Latin American Journal of Central Banking 4 (2023) 100098

Contents lists available at ScienceDirect

Latin American Journal of Central Banking

journal homepage: www.elsevier.com/locate/latcb

Redundancy of Centrality Measures in Financial Market Infrastructures

Constanza Martínez-Ventura[†,*], Ricardo Mariño-Martínez, Javier Miguélez-Márquez

Financial Infrastructure Oversight Department, Banco de la República (Central bank of Colombia), Colombia

ARTICLE INFO

JEL:
G20
C38
C23

Keywords:
Centrality
Robust principal component analysis
Redundancy analysis
Clustering analysis

ABSTRACT

The concept of centrality is widely used to monitor systems with a network structure because it allows identifying their most influential participants. This monitoring task can be difficult if the number of system participants is considerably large or if the wide variety of centrality measures currently available produce non-coincident (or mixed) signals. This document uses robust principal component analysis to evaluate a set of centrality measures calculated for the financial institutions that participate in Colombia's four financial market infrastructures. The results obtained are used to construct general indices of centrality, using the most robust measures of centrality as inputs and leaving aside those considered redundant.

1. Introduction

Financial market infrastructures (FMIs) are the systems through which the clearing, settling, and recording of transactions (payments, securities, and derivatives) take place (BIS-PFMI, 2012). The timely and efficient performance of these market infrastructures ensures the smooth functioning of the payment system, the financial system, and, therefore, the economy as a whole.[1] However, the changes that these market infrastructures may experience, caused by variations in the activity of the institutions that participate in these systems or by unexpected exogenous shocks, have encouraged central banks and supervisory authorities to conduct monitoring activities aimed at identifying the cases or system participants that need to be studied in more detail.

Systems with a network structure, like the FMIs and financial markets, are often studied using measures of centrality since these allow them to identify their most influential participants. Centrality measures have been used to establish the systemic importance of financial institutions in payment systems (Soramäki and Cook, 2013; Baek et al., 2014) and unsecured interbank markets (Temizsoy et al., 2017; Rovira and Spelta, 2019). These measures have also been used to assess systemic risk (Battiston et al., 2012; Dungey et al., 2014), monitor systemic risk (Battiston et al., 2016), identify liquidity providers and liquidity hoarders in payment systems (Soramäki and Cook, 2013), and differentiate networks of financial institutions (León et al., 2021).

* Corresponding author.
E-mail addresses: amartive@banrep.gov.co (C. Martínez-Ventura), jmarinma@banrep.gov.co (R. Mariño-Martínez), jmiguema@banrep.gov.co (J. Miguélez-Márquez).

† The authors thank Pamela Cardozo, Clara Machado, Carlos León, Freddy Cepeda, Fredy Gamboa, and the two referees for their comments and suggestions. They also want to thank Fabio Ortega for suggesting this research topic, and Jefferson Marin and Juan Camilo Montañez for preparing and sharing data.

[1] According to the BIS, the safe and efficient functioning of FMIs is essential to guarantee the reliable transfer of funds and securities between the participants of the financial system, and to assure that the implementation of the monetary policy can spread quickly throughout the economy (BIS-PFMI, 2012).

https://doi.org/10.1016/j.latcb.2023.100098
Received 15 September 2022; Received in revised form 18 April 2023; Accepted 9 May 2023

Similarly to Soramäki and Cook (2013) and Baek et al. (2014), this paper uses centrality measures to identify the systemic importance of financial institutions in FMIs, since these institutions are precisely those that, in the event of non-complying to their commitments (e.g., payments or collateral delivery), could have a considerable impact on their respective network. Centrality measures can facilitate the monitoring tasks that must be carried out to detect warning signs in the normal functioning of these market infrastructures. This goal can be achieved by identifying changes in the relative importance of the most influential participants in a system. Still, the wide variety of centrality measures currently available can make monitoring each financial institution under all measures an endless challenge, either because the total number of participants in the system is considerably large and/or because the signals extracted from the measures do not coincide.

We propose to simplify this monitoring task by constructing general centrality indices using robust principal component analysis —RPCA— (Croux and Haesbroeck, 2000), which is a multivariate statistical technique that allows separating the most relevant variables (i.e., centrality measures) from those considered redundant while producing results robust to atypical data. The general centrality indices are constructed using the results of the first principal component, as this explains the largest variance of the dataset. Since the information contained in the redundant variables is captured by the retained (non-redundant) variables or by linear combinations of the latter (represented by the first principal component), the centrality indices discard all redundant variables.

RPCA is used to examine the centrality of the financial institutions that participate in four financial market infrastructures of Colombia: the large-value payment system, the foreign exchange clearing house, and two central securities depositories. The first of these FMIs (CUD) is the conduit all financial institutions use to transfer large-value payments. The second FMI is the foreign exchange clearing house (CCDC) in charge of the clearing and settling of peso/dollar transactions. The remaining FMIs are central securities depositories: the Central Securities Depository —DCV— provides account and custody services for sovereign debt securities, and Deceval provides these services for corporate and non-sovereign government securities and equities.[2]

In this study, each FMI is envisioned as a network of financial institutions connected by transactions defined by payments, trades, or similar agreements. The study period goes from January 2, 2019, to August 31, 2020, and includes daily data on the respective transactions. A total of twenty-six centrality measures with a daily frequency were computed and used to construct indices of centrality for each FMI. These indices include nineteen measures in the case of CUD, twenty measures for the CCDC, sixteen measures for DCV, and eleven measures for Deceval.[3] Centrality indices provide relevant information on system participants (i.e., ranking of centrality) and can therefore be used as tools to monitor FMIs: they can facilitate the identification of the system participants that can produce substantial impacts on the network or considerable changes in its stability. This topic is relevant for central banks, as they are the authorities responsible for monitoring these FMIs and ensuring the safe and efficient functioning of the payment system.

Some robustness checks were implemented to evaluate how well RPCA identifies redundant variables, which compared several statistical methods that also allowed to determine which variables should be retained and which should be discarded. These checks include two versions of RPCA based on subsamples of data, the redundancy analysis —RDA— (Kelley, 1940; Rao, 1964; van den Wollenberg, 1977), and a clustering method. Centrality measures identified as redundant by applying RPCA to the entire sample of observations remain nearly the same with these alternative methods of data reduction.

Besides these robustness checks, we evaluated possible changes in the ranking of the most central financial institutions during the lockdown period introduced by the government at the beginning of 2020 to contain the spread of the COVID-19 virus. Some changes represent the reorganization of the same financial institutions in the ranking, and others the entry of new institutions into the ranking.

The paper is organized as follows: Section 2 provides a general description of the centrality measures, and Section 3 presents a brief explanation of the statistical methods used in this study: RPCA, RDA, and cluster analysis. Section 4 describes the FMIs, and Section 5 presents the main results and the rankings of the most central participants in each network.

2. Network centrality measures

In the theory of graphs, a network is a graphical representation of a complex system composed of nodes that may (or not) be connected by edges or links. The links connecting pairs of nodes vary only when there are differences in the strength of the relationships. These cases correspond to weighted edges and allow the analyst to identify stronger from weaker links. In contrast, the strength of relationships in unweighted links does not vary for the nodes that compose the network. Similarly, the nodes in a network can be the same size when there are no differences between them, or they can have different sizes when scaled by a criterion pre-established for this purpose.

A network can be described mathematically by an adjacency matrix (A) of $N \times N$ dimension, representing pairs of connected nodes (i and j) with nonzero elements and nonconnected nodes with zeroes. In the binary case, the connected pairs are denoted with elements equal to one (A_{ij} =1), and the nonconnected are again represented with zeroes (A_{ij} =0). In directed networks, a nonzero A_{ij} element represents a link pointing from j to i, independent of a link pointing in the opposite direction A_{ji} (Newman, 2008). The matrix is symmetric in undirected networks, hence, $A_{ij} = A_{ji}$.

The centrality concept was initially postulated by Camille Jordan in 1869 to identify the most influential node (element) in a network (Hage and Harary, 1995). Since then, several centrality measures have been proposed, ranging from count methods to

[2] Central securities depositories aim at ensuring the integrity of securities in transactions (see BIS-CPSS, 2003).

[3] The Financial Infrastructure Oversight Department is using a centrality index obtained from applying PCA to six centrality measures (i.e., in-degree, out-degree, hub, authority, in-strength, and out-strength).

algorithms used to study financial networks. This section briefly explains some of these centrality measures, assuming the pairs of nodes are in the adjacency matrix A and denoted as A_{ij}. This document's adjacency matrix A transposed will be denoted as A^T.

Degree centrality measures the connectedness of a node by the number of links to its neighbors (Barrat et al., 2004). In its simplest form, degree centrality applies to undirected networks. Still, there are two additional variants for directed networks: one that counts the links that point towards a node (i.e., in-degree) and another that counts the links that point outside the node (i.e., out-degree).[4] In all versions of the degree centrality measure, the count of neighbors should be adjusted by the total number of nodes ($N - 1$) with which the node under study (i) can have an interaction in the network (G):

$$Degree(i) = \sum_{j \epsilon G} \frac{A_{ij}}{N-1}. \tag{1}$$

Thus, the most central node has the most links (Newman, 2008).

Closeness centrality indicates how near a node is to the other nodes of the network based on the length of the shortest paths (Goldbeck, 2013). This measure is formally calculated by the inverse function of the sum of the average shortest distances between a node and all the other nodes in the network (Bavelas, 1950).

$$Closeness(i) = \frac{1}{\sum distance\ (i, j)} \tag{2}$$

In this measure, the most central nodes have the highest centrality results.

Betweenness centrality determines how often a path between two nodes must go through a given node. Betweenness centrality is calculated as the ratio of the total number of shortest paths (σ_{st}) between a pair of nodes (s and t) and the number of those paths that go through node i ($\sigma_{st}(i)$) (Batool and Niazi, 2014).[5]

$$Betweenness(i) = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}} \tag{3}$$

As a result, the most central node will be the highest betweenness result.

Eccentricity centrality identifies the node that will cause the highest propagation of an effect in a network (Hage and Harary, 1995; Batool and Niazzi, 2014).

$$Eccentricity(i) = \frac{1}{\max\{distance\ (i,\ j)\}} \tag{4}$$

This measure is assessed in three steps. The first step quantifies the shortest path between a given node i and all the other nodes j in the network (*distance* (i, j)). The second step compares the resultant distances for every pair of nodes and identifies the maximum distance per node (i). The third step computes the inverse of the maximum distance found in the previous step. The node with the highest result will be the most central in the network.

Eigenvector centrality defines a node's importance (centrality) by connecting the most influential nodes in an undirected network (Pozzi et al., 2017). This iterative method of linear algebra is used to solve a general equation that, in its matrix notation, is of the form:

$$A^T v = \lambda v \tag{5}$$

In the first iteration, this method assumes the eigenvalue (λ) is a vector of ones, and from there on, it replaces the eigenvalues with the eigenvectors (v) obtained in the previous iteration until the solution (eigenvector) converges, producing n solutions corresponding to the n values of λ. As this measure defines the importance of a node as a function of the nodes with which the former is interacting, the most central will be that with the highest eigenvector centrality (Bonacich and Lloyd, 2001).

Hypertext-induced topic search (HITS) is an algorithm used to analyze search methods in the world wide web that recursively identify hubs (i.e., web pages that point to many other web pages) and authorities (i.e., web pages to which other several web pages point to). The HITS algorithm updates separate operations on the weights: one for authorities and another for the hubs. The equilibrium weights are found alternating these operations until a fixed point is reached. To this aim, an eigenvector algorithm based on matrix products defined on an adjacency matrix A of G_σ (i.e., subgraph associated with a query string σ) is computed, where the optimal weights are determined recursively, starting from initial vectors, and updating these weights with the eigenvectors computed for $A^T A$ and AA^T. The solution to the matrix product $A^T A$ defines the authorities, and AA^T determines the hubs (Kleinberg, 1999). This algorithm is used in network analysis by redefining web pages as nodes.

PageRank is an algorithm based on the connections of web pages to the most influential web pages. It measures the probability of visiting a web page as a function of its incoming links (Brin and Page, 1998). By redefining web pages as nodes, this algorithm describes the probability that starting from node j, node i will be visited, as can be seen in Eq. (6):

$$PageRank(i) = d \sum_{j \to i} \frac{page - rank(j)}{N_j} + \underbrace{(1-d)}_{dE(i)} \tag{6}$$

[4] Two alternative measures can be obtained by weighting these measures: in-strength (i.e., in-degree weighted) and out-strength (i.e., out-degree weighted).

[5] According to Newman (2008), a path is a sequence of nodes crossed by following links across a network.

Hence, the visiting probability depends on the eigenvector of adjacency matrix (A) denoted as page-rank(j), the overall number of links from node j that point to node i (N_j), the uniform distribution function (E) that makes the probability of jumping to a random node equally likely for all nodes in the network, and a parameter between zero and one (d) included to avoid traps in which sink nodes impede finding a solution.[6] The page-rank algorithm is defined recursively, computing the eigenvector P_i of matrix A at the maximal eigenvalue, with eigenvectors considered probabilities. Therefore, a node will have a high page rank if the sum of the ranks of its incoming links is high. This algorithm produces a ranking of the global importance of nodes called PageRank, where the nodes with high-rank probabilities will be considered the most central.

CheiRank is an algorithm that follows the same idea as PageRank in that it calculates the ranking of nodes based on their connections but uses the links in the opposite direction. To produce a new matrix (A^*), the page-rank's adjacency matrix (A) is utilized with all link directions reversed. As the page-rank algorithm, CheiRank consists in computing the eigenvector $P_i{}^*$ of A^* with the maximal eigenvalue, with eigenvectors considered probabilities and used to rank the nodes as a function of the outgoing links (see Chepelianskii, 2010). CheiRank eigenvectors are also considered as probabilities of visiting nodes.

Random Walk Betweenness counts how often a node is traversed by a random walk between two other nodes (Newman, 2005). Unlike the betweenness measure, which is based on shortest paths, this measure uses a random walk to generalize the betweenness idea to all nodes in the network. This measure is given by the inverse of the mean first-passage (m_{ij}) from node i to j which depends on the expected number of steps (n) taken until the first arrival to node j starting in node i and the probability that the Markov chain (i.e., probability of transitioning from node j to node i) first returns to node j ($f_{ij}^{(n)}$) in exactly n steps: $m_{ij} = \sum_{n=1}^{\infty} n f_{ij}^{(n)}$. Thus, the average importance of node j relative to the set of all nodes (R) is:

$$I(j|R) = \frac{1}{\frac{1}{|R|}\sum_{i\epsilon R} m_{ij}}. \tag{7}$$

This measure, called Markov centrality, produces a ranking that designates the most central nodes as those with the highest results (White and Smith, 2003).

SinkRank is an algorithm that identifies systemically important banks in a payment system by executing a simulated failure of a bank and identifying the most affected counterparts (Soramäki and Cook, 2013). The SinkRank node is calculated as the absorbing node, and all the others are the non-absorbing nodes (Baek et al., 2014). SinkRank depends on the likelihood (transition probability) that a random walk moves from one node to another, which will differ from zero for non-absorbing nodes. For absorbing nodes, that probability will be zero as it defines the termination of the walk. The SinkRank measure is formally given by the inverse of the average sink distance of each non-absorbing node, which is the same ratio between the number of non-absorbing nodes (n - m) and the sink distance of a node ($\Sigma_i\Sigma_j q_{ij}$):

$$SinkRank = \frac{n-m}{\sum_i \sum_j q_{ij}}. \tag{8}$$

In Eq. (8), q_{ij} denotes an element of the fundamental matrix Q that defines the number of times the random walk at state i is expected to visit node j before being absorbed by a node. Thus, high sink ranks correspond to the most central nodes, identified as those for which the simulated failure cause the strongest impact on the system.

SourceRank is a centrality measure that identifies the liquidity providers in a payment system. In general terms, source-rank is the opposite criterion to sink-rank, aimed at identifying liquidity hoarders (Soramäki and Cook, 2013).

3. Some basics of statistical discarding methods

The most central financial institutions in FMIs should be monitored closely due to the adverse effects they may cause in these systems if they fail to meet their obligations or commitments. We implement robust principal component analysis to identify these institutions, using the centrality measures described in the previous section. As this procedure allows the construction of a general index using individual measures of centrality as inputs, those considered redundant will be discarded. Below is a brief explanation of the statistical techniques used to establish the centrality measures that should be retained or discarded.

3.1. Robust principal component analysis

The analysis of principal components is a multivariate statistical technique proposed by Pearson (1901) and formally developed by Hotelling (1933) to transform a set of 'n' variables into a new set of 'p' synthetic variables that are linear combinations of the original ones, for a given matrix Y, of $n \times n$ dimension, principal component analysis (PCA) consists in computing the eigenvalues and eigenvectors of Y^TY. This statistical technique reduces data by removing redundant (i.e., multicollinear) variables from the dataset while creating new synthetic variables —called the principal components—. These 'p' new variables ($p << n$) explain, in decreasing order, the variation in the data not explained by the previous component. Thus, the first principal component explains the maximum amount of variance of the original dataset. The second principal component explains the second highest amount of variance not

[6] This algorithm could be affected by traps that accumulates rank but never distributes any rank. The adjustment term dE(i) that is included to overcome this problem, usually takes a value of 0.15 (Brin and Page, 1998).

explained by the first component, and so forth. Thus, the selected components jointly represent the highest variance share of the original variables.

A frequent problem in PCA implementation relates to outliers, corruption, and measurement errors in the dataset, as these atypical data points contaminate the principal components. The most common solution to this problem is replacing the classical covariance or correlation matrix with a robust estimator. This variation on the classical version of PCA is known as robust principal components analysis (RPCA). It consists of splitting the general data matrix into a low-rank matrix well described by a few patterns and a sparse matrix containing all atypical observations of the dataset. As there are many possible ways to partition the general matrix, the problem then consists of finding the solution that produces the optimal robust estimators of the eigenvalues and eigenvectors of the sample covariance or correlation matrix (see Croux and Haesbroeck, 2000).

When multivariate atypical observations are present, the estimate of Mahalanobis distances used to identify these data points is distorted. A correct identification of atypical data is possible if the estimate of the sample mean and the correlation or covariance matrix is robust; that is, if they are not excessively affected by these observations (see Verardi and Dehon, 2010).[7]

In this document, we derive the principal components from the correlation matrix and use the robust estimator of multivariate outlier data and scatter matrix (i.e., correlation matrix) given by the minimum covariance determinant (MCD) estimator proposed by Rousseeuw (1985).[8] To identify the number of components to retain in the empirical implementation of RPCA, we use the scaled Kaiser-Guttman test that selects the components that exceed 70% of the average eigenvalue (see Jolliffe, 1972).

3.2. Redundancy analysis

The redundancy analysis (RDA) is a multivariate statistical technique proposed by Kelley (1940) and developed by Rao (1964) and van den Wollenberg (1977) that allows obtaining a reduction in the dimensionality of data by implementing PCA on the projection of the dependent variables on the space spanned by the explanatory variables (Isräels, 1992). Since RDA combines the linear regressions and PCA, these data reduction techniques only diverge on the fact that PCA is a univariate method while RDA is multivariate. Therefore, the former method is frequently considered a particular case of the latter (see van den Wollenberg, 1977).

As robustness checks, we use RPCA in the second step of the RDA procedure. This variation to the classical RDA is intended to correct results for the possible existence of outliers and other atypical data points and make them comparable to those obtained with the plain RPCA. This latter technique will, again, be based on the correlation matrix and the MCD robust estimator.

3.3. Clustering analysis

Another statistical technique used in this paper is the clustering analysis. According to Jolliffe (1973), most clustering methods produce similar results for data reduction. Among these methods, single-linkage clustering is faster than average-linkage clustering, but neither outperforms the other. In our robustness checks, we use the average-linkage method defined on outer clustering, which selects the last variable to join the main group(s) in the cluster tree (i.e., dendrograms). This method will be implemented using the average dissimilarity of observations (i.e., the Euclidean distance) between pairs of centrality measures.[9]

4. Data description

Financial market infrastructures (FMIs) are multilateral systems that carry out the clearing and settlement of payments, securities, derivatives, and other financial transactions (BIS-PFMI, 2012). Monitoring activities on these market infrastructures may help financial authorities to detect warning signs that deserve a more profound analysis and probably specific actions to contain their potential negative effects on the system. This section briefly describes the FMIs on which the statistical analysis will be performed.

4.1. The large-value payment system (CUD)

The FMI that provides clearing and settlement services to institutions that send large-value payments in local financial markets is CUD, owned and operated by Banco de la República. This large-value payment system works in a real-time gross settlement mode, settling each transaction immediately and at its gross value, subject to the condition that the balances in the sender's account are sufficient to cover its payment orders.

Several types of payments are sent through the CUD, among which are found intraday interbank funding, payments related to other clearing and settlement systems (i.e., the Foreign Exchange Clearing House (CCDC) and central securities depositories (DCV and Deceval)), government payments, and payments related to the implementation of the monetary policy.

CUD's network is presented in Fig. 1, with nodes symbolizing the system's participants and links representing the average value of daily payments in a two-month period of 2020. All participants have direct access to the system; therefore, each can initiate payments without having to resort to any other institution. The average number of system participants during 2019 was 138, of which more than 80% were financial institutions (primarily banks, trust companies, brokerage firms, and commercial financing companies).

[7] When there are multivariate outliers in the dataset, the classical estimate of the sample mean is affected and the covariance matrix is inflated, which generates results substantially biased (Verardi and Dehon, 2010).

[8] Other robust estimators of multivariate location and scatter are the Huber M-estimator and the S-estimator (see Croux and Haesbroeck, 2000).

[9] The Euclidean distance is the shortest path between two points, since it is defined by a line that joins them.

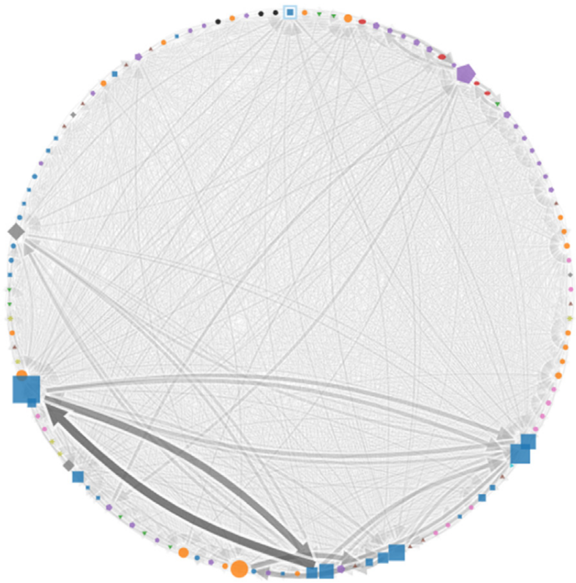

Fig. 1. CUD's network with system participants represented by nodes and their transactions by links pointing to the institutions receiving payments. Squared nodes represent banks, brokerage firms are circles, mutual funds are pentagons, and commercial financing companies are rhombuses. The size of the nodes corresponds to the average centrality score obtained from the hub measure. More comprehensive links denote higher values of payments.

4.2. The foreign exchange clearing house of colombia (CCDC)

The market infrastructure that provides multilateral netting and settlement services for foreign exchange transactions is the Foreign Exchange Clearing House of Colombia (CCDC). This FMI serves two functions in the foreign exchange market. Firstly, it mitigates the risks related to foreign exchange transactions of Colombian pesos and US dollars settled on the same day ($t + 0$) and up to three days after the trade ($t + 1$, $t + 2$, and $t + 3$). To this aim, the CCDC handles the counterparty risk by employing the payment-versus-payment mechanism.

For the market risk, it requires guarantees from the participants, and for the liquidity risk, it uses credit lines acquired with local financial institutions. The CCDC is not a central counterparty, and hence, in response to extreme events (such as liquidity deficits that the mentioned risks mitigation mechanisms cannot cover, multiple non-compliance in payment of multilateral obligations from the participants, or the impossibility that this FMI provides its services), the system participants will have to settle transactions bilaterally.[10] Secondly, CCDC facilitates the liquidity savings that result from multilateral netting.[11]

The CCDC's network representation is shown in Fig. 2, with its participants as nodes, connected through links that denote the average daily bilateral US dollar sold amount in a two-month period of 2020. In CCDC participate 33 financial institutions (primarily banks and brokerage firms), all with direct access to the system.

4.3. The central securities depository (DCV)

Two central securities depositories are responsible for the clearing and settlement services of securities transactions in the domestic market: DCV and Deceval.[12] The securities depository for local sovereign securities, DCV, is a settlement system owned and administered by Banco de la República. In this system, the settlement of transactions is based on the delivery versus payment mechanism and is conducted in real time on the large-value payment system (CUD). The central bank also determines the access of financial institutions to DCV and takes one of two forms: as a direct depositor (i.e., accepted as holder of securities in their position or the position of third parties) or as an indirect depositor (i.e., accepted as the holder of a subaccount) through one of the direct depositors.[13]

---

[10] From December 14, 2020, the FMI in charge of providing the clearing services for the peso/dollar transactions is the Colombian Central Counterparty (i.e., Cámara de Riesgo Central de Contraparte S.A.). However, it was not until February 1st, 2021, that these services began to operate through the novation process. As this document started long before that change, the results for these transactions are solely based on data from the CCDC.

[11] During 2019, the average daily liquidity savings that emerge as a result of the multilateral netting procedure was 86%, which signifies that system's participants paid only 14% of the gross value of transactions.

[12] These securities are currently mostly represented in a dematerialized (electronic) form.

[13] During 2019, DCV settled a daily average of 2,122 operations and COP 39 trillion in nominal value. Of this, six billion corresponded to the primary market, COP 19 trillion to the secondary market and COP 20 trillion to monetary operations (services provided to the Banco de la República, which

Fig. 2. CCDC's network with participants represented by nodes and their bilateral transactions by links pointing to the institutions that receive the amount of US dollars purchased. Squared nodes represent banks, brokerage firms are circles, and financial corporations are rhombuses. The size of the nodes corresponds to the average centrality score obtained from the hub measure. More comprehensive links denote higher gross values of foreign exchange transactions.

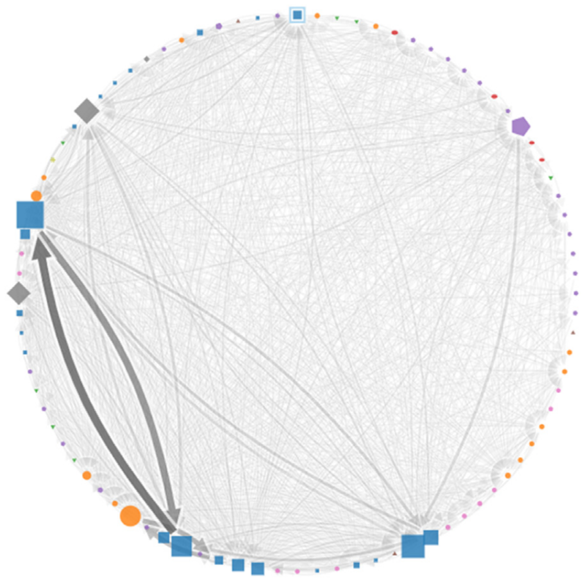

Fig. 3. DCV's network with participants represented by nodes and their securities transactions by links pointing to the institutions receiving liquidity. Squared nodes represent banks, circles are brokerage firms, pentagons are mutual funds, and rhombuses are financial corporations. The size of the nodes corresponds to the average centrality score obtained from the hub measure. More comprehensive links denote higher values of transactions.

The network representation of this FMI, shown in Fig. 3, was constructed using daily average data of the securities transactions in a two-month period of 2020. The average number of direct depositors in DCV during December 2019 was 126, represented mainly by pension and severance funds, banks, trust companies, and the public sector.

involve open market operations and liquidity provisions to the large-value payment system). This depository held COP 323 trillion at the end of 2019, of which 97% corresponded to securities issued by the national government and the remaining to securities issued by the Fund for Financing the Agricultural Sector (3%).

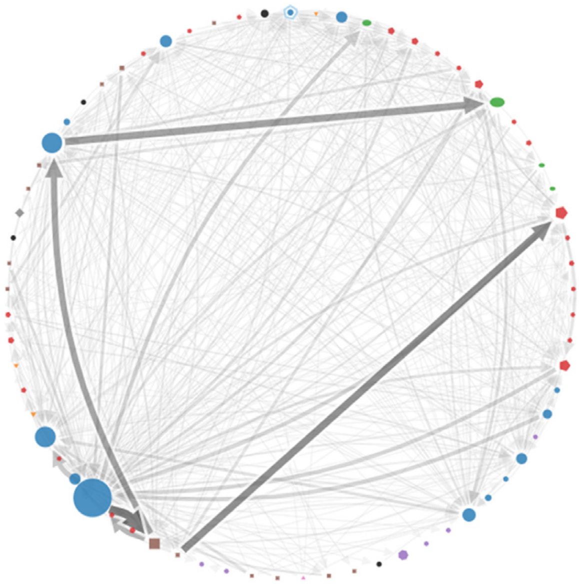

Fig. 4. Deceval's network with participants represented by nodes and their securities transactions by links pointing to the institutions receiving liquidity. Squared nodes represent banks, circles are brokerage firms, pentagons are mutual funds, and rhombuses are financial corporations. The size of the nodes corresponds to the average centrality score obtained from the hub measure. More comprehensive links denote higher values of transactions.

4.4. The centralized securities depository of colombia (Deceval)

Deceval is the privately-owned securities depository and securities settlement system, which provides deposit, clearing, and settlement services for corporate and non-sovereign government securities and depository services for the equity market. This depository classifies its depositors in direct (i.e., financial institutions supervised by the Financial Superintendency of Colombia, public entities, issuing entities with securities registered in the national securities registry, intermediaries that have entered into a deposit agreement with Deceval, and other centralized securities depositors) and indirect (i.e., persons who cannot be direct depositors by regulation and can only sign a contract with a direct depositor).[14]

Deceval's network is shown in Fig. 4, with nodes and links between nodes representing daily average data for a two-month period of 2020. As in the former FMIs, participants are represented by nodes and their transactions by links connecting pairs of nodes. The average number of direct participants in Deceval during 2019 was 71.

We examine twenty-six centrality measures for each FMI and use them to construct general centrality indices. These measures were calculated using daily data (from January 2, 2019, to August 31, 2020) of the payments settled through the CUD, the gross value of peso/dollar transactions for the CCDC, the securities transactions for DCV, and the transactions' value of purchase and sales, sell/buy-backs, and repurchase agreements for Deceval. The entire set of centrality measures corresponds to the alternatives the theory supports. Their calculation considered whether they should be based on directed or undirected networks, as explained in Section 2. These measures are considered in both their weighted and unweighted forms. The former versions of these measures were calculated using the value of daily transactions between system participants as weights.

Summary statistics of these measures are provided in Table A.1 in the appendix. As can be seen in that table, the total number of observations for the sample period is larger for CUD (46,303) than for CCDC (12,939), DCV (24,135), and Deceval (18,999). However, the effective number of observations is slightly shortened due to the existence of gaps in some centrality measures.[15] All these measures were transformed into their percent participation per day to avoid that differences in the measurement units will alter the results.[16]

[14] In 2019 the transactions carried out in Deceval, including primary and secondary market operations (fixed and variable income) and money market operations (repos, sell/buy-backs and securities lending) with their respective reverse transactions and cash guarantees, represented a daily average of 5,239 operations and COP 3.75 trillion. As a depository, at the end of 2019 this system held COP 561 trillion, 58% of which corresponded to equities (ordinary and preferential), 23% to term certificates of deposits, 10% to ordinary bonds, and 9% to other instruments (commercial papers, acceptances, among others).

[15] Data loss may arise from networks with links that do not allow the calculation of centrality measures given that the network's data do not adjust to the algorithm and/or its restrictions. That is the case of weakly connected nodes, for which some centrality measures either cannot be computed or produce a result that tends to infinitum.

[16] Comparable results were obtained when the usual standardization (i.e., zero mean and unit variance) was applied to the data set.

Table 1
Robust Principal Components for CUD.

Centrality measure	Components					Unexplained variance
	PC1	PC2	PC3	PC4	PC5	
Degree	0.236	0.162	−0.094	0.023	−0.039	1.08%
Degree weighted	0.210	−0.266	0.023	−0.003	−0.054	2.21%
Indegree	0.223	0.113	−0.138	0.328	0.111	2.31%
Indegree weighted	0.210	−0.263	0.024	0.011	−0.053	2.85%
Outdegree	0.217	0.187	−0.039	−0.268	−0.176	1.92%
Outdegree weighted	0.209	−0.267	0.022	−0.017	−0.055	2.58%
Closeness	0.157	0.182	0.246	−0.216	0.323	23.73%
Closeness weighted	0.030	0.205	0.544	0.234	−0.383	14.79%
Betweennes	0.165	0.203	−0.151	0.022	−0.154	34.86%
Eccentricity	0.108	0.081	0.235	−0.265	0.725	12.64%
Eccentricity weighted	0.045	0.148	0.656	0.236	−0.013	14.10%
Eigenvector	0.231	0.176	−0.066	−0.017	−0.031	4.01%
Eigenvector weighted	0.217	−0.219	0.082	−0.007	0.019	6.65%
Authority	0.222	0.076	−0.124	0.314	0.144	6.73%
Authority weighted	0.162	−0.185	0.024	0.032	−0.038	45.00%
Hub	0.212	0.172	−0.008	−0.293	−0.121	8.02%
Hub weighted	0.151	−0.186	0.009	−0.080	0.023	49.81%
PageRank	0.208	0.165	−0.158	0.350	0.134	3.14%
PageRank weighted	0.210	−0.238	0.101	0.039	0.020	6.53%
CheiRank	0.217	0.177	−0.025	−0.276	−0.181	3.20%
CheiRank weighted	0.210	−0.241	0.053	−0.014	−0.054	7.00%
Random-walk betweenness	0.219	0.210	−0.020	−0.034	−0.003	7.69%
SinkRank	0.208	0.166	−0.158	0.347	0.134	3.08%
SinkRank weighted	0.211	−0.234	0.103	0.035	0.025	6.51%
SourceRank	0.217	0.177	−0.025	−0.276	−0.179	3.05%
SourceRank weighted	0.211	−0.237	0.052	−0.016	−0.051	7.04%
% of variance explained	**59.77%**	**15.51%**	**6.09%**	**4.00%**	**3.68%**	

Notes: authors' calculations.

5. Redundant centrality measures

One of the most used tools for monitoring networks is the general centrality indices, as they facilitate the identification of nodes that should be examined more carefully due to their relative importance in the system. Considering the level of information they provide, as these indices are usually based on an extensive dataset, it is necessary to know which variables should be retained and which should be discarded. We use RPCA on the correlation matrix of these measures to identify and discard those considered redundant and use the retained ones to construct general indices of centrality.[17]

5.1. Results based on robust principal components analysis

Prior to the statistical analysis with robust PCA (henceforth, RPCA), we checked whether this method could be applied to these FMIs by calculating the overall sampling adequacy test (i.e., Kaiser-Meyer-Olkin (Kaiser, 1974)) on the centrality measures. All results are close to unity (CUD: 0.9172, CCDC: 0.9150, DCV: 0.9163, Deceval: 0.8759), indicating that the correlation of these measures is high enough to factor the matrix of correlation coefficients. To determine the appropriate number of components to keep, the scaled Kaiser-Guttman test was utilized. The test measured the total amount of explained variance, and principal components exceeding the 70% threshold were selected.

The principal components (henceforth PC) are linear combinations (i.e., eigenvalues) of the original variables, represented by the centrality measures. Among these linear combinations, the retained components jointly explain a large part of the variance of these measures. These components for the CUD network explain 89.05% of the centrality measure total variance. Hence, the portion of variance that remains unexplained (i.e., 1 – 0.8905) arises because the components retained do not contain all the information about the centrality in this FMI.

The percentage of unexplained variance per measure is presented in the last column of Table 1 (10.95%). The centrality measures with the highest percentage of unexplained variance are hub weighted (49.81%), authority weighted (45.00%), betweenness (34.86%), closeness (23.73%), closeness weighted (14.79%), eccentricity weighted (14.10%), and eccentricity (12.64%).

For the CCDC, the four components retained jointly explain 91.05% of the total variance of transactions. Hence, the variance of the centrality measures not explained by the selected components is 8.95%. The eigenvectors corresponding to these principal components are reported in Table 2, along with the percentage of unexplained variance reported in the last column. The highest

[17] According to Jolliffe (2002), when all measures are in the same units, the covariance matrix might be more appropriate to implement PCA. Since that is not the case for our dataset, we use correlation matrices.

Table 2
Robust Principal Components for CCDC.

Centrality measure	Components				Unexplained variance
	PC1	PC2	PC3	PC4	
Degree	0.216	0.119	−0.113	−0.045	0.75%
Degree weighted	0.208	−0.203	0.132	0.032	0.24%
Indegree	0.210	0.107	−0.061	−0.342	1.10%
Indegree weighted	0.204	−0.199	0.143	−0.108	2.69%
Outdegree	0.208	0.124	−0.159	0.264	2.28%
Outdegree weighted	0.204	−0.201	0.117	0.172	2.39%
Closeness	0.210	0.149	−0.107	−0.022	3.62%
Closeness weighted	0.076	0.453	0.423	0.019	16.33%
Betweennes	0.184	0.005	−0.122	−0.106	30.20%
Eccentricity	0.147	0.195	−0.145	0.224	41.25%
Eccentricity weighted	0.057	0.405	0.664	0.101	10.36%
Eigenvector	0.210	0.158	−0.126	−0.029	2.60%
Eigenvector weighted	0.207	−0.195	0.123	0.039	1.71%
Authority	0.207	0.120	−0.062	−0.329	3.07%
Authority weighted	0.195	−0.216	0.175	−0.086	7.51%
Hub	0.205	0.135	−0.157	0.267	3.72%
Hub weighted	0.194	−0.219	0.157	0.153	7.12%
PageRank	0.210	0.099	−0.057	−0.343	1.25%
PageRank weighted	0.205	−0.188	0.124	−0.129	3.19%
CheiRank	0.208	0.115	−0.165	0.270	2.86%
CheiRank weighted	0.205	−0.186	0.086	0.190	3.25%
Random-walk betweenness	0.201	0.180	−0.134	−0.008	7.92%
SinkRank	0.211	0.096	−0.058	−0.343	1.15%
SinkRank weighted	0.205	−0.189	0.125	−0.132	3.18%
SourceRank	0.208	0.112	−0.164	0.264	2.80%
SourceRank weighted	0.204	−0.188	0.089	0.189	3.32%
% of variance explained	**72.16%**	**11.62%**	**4.26%**	**3.01%**	

Notes: authors' calculations.

results of that unexplained variance correspond to the measures of eccentricity (41.25%), betweenness (30.20%), closeness weighted (16.33%), and eccentricity weighted (10.36%).

The seven principal components retained for DCV explain 89.00% of the total variation of transactions in this securities depository (Table 3). The first component explains 44.10%, while the other components contribute with values below 16%. The average percentage of unexplained variance (11.01%) is mainly concentrated on measures like random walk betweenness (42.39%), hub weighted (37.66%), betweenness (33.79%), eccentricity (27.99%), closeness weighted (21.80%), and eccentricity weighted (15.99%).

For Deceval, the retained components explain 88.22% of the variance of the securities transactions in this depository (Table 4). The complementary percentage (11.78%), representing the percentage of variance not explained by this subset of components, depends fundamentally on eccentricity (47.79%), hub weighted (38.82%), authority weighted (37.21%), closeness (25.49%), random walk betweenness (17.90%), and closeness weighted (17.76%).

The eigenvectors corresponding to the first principal component are used to identify redundant centrality measures for each FMI. This component explains the highest degree of co-movement across centrality measures. These eigenvectors (scores) are standardized (by subtracting the mean and dividing them by the standard deviation) and reorganized in decreasing order to quickly identify the measures with the lowest contribution to the centrality of the FMI. The result for the centrality measures whose contribution to the first component is below the average scores of the retained measures is presented in Table 5.

The subset of redundant centrality measures for CUD and CCDC includes closeness weighted, betweenness, eccentricity (weighted and unweighted), hub (weighted and unweighted), and authority (weighted and unweighted). For the securities depository (DCV), this subset additionally contains indegree weighted, chei-rank weighted, random walk betweenness, and source-rank weighted. Deceval also includes sink-rank weighted, page-rank weighted, eigenvector weighted, degree weighted, out-degree weighted, and closeness.

RPCA penalizes the contribution of the redundant measures to the general results of centrality and, therefore, ends up discarding them. These measures are represented mainly by weighted versions of the original criteria, which suggests that they are well represented by their unweighted counterparts or by linear combinations of the retained ones, all represented by the first principal component. In addition, the weights utilized to calculate these measures are nearly identical because the quantity, frequency, or combination of transactions produce comparable outcomes.

The redundant measures are precisely those that exhibit the lowest eigenvectors in the first principal component (see column PC1 in Tables 1-4), a finding that coincides with the results of Jolliffe (1973) in that PCA rejects the variables associated with the last principal components. This premise also applies to alternative versions of this method of data reduction, like RPCA.

5.2. Robustness checks

The redundancy of the centrality measures is corroborated in four ways, using: *i)* RPCA on subsamples of measures, *ii)* RPCA on subsamples of system participants, *iii)* a constrained method of data reduction (i.e., RDA), and *iv)* a clustering method. For the first

Table 3
Robust Principal Components for DCV.

Centrality measure	Components							Unexplained variance
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	
Degree	0.274	0.119	−0.036	−0.186	0.015	−0.042	−0.102	1.28%
Degree weighted	0.213	−0.322	0.011	−0.053	−0.011	0.116	0.035	4.35%
Indegree	0.223	0.066	−0.333	−0.078	0.058	−0.143	−0.189	2.11%
Indegree weighted	0.175	−0.328	−0.150	−0.062	0.176	0.201	0.122	5.22%
Outdegree	0.227	0.130	0.273	−0.229	−0.033	0.073	0.021	2.66%
Outdegree weighted	0.204	−0.235	0.191	−0.027	−0.226	−0.017	−0.076	13.04%
Closeness	0.207	0.162	0.000	0.276	0.066	0.311	−0.131	14.79%
Closeness weighted	0.070	0.405	−0.044	0.158	0.000	0.055	0.184	21.40%
Betweennes	0.162	0.106	−0.028	−0.213	0.125	−0.320	0.376	33.79%
Eccentricity	0.148	0.137	0.007	0.317	0.042	0.401	−0.245	27.32%
Eccentricity weighted	0.066	0.337	−0.018	0.386	0.017	0.172	0.268	15.99%
Eigenvector	0.262	0.150	−0.039	−0.130	0.087	−0.007	−0.044	7.92%
Eigenvector weighted	0.221	−0.270	0.044	0.151	0.086	0.048	0.129	8.10%
Authority	0.222	0.051	−0.316	−0.056	0.043	−0.098	−0.288	5.31%
Authority weighted	0.009	−0.194	−0.103	−0.194	0.683	0.338	−0.005	6.56%
Hub	0.224	0.113	0.262	−0.211	−0.078	0.141	−0.068	6.41%
Hub weighted	0.142	−0.180	0.139	−0.081	−0.375	0.074	−0.175	37.66%
PageRank	0.230	0.078	−0.313	0.011	−0.078	−0.155	−0.131	3.34%
PageRank weighted	0.208	−0.226	−0.142	0.196	−0.218	0.062	0.296	5.45%
CheiRank	0.228	0.128	0.294	−0.142	0.018	0.061	0.024	4.57%
CheiRank weighted	0.148	−0.119	0.277	0.347	0.256	−0.365	−0.145	1.88%
Random-walk betweenness	0.168	0.104	−0.010	−0.071	0.141	−0.143	0.464	42.39%
SinkRank	0.231	0.081	−0.311	0.009	−0.077	−0.153	−0.127	3.16%
SinkRank weighted	0.210	−0.219	−0.136	0.200	−0.220	0.073	0.299	5.40%
SourceRank	0.229	0.132	0.292	−0.132	0.023	0.063	0.028	4.35%
SourceRank weighted	0.142	−0.112	0.261	0.362	0.259	−0.395	−0.153	1.84%
% of variance explained	**44.10%**	**15.33%**	**11.31%**	**6.41%**	**4.58%**	**4.52%**	**2.75%**	

Notes: authors' calculations.

Table 4
Robust Principal Components for Deceval.

Centrality measure	Components						Unexplained variance
	PC1	PC2	PC3	PC4	PC5	PC6	
Degree	0.306	−0.097	−0.062	−0.168	−0.017	−0.051	0.85%
Degree weighted	0.150	0.245	0.310	0.035	0.115	−0.010	4.04%
Indegree	0.222	−0.269	0.037	−0.152	0.271	−0.047	1.10%
Indegree weighted	0.057	−0.152	0.414	0.088	−0.202	−0.006	7.44%
Outdegree	0.282	0.104	−0.136	−0.126	−0.290	−0.037	2.42%
Outdegree weighted	0.132	0.349	0.113	−0.010	0.234	−0.008	4.17%
Closeness	0.142	−0.028	−0.076	0.475	0.180	0.007	25.49%
Closeness weighted	0.123	−0.126	−0.254	0.382	0.001	0.029	17.76%
Betweennes	0.114	−0.067	0.034	−0.075	0.019	0.966	2.00%
Eccentricity	0.167	0.005	−0.039	0.337	0.029	−0.085	47.79%
Eccentricity weighted	0.100	−0.079	−0.196	0.491	0.008	0.025	17.63%
Eigenvector	0.314	−0.049	−0.048	−0.088	−0.063	−0.012	5.85%
Eigenvector weighted	0.151	0.245	0.299	0.067	0.100	0.019	6.08%
Authority	0.223	−0.262	0.021	−0.157	0.268	−0.098	2.40%
Authority weighted	0.049	−0.119	0.333	0.070	−0.233	−0.088	37.21%
Hub	0.276	0.095	−0.133	−0.120	−0.261	−0.097	7.78%
Hub weighted	0.097	0.270	0.091	−0.015	0.268	−0.101	38.82%
PageRank	0.225	−0.270	0.036	−0.128	0.261	−0.057	1.60%
PageRank weighted	0.068	−0.165	0.389	0.141	−0.191	−0.009	9.81%
CheiRank	0.280	0.118	−0.135	−0.084	−0.300	−0.011	3.27%
CheiRank weighted	0.144	0.333	0.078	0.033	0.122	0.041	13.37%
Random-walk betweenness	0.283	−0.025	−0.043	0.217	0.005	0.062	14.90%
SinkRank	0.226	−0.269	0.035	−0.130	0.259	−0.053	1.57%
SinkRank weighted	0.069	−0.163	0.390	0.139	−0.196	−0.004	9.90%
SourceRank	0.281	0.116	−0.132	−0.086	−0.301	−0.006	3.21%
SourceRank weighted	0.145	0.331	0.081	0.031	0.124	0.045	13.61%
% of variance explained	**33.87%**	**22.98%**	**15.44%**	**8.53%**	**3.92%**	**3.48%**	

Notes: authors' calculations.

Table 5
Redundant centrality measures.

	CUD	CCDC	DCV	Deceval
Degree weighted				−0.341
Indegree weighted			−0.196	−1.461
Outdegree weighted				−0.554
Closeness	−0.614			−0.435
Closeness weighted	−3.041	−3.120	−1.883	−0.668
Betweenness	−0.455	−0.232	−0.437	−0.772
Eccentricity	−1.558	−1.093	−0.631	−0.126
Eccentricity weighted	−2.760	−3.379	−1.948	−0.935
Eigenvector weighted				−0.327
Authority weighted	−0.518	−0.142	−2.929	−1.557
Hub weighted	−0.728	−0.210	−0.664	−0.979
PageRank weighted				−1.325
CheiRank weighted			−0.618	−0.407
Random-walk betweenness			−0.293	
SinkRank weighted				−1.320
SourceRank weighted			−0.705	−0.399

Notes: This table presents the centrality measures considered redundant and their respective rescaled scores that represent their contribution to the index. The lower the value, the lower the contribution of the centrality measure to the general index.

robustness check, we separate the centrality measures according to the direction the nodes point on the networks. One subset includes measures based on links pointing towards the network (i.e., incoming criterion), and the other contains measures with links pointing out of the network (i.e., outgoing criterion). The measures that do not depend on the direction the nodes point (e.g., closeness, betweenness, eccentricity, eigenvector centrality, and random-walk betweenness) are included in both subgroups to ensure that our results consider all measures.[18] Table 6 presents the redundant measures identified with RPCA for these subsamples in columns (2) and (3). The redundant measures identified in the previous section are shown in column (1) for comparison purposes. The second robustness check uses two subsamples of system participants: banks and nonbanks, and their redundant measures are reported in columns (4) and (5).

The third robustness check is based on RDA, previously explained in Section 3.2. This parametric approach was implemented in panel data, setting each centrality measure (y) as a function of daily indicators (X) considered the main drivers of transactional activity in each FMI and institution's fixed effects. A description of these explanatory variables, along with the results obtained from the first step of this method, are presented in Appendix B.

The measures identified as redundant using RDA are shown in column (6). Lastly, we study the average-linkage clustering method, previously explained in Section 3.3. The redundant measures are presented in column (7), and the corresponding dendrograms (i.e., cluster trees) are shown in Appendix C.

Most redundant measures remain unchanged, even considering RPCA on reduced data sets, a model-based approach (RDA), and a clustering method. As expected, these measures exhibit some changes when subsamples of data are examined (see columns 2 and 3) because there is a loss of information caused by not evaluating the entire set of measures. This does not occur when we split the sample of system participants into banks and nonbanks or when we use RDA. Indeed, the results obtained from this model-based method are qualitatively almost the same as those obtained with RPCA.

The same occurs from applying the selected clustering method. In all FMIs, the last variables to join the primary (or main) groups in dendrograms are the same variables identified as redundant with the other statistical methods. Indeed, these measures exhibit the highest average Euclidean distance towards the other centrality measures, as seen in Table C.1 in Appendix C. Hence, we can be confident that the centrality measures identified as redundant with RPCA are roughly the same under other criteria. These variables barely contribute to determining centrality so that they can be consistently discarded from the general indices. The number of variables retained for constructing the general indices of centrality is nineteen for CUD, twenty for CCDC, sixteen for DCV, and eleven for Deceval.

5.3. Centrality indices for financial market infrastructures

The composite indices of centrality are synthetic measures that encompass different definitions of centrality. As such, they are more effective than individual criteria in identifying participants who can affect the stability of the network by either entering or leaving the ranking of institutions with the highest centrality.

[18] The subgroup representing the incoming criterion includes sixteen measures: Indegree (weighted and unweighted), Closeness (weighted and unweighted), Betweenness, eccentricity (weighted and unweighted), Eigenvector centrality (weighted and unweighted), Authority (weighted and unweighted), PageRank (weighted and unweighted), Random-walk betweenness, and SinkRank (weighted and unweighted). The subgroup for the outgoing criterion also includes sixteen measures: Outdegree (weighted and unweighted), Closeness (weighted and unweighted), Betweenness, Eccentricity (weighted and unweighted), Eigenvector centrality (weighted and unweighted), Hub (weighted and unweighted), CheiRank (weighted and unweighted), Random-walk betweenness, and SourceRank (weighted and unweighted).

Table 6
Robustness checks.

		(1) RPCA All measures	(2) RPCA Incoming	(3) RPCA Outgoing	(4) RPCA Banks	(5) RPCA Non-banks	(6) RDA All measures	(7) Clustering All measures
CUD	Degree weighted						−0.057	X
	Indegree weighted						−0.479	X
	Outdegree weighted						−0.162	X
	Closeness	**−0.614**	−0.283	−0.094	−0.558	−0.200		
	Closeness weighted	**−3.041**	−2.421	−2.290	−2.679	−3.127	−2.887	X
	Betweenness	**−0.455**	−0.171	−0.063	−0.224		−0.012	X
	Eccentricity	**−1.558**	−1.142	−1.094	−2.127	−1.529	−0.017	
	Eccentricity weighted	**−2.760**	−2.033	−2.065	−2.896	−3.055	−2.266	X
	Authority weighted	**−0.518**	−0.307		−0.406	−0.481	−1.661	X
	Hub weighted	**−0.728**		−0.668	−0.654	−0.420	−1.757	X
	PageRank weighted	**−0.518**	−0.307		−0.406	−0.481	−0.033	
CCDC	Closeness weighted	**−3.120**	−2.119	−2.191	−2.978	−3.022	−3.525	X
	Betweenness	**−0.232**	−0.178	−0.110	−0.413	−0.260		X
	Eccentricity	**−1.558**	−0.893	−0.745	−0.671	−1.408		X
	Eccentricity weighted	**−2.760**	−2.682	−2.716	−3.625	−3.371	−3.368	X
	Authority weighted	**−0.142**			−0.184	−0.004		X
	Hub weighted	**−0.728**			−0.045	−0.001	−0.009	X
DCV	Indegree weighted	**−0.196**			−0.281	−0.137	−0.952	
	Closeness							X
	Closeness weighted	**−1.883**	−1.380	−1.605	−2.029	−1.799	−1.703	X
	Betweenness	**−0.437**	−0.469	−0.515	−0.095	−0.579	−0.036	X
	Eccentricity	**−0.631**	−0.415	−0.621	−2.062	−0.336	−0.283	X
	Eccentricity weighted	**−1.948**	−1.337	−1.656	−2.964	−1.736	−2.125	X
	Eigenvalue weighted			−0.515				
	Authority weighted	**−2.929**	−2.484		−1.697	−3.140	−3.394	X
	Hub weighted	**−0.664**		−0.785	−0.559	−0.569	−0.848	
	CheiRank weighted	**−0.618**		−0.352		−0.634		
	Random-walk betw.	**−0.293**	−0.379	−0.332		−0.578		
	SourceRank weighted	**−0.705**		−0.515		−0.717		
Deceval	Degree weighted	**−0.341**			−0.861	−0.850		X
	Indegree weighted	**−1.461**			−1.171	−1.170	−0.075	X
	Outdegree weighted	**−0.554**			−1.121	−1.116	−0.202	X
	Closeness	**−0.435**	−0.907	−0.945	−0.341	−0.354	−1.695	X
	Closeness weighted	**−0.668**	−0.935	−1.316	−0.318	−0.347	−3.140	X
	Betweenness	**−0.772**	−1.025	−2.191	−0.215	−0.179		X
	Eccentricity	**−0.126**	−0.862	−0.544	−0.039	−0.058		X
	Eccentricity weighted	**−0.935**	−0.894	−1.174	−0.714	−0.747	−2.726	X
	Eigenvector		−0.965	−0.741				
	Eigenvector weighted	**−0.327**	−1.408		−0.889	−0.880		X
	Authority weighted	**−1.557**			−1.273	−1.277	−0.165	X
	Hub weighted	**−0.979**		−0.057	−1.363	−1.364	−0.493	X
	PageRank weighted	**−1.325**			−1.015	−1.014		
	CheiRank weighted	**−0.407**			−0.942	−0.931		
	Random-walk betw.		−0.814					X
	SinkRank weighted	**−1.320**			−1.015	−1.014		
	SourceRank weighted	**−0.399**			−0.932	−0.920		

Notes: This table presents the centrality measures considered redundant and their respective rescaled scores that represent their contribution to the index. The lower the value, the lower the contribution to the general index. For the clustering method, the variables identified with an X are considered redundant as they exhibit the highest average Euclidean distance to the other centrality measures.

The Top-10 most central participants obtained from centrality indices calculated for the entire sample of measures and financial institutions are presented in Table 7. These rankings present in descending order the score (i.e., the average score for the study period, weighted by the scores obtained from RPCA) obtained by each financial institution. Due to statistical reserve, the names of these institutions are undisclosed. However, we identify them using B for banks, BF for brokerage firms, MF for mutual funds, and FC for financial corporations, along with a code representing each financial institution.

The results for the CUD, presented in columns (1) and (2), identify eight banks (B), one brokerage firm (BF), and one mutual fund (MF) as the most central financial institutions. In the first five positions appear the participants that most contribute with payments in value and the number of transactions. A salient feature of this ranking is the considerable distance in the scores obtained by the banks in the first and second positions, which suggests that bank B9 is, to a considerable extent, the most central participant in the system.

In CCDC, the participants with the highest centrality scores are five banks, three brokerage firms, and two financial corporations (columns (3) and (4)). The banks in the first three positions are the most active peso/dollar market participants. However, the

Table 7
Top-10 most central participants.

(1) System participant	(2) Score in the index	(3) System participant	(4) Score in the index	(5) System participant	(6) Score in the index	(7) System participant	(8) Score in the index
B9	11.69	B15	1.98	BF30	7.87	BF30	6.09
B53	8.09	B9	1.80	B9	6.39	BF64	5.38
B15	6.96	B25	1.72	B15	6.00	BF16	4.61
B3	6.39	BF05	1.64	B21	5.10	BF53	4.55
MF26	5.30	BF30	1.60	B3	4.82	BF28	4.34
B8	5.30	FC13	1.58	B11	4.71	BF24	3.62
B14	5.09	B3	1.51	B53	4.51	BF78	3.16
BF30	4.82	B11	1.41	FC13	4.20	BF05	2.79
B11	4.68	BF78	1.37	B42	3.73	BF99	2.55
B21	4.48	FC43	1.36	FC43	3.47	BF5006	2.31

Notes: authors' calculations. The letter B is used to identify banks, BF for brokerage firms, MF for mutual funds, and FC for financial corporations.

Table 8
Top 10 most central participants before and during the Lockdown period.

CUD			CCDC			DCV			Deceval		
(1) Entire period	(2) Before the lockdown period	(3) During the lockdown period	(1) Entire period	(2) Before the lockdown period	(3) During the lockdown period	(1) Entire period	(2) Before the lockdown period	(3) During the lockdown period	(1) Entire period	(2) Before the lockdown period	(3) During the lockdown period
B9	B9	B9	B15	B15	B15	BF30	BF30	B15	BF30	BF64	BF30
B53	B53	B53	B9	B9	B9	B9	B9	BF30	BF64	BF30	BF16
B15	B15	B3	B25	B25	B25	B15	B15	B9	BF16	BF53	BF28
B3	B3	B15	BF05	BF30	BF05	B21	B3	B21	BF53	BF16	BF64
MF26	B10	B14	BF30	FC11	B12	B3	B21	B42	BF28	BF28	BF5026
B8	MF26	B21	FC13	BF05	B3	B11	B11	B3	BF24	BF24	BF53
B14	B14	MF26	B3	B3	BF30	B53	B53	B11	BF78	BF78	BF24
BF30	BF30	B11	B11	B11	BF53	FC13	FC13	B53	BF05	BF05	BF99
B11	B11	BF30	BF78	BF16	FC13	B42	B8	FC13	BF99	BF5006	BF05
B21	B25	B25	FC43	BF78	FC43	FC43	B42	FC43	BF5006	BF99	B11

Notes: authors' calculations. The letter B is used to identify banks, BF for brokerage firms, MF for mutual funds, and FC for financial corporations.

differences in the scores they achieved are minor, indicating that the second and third participants are not far from the bank in the first position. Similarly, slight differences are observed in the remaining places, which suggests that monitoring activities on this FMI should emphasize all the financial institutions in this ranking.

For the central securities depository for sovereign debt securities (DCV), the most central participants are seven banks, one brokerage firm, and two financial corporations. The results, reported in columns (5) and (6), reveal that the participants in the Top-3 positions are remarkably close to one another and therefore, should be closely monitored. Columns (7) and (8) report the results for Deceval, where brokerage firms occupy the top ten positions. As in DCV, the most central participant is the brokerage firm BF30 and the score it obtains is very far from the participant in the second position. Hence, it can be said that for the study period, this brokerage firm was very active in carrying out transactions, either using collateral sovereign debt, equities, bonds, or term deposit certificates.

When comparing the results between networks, it can be observed that the most central participants in CCDC, DCV, and Deceval also appear in the ranking of CUD. This correspondence is a finding we already expected, given that the cash leg of the transactions in these FMIs is settled in the large-value payment system.

5.4. Centrality indices before and after the COVID-19

The emergence of COVID-19 in the first quarter of 2020 led the government to introduce lockdown measures to contain the spread of the virus. We test whether this period influenced economic activities such as payments and other peso-denominated transactions by examining changes in the ranking of the most central FMIs participants. To this end, centrality indices are calculated with RPCA for subsamples of data representing the periods before and during this episode. The dates on which the Colombian government introduced the lockdown measures (March 25 to September 1) are used to construct the subperiods mentioned above. Table 8 presents the Top-10 most central financial institutions for each FMI for these subperiods, accompanied by the results obtained for the entire sample period (for comparison purposes).

As seen in Table 8, some participant shifts are considered systemically important, especially from the third position. That is the case of CUD, CCDC, and DCV, where, in general terms, the financial institutions in the ranking remain the same, but minor changes are observed in the positions they occupy. In contrast, Deceval exhibit changes not only in the participants in the first positions but also in those that represent this group. This result can be identified by the entrance of new financial institutions to the ranking of the most central.

Overall, any changes in financial institutions that are deemed systemically important should be seen as signals that require a closer examination of the factors driving them. In this paper, we study this topic through general centrality indices constructed with RPCA, which can be easily updated and can, therefore, facilitate monitoring activities on these FMIs. These indices can become a valuable tool for examining the stability of the ranking of the most central institutions.

6. Conclusions

The concept of centrality is commonly used to identify the participants that play the most relevant role in systems with a network structure, which is critical for monitoring purposes. Monitoring all financial institutions under every criterion can be a challenging task for network analysts due to the wide range of centrality measures. This becomes even more daunting when there are a significant number of system participants involved. We use a robust version of principal components analysis to reduce the information extracted from twenty-six centrality measures applied to transactions data of the large-value payment system (CUD), the foreign exchange clearing house (CCDC), and two securities depositories (DCV and Deceval). As a result, we obtain general centrality indices for these FMIs, represented primarily through the unweighted versions of these measures. These indices can facilitate the monitoring activities on these FMIs, as they reduce the complexity of the dataset by eliminating the redundant information already represented by linear combinations of the retained measures.

A common understanding on this topic suggests that system participants (nodes) identified as the most central in a network should be closely monitored since they are the ones that could produce considerable impacts on the stability of the network if they fail to comply with the activities that connect them with other participants. Considering that premise, an interesting extension of this work could include implementing simulation exercises to assess whether those identified as the most central would substantially impact the network (FMI) stability if they fail to meet their commitments (payments, trade, or similar transactions). One way to accomplish this is by considering simulated shocks to the most central participants through a direct-linkage contagion simulation model, like DebtRank.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.latcb.2023.100098.

References

Baek, S., Soramäki, K., Yoon, J., 2014. Network indicators for monitoring intraday liquidity in BOK-Wire+. Journal of Financial Market Infrastructures 2 (3), 37–66.
Batool, K., Niazi, M., 2014. Towards a Methodology for Validation of Centrality Measures in Complex Networks. PLoS ONE 9 (5), e98379 1-14.
Barrat, A., Barthélemy, M., Pastor-Satorras, R., and Vespignani, A. (2004). “The architecture of complex weighted networks”. Proceedings of the National Academy of Sciences of the United States of America, 3747–3752.
Battiston, S., Puliga, M., Kaushik, R., Tasca, P., Caldarelli, G., 2012. DebtRank: too central to fail?”, Financial Networks, the FED and systemic risk. Sci Rep 2.
Battiston, S., Caldarelli, G., D’Ericco, M., Gurciullo, S, 2016. Leveraging the network: a stress-test framework based on DebtRank. Statistics and Risk Modelling 33 (3–4), 117–138.
Bavelas, A., 1950. Communication patterns in task-oriented groups. Journal of the Acoustical Society of America 22, 725–730.
BIS-PFMI. (2012). Principles for financial market infrastructures.
BIS-CPSS. (2003). A glossary of terms used in payments and settlement systems. March.
Bonacich, P., Lloyd, P, 2001. Eigenvector-like measures of centrality for asymmetric relations. Soc Networks 23 (3), 191–201.
Brin, S., Page, L., 1998. The anatomy of large-scale hyper-textual Web search engine. Computer Networks and ISDN 30, 107–117.
Chepelianskii, A. (2010). Towards physical laws of software architecture, arXiv:1003.545
Croux, C., Haesbroeck, G., 2000. Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies. Biometrika 87 (3), 603–618.
Dungey, M., Luciani, M., and Veredas, D. (2014). “Googling SIFIs”, Centre for International Finance and Regulation, Working Paper 003/2013.
Goldbeck, J., 2013. Analyzing the Social Web. Elsevier.
Hage, P., Harary, F., 1995. Eccentricity and Centrality in Networks. Soc Networks 17 (1), 57–63.
Hotelling, H., 1933. Analysis of Complex Statistical Variables into Principal Components. Journal of Education Phycology 24, 417–441.
Isräels, A., 1992. Redundancy Analysis for various types of variables. Statistica Applicata 4 (4), 531–542.
Jolliffe, I.T., 1972. Discarding variables in a principal components analysis I: artificial Data. Appl Stat 22 (2), 160–173.
Jolliffe, I.T., 1973. Discarding variables in a principal components analysis II: real data. Appl Stat 22 (1), 21–31.
Jolliffe, I.T., 2002. Principal components analysis. Springer Series in Statistics. Springer, New York NY.
Kaiser, H.F., 1974. An index of factorial simplicity. Phsychometrika 39, 31–36.
Kelley, T., 1940. Talents and tasks: Their Conjunction in a Democracy For Wholesome Living and National defense. Harvard Education Press no.1. Graduate School of Education, Harvard University.
Kleinberg, J., 1999. Authoritative sources in a hyperlinked environment. Journal of the ACM 46 (5), 604–632.
León, C., Mariño, R., Cadena, C., 2021. Do central counterparties reduce counterparty and liquidity risk? Empirical evidence. Algorithmic Finance 9, 25–34.
Newman, M., 2005. A measure of betweenness centrality based on random walks. Soc Networks 27 (1), 39–54.
Newman, M., 2008. The mathematics of networks. In: Blume, L (Ed.), The New Palgrave Encyclopedia of Economics, 2nd edition editor. SD, editors. Palgrave Macmillan, Basingstoke.
Pearson, K., 1901. On Lines and Planes of Closest Fit to Systems of Points in Space. Philosophical Magazine 2 (11), 559–572.

Pozzi, F., Fersini, E., Messin, E., Liu, B., 2017. Sentiment Analysis in Social Networks. Elsevier.
Rao, C.R., 1964. The Use and Interpretation of Principal Components Analysis in Applied Research. Sankhyā: The Indian Journal of Statistics 26 (4), 329–358.
Rousseeuw, P., 1985. Multivariate estimation with high breakdown point. In: Grossmann, W., Pflug, G., Vincze, I., Wertz, W. (Eds.), Mathematical Statistics and Applications. Reidel Publishing Company, Dordrecht.
Rovira, P., Spelta, A., 2019. Identifying systemically important financial institutions. Computational Management Science 16, 155–185.
Soramäki, K., Cook, S., 2013. SinkRank: an Algorithm for Identifying Systemically Important Banks in Payment Systems. Open-Assessment E-Journal 7, 2013–2028.
Temizsoy, A., Iori, G., Montes-Rojas, G., 2017. Network centrality and funding rates in the e-MID interbank market. Journal of Financial Stability 33, 346–365.
van den Wollenberg, A., 1977. Redundancy Analysis- An alternative for canonical analysis. Psychometrika 42 (2), 207–219.
Verardi, V., Dehon, C., 2010. Multivariate outlier detection in Stata. Stata J 10 (2).
White, S., and Smith, P. (2003). Algorithms for estimating relative importance in networks. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.