Statistical Validation of Contagion Centrality in Financial Networks

Agathe Sadeghi* Zachary Feinstein†

Abstract

In this paper, we introduce a novel centrality measure to evaluate shock propagation on financial networks capturing a notion of contagion and systemic risk contributions. In comparison to many popular centrality metrics (e.g., eigenvector centrality) which provide only a relative centrality between nodes, our proposed measure is in an absolute scale permitting comparisons of contagion risk over time. In addition, we provide a statistical validation method when the network is estimated from data, as is done in practice. This statistical test allows us to reliably assess the computed centrality values. We validate our methodology on simulated data and conduct empirical case studies using financial data. We find that our proposed centrality measure increases significantly during times of financial distress and is able to provide insights in to the (market implied) risk-levels of different firms and sectors.

Keywords: Finance, Contagion Risk, Financial Stability, Network Analysis, Statistical Validation

*Stevens Institute of Technology, School of Business, Hoboken, NJ 07030. *asadeghi@stevens.edu* Corresponding author.
†Stevens Institute of Technology, School of Business, Hoboken, NJ 07030.

1 Introduction

The growing complexity of financial systems requires the creation of a framework that can assess the interconnections among firms, see e.g. (Feinstein and Hałaj, 2023, Figure 1). These connections play a crucial role in capturing systemic events and shocks that can lead to system instability. Researchers have shown growing interest in modeling financial system using network representations to gain a deeper understanding of these effects (Huang et al., 2013; Jackson and Pernoud, 2020). Emerging as a significant driver of instability, contagion, often described as the propagation of information or shocks across nodes, results in cascading effects that spread throughout the entire network, akin to an epidemic (Jalili and Perc, 2017). Within this paper, we construct a novel network centrality measure designed to capture financial contagion. Due to the proposed construction, we are able to, e.g., statistically test our centrality measure; as far as the authors are aware, this is the first network centrality that can be tested.

The structure of this paper is as follows: In this section, we provide an introduction to the research topic with relevant literature. Section 2 outlines the approach used to construct the network and introduces the Leontief inverse centrality measure, referred to as the "Leontief Centrality Measure (LCM)" hereafter. With this definition, we detail the statistical framework for the LCM in Section 3 and validate this framework using simulated data. We apply the LCM approach to financial data scenarios in Section 4 to assess framework's applicability to capture stress events. Finally, in Section 5, we summarize the key findings of the paper and propose potential avenues for future research.

1.1 Literature Review

Evaluating a network involves two key aspects: the topology of the network and the determination of the importance of each node within it, referred to as the centrality measure. Markose (2012) emphasizes that the network's topology plays a pivotal role in assessing its vulnerability.

Due to the recent financial events, it is important to view the financial system as a complex network, where firms serve as nodes and financial dependencies as links (Battiston et al., 2012). There are multiple perspectives on financial networks, e.g., (i) informational (Fink et al., 2016), (ii) interbank or default contagion (Acemoglu et al., 2016), (iii) portfolio overlap or price-mediated contagion (Cont and Schaanning, 2017), and (iv) cross-ownership (Fichtner et al., 2017). The construction of financial networks involves various methods, including correlation (Pacreau et al., 2021), partial correlation (Millington and Niranjan, 2020), and analysis of financial dependencies based on balance sheets and contracts between institutions (Eisenberg and Noe, 2001).

Failures or contagion can be triggered by common factors, such as market declines, or by idiosyncratic shocks affecting individual institutions, like spread widening (Benoit et al., 2017). The consequences of a shock can be broken down into two parts: the immediate failure of an institution triggered by a sufficiently large shock, and the spread of the shock throughout the system due to network ef-

fects. While many studies focus on the direct collapse of a firm (fundamental default) resulting from an external shock, they often overlook the aggregated impact of network-induced contagion (default by contagion) as noted by Cont et al. (2013). It is important to distinguish between the first-order effect, which could be the default of a firm or the occurrence of a shock impacting the firm, and the higher-order impacts due to these stresses. As highlighted by di Iasio et al. (2013); Veraart (2020), distress contagion can propagate through the network, causing widespread losses, even when there are no defaults within financial system. To measure these effects, it is essential to define a suitable centrality measure.

Centrality measures serve as valuable tools to quantitatively assess the structural significance of nodes within a network. As stated by D'Arcangelis and Rotundo (2016), a higher centrality score indicates that a particular property is more fitting for a given node. Different centrality measures offer distinct insights into various centrality dimensions (Hevey, 2018). In addition, the centrality notion becomes more crucial when the networks are not symmetric (Capponi and Larsson, 2015) which is the case in financial systems. When examining propagation within networks, Pacreau et al. (2021) highlights that the centrality of a node indicates its sensitivity to market fluctuations and reveals its connectivity within the network. Nodes with high centrality are often connected to many other nodes or are neighbors to densely interconnected subnetworks. Along this line, Battiston et al. (2012) introduces DebtRank, utilizing feedback-centrality to measure the impact of a shock on a single node. Building on this concept, Puliga et al. (2014) presents Group Debtrank to quantify the centrality of nodes when small shocks affect firms and concludes that during crisis periods, the centrality measure tends to increase.

It is crucial to select an appropriate centrality measure that aligns with the attribute under consideration (Singh et al., 2020). In this work, we introduce a simple yet innovative centrality measure that reflects the spread of a shock over the network structure based on the Leontief inverse (Leontief, 1936, 1986). Widely used in input-output analysis in economics (Antràs et al., 2012; Antràs and Chor, 2013, 2018), the Leontief inverse has also found applications in network science more recently (Moran and Bouchaud, 2019). Certain studies in the financial network field have employed the Leontief inverse to establish the network structure directly (Elliott et al., 2014; D'Arcangelis and Rotundo, 2016), rather than applying it to compute the centrality.

Prior studies have typically produced a single value as the centrality measure. To ensure the validity of research findings, it becomes essential to gain a comprehensive understanding of the sensitivity of centrality measures (Borgatti et al., 2006). This need for accuracy analysis is further highlighted in the context of large and complex networks, where potential issues like missing data or the presence of hidden variables (confounders) may arise (Breiger et al., 2003). The accuracy of centrality measures in network theory is influenced primarily by the deviation between the estimated network and the true underlying network. Existing literature on error analysis is limited by its focus on common centrality measures and tends to be case-specific, considering scenarios with predefined error types, levels, densities, and node numbers (Smith and Moody, 2013; Lee and Pfeffer, 2015; Niu et al., 2015). Moreover, many

of these measurement errors are qualitatively assessed (Martin and Niemeyer, 2019) and are primarily restricted to undirected and unweighted networks. Additionally, a common assumption made is that the true network is known, allowing for the introduction of various error types. However, this assumption is often unrealistic, as the true network is typically unknown in practice.

1.2 Primary Contributions

In this paper, we construct weighted directed networks using a vector autoregressive (VAR) model, where nodes represent firms and the links depict interconnections between them. Rather than utilizing the Leontief inverse notion solely as a generative method for networks, we redefine it as the Leontief centrality measure, broadening its applicability across different network structures. This measure, referred to as LCM, quantifies node centrality during shock scenarios, illustrating a node's position when hit by a shock. It considers the propagation and contagion risks inherent in the network's topology, and boils down the shock propagation dynamics to a closed form.

We define LCM at pair-, node-, and system-levels, offering different granularities for analysis. Pair-level LCM between two firms demonstrates the first firm's position when the second firm experiences a shock. Node-level LCM of a firm is the total resulting shock when it is stressed. System-level LCM aggregates multiple scenarios to portray overall centrality changes under equally probable situations. We also allow the association of weights with node- and system-level centralities.

Our approach involves statistical analysis to assess the robustness of the centrality measure. We derive the asymptotic distribution of LCM through its Taylor expansion and form hypotheses for testing – specifically focusing on whether LCM significantly differs from zero. We validate our framework using simulated data, demonstrating the theoretical distribution's compatibility with empirical observations. In this simulated setup, the true value of LCM consistently falls within the confidence intervals of the theoretical distribution at a 95% confidence level.

Applicability in financial contexts is a key aspect of our paper. We evaluate the performance of our framework using two financial datasets – Credit Default Swap Index (CDX) and equity tick data. In the CDX case study, LCM effectively captures underlying market dynamics, outperforming both degree centrality and the leading eigenvalue. In the case of tick data, LCM provides insights into the market and institutions' states but does not suffice for standalone investment decisions. However, a blend of node- and system-level LCM dynamics along with institutional background information can guide investment strategies effectively.

2 Network Centrality

A network consists of a set of nodes and links. The connectivity pattern between nodes is captured by the adjacency matrix $A \in \mathbb{R}^{M\times M}$, where M denotes the number of nodes, and its elements a_{ij}, for

$i, j = 1, ..., M$, indicate the presence or absence of connections between nodes. For our purpose, the elements a_{ij} represent the connectivity weights between nodes i and j (Papana et al., 2017).

In this paper, we consider the nodes to be the firms and the links to be the connections based on a linear model. Our network construction follows a VAR(l) structure considering only lag l. The structure is as follows:

$$X_t = X_{t-l}A + \mathcal{E}_t\,, \tag{1}$$

where $X_t \in \mathbb{R}^M$ is a vector (of, e.g., market returns of each firm) at time t, while $\mathcal{E}_t \in \mathbb{R}^M$ denotes the independent and identically distributed (iid) noise. As a weighted directed matrix, A serves as the adjacency matrix of the network, showcasing the interconnectedness of different variables. We refer to this system as the linear model.

Assumption 1 (Stationarity)**.** *For the remainder of this work, we assume that the spectral norm of the adjacency matrix is strictly less than one, i.e., the VAR(l) model is stationary.*

Remark 1. *If one wants to explore the idiosyncratic relations among time series, dimensionality reduction to eliminate the common factors can be performed. The dimension reduction can be achieved using various methods such as Singular Value Decomposition (SVD) or spectral reduction (Ricciardi et al., 2022).*

We introduce a novel centrality measure that overcomes the limitations of relative measures such as the eigenvector centrality (which can be scaled arbitrarily; see, e.g., Markose et al. (2012); Chan-Lau (2018); Olmo (2021)) to allow for direct comparison between different time windows. This measure defined as in Equation 2 below, is readily interpretable in the context of financial contagion.

As an example, in Figure 1 we present a schematic network diagram along with its shock propagation dynamics. Consider four time series, corresponding to four nodes denoted as W, X, Y and Z, which are connected as illustrated in the figure on the left. The shock propagation pattern is shown in the figure on the right side, under the scenario in which node W experiences shock ϵ at time $t = 0$. Due to the connections with nodes X and Y, this shock subsequently propagates and affects these nodes at time $t = 1$. Following the network, the shock then cycles back to node W and hits Z at $t = 2$. This cascading effect goes on until the shock dissipates.

In order to formulate these dynamics more generally, we assume M time series and define an initial shock vector $\boldsymbol{\epsilon} \in \mathbb{R}^M$, where element i represents the magnitude of the shock to node i. The shock propagation dynamics can be shown as:

$$S_t = S_{t-1}A, \quad S_0 = \boldsymbol{\epsilon} \quad \Rightarrow \quad S_t = \boldsymbol{\epsilon}A^t,$$

where S_t is the state of the system at time t and A is the adjacency matrix. By applying the adjacency matrix A to each state iteratively, we can trace the propagation of the shock over time. Specifically,

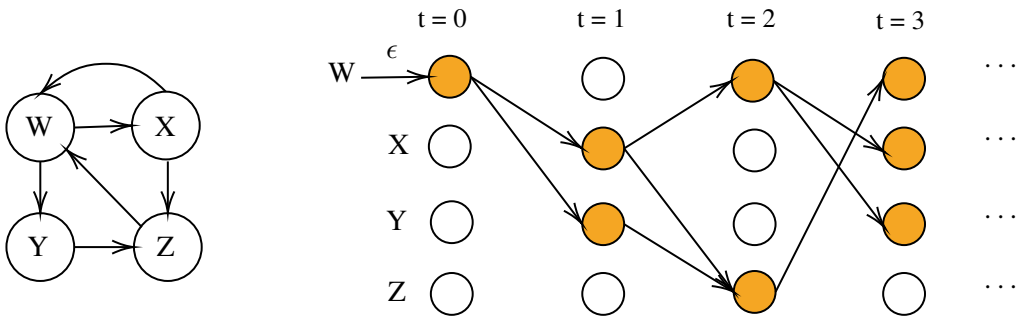


Figure 1: **Left** schematic network **Right** shock propagation through network over time

$S_1 = S_0 A$, $S_2 = S_1 A = S_0 A^2$, and so on. Each term in the series $\boldsymbol{\epsilon} A^t$ represents the effect of the shock propagating through the network over different time intervals, accounting for higher-order interactions and contagion effects. The summation

$$\sum_{t=1}^{\infty} S_t = \boldsymbol{\epsilon} A + \boldsymbol{\epsilon} A^2 + ... = \boldsymbol{\epsilon} \underbrace{(A + A^2 + ...)}_{\text{pair-level LCM}}$$

captures the cumulative effect of the shock propagation and represents the centrality of the system. Each element in this matrix represents the pair-level LCM, i.e., the centrality of node j (corresponding to the column index) when node i (corresponding to the row index) is hit by a shock. Note that this pair-level LCM is independent of the initial shock $\boldsymbol{\epsilon}$.

Proposition 2. *Under Assumption 1, the pair-level LCM is well-defined and can be simplified via the Leontief inverse:*

$$\sum_{t=1}^{\infty} A^t = (I - A)^{-1} - I \,. \tag{2}$$

where I is the $M \times M$ identity matrix.

Proof. $\sum_{t=0}^{\infty} A^t = I + \sum_{t=1}^{\infty} A^t = (I - A)^{-1}$ per Leontief (1936). □

Remark 2. *This methodology can handle networks with cycles and does not impose the constraint of acyclic graph structures, as seen in some prior research (see, e.g., Teteryatnikova (2014)).*

Remark 3. *LCM is not solely influenced by the strength of the links but also by the overall structure of the network. A densely connected network with comparatively weaker links may yield a higher centrality measure than a sparsely connected network with stronger links.*

To compute the node-level centrality when node i is shocked, we consider the weighted summation of each row of the centrality measure matrix:

$$C = [\,(I - A)^{-1} - I\,]\,\boldsymbol{w}\,, \tag{3}$$

where C is the vector of centrality and $\boldsymbol{w} = (w_1, ..., w_M)^\top$ are node weights. This weight vector can be customized based on the specific context of the research. For example, it can be defined as the ratio of liabilities to assets of the institutions or the market capitalizations of the firms. Incorporating weights allows us to assign varying degrees of importance to each node when assessing the system's centrality. As highlighted by Ali et al. (2016), conventional centrality measures may not fully capture the systemic importance of financial firms as the size and structure of their balance sheets play a crucial role in understanding the potential risks they pose to the entire system. Larger values in this centrality measure indicate a higher level of potential spillover, particularly during times of crisis.

Remark 4. *Though we constructed the above with a static network, we note that both the adjacency matrix and the weight vector for the node-level LCM can be time-varying. This will be used in the financial case studies of Section 4.*

Remark 5. *System-level centrality is defined as the total sum of the node-level centralities C. This value provides a useful proxy for the total systemic risk captured in financial data.*

3 Statistical Analysis

We go beyond traditional centrality calculations by deriving the asymptotic distribution of the LCM, enabling us to compute confidence intervals and perform statistical tests. As we calibrate the adjacency matrix as the OLS estimates, we recall this construction for a generic multivariate model. Given two matrices $X, Y \in \mathbb{R}^{T\times M}$, where T represents the number of time steps and M denotes the number of variables, and an iid noise term $\mathcal{E} \in \mathbb{R}^{T\times M}$, we can adopt a linear regression framework as below:

$$\begin{aligned} Y &= XA + \mathcal{E}\,, \\ \hat{A} &= (X^\top X)^{-1} X^\top Y\,, \end{aligned} \tag{4}$$

where A is the true coefficient matrix that governs the relationships between the elements of X and Y. On the other hand, $\hat{A}$ represents the estimated coefficient matrix, which is obtained from the regression analysis.

3.1 Distribution of the LCM

In practice, a discrepancy between the true adjacency matrix A and its estimated counterpart $\hat{A}$ potentially exists. We denote this difference as Δ, such that $\hat{A} = A + \Delta$. Taking advantage of the Taylor series expansion, we can express the LCM in terms of the discrepancy as follows:

$$(I - \hat{A})^{-1} = (I - A)^{-1} + (I - A)^{-1}\Delta(I - A)^{-1} + O\,(\frac{1}{T})\,. \tag{5}$$

Theorem 3 (LCM Distribution)**.** *Assume the data follows a linear model. Let ρ_{lk} denote the correlation coefficient between time series l, k. Let σ_k denote the standard deviation of time series k. Define $Q := \lim_{T\to\infty} T^{-1}X^\top X$ (with limit taken in probability as the number of data points grows to infinity). Set $\Sigma_{lk} := \rho_{lk}\sigma_l\sigma_k Q^{-1}$ for each l, k. Then for every i, j we have the following asymptotic distribution for the pair-level LCM:*

$$\begin{aligned}&\sqrt{T}\left[(I-\hat{A})^{-1}_{ij} - (I-A)^{-1}_{ij}\right]\\ &\quad\sim N\left(0, \left[(I-A)^{-1}_{\cdot j}\right]^\top \left[(I-A)^{-1}_{i\cdot}\Sigma_{lk}\left[(I-A)^{-1}_{i\cdot}\right]^\top\right]_{lk}(I-A)^{-1}_{\cdot j}\right).\end{aligned}$$

This result follows from the normality of the error terms $\Delta_{\cdot j} \sim N(0, \Sigma_{jj})$ with covariance Σ_{ij} between $\Delta_{\cdot i}, \Delta_{\cdot j}$. The detailed proof of Theorem 3 is given in the Appendix A.1.

Corollary 4. *Consider the same setting as in Theorem 3. The node-level LCM, defined in Equation 3, asymptotically follows the normal distribution:*

$$\begin{aligned}&\sqrt{T}\left[(I-\hat{A})^{-1}_{i\cdot}\boldsymbol{w} - (I-A)^{-1}_{i\cdot}\boldsymbol{w}\right]\\ &\quad\sim N\left(0, \left[(I-A)^{-1}\boldsymbol{w}\right]^\top \left[(I-A)^{-1}_{i\cdot}\Sigma_{lk}\left[(I-A)^{-1}_{i\cdot}\right]^\top\right]_{lk}(I-A)^{-1}\boldsymbol{w}\right)\end{aligned}$$

for every $i = 1, ..., M$ where the weight vector is denoted as $\boldsymbol{w} = (w_1, ..., w_M)^\top$.

The proof of Corollary 4 is given in Appendix A.2.

3.2 Hypothesis Testing

In this section, we aim to explore a potential application of the node-level LCM distribution as outlined in Corollary 4. Specifically, we propose a statistical test to determine whether the node i LCM significantly exceeds zero, i.e., the node contributes significantly to the propagation of a shock within the network.

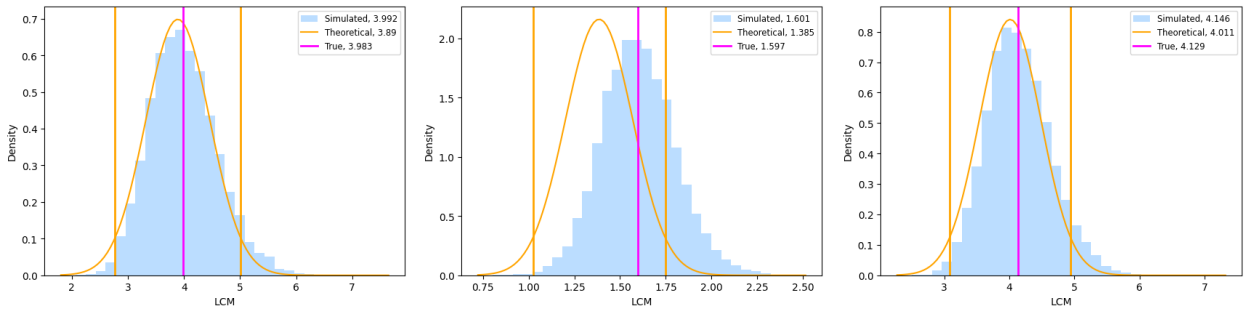
$$\begin{aligned}H_0 &: (I-A)^{-1}_{i\cdot}\boldsymbol{w} - w_i = 0\,,\\ H_1 &: (I-A)^{-1}_{i\cdot}\boldsymbol{w} - w_i > 0\,.\end{aligned} \tag{6}$$

We note that when working with networks, constructing them effectively requires a substantial number of data points. Therefore, the asymptotic distribution of the network is sufficient.

Corollary 5 (Test Statistic and Distribution)**.** *Consider the setting of Theorem 3 with the test statistic:*

$$Z := \frac{\sqrt{T}\left[(I-\hat{A})^{-1}_{i\cdot}\boldsymbol{w} - w_i\right]}{\left(\left[(I-\hat{A})^{-1}\boldsymbol{w}\right]^\top \left[(I-\hat{A})^{-1}_{i\cdot}\Sigma_{lk}\left[(I-\hat{A})^{-1}_{i\cdot}\right]^\top\right]_{lk}(I-\hat{A})^{-1}\boldsymbol{w}\right)^{\frac{1}{2}}}.$$

Under the null hypothesis, $Z \sim N(0, 1)$ asymptotically.



(a) node 1 (b) node 2 (c) node 3

Figure 2: Distribution of node-level LCM.

This result directly follows from Corollary 4 the Central Limit Theorem and Slutsky's Theorem. The formal proof of Corollary 5 is given in Appendix A.3.

3.3 Numerical Validation

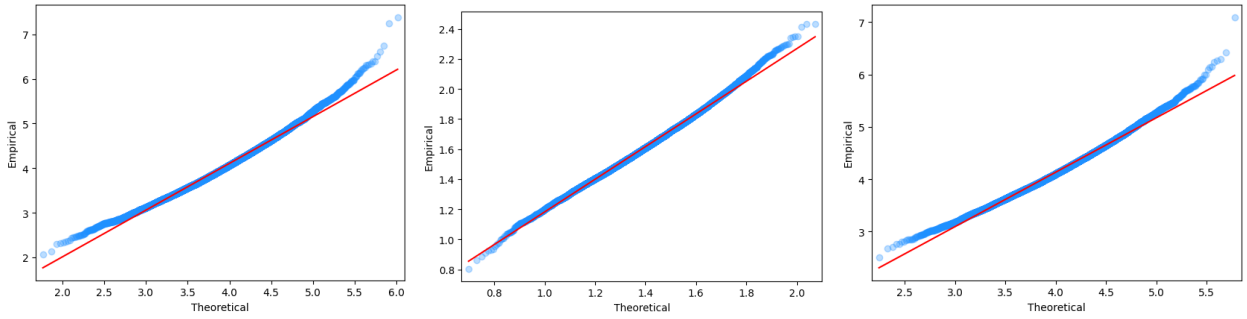
To evaluate the validity of the LCM, i.e., alignment between the theoretical and empirical distributions and their true counterparts, we conduct a case study using simulated data. We generate a dataset of 3 time series with iid noise terms and length of 600. The empirical analysis involves 10,000 simulations to obtain a robust representation. Subsequently, we calculate the mean and variance of the theoretical and empirical distributions of Leontief centrality measure and compare these distributions with the true values of the centrality.

Figure 2 presents the node-level LCM distributions. A close examination of the figures reveals a strong alignment between all three sets of values. The true centrality value falls within the 95% confidence interval calculated from the theoretical distribution. We note that the theoretical distribution may exhibit a slight bias due to the theoretical distribution's mean being estimated from one realization of the adjacency matrix (see Figure 2b).

The QQ plots in Figure 3 provide a visual comparison between the theoretical and empirical distributions. The points in the plots are close to the red regressed line, indicating similarity between the two distributions. However, slight deviations can be observed at the tails of the distributions. This is due to our use of the asymptotic theoretical distribution. Consequently, some bias is expected when finite data is considered, e.g., from the errors of order $1/T$ in the Taylor expansion of Equation 5.

4 Financial Case Studies

In this section we present two case studies in order to demonstrate the financial implications of our framework. In the first case study, we utilize CDX spread data to construct networks, from which we analyze the LCM for a spanning history of over 11 years. In the second one, we leverage high-frequency



(a) node 1 (b) node 2 (c) node 3

Figure 3: QQ plot of node-level LCM comparing the theoretical and empirical distribution.

equity tick data to capture the information in the stock market. We highlight the years with turmoils to demonstrate how the LCM reflects market behavior.

4.1 Credit Default Swap Index Market Contagion

In most of the existing literature, financial networks are typically constructed based on contracts between institutions, particularly in the banking sector, where a core of large banks are connected to peripheral ones. Numerous studies have explored this core-periphery network structure, mostly in the context of interbank lending (Battiston et al., 2015; Bardoscia et al., 2016; Bichuch and Feinstein, 2018; Blasques et al., 2018; Farboodi, 2021). However, only a limited number of studies have delved into derivative markets, specifically Credit Default Swaps (CDS) (Brunnermeier et al., 2013; Blasques et al., 2016).

Our dataset is obtained from Bloomberg, breaking down Markit's Investment Grade (IG) and High Yield (HY) CDX spread data[1] into sectors based on the Global Industry Classification Standard (GICS)[2]. This breakdown yields 10 sectors for each grade, totaling 20 time series. These indices have been reporting since January 1, 2011. Our analysis spans from this starting date to August 31, 2023. We use a one-year time window, sliding it monthly, resulting in 146 time windows. This approach allows us to assess the contagion risk across different periods, enabling a deeper understanding of the derivative market's network's dynamics over time.

To make the time series stationary and reduce variability, we compute the daily log return of the time series $r_{i,t} = log(\frac{s_{i,t+1}}{s_{i,t}})$ where $s_{i,t}$ is the CDX spread of the sector i at time t. Using the framework with 1-day lag, we build our network for each time window. Afterwards, we calculate the LCM for the system. For simplicity, we perform the analysis and statistical tests based on the unweighted centrality setting ($\boldsymbol{w} = \mathbf{1}$).

Visualizing the financial networks for February and March 2020 in Figure 4, we observe that the network in March 2020 exhibits thicker links compared to the network in February 2020, indicating

[1]For more information on these indices, we refer the interested reader to Appendix B.
[2]The breakdown is at level 1 and excludes real estate.

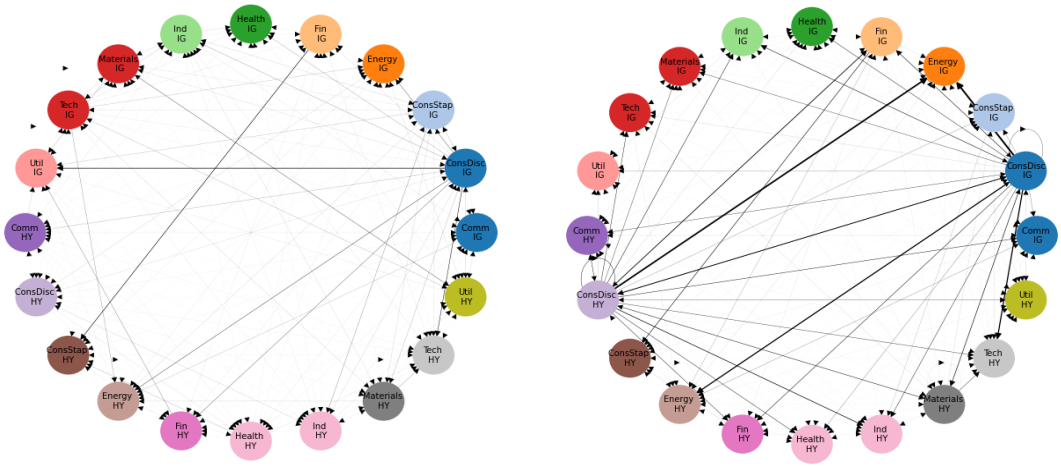


Figure 4: CDX network. **Left** February 2019-2020 **Right** March 2019-2020

stronger connectivity. Moreover, the March network appears to be more densely connected. Focusing on the March network, we observe a core-periphery like structure. The core nodes, characterized by more and thicker links either incoming or outgoing, feature the Consumer Discretionary and Energy indices. A more detailed analysis on what this entails is given below.

In Figure 5, we present the system-level Leontief centrality measure as a moving average across different time windows. The visualization indicates the overall level of contagion risk and sector instability, containing indices from both high yield and investment grade sectors. One key insight derived from the figure is the surge in LCM during and after the Covid-19 pandemic, indicating a high contagion risk during this period (Yu et al., 2021). Comparing the time window from March 2019 to March 2020 with the preceding years, the measure obtains a value of about five times greater, reaching its peak in late 2020. Remarkably, when March is added to the time window, a sharp increase in LCM is triggered instantly, while it gradually starts to drop close to the end of 2020. These dynamics are primarily influenced by several events related to the outbreak of Covid-19 in early 2020 leading to a widespread economic disruptions. The uncertainty triggered significant market volatility with concerns about the economic fallout, indicating a higher perceived risk. Additionally, there were vulnerabilities specific to sectors that were captured in the CDX data. Afterwards, the US government and Federal Reserve responded with monetary policy interventions and fiscal stimulus in order to support financial markets. Signs of recovery were shown also when vaccines were rolled out, as optimism grew regarding the potential economic rebound. Economic indicators improved and reduced the uncertainty about the long-term pandemic's impact. All these led to the gradual decline of risk and hence the LCM measure.

We also notice a ramp up upon including mid-2022 data. This is mainly due to the increasing interest rates by Federal Reserve and some bankruptcy, merger and near-default events. It is worth mentioning that the elevated LCM in late 2012 may be due to the lingering effects of the Great Recession and the

Table 1: Percentage change of centrality measures and leading eigenvalue during two periods of financial stress

Date	LCM	Degree Centrality	Leading Eigenvalue
2020-2021	4.78283	0.07169	0.23087
2022-2023	5.52444	0.00096	-0.00312

2011 European sovereign debt crisis. This analysis highlights the measure's robustness to economic events and interventions, showing its ability to capture the sector instability and contagion risk.

We plot the leading eigenvalue and thresholded degree centrality alongside LCM in Figure 5. We see that the leading eigenvalue exhibits some level of noise, displaying marginal increase during Covid-19 period. Towards the end of the date range, this value maintains a consistent trend without any notable elevation. Similarly, the degree centrality line appears rather stable throughout the figure. This suggests that both leading eigenvalue and degree centrality struggle to capture the contagion risk effectively. For more in-depth analysis, Table 1 shows the percentage changes in these three metrics during the two periods of financial stress.

Figure 6 shows LCM over time for the Consumer Discretionary sector. We display the unvalidated LCM along with the 97.5% confidence interval. When the confidence interval intersects the zero line, the validated LCM is considered zero. In time window March 2019 to 2020, a notable increase in both investment grade and high yield CDX within the sector becomes evident. This time period coincides with the beginning and the outbreak of Covid-19 pandemic in U.S., disrupting economic activities across multiple sectors. Moreover, the lockdowns, social distancing and job losses affected the consumer behavior, leading to decreased spending on non-essential expenses such as leisure, travel and luxury items.

On top of this, the global spread of Covid-19, which had already affected China, disrupted sup-

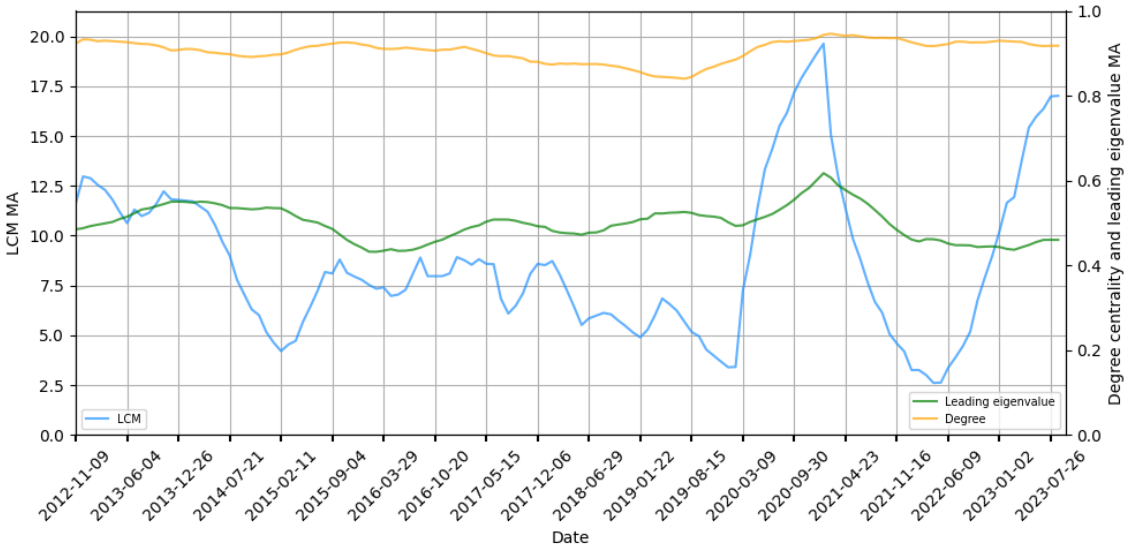


Figure 5: Moving average of CDX LCM, degree centrality and leading eigenvalue

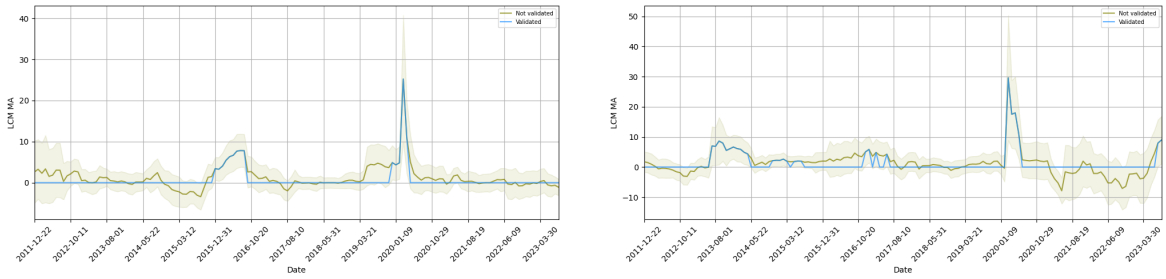


Figure 6: Consumer Discretionary sector CDX LCM **Left** Investment grade **Right** High yield

ply chains worldwide. Disruptions in the production and distribution of goods created challenges for companies within the discretionary sector, hindering their access to raw materials and logistics. Consequently, many businesses in non-essential retail sector either closed or experienced a major decline in sales. Hence, due to the pandemic, consumer behavior shifted, supply chain systems had disruptions and economic uncertainties escalated. Combination of these factors resulted into a notable increase in the contagion risk and vulnerability in the U.S. consumer discretionary sector. For a comprehensive view of sector-specific figures of CDX LCM, we refer the interested reader to Figure 9 in the Appendix.

4.2 Equity Tick Data

The choice of data frequency is a critical consideration in network analysis. To ensure that dependencies are captured in the model, we must strike the right balance between a frequency that is not too low or too high. Given that stock prices swiftly adjust to market changes, especially to similar stock price fluctuations, and trading occurs at high frequency to capitalize on these changes, we opt for tick data. We use tick log return data for 20 Financials sector companies (provided in Table 3 of Appendix) as well as both Lehman Brothers (LEH) and Bear Stearns (BSC). The tick data is obtained from Refinitiv, collected at one-second intervals, and spans from January 1, 2005 to December 31, 2010.

In Figure 7, we focus on the period surrounding the 2008 Financial Crisis to gain a deeper understanding of the centrality measure's behavior in the years preceding, during, and after the turmoil. We consider two versions of the system-level LCM: (i) for the network including LEH and BSC, for which we run the analysis for as long as the data is available (October 8, 2008 indicated by a red dashed line in the figure) and (ii) for the network excluding these two firms. In both versions, we observe an upward trend starting from early 2007 with the peak LCM reached in August 2008. A gradual decline is then observed when excluding LEH and BSC. This signifies that the contagion risk increased until reaching the peak of turmoil, subsequently decreasing after government interventions and as the economy and markets slowly returned to normalcy. This system-level trend closely aligns with the timeline of the financial crisis: the acquisition of Bear Sterns by JPMorgan on March 16, 2008, and the subsequent default of Lehman Brothers on September 15, 2008. The Wall Street Bailout law was passed on October 3,

2008, which included the $700 billion Troubled Asset Relief Program (TARP) aimed at rescuing banks from defaulting.

We plot the node-level LCM for three specific firms (American International Group (AIG), JPMorgan (JPM) and Goldman Sachs (GS)) alongside Lehman Brothers and Bear Stearns in Figure 7. We note that the node-level LCM for LEH and BSC have similar behavior to the system-level LCM until May 2008. Afterwards, they start to decline, due to which we experience a slight decline in the system-level LCM as well. The node-level drop is aligned with recognition of the substantial risk these two companies carried by investors. Due to this market behavior, investors treated these companies idiosyncratically thus reducing the spillover effects to the rest of the system from shocks they might experience. This is why, despite the systemic contagion risk rising, the LCM for LEH and BSC move inversely. This detachment is also reflected in the network structure. Figure 8 shows the interdependencies between LEH and BSC and all other firms for mid March and September 2008 time windows. As is evident, the connectivity of these two companies diminishes in both quantity and significance during their corresponding drop in their LCM.

Analyzing node-level LCMs reveals some similarities and differences between the three tickers, AIG, JPM and GS, and the delisted ones. We notice that AIG, prominent in the insurance sector, exhibits a similar behavior to LEH and BSC; its LCM took a downward turn following the bankruptcy and merger of LEH and BSC respectively. During the crisis, $182 billion was injected into AIG to prevent its default by the Treasury and Federal Reserve Bank of New York, as it was believed that its collapse would jeopardize the integrity of other firms. Essentially, AIG came under government ownership and was safeguarded, which is reflected in its LCM as it reaches nearly zero by early 2009. On the other hand, JPM and GS, for which the LCM consistently increases even after October 8, 2008, diverge from the system-level LCM. As these two companies were considered safer by investors, they maintained investments despite the systemic turmoil. However, under a shock scenarios to these companies, there would have been a high contagion risk. Looking into the stock prices of these tickers, we observe a significant decline in AIG's price, whereas JPM and GS experience a decrease but recover shortly afterwards. These patterns are aligned with the fluctuations seen in the LCM, indicating its ability to mirror the dynamics of the market.

Table 2 provides a concise summary of the logic above. When both node- and system-level LCMs are on the rise, more instability is introduced coming from individual institutions. Conversely, a declining trend at the node-level while the system-level LCM is increasing demands a closer look at the specific firm. Such divergence implies the firm is treated independently from the network and can imply different conditions: if the firm is struggling, portfolio detachment is triggered; if the firm is in a good shape, it is considered a safe investment opportunity. Similarly, when the system-level LCM is decreasing while the node-level is increasing, it holds implications based on the firm's condition; a struggling firm means a high likelihood of instability, while a healthy one becomes densely interconnected within the network,

drawing heavy reliance from investors. This interconnectedness elevates the risk of contagion under a shock scenario to these robust entities.

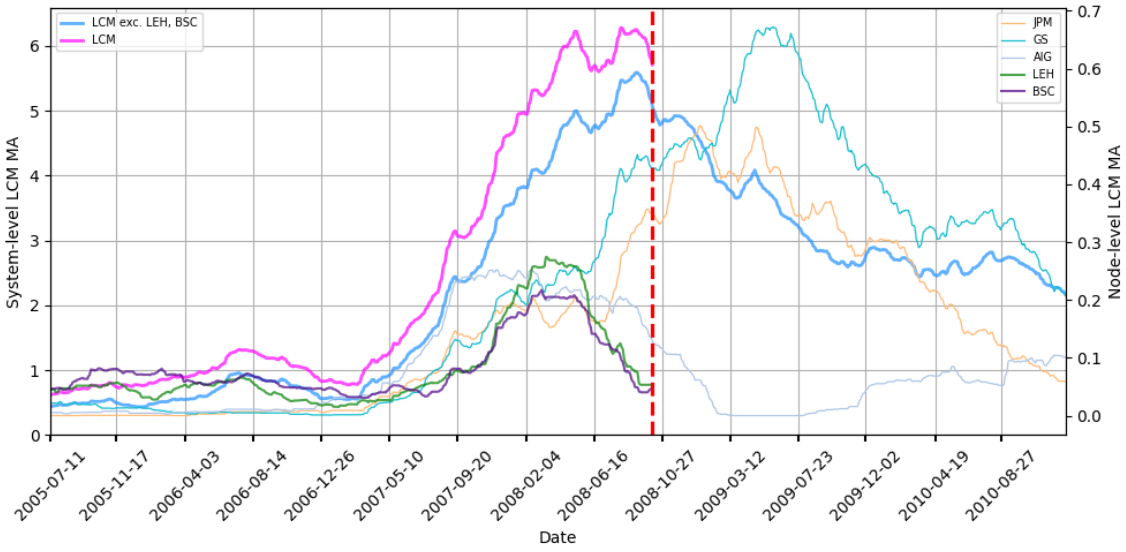


Figure 7: Node- and system-level LCM

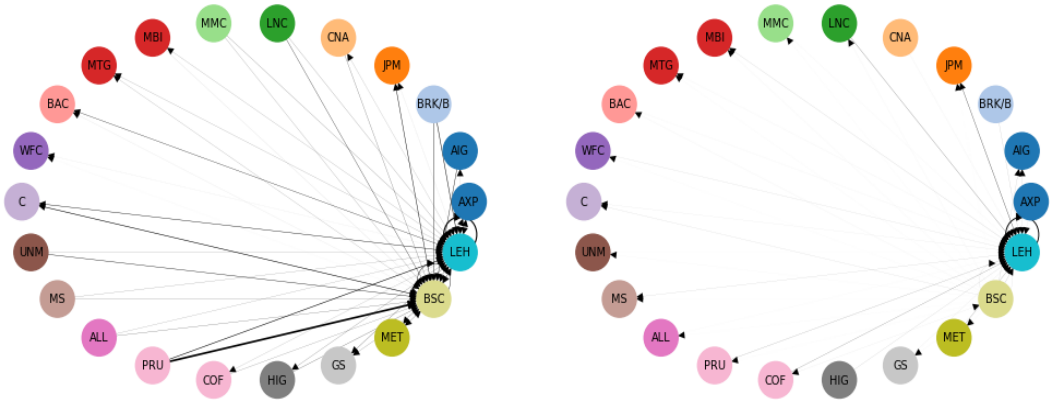


Figure 8: LEH and BSC network. **Left** Mid March 2008 **Right** Mid September 2008

Table 2: Node- and system-level LCM dynamics

System-level	Node-level	Intuition
↑	↑	More instability and risk in system (can result in a crisis)
↑	↓	Idiosyncratic treatment for the firm from investors
↓	↑	Heavy reliance on the firm
↓	↓	More stability and less risk in system (safe)

5 Conclusion

This paper presents a framework for constructing a weighted directed network, computing the absolute centrality measure, and statistically validating it. This offers a robust approach to assess contagion risk and vulnerability within financial systems. The Leontief centrality measure serves as a reliable proxy for contagion risk, allowing for meaningful comparisons across different time periods. To further enhance analysis of the network, we derive the distribution of the centrality measure, enabling the computation of confidence intervals and hypothesis testing. We show the alignment of the theoretical and empirical centrality distributions with the true centrality measure using simulated data. Leveraging real world financial data, CDX spread and equity tick data, we construct financial networks and evaluate the centrality of the system. During distressed years, we observe higher LCM values, underscoring the measure's ability to capture systemic risk within the network and assess the investors' beliefs. LCM can be used to form clusters of firms based on their susceptibility to contagion risk, thereby being useful for factor investing strategies and risk management practices. Central institutions through this measure can be identified to determine which firms require more rigorous oversight.

Future research can expand the scope of this work to, e.g., incorporate different numbers of lags in building the network. Such a construction would introduce a multilayered centrality metric. This extension could offer a more comprehensive understanding of the complex dynamics and interdependencies observed in financial data.

References

Acemoglu, D., Akcigit, U., and Kerr, W. (2016). Networks and the macroeconomy: An empirical exploration. *NBER Macroeconomics Annual*, 30:273–335.

Ali, R., Vause, N., and Zikes, F. (2016). Systemic risk in derivatives markets: a pilot study using cds data. Bank of England Financial Stability Papers 38, Bank of England.

Antràs, P. and Chor, D. (2013). Organizing the global value chain. *Econometrica*, 81(6):2127–2204. From March 2013.

Antràs, P. and Chor, D. (2018). *On the Measurement of Upstreamness and Downstreamness in Global Value Chains*, pages 126–194. Taylor & Francis Group.

Antràs, P., Chor, D., Fally, T., and Hillberry, R. (2012). Measuring the upstreamness of production and trade flows. *American Economic Review*, 102(3):412–16.

Bardoscia, M., Caccioli, F., Perotti, J. I., Vivaldo, G., and Caldarelli, G. (2016). Distress propagation in complex networks: The case of non-linear debtrank. *PLOS ONE*, 11(10):1–12.

Battiston, S., D'Errico, M., Gurciullo, S., and Caldarelli, G. (2015). Leveraging the network: a stress-test framework based on debtrank.

Battiston, S., Puliga, M., Kaushik, R., Tasca, P., and Caldarelli, G. (2012). Debtrank: too central to fail? financial networks, the fed and systemic risk. *Scientific reports*, 2(1):541–541.

Benoit, S., Colliard, J.-E., Hurlin, C., and Pérignon, C. (2017). Where the risks lie: A survey on systemic risk. *Review of Finance*, 21(1):109 – 152. Cited by: 193; All Open Access, Bronze Open Access, Green Open Access.

Bichuch, M. and Feinstein, Z. (2018). Optimization of fire sales and borrowing in systemic risk.

Blasques, F., Bräuning, F., and Lelyveld, I. v. (2018). A dynamic network model of the unsecured interbank lending market. *Journal of Economic Dynamics and Control*, 90(C):310–342.

Blasques, F., Koopman, S. J., Lucas, A., and Schaumburg, J. (2016). Spillover dynamics for systemic risk measurement using spatial financial time series models. *Journal of Econometrics*, 195(2):211–223.

Borgatti, S. P., Carley, K. M., and Krackhardt, D. (2006). On the robustness of centrality measures under conditions of imperfect data. *Social Networks*, 28(2):124–136.

Breiger, R., Carley, K., and Pattison, P. (2003). Dynamic social network modelling and analysis: Workshop summary and papers. *J. Artificial Societies and Social Simulation*, 6.

Brunnermeier, M., Clerc, L., and Scheicher, M. (2013). Assessing contagion risks in the cds market. *Financial Stability Review*, (17):123–134.

Capponi, A. and Larsson, M. (2015). Price Contagion through Balance Sheet Linkages. *The Review of Asset Pricing Studies*, 5(2):227–253.

Chan-Lau, J. A. (2018). Systemic centrality and systemic communities in financial networks. *Quantitative Finance and Economics*, 2(2):468–496.

Cont, R., Moussa, A., and Santos, E. B. (2013). *Network Structure and Systemic Risk in Banking Systems*, page 327–368. Cambridge University Press.

Cont, R. and Schaanning, E. (2017). Fire sales, indirect contagion and systemic stress testing. Working Paper 2017/2, Norges Bank.

di Iasio, G., Battiston, S., Infante, L., and Pierobon, F. (2013). Capital and Contagion in Financial Networks. MPRA Paper 52141, University Library of Munich, Germany.

D'Arcangelis, A. M. and Rotundo, G. (2016). Complex networks in finance. pages 209–235. Springer.

Eisenberg, L. and Noe, T. H. (2001). Systemic risk in financial systems. *Management Science*, 47(2):236–249.

Elliott, M., Golub, B., and Jackson, M. O. (2014). Financial networks and contagion. *American Economic Review*, 104(10):3115–53.

Farboodi, M. (2021). Intermediation and Voluntary Exposure to Counterparty Risk. NBER Working Papers 29467, National Bureau of Economic Research, Inc.

Feinstein, Z. and Hałaj, G. (2023). Interbank asset-liability networks with fire sale management. *Journal of Economic Dynamics and Control*, 155:104734.

Fichtner, J., Heemskerk, E. M., and Garcia-Bernardo, J. (2017). Hidden power of the big three? passive index funds, re-concentration of corporate ownership, and new financial risk. *Business and Politics*, 19(2):298–326.

Fink, K., Krüger, U., Meller, B., and Wong, L.-H. (2016). The credit quality channel: Modeling contagion in the interbank market. *Journal of Financial Stability*, 25(C):83–97.

Hevey, D. (2018). Network analysis: a brief overview and tutorial. *Health Psychology and Behavioral Medicine*, 6:301 – 328.

Huang, X., Vodenska, I., Havlin, S., and Stanley, H. E. (2013). Cascading failures in bi-partite graphs: Model for systemic risk propagation. *Scientific Reports*, 3(1).

Jackson, M. O. and Pernoud, A. (2020). Systemic risk in financial networks: A survey.

Jalili, M. and Perc, M. (2017). Information cascades in complex networks. *Journal of Complex Networks*, 5(5):665–693.

Lee, J.-S. and Pfeffer, J. (2015). Robustness of network centrality metrics in the context of digital communication data. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2015:1798–1807.

Leontief, W., editor (1986). *Input-Output Economics*. Number 9780195035278 in OUP Catalogue. Oxford University Press.

Leontief, W. W. (1936). Quantitative input and output relations in the economic systems of the united states. *The Review of Economics and Statistics*, 18(3):105–125.

Markose, M. S. M. (2012). Systemic Risk from Global Financial Derivatives: A Network Analysis of Contagion and Its Mitigation with Super-Spreader Tax. IMF Working Papers 2012/282, International Monetary Fund.

Markose, S., Giansante, S., and Shaghaghi, A. R. (2012). 'too interconnected to fail' financial network of us cds market: Topological fragility and systemic risk. *Journal of Economic Behavior Organization*, 83(3):627–646. The Great Recession: motivation for re-thinking paradigms in macroeconomic modeling.

Martin, C. and Niemeyer, P. (2019). Influence of measurement errors on networks: Estimating the robustness of centrality measures. *Network Science*, 7(2):180–195.

Millington, T. and Niranjan, M. (2020). Partial correlation financial networks. *Applied Network Science*, 5(1):11.

Moran, J. and Bouchaud, J.-P. (2019). May's instability in large economies. *Phys. Rev. E*, 100:032307.

Niu, Q., Zeng, A., Fan, Y., and Di, Z. (2015). Robustness of centrality measures against network manipulation. *Physica A: Statistical Mechanics and its Applications*, 438.

Olmo, J. (2021). Optimal portfolio allocation and asset centrality revisited. *Quantitative Finance*, 21(9):1475–1490.

Pacreau, G., Lezmi, E., and Xu, J. (2021). Graph neural networks for asset management.

Papana, A., Kyrtsou, C., Kugiumtzis, D., and Diks, C. (2017). Financial networks based on granger causality: A case study. *Physica A: Statistical Mechanics and its Applications*, 482:65–73.

Puliga, M., Caldarelli, G., and Battiston, S. (2014). Credit default swaps networks and systemic risk. *Scientific Reports*, 4.

Ricciardi, G., Montagna, G., Caldarelli, G., and Cimini, G. (2022). Dimensional reduction of solvency contagion dynamics on financial networks.

Singh, A., Singh, R., and Iyengar, S. (2020). Node-weighted centrality: a new way of centrality hybridization. *Computational Social Networks*, 7:6.

Smith, J. and Moody, J. (2013). Structural effects of network sampling coverage i: Nodes missing at random. *Social Networks*, 35:652–668.

Teteryatnikova, M. (2014). Systemic risk in banking networks: Advantages of "tiered" banking systems. *Journal of Economic Dynamics and Control*, 47(C):186–210.

Veraart, L. A. M. (2020). Distress and default contagion in financial networks. *Mathematical Finance*, 30(3):705–737.

Yu, H., Chu, W., Ding, Y., and Zhao, X. (2021). Risk contagion of global stock markets under covid-19:a network connectedness method. *Accounting & Finance*, 61(4):5745–5782.

A Proofs

A.1 Proof of Pair-Level LCM Distribution (Theorem 3)

As we are focused on the asymptotic distribution of $(I - \hat{A})^{-1}_{ij}$, we consider the distribution associated with the first-order Taylor expansion

$$(I - \hat{A})^{-1}_{ij} = (I - A)^{-1}_{ij} + (I - A)^{-1}_{i\cdot}\Delta(I - A)^{-1}_{\cdot j}.$$

Under the assumptions of this theorem, the noise of the model Δ in Equation 4 follows a multivariate normal distribution. In fact, due to the linear relation between the (first-order approximation of the) pair-level LCM $(I - \hat{A})^{-1}_{ij}$ and the noise Δ, it trivially follows that the pair-level LCM must also follow a normal distribution. Thus, to complete this proof we only need to consider the mean and variance of this linear form.

First, we note that the mean of the noise $E[\Delta] = 0$ by assumption. Therefore, for any $T > 0$,

$$\sqrt{T}E\left[(I - \hat{A})^{-1}_{ij} - (I - A)^{-1}_{ij}\right] = 0\,.$$

Second, consider the variance of the noise, i.e., $var(\Delta_{\cdot j}) = \sigma_j^2(X^\top X)^{-1}$ where σ_j is the standard deviation of the noise of the j^{th} time series by standard OLS arguments. Therefore, by the first-order Taylor expansion specified above,

$$\begin{aligned}
var\left(\sqrt{T}[(I - \hat{A})^{-1}_{ij} - (I - A)^{-1}_{ij}]\right) &= T\,var\left((I - A)^{-1}_{i\cdot}\Delta(I - A)^{-1}_{\cdot j}\right) \\
&= T\left[(I - A)^{-1}_{\cdot j}\right]^\top var\left((I - A)^{-1}_{i\cdot}\Delta\right)(I - A)^{-1}_{\cdot j} \\
&= \left[(I - A)^{-1}_{\cdot j}\right]^\top \left[(I - A)^{-1}_{i\cdot}\Sigma^T_{lk}\left[(I - A)^{-1}_{i\cdot}\right]^\top\right]_{lk}(I - A)^{-1}_{\cdot j}
\end{aligned}$$

where $\Sigma^T_{lk} = \rho_{lk}\sigma_l\sigma_k Q_T^{-1}$ provides the covariance between $\Delta_{\cdot l}$ and $\Delta_{\cdot k}$ (with multiplier $\sqrt{T}$). For this formulation we specify $Q_T = T^{-1}X^\top X$. As $Q = \lim_{T\to\infty} Q_T$, the result follows.

A.2 Proof of Node-Level LCM Distribution (Corollary 4)

This result follows trivial as the distribution of a linear combination of a multivariate normal is itself normally distributed. Following directly from Theorem 3, asymptotically, we find:

$$\sqrt{T}E\left[(I - \hat{A})^{-1}_{i\cdot}\boldsymbol{w} - (I - A)^{-1}_{i\cdot}\boldsymbol{w}\right] = 0\,,$$

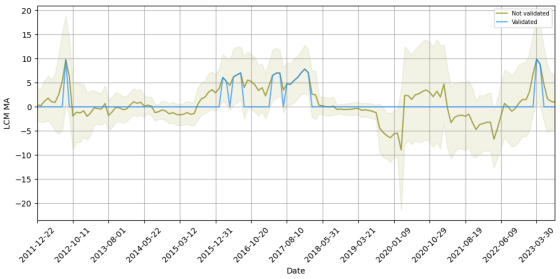
$$T\,var\left((I - \hat{A})^{-1}_{i\cdot}\boldsymbol{w} - (I - A)^{-1}_{i\cdot}\boldsymbol{w}\right) = \left[(I - A)^{-1}\boldsymbol{w}\right]^\top \left[(I - A)^{-1}_{i\cdot}\Sigma_{lk}\left[(I - A)^{-1}_{i\cdot}\right]^\top\right]_{lk}(I - A)^{-1}\boldsymbol{w}\,.$$

A.3 Proof of Test Statistic's Distribution (Corollary 5)

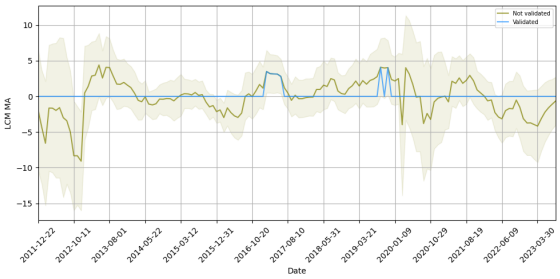
First, taking advantage of the weak law of large numbers, the estimated adjacency matrix $\hat{A}$ converges in probability to the true adjacency matrix A. By the continuous mapping theorem, the estimated Leontief inverse $(I - \hat{A})^{-1}$ converges in probability to the true Leontief inverse $(I - A)^{-1}$. Therefore, based on the Central Limit Theorem and Slutsky's Theorem, the test statistic Z asymptotically follows a standard normal distribution. In addition, we note that this convergence does not introduce any bias due to the asymptotics used herein.

B Additional Details and Figures for Section 4.1

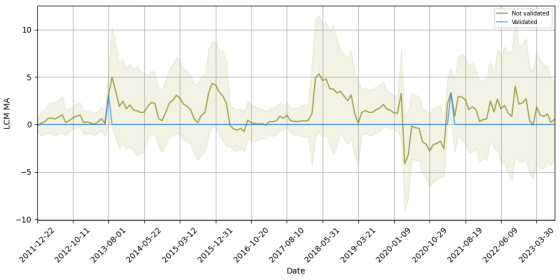
Data CDX spread data, which are available as Markit's North American Investment Grade and High Yield indices, closely reflect the credit quality and direction of the underlying basket in one tradable instrument. These indices are domiciled in North America, have 5 year tenor, USD denomination, quarterly coupon frequency, without restructuring, and semi-annual roll periods in March and September. The CDX IG is composed of 125 equally weighted credit default swaps on investment grade entities traded in the CDS market. Meanwhile, the CDX HY comprises 100 non-investment grade liquid entities within the CDS market.



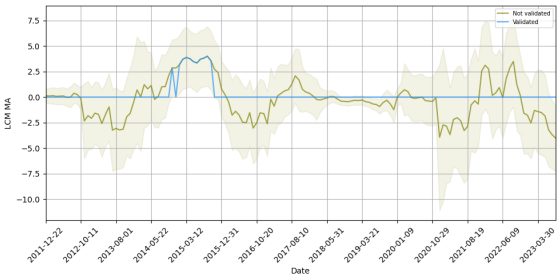
(a) Communication Services IG



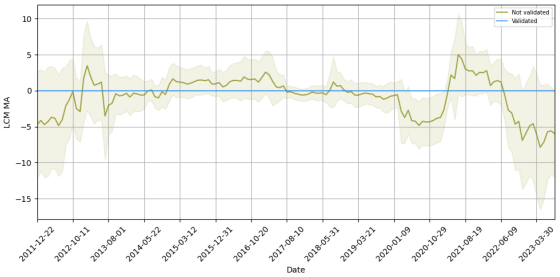
(b) Communication Services HY



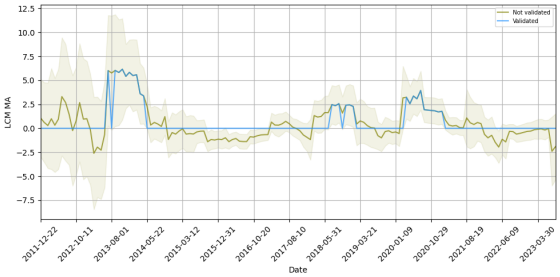
(c) Consumer Staples IG



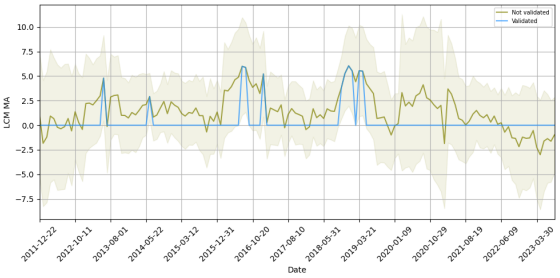
(d) Consumer Staples HY



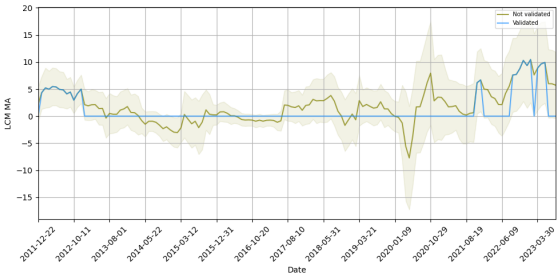
(e) Energy IG



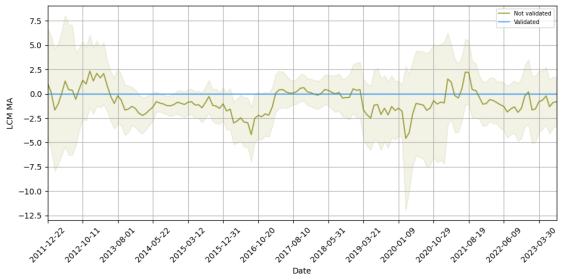
(f) Energy HY



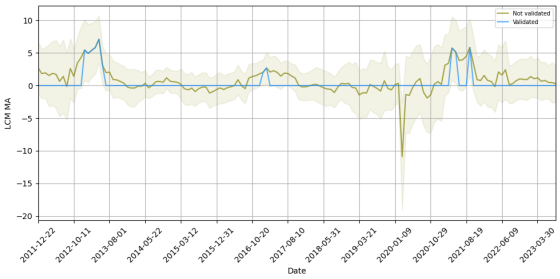
(g) Financials IG



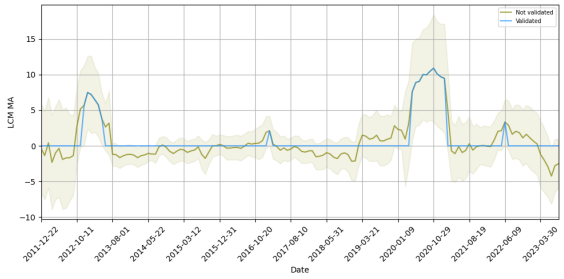
(h) Financials HY



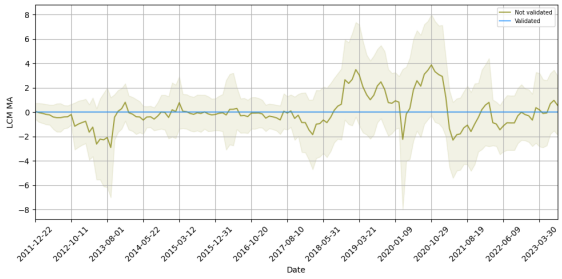
(i) Healthcare IG



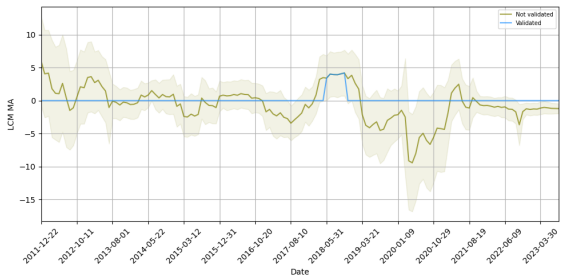
(j) Healthcare HY



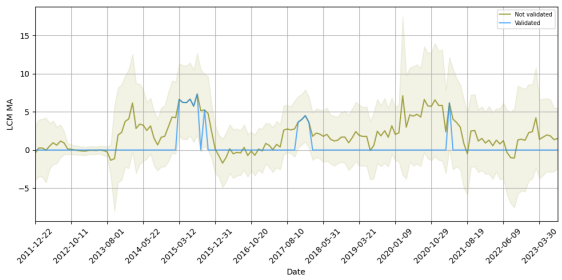
(k) Industrials IG



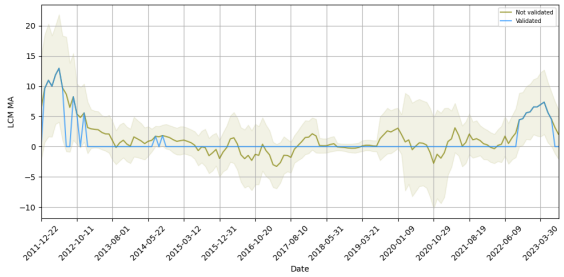
(l) Industrials HY



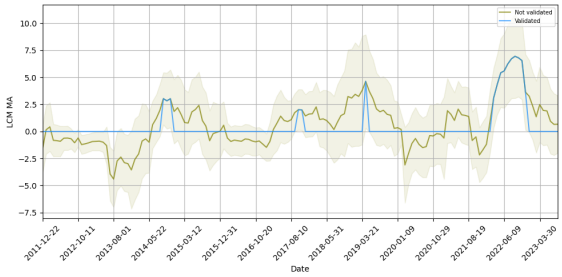
(m) Materials IG



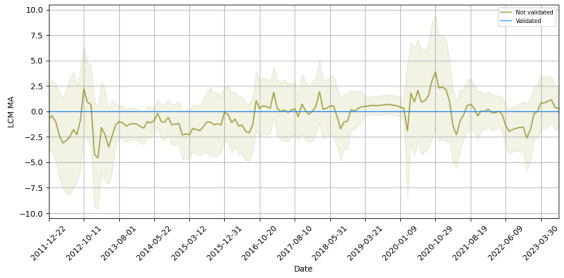
(n) Materials HY



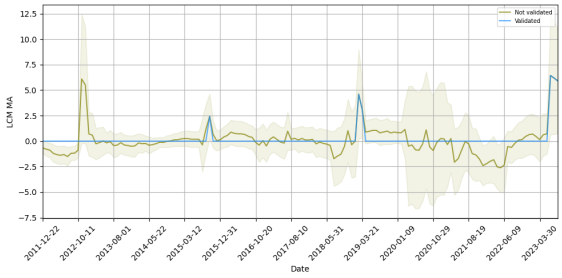
(o) Technology IG



(p) Technology HY



(q) Utilities IG



(r) Utilities HY

Figure 9: Validated, not validated and shaded confidence interval of CDX LCM.

C Details of the Data Set for Section 4.2

Table 3: Financial sector companies

Ticker	Short name	Industry Group[3]
AXP	American Express Co	Financial Services
AIG	American International Group	Insurance
BRK/B	Berkshire Hathaway Inc-CL B	Insurance
JPM	JPMorgan Chase & Co	Banking
CNA	CNA Financial Corp	Insurance
LNC	Lincoln National Corp	Insurance
MMC	Marsh & Mclennan Cos	Insurance
MBI	Mbia Inc	Insurance
MTG	Mgic Investment Corp	Financial Services
BAC	Bank of America Corp	Banking
WFC	Wells Fargo & Co	Banking
C	Citigroup Inc	Banking
UNM	Unum Group	Insurance
MS	Morgan Stanley	Financial Services
ALL	Allstate Corp	Insurance
PRU	Prudential Financial Inc	Insurance
COF	Capital One Financial Corp	Financial Services
HIG	Hartford Financial SVCS GRP	Insurance
GS	Goldman Sachs Group Inc	Financial Services
MET	Metlife Inc	Insurance

[3] Based on Bloomberg Industrial Classification Standard (BICS) level 2.