

CORVINUS UNIVERSITY OF BUDAPEST

EÖTVÖS LORÁND UNIVERSITY

**A NEW APPROACH TO MEASURE CAUSAL
NETWORK CONNECTEDNESS**

MSc Thesis

Márton Espán

MSc in Actuarial and Financial Mathematics
Quantitative Finance Major

Supervisor:

Milán Csaba Badics



Budapest, 2025

Acknowledgements

I would like to express my gratitude first and foremost to Milán Badics, who supported my work with numerous articles and ideas, in addition to the regular weekly consultations, and kept me motivated throughout the process. I am also deeply thankful to everyone who, over the past two years, helped me stay strong and supported me in completing my studies. Furthermore, I would like to thank Dávid Apagyi for his insights and advice regarding the use of the LaTeX typesetting system.

Contents

- 1 Introduction** **4**

- 2 Methodology** **6**
 - 2.1 The Diebold–Yilmaz framework 6
 - 2.2 The IRF network 9
 - 2.2.1 Decomposition of the $IRF^{(H)}$ matrix. 11
 - 2.2.2 A toy example 14
 - 2.3 Comparing Diebold–Yilmaz and IRF-based Networks 17
 - 2.3.1 Contemporaneous-only effects 18
 - 2.3.2 Lagged-only effects 21
 - 2.3.3 Lagged and contamporenaous effects as well 25

- 3 Empirical analysis on finacial data** **29**
 - 3.1 Rolling network 34
 - 3.2 Robustness check 35

- 4 Conclusion** **37**

- Appendix** **38**

Chapter 1

Introduction

A key challenge in macroeconomics and finance is to understand the propagation of shocks across interconnected systems. In recent years, network-based measures of connectedness have emerged as powerful tools to quantify spillover effects and interdependencies among variables. One of the most prominent approaches in this literature is the framework introduced by Diebold and Yilmaz (Diebold and Yilmaz 2009; Diebold and Yilmaz 2012), which constructs networks based on forecast error variance decompositions derived from vector autoregressions. While this method has been widely applied and extended, it relies on generalized impulse response functions, which do not depend on the causal ordering of variables. As a result, the resulting networks often lack structural interpretability.

This paper proposes a new approach to connectedness measurement based on structural impulse response functions. By recovering the contemporaneous causal structure of the system using the LiNGAM algorithm (Shimizu et al. 2006), we identify a structural VAR (SVAR) model and derive a sequence of IRFs that fully reflects the dynamics and directionality of shock transmission. Unlike the GIRF-based method, our framework allows for a clear distinction between direct and indirect, as well as contemporaneous and lagged effects, offering a richer and more interpretable view of connectedness.

We construct a cumulative impulse response matrix and normalize it in a way that preserves the relative magnitudes of off-diagonal interactions, thereby avoiding the loss of interpretive information often associated with row-wise normalization in the DY framework. Based on this matrix, we define directional measures such as TO, FROM, NET, and the Total Connectedness Index, analogous to the DY measures but grounded in structural dynamics.

Through a series of simulated experiments, we compare our IRF-based network with

the Diebold–Yilmaz network under various data-generating processes. The simulations demonstrate the advantages of our method in capturing the correct causal structure and spillover channels. We also apply the proposed framework to the `dy2012` dataset (Diebold and Yilmaz 2012), where we recover economically plausible causal relations among key asset classes and find that our IRF-based network consistently produces higher and more informative connectedness measures.

The remainder of the paper is structured as follows. Section 2 reviews the Diebold–Yilmaz framework, introduces our IRF-based network methodology and also presents simulated examples highlighting the strengths of the proposed approach. Section 3 provides an empirical analysis using financial market data while Section 5 concludes.

The full replication code, including all simulations and empirical analyses, is available at the accompanying GitHub repository.¹

¹github.com/Espanm/R2DAG/blob/master/TDK

Chapter 2

Methodology

In this section, we first summarize the Diebold–Yilmaz (DY) connectedness framework as a benchmark approach. Then, we introduce our novel network construction based on structural impulse response functions, followed by a comparison between the two methodologies.

2.1 The Diebold–Yilmaz framework

The connectedness framework developed by Diebold and Yilmaz (2009) and extended by Diebold and Yilmaz (2012) is based on forecast error variance decompositions derived from a reduced-form Vector Autoregression (VAR) model. It quantifies how much of the forecast error variance of one variable can be attributed to shocks in another, resulting in a directed, weighted network of spillovers.

We consider a stationary VAR(p) process of dimension N :

$$y_t = \sum_{i=1}^p \Phi_i y_{t-i} + u_t, \quad u_t \sim (0, \Sigma_u)$$

This system admits a moving average representation:

$$y_t = \sum_{i=0}^{\infty} A_i u_{t-i}$$

Following Diebold and Yilmaz (2009) and Diebold and Yilmaz (2012), the generalized forecast error variance decomposition (GFEVD) is defined as:

$$\theta_{ij}^{(H)} = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' A_h \Sigma_u e_j)^2}{\sum_{h=0}^{H-1} (e_i' A_h \Sigma_u A_h' e_i)}$$

where σ_{jj} is the variance of shock j , e_i is the selection vector for variable i , and H denotes the forecast horizon.

The normalized matrix is then:

$$\tilde{\theta}_{ij}^{(H)} = \frac{\theta_{ij}^{(H)}}{\sum_{j=1}^N \theta_{ij}^{(H)}}, \quad \sum_{j=1}^N \tilde{\theta}_{ij}^{(H)} = 1$$

From this, directional connectedness measures are computed as:

$$\begin{aligned} TO_i &= \sum_{j \neq i} \tilde{\theta}_{ij}^{(H)} \\ FROM_i &= \sum_{j \neq i} \tilde{\theta}_{ji}^{(H)} \\ NET_i &= TO_i - FROM_i \\ TCI &= \frac{1}{N} \sum_{i=1}^N \sum_{j \neq i} \tilde{\theta}_{ij}^{(H)} \\ NPDC_{ij} &= \tilde{\theta}_{ij}^{(H)} - \tilde{\theta}_{ji}^{(H)} \end{aligned}$$

The matrix $\tilde{\theta}^{(H)}$ serves as the basis for constructing a directed network, where nodes represent variables and edges capture the forecast-based spillovers. The TCI summarizes the overall level of connectedness in the system, while the directional and net measures allow detailed analysis of transmission and reception of shocks among variables.

Algorithm 1 Diebold–Yilmaz connectedness measures

Require: Time series data $\{y_t\}_{t=1}^T$ with N variables, forecast horizon H

Ensure: Generalized FEVD matrix $\tilde{\Theta}^{(H)}$, and connectedness measures

- 1: Estimate a VAR(p) model on $\{y_t\}$
- 2: Compute moving average coefficients $\{A_h\}_{h=0}^{H-1}$
- 3: Compute residual covariance matrix Σ_u
- 4: **for** $i = 1$ to N **do**
- 5: **for** $j = 1$ to N **do**
- 6: Compute GFEVD entry $\theta_{ij}^{(H)}$ using:

$$\theta_{ij}^{(H)} = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' A_h \Sigma_u e_j)^2}{\sum_{h=0}^{H-1} (e_i' A_h \Sigma_u A_h' e_i)}$$

- 7: **end for**
- 8: **end for**
- 9: Normalize rows:

$$\tilde{\theta}_{ij}^{(H)} = \frac{\theta_{ij}^{(H)}}{\sum_{j=1}^N \theta_{ij}^{(H)}}$$

- 10: Calculate TO, FROM, NET, NPDC and TCI values
-

2.2 The IRF network

In this subsection, we introduce our proposed method for constructing a network based on structural impulse response functions.

We assume that the observed data is generated by a structural vector autoregression (SVAR) process of order p , defined as:

$$A_0 Y_t = \sum_{i=1}^p A_i Y_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim (0, I)$$

where Y_t is an $n \times 1$ vector of endogenous variables, A_0 is an $n \times n$ contemporaneous impact matrix, A_i are structural lag coefficient matrices, and ε_t is a vector of structural shocks that are assumed to be mutually uncorrelated and standardized.

Although the SVAR model provides a causal interpretation of shock transmission, it cannot be directly estimated from the data. This is because the model is not identified without additional assumptions regarding the contemporaneous causal ordering of the variables. Specifically, identifying the structural impact matrix A_0 requires knowledge of the underlying causal structure, which is not observable.

So, we begin by considering a reduced-form Vector Autoregression (VAR) model of order p for an n -dimensional vector of endogenous variables:

$$Y_t = \sum_{i=1}^p B_i Y_{t-i} + u_t, \quad u_t \sim (0, \Sigma_u)$$

Here, Y_t is an $n \times 1$ vector of jointly determined time series variables, B_i are $n \times n$ coefficient matrices corresponding to lag i , and u_t is an error term with covariance matrix Σ_u .

The reduced-form VAR parameters are related to the structural SVAR parameters through the following transformation:

$$B_i = A_0^{-1} A_i \quad \text{and} \quad \Sigma_u = A_0^{-1} (A_0^{-1})'$$

To identify the contemporaneous causal ordering required for estimating the structural matrix A_0 , we employ the Linear Non-Gaussian Acyclic Model (LiNGAM) algorithm proposed by Shimizu et al. (2006). LiNGAM provides a data-driven approach to uncovering directed acyclic graphs (DAGs) based on the non-Gaussianity of structural shocks.

This methodology is particularly well-suited for financial and economic data, which often deviate from the assumptions of classical linear models. As noted in the literature,

residuals from linear regressions fitted to financial time series typically exhibit properties such as heavy tails, skewness, excess kurtosis, and volatility clustering (Cont 2001; Nolan and Ojeda-Revah 2013). These features violate the Gaussian assumptions underpinning traditional identification schemes, rendering them unreliable in such contexts.

LiNGAM exploits these violations by leveraging the statistical independence and non-Gaussian distribution of structural errors to recover the underlying causal structure. Unlike conventional approaches, it does not require a priori ordering or exclusion restrictions; instead, it derives the causal ordering directly from the data.

Once the contemporaneous causal ordering of the variables has been recovered using the LiNGAM algorithm, we can use this information to estimate the structural impact matrix A_0 . The identified ordering allows us to impose a recursive structure—typically a (possibly permuted) lower triangular form—on A_0 , enabling full identification of the SVAR model.

Given the identified structural matrices A_0 and $\{A_i\}_{i=1}^p$, we can compute the structural impulse response functions (IRFs), which trace the dynamic effects of one-unit structural shocks on the system over time. Following standard SVAR methodology (Lütkepohl 2005, Chapter 11), the IRFs are computed recursively as follows:

At horizon $h = 0$, the immediate response is:

$$IRF_0 = A_0^{-1}$$

For $h \geq 1$, the responses are given by:

$$IRF_h = \left(\sum_{j=1}^{\min(h,p)} A_0^{-1} A_j \cdot IRF_{h-j} \right) A_0^{-1}$$

This recursive formulation captures how structural shocks propagate through the system via both contemporaneous interactions (encoded in A_0) and lagged dynamics (captured by $\{A_j\}$). The resulting sequence $\{IRF_h\}_{h=0}^H$ forms the foundation of our impulse response-based network, where each IRF_h matrix quantifies the effect of shocks at time t on outcomes at time $t + h$.

Analogously to the $\theta^{(H)}$ spillover table used in the Diebold–Yilmaz framework, our approach constructs an impulse response-based matrix denoted by $IRF^{(H)}$. This matrix aggregates the dynamic effects of shocks across time and is computed as the cumulative sum of the structural impulse response functions up to horizon H :

$$IRF^{(H)} = \sum_{h=0}^H |IRF_h|$$

To obtain a comparable and interpretable representation, we normalize $IRF^{(H)}$ by dividing all its elements by the maximum row sum of the matrix. The resulting normalized matrix is denoted by $\widetilde{IRF}^{(H)}$:

$$\widetilde{IRF}^{(H)} = \frac{IRF^{(H)}}{\max_i \sum_{j=1}^n IRF_{ij}^{(H)}}$$

This row-wise normalization ensures that each row sum lies between 0 and 1, while preserving the relative magnitudes of the off-diagonal elements within each row.

Following the approach proposed in Caloia, Cipollini, and Muzzioli (2018), we define the following connectedness measures based on the normalized matrix $\widetilde{IRF}^{(H)}$:

$$TO_i = \frac{\sum_{j \neq i} \widetilde{IRF}_{ij}^{(H)}}{\sum_{j=1}^n \widetilde{IRF}_{ij}^{(H)}}$$

$$FROM_i = \frac{\sum_{j \neq i} \widetilde{IRF}_{ji}^{(H)}}{\sum_{j=1}^n \widetilde{IRF}_{ji}^{(H)}}$$

$$NET_i = TO_i - FROM_i$$

$$NPDC_{ij} = \widetilde{IRF}_{ij}^{(H)} - \widetilde{IRF}_{ji}^{(H)}$$

$$TCI = \frac{1}{n} \sum_{i=1}^n FROM_i$$

2.2.1 Decomposition of the $IRF^{(H)}$ matrix.

An important property of the $IRF^{(H)}$ matrix—and consequently of its normalized form $\widetilde{IRF}^{(H)}$ —is that it can be decomposed into the sum of four interpretable components:

- C_d : Contemporaneous direct effects,
- C_i : Contemporaneous indirect effects,

- L_d : Lagged direct effects,
- L_i : Lagged indirect effects.

The contemporaneous direct effects (C_d) are given by the absolute value of A_0^{-1} , as these represent the immediate structural responses due to direct connections among variables. All other effects present in IRF_0 are classified as contemporaneous indirect effects (C_i), since they emerge through intermediary pathways in the recursive system.

Among the lagged effects, IRF_1 captures the first-order dynamic propagation and is thus categorized as the lagged direct component (L_d). All remaining matrices in the cumulative response, i.e., IRF_h for $h \geq 2$, are considered lagged indirect effects (L_i), reflecting multi-step propagation through the system.

The full decomposition is:

$$IRF^{(H)} = C_d + C_i + L_d + L_i$$

Table 2.1: Decomposition of $IRF^{(H)}$ into structural components

Component	Definition
C_d	$ A_0 $
C_i	$ IRF_0 - A_0 $
L_d	$ IRF_1 $
L_i	$\sum_{h=2}^H IRF_h $

Algorithm 2 Impulse response-based network construction

Require: Multivariate time series data $\{Y_t\}_{t=1}^T$, forecast horizon H

Ensure: Connectedness measures based on structural impulse responses

- 1: Estimate a reduced-form VAR(p) model on $\{Y_t\}$
- 2: Apply the LiNGAM algorithm to identify the contemporaneous causal ordering
- 3: Estimate the structural impact matrix A_0 based on the LiNGAM-implied ordering
- 4: Obtain the structural lag matrices $\{A_i\}_{i=1}^p$
- 5: Compute structural impulse response matrices:

$$IRF_0 = A_0^{-1}$$

$$IRF_h = \left(\sum_{j=1}^{\min(h,p)} A_0^{-1} A_j \cdot IRF_{h-j} \right) A_0^{-1}, \quad \text{for } h = 1, \dots, H$$

- 6: Construct the cumulative impulse response matrix:

$$IRF^{(H)} = \sum_{h=0}^H |IRF_h|$$

- 7: Normalize the matrix by dividing all entries by the maximum row sum:

$$\widetilde{IRF}^{(H)} = \frac{IRF^{(H)}}{\max_i \sum_{j=1}^n IRF_{ij}^{(H)}}$$

- 8: Calculate the FROM, TO, NET, NPDC and TCI values
 - 9: Decompose the $IRF^{(H)}$ to C_d, C_i, L_d and L_i
-

2.2.2 A toy example

To illustrate the construction of our impulse response-based network, we consider a simple three-dimensional SVAR(1) process with $H = 1$. This toy example allows us to demonstrate each computational step explicitly, from the structural specification to the reduced-form VAR and ultimately to the network representation.

The data-generating process is defined as:

$$A_0 Y_t = A_1 Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim t_{10}(0, \Sigma_\varepsilon)$$

Since a standard multivariate Student- t distribution with $\nu = 10$ degrees of freedom has covariance $\frac{10}{8}I = \frac{5}{4}I$, we rescale the shocks to ensure unit variance, thus

$$\Sigma_\varepsilon = I$$

The structural matrices are given by:

$$A_0 = \begin{pmatrix} 1 & 0 & 0 \\ -0.5 & 1 & 0 \\ 0 & -0.4 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0.7 & 0 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0 & 0.7 \end{pmatrix}$$

To derive the reduced-form VAR(1) model:

$$Y_t = B_1 Y_{t-1} + u_t, \quad u_t = A_0^{-1} \varepsilon_t, \quad u_t \sim (0, \Sigma_u)$$

we compute the inverse of A_0 :

$$A_0^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.2 & 0.4 & 1 \end{pmatrix}$$

Then the estimated reduced-form coefficient matrix becomes:

$$B_1 = A_0^{-1} A_1 = \begin{pmatrix} 0.7 & 0 & 0 \\ 0.35 & 0.7 & 0 \\ 0.14 & 0.28 & 0.7 \end{pmatrix}$$

The reduced-form error covariance matrix is:

$$\Sigma_u = A_0^{-1}(A_0^{-1})' = \begin{pmatrix} 1 & 0.5 & 0.2 \\ 0.5 & 1.25 & 0.5 \\ 0.2 & 0.5 & 1.2 \end{pmatrix}$$

We obtain the following adjacency matrix from the LiNGAM algorithm:

$$\text{Adj} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Hence, the causal ordering among the variables is:

$$Y_1 \prec Y_2 \prec Y_3$$

To ensure that the structural shocks are orthogonal, we require the structural impact matrix A_0 to satisfy the following condition:

$$A_0 \Sigma_u A_0' = D$$

where D is a diagonal matrix. Since Σ_u is symmetric, it has 6 unique elements, of which 3 off-diagonal constraints must be zero to ensure orthogonal structural shocks. To match these, we require at least 3 free parameters in A_0 , which—given the LiNGAM ordering $Y_1 \prec Y_2 \prec Y_3$ —we place in the lower triangular part of a recursive structure:

$$A_0 = \begin{pmatrix} 1 & 0 & 0 \\ a_{21} & 1 & 0 \\ a_{31} & a_{32} & 1 \end{pmatrix}$$

Substituting into the condition $A_0 \Sigma_u A_0' = D$, we impose that all off-diagonal elements of the resulting product vanish. This leads to the following system of equations:

$$\begin{aligned} a_{21} + 0.5 &= 0 \\ a_{31} + 0.5a_{32} + 0.2 &= 0 \\ a_{21}(a_{31} + 0.5a_{32} + 0.2) + 0.5a_{31} + 1.25a_{32} + 0.5 &= 0 \end{aligned}$$

Solving this system yields:

$$a_{21} = -0.5, \quad a_{31} = 0, \quad a_{32} = -0.4$$

Therefore, the fully identified A_0 matrix is:

$$A_0 = \begin{pmatrix} 1 & 0 & 0 \\ -0.5 & 1 & 0 \\ 0 & -0.4 & 1 \end{pmatrix}$$

Having completed the identification, we can now compute the first two structural impulse response matrices.

The immediate impact matrix is:

$$IRF_0 = A_0^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.2 & 0.4 & 1 \end{pmatrix}$$

The one-step response is given by:

$$IRF_1 = A_0^{-1}A_1A_0^{-1} = \begin{pmatrix} 0.7 & 0 & 0 \\ 0.7 & 0.7 & 0 \\ 0.42 & 0.56 & 0.7 \end{pmatrix}$$

Summing the absolute values elementwise:

$$IRF^{(1)} = |IRF_0| + |IRF_1| = \begin{pmatrix} 1.7 & 0 & 0 \\ 1.2 & 1.7 & 0 \\ 0.62 & 0.96 & 1.7 \end{pmatrix}$$

The maximum row sum of the cumulative absolute IRF matrix is:

$$\max_i \sum_j (IRF^{(1)})_{ij} = 3.28$$

Normalizing by this value, we obtain the scaled matrix:

$$\widetilde{IRF}^{(1)} = \begin{pmatrix} 0.5183 & 0 & 0 \\ 0.3659 & 0.5183 & 0 \\ 0.1890 & 0.2927 & 0.5183 \end{pmatrix}$$

From this normalized matrix, the *TO*, *FROM*, *NET*, *NPDC*, and *TCI* measures can be easily derived using the definitions provided earlier.

The decomposition is as follows:

-

$$C_d = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0 & 0.4 & 1 \end{pmatrix}$$

$$C_i = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.2 & 0 & 0 \end{pmatrix}$$

$$L_d = \begin{pmatrix} 0.7 & 0 & 0 \\ 0.7 & 0.7 & 0 \\ 0.42 & 0.56 & 0.7 \end{pmatrix}$$

$L_i = 0$, since $H = 1$ and there are no responses beyond horizon 1.

The contemporaneous causal structure in this example is depicted in the following graph, which provides an intuitive basis for distinguishing between direct and indirect effects.

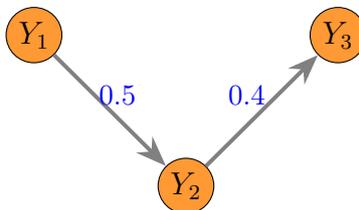


Figure 2.1: Contemporaneous structure

It is important to emphasize that these are theoretical values; in empirical applications, model estimates are inherently subject to approximation and are never perfectly accurate.

2.3 Comparing Diebold–Yilmaz and IRF-based Networks

To assess the practical implications of our proposed impulse response-based network framework, we compare it with the widely used Diebold–Yilmaz connectedness approach across several simulated four-dimensional data-generating processes.

While both methods rely on the estimation of a vector autoregressive (VAR) model, the underlying principles and assumptions differ significantly. The Diebold–Yilmaz framework builds on generalized impulse response functions (GIRFs), which, due to their invariant ordering property, sacrifice interpretability. In contrast, structural impulse response functions (IRFs) provide more tangible and interpretable dynamics, especially when the contemporaneous causal structure is known.

Thanks to the LiNGAM algorithm, we are able to identify the contemporaneous causal ordering, which allows for a specification of the structural system. This additional layer of information enables us to distinguish between direct and indirect, contemporaneous and lagged effects in a meaningful way.

Moreover, unlike the Diebold–Yilmaz approach—which applies normalization row-wise—the IRF-based normalization is performed using the maximum row sum of the cumulative impulse response matrix. As a result, the relative magnitudes of the matrix elements remain intact, preserving the internal structure of directional influences in the network.

2.3.1 Contemporaneous-only effects

To begin our comparison, we consider a simple setting in which the data-generating process includes only contemporaneous effects, with no lagged interactions. This setup isolates the role of contemporaneous structural dependence in determining connectedness.

The structural model is specified as:

$$A_0 Y_t = \varepsilon_t, \quad \varepsilon_t \sim t_{10}(0, I)$$

where A_0 is a lower triangular matrix encoding the contemporaneous causal structure:

$$A_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -0.5 & -0.5 & 1 & 0 \\ -0.5 & 0 & -0.5 & 1 \end{pmatrix}$$

The innovations ε_t follow a multivariate Student- t distribution with 10 degrees of freedom and identity covariance matrix. The sample size is 1000, and the horizon is set to $H = 5$.

The resulting contemporaneous causal structure is illustrated in the following directed acyclic graph, where solid arrows indicate direct effects, dashed arrows represent indirect ones.

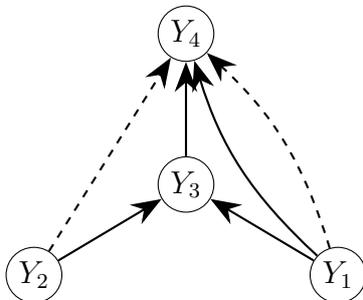


Figure 2.2: Causal DAG implied by A_0 .

The results of our impulse response-based network, applied to a structural system with contemporaneous effects only, are summarized in the table below. Each cell captures the relative strength of influence based on the normalized cumulative impulse response matrix $\widetilde{IRF}^{(5)}$. The final two rows report the total incoming connectedness (FROM) and the net directional connectedness (NET) for each variable, while the final column shows the total outgoing connectedness (TO). The Total Connectedness Index (TCI) is shown as percentage.

	Y1	Y2	Y3	Y4	FROM
Y1	0.4063	0.0061	0.0056	0.0034	0.0357
Y2	0.0131	0.4093	0.0190	0.0011	0.0750
Y3	0.2019	0.2044	0.4066	0.1870	0.5934
Y4	0.3119	0.1263	0.0006	0.4205	0.5106
TO	0.5646	0.4514	0.0583	0.3129	TCI
NET	0.5289	0.3764	-0.5351	-0.1977	30.36%

Table 2.2: Normalized IRF-based connectedness table for contemporaneous-only system

For comparison, the results of the Diebold–Yilmaz framework are shown below.

	Y1	Y2	Y3	Y4	FROM
Y1	0.6981	0.0005	0.1051	0.1963	0.3019
Y2	0.0005	0.8492	0.1148	0.0355	0.1508
Y3	0.0915	0.0814	0.6083	0.2188	0.3917
Y4	0.1679	0.0249	0.2135	0.5938	0.4062
TO	0.2600	0.1068	0.4334	0.4506	TCI
NET	-0.0420	-0.0441	0.0417	0.0443	31.27%

Table 2.3: Diebold–Yilmaz connectedness table for the contemporaneous-only system

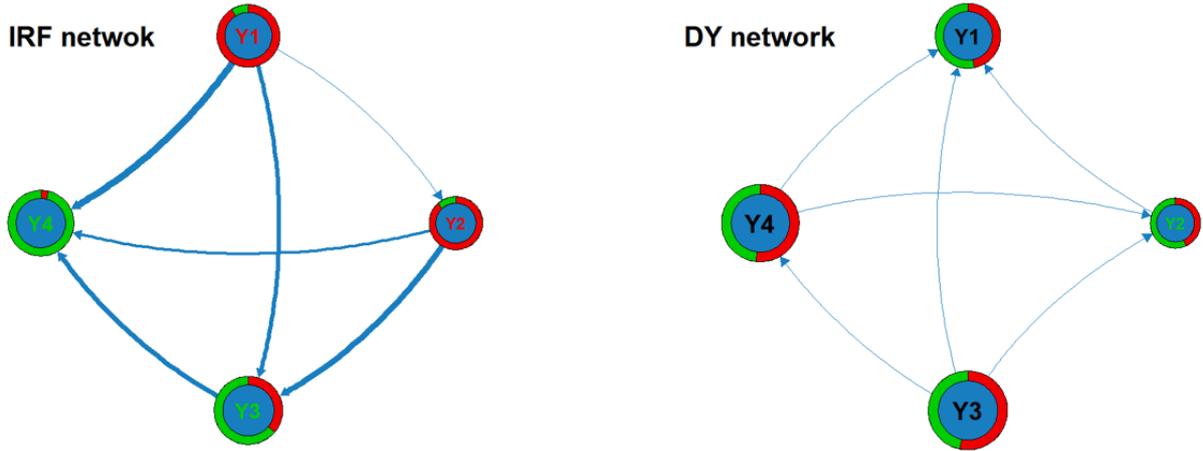


Figure 2.3: Visualization of the networks in case of contemporaneous-only effects. The size of each node reflects its FROM+TO value. In the pie charts, the green (red) segments correspond to the proportion of FROM (TO) spillovers within the total for that series. Arrows indicate the direction of net spillovers, with their thickness representing the magnitude. Nodes labeled in green denote time series with FROM/SUM ratios of at least 55%, while red labels indicate TO/SUM ratios of 55% or more.

A key distinction between the Diebold–Yilmaz and IRF-based networks lies in their treatment of structural information. The Diebold–Yilmaz framework does not incorporate the contemporaneous causal structure of the system; instead, it relies solely on the reduced-form covariance matrix of innovations. As a result, the derived connectedness matrix tends to be close to symmetric, and the net connectedness values (NET) for each variable typically hover around zero.

In contrast, the IRF-based network directly incorporates the causal ordering identified by the structural model. This leads to a more informative and interpretable structure: in the contemporaneous-only example, the IRF matrix clearly reflects that variables Y_1 and Y_2 act primarily as transmitters of shocks (positive NET), while Y_3 and Y_4 behave as receivers (negative NET). Despite this directional asymmetry, the Total Connectedness Index values in both approaches remain close in magnitude, indicating comparable overall spillover intensity.

These differences are also visually evident when comparing the two resulting network graphs side by side:

Since the structural system in this example contains no lagged dynamics, both the lagged direct (L_d) and lagged indirect (L_i) components of the cumulative impulse response matrix are zero by construction. As a result, the entire response is driven by contemporaneous effects, which we decompose into direct (C_d) and indirect (C_i) components.

The direct contemporaneous effects, captured by the absolute value of the structural impact matrix A_0 .

$$C_d = \begin{pmatrix} 1.0000 & 0 & 0 & 0 \\ 0.0232 & 1.0000 & 0 & 0 \\ 0.5106 & 0.4782 & 1.0000 & 0 \\ 0.4936 & 0.0701 & 0.4687 & 1.0000 \end{pmatrix}$$

The indirect contemporaneous effects are the following:

$$C_i = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -0.0111 & 0 & 0 & 0 \\ 0.2325 & 0.2241 & 0 & 0 \end{pmatrix}$$

Note that these are non-normalized values computed directly from the $IRF^{(5)}$ matrix. It is also important to emphasize that the entries $(C_d)_{21}$, $(C_d)_{42}$, and $(C_i)_{31}$ deviate slightly from zero solely due to estimation error.

2.3.2 Lagged-only effects

As a contrasting case, we now consider a system in which the structural matrix A_0 is set to the identity matrix. In this setting, the process corresponds to a reduced-form VAR

model rather than a structural VAR (SVAR), and no contemporaneous causal structure is imposed or recoverable.

Since the structural matrix is the identity, the model takes the form:

$$Y_t = A_1 Y_{t-1} + \varepsilon_t$$

with the lag matrix defined as:

$$A_1 = \begin{pmatrix} 0.5 & 0 & 0 & 0 \\ 0.7 & 0.5 & 0 & 0 \\ 0 & 0.7 & 0.5 & 0 \\ 0 & 0.2 & 0 & 0.5 \end{pmatrix}$$

The innovations ε_t follow a multivariate Student- t distribution with 10 degrees of freedom and identity covariance matrix. The sample size is 1000, and the horizon is set to $H = 5$.

The resulting lagged causal structure is illustrated in the following directed acyclic graph, where solid arrows indicate direct effects, dashed arrows represent indirect ones.

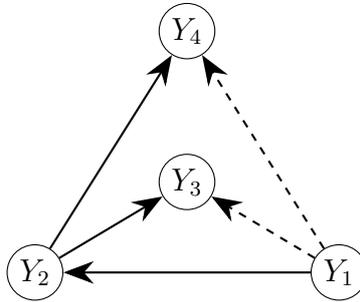


Figure 2.4: Causal DAG implied by A_1 .

Having introduced the VAR model with lagged dependencies and no contemporaneous structure, we now present the connectedness results obtained from both the IRF-based and Diebold–Yilmaz frameworks. The following tables summarize the directional influences, total effects, and net positions across variables.

	Y1	Y2	Y3	Y4	FROM
Y1	0.2680	0.0068	0.0030	0.0119	0.0751
Y2	0.3566	0.2754	0.0131	0.0092	0.5791
Y3	0.3848	0.3364	0.2737	0.0051	0.7263
Y4	0.1150	0.1024	0.0087	0.2538	0.4712
TO	0.7616	0.6180	0.0833	0.0937	TCI
NET	0.6865	0.0390	-0.6429	-0.3775	46.29%

Table 2.4: Normalized IRF-based connectedness table for the lagged-only system

	Y1	Y2	Y3	Y4	FROM
Y1	0.9932	0.0021	0.0002	0.0045	0.0068
Y2	0.5413	0.4542	0.0001	0.0044	0.5458
Y3	0.3920	0.3305	0.2745	0.0029	0.7255
Y4	0.0841	0.0831	0.0003	0.8325	0.1675
TO	1.0174	0.4158	0.0006	0.0119	TCI
NET	1.0106	-0.1301	-0.7249	-0.1556	36.14%

Table 2.5: Diebold–Yilmaz connectedness table for the lagged-only system

An important consistency check is the comparison of the diagonal elements across the two connectedness tables. These values reflect the own-variable cumulative response and should theoretically be equal across both frameworks. However, due to the row-wise normalization employed in the Diebold–Yilmaz approach, substantial discrepancies can be observed. For instance, in Table 2.5, the diagonal entries range from as low as 0.27 to 0.99, whereas our impulse response-based network preserves consistency in these values, yielding identical diagonal elements. A similar distortion appears in specific off-diagonal entries such as (2, 1) and (3, 2), where the relative magnitudes are notably altered in the Diebold–Yilmaz matrix due to normalization artifacts.

Moreover, our framework enables a clear decomposition of the total impact into direct and indirect components, in the case of purely lagged dependencies as well.

The lagged direct and indirect matrices computed from the cumulative structural impulse responses exhibit a clear correspondence with the causal structure encoded in the

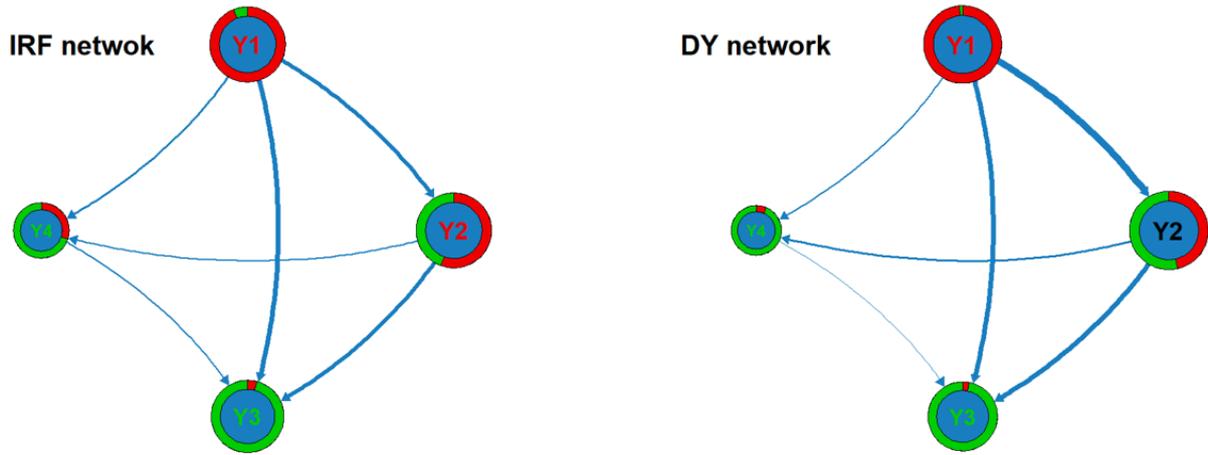


Figure 2.5: Visualization of the networks in case of lagged-only effects. The size of each node reflects its FROM+TO value. In the pie charts, the green (red) segments correspond to the proportion of FROM (TO) spillovers within the total for that series. Arrows indicate the direction of net spillovers, with their thickness representing the magnitude. Nodes labeled in green denote time series with FROM/SUM ratios of at least 55%, while red labels indicate TO/SUM ratios of 55% or more.

data-generating process.

L_d , which reflects the one-step, first-order effects derived from the lag matrix A_1 , is given by:

$$L_d = \begin{pmatrix} 0.5097 & 0.0006 & 0.0055 & 0.0255 \\ 0.7193 & 0.5188 & 0.0207 & 0.0258 \\ 0.0777 & 0.6832 & 0.5133 & 0.0053 \\ 0.0342 & 0.2001 & 0.0113 & 0.4934 \end{pmatrix}$$

L_i , which aggregates the multi-step propagated effects, is:

$$L_i = \begin{pmatrix} 0.5358 & 0.0515 & 0.0176 & 0.0655 \\ 1.9658 & 0.5832 & 0.0680 & 0.0444 \\ 2.8451 & 1.8842 & 0.5759 & 0.0337 \\ 0.7702 & 0.5776 & 0.0526 & 0.4436 \end{pmatrix}$$

These matrices faithfully reflect the underlying causal dependencies specified in the data-generating process.

2.3.3 Lagged and contemporaneous effects as well

To explore the behavior of our connectedness framework in a more complex and realistic setting, we now turn to a structural system that includes both contemporaneous and lagged dependencies.

The data-generating process is specified as follows:

$$A_0 Y_t = A_1 Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim t_{10}(0, I)$$

where the structural impact matrix A_0 encodes contemporaneous causal relationships, and A_1 governs the lagged dynamic effects. The specific matrices are given by:

$$A_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -0.5 & -0.5 & 1 & 0 \\ -0.5 & 0 & -0.5 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0.5 & 0 & 0 & 0 \\ 0.7 & 0.5 & 0 & 0 \\ 0 & 0.7 & 0.5 & 0 \\ 0 & 0.2 & 0 & 0.5 \end{pmatrix}$$

The innovations ε_t follow a multivariate Student- t distribution with 10 degrees of freedom and identity covariance matrix. The sample size is 1000, and the horizon is set to $H = 5$.

We now compare the outcomes of the IRF-based and Diebold–Yilmaz connectedness frameworks in a structural system that contains both contemporaneous and lagged dependencies. The following tables present the connectedness metrics derived from each approach.

	Y1	Y2	Y3	Y4	FROM
Y1	0.1237	0.0089	0.0020	0.0002	0.0821
Y2	0.1696	0.1112	0.0053	0.0056	0.6189
Y3	0.4343	0.2405	0.1249	0.0053	0.8449
Y4	0.5148	0.2468	0.1229	0.1156	0.8844
TO	0.9004	0.8170	0.5105	0.0877	TCI
NET	0.8184	0.1981	-0.3344	-0.7967	60.76%

Table 2.6: Normalized IRF-based connectedness table for the mixed system

	Y1	Y2	Y3	Y4	FROM
Y1	0.6992	0.0020	0.1118	0.1870	0.3008
Y2	0.3657	0.3094	0.1819	0.1430	0.6906
Y3	0.3519	0.1623	0.2882	0.1976	0.7118
Y4	0.3493	0.1182	0.2438	0.2887	0.7113
TO	1.0668	0.2826	0.5376	0.5275	TCI
NET	0.7661	-0.4080	-0.1742	-0.1839	60.36%

Table 2.7: Diebold–Yilmaz connectedness table for the mixed system

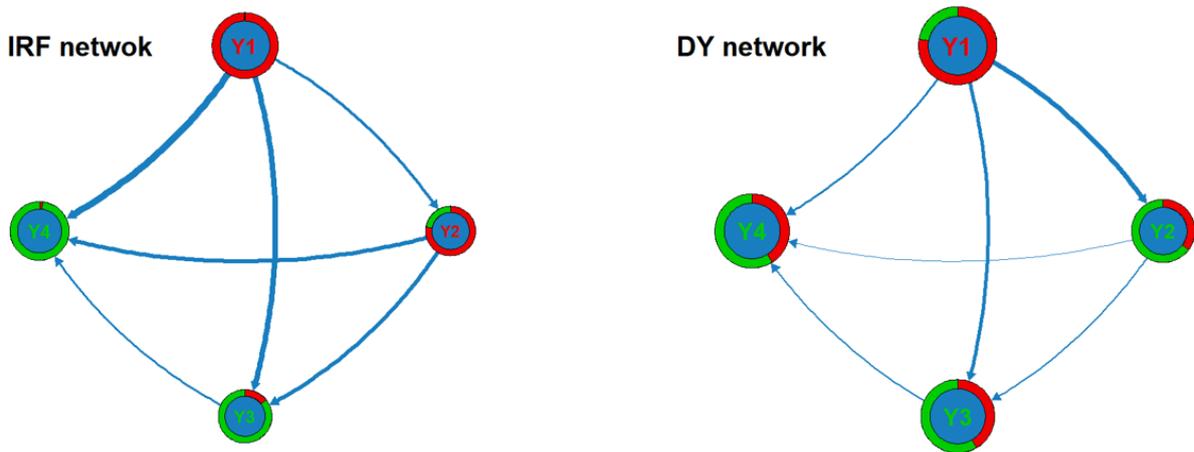


Figure 2.6: Visualization of the networks in case of both lagged and contemporaneous effects. The size of each node reflects its FROM+TO value. In the pie charts, the green (red) segments correspond to the proportion of FROM (TO) spillovers within the total for that series. Arrows indicate the direction of net spillovers, with their thickness representing the magnitude. Nodes labeled in green denote time series with FROM/SUM ratios of at least 55%, while red labels indicate TO/SUM ratios of 55% or more.

This example clearly illustrates the combined advantages of the impulse response-based network. The Diebold–Yilmaz approach once again fails to incorporate the contemporaneous structure: the resulting connectedness matrix is not lower triangular, despite the underlying form of the system. Moreover, the diagonal elements differ significantly in magnitude across variables, highlighting once again the distorting effect of row-wise normalization.

In contrast, the IRF-based network accurately reflects the causal structure through a lower triangular form, and the diagonal elements remain consistent across variables. This highlights the superiority of our framework in capturing both contemporaneous and dynamic dependencies in a coherent and interpretable manner.

This case also shows more clearly than any previous example how well the impulse response-based network can separate different types of effects. Since the system includes both contemporaneous and lagged connections, we can observe direct and indirect effects happening both within the same time period and across time. Our method makes it easy to separate and measure these effects one by one, giving a more complete and understandable picture of how shocks move through the system. The Diebold–Yilmaz approach cannot do this, because it does not use the structural relationships and mixes all effects together.

Contemporaneous Direct Effects:

$$C_d = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0445 & 1.0000 & 0.0000 & 0.0000 \\ 0.4810 & 0.5192 & 1.0000 & 0.0000 \\ 0.4552 & 0.0265 & 0.5303 & 1.0000 \end{pmatrix}$$

Contemporaneous Indirect Effects:

$$C_i = \begin{pmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0231 & 0.0000 & 0.0000 & 0.0000 \\ 0.2685 & 0.2753 & 0.0000 & 0.0000 \end{pmatrix}$$

Lagged Direct Effects:

$$L_d = \begin{pmatrix} 0.5451 & 0.0301 & 0.0071 & 0.0005 \\ 0.7657 & 0.4457 & 0.0446 & 0.0296 \\ 0.9240 & 1.1735 & 0.4670 & 0.0304 \\ 1.1125 & 0.8770 & 0.5107 & 0.4875 \end{pmatrix}$$

Lagged Indirect Effects:

$$L_i = \begin{pmatrix} 0.4327 & 0.1119 & 0.0249 & 0.0024 \\ 1.9021 & 0.3319 & 0.0407 & 0.0597 \\ 5.5169 & 2.1531 & 0.5300 & 0.0550 \\ 6.3952 & 2.7677 & 0.9241 & 0.3604 \end{pmatrix}$$

It is important to emphasize that the structural matrix A_0 must always be lower triangular, possibly after an appropriate row and column permutation, and its diagonal elements are normalized to one by definition.

In contrast, the lagged coefficient matrix A_1 is not subject to such structural restrictions. It does not need to be lower triangular, and its form depends entirely on the dynamics of the system being modeled.

In our examples, we deliberately selected cases where A_1 is lower triangular in order to make the properties of the IRF-based network more visually interpretable and easier to explain. However, the framework is fully applicable even when A_1 has a more general structure.

Chapter 3

Empirical analysis on financial data

In this section, we apply our methodology to real-world financial data in order to evaluate volatility spillovers across major US asset classes. Our empirical investigation builds on the dataset originally used by Diebold and Yilmaz (2012), which includes daily observations from the US stock, bond, foreign exchange, and commodity markets.

Specifically, we use data on the S&P 500 index (equities), the 10-year US Treasury bond yield (bonds), the New York Board of Trade USD index futures (foreign exchange), and the Dow Jones/UBS Commodity Index (commodities). The sample period spans from January 25, 1999, to January 29, 2010, yielding a total of 2771 daily observations.

To quantify market uncertainty, we calculate daily volatilities using the Parkinson estimator (Parkinson 1980), which captures intraday price variation based on the high and low prices. For market i on day t , the estimator is given by:

$$\hat{\sigma}_{it}^2 = 0.361 [\ln(P_{it}^{\max}) - \ln(P_{it}^{\min})]^2,$$

where P_{it}^{\max} and P_{it}^{\min} denote the daily high and low prices, respectively. This provides an estimate of the daily return variance.

Following Diebold and Yilmaz (2012), we annualize the daily standard deviations by computing:

$$\hat{\sigma}_{it}^{\text{ann}} = 100 \cdot \sqrt{365 \cdot \hat{\sigma}_{it}^2},$$

and apply a logarithmic transformation.

The descriptive statistics of the log-transformed annualized volatilities for the four asset classes are reported in the table below, as originally presented in Diebold and Yilmaz (2012).

	Stocks	Bonds	Commodities	FX
Mean	-9.70	-9.44	-10.69	-11.00
Median	-9.74	-9.44	-10.50	-10.99
Maximum	-5.45	-4.23	-6.34	-7.62
Minimum	-13.09	-13.79	-18.33	-16.86
Std. deviation	1.19	1.19	1.54	0.98
Skewness	0.21	0.019	-0.73	-0.21
Kurtosis	3.18	3.16	4.21	3.87

Table 3.1: Log volatility summary statistics for the four asset classes. Source: Diebold and Yilmaz (2012).

We estimate a (VAR(1)) model on the log-transformed, annualized volatilities of the four asset classes. The estimated coefficient matrix B_1 is:

$$B_1 = \begin{pmatrix} 0.5514 & 0.0461 & -0.0358 & 0.0521 \\ 0.1556 & 0.3618 & 0.1120 & 0.0259 \\ -0.1326 & 0.2007 & 0.4448 & 0.1122 \\ 0.1211 & 0.0352 & 0.0720 & 0.2062 \end{pmatrix}$$

The residual correlation matrix from the estimated VAR model is as follows:

$$\text{Corr}(\hat{u}_t) = \begin{pmatrix} 1.0000 & 0.2984 & 0.0419 & 0.2087 \\ 0.2984 & 1.0000 & 0.0988 & 0.2725 \\ 0.0419 & 0.0988 & 1.0000 & 0.0995 \\ 0.2087 & 0.2725 & 0.0995 & 1.0000 \end{pmatrix}$$

This correlation structure among the reduced-form residuals reflects potential contemporaneous relationships between asset classes, which cannot be interpreted causally without structural identification.

As evident from the matrix, the most relevant dependencies involve stocks, bonds, and foreign exchange. Determining the correct causal ordering among these three variables is therefore central to identifying the structural system.

To assess the stability of the identified structure, we implement a bootstrap procedure. Specifically, we repeatedly draw samples of the same size as the original dataset using sampling with replacement. For each bootstrap sample, we run the LiNGAM al-

gorithm and extract the resulting binary contemporaneous adjacency matrix. This procedure is repeated 1000 times, and the resulting adjacency structures are aggregated by their frequency of occurrence.

The three most frequently observed adjacency matrices and their relative frequencies are presented below. Each matrix corresponds to a directed graph where entry $(i, j) = 1$ indicates a contemporaneous effect from variable j to variable i :

$$\text{Adjacency Matrix 1 (Frequency: 33.9\%)} \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\text{Adjacency Matrix 2 (Frequency: 29.6\%)} \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\text{Adjacency Matrix 3 (Frequency: 16.8\%)} \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Based on the dominant patterns in the bootstrap results, we infer a causal ordering of **FX** \rightarrow **Stocks** \rightarrow **Bonds** \rightarrow **Commodities**, with **Commodities** did not exhibit any consistent incoming or outgoing contemporaneous effects across the bootstrap replications. Therefore, its position in the contemporaneous structure remains ambiguous and can be considered weakly connected.

Using the bootstrap-selected causal ordering, we constructed the contemporaneous structural matrix A_0 and the lagged coefficient matrix A_1 .

$$A_0 = \begin{pmatrix} 1.0000 & 0 & -0.0155 & -0.2168 \\ -0.2689 & 1.0000 & -0.0524 & -0.2399 \\ 0 & 0 & 1.0000 & 0 \\ 0 & 0 & -0.0692 & 1.0000 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} 0.5272 & 0.0354 & -0.0583 & 0.0056 \\ -0.0148 & 0.3305 & 0.0810 & -0.0434 \\ -0.1326 & 0.2007 & 0.4448 & 0.1122 \\ 0.1303 & 0.0213 & 0.0412 & 0.1985 \end{pmatrix}$$

Based on the estimated structural matrices A_0 and A_1 , we compute the structural impulse response functions up to a horizon of $H = 5$. The resulting cumulative impulse response matrix $IRF^{(5)}$ serves as the basis for the connectedness measures presented below.

	Stocks	Bonds	Commodities	FX	FROM
Stocks	0.6036	0.0347	0.0180	0.1742	0.2732
Bonds	0.2442	0.4440	0.1207	0.1911	0.5560
Commodities	0.0387	0.1490	0.5226	0.0964	0.3522
FX	0.0944	0.0361	0.0736	0.3740	0.3530
TO	0.3846	0.3312	0.2889	0.5525	TCI
NET	0.1114	-0.2249	-0.0633	0.1995	38.36%

Table 3.2: Normalized IRF-based connectedness table for dy2012 data

For comparison, we also compute the connectedness measures using the Diebold–Yilmaz framework.

	Stocks	Bonds	Commodities	FX	FROM
Stocks	0.8506	0.0960	0.0014	0.0520	0.1494
Bonds	0.1288	0.7528	0.0445	0.0739	0.2472
Commodities	0.0015	0.0562	0.9112	0.0311	0.0888
FX	0.0767	0.0829	0.0281	0.8122	0.1878
TO	0.2070	0.2352	0.0740	0.1570	TCI
NET	0.0576	-0.0120	-0.0148	-0.0308	16.83%

Table 3.3: Diebold–Yilmaz connectedness table for dy2012 data

The main distinction between the two frameworks lies in the behavior of the diagonal elements. In the Diebold–Yilmaz network, the diagonals are much more dominant,

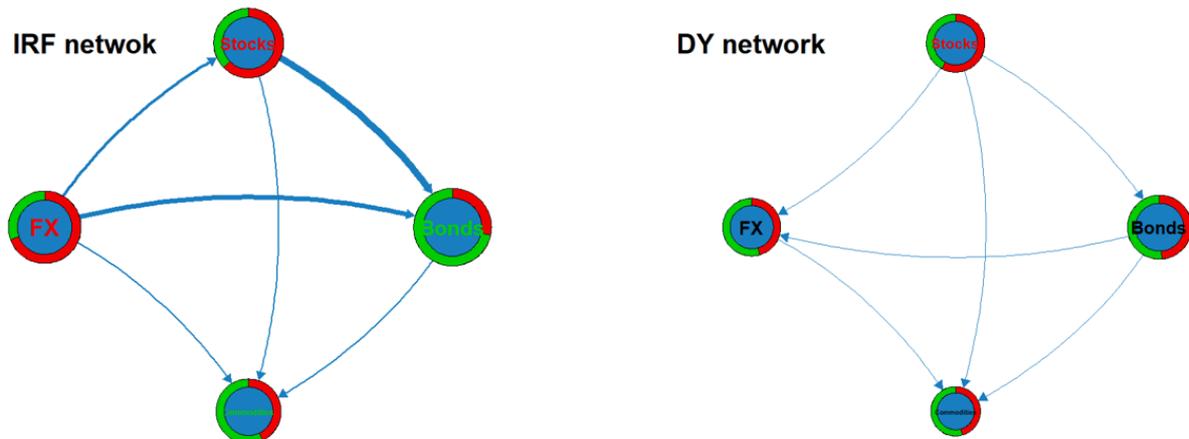


Figure 3.1: Visualization of the networks in case of dataset `dy2012`. The size of each node reflects its FROM+TO value. In the pie charts, the green (red) segments correspond to the proportion of FROM (TO) spillovers within the total for that series. Arrows indicate the direction of net spillovers, with their thickness representing the magnitude. Nodes labeled in green denote time series with FROM/SUM ratios of at least 55%, while red labels indicate TO/SUM ratios of 55% or more.

resulting in a substantially lower Total Connectedness Index.

Another striking difference is observed in the net connectedness of the foreign exchange market. In the DY network, the FX variable appears nearly neutral, with a net value close to zero. In contrast, under the IRF-based network—where the structural causal ordering is explicitly identified—FX emerges as a net transmitter of volatility, highlighting its role as an influential driver in the system.

To further investigate the underlying structure of the IRF-based network, we observe the decomposition of the cumulative impulse response matrix $IRF^{(5)}$. This decomposition enables a more nuanced comparison of how shocks propagate across markets in both timing and type. The matrices below are reported in their original, non-normalized form:

$$C_d = \begin{pmatrix} 1.0000 & 0.0000 & 0.0155 & 0.2168 \\ 0.2689 & 1.0000 & 0.0524 & 0.2399 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0692 & 1.0000 \end{pmatrix}$$

$$C_i = \begin{pmatrix} 0.0000 & 0.0000 & 0.0150 & 0.0000 \\ 0.0000 & 0.0000 & 0.0248 & 0.0583 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{pmatrix}$$

$$L_d = \begin{pmatrix} 0.5638 & 0.0461 & 0.0118 & 0.1854 \\ 0.2529 & 0.3618 & 0.1465 & 0.1676 \\ 0.0786 & 0.2007 & 0.4640 & 0.1433 \\ 0.1306 & 0.0352 & 0.0927 & 0.2430 \end{pmatrix}$$

$$L_i = \begin{pmatrix} 0.7190 & 0.0851 & 0.0258 & 0.2566 \\ 0.4016 & 0.3173 & 0.2329 & 0.2570 \\ 0.0677 & 0.3628 & 0.5124 & 0.2214 \\ 0.2264 & 0.1014 & 0.1163 & 0.1716 \end{pmatrix}$$

This decomposition shows that contemporaneous indirect effects (C_i) are negligible. In contrast, among the lagged components, the indirect effects (L_i) are more pronounced, indicating that the majority of dynamic spillovers propagate over time through indirect channels.

3.1 Rolling network

To explore the time-varying nature of connectedness, we conduct a rolling window analysis based on the first seven years of the `dy2012` dataset. We use a window size of 252 observations, corresponding approximately to one calendar year of daily data.

For each rolling window, we re-estimate the entire network structure separately for both the Diebold–Yilmaz and the IRF-based frameworks. This includes re-fitting the VAR model, identifying the contemporaneous structure (in the case of the IRF network), and computing the connectedness measures. We then extract the Total Connectedness Index from each window and visualize its evolution over time for both approaches.

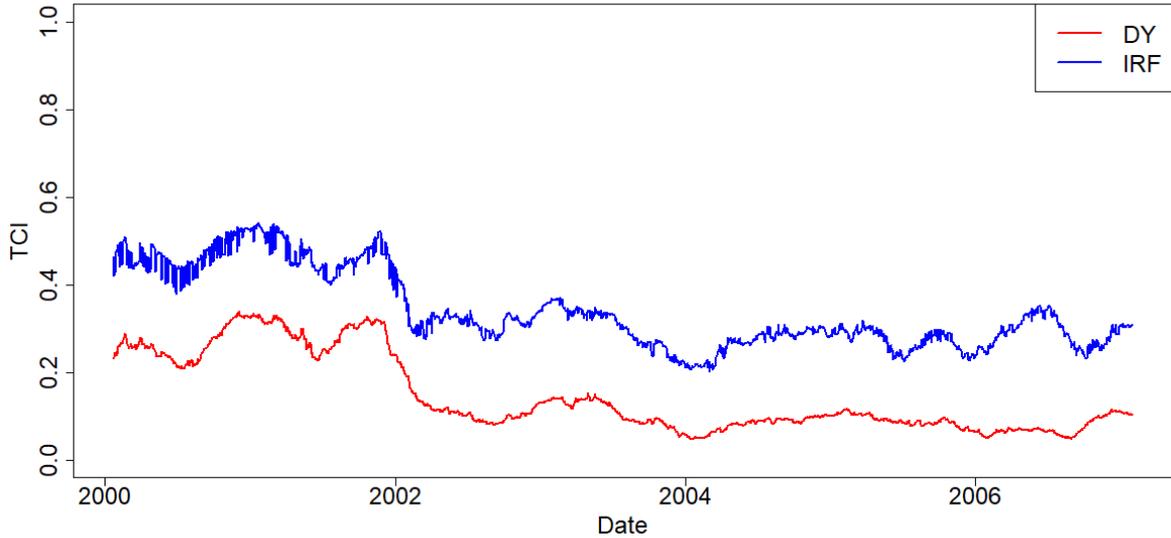


Figure 3.2: Rolling Total Connectedness Index computed from IRF and Diebold–Yilmaz networks using a 252-day moving window.

Figure 3.2 illustrates the evolution of the Total Connectedness Index over time, as computed from both the IRF-based and Diebold–Yilmaz frameworks. It is evident that the IRF-based network consistently yields higher TCI values.

In periods where the underlying causal structure becomes less stable or undergoes a transition, the LiNGAM algorithm finds it more difficult to establish a clear ordering. This results in small fluctuations in the IRF-based TCI. One such example is the interval between 2000 and 2001, where the rolling TCI briefly dips and rebounds, reflecting underlying structural uncertainty.

3.2 Robustness check

To assess the robustness of our IRF-based network to the choice of forecast horizon, we perform a rolling-window analysis using different values of H . The analysis is conducted over the first seven years of the `dy2012` dataset, applying a 252-day moving window, as before. For each window, we compute the Total Connectedness Index (TCI) using the IRF-based approach for horizons $H = 3$, $H = 5$, and $H = 7$.

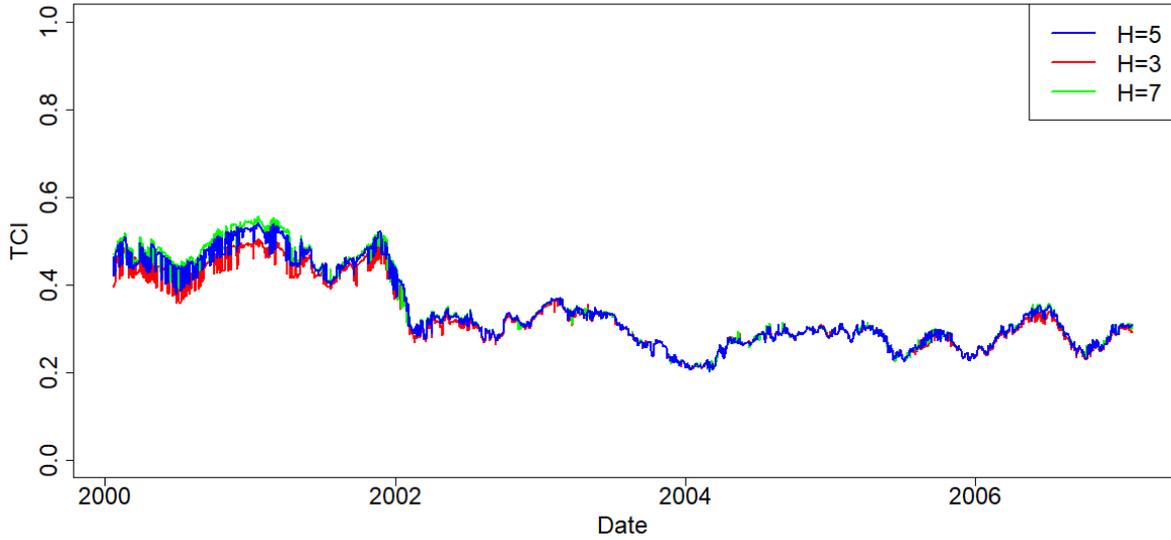


Figure 3.3: Rolling IRF-based TCI computed for different impulse response horizons ($H = 3$, $H = 5$, $H = 7$)

As shown in Figure 3.3, the TCI trajectories for the three horizon lengths are nearly indistinguishable throughout the sample period. While minor deviations occur at certain points, the overall pattern suggests a high degree of robustness to the choice of H .

To further support this observation, Table 3.4 reports the average TCI values computed over all rolling windows for each horizon. The differences are marginal, with only a slight upward trend as the horizon increases.

Horizon	Average TCI
$H = 3$	0.3313
$H = 5$	0.3404
$H = 7$	0.3431

Table 3.4: Average IRF-based Total Connectedness Index for different horizons

This phenomenon is consistent with the interpretation that, as the forecast horizon increases, indirect effects accumulate and become more pronounced. These effects typically involve off-diagonal propagation paths, thereby increasing total connectedness without inflating the diagonal elements.

Chapter 4

Conclusion

In this paper, we introduced a novel network-based connectedness framework built upon structural impulse response functions. By explicitly incorporating contemporaneous causal relationships—identified through the LiNGAM algorithm—our method provides a more interpretable and structurally grounded view of how shocks propagate across variables.

Using a series of simulated data-generating processes, we demonstrated that the IRF-based network is capable of distinguishing between direct and indirect, as well as contemporaneous and lagged effects. Such decomposition is not available in the Diebold–Yilmaz framework, which relies on generalized impulse responses and does not account for causal ordering. As a result, the DY network tends to produce more symmetric connectedness matrices with suppressed off-diagonal and net effects.

In our empirical application, we revisited the well-known dataset from Diebold and Yilmaz (2012) (Diebold and Yilmaz 2012), referred to as `dy2012`. The IRF-based network consistently delivered higher Total Connectedness Index values and clearer directional roles—particularly for the foreign exchange market. Rolling window and robustness analyses further confirmed the stability of the IRF-based network and its resilience to variations in the forecast horizon.

Overall, our results suggest that IRF-based networks offer a theoretically sound and empirically robust alternative to traditional connectedness measures, particularly when the identification of structural shocks is feasible.

Appendix: Accuracy of the LiNGAM Algorithm

This appendix presents simulation results evaluating the accuracy of the LiNGAM algorithm under varying data conditions, using a fixed structural model. Specifically, we assess how the algorithm performs in recovering the correct contemporaneous causal structure when sample size and the heaviness of tails (captured by the degrees of freedom of the t -distribution) change.

The structural matrix used for data generation is:

$$A_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -0.5 & -0.5 & 1 & 0 \\ -0.5 & 0 & -0.5 & 1 \end{pmatrix}$$

Thus, the corresponding adjacency matrix to be recovered by the LiNGAM algorithm is:

$$\text{Adjacency matrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

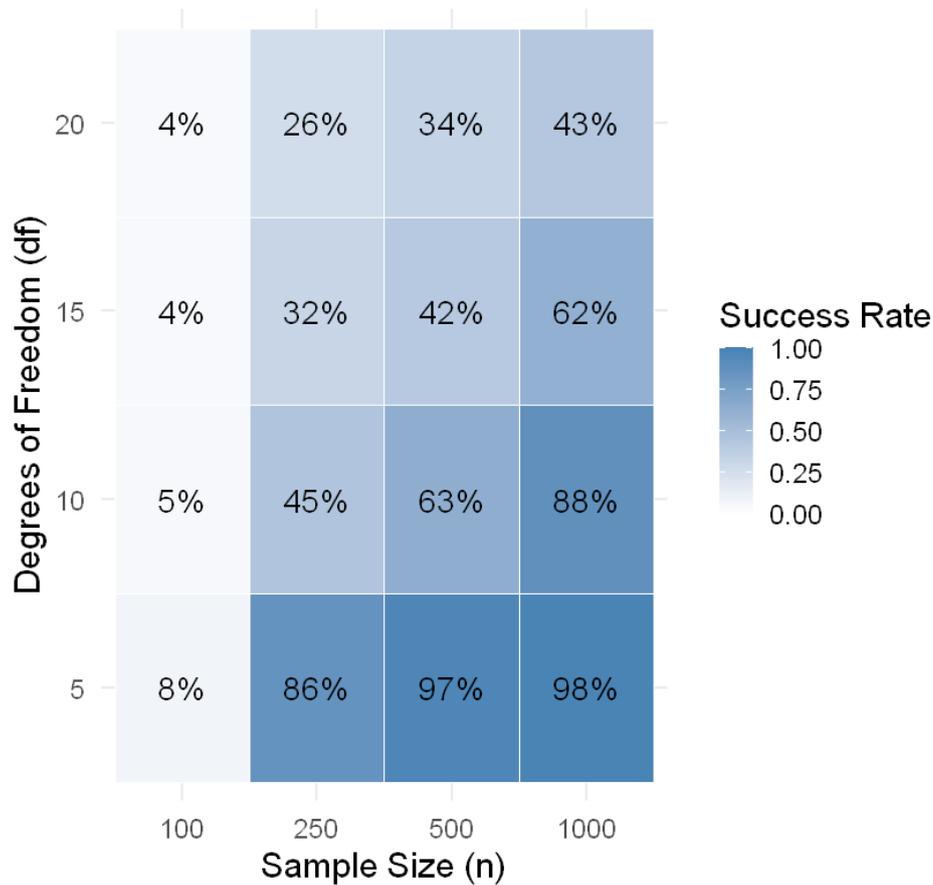


Figure 4.1: Accuracy of the LiNGAM algorithm in recovering the structure of the fixed A_0 matrix across varying sample sizes and degrees of freedom

Figure 4.1 displays a heatmap summarizing the percentage of correct identifications of the exact causal structure by the LiNGAM algorithm, based on 1000 simulation runs for each combination of sample size and degrees of freedom.

The results reveal two main insights:

- **Larger sample sizes** significantly improve accuracy, as expected.
- **Lower degrees of freedom**—i.e., heavier tails—also enhance recovery. This is particularly relevant for financial data, where error terms often follow a Student- t distribution with 4–6 degrees of freedom, aligning well with the favorable regions of the heatmap.

Furthermore, even when the full adjacency matrix is not perfectly recovered, the correct *causal ordering* can still be inferred. As demonstrated in our empirical application, combining LiNGAM with **bootstrap aggregation** helps stabilize the results and enhances the robustness of the recovered structure.

Magyar nyelvű összefoglaló

A makrogazdaságtanban és a pénzügyek területén kulcsfontosságú kérdés, hogy hogyan hatnak egymásra az egyes szereplők az összekapcsolódó rendszerekben. Az elmúlt években egyre népszerűbbé váltak a hálózatalapú kapcsolati mutatók, amelyek segítségével mérhetővé váltak az egyes változók közötti kölcsönhatások. A legismertebb módszerek közé tartozik a Diebold–Yilmaz-féle keretrendszer, amely vektor-autoregressziós modellekből származó előrejelzési hibavariancia-felbontásokon alapul. Bár ez a megközelítés széles körben elterjedt, az alkalmazott általánosított impulzusválasz függvények nem veszik figyelembe a kauzális sorrendet, így a kapott hálózatok gyakran ellentmondásosak vagy nehezen értelmezhetőek.

A dolgozat egy új, strukturális impulzusválasz-függvényeken alapuló hálózati módszert javasol, amely lehetővé teszi az oksági viszonyok figyelembevételét. A LiNGAM algoritmus segítségével feltárjuk a rendszeren belüli egyidejű oksági kapcsolatokat, majd egy strukturális VAR (SVAR) modellből származtatjuk a sokterjedést leíró válaszfüggvényeket. Az új módszerben lehetőség nyílik a közvetlen és közvetett, valamint a késleltetett és azonos idejű hatások szétválasztására, ami részletesebb és értelmezhetőbb eredményeket ad.

A módszer részeként egy kumulatív impulzusválasz-mátrixot definiálunk, melyet úgy normalizálunk, hogy megőrizze a változók közötti kapcsolatok relatív nagyságát. Ez elkerüli a DY-módszer egyik problémáját, a soronkénti normalizálás okozta információvesztést. A mátrix alapján a DY rendszerhez hasonlóan különböző összefüggőségi mutatókat – például *TO*, *FROM*, *NET* és a teljes összefüggőségi indexet (*Total Connectedness Index*, *TCI*) – vezetünk be.

A módszer hatékonyságát szimulációs kísérletek során vizsgáltuk, különböző adatgeneráló folyamatok mellett. Az IRF-alapú hálózat rendre pontosabban tükrözte a valódi oksági viszonyokat, és érzékenyebben jelezte a spillover-hatásokat. Az empirikus vizsgálatban a Diebold és Yilmaz (2012) által használt *dy2012* adathalmazt elemeztük, ahol az

IRF-hálózat magasabb TCI-értékeket eredményezett. A gördülő ablakos és robusztitási vizsgálatok megerősítették a módszer stabilitását és megbízhatóságát.

Összességében az eredmények azt sugallják, hogy az IRF-alapú hálózati megközelítés elméletileg megalapozott és empirikusan robusztus alternatívát kínál az összefüggési viszonyok mérésére, különösen akkor, ha az azonos idejű kapcsolatok azonosítása lehetséges.

Angol–magyar szakszószedet

Angol kifejezés	Magyar megfelelő
shock propagation	sokkterjedés
network-based measure	hálózatalapú mutató
spillover effect	spillover-hatás / áthúzódó hatás
connectedness	kapcsolódás / összefonódottság
forecast error variance decomposition	előrejelzési hibavariancia-felbontás
vector autoregression	vektor-autoregressziós modell
impulse response function	impulzusválasz függvény
generalized impulse response	általánosított impulzusválasz-függvény
structural VAR	strukturális VAR modell
causal ordering	kauzális sorrend
contemporaneous effect	azonos idejű hatás
lagged effect	késleltetett hatás
cumulative impulse response matrix	kumulatív impulzusválasz-mátrix
Total Connectedness Index	teljes összefüggőségi index
rolling window	gördülő ablak
identification of shocks	sokkok azonosítása
LiNGAM algorithm	LiNGAM algoritmus
structural interpretability	strukturális értelmezhetőség

MI-eszközhasználati nyilatkozat

Alulírott, **Espán Márton**, nyilatkozom, hogy szakdolgozatom elkészítése során az alább felsorolt feladatok elvégzésére a megadott MI-alapú eszközöket alkalmaztam:

Feladat	Felhasznált eszköz	Felhasználás helye	Megjegyzés
LaTeX táblázatok készítése, R kódok írása és nyelvhelyesség ellenőrzése	ChatGPT 4o	Teljes dolgozat	Az MI eszközt a letisztultabb kódoláshoz, táblázatok készítésére és a szellemi munkám nyelvhelyességének ellenőrzésére használtam.

A felsoroltakon túl más MI-alapú eszközt nem használtam.

Bibliography

- Caloia, Francesco Giuseppe, Andrea Cipollini, and Silvia Muzzioli (2018). “Asymmetric semi-volatility spillover effects in EMU stock markets”. In: *International Review of Financial Analysis* 57, pp. 221–230.
- Cont, Rama (2001). “Empirical properties of asset returns: stylized facts and statistical issues”. In: *Quantitative finance* 1.2, p. 223.
- Diebold, Francis X and Kamil Yilmaz (2009). “Measuring financial asset return and volatility spillovers, with application to global equity markets”. In: *The Economic Journal* 119.534, pp. 158–171.
- (2012). “Better to give than to receive: Predictive directional measurement of volatility spillovers”. In: *International Journal of forecasting* 28.1, pp. 57–66.
- Lütkepohl, Helmut (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.
- Nolan, John P and Diana Ojeda-Revah (2013). “Linear and nonlinear regression with stable errors”. In: *Journal of Econometrics* 172.2, pp. 186–194.
- Parkinson, Michael (1980). “The extreme value method for estimating the variance of the rate of return”. In: *Journal of business*, pp. 61–65.
- Shimizu, Shohei et al. (2006). “A linear non-Gaussian acyclic model for causal discovery.” In: *Journal of Machine Learning Research* 7.10.