Spectral Networks for Times Series

Petre Caraiani* Maria Grith†

Abstract

We represent an autoregressive panel of time series as a sum of (orthogonal) scale-specific autoregressive panels capturing interactions between time-series components at different frequencies. This leads to scale-specific autoregressive matrices and inverse covariance matrices of the innovations that are assumed to be sparse, and can - but are not forced to - have different sparsity patterns across scales. The decomposition leads to a novel representation of the system as a series of (orthogonal) scale-specific networks in terms of a multilayer directed graph representing predictive Granger relations and a multilayer undirected graph representing contemporaneous partial correlations at different frequencies. A method is introduced to estimate the model. Building on these insights, we apply the methodology to analyze a panel of volatility measures in the financial sector. We illustrate how the methods may lead to economically meaningful new results on spillover effects and systemically important institutions at different frequencies.

Keywords: Networks, Multivariate Time Series, Long Run Covariance, Frequency, GLASSO, LARS, Systemic risk, Volatility Spillover
JEL classification: C01 C32 G1 E32

1 Introduction

The interaction between the volatility of different stocks is a key element in understanding the propagation of risk in the financial markets. Among various approaches to tackle this topic, network analysis - concerned with examining the relationships between a group of individual entities - has emerged as a useful methodological tool for time series analysis. In these networks,

*Faculty of Business Administration in Foreign Languages, Bucharest University of Economic Studies; Institute for Economic Forecasting, Romanian Academy

†E-mail: grith@ese.eur.nl. Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands

nodes denote univariate time series components, and edges describe different forms of conditional dependence among components. The interactions in a dynamic system change with the subset analyzed but may also change across frequencies. While most of the macro and finance literature has focused on the observational domain, more recent evidence suggests that it is crucial to understand the interactions between the subcomponents of time series that evolve within cycles of different lengths, i.e., frequency bands.

The main contribution of this article is to propose a novel network model that captures interactions in a panel of time series for different frequency bands. Our approach relies on two strands of literature: one on modeling the linear interactions between multivariate time series using networks or graphs, see Barigozzi and Brownlees (2019); and the second one on using orthogonal projections to decompose time series in components with different resolutions, colloquially named scales. The Extended Wold Decomposition (EWD) proposed by Ortu et al. (2020) is particularly appealing because it fully restricts interactions between components at different scales. Consequently, the joint interactions can be decomposed into scale-specific interactions, that are fully separable among scales and lead to multilayer networks or graphs. This allows us to represent the system as a sum of orthogonal scale-specific sub-systems capturing interactions between time-series components at different frequencies. For identification, we focus on linear autoregressive panels that are characterized by scale-specific autoregressive matrices and inverse covariance matrices of the innovations that are assumed to be sparse, and can - but are not forced to - have different sparsity patterns across scales. The representation leads to a series of (orthogonal) scale-specific networks in terms of a multilayer directed graph representing predictive Granger relations and a multilayer undirected graph representing contemporaneous partial correlations at different frequencies. This representation allows us to incorporate dependence between components corresponding to long, medium, and short-term shocks.

A second contribution is the estimation of the networks. We estimate contemporaneous and Granger causality networks for both the actual financial series and the details obtained through the EWD decomposition. We use the graphical LASSO (GLASSO) to estimate the contemporaneous network and the LARS algorithm to estimate the Granger causality network.

We also contribute to the understanding of the risk in the scale domain. We show that the central nodes of the financial networks are different for short-term, medium-term and long-term components. The degree distribution is also different for the different frequencies.

Few papers have addressed this. One notable contribution is due to Baruník and Křehlík (2018). The study looks at connectedness along decomposed components obtained through

spectral representation of variance decomposition. Shocks for high frequency are not persistent, in contrast to those at lower frequencies that have more durable effects.

Another contribution on this topic is due to Teply and Kvapilikova (2017), who proposed Wavelet Conditional Value at Risk to determine short-term, medium-term, and long-term financial connectedness.

2 Literature

This article analyzes the network characterization of the interactions between multivariate time series when the scales (or frequencies) are taken into account. To develop our model, we adopt established techniques from the network/graphs and time series literature. We also contribute methodologically with an estimation method for the multilevel graphs. We apply the methodology to analyze a panel of volatility measures in the financial sector and illustrate how it leads to economically meaningful new results on spillover effects and systemically important institutions at different frequencies. Our work is related to several strands of literature.

The first strand of literature builds on the Extended Wold Decomposition (EWD) literature. Cerreia-Vioglio et al. (2023) propose a Hilbert A-module approach for the Multivariate Wold decompositions that features a notion of orthogonality that permits to easily retrieve two important orthogonal decompositions for weakly stationary vector processes. One is the celebrated multivariate classical Wold decomposition (MCWD, henceforth). The other is the multivariate extended Wold decomposition (MEWD, henceforth), which constitutes the multivariate version of the extended Wold decomposition of Ortu et al. (2020), and it is used by Bandi et al. (2019, 2021) in financial economics settings, and extended by Baruník and Vacha (2023) for locally stationary processes.

The second strand of literature focuses on network representations of time series. One of the most known ones is due to Diebold and Yılmaz (2014a) who proposed constructing a connectedness measure based on the variance decomposition from an estimated VAR model. In a related contribution, Billio et al. (2012) suggested alternative measures of connectedness using both principal components and Granger-causality measures. Using data on hedge funds, banks, brokers and insurance companies they found higher connectedness between them and that banks are central in the transmission of financial shocks. Our approach is similar in the sense of contructing measures of "connectedness" based on various network estimations. There are also several differences with their work. First, we estimate both a contemporaneous and a Granger-causality network. Second, we also look at different scales, while Diebold and Yılmaz

(2014a) only looked at aggregate series.

The third strand of literature contributes to the understanding of systemic risk using networks. The recent work by Jackson and Pernoud (2021) reviews the contributions done using networks on the topic of systemic risk. The analysis discusses two types of systemic risk, one comes from contagion coming from direct externalities [1]. The second type of systemic risk is based on feedback effects. Our work is based on the first type of systemic risk, based on contagion effects. In another contribution, Cont et al. (2013) used data on financial institutions from Brazil to construct a financial contagion index. Their main finding was that systemic risk comes from several central nodes/institutions from the network.

A pertinent question would be how to differentiate between systemic risk, spillover effects, and contagion. Since, we work within a network based approach, these notions are all modelled within this framework. Our contribution links these topics. As shown by Jackson and Pernoud (2021), there are two types of systemic risk within a network, and it is specifically on the one related to contagion that we focus on. Furthermore, it is through spillovers that the contagion makes its effects. **?** propose a modelling framework for spillover effects with mixed-frequency variables that allow for differences in sampling frequency between financial and macroeconomic data. Baruník and Křehlík (2018) explores connectedness among financial variables that arise due to heterogeneous frequency responses to shocks, where the frequency components are defined as a Fourier transform of $MA(\infty)$ filtered series.

The fourth strand of literature: regards volatility modeling. Volatility networks. As in Diebold and Yılmaz (2014a), we study volatility networks. As outlined by the authors of this study, volatility connectedness has an intuitive meaning as it can be understood as a sort of "investor fear". The second reason one should focus on volatility measures is that it is strongly related to the crisis.

3 Methodology

3.1 Extended Wold Decomposition

Extended Wold Decomposition (EWD) in Ortu et al. (2020) applies the transform to the infinite moving-average parameters and innovations of a covariance stationary process. Consequently, each resulting scale-specific component supports a Wold decomposition at a particular scale.

[1] This happens when a negative shock on one bank, be it a default or a firesale of assets leads to a negative impact for other banks

3.1.1 Univariate case

Let's consider a zero-mean covariance stationary one-dimensional time series process $\boldsymbol{y} = \{y_t\}_{t\in\mathbb{Z}}$. Then, it admits the classical Wold-type decomposition (Wold (1938)):

$$y_t = \sum_{h=0}^{+\infty} \alpha_h \varepsilon_{t-h}$$

with variance white noise $\boldsymbol{\varepsilon} = \{\varepsilon_t\}_{t\in\mathbb{Z}}$, where the coefficients α_h are square-summable, independent of t, and $\alpha_h = \mathbb{E}[y_t \varepsilon_{t-h}]$. The process $\boldsymbol{\varepsilon}$ is commonly called the sequence of fundamental innovations of $\boldsymbol{y}$.

We use up to 25 lags for each univariate AR(p) process, following Ortu et al. (2020). We estimate for each individual series an AR(p) model.

To define scale-specific or spectral decomposition of time series, we define the following processes by introducing a scale j that is a measure of aggregation of the coefficients and processes:

(i) **Scale-specific innovations:** for any fixed $j \in \mathbb{N}$, the process $\boldsymbol{\varepsilon}^{(j)} = \left\{\varepsilon_t^{(j)}\right\}_{t\in\mathbb{Z}}$ with

$$\varepsilon_t^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-i} - \sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-2^{j-1}-i} \right) \tag{3.1}$$

is a MA $(2^j - 1)$ with respect to the classical Wold innovations of $\boldsymbol{y}$ and $\left\{\varepsilon_{t-k2^j}^{(j)}\right\}_{k\in\mathbb{Z}}$ is a unit variance white noise;

(ii) **Scale-specific MA parameters:** for any $j \in \mathbb{N}, k \in \mathbb{N}_0$, the coefficients $\psi_k^{(j)}$ are uniquely determined via

$$\psi_k^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \alpha_{k2^j+i} - \sum_{i=0}^{2^{j-1}-1} \alpha_{k2^j+2^{j-1}+i} \right) \tag{3.2}$$

hence, they are independent of t and $\sum_{k=0}^{\infty} \left(\psi_k^{(j)}\right)^2 < +\infty \forall j \in \mathbb{N}$;

(iii) **Scale-specific component** of process y_t

$$y_t^{(j)} = \sum_{k=0}^{+\infty} \psi_k^{(j)} \varepsilon_{t-k2^j}^{(j)}$$

for any $j, l \in \mathbb{N}, p, q, t \in \mathbb{Z}, \mathbb{E}\left[x_{t-p}^{(j)} x_{t-q}^{(l)}\right]$ depends at most on $j, l, p-q$. Moreover,

$\mathbb{E}\left[x^{(j)}_{t-m2^j} x^{(l)}_{t-n2^l}\right] = 0$ for all $j \neq l, m, n \in \mathbb{N}_0$, and $t \in \mathbb{Z}$.

Putting equations together, y_t can be expressed as the infinite orthogonal sum of scale-specific components. We call this the extended Wold decomposition:

$$y_t = \sum_{j=1}^{\infty} y_t^{(j)} = \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} \psi_k^{(j)} \varepsilon_{t-k2^j}. \tag{3.3}$$

For each j the process $\{\varepsilon_{t-k2^j},\, k \in \mathbb{Z}^+\}$ has a zero mean and unit variance and $(\psi_k^{(j)},\, k = 1, \ldots, T)$ form a system of orthogonal basis.

3.1.2 Multivariate case

Consider the zero mean, covariance stationary, n-variate process $\boldsymbol{y} = \{\boldsymbol{y}_t\}_{t\in\mathbb{Z}}$, with $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{nt})'$. Its Wold representation is

$$\boldsymbol{y}_t = \sum_{k=0}^{+\infty} \boldsymbol{\alpha}_k \boldsymbol{\varepsilon}_{t-k}, \tag{3.4}$$

where $\boldsymbol{\varepsilon}_t$, $t \in \mathbb{Z}$ is a n-dimensional vector of (possibly) cross-correlated white noise shocks and the Wold coefficients $\boldsymbol{\alpha}_k$ are, for all $k \in \mathbb{N}_0, n \times n$ matrices. By analogy with the scalar case, we can now write

$$\boldsymbol{y}_t = \sum_{j=1}^{\infty} \boldsymbol{y}_t^{(j)} = \sum_{j=1}^{+\infty} \sum_{k=0}^{+\infty} \boldsymbol{\Psi}_k^{(j)} \boldsymbol{\varepsilon}_{t-k2^j}^{(j)}. \tag{3.5}$$

For any $j \in \mathbb{N}$, the $N \times N$ matrices $\boldsymbol{\Psi}_k^{(j)}$ are unique Discrete Haar Transforms (DHTs) of the original Wold coefficients:

$$\boldsymbol{\Psi}_k^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \boldsymbol{\alpha}_{k2^j+i} - \sum_{i=0}^{2^{j-1}-1} \boldsymbol{\alpha}_{k2^j+2^{j-1}+i} \right) \tag{3.6}$$

Similarly,

$$\boldsymbol{\varepsilon}_t^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \boldsymbol{\varepsilon}_{t-i} - \sum_{i=0}^{2^{j-1}-1} \boldsymbol{\varepsilon}_{t-2^{j-1}-i} \right). \tag{3.7}$$

We emphasize that the components of the extended Wold are uncorrelated for all t.

3.2 Identification

To define an economically meaningful network, we model $\boldsymbol{y}_t$ as an autoregressive of order p process VAR(p)

$$\boldsymbol{y}_t = \sum_{k=1}^{p} \boldsymbol{A}_k \boldsymbol{y}_{t-k} + \boldsymbol{\epsilon}_t \quad \boldsymbol{\epsilon}_t \sim wn\left(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}\right), \tag{3.8}$$

where $\boldsymbol{A}_k$ and $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ are $n \times n$ matrices. Throughout, the VAR is assumed to be stable and $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ to be positive definite. Under standard assumptions, a stable VAR is covariance stationary and therefore, admits a Wold representation. This implies that this representation of $\boldsymbol{y}_t$ allows us to identify parameters $\boldsymbol{\Psi}_k^{(j)}$ and innovations $\boldsymbol{\varepsilon}_t^{(j)}$ in Equation (3.5).

The VAR(p) process in Equation (3.8) can be written in a companion form, as an np-dimensional VAR(1) process by stacking p consecutive $\boldsymbol{y}_t$ variables in an $(np \times 1)$-dimensional vector

$$\underbrace{\begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{y}_{t-1} \\ \boldsymbol{y}_{t-2} \\ \vdots \\ \boldsymbol{y}_{t-p+1} \end{bmatrix}}_{\boldsymbol{Y}_t} = \underbrace{\begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}}_{\boldsymbol{\Gamma}} \underbrace{\begin{bmatrix} \boldsymbol{y}_{t-1} \\ \boldsymbol{y}_{t-2} \\ \boldsymbol{y}_{t-3} \\ \vdots \\ \boldsymbol{y}_{t-p} \end{bmatrix}}_{\boldsymbol{Y}_{t-1}} + \underbrace{\begin{bmatrix} \boldsymbol{\epsilon}_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\boldsymbol{U}_t}$$

Thus, we have

$$\boldsymbol{Y}_t = \boldsymbol{\Gamma} \boldsymbol{Y}_{t-1} + \boldsymbol{U}_t. \tag{3.9}$$

The stability of $\boldsymbol{Y}_t$ is contingent upon the absolute value of all eigenvalues of $\boldsymbol{\Gamma}$ being less than one. It can be shown that if the process $\boldsymbol{y}_t$ is stable, then $\boldsymbol{Y}_t$ is also stable. The stability of the process $\boldsymbol{Y}_t$ ensures the existence of the inverse VAR operator $\left(\boldsymbol{I}_{np} - \boldsymbol{\Gamma} L\right)^{-1} = \sum_{k=0}^{\infty} \boldsymbol{\gamma}_k L^k$, where $\boldsymbol{\gamma}_k = \boldsymbol{\Gamma}^k$. As a result, one can obtain the following Wold moving average representation of $\boldsymbol{Y}_t$:

$$\boldsymbol{Y}_t = \sum_{k=0}^{\infty} \boldsymbol{\gamma}_k \boldsymbol{U}_{t-k}. \tag{3.10}$$

The long-run concentration matrix of $\boldsymbol{Y}_t$ is

$$\boldsymbol{\Sigma}_{\boldsymbol{Y}}^{L} = \left(\boldsymbol{I}_{np} - \boldsymbol{\Gamma}'\right) \boldsymbol{\Sigma}_{\boldsymbol{U}}^{*,-1} \left(\boldsymbol{I}_{np} - \boldsymbol{\Gamma}\right) \tag{3.11}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{U}}^{*,-1}$ is the Moore–Penrose inverse of $\boldsymbol{\Sigma}_{\boldsymbol{U}}$.

Define $\boldsymbol{e}_{np}^{j,n}$, $j = 1, \cdots, p$ an $np \times n$ vector whose j-th block values is the $n \times n$ identity matrix

and whose values are zero otherwise, with $(\boldsymbol{e}_{np}^{j,n})'\boldsymbol{e}_{np}^{j,n} = \boldsymbol{I}$ and $\boldsymbol{e}_{np}^{j,n}(\boldsymbol{e}_{np}^{j,n})' = \boldsymbol{I} + \boldsymbol{0}_{np\times np}$.[2] Now, we can derive the long-run covariance matrix of $\boldsymbol{y}_t$ (also known as the long-run concentration matrix $\boldsymbol{K}$) by focusing on the first n components of $\boldsymbol{Y}_t$.

$$\begin{aligned}\boldsymbol{K} := \boldsymbol{\Sigma}_{\boldsymbol{y}}^{L} = (\boldsymbol{e}_{np}^{j,n})'\boldsymbol{\Sigma}_{\boldsymbol{L}}^{L}\boldsymbol{e}_{np}^{j,n} &= (\boldsymbol{e}_{np}^{j,n})'\left(\boldsymbol{I}-\boldsymbol{\Gamma}'\right)\boldsymbol{\Sigma}_{\boldsymbol{U}}^{*,-1}\left(\boldsymbol{I}-\boldsymbol{\Gamma}\right)\boldsymbol{e}_{np}^{j,n} \\ &= \left(\boldsymbol{I}-(\boldsymbol{e}_{np}^{j,n})'\boldsymbol{\Gamma}'\boldsymbol{e}_{np}^{j,n}\right)(\boldsymbol{e}_{np}^{j,n})'\boldsymbol{\Sigma}_{\boldsymbol{U}}^{*,-1}\boldsymbol{e}_{np}^{j,n}\left(\boldsymbol{I}-(\boldsymbol{e}_{np}^{j,n})'\boldsymbol{\Gamma}\boldsymbol{e}_{np}^{j,n}\right) \\ &= \left(\boldsymbol{I}-\boldsymbol{G}'\right)\boldsymbol{C}(\boldsymbol{I}-\boldsymbol{G})\end{aligned} \tag{3.12}$$

where $\boldsymbol{C} = \boldsymbol{\Sigma}_{\epsilon}^{-1} = (\boldsymbol{e}_{np}^{j,n})'\boldsymbol{\Sigma}_{\boldsymbol{U}}^{*,-1}\boldsymbol{e}_{np}^{j,n}$ and $\boldsymbol{G} = (\boldsymbol{e}_{np}^{j,n})'\boldsymbol{\Gamma}\boldsymbol{e}_{np}^{j,n}$, with notations similar to Barigozzi and Brownlees (2019) to define Contemporaneous undirected and Granger directed networks (graphs).

3.3 Spectral Networks for Time Series

To define scale-specific networks for time series, we will use VAR(1) specification of $\boldsymbol{Y}_t$

$$\boldsymbol{Y}_t = \sum_{k=0}^{\infty}\boldsymbol{\Gamma}^k\boldsymbol{U}_{t-k}. \tag{3.13}$$

The network representation of the scale-specific system is

$$\boldsymbol{K}^{(j)} = \left(\boldsymbol{I}-(\boldsymbol{G}^{(j)})'\right)\boldsymbol{C}^{(j)}(\boldsymbol{I}-\boldsymbol{G}^{(j)}) \tag{3.14}$$

with $\boldsymbol{G}^{(j)} = (\boldsymbol{e}_{np}^{j,n})'\boldsymbol{\Gamma}^{2k}\boldsymbol{e}_{np}^{j,n}$ the directed Granger and $\boldsymbol{C}^{(j)} = (\boldsymbol{e}_{np}^{j,n})'\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{U}}^{(j),*,-1}\boldsymbol{e}_{np}^{j,n}$ the undirected Contemporaneous spectral networks.

A proof and definition of parameters is given in Appendix A.2.

4 Estimation

We combine several approaches to estimate the two key matrices: the Granger causality network G and the contemporaneous network C. While (Barigozzi and Brownlees, 2019) uses bases the estimations of the two matrices on equation (3.4), we estimate first the C matrix following the Graphical Lasso approach, see (Yuan and Lin, 2006) for the representation which we describe

[2]Here we use an abuse of notations to express the structure of the matrix.

below:

$$\hat{\boldsymbol{K}} = \arg\min_{\boldsymbol{K}\in S}\{tr(\bar{\Sigma}\boldsymbol{K} - log\ det\ \boldsymbol{K}) + \lambda \sum_{i\neq j} |k_{ij}|\} \tag{4.1}$$

In the estimation, we follow (Friedman et al., 2007) who recasted the estimation as a sequence of basic LASSO regressions.

To select the λ, we set a grid of values and select the optimal λ using the BIC computed as below:

$$BIC(\lambda) = log(RSS(\lambda)) + \frac{log(T)}{T} \times n \tag{4.2}$$

λ is selected optimally for the actual series, as well as for each frequency component (be it short-run, long-run or medium-run). To select λ, we use the BIC criterion.

In order to determine the Granger causality network, we follow a two-step approach. First, we estimate A VAR(1) model as in equation (3.4), using the LARS algorithm, see (Efron et al., 2004). The optimal model is selected by minimizing the BIC criterion. Then we construct an adjacency matrix setting that a connection exists whenever the coefficient determined in the previous step is not zero.

LARS is a stepwise forward algorithm which starts from a different assumption as compared to LASSO. For LARS, initially, all coefficients are assumed to be zero. Then the procedure activates the variables with the strongest correlation with the residual. The algorithm continues until all variables become active.

The algorithm consists in the following steps:

1. The variables X are standardized to have zero mean and a variance of 1.
2. All β's are set to zero, while the residuals are set at $r_0 = y_t$
3. We determine the variable from X with the highest correlation with r_0
4. We further activate the first β_i based on the least square values resulting from $X_j' r_0$, until we get another variable X_i maximally correlated with the current value of residuals, as given by r_k
5. We repeat the previous step until another variables becomes active
6. The algorithm is repeated until all coefficients are non-zero.

Note that we use the same algorithm also to estimate $\boldsymbol{K}^{(j)}$ and its components.

5 Data

We use data from the financial sector from S&P 500. We have daily returns from 2000 to 2015. Focusing on the financial sector helps with comparing our results with related work from Diebold and Yılmaz (2014b) or Billio et al. (2012). Appendix A shows the sample of companies.

We present here the results for the simulations and estimations. The aggregate series (for real data or simulated data) have been decomposed into 9 different components. From here we compute the short-term, medium-term and long-term components as given by:

$$y_{long} = \sum_{j=7}^{9} y_j \tag{5.1}$$

$$y_{med} = \sum_{j=4}^{6} y_j \tag{5.2}$$

$$y_{short} = \sum_{j=1}^{3} y_j \tag{5.3}$$

We perform this decomposition in order to reveal the different components at different frequencies. Each j component corresponds to $2^{(j-1)} - 2^{(j)}$ days frequencies. Thus, the short run corresponds to 1. Thus, the short-term component corresponds to 1 day to 8 days of trading, the medium-run component corresponds to 8 days to $2^6 = 64$ trading days, while the long-run component corresponds to 64 trading days to 512 days (about two years).

6 Results

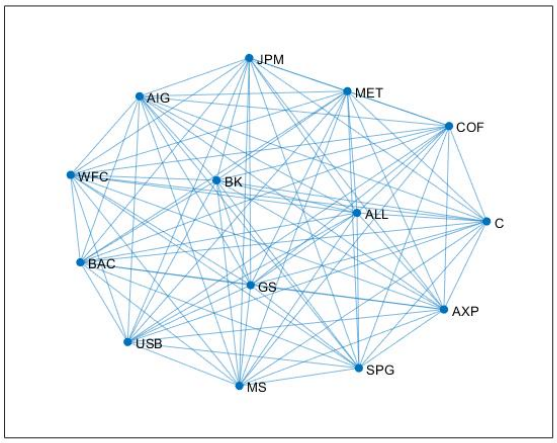
We first look at the results for the aggregate series, see Figures 5a and 6a below. We also look at the results for the decomposed series, see Figures 5 and 6, b, c and d. While the contemporaneous correlation matrices look at a first look quite similar, there are some significant differences with respect to the long-run Granger causality matrices.

First, we can see that the short-run Granger causality network is much sparse than the medium or the long run. This might come from the fact that the short-run components are more erratic and harder to predict. This is in line with previous findings on the fact that short-run components are harder to be attracted. For example, Caraiani (2017) showed that the short-run component of exchange rates, as identified using wavelets, is hardest to be predicted.

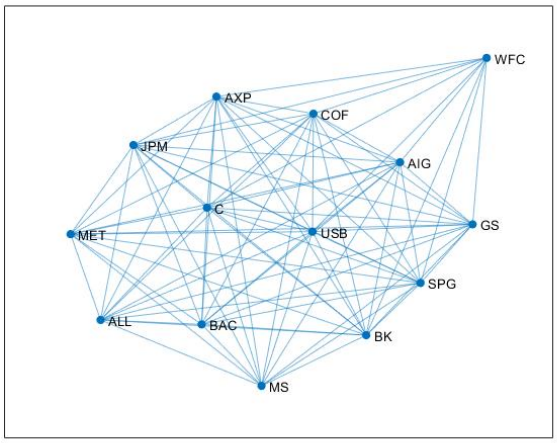
A second key finding is that the central nodes differ among the different components. This

outlines the fact that the different components lead to different patterns. This is a finding that has been also outlined in the literature using wavelet decomposition. A study in this direction was done by Caraiani (2015) who estimated a dynamic stochastic general equilibrium model along different components, obtaining different structural estimations for the different components.

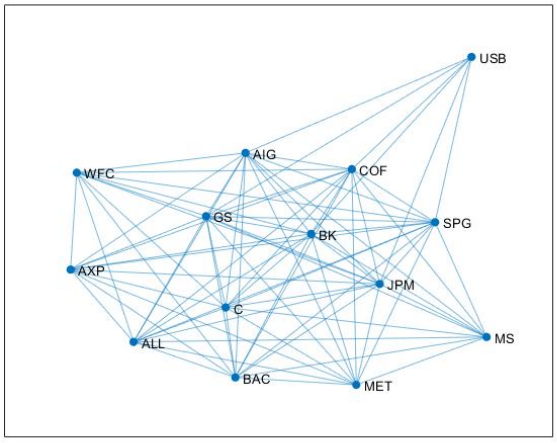
Finally, we can also see that, even for the contemporaneous network, the degree out-degree distribution has different patterns. While the aggregate series has the same number of degrees for each node, for the other components we can see some differences. Especially for the medium and short term components, we can see that there are more nodes having less than the maximum number of degrees.



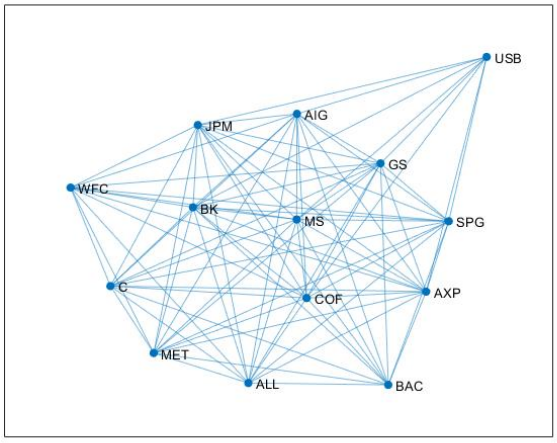
(a) Aggregate



(b) Long-run

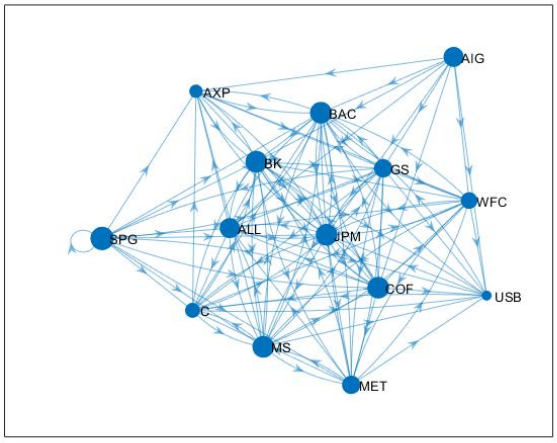


(c) Medium-run

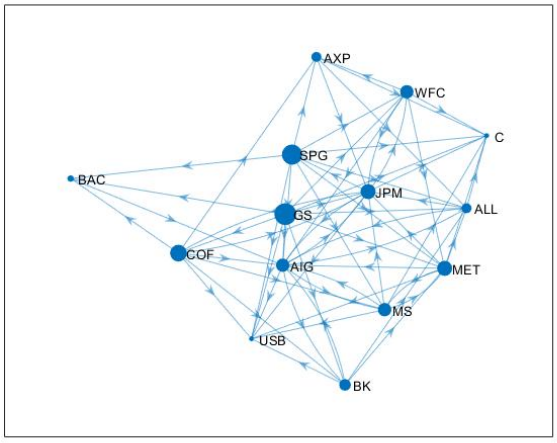


(d) Short-run

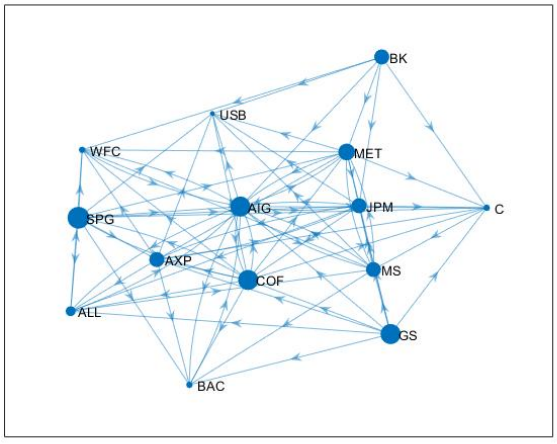
Figure 1: Contemporaneous networks



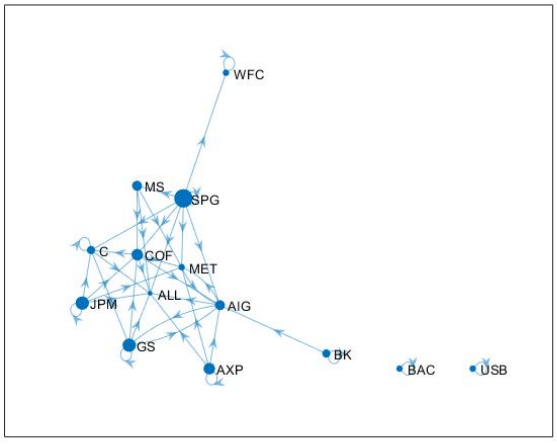
(a) Aggregate



(b) Long-run



(c) Medium-run



(d) Short-run

Figure 2: Granger-causality networks: a) aggregate series; b) long-term; c) medium-term; d) short-term. Notes: Each node stands for a company.

7 Conclusion

In this paper, we approached the topic of modelling connectedness in volatility networks using estimates of contemporaneous and Granger causality networks. While the main literature on connectedness, following Diebold and Yılmaz (2014a) or Billio et al. (2012) focused on a single type of network, we followed Barigozzi and Brownlees (2019) and estimated both contemporaneous and Granger causality networks. In contrast however to the former and latter, we performed the estimations also along different frequencies as obtained using EWD.

The research leads to several key findings. First, the central nodes differ between the networks based the actual financial series and those based on short-run, medium-run and long-run EWD details. In other words, the different frequencies have different central financial institutions as central.

Second, the sparsity of the networks is different along the frequencies. Particularly, the short-run networks, especially the Granger causality ones, are much sparse than the network corresponding to the long-run component or the actual series.

Third, our finding underscore that systemic risk (understood in the sense of contagion) is affected by the different scales, a finding not very well studied in the literature, see Baruník and Křehlík (2018) for a previous contribution in this literature.

References

Bandi, F.M., Chaudhuri, S.E., Lo, A.W., Tamoni, A., 2021. Spectral factor models. Journal of financial economics 142, 214–238.

Bandi, F.M., Perron, B., Tamoni, A., Tebaldi, C., 2019. The scale of predictability. Journal of econometrics 208, 120–140.

Barigozzi, M., Brownlees, C., 2019. NETS: Network estimation for time series. Journal of applied economics 34, 347–364.

Baruník, J., Křehlík, T., 2018. Measuring the frequency dynamics of financial connectedness and systemic risk. Journal of Financial Econometrics 16, 271–296.

Baruník, J., Vacha, L., 2023. The dynamic persistence of economic shocks.

Baruník, J., Křehlík, T., 2018. Measuring the Frequency Dynamics of Financial Connectedness and Systemic Risk*. Journal of Financial Econometrics 16, 271–296.

Billio, M., Getmansky, M., Lo, A.W., Pelizzon, L., 2012. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. Journal of Financial Economics 104, 535–559. Market Institutions, Financial Market Risks and Financial Crisis.

Caraiani, P., 2015. Estimating dsge models across time and frequency. Journal of Macroeconomics 44, 33–49.

Caraiani, P., 2017. Evaluating exchange rate forecasts along time and frequency. International Review of Economics Finance 51, 60–81.

Cerreia-Vioglio, S., Ortu, F., Severino, F., Tebaldi, C., 2023. Multivariate wold decompositions: a hilbert a-module approach. Decisions in Economics and Finance. A Journal of Applied Mathematics 46, 45–96.

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

Diebold, F.X., Yılmaz, K., 2014a. On the network topology of variance decompositions: Measuring the connectedness of financial firms. Journal of econometrics 182, 119–134.

Diebold, F.X., Yılmaz, K., 2014b. On the network topology of variance decompositions: Measuring the connectedness of financial firms. Journal of econometrics 182, 119–134.

Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., 2004. Least angle regression. The Annals of Statistics 32, 407–451.

Friedman, J., Hastie, T., Tibshirani, R., 2007. Sparse inverse covariance estimation with the graphical lasso. Biostatistics 9, 432–441.

Jackson, M.O., Pernoud, A., 2021. Systemic risk in financial networks: A survey. Annual Review of Economics 13, 171–202.

Ortu, F., Severino, F., Tamoni, A., Tebaldi, C., 2020. A persistence-based wold-type decomposition for stationary time series. Quantitative Economics 11, 203–230.

Teply, P., Kvapilikova, I., 2017. Measuring systemic risk of the us banking sector in time-frequency domain. The North American Journal of Economics and Finance 42, 461–472.

Yuan, M., Lin, Y., 2006. Model selection and estimation in regression with grouped variables. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 68, 49–67.

A Appendix

A.1 Networks for Time Series

Following Barigozzi and Brownlees (2019), consider a zero-mean stationary n-dimensional multivariate time series $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{nt})'$ generated by a p th-order VAR:

$$\boldsymbol{y}_t = \sum_{k=1}^{p} \boldsymbol{A}_k \boldsymbol{y}_{t-k} + \boldsymbol{\epsilon}_t \quad \boldsymbol{\epsilon}_t \sim wn\left(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}\right) \tag{A.1}$$

where $\boldsymbol{A}_k$ and $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ are $n \times n$ matrices. Throughout, the VAR is assumed to be stable and $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ to be positive definite. The long-run concentration matrix of the VAR approximation is

$$\begin{aligned} \boldsymbol{K}_L &= \left(\boldsymbol{I}_{np} - \sum_{k=1}^{p} \boldsymbol{A}_k'\right) \boldsymbol{\Sigma}_{\epsilon}^{-1} \left(\boldsymbol{I}_{np} - \sum_{k=1}^{p} \boldsymbol{A}_k\right) \\ &= \left(\boldsymbol{I} - \boldsymbol{G}'\right) \boldsymbol{C} (\boldsymbol{I} - \boldsymbol{G}) \end{aligned} \tag{A.2}$$

where $\boldsymbol{G} = \sum_{k=1}^{p} \boldsymbol{A}_k$ - as in Granger, and $\boldsymbol{C} = \boldsymbol{\Sigma}_{\epsilon}^{-1}-$ as in Contemporaneous.[3]

We work under the assumption that the VAR approximation is sparse. This, in turns, determines the sparsity of $\boldsymbol{G}, \boldsymbol{C}$ and $\boldsymbol{K}_L$. The matrix $\boldsymbol{G}$ can be associated to a long run Granger network (directed) expressing long predictive relations of the system a and the matrix $\boldsymbol{C}$ can be associated to a Contemporaneous partial correlation network of the system innovations. The Long Run Partial Correlation network is a (nontrivial) combination of the Granger and Contemporaneous networks.

A.2 Scale-decomposition on VAR(1) processes

Example VAR(1): To discuss this issue, assume for simplicity that the vector autoregressive process $\boldsymbol{y}_t$ is of order 1 and denote $A := A_1$ i.e.,

$$\boldsymbol{y}_t = \sum_{k=0}^{+\infty} A^k \boldsymbol{\epsilon}_{t-k} \tag{A.3}$$

Drawing a parallel to the earlier notations used to introduce the EWD, for VAR(1) process we have $\boldsymbol{\alpha}_k = A^k$. We can now obtain the extended Wold representation for a VAR(1) process

[3] $\boldsymbol{C} = \boldsymbol{\Sigma}_{\epsilon}^{-1}-$ is also known as concentration matrix of the innovations.

with the following coefficients $\Psi_k^{(j)}$:

$$\begin{aligned}
\boldsymbol{\Psi}_0^{(1)} &= \frac{I_2 - A}{\sqrt{2}} \\
\boldsymbol{\Psi}_1^{(1)} &= \frac{A^2 - A^3}{\sqrt{2}} = \frac{I_2 - A}{\sqrt{2}} A^2, \\
&\cdots \\
\boldsymbol{\Psi}_k^{(1)} &= \frac{A^{2k} - A^{2k+1}}{\sqrt{2}} = \frac{I_2 - A}{\sqrt{2}} A^{2k},
\end{aligned}$$

for $j = 1$. Similarly, for a generic $j = 1, 2, \ldots$, and for $k = 0$:

$$\begin{aligned}
\boldsymbol{\Psi}_0^{(j)} =& \frac{\underbrace{I_2 + A + \cdots + A^{2^{(j-1)}-1}}_{2^{(j-1)} \text{ terms}} - (\underbrace{A^{2^{(j-1)}} + \cdots + A^{2^j-1}}_{2^{(j-1)} \text{ terms}})}{\sqrt{2^j}} \\
=& \frac{\left(I_2 - A^{2^{(j-1)}}\right)(I_2 - A)^{-1} - \left(I_2 - A^{2^{(j-1)}}\right)(I_2 - A)^{-1} A^{2^{(j-1)}}}{\sqrt{2^j}} \\
=& \frac{\left(I_2 - A^{2^{(j-1)}}\right)^2 (I_2 - A)^{-1}}{\sqrt{2^j}}.
\end{aligned}$$

For $j = 1, 2, \ldots$, and for $k > 0$:

$$\begin{aligned}
\boldsymbol{\Psi}_k^{(j)} =& \frac{\underbrace{I_2 + A + \cdots + A^{2^{(j-1)}-1}}_{2^{(j-1)} \text{ terms}} - (\underbrace{A^{2^{(j-1)}+\cdots+A^{2^j-1}}}_{2^{(j-1)} \text{ terms}})}{\sqrt{2^j}} \times A^{k2^j} \\
=& \frac{\left(I_2 - A^{2^{(j-1)}}\right)^2 (I_2 - A)^{-1}}{\sqrt{2^j}} \times A^{k2^j}.
\end{aligned}$$

Notice in the derivations above that $\boldsymbol{\Psi}_k^{(j)} = \boldsymbol{\Psi}_0^{(j)} \times A^{k2^j}$. Hence, the structure of the coefficients $\boldsymbol{\Psi}_k^{(j)}$ is that of a VAR(1) defined on the support $S_t^{(j)} = \{t - k2^j : k \in \mathbb{Z}\}$. If we define $\widehat{\boldsymbol{\varepsilon}}^{(j)} = \boldsymbol{\Psi}_0^{(j)} \boldsymbol{\varepsilon}^{(j)}$, we can define the scale specific components in the following way

$$\boldsymbol{y}_t^{(j)} = \sum_{k=0}^{+\infty} \boldsymbol{\Psi}_k^{(j)} \boldsymbol{\varepsilon}_{t-k2^j}^{(j)} = \widehat{\boldsymbol{\varepsilon}}_t^{(j)} + A^{2^j} \widehat{\boldsymbol{\varepsilon}}_{t-2^j}^{(j)} + A^{2\times 2^j} \widehat{\boldsymbol{\varepsilon}}_{t-2\times 2^j}^{(j)} + \cdots \tag{A.4}$$

After some calculations, the previous equation simplifies to

$$\boldsymbol{y}_t^{(j)} = A^{2^j} \boldsymbol{y}_{t-2^j}^{(j)} + \widehat{\boldsymbol{\varepsilon}}_t^{(j)}, \quad \widehat{\boldsymbol{\varepsilon}}_t^{(j)} \sim wn\left(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}^{j}\right), \tag{A.5}$$

with $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\varepsilon}}^{(j)} = \mathrm{E}[\widehat{\boldsymbol{\varepsilon}}^{(j)}(\widehat{\boldsymbol{\varepsilon}}^{(j)})'] = \mathrm{E}[\boldsymbol{\Psi}_0^{(j)}\boldsymbol{\varepsilon}^{(j)}(\boldsymbol{\Psi}_0^{(j)}\boldsymbol{\varepsilon}^{(j)})']$.

The network representation of the scale-specific system is

$$\begin{aligned} \boldsymbol{K}_L^{(j)} &= \left(\boldsymbol{I} - \sum_{k=1}^{p_j} (\boldsymbol{A}_k^{(j)})'\right) (\Sigma_\epsilon^{(j)})^{-1} \left(\boldsymbol{I} - \sum_{k=1}^{p_j} \boldsymbol{A}_k^{(j)}\right) \\ &= \left(\boldsymbol{I} - (\boldsymbol{G}^{(j)})'\right) \boldsymbol{C}^{(j)} (\boldsymbol{I} - \boldsymbol{G}^{(j)}) \end{aligned} \tag{A.6}$$

Stock	Name	Sector	Sample
AIG	American International Group	Financials	2004/01/02-2015/12/13
ALL	Allstate Corp	Financials	2004/01/02-2015/12/13
AXP	American Express Co	Financials	2004/01/02-2015/12/13
BAC	Bank of America Corp	Financials	2004/01/02-2015/12/13
BK	The Bank of New York Mellon Corp.	Financials	2004/01/02-2015/12/13
C	Citigroup Inc.	Financials	2004/01/02-2015/12/13
COF	Capital One Financial	Financials	2004/01/02-2015/12/13
GS	Goldman Sachs Group	Financials	2004/01/02-2015/12/13
JPM	JPMorgan Chase & Co.	Financials	2004/01/02-2015/12/13
MET	MetLife Inc.	Financials	2004/01/02-2015/12/13
MS	Morgan Stanley	Financials	2004/01/02-2015/12/13
SPG	Simon Property Group Inc	Financials	2004/01/02-2015/12/13
USB	U.S. Bancorp	Financials	2004/01/02-2015/12/13
WFC	Wells Fargo	Financials	2004/01/02-2015/12/13

Table 1: Data Sample

A.3 Financial companies

A.4 Degree Distribution

Figure 3: Contemporaneous Network: long-term component

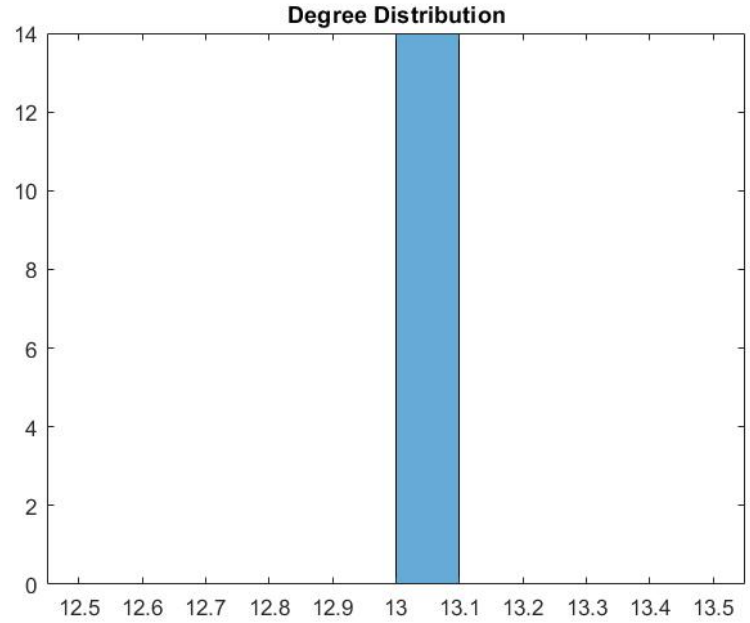


Figure 4: Granger Network: long-term component

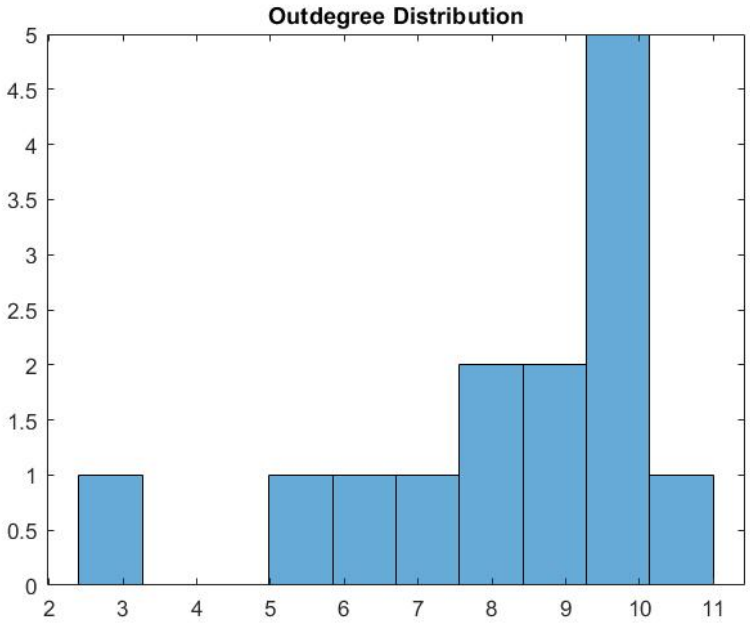


Figure 5: Contemporaneous Network: long-term component

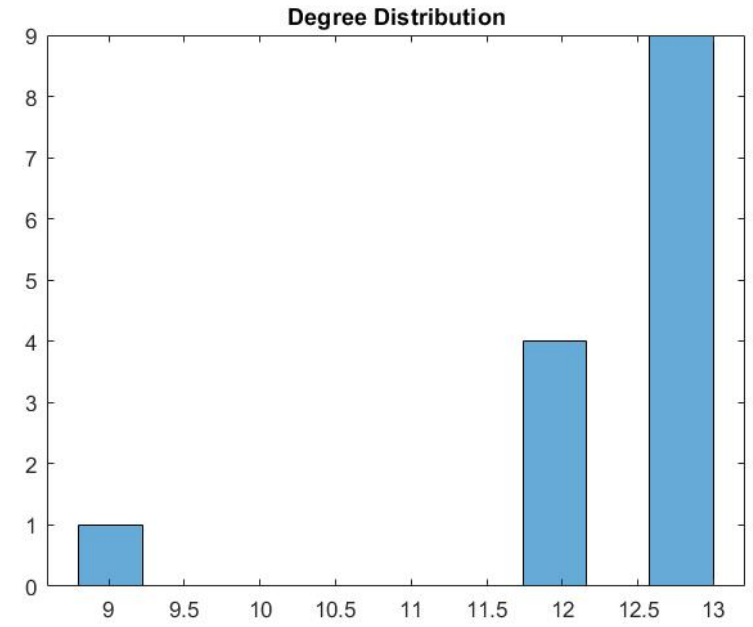


Figure 6: Granger Network: long-term component

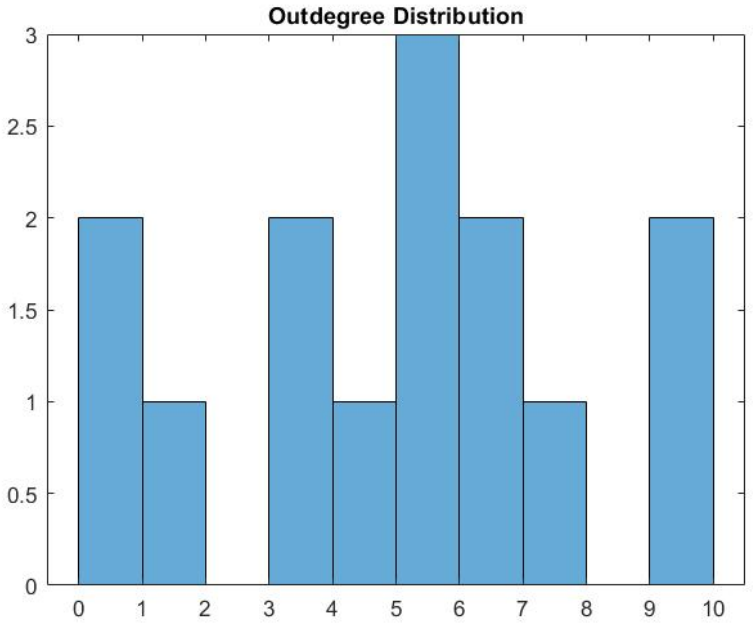


Figure 7: Contemporaneous Network: medium-term component

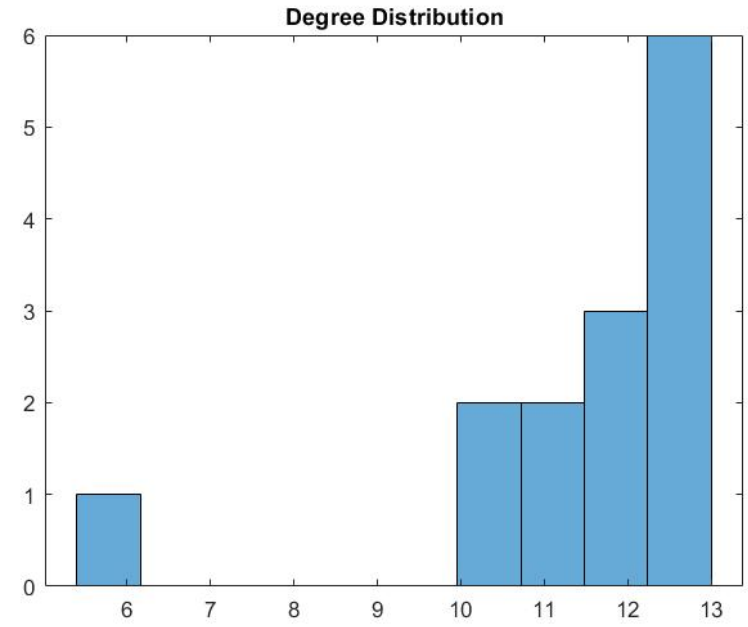


Figure 8: Granger Network: medium-term component

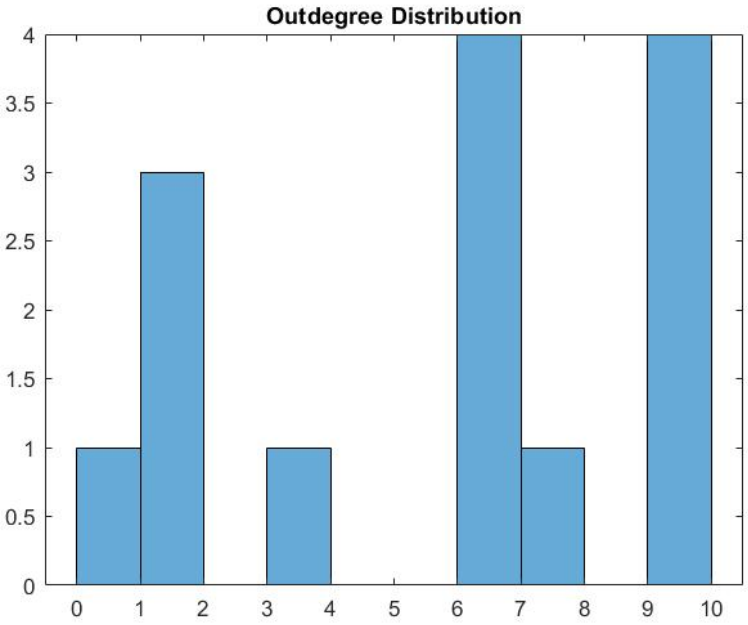


Figure 9: Contemporaneous Network: short-term component

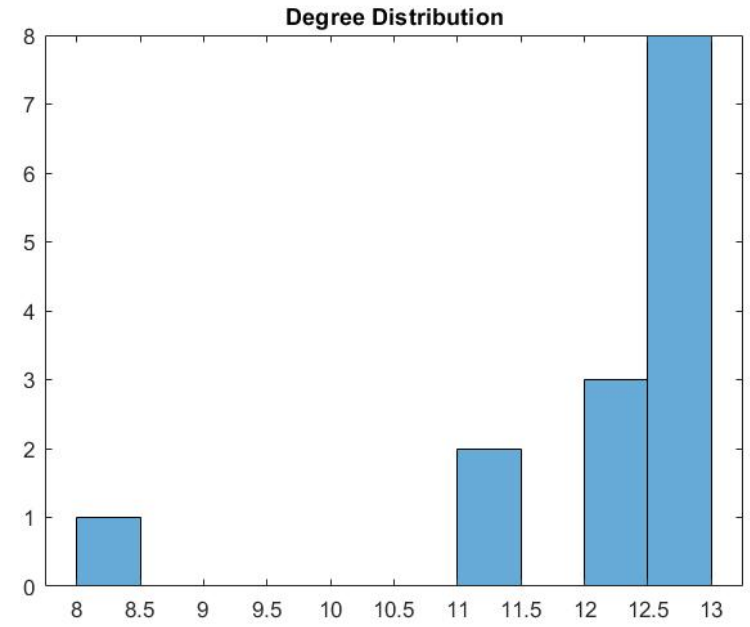
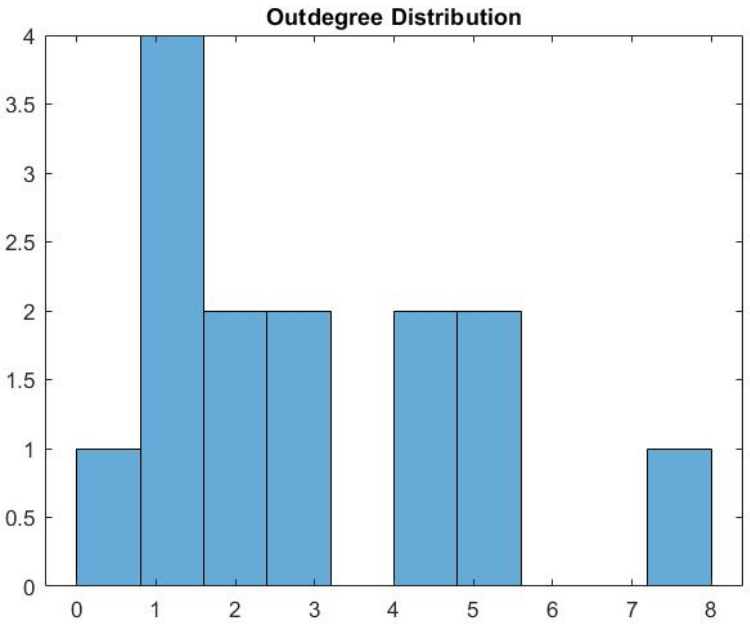


Figure 10: Granger Network: short-term component



A.5 Simulated Data

We simulate data for a sparse VAR using a degree of sparsity of 0.10, and an autocorrelation coefficient of $\rho = 0.75$. We simulate a VAR(p) model and obtain

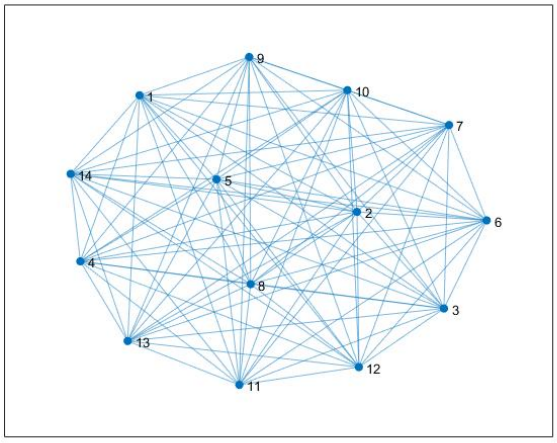
3021, that correspond to the number of observations for the actual financial series. The lag p in the VAR model is set to 1 (corresponding basically with the approach in Barigozzi and Brownlees (2019)

We simulate 14 series, each having the same number of observations as the real data. We feed the simulated VAR(1) model further in the estimation, for the aggregate series and after getting the decomposed EWD components (using the short-term, medium-term and long-term components).

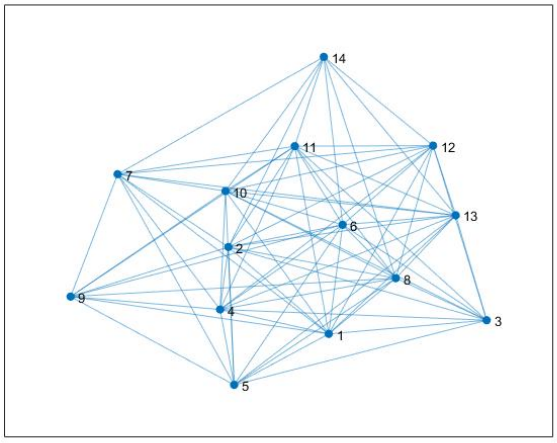
The estimation procedure is similar to the one for the actual series, see Estimation. λ is set using the BIC criterion.

We show below the contemporaneous network and the Granger causality network for the aggregate series, see Figure C.1 and C.2 below.

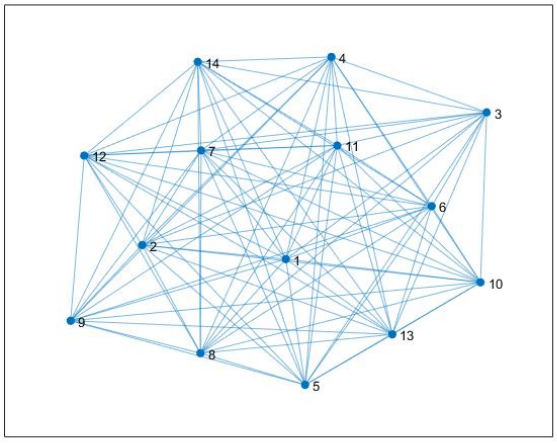
We also present the results for the short term component, see Figures C.1 and C.2.



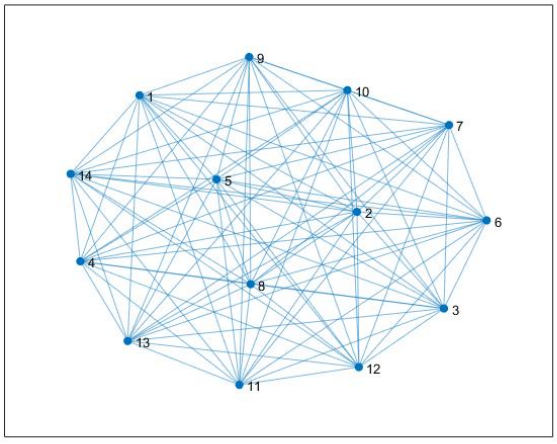
(a) Aggregate



(b) Long-run



(c) Medium-run

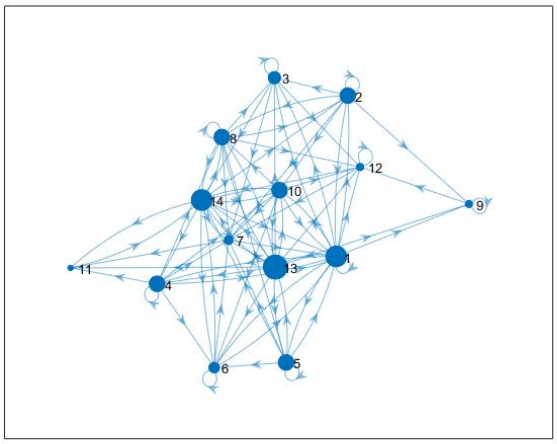


(d) Short-run

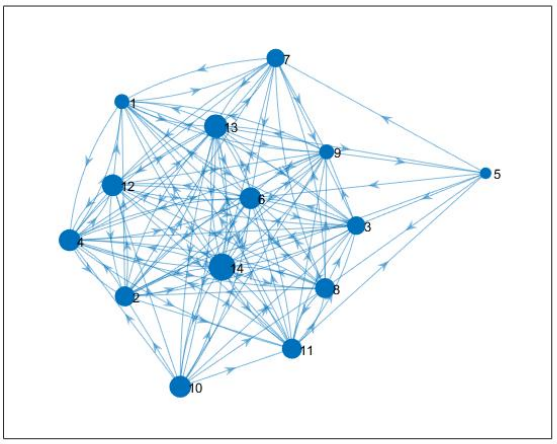
Figure 11: Contemporaneous networks: a) aggregate series; b) long-term; c) medium-term; d) short-term. Notes: Each node stands for a company.

The results suggest some differences between the short-run component and the aggregate series. First, for the contemporaneous matrix, the short-run component leads to sligthly different central nodes.

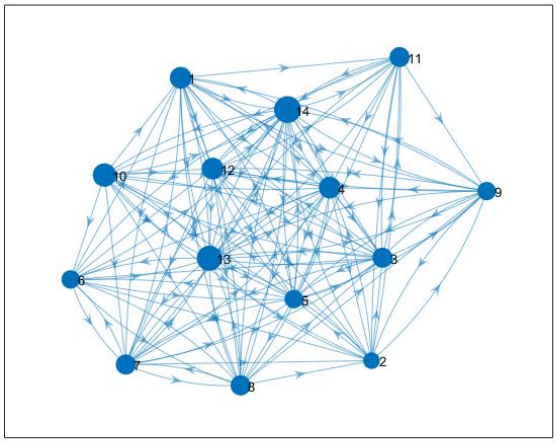
Furthermore, the differences are more significant for the case of the Granger causality network. Here, Figure 4 shows a much sparser network, while also indicating a different subset of central nodes.



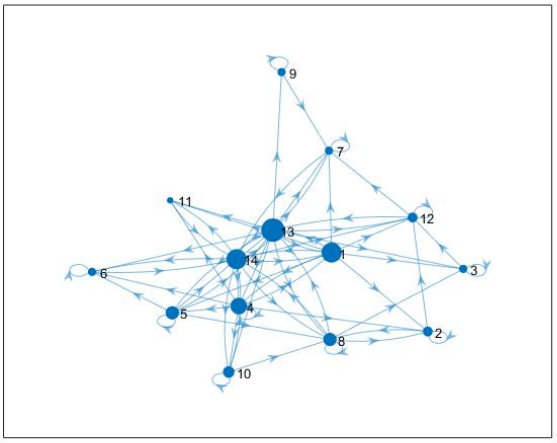
(a) Aggregate



(b) Long-run



(c) Medium-run



(d) Short-run

Figure 12: Contemporaneous networks: a) aggregate series; b) long-term; c) medium-term; d) short-term. Notes: Each node stands for a company.