Let $\mathbf{M} = -(\sum_{i=1}^n y_i) \cdot \left(\mathbf{I} - \mathbf{Q} + \frac{1}{\sum_{i=1}^n y_i} \cdot \mathbf{Y}\right)^{-1}$ and the matrix $\mathbf{D}$ the columns of which are $\mathbf{d}_i = \left[-\mathbf{v}_0 \cdot \frac{m_{i,i}}{v_{0,i}}\right]$, where $\mathbf{v}_0$ is the eigenvector corresponding to eigenvalue 0 of matrix $(\mathbf{I} - \mathbf{Q})$. $\mathbf{D}$ will be responsible for the diagonal restriction of harmonic distances.

Proposition 1. *The matrix* $\mathbf{H} = \{h_{i,j}\}_{i,j=1}^n$ *of pairwise size-adjusted harmonic distances is explicitly given by*

$$\mathbf{H} = -\left(\sum_{i=1}^n y_i\right) \cdot \left(\mathbf{I} - \mathbf{Q} + \frac{1}{\sum_{i=1}^n y_i} \cdot \mathbf{Y}\right)^{-1} + \mathbf{D},$$

if and only if there is no non-borrowing node in the directed network.

For proof, see Appendix A.

5 Extended harmonic distance

This section presents a new extension of the original definition of harmonic distance. Results and calculations remain very similar. The problem with the original definition is that liquid assets are scaled with the same scaling factor θ_i as total liabilities for bank i. Having seen in the definition of equation (5), it is straightforward to take a further step in the usage of the scaling factor. Thus I define the *extended harmonic distance* as

$$h_{i,j} = e_i + \sum_{k \neq j} \left(\frac{y_{i,k}}{y_k}\right) \cdot h_{k,j}, \tag{6}$$

with the restriction that $h_{i,i} = 0$ for all i. Now e_i is the liquidity surplus or liquid assets of bank i. With the help of this definition, the following statement holds the proof of which closely follows the proof of Proposition 12 in Acemoglu et al. (2015b).

Proposition 2. *Suppose that bank j is hit with a negative shock $\varepsilon > \sum_{i=1}^n e_i$. Then*

1. *bank j defaults*
2. *all other banks also default if and only if $h_{i,j} < y_i$ for all i.*

For proof, see Appendix A.

This proposition very similarly states that if bank j is hit with a large enough shock it immediately defaults. Its default triggers a cascade of defaults to all other banks in the network if and only if the harmonic distances of the others is below their total liabilities. This version of the statement is quite clear and intuitive but it still remains an open question what happens when these specific conditions do not hold exactly.

Note that Proposition 1 remains valid for the calculation of extended harmonic distance by changing matrix $\mathbf{Y}$ with $\mathbf{E}$ where the ith row of $\mathbf{E}$ is $(e_i, \ldots, e_i)$.

Proposition 3. *The matrix* $\mathbf{H} = \{h_{i,j}\}_{i,j=1}^n$ *of pairwise size-adjusted harmonic distances is explicitly given by*

$$\mathbf{H} = -\left(\sum_{i=1}^n e_i\right) \cdot \left(\mathbf{I} - \mathbf{Q} + \frac{1}{\sum_{i=1}^n e_i} \cdot \mathbf{E}\right)^{-1} + \mathbf{D},$$

if and only if there is no non-borrowing node in the directed network.

For proof, see Appendix A.

Now one is able to calculate pairwise harmonic distances and is motivated by Proposition 2 to estimate systemic importance with the help of these distances. Though it is shown in Acemoglu et al. (2015b) that the sum of harmonic distances from bank i to all other nodes is constant, it only holds when total repayments are identical. Given this and the fact that banks that are closer in harmonic distance to the defaulted bank are more vulnerable it is straightforward to calculate $\sum_{i=1}^n h_{i,j}$ as a measure of network systemic importance of bank j which is the sum of the jth column of matrix $\mathbf{H}$. One could argue that any change in $\sum_{i=1}^n h_{i,j}$ is only due to the changes of institution sizes. One can simply check that networks with identical total transaction volumes produce different harmonic distances, furthermore, when total liabilities of nodes are heterogenized in the same way but with different transaction structures, $\sum_{i=1}^n h_{i,j}$ behaves differently again.

One can see from the definition that increasing the size of liabilities leads to an increase in harmonic distances. However, one can control for this kind of change with the help of Proposition 2. It states that all other banks than j default if and only if $h_{i,j} < y_i$ for all i. To get rid of institution sizes, these conditions can be rearranged to $\frac{h_{i,j}}{y_i} < 1$ for all i. This transformation will also be applied in Subsection 8.3 when I analyze the time evolution of harmonic distances in a real life financial network.

6 Artificial interbank networks

This section is related to the interbank networks that I will compare systemic importance measures on. After disclosing some technical difficulties and their solution in Subsection 6.1, I turn to the generation process of artificial interbank networks in Subsection 6.2. In the lack of empirical liquidity defaults, I generate artificial payment networks having similar network topology to an interbank network. I follow the method of Soramäki and Cook (2013). The reason one chooses to generate large numbers of networks from the same family of graphs is that robustness of results is guaranteed only in this case, while real life networks change in time. After obtaining a network, I calculate the payment equilibrium and induce individual defaults in the system.

6.1 Dealing with weakly connected networks, absorbing and transient nodes

I restrict the analysis to weakly connected networks, where any node can be reached on an undirected path from any node. In the proof of Proposition 1 and 3, I used that the sum of a column of $\mathbf{Q}$ is 1 but it does not hold if there is an absorbing state because the elements of the corresponding column are 0. This also motivates the analysis of absorbing and transient nodes separately. Two nodes are *strongly connected* if they are reachable from one another on a directed path. In the strongly connected component, any two nodes are strongly connected. In this subsection I assume that there is only one strongly connected component (SCC). If there exists a node which cannot be reached from the strongly connected component but can reach the SCC, it is said to be *transient*. Nodes that can be reached from the SCC nodes are called *absorbing* nodes. It is easy to verify that absorbing nodes cannot reach transient nodes otherwise absorbing nodes would also be strongly connected. Furthermore, transient nodes may reach absorbing nodes directly. See Figure 3 for demonstration.

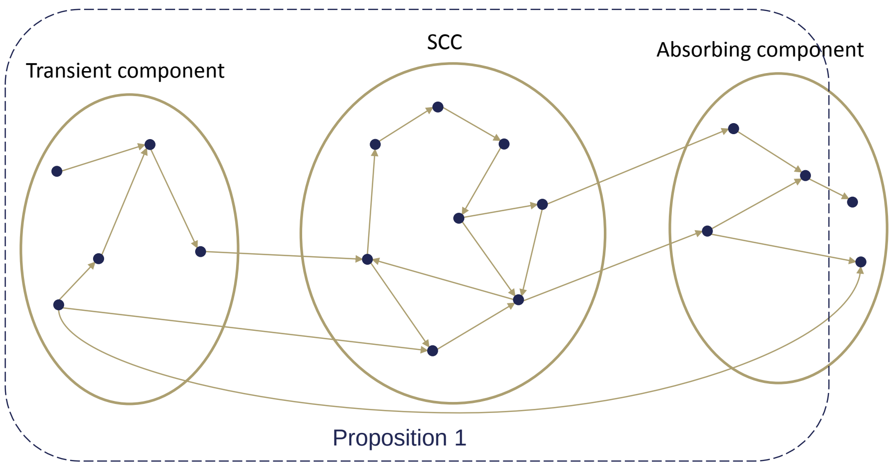

Figure 3: A decomposed weakly connected network and the usefulness (dashed blue) of Proposition 1 and 3

From the above considerations, it is clear that

$$y_{transient,SCC} = y_{transient,absorbing} = y_{SCC,absorbing} = 0.$$

To overcome the usefulness of Proposition 1 and stay between the strongly connectedness assumptions of Acemoglu et al., I add virtual payments of 1 unity of money to all possible directed edges, that is, I add the adjacency matrix of a complete directed network to the original payment network, i.e. $y'_{i,j} = y_{i,j} + 1$ for all i, j. This operation is rather technical because the ratio of virtual payments to real payments is below 10^{-9} therefore contagion will not take effect through virtual edges and strictly viewing equation (5), only the negligible virtual outpayments of absorbing nodes will have effect on harmonic distances.

6.2 An extended Barabási – Albert network algorithm

Soramäki and Cook (2013) use an extended version of preferential attachment algorithm by Albert and Barabási (2002). While the original algorithm attaches new edges to the network undirectedly, one needs directed edges and edge weights also. Note that the algorithm in Soramäki and Cook (2013) does not guarantee that newly added nodes are connecting to the existing network and therefore at termination the number of borrowing or lending banks can be much smaller than the desired number, n. I modify their algorithm so that it terminates only if I have the satisfactory number of active nodes in the network. In the following pseudo-code, n is the total desired number of banks, n_0 is the initial number of banks in the network. α is the preferential attachment parameter which can be interpreted as a kind of gravity parameter: a higher value of α increases the likelihood of connecting to already existing nodes having more connections. m is the number of edges attached at an iteration step. h_i will be the "strength" of node i which determines the probability of being selected as an endpoint of an edge.

In words, the algorithm works as follows. It starts with n_0 initial banks in the network, adds a new node at every iteration step and connects m directed edges by selecting the starting point (sender) and ending point (receiver) of the directed edge (payment) with probability proportional to the existing nodes' relative strength in the network. The algorithm terminates if there are n active banks in the network sending or receiving payments. It is possible that there will be more than 1 connected components.

`FOR` $i = 1..n_0$ (add initial banks/nodes)
    `SET` $h_i = 1$
`END FOR`
    `SET` $active = 0$ (initial number of active banks in the network)
    `SET` $k = n_0 + 1$ (first new bank)
`WHILE` $active < n$
    `FOR` $l = 1..m$ (average number of payments per bank)
        `SELECT` random sender $i \in \{1, \ldots, k\}$ such that bank i has the probability $\frac{h_i}{\sum_l h_l}$ of
        `SET` $h_i = h_i + \alpha$ (update preferential attachment strength)
        `SELECT` random receiver $j \in \{1, \ldots, k\}$ such that bank j has the probability $\frac{h_j}{\sum_l h_l}$ of being selected as recipient of the payment
        `SET` $h_j = h_j + \alpha$ (update preferential attachment strength)
        `SET` $y_{j,i} = y_{j,i} + 1$ (create payment/link)
    `END FOR`
    `IF` $k \leq n$ `SET` $h_k = 1$ `AND SET` $k = k + 1$ (create new bank/node)
    `SET` *active* as the number of nodes sending or receiving any payments
`END WHILE`

Furthermore, I need edge weights, i.e. values of repayments. Following again Soramäki and Cook (2013), I set values proportionally to the minimum of in-degree and out-degree of an edge and the number of payments $y_{j,i}$ (which are understood as multiple edges in the network) multiplied by a random variable drawn from a log-normal distribution:

$$edgeweight = \min(indegree, outdegree) \cdot \exp(N(0,1))$$

I also need an operating cash flow for all banks in the network high enough so that there is no default in the baseline payment equilibrium. The cash flow vector $\mathbf{e} = c \cdot (\mathbf{y} - \mathbf{Q} \cdot \mathbf{y}), c \geq 1$ is a suitable choice as the incoming payments $\mathbf{Q} \cdot \mathbf{y}$ together with liquid assets $\mathbf{e}$ will be at least the amount of outgoing payments $\mathbf{y}$.

6.3 Complete networks

To cross-check with BA network results, I also examined complete networks and set edge weights in a randomized manner similar to the BA network:

$$edgeweight = \lambda \cdot \exp(N(0,1)).$$

6.4 Payment equilibrium

The payment equilibrium vector is the vector of outgoing payments that satisfies

$$\mathbf{x}^* = [\min\{\mathbf{Q}\mathbf{x}^* + \mathbf{e}, \mathbf{y}\}]^+ ,$$

where $\mathbf{x}^* \in \mathcal{H} = \prod_{i=1}^n [0, y_i]$ is the Cartesian product of closed intervals. Defining the mapping $\Phi : \mathcal{H} \to \mathcal{H}$,

$$\Phi(\mathbf{x}^*) = [\min\{\mathbf{Q}\mathbf{x}^* + \mathbf{e}, \mathbf{y}\}]^+ ,$$

where the minimum operation (and also the maximum below) is element-wise in a vector. The payment equilibrium or clearing vector is $\mathbf{x}^* = \Phi(\mathbf{x}^*)$ and individual payments are given by $x^*_{i,j} = q_{i,j} \cdot x^*_j$. This payment equilibrium is exactly the same as in Eisenberg and Noe (2001), therefore one is able to implement their fictitious default algorithm which is more deeply analyzed and extended with fire sales in Cecchetti et al. (2016). Following them, there exists a unique greatest clearing vector $\mathbf{x}^*$, that is, $\mathbf{x}^* = \Phi(\mathbf{x}^*)$ and if $\mathbf{x} = \Phi(\mathbf{x})$ then $\mathbf{x}^* \geq \mathbf{x}$. Let us denote the default set

$$D(\mathbf{x}) = \{i \in 1 \dots n : \Phi(\mathbf{x})_i < y_i\}$$

and define the matrix

$$\Lambda(\mathbf{x})_{i,j} = \begin{cases} 1 & \text{if } i = j \text{ and } i \in D(\mathbf{x}) \\ 0, & \text{otherwise.} \end{cases}$$

With these notations, further let

$$F_{x'}(\mathbf{x}) = \Lambda(\mathbf{x}')(\min\{\mathbf{y}, \max\{0, \mathbf{Q} \cdot (\Lambda(\mathbf{x}')\mathbf{x} + (\mathbf{I} - \Lambda(\mathbf{x}')) \cdot \mathbf{y}) + \mathbf{e}\}\}) + (\mathbf{I} - \Lambda(\mathbf{x}')) \cdot \mathbf{y}.$$

One can check that $F_{\mathbf{x}}(\mathbf{x}) = \Phi(\mathbf{x})$ for all $\mathbf{x}$. The algorithm works as follows. Define the sequence $\mathbf{y}_0 = \mathbf{y}, \mathbf{y}_j = f(\mathbf{y}_{j-1})$, where $f(\mathbf{y}_{j-1})$ is a fixed point of $F_{\mathbf{y}_{j-1}}(\cdot)$. This iteration terminates in at most n iterations and $\mathbf{y}_n$ becomes the greatest clearing vector $\mathbf{x}^*$ for network $(\mathbf{Y}, \mathbf{e})$. At every iteration step I am looking for the fixed point of $F_{\mathbf{x}'}(\cdot)$, say a function $f(\mathbf{x}')$ for which $F_{\mathbf{x}'}(f(\mathbf{x}')) = f(\mathbf{x}')$. The fixed point of $F_{\mathbf{y}_{j-1}}(\cdot)$ is calculated by another iteration: let $\mathbf{Y}_0 = \mathbf{y}$, $\mathbf{Y}_n = F_{\mathbf{y}_{j-1}}(\mathbf{Y}_{n-1})$ for $n \geq 1$. $\{\mathbf{Y}_n\}$ converges to $f(\mathbf{y}_{j-1})$, the iteration is terminated if $\|F_{\mathbf{y}_{j-1}}(\mathbf{Y}_n) - \mathbf{Y}_n\|_2 < \varepsilon$ for a predefined tolerance level ε. It is clear that if $x_j^* < y_j$ then bank j is insolvent and one can also differentiate fundamental and contagious defaults; bank i suffers fundamental default if

$$\sum_{j=1}^{n} q_{i,j} \cdot y_j + e_i - y_i < 0$$

and it suffers contagious default if

$$\sum_{j=1}^{n} q_{i,j} \cdot y_j + e_i - y_i \geq 0,$$

but $\sum_{j=1}^{n} q_{i,j} \cdot x_j^* + e_i - y_i < 0$.

In our setup, the initial default of a bank i deletes its interbank liabilities (outgoing payments) from the network, i.e. the ith column is deleted from matrix $\{y_{i,j}\}_{i,j=1}^{n}$. To measure the effect of this default in the network I simulate contagion with the method above and get the equilibrium payment vector $\mathbf{x}^*$. Thus compared to the original vector of liabilities $\mathbf{y}$, the overall loss induced in the financial system by the default of bank j is the sum of the elements of $\mathbf{y} - \mathbf{x}^*$,

$$losses_j = \sum_{i=1}^{n} (y_i - x_i^*).$$

This amount is our benchmark measure of systemic importance of a bank. The number of defaults is given by $defaults = \sum_{i=1}^{n} \mathbf{1}\{y_i > x_i^*\}$.

7 Application to systemically important institutions

The first straightforward question that comes to one's mind is that how well harmonic distances and concentration centrality describe the systemic importance of individual financial institutions. This question is closely related to the application of centrality measures in the systemic risk literature and may be interpreted as a cross-section analysis of measures under consideration.

This section summarizes the simulation results on our theoretically established measures compared to some "off-the-shelf" measures. The analysis was carried out as follows. I generated a given number of networks with fixed parameters preferential attachment, number of banks, number of initial banks, number of payments and cash flow parameter. I set the cash flow vector (liquid assets) so that there is no contagion. After this, I erased outgoing payments for every node one-by-one and ran the Eisenberg - Noe algorithm and calculated $\mathbf{y} - \mathbf{x}^*$ and the above defined measure *losses* for every initial default. Then I calculated the importance measures. In case of BA networks and harmonic distances, I applied the "trick" of Subsection 6.1, i.e. $y'_{i,j} = y_{i,j} + 1$ for all i, j.

7.1 Results on BA networks

7.1.1 Observed behaviour

Our first remark is that harmonic distances range between extremely large scales compared to centrality measures and losses also. Furthermore, important nodes have low harmonic distances while their centralities are large. Based on these experiences, I set the importance measure to $\frac{1}{\sum_{i=1}^{n} h_{i,j}}$, see Figures 4 and 5.

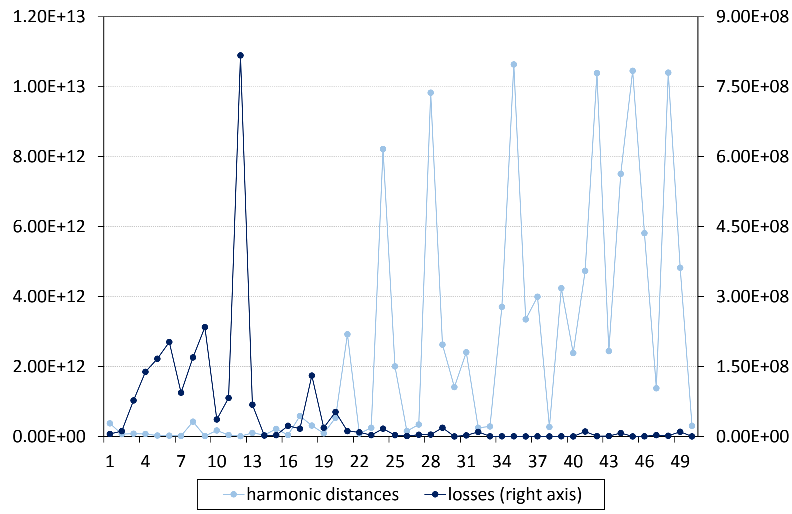

Figure 4: The sum of harmonic distances and implied losses of nodes in a BA network with parameters $n = 50$, $n_0 = 5$, $m = 4$, $\alpha = 0.1$

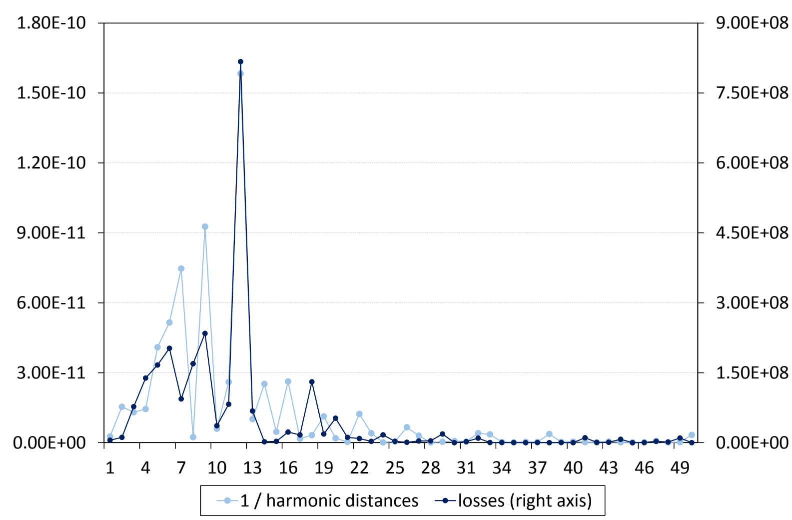

Figure 5: The reciprocal of sum of harmonic distances and implied losses of nodes in a BA network with parameters $n = 50$, $n_0 = 5$, $m = 4$, $\alpha = 0.1$

In the following, I generated 1000 networks for fixed parameter set $(\alpha, n = 50, n_0 = 5, m = 4, c)$, calculated systemic importance measures and ran the payment equilibrium algorithm to obtain losses implied by the failure of any node. Table 1 contains the average correlations of importance measures with normed losses

$$losses = \sum_{i=1}^{n} \frac{(y_i - x_i^*)}{\max_{i,j} y_{i,j}}$$

and the standard deviations of correlations.

One can conclude that though our suggested harmonic distances perform better than the original harmonic distances, they are still weaker than classical weighted degree and eigenvector centralities. Both harmonic measures, Bonacich and concentration centrality's correlation with losses have really high standard deviation. Except for closeness and betweenness, all measures performance improve by the increase of preferential attachment parameter α. This is important in the sense that the structure of network is a key issue in the analysis. One can note that results do not change significantly along parameter c which may seem surprising: banks with slightly higher amount of liquid assets have similar loss patterns when an institution fails. As an additional check, in the following experiments I generated 10 000 BA networks for fixed parameter set $(\alpha, n = 50, n_0 = 5, m = 4)$ while $c \in (1, 3)$ was uniformly random for each bank individually. See results in Table 2. Results are similar to the previous ones.

Rank correlations are available in Appendix B. It is important to note that rank correlations are not the proper measures to quantify the linear co-movement of variables. Though they are higher and deviations in case of harmonic distances are significantly lower. These results highlight that harmonic distances are useful for identifying and ranking systemic institutions, but loss patterns might differ. Bonacich and concentration centralities are still weaker than extended harmonic distances.

7.1.2 "Mean" behaviour

In the previous subsection we have seen that both harmonic distances correlation to losses have high standard deviation. Now I investigate this behaviour when a large number of network measures are averaged. This could be interpreted as an expected behaviour of systemic importance measures where expectation is through random networks. I repeat calculations, but this time all measures and losses will be averaged along networks for a given parameter set. More formally, if $c_{i,t}$ is a centrality of bank i in network t, then the average centrality of bank i will be $\sum_{t=1}^{T} \frac{c_{i,t}}{T}$, the method is the same for implied losses. The averaging method can be used if one would like to take a usual interbank network when imposing a capital buffer for mitigating contagion risks. This is the method that the MNB (the central bank of Hungary) applies to analyze the systemic importance of financial institutions in interbank networks. Weekly network measures are averaged in a yearly time horizon to indentify important participators in interbank markets.

The results in Table 3 show that harmonic distances become an appropriate measure of systemic importance along with weighted degree, eigenvector, Bonacich and concentration centrality, though for higher preferential attachment parameters the performance of harmonic distances decreases.

7.2 Results on complete networks

As importance measures are defined not specially for BA networks, it is reasonable to compare them on other network structures. To cross-check the results on BA networks I ran the algorithm

harmonic distances:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.406	0.215	0.423	0.219	0.410	0.215
0.2	0.524	0.229	0.529	0.226	0.531	0.236
0.4	0.665	0.223	0.675	0.223	0.669	0.236
0.6	0.703	0.215	0.721	0.220	0.733	0.227

extended harmonic distances:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.443	0.229	0.460	0.236	0.443	0.227
0.2	0.579	0.247	0.587	0.242	0.587	0.252
0.4	0.733	0.224	0.738	0.226	0.732	0.239
0.6	0.773	0.207	0.778	0.215	0.794	0.217

weighted degree:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.783	0.063	0.793	0.070	0.796	0.067
0.2	0.805	0.071	0.816	0.073	0.817	0.072
0.4	0.839	0.076	0.848	0.077	0.846	0.082
0.6	0.849	0.079	0.861	0.080	0.866	0.080

eigenvector:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.746	0.078	0.757	0.078	0.757	0.078
0.2	0.782	0.076	0.792	0.079	0.792	0.079
0.4	0.821	0.080	0.830	0.079	0.831	0.084
0.6	0.836	0.084	0.843	0.086	0.849	0.083

Bonacich:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.397	0.188	0.380	0.187	0.375	0.187
0.2	0.466	0.212	0.450	0.192	0.451	0.198
0.4	0.538	0.219	0.511	0.208	0.505	0.214
0.6	0.505	0.214	0.509	0.225	0.503	0.221

concentration:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.386	0.193	0.369	0.193	0.364	0.192
0.2	0.456	0.216	0.441	0.198	0.441	0.201
0.4	0.527	0.222	0.502	0.211	0.490	0.219
0.6	0.490	0.219	0.493	0.225	0.489	0.220

...

closeness:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.431	0.095	0.422	0.096	0.425	0.092
0.2	0.403	0.100	0.402	0.101	0.401	0.102
0.4	0.364	0.117	0.355	0.119	0.355	0.120
0.6	0.316	0.128	0.314	0.130	0.315	0.128

betweenness:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.373	0.168	0.369	0.162	0.367	0.162
0.2	0.345	0.168	0.357	0.173	0.356	0.177
0.4	0.300	0.182	0.289	0.183	0.292	0.189
0.6	0.230	0.188	0.232	0.191	0.234	0.188

Table 1: Average correlation of centrality measures and losses generated by the failure of single nodes and standard deviation of correlations.

harmonic distances

α	avg.corr.	std.dev.
0.1	0.416	0.217
0.2	0.540	0.234
0.4	0.666	0.232
0.6	0.722	0.223

extended harmonic distances

α	avg.corr.	std.dev.
0.1	0.450	0.231
0.2	0.591	0.248
0.4	0.724	0.235
0.6	0.778	0.219

weighted degree

α	avg.corr.	std.dev.
0.1	0.794	0.071
0.2	0.819	0.075
0.4	0.847	0.079
0.6	0.864	0.080

eigenvector

α	avg.corr.	std.dev.
0.1	0.759	0.078
0.2	0.796	0.078
0.4	0.828	0.082
0.6	0.846	0.083

Bonacich

α	avg.corr.	std.dev.
0.1	0.377	0.185
0.2	0.455	0.197
0.4	0.502	0.214
0.6	0.508	0.229

concentration

α	avg.corr.	std.dev.
0.1	0.365	0.190
0.2	0.444	0.202
0.4	0.491	0.217
0.6	0.494	0.229

closeness

α	avg.corr.	std.dev.
0.1	0.427	0.094
0.2	0.402	0.100
0.4	0.356	0.117
0.6	0.313	0.127

betweenness

α	avg.corr.	std.dev.
0.1	0.433	0.164
0.2	0.428	0.181
0.4	0.293	0.185
0.6	0.303	0.195

Table 2: Average correlation of centrality measures compared to losses generated by the failure of single nodes, randomized liquid assets.

harmonic distances

c	1	2	3
α	corr.	corr.	corr.
0.1	0.993	0.994	0.991
0.2	0.985	0.979	0.987
0.4	0.947	0.915	0.947
0.6	0.880	0.839	0.900

extended harmonic distances

c	1	2	3
α	corr.	corr.	corr.
0.1	0.993	0.995	0.991
0.2	0.986	0.980	0.988
0.4	0.950	0.916	0.949
0.6	0.877	0.811	0.881

weighted degree

c	1	2	3
α	corr.	corr.	corr.
0.1	0.998	0.999	0.998
0.2	0.999	0.999	1.000
0.4	0.999	0.999	0.998
0.6	0.996	0.997	0.995

eigenvector

c	1	2	3
α	corr.	corr.	corr.
0.1	0.995	0.996	0.996
0.2	0.999	0.998	0.999
0.4	0.999	0.998	0.999
0.6	0.997	0.997	0.997

Bonacich

c	1	2	3
α	corr.	corr.	corr.
0.1	0.993	0.993	0.992
0.2	0.996	0.993	0.995
0.4	0.997	0.996	0.995
0.6	0.992	0.995	0.994

concentration

c	1	2	3
α	corr.	corr.	corr.
0.1	0.993	0.993	0.992
0.2	0.996	0.993	0.995
0.4	0.997	0.996	0.995
0.6	0.992	0.995	0.994

closeness

c	1	2	3
α	corr.	corr.	corr.
0.1	0.864	0.847	0.853
0.2	0.800	0.798	0.793
0.4	0.719	0.710	0.714
0.6	0.651	0.635	0.662

betweenness

c	1	2	3
α	corr.	corr.	corr.
0.1	0.841	0.819	0.817
0.2	0.726	0.726	0.724
0.4	0.603	0.586	0.600
0.6	0.511	0.477	0.512

Table 3: Correlation of averaged network measures and average induced losses.

on 1000 complete networks for fixed parameter c generated according to 6.3, see results in Table 4. Note that betweenness in a complete network is 0 by definition and closeness is also constant, therefore the correlations are not interpretable. The main finding is that harmonic distances behave very differently compared to real losses, though extended harmonic distances still perform slightly better than harmonic distances. Bonacich and concentration centralities perform even worse. I also note that for complete networks the reciprocal transformation did not result any significant change in correlations.

These results are even more surprising concerning that also weighted degree and eigenvector centrality become much worse. It also confirms that centrality performances are highly dependent on the network structure itself therefore calculating the real payment equilibrium is always useful and recommended and centrality measures are to be used for supplementary analysis.

	c					
	1		2		3	
	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
harmonic distances	0.067	0.151	0.024	0.148	0.020	0.141
extended harmonic distances	0.168	0.154	0.136	0.153	0.129	0.146
weighted degree	0.654	0.087	0.680	0.076	0.677	0.077
eigenvector	0.642	0.095	0.672	0.076	0.669	0.075
Bonacich	0.040	0.151	0.002	0.147	-0.001	0.142
concentration	-0.053	0.143	0.029	0.143	-0.029	0.150
closeness	0	0	0	0	0	0
betweenness	N/A	N/A	N/A	N/A	N/A	N/A

Table 4: Correlations for complete networks.

8 Application as a financial stress indicator

In the previous section I analyzed harmonic distances and centrality measures in cross-section without time-evolution. Now I turn to a real-life financial network, the unsecured interbank lending of Hungarian financial institutions to see how these measures behave in time. It turns out to be interesting to look at these measures aggregated as an indicator of the state of the interbank market.

Hungarian financial institutions including banks, saving cooperatives, building societies and financial undertakings have to report their unsecured interbank lending and deposit transactions to the MNB thus the time evolution of the whole network is known. Data starts on 2 January 2008 and lasts until 31 December 2015. Banks with several subsidiaries are handled as consolidated banking groups. A daily network of transactions is usually sparse but one is able to calculate centrality measures on connected networks. To obtain at least a weakly connected network, I aggregated daily networks in 5-day, non-overlapping windows. This aggregation resulted networks in which the size of the largest weakly connected component is almost always equal to the number of active banks in the corresponding 5-day period. Therefore, a network contains the weekly transactions of financial institutions. Figure 6 shows the number of financial institutions and their total value of transactions in every week in the given time period.

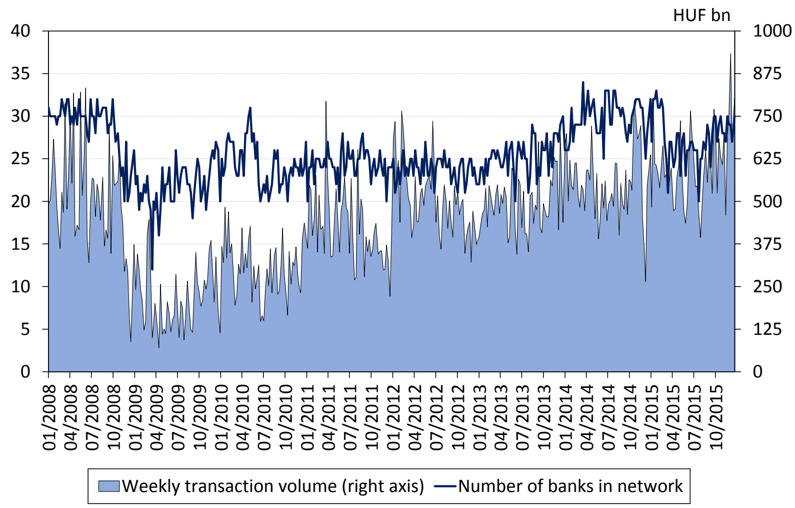

Figure 6: Number of banks and transaction volume in weekly networks

8.1 Scale-free property

As a first side-result, one has to check whether the Hungarian interbank lending network is a scale-free network which is a missing result from previous papers on Hungarian data. This fact, on the other hand justifies the transformation $\frac{1}{\sum_{i=1}^{n} h_{i,j}}$ also for Hungarian data that was applied in Subsection 7.1. For this, one has to empirically check that the degree distribution follows a power law distribution: $\mathbb{P}(k) \sim k^{-\gamma}$ for some γ typically lying between 2 and 3. Figure 7 shows the degree distribution for the 405 weekly networks. Colours change across 405 weeks as indicated on the colour bar.

I fit power-law distributions with the MATLAB codes of Clauset et al. (2009). The code estimates parameters with maximum likelihood method and tests the null-hypothesis of being drawn from a power-law distribution with a Kolmogorov – Smirnov test. I accept a network to be power-law distributed with p-value greater than 0.1 as in Clauset et al. (2009). This is approximately 80% of our sample, 78 networks failed the test out of 405 networks. Table 5 shows the numbers of networks and the average, minimum, maximum p-value and γ parameters of degree distributions. Rejected networks are distributed roughly uniformly in the time series.

Summarizing this subsection, I showed that most of the Hungarian interbank networks in the time series from 2008 to 2015 are scale-free. For similar results on the Austrian interbank network, see Boss et al. (2004). For a statistical test based approach on US payment flows, see Soramäki et al. (2007).

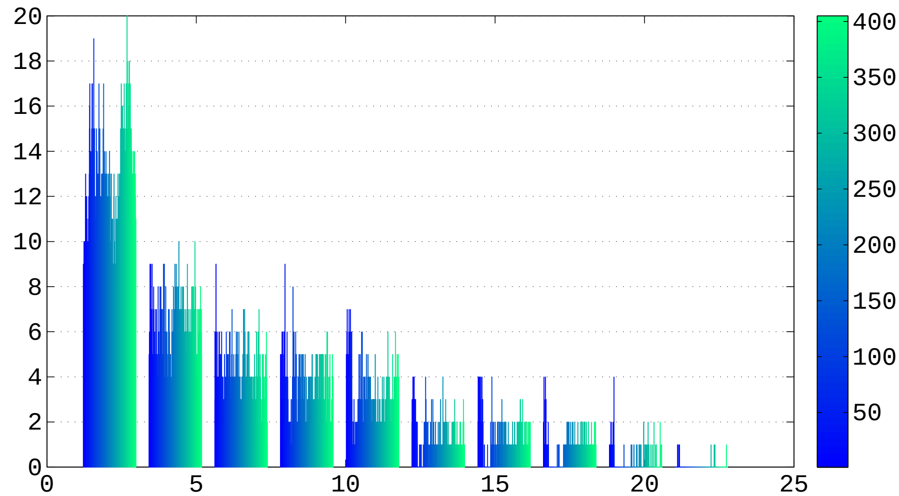

Figure 7: Degree distribution of the Hungarian interbank lending network between 2008 and 2015

		p-value			γ		
	no. of networks	min	max	avg	min	max	avg
scale-free	327	0.100	0.993	0.477	1.886	3.344	2.906
non-scale-free	78	0.000	0.098	0.038	1.499	3.321	2.117

Table 5: Accepted (scale-free) and rejected (non-scale-free) networks' p-values and γ parameters of degree distributions

8.2 Cross-section analysis

Similarly to Subsection 7.1. and 7.2, I calculate correlations along institutions for every week between 2008 and 2015. In contrast with our simulated scale-free networks where out-degree is very close to in-degree because of symmetric edge attachment, weighted degree is substituted with weighted out-degree as an ex-ante failure of an institution causes contagion if only if it is intended to pay back its interbank loans. In this analysis, I accept the fact from Subsection 7.1 that weighted degree fits best to generated loss patterns. I do not drop rejected (non-scale-free) networks from our sample because correlations do not improve significantly. Thus, one can think of rejected networks as nearly scale-free and apply the reciprocal transformation of harmonic distances. Correlations on Figure 7 confirms our results in Subsection 7.1. Weighted degrees and eigenvector centralities are highly correlated, the correlation of transformed harmonic distances to eigenvector centrality is close to 0.5 but extremely volatile with standard deviation around 0.2 as can be seen on Figure 8. Table 6 shows average correlations and standard deviations of correlations. I also present the correlations on the average network in Table 7 which reflects to Subsection 7.1.2 and the results are similar to that subsection. Therefore I conclude that in case of averaging networks or measures in time aggregated harmonic distances become good measures of systemic importance. This is an important result from a practitioner's point of view since as mentioned earlier, at the MNB we identify systemic importance of banks in the financial network by averaging in time.

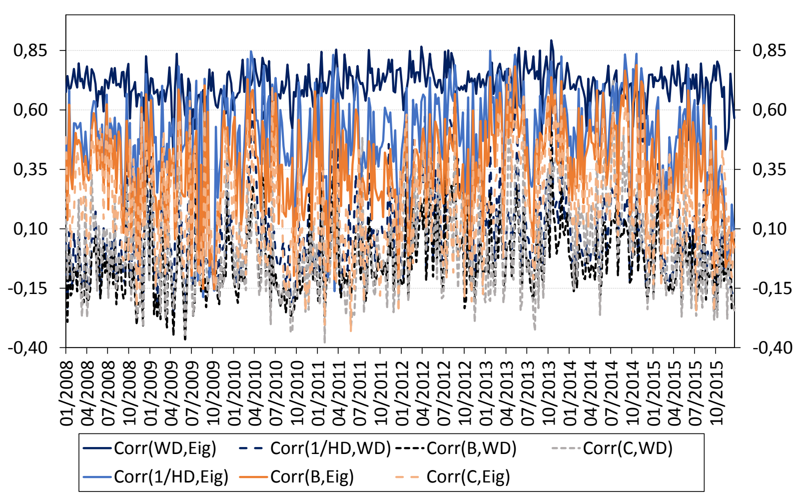

Figure 8: Evolution of correlations across institutions of weighted degrees (WD), eigenvector centralities (Eig), the reciprocal of harmonic distances (1/HD), Bonacich centralities (B) and concentration centralities (C)

8.3 Evolution in time

To be able to capture changes also in transaction volumes and see the variation of risk in time, I use un-normalized centrality measures along the time horizon: weighted out-degree is replaced

	avg.	std.dev.
Corr (WD, Eig)	0.704	0.076
Corr $(1/HD, WD)$	0.079	0.187
Corr $(1/HD, Eig)$	0.488	0.197
Corr (B, WD)	-0.041	0.142
Corr (B, Eig)	0.361	0.196
Corr (C, WD)	0.019	0.208
Corr (C, Eig)	0.189	0.217

Table 6: Average correlations and standard deviations across institutions.

Corr (WD, Eig)	0.881
Corr $(1/HD, WD)$	0.847
Corr $(1/HD, Eig)$	0.937
Corr (B, WD)	0.790
Corr (B, Eig)	0.922
Corr (C, WD)	0.259
Corr (C, Eig)	0.177

Table 7: Correlations on the averaged network.

with total liabilities. Figure 9 shows the sum of liabilities and harmonic distances of an O-SII in Hungary. Similar graphs for two further systemically important institutions are available in Appendix B.

Harmonic distance obtains its peak value during the global financial crisis, while afterwards its volatility becomes lower. This result is similar to other institutions, thus it is promising to define a financial network stability index as

$$I(t) = \sum_{j=1}^{n} \frac{1}{\sum_{i=1}^{n} h_{i,j}(t)}.$$

The behaviour of this index can be seen on Figure 10 together with interbank transaction volume (i.e, the sum of all liabilities) and the sum of previous other measures. One can see on Figure 6 that neither the number of banks in the network, nor the above mentioned measures showed as spectacularly the volatility change around the financial crisis. Harmonic distances seem to catch the systemic instability around a systemic distress situation or the role of an institution in systemic instability. Although one could argue that if harmonic distances performed worse than other measures in the cross-section analysis, why one should apply the aggregation to make implications about the state of systemic stability in the network.

A critique of $I(t)$ might be that a decrease in transaction volumes leads to an increase in transformed harmonic distances. This may lead to the behaviour of harmonic distances during the crisis. Though one can also interpret it as a good result as the drying up of a market is a good indication of a stress situation, to control for the changes in transactions, I define a modified

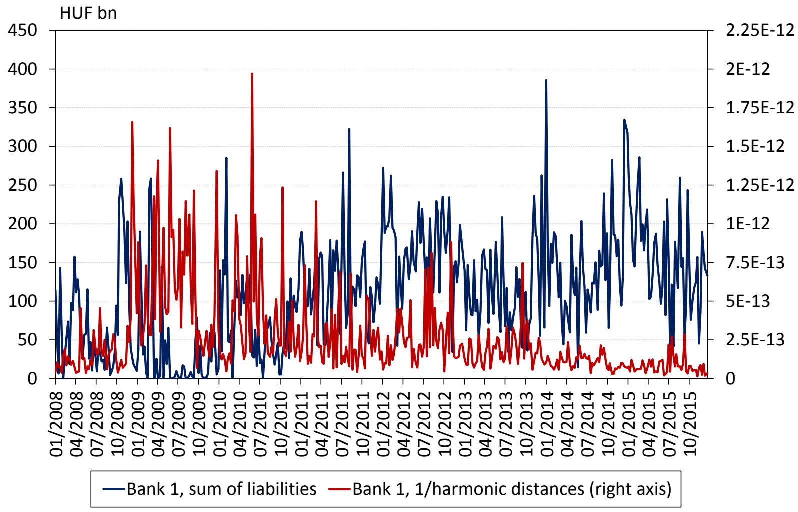

Figure 9: Time evolution of interbank liabilities and harmonic distances of Bank 1 (O-SII) in Hungary

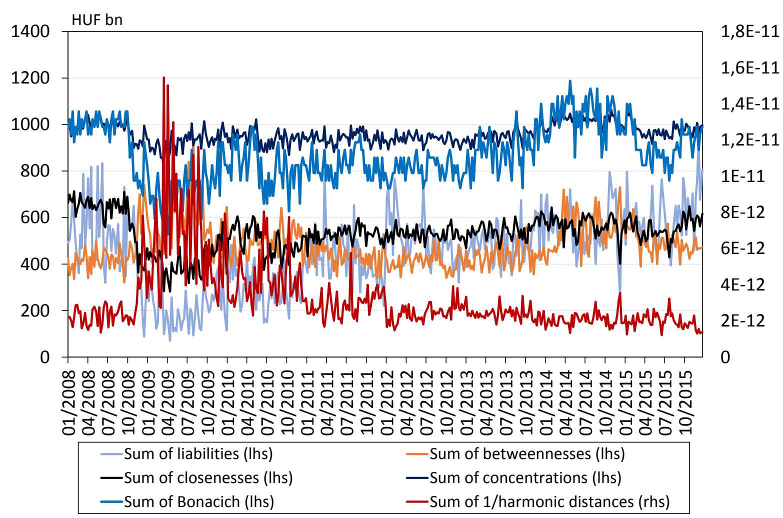

Figure 10: Centrality measures and the sum of all harmonic distances $I(t)$ in the network

version of $I(t)$ inspired by the statement of Proposition 2. Since all banks default if and only if $h_{i,j} < y_i$ for all i, it may be reasonable to rearrange these inequalities to $\frac{h_{i,j}}{y_i} < 1$ for all i. One can see that the left side of this inequality is independent of size as it is compared to a constant 1 in the statement. Let the modified version

$$\hat{I}(t) = \sum_{j=1}^{n} \frac{1}{\sum_{i=1}^{n} \frac{h_{i,j(t)}}{y_i(t)}}.$$

On Figure 11, one can see that when $I(t)$ obtained its peak, the modified version $\hat{I}(t)$ also significantly increased and its volatility also became higher. Nevertheless, it is easy to see that higher transaction volumes make the original index $I(t)$ smoother.

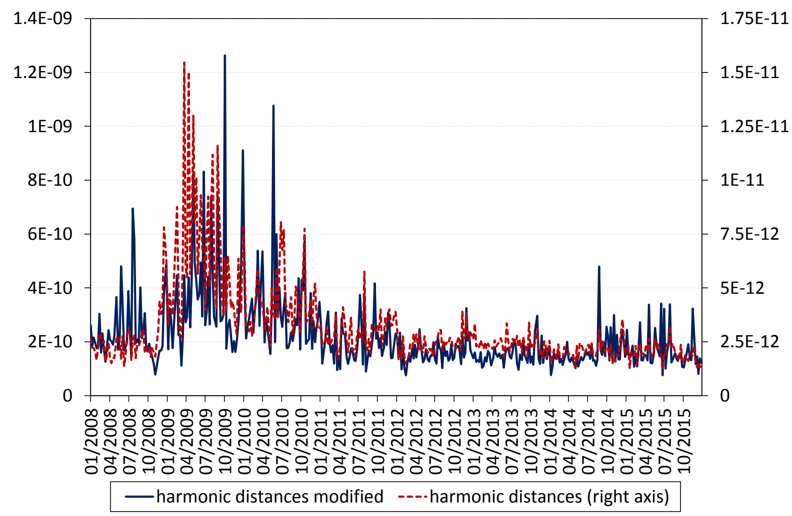

Figure 11: The graph of $I(t)$ and its modification $\hat{I}(t)$

8.4 Performances in a factor model

The idea of using network measures to identify systemic stress can be achieved by including them in a systemic stress index (Hałaj and Kok, 2013). The quantification of their usefulness is possible if we apply a factor model approach. Factor modelling is a dimension reduction tool which creates common components from a dataset by linear operations. If a few number of factors explain the variance of a large data set and explained variance is high enough, the role of variables can be identified in each factor. In this subsection, I quantify the performance of different network measures in a static factor model approach. I use the dataset and preliminary methodology of Szendrei and Varga (2017). They rethink the model of Holló et al. (2012) in a statistically rigorous way by using a dynamic factor model to create an index of financial systemic stress. Usually, the creation of a sophisticated factor model starts with the simple static factor model. There are five sets of variables: government bond market, interbank market, banking sector, FX market and capital market. The exact names of the variables in these sets are indicated in Table 8.

These variables form the vector $\mathbf{y}_t$ at time t, and the standard static factor model is given by

$$\mathbf{y}_t = \lambda \cdot \mathbf{f}_t + \epsilon_t,$$

where the information in $\mathbf{y}_t$ is compressed into the factors $\mathbf{f}_t$, $\mathbf{y}_t$ is n-dimensional, $\mathbf{f}_t$ is q-dimensional and $q < n$. Factors and data are normally distributed, i.e. $\mathbf{f}_t \sim N(\mathbf{0}, \mathbf{I}_q)$, $\epsilon_t \sim N(\mathbf{0}, \mathbf{\Sigma})$ are iid, λ is a $n \times q$ matrix of factor loadings. In this example, the number of variables is $n = 19$ and the number of factors is $q = 4$. The variance is then given by

$$\mathbf{Var}\left(\mathbf{y}_t\right) = \lambda \cdot \lambda^T + \mathbf{\Sigma}.$$

Szendrei and Varga (2017) identify four intuitive factors that drive the variance of the dataset. Since the variance-covariance matrix of $\mathbf{f}_t$ is assumed to be diagonal, the factors are orthogonal to each other in the preliminary, static factor model. It turns out that the inclusion of a network measure makes the fourth factor much better interpretable. The corresponding factor seems to be dominated by the largest Hungarian bank's PD in the banking sector variable set. One can see that network measures play an important role in the third and fourth factor and harmonic distances have a quite high coefficient of 0.3185 in the fourth factor. Furthermore, explained variance increases by approximately 2.7% by including network measures. The exact loading matrices are available in Table 9. Regarding the explained variance, models that include harmonic distances, concentration and Bonacich centralities, betweennesses, closenesses and degrees explain 84.3 percent, 83.4 percent, 84.3 percent, 82.9 percent, 84.7 percent and 84.6 respectively. Note that all measures have been transformed by the reciprocal transformation in order to obtain similar functional shapes.

One can conclude that the inclusion of network measures into a systemic stress indicator is reasonable and improves the explained variance of the information in the dataset.

government bond market	bond yields (3-month and 10-year) CDS (5-year bond)
interbank market	BUBOR (3-month) HUFONIA overnight rate HUFONIA trading volume
banking sector	bank PDs: from market price (Merton model) *network measure*
FX market	bid-ask spreads: HUF/EUR + HUF/USD volatilities: HUF/EUR, HUF/USD, HUF/GBP, HUF/CHF
capital market	CMAX: BUX, BUMIX, CETOP20, DAX implied volatility: VDAX

Table 8: Variables in the factor model (Szendrei and Varga, 2017).

		F1	F2	F3	F4
government bond market	Benchmark yield 3m	0,2852	0,0369	0,2248	0,0081
	Benchmark yield 10y	0,2228	0,0872	0,1886	0,0659
	CDS HUN-GER spread	0,2126	-0,3522	-0,2758	-0,0684
interbank market	BUBOR 3m	0,1661	-0,2024	0,3426	0,3867
	HUFONIA rate	0,2818	-0,0448	0,1635	0,0085
	HUFONIA vol	0,2219	0,2428	-0,3142	0,1235
banking sector	PD Bank1	0,2266	0,3064	-0,2133	0,1385
	PD Bank2/Rest	0,1845	0,3523	-0,2394	0,1004
	Harmonic distances	0,2692	-0,0794	0,1260	-0,3877
FX market	bidask spot	0,1942	-0,3607	-0,2985	-0,0022
	Vol EURO	0,2103	-0,3560	-0,2758	-0,0491
	Vol USD	-0,1444	0,2339	-0,1786	-0,4073
	Vol GBP	0,2572	0,1046	0,2122	0,1018
	Vol CHF	0,2880	-0,0050	0,2049	0,0245
capital market	BUX CMAX	0,2497	-0,2901	-0,1034	-0,1851
	BUMIX CMAX	0,2514	0,2614	-0,0228	0,0996
	CETOP CMAX	0,2390	0,0682	0,0341	0,0818
	DAX CMAX	0,1565	0,1396	0,2957	-0,6378
	VDAX	0,2252	0,2041	-0,3039	-0,0975
	Explained variance	0,5247	0,1611	0,1050	0,0521

	F1	F2	F3	F4
Benchmark yield 3m	0,2853	0,0441	0,2150	0,1441
Benchmark yield 10y	0,2216	0,0963	0,1499	0,3015
CDS HUN-GER spread	0,2104	-0,3807	-0,2314	-0,0644
BUBOR 3m	0,1695	-0,0753	0,3643	-0,2073
HUFONIA rate	0,2804	-0,0411	0,1521	0,1738
HUFONIA vol	0,2227	0,2236	-0,3437	-0,0149
PD Bank1	0,2283	0,2939	-0,2465	0,0123
PD Bank2/Rest	0,1864	0,3400	-0,2753	-0,0678
Concentration	0,2708	-0,0884	0,1744	-0,3449
bidask spot	0,1917	-0,3884	-0,2592	-0,0074
Vol EURO	0,2079	-0,3836	-0,2341	-0,0459
Vol USD	-0,1393	0,2149	-0,1404	-0,3871
Vol GBP	0,2563	0,1170	0,1759	0,2108
Vol CHF	0,2876	0,0014	0,1962	0,1551
BUX CMAX	0,2481	-0,3089	-0,0596	-0,1803
BUMIX CMAX	0,2519	0,2607	-0,0640	0,0828
CETOP CMAX	0,2394	0,0659	0,0246	0,1233
DAX CMAX	0,1619	0,1387	0,3557	-0,6008
VDAX	0,2275	0,1785	-0,3016	-0,2260
Explained variance	0,5253	0,1569	0,1055	0,0465

		F1	F2	F3	F4
government bond market	Benchmark yield 3m	0,2859	0,0475	0,1988	0,1251
	Benchmark yield 10y	0,2226	0,0958	0,1381	0,2447
	CDS HUN-GER spread	0,2070	-0,3917	-0,2107	-0,0672
interbank market	BUBOR 3m	0,1878	-0,0054	0,4273	-0,0495
	HUFONIA rate	0,2805	-0,0410	0,1429	0,1421
	HUFONIA vol	0,2203	0,2095	-0,3537	0,0119
banking sector	PD Bank1	0,2266	0,2837	-0,2640	0,0461
	PD Bank2/Rest	0,1856	0,3295	-0,2844	-0,0384
	Bonacich	0,2704	-0,0856	0,1609	-0,3739
FX market	bidask spot	0,1882	-0,4000	-0,2357	-0,0025
	Vol EURO	0,2046	-0,3948	-0,2118	-0,0482
	Vol USD	-0,1398	0,2125	-0,1547	-0,4094
	Vol GBP	0,2572	0,1188	0,1598	0,1943
	Vol CHF	0,2879	0,0041	0,1824	0,1372
capital market	BUX CMAX	0,2459	-0,3143	-0,0503	-0,1940
	BUMIX CMAX	0,2522	0,2552	-0,0746	0,0823
	CETOP CMAX	0,2389	0,0640	0,0139	0,1384
	DAX CMAX	0,1634	0,1499	0,3238	-0,6458
	VDAX	0,2253	0,1660	-0,3101	-0,2160
	Explained variance	0,5287	0,1562	0,1129	0,0456

	F1	F2	F3	F4
Benchmark yield 3m	0,2850	0,0453	0,2282	0,1433
Benchmark yield 10y	0,2219	0,0979	0,1710	0,3034
CDS HUN-GER spread	0,2129	-0,3778	-0,2326	-0,0551
BUBOR 3m	0,1494	-0,1189	0,2848	-0,4168
HUFONIA rate	0,2804	-0,0379	0,1627	0,1925
HUFONIA vol	0,2243	0,2261	-0,3511	-0,0125
PD Bank1	0,2296	0,2946	-0,2485	-0,0128
PD Bank2/Rest	0,1880	0,3395	-0,2755	-0,0988
Betweenness	0,2710	-0,0877	0,1860	-0,2856
bidask spot	0,1941	-0,3853	-0,2635	-0,0093
Vol EURO	0,2103	-0,3806	-0,2361	-0,0373
Vol USD	-0,1384	0,2105	-0,1321	-0,4066
Vol GBP	0,2560	0,1186	0,1887	0,2034
Vol CHF	0,2874	0,0029	0,2078	0,1558
BUX CMAX	0,2498	-0,3066	-0,0559	-0,1452
BUMIX CMAX	0,2527	0,2613	-0,0556	0,0624
CETOP CMAX	0,2398	0,0669	0,0305	0,0902
DAX CMAX	0,1609	0,1357	0,3752	-0,5242
VDAX	0,2293	0,1797	-0,3036	-0,2039
Explained variance	0,5221	0,1579	0,1003	0,0488

		F1	F2	F3	F4
government bond market	Benchmark yield 3m	0,2870	0,0589	0,1868	0,0810
	Benchmark yield 10y	0,2245	0,1020	0,1305	0,1721
	CDS HUN-GER spread	0,2052	-0,4104	-0,1641	-0,0502
interbank market	BUBOR 3m	0,1732	0,1009	0,4598	0,0978
	HUFONIA rate	0,2821	-0,0328	0,1530	0,0959
	HUFONIA vol	0,2219	0,1749	-0,3633	0,0515
banking sector	PD Bank1	0,2280	0,2551	-0,2921	0,0753
	PD Bank2/Rest	0,1870	0,3003	-0,3179	0,0063
	Closeness	0,2704	-0,0772	0,1545	-0,3933
FX market	bidask spot	0,1864	-0,4199	-0,1850	0,0197
	Vol EURO	0,2029	-0,4131	-0,1629	-0,0297
	Vol USD	-0,1407	0,1958	-0,1989	-0,4329
	Vol GBP	0,2592	0,1276	0,1486	0,1552
	Vol CHF	0,2888	0,0145	0,1760	0,0947
capital market	BUX CMAX	0,2447	-0,3208	-0,0192	-0,1894
	BUMIX CMAX	0,2542	0,2435	-0,0989	0,0841
	CETOP CMAX	0,2402	0,0617	0,0109	0,1327
	DAX CMAX	0,1634	0,1699	0,2708	-0,6819
	VDAX	0,2260	0,1341	-0,3272	-0,1797
	Explained variance	0,5259	0,1572	0,1180	0,0460

	F1	F2	F3	F4
Benchmark yield 3m	0,2866	0,0569	0,1873	0,0801
Benchmark yield 10y	0,2243	0,1010	0,1334	0,1719
CDS HUN-GER spread	0,2050	-0,4085	-0,1720	-0,0512
BUBOR 3m	0,1790	0,0885	0,4523	0,0995
HUFONIA rate	0,2818	-0,0343	0,1526	0,0953
HUFONIA vol	0,2217	0,1797	-0,3612	0,0522
PD Bank1	0,2278	0,2591	-0,2883	0,0760
PD Bank2/Rest	0,1866	0,3041	-0,3165	0,0051
Degrees	0,2701	-0,0789	0,1533	-0,3934
bidask spot	0,1862	-0,4178	-0,1932	0,0187
Vol EURO	0,2028	-0,4112	-0,1706	-0,0305
Vol USD	-0,1407	0,1983	-0,1967	-0,4343
Vol GBP	0,2589	0,1262	0,1515	0,1553
Vol CHF	0,2885	0,0127	0,1768	0,0944
BUX CMAX	0,2445	-0,3205	-0,0241	-0,1895
BUMIX CMAX	0,2538	0,2447	-0,0974	0,0827
CETOP CMAX	0,2400	0,0619	0,0126	0,1333
DAX CMAX	0,1633	0,1672	0,2759	-0,6807
VDAX	0,2257	0,1383	-0,3268	-0,1798
Explained variance	0,5270	0,1570	0,1169	0,0460

Table 9: Loading matrices and explained variances of the preliminary, static factor model (Szendrei and Varga, 2017) with different network measures included.

9 An explanation of results

We saw in Subsection 7.1 that the best fitting measure is weighted degree. That is the consequence of the fact that first order losses (i.e. the sum of interbank liabilities which is equivalent to weighted degree) largely dominate higher order losses and higher order losses only add a shift to first order losses. Figure 12 shows the amount of first order and higher order losses as an output of the Eisenberg – Noe algorithm in a typical scale-free network. In complete networks, higher order losses are larger which explains lower correlations in Subsection 7.2. One can conclude that we are still not able to approximate higher order losses without the iterative algorithm of Eisenberg and Noe.

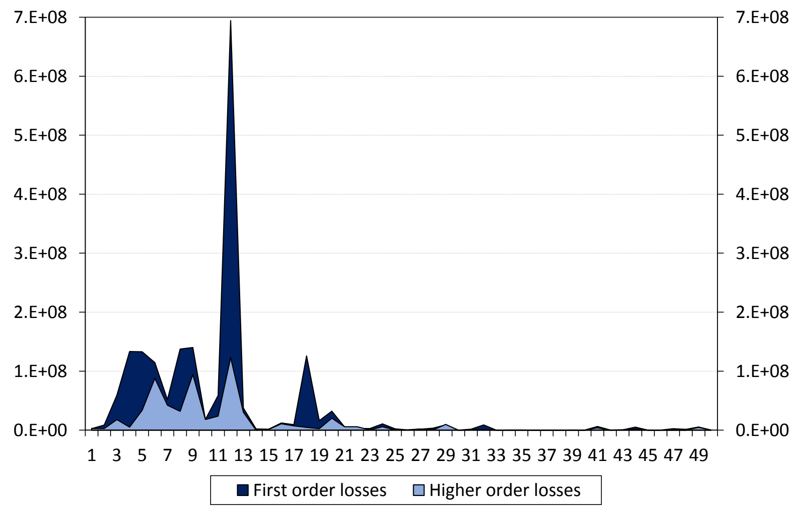

Figure 12: First order and higher order losses induced by the initial default of banks

Arriving to the not so attractive results of harmonic distances, this behaviour may be the consequence of the weakness of Proposition 2 and its pair Proposition 12 in Acemoglu et al. (2015b). These statements linearize the non-linear problem

$$\mathbf{x}^* = [\min\{\mathbf{Q}\mathbf{x}^* + \mathbf{e}, \mathbf{y}\}]^+$$

by requiring all banks in the network to default if a single bank initially fails. In real life it is not realistic that all other institutions fail even if the initial default is of a dominant systemic bank. The default of all nodes in the presence of a large shock could be possible in case of a complete network because of dense interconnections as theoretically stated in Acemoglu et al. (2015b). In Subsection 7.2, I analyzed contagion in complete networks and did not experience the default of all banks in case of simulated networks. Results for Bonacich and concentration centralities might be explained by the Taylor series expansion method, a series expansion around 0 (Acemoglu et al., 2015a) might be inappropriate for our aims.

Another observation was the behaviour of these measures in time using the daily data of lending and deposit transactions of Hungarian financial institutions. Correlations were found to be similar to the case of simulated networks. The transformation of harmonic distances peaked around the crisis. As it was mentioned in Section 5, this is partly a consequence of the decline of transaction volumes.

10 Conclusion

The main goal of this paper was to provide the first numerical results on harmonic distances by the extension of the model of Acemoglu et al. (2015b), also analyzing the concentration centrality of Acemoglu et al. (2015a). The first set of results are based on simulated interbank networks, while the second set of results shows the behaviour in a real financial network. I gave the explicit formula for the calculation of harmonic distances and extended the results for any size of liquid assets. To perform the analysis, I generated large numbers of Barabási – Albert type and complete networks and induced institutional failures by the deletion of their outgoing payments. The results showed that traditionally used network measures like weighted degree and eigenvector centrality can catch the importance of individual nodes as default-implied losses of a node correlates highly with these measures. Despite the theoretical grounding, harmonic distances, Bonacich and concentration centralities could not outperform the previously mentioned "off-the-shelf" measures but the application of averaging methods resulted in the similar performance of the above as well. I note that the extended version of harmonic distance performed better than the original one. In case of complete networks even the previously good measures became much poorer. This fact underpins that one has to strictly investigate the application purpose of centrality measures.

I repeated the analysis for averaged Barabási – Albert type networks. This method showed very good results for harmonic distances, Bonacich and concentration centralities as well. One could conclude (carefully interpreting this result) that the pattern of expected losses (where expectation is along networks) is very well described by the two main traditional measures and the two harmonic distances. I can conclude that for specific policy applications all these measures are suitable when we examine time-averaged real life networks or time averages of centrality measures.

For empirical results, I also calculated network measures for the time series of Hungarian interbank lending networks. After showing that most of these networks are scale-free, I calculated correlations between the measures and found a similar phenomenon as in case of simulated networks, for observed and "mean" behaviour as well. One finding was that the behaviour of harmonic distances is very stable in "normal times" and it peaks around the financial crisis. This makes it a possible indicator of financial network stress besides other measures as a factor model approach shows it on Hungarian data (Szendrei and Varga, 2017). This phenomenon is partly due to the sharp drop in total transaction volumes. A modified measure behaves similarly with higher volatility.

This work pointed out that the behaviour of financial networks in a contagion situation is extremely complicated and hard to approximate despite the large amount of research carried out in recent years. The structure of the underlying network is a key component of the solution when one applies different importance measures. I found that the information in a single random or real network cannot be compressed into a centrality measure when the area of analysis is direct contagion. When looking at an average network along a time horizon, the compression of information becomes very good. The decrease of volatility is related to the law of large numbers.

References

Acemoglu, D., A. Ozdaglar and A. Tahbaz-Salehi (2015a), 'Networks, shocks and systemic risk', *The Oxford Handbook on the Economics of Networks*,

Acemoglu, D., A. Ozdaglar and A. Tahbaz-Salehi (2015b), 'Systemic risk and stability in financial networks', *American Economic Review*, vol. 105 no. 2, pp. 564–608

Albert, R. and A. L. Barabási (2002), 'Statistical mechanics of complex networks', *Reviews of Modern Physics*, vol. 74 no. 1, pp. 47

Allen, F. and D. Gale (2000), 'Financial contagion', *Journal of Political Economy*, vol. 108 no. 1, pp. 1–33

Alter, A., B. Craig and P. Raupach (2015), 'Centrality-Based Capital Allocations', *International Journal of Central Banking*, vol. 11 no. 3, pp. 329–379

Battiston, S., M. Puliga, R. Kaushik, P. Tasca and G. Caldarelli (2012), 'DebtRank: Too central to fail? Financial networks, the fed and systemic risk', *Scientific Reports*, vol. 2,

Bonacich, P. (1987), 'Power and Centrality: A Family of Measures', *American Journal of Sociology (University of Chicago Press)*, no. 92, pp. 1170–1182

Bonacich, P. and P. Lloyd (2001), 'Eigenvector-like measures of centrality for asymmetric relations', *Social Networks*, vol. 23 no. 3, pp. 191–201

Borgatti, S. (2005), 'Centrality and network flow', *Social Networks*, no. 27, pp. 55–71

Boss, M., H. Elsinger, M. Summer and S. Thurner (2004), 'The Network Topology of the Interbank Market', *Quantitative Finance*, vol. 4 no. 6, pp. 677–684

Cecchetti, S., M. Rocco and L. Sigalotti (2016), *Contagion and fire sales in banking networks*, Temi di Discussione (Working Paper) 1050, Bank of Italy.

Clauset, A., C. Shalizi and M. Newman (2009), 'Power-law distributions in empirical data', *SIAM Review*, vol. 51 no. 4, pp. 661–703

Eisenberg, L. and T. Noe (2001), 'Systemic risk in financial systems', *Management Science*, vol. 47 no. 2, pp. 236–249

Fink, K., U. Krüger, B. Meller and L.-H. Wong (2015), *The credit quality channel: Modeling contagion in the interbank market*, Discussion Paper No 38, Deutsche Bundesbank.

Freixas, X., B. M. Parigi and J. C. Rochet (2000), 'Systemic risk, interbank relations, and liquidity provision by the central bank', *Journal of Money, Credit and Banking*, vol. 32 no. 3, pp. 611–638

Hałaj, G. and C. Kok (2013), 'Assessing interbank contagion using simulated networks', *Computational Management Science*, vol. 10 no. 2-3, pp. 157–186

Holló, D., M. Kremer and M. Lo Duca (2012), *CISS-a composite indicator of systemic stress in the financial system*, Working Paper series 1426, European Central Bank.

Levin, D., Y. Peres and E. Wilmer (2009), *Markov chains and mixing times*, American Mathematical Society.

Lovász, L. (1993), 'Random walks on graphs: A survey', *Combinatorics, Paul Erdős is eighty*, vol. 2 no. 1, pp. 1–46

Soramäki, K. and S. Cook (2013), 'SinkRank: An algorithm for identifying systemically important banks in payment systems', *Economics*, vol. 7 no. 28, pp. 1

Soramäki, K., M. L. Bech, J. Arnold, R. J. Glass and W. E. Beyeler (2007), 'The topology of interbank payment flows', *Physica A: Statistical Mechanics and its Applications*, vol. 379 no. 1, pp. 317–333

Szendrei, T. and K. Varga (2017), *FISS – A Factor Based Index of Systemic Stress in the Financial System*, mimeo, Magyar Nemzeti Bank.

Upper, C. (2011), 'Simulation methods to asses the danger of contagion in interbank markets', *Journal of Financial Stability*, vol. 7 no. 3, pp. 111–125

Vivier-Lirimont, S. (2006), *Contagion in interbank debt networks*, Working Paper,

A Theory and proofs

A matrix $\mathbf{A}$ with non-negative elements is called *irreducible* if for any i, j there is a k such that $\mathbf{A}^k_{i,j} > 0$.

Theorem (Perron – Frobenius)**.** *Let* $\mathbf{A}$ *be an irreducible square matrix. Then*

1. $\mathbf{A}$ *has a positive real eigenvalue* λ_{max} *such that all other eigenvalues satisfy* $|\lambda| \leq \lambda_{max}$.
2. λ_{max} *has multiplicity 1 and has an eigenvector* $\mathbf{v} > 0$.
3. *Any positive eigenvector is a multiple of* $\mathbf{v}$.
4. *If* $\mathbf{w} \geq 0$, $\mathbf{w} \neq 0$ *and* μ *is a number such that* $\mathbf{A}\mathbf{w} \leq \mu\mathbf{w}$, *then* $\mathbf{w} > 0$ *and* $\mu \geq \lambda_{max}$. $\mu = \lambda_{max}$ *if and only if* $\mathbf{w}$ *is a multiple of* $\mathbf{v}$. □

Proof of Lemma 1 (Bonacich and Lloyd (2001))**.** For the sake of generality, I do not assume that $\mathbf{A}$ is symmetric. Let $\mathbf{V}$ be the matrix of eigenvectors $\mathbf{v}_i$ of $\mathbf{A}$. Then $\mathbf{AV} = \mathbf{V}\lambda$ and $\mathbf{A}^k = \mathbf{V}\lambda^k V^{-1}$. Let $\mathbf{w}_i$ be the ith row of $\mathbf{V}^{-1}$. Then $\mathbf{A}^k = \sum_{i=1}^n \lambda_i^k \cdot \mathbf{v}_i \cdot \mathbf{w}_i$. Therefore Bonacich centrality for $\alpha = 1$ becomes

$$\mathbf{b} = (\mathbf{I} - \beta\mathbf{A})^{-1} \cdot \mathbf{1} = \left(\sum_{k=0}^{\infty} \beta^k \mathbf{A}^k\right) \cdot \mathbf{1} = \left(\sum_{k=0}^{\infty} \beta^k \sum_{i=1}^{n} \lambda_i^k \cdot \mathbf{v}_i \cdot \mathbf{w}_i\right) \cdot \mathbf{1}$$
$$= \left(\sum_{i=1}^{n} \sum_{k=0}^{\infty} \beta^k \lambda_i^k \cdot \mathbf{v}_i \cdot \mathbf{w}_i\right) \cdot \mathbf{1} = \sum_{i=1}^{n} \frac{\mathbf{w}_i \cdot \mathbf{v}_i}{1 - \beta\lambda_i} \cdot \mathbf{1},$$

thus using the fact that λ_1 is the largest eigenvalue of $\mathbf{A}$, it is easy to see that the second term disappears in limit, $\lim_{\beta \to 1/\lambda_1 -} \mathbf{b}(1, \beta) \cdot (1 - \beta\lambda_1) = (\mathbf{w}_1 \cdot \mathbf{1})\mathbf{v}_1$, what one needed to show. □

Proof of Lemma 2. First, I show that 1 is an eigenvalue of a stochastic matrix denoted by $\mathbf{W}$. The rows of $\mathbf{W}$ sum up to 1, therefore $\mathbf{W} \cdot \mathbf{1} = 1 \cdot \mathbf{1}$, 1 is indeed an eigenvalue. Secondly, I show that there is no larger eigenvalue than 1. Suppose that $\mathbf{W} \cdot \mathbf{x} = \lambda\mathbf{x}$ for a $\lambda > 1$. Then $\mathbf{W} \cdot \mathbf{x}$ is a vector with elements smaller than the largest element of $\mathbf{x}$. On the other hand, at least one element of $\lambda\mathbf{x}$ is greater than the largest element of $\mathbf{x}$. This is a contradiction. □

Proof of Proposition 1. Let us first rewrite equation (5) in matrix form. Let $\mathbf{Y}$ be the matrix of total repayments, where the ith row is $(y_i, \ldots, y_i)$. Then equation (5) turns into

$$(\mathbf{I} - \mathbf{Q}) \cdot \mathbf{H} = \mathbf{Y} - \left(\sum_{i=1}^{n} y_i\right) \cdot \mathbf{I}. \tag{7}$$

1st case: If $(\mathbf{I} - \mathbf{Q})$ is not invertible (or equivalently, 0 is an eigenvalue of $(\mathbf{I} - \mathbf{Q})$), one cannot simply obtain $\mathbf{H}$. I show that $\mathbf{M} = -\left(\sum_{i=1}^n y_i\right) \cdot \left(\mathbf{I} - \mathbf{Q} + \frac{1}{\sum_{i=1}^n y_i} \cdot \mathbf{Y}\right)^{-1}$ solves equation (7).

$$
\begin{aligned}
(\mathbf{I}-\mathbf{Q})\cdot\mathbf{M} &= -(\mathbf{I}-\mathbf{Q})\cdot\left(\sum_{i=1}^{n} y_i\right)\cdot\left(\mathbf{I}-\mathbf{Q}+\frac{1}{\sum_{i=1}^{n} y_i}\cdot\mathbf{Y}\right)^{-1}\\
&= -\left(\mathbf{I}-\mathbf{Q}+\frac{1}{\sum_{i=1}^{n} y_i}\cdot\mathbf{Y}-\frac{1}{\sum_{i=1}^{n} y_i}\cdot\mathbf{Y}\right)\cdot\left(\sum_{i=1}^{n} y_i\right)\cdot\left(\mathbf{I}-\mathbf{Q}+\frac{1}{\sum_{i=1}^{n} y_i}\cdot\mathbf{Y}\right)^{-1}\\
&= -\left(\sum_{i=1}^{n} y_i\right)\cdot\mathbf{I}+\mathbf{Y}\cdot\left(\mathbf{I}-\mathbf{Q}+\frac{1}{\sum_{i=1}^{n} y_i}\cdot\mathbf{Y}\right)^{-1}.
\end{aligned}
$$

It remained to check that $\mathbf{Y}\cdot\left(\mathbf{I}-\mathbf{Q}+\frac{1}{\sum_{i=1}^{n} y_i}\cdot\mathbf{Y}\right)^{-1}=\mathbf{Y}$, which is equivalent with $\mathbf{Y}=\mathbf{Y}\cdot\left(\mathbf{I}-\mathbf{Q}+\frac{1}{\sum_{i=1}^{n} y_i}\cdot\mathbf{Y}\right)$. After noticing that $\mathbf{Y}\cdot(\mathbf{I}-\mathbf{Q})=\mathbf{0}$, it is enough to verify that $\mathbf{Y}\cdot\frac{1}{\sum_{i=1}^{n} y_i}\cdot\mathbf{Y}=\frac{1}{\sum_{i=1}^{n} y_i}\cdot\mathbf{Y}\cdot\sum_{i=1}^{n} y_i=\mathbf{Y}$.

$\mathbf{M}$ satisfies equation (7), but there is one more restriction; the diagonals have to be zero according to the definition. This can be reached by adding a matrix $\mathbf{D}$ to $\mathbf{M}$ for which $(\mathbf{I}-\mathbf{Q})\cdot\mathbf{D}=\mathbf{0}$ and $d_{i,i}=-m_{i,i}$.

According to that every column of $\mathbf{D}$ is an eigenvector of $(\mathbf{I}-\mathbf{Q})$ corresponding to 0 eigenvalue (let it be $\mathbf{v}_0$). This eigenvector multiplied by a scalar is still an eigenvector since $(\mathbf{I}-\mathbf{Q})\cdot\lambda\mathbf{v}_0=\lambda(\mathbf{I}-\mathbf{Q})\mathbf{v}_0=\lambda\cdot 0\cdot\mathbf{v}_0=\mathbf{0}$. Thus for the ith column of $\mathbf{D}$, let $\mathbf{d}_i=-\mathbf{v}_0\cdot\frac{m_{i,i}}{v_{0,i}}$, then $(\mathbf{I}-\mathbf{Q})\cdot\mathbf{d}_i=\frac{-m_{i,i}}{v_{0,i}}(\mathbf{I}-\mathbf{Q})\cdot\mathbf{v}_0=\mathbf{0}$ and $d_{i,i}=-m_{i,i}$. $\mathbf{H}=\mathbf{M}+\mathbf{D}$ solves equation (7) with the restriction on the diagonals.

2nd case: If $(\mathbf{I}-\mathbf{Q})$ is invertible then 0 is not an eigenvalue of $(\mathbf{I}-\mathbf{Q})$. Simply $\mathbf{M}=(\mathbf{I}-\mathbf{Q})^{-1}\cdot(\mathbf{Y}-(\sum_{i=1}^{n} y_i)\cdot\mathbf{I})$. For the diagonal restriction one again needs a matrix $\mathbf{D}$ for which $(\mathbf{I}-\mathbf{Q})\cdot\mathbf{d}_i=\mathbf{0}$ for all i. This would mean that all columns of $\mathbf{D}$ are eigenvectors corresponding to 0 eigenvalue. This is a contradiction.

In fact, the assumption that there is no non-borrowing node in the network is equivalent with the invertibility of $(\mathbf{I}-\mathbf{Q})$: if there is no non-borrowing node then the sum of all columns is 0 in $(\mathbf{I}-\mathbf{Q})$, $(\mathbf{I}-\mathbf{Q})^T\cdot\mathbf{1}=\mathbf{0}$, 0 is an eigenvalue of $(\mathbf{I}-\mathbf{Q})^T$. □

Proof of Proposition 2. 1. Suppose $\varepsilon>\sum_{i=1}^{n} e_i$ but no bank defaults. Then from the definition of payment equilibrium,

$$
\sum_{k\neq i} x_{i,k}+e_i\geq\sum_{k\neq i} x_{k,i}
$$

for all banks i. Summing over all i and using that j is shocked with ε,

$$
\sum_{i}\sum_{k\neq i} x_{i,k}+e_i\geq\sum_{i}\sum_{k\neq i} x_{k,i},
$$

and by the fact that the sums on the two sides are equal, $\sum_{i\neq j} e_i+e_j-\varepsilon\geq 0$. Furthermore, since only j is shocked, I equivalently get $\sum_i e_i\geq\varepsilon$. This is a contradiction.

2. In the presence of a large shock to bank j, all other banks default if and only if $x_i < y_i$ for all i, where x_i's are the solutions of the following equations:

$$x_i = e_i + \sum_{k \neq j} q_{i,k} \cdot x_k.$$

Comparing it to equation (6), it is clear that $x_i = h_{i,j}$. All banks default if and only if $h_{i,j} < y_i$. $\square$

Proof of Proposition 3. The proof is identical to that of Proposition 1 by interchanging **Y** with **E**. $\square$

B Additional figures and tables

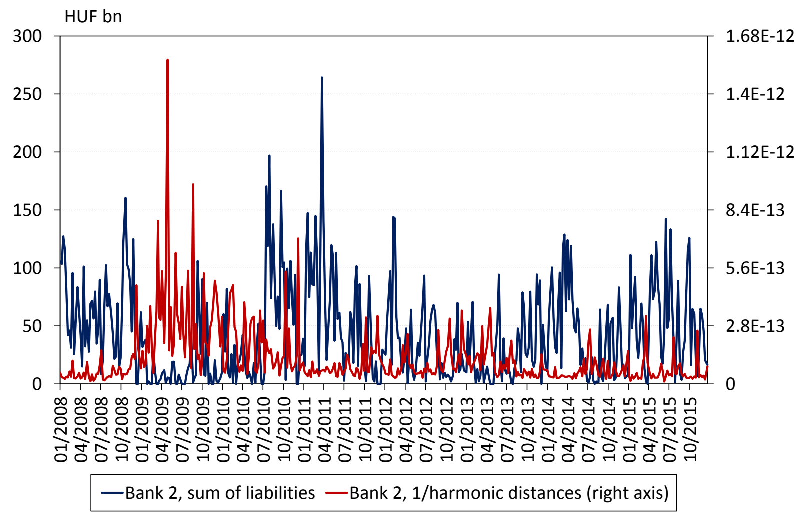

Figure 13: Time evolution of sum of interbank liabilities and harmonic distances of Bank 2 (O-SII) in Hungary

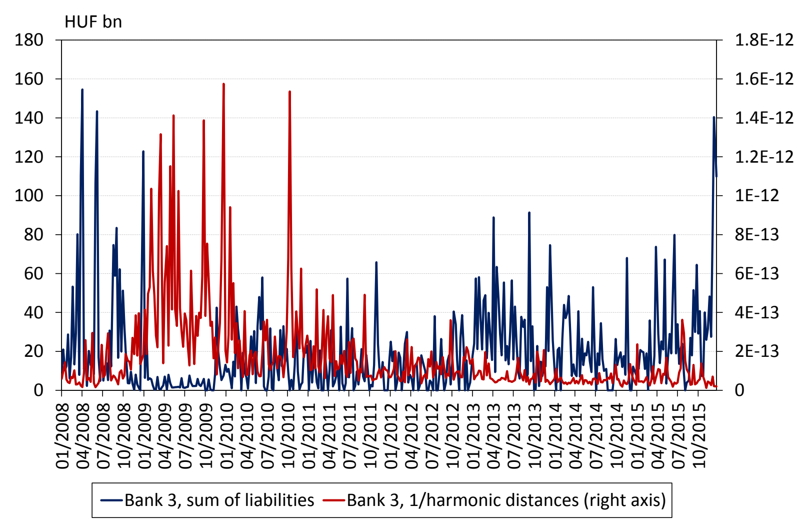

Figure 14: Time evolution of sum of interbank liabilities and harmonic distances of Bank 3 (O-SII) in Hungary

harmonic distances:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.588	0.094	0.589	0.092	0.647	0.081
0.2	0.653	0.081	0.654	0.079	0.661	0.083
0.4	0.750	0.071	0.744	0.071	0.750	0.074
0.6	0.809	0.066	0.811	0.068	0.809	0.068

extended harmonic distances:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.592	0.094	0.592	0.092	0.594	0.096
0.2	0.655	0.080	0.656	0.079	0.649	0.080
0.4	0.751	0.071	0.746	0.071	0.751	0.073
0.6	0.810	0.066	0.812	0.068	0.810	0.068

weighted degree:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.864	0.040	0.867	0.037	0.867	0.039
0.2	0.885	0.037	0.886	0.037	0.885	0.037
0.4	0.925	0.031	0.917	0.032	0.920	0.030
0.6	0.940	0.028	0.941	0.028	0.940	0.029

eigenvector:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.825	0.050	0.823	0.048	0.826	0.050
0.2	0.851	0.046	0.852	0.046	0.850	0.047
0.4	0.895	0.039	0.892	0.040	0.896	0.039
0.6	0.923	0.033	0.924	0.035	0.922	0.036

Bonacich:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.575	0.098	0.576	0.095	0.576	0.105
0.2	0.645	0.083	0.645	0.081	0.636	0.094
0.4	0.743	0.072	0.739	0.072	0.744	0.075
0.6	0.744	0.075	0.801	0.066	0.799	0.066

concentration:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.588	0.111	0.588	0.109	0.589	0.114
0.2	0.653	0.092	0.653	0.093	0.643	0.098
0.4	0.713	0.082	0.712	0.085	0.710	0.083
0.6	0.710	0.083	0.724	0.092	0.725	0.089

...

closeness:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.827	0.045	0.825	0.047	0.824	0.045
0.2	0.845	0.042	0.843	0.055	0.845	0.042
0.4	0.867	0.052	0.864	0.060	0.862	0.055
0.6	0.860	0.096	0.856	0.100	0.858	0.091

betweenness:

c	1		2		3	
α	avg.corr.	std.dev.	avg.corr.	std.dev.	avg.corr.	std.dev.
0.1	0.827	0.045	0.825	0.047	0.824	0.045
0.2	0.845	0.042	0.843	0.055	0.845	0.042
0.4	0.867	0.052	0.864	0.060	0.862	0.055
0.6	0.860	0.096	0.856	0.100	0.858	0.091

Table 10: Rank correlations corresponding to Table 1

harmonic distances

α	avg.corr.	std.dev.
0.1	0.588	0.093
0.2	0.654	0.080
0.4	0.747	0.073
0.6	0.811	0.066

extended harmonic distances

α	avg.corr.	std.dev.
0.1	0.591	0.093
0.2	0.656	0.080
0.4	0.748	0.073
0.6	0.812	0.066

weighted degree

α	avg.corr.	std.dev.
0.1	0.867	0.039
0.2	0.887	0.036
0.4	0.919	0.032
0.6	0.941	0.027

eigenvector

α	avg.corr.	std.dev.
0.1	0.823	0.050
0.2	0.852	0.046
0.4	0.894	0.041
0.6	0.924	0.034

Bonacich

α	avg.corr.	std.dev.
0.1	0.575	0.099
0.2	0.645	0.088
0.4	0.741	0.075
0.6	0.802	0.066

concentration

α	avg.corr.	std.dev.
0.1	0.587	0.111
0.2	0.649	0.097
0.4	0.711	0.084
0.6	0.728	0.088

closeness

α	avg.corr.	std.dev.
0.1	0.825	0.046
0.2	0.845	0.044
0.4	0.864	0.057
0.6	0.861	0.089

betweenness

α	avg.corr.	std.dev.
0.1	0.738	0.066
0.2	0.770	0.059
0.4	0.778	0.078
0.6	0.801	0.086

Table 11: Rank correlations corresponding to Table 2.