Network Analysis of the Financial Sector: A Comprehensive Perspective with Adaptive Joint LASSO Method

Badics Milán Csaba

Pénzügy Intézet

Témavezető:
Huszár Zsuzsa Réka Ph.D.
Egyetemi docens

Budapesti Corvinus Egyetem

Közgazdasági és Gazdaságinformatikai Doktori Iskola

Network Analysis of the Financial Sector: A Comprehensive Perspective with Adaptive Joint LASSO Method

Doktori értekezés

Badics Milán Csaba

Budapest, 2023.11.20.

Contents

List of Figures

List of Tables

Acknowledgements

I was a first-year Finance Msc student when I first learned about Vector autoregressive methods. At this point in my life, I didn't think that twelve years later, this approach would be a fundamental tool of my dissertation. It has been a long way until I write these sentences, but I have learned a lot over these years. I can't name everyone who helped me on this path, but I will do my best!

First, I am very thankful for my supervisor, Zsuzsa Réka Huszar, who helped me finish my dissertation on time. This thesis would not have been completed without her. I also owe gratitude to Edina Berlinger, who has been my supervisor in the early stages of my dissertation. I am grateful for her encouragement and support even after she finished my supervision. I always felt like I could tell her about my problems, and she had a lot of great advice on dealing with them. I would also like to acknowledge all the members of my thesis committee. Without the detailed thesis reviews, my dissertation would not have gained the final form. I extend my acknowledgment to my co-authors, Áron Dénes Hartvig and Balázs Bence Kotro as well.

I also feel obliged to mention two former teachers of mine, Zsolt Darvas, who helped me gain access to bootstrap methods, and Gábor Uliha, who first introduced me regularization methods.

While writing this dissertation, I have received much support from my colleagues. I am thankful for the whole Finance Institute, especially to Péter Csóka, Kata Váradi, Dávid Szabó, Barbara Dömötör, and Zsolt Bihary. Special thanks to Lilla Keresztúri, Péter Kerényi, Attila Víg, and Ferenc Illés, whose opinions always meant a lot to me. I express my sincere thanks to Gábor Neszveda, Mihály Ormos, Róbert Lieli, and András Fülöp, who participated in reviewing my articles and added feedback on them at conferences. Áron Szakály, Márk Bérchegyi, Tamás Jászai, Péter Simon, Bálint Plangár, and Dániel Bíró, it was great to talk to you about the mathematical background of the methods that I used in my dissertation. I also want to thank my former students, Áron Szakály and Péter Simon, for the shared R codes.

I acknowledge my former manager, Imre Koncz, and my current co-owners, Balázs Márkus, and Zsolt Dányi, for allowing my professional and academic career to progress parallelly. I also want to thank my friends who have supported me through this route, especially Misi and Zoli.

My family has always supported me during my PhD. I thank my parents for their love and support. Finally, I am most grateful to my wife Vera and my son Matyi, whose support, patience, and encouragement have strengthened me in challenging moments. This thesis is dedicated to them.

1 Chapter 1. Introduction

The emergence of systemic risk from the 2008 Global Financial Crisis (GFC) and the recent energy market shocks following the COVID-19 pandemic and the onset of the Russo-Ukrainian war highlighted the importance of understanding shock spillovers for policymakers and academics alike. Specifically, shock transmissions in financial network settings warrant attention to support regulatory or policy interventions for effectively mitigating or preventing the transmission of systemic risk.

The thesis focuses on this important topic from various angles to answer the following questions:

- How can machine learning techniques help to improve financial network studies, the analysis of high dimensional time series?

- How can structural changes in financial networks efficiently examined using event analysis framework?

My thesis provides methodological and conceptual contributions to financial network analysis. The **contribution of my thesis** to the existing literature is threefold:

1. I propose a new regularization method, the adaptive joint least absolute shrinkage and selection operator (AJ LASSO). The innovation in this method is that it accounts for possible sparsity both in the coefficient and the covariance matrix of the estimated Vector autoregressive (VAR) model. The method is especially suitable for high-dimension network analysis, where parameter estimation is a critical issue.

2. I extend the Diebold-Yilmaz (DY) framework with an event study tool to facilitate an in-depth analysis of the contagion channels during critical local and global shocks in financial networks. Integrating a moving-block bootstrap method (MBB) into the framework, I can investigate how the observed shocks transform the financial networks.

3. My conceptual contribution is that I characterize and analyze illiquidity networks. The liquidity concerns are globally recognized and the focal point of the newly revised Basel IV regulatory guidelines; thus, intuitively and from a regulatory perspective, illiquidity spillover analysis in the financial sector is valuable. I analyze the illiquidity connectedness of financial institutions (FIs) and show that the illiquidity network better tracks the dominant shock transmitters in the system during financial turmoil than the volatility-based networks. I show that the DY framework extended with the MBB method is a powerful tool to identify troubled financial institutions and contagion channels.

The thesis is structured in 7 Chapters. Chapter 2 briefly surveys the financial networks-based systemic risk literature, emphasizing empirical studies. I summarize the literature, highlighting the main findings and the applied methods. In the second part of the chapter, I present the popular Diebold-Yilmaz framework (Diebold and Yilmaz, 2012; Diebold and Yılmaz, 2014) and compare the other network models: Granger-causality network and CoVaR-based networks. I discuss the framework's applicability for measuring spillovers in financial networks. Lastly, I present examples of the DY spillover table and network graphs to visualize the in-depth insights from the DY framework for financial network.

In Chapter 3, I highlight the limitations of the Diebold-Yilmaz framework and the room for improvement. I summarize the shortcomings of the empirical network literature and present the contribution of my thesis in detail.

In Chapter 4, I discuss the machine learning regularization approaches, specifically the LASSO (Tibshirani, 1996; Bühlmann and Van De Geer, 2011) and adaptive LASSO (Zou, 2006) methods. In the first part of the chapter, I summarize the advantages of LASSO-based VAR models in high-dimension time series environments. Next, I compare two LASSO-based VAR estimators (row-wise LASSO and penalized maximum likelihood-based LASSO). At the end of this chapter, I explain the connection between the regularization methods and the DY framework.

In the first part of Chapter 5, I propose a new regularization method that is useful in the Diebold-Yilmaz framework, as it is able to shrink and select the essential coefficients of the VAR and simultaneously account for possible sparsity in the distribution of the VAR model errors. I devise a novel extended penalized maximum likelihood estimator of the VAR coefficient and covariance matrix, and I refer to this method as the adaptive joint LASSO algorithm. In the second part of the chapter, I conduct an extensive simulation study to analyze my proposed estimator's statistical properties and compare it to the most commonly applied regularization methods in the Diebold-Yilmaz literature. Based on the earlier literature, I consider two alternative autoregressive coefficient matrices for the VAR(1) and three different covariances matrixes, and compare the performance of the adaptive joint LASSO to the estimators obtained from established alternative models, such as ordinary least squares (OLS), LASSO, and adaptive LASSO methods. I evaluate the robustness of the results to the choice of the length and number of time series and the contemporaneous correlation. Based on the results, it can be confidently stated that the adaptive joint LASSO outperforms the commonly used LASSO and the adaptive LASSO methods in various settings where a high dimension is present. My estimator yields superior sparsity pattern recognition, lower model error, and Diebold-Yilmaz spillover table estimation error.

Combining the DY spillover index method for network modeling with event study methodology may be useful at turbulent times. In the first part of the Chapter 6, I extend the DY framework with event study tools to provide more insights into the contagion channels appearing in the financial networks during turbulent times. I build an event study framework to identify critical events and linkages in the evolution of systemic risk. One of the limitations of the original DY framework is the absence of a formal statistic to test whether the changes in the spillover matrix are economically and statistically significant. To address at least the statistical shortcoming, I introduce a formal test to compare the distribution of the SUM, FROM, TO, and NET spillover indices over time. Integrating the residual-based moving-block bootstrap method into the DY framework, I can investigate how the observed shocks transform the financial networks.

To address the gap in the empirical literature, in the second part of Chapter 6, I present a characterization of the illiquidity connectedness of a US financial network with my proposed AJ LASSO method in the Diebold-Yilmaz framework. Besides investors' fear (the volatility spillover-based network), the illiquidity contagion can be followed by these spillover indices. I use daily data to cover a more than 15-year-long horizon, covering two major crises with strikingly different attributes and impacts, the GFC and the European Sovereign Debt Crisis (ESDC). I investigate the density of networks during calm and turbulent periods. I analyze the network on macro (system analysis) and micro levels (node-wise analysis), as the dominant nodes (key participants) of the system have been

identified only in a few papers. The empirical results suggest that the illiquidity total spillover indices are also relevant in the analysis of systemic risk as they behave differently than total volatility connectedness. On a macro level, both total spillover indices react to shocks in the financial system, but even in tranquil periods, the dynamics might vary.

In the last part of the Chapter 6, I consider the critical financial events of the GFC, as Diebold and Yılmaz (2014), with the analytical purpose of testing the performance of my new event study framework. Applying the MBB method, I analyze the volatility and liquidity shocks, focusing on four key events of the GFC. Besides that, I provide daily snapshots of the financial network to illustrate that illiquidity spillovers act as contagion channels during turmoil time. During financial turmoil, the troubled institutions transmit illiquidity shocks to the other members of the system, and as a result, their pairwise connections strengthen, and average connectedness increases in the network. I show that illiquidity connectedness outperforms volatility in signaling financial shocks during these events, and besides that, illiquidity connectedness better tracks the dominant shock transmitters in the financial system. NET pairwise illiquidity connections act as contagion channels during turmoil and promote cascades of failures. From these linkages, we can identify the FIs that are the most affected by the failure of another FI in the network. I conclude that the DY framework extended with the residual-based MBB method is a powerful tool to execute event study analysis and can identify troubled FIs and contagion channels. Also, it can be utilized for monitoring the financial system on a daily basis.

In the first part of Chapter 7, I summarize my theoretical contribution to the existing network literature. I summarize the advantages of my AJ LASSO method in high-dimension network modeling and reflect on my proposed event study analysis integration into the DY framework. In the second part of the chapter, I summarize my empirical contribution to the existing systemic risk literature. In terms of my conceptual contribution, I highlight why illiquidity connectedness analysis can outperform volatility spillover analysis in signaling financial shocks during financial crisis periods. In addition, I conclude that the Diebold-Yilmaz framework extended with the residual-based moving-block bootstrap method is a powerful tool to execute event study analysis and can identify troubled financial institutions and contagion channels.

2 Chapter 2. Systemic risk modeling and Diebold-Yilmaz framework

In this chapter, I provide a compact survey of the extensive systemic risk literature which exploded since the GFC. I specifically focus on emphasizing empirical studies, those with financial network analysis where the research question is understanding risk contagion in the network. In the second part of the chapter, I summarize the popular Diebold-Yilmaz framework (Diebold and Yilmaz, 2012; Diebold and Yılmaz, 2014) and compare it to other network models, such as the Granger-causality network and CoVaR-based networks. I discuss the framework's applicability for measuring connectedness and spillovers in financial networks. Lastly, I present examples of the DY spillover table and network graphs to visualize the in-depth insights from the DY framework for financial network.

2.1 Literature review of network-based systemic risk modeling

The first part of this section (the empirical systemic risk literature review) is based on two working paper's, as a result of my collaboration with Áron Dénes Hartvig. Earlier versions of the research were presented at the 11th and 12th Annual Financial Market Liquidity (AFML) Conference[1] and the 2021 Annual Hungarian MKE Conference[2]. Major modifications are made to align the paper's content with the dissertation format.

It is somewhat surprising in hindsight that network analysis of financial systems has become widely recognized as a critical regulatory issue only after the GFC. Pairwise and group connections among FIs can both stabilize or destabilize the system. While on the one hand, FIs' interconnectivity aids flexibility to investment and financing in the economy, on the other hand, these connections may contribute to risk propagation during crisis periods. Furthermore, high connectedness (e.g., strong liquidity dependency) between FIs promotes the sudden transformation of the network architecture (Elliott et al., 2014; Acemoglu et al., 2015).

Pairwise and groupwise spillovers in the financial network play a crucial role in systemic risk assessment. Furthermore, the strength of connections is non-constant, and significantly varies over time, sharply increasing during crisis periods, as documented in a GFC study (Diebold and Yılmaz, 2014). In the network the FIs' connections are especially critical with the liquidity provider or strong network participants because many smaller institutions rely on the large institutions for financing, which is especially critical during turbulent times. For this reason, the regulators need to monitor and analyze the structural changes in financial networks and identify the systematically important financial institutions (SIFIs).

Motivated by the GFC, the financial network literature exploded, with studies attempting to better understand or model the connections between FIs in the US and globally. Acemoglu et al. (2015) and Elliott et al. (2014) derived financial system connectedness from the crossholdings of shares, debt, or obligations, and investments. These

---

[1]AFML Previous Conferences: https://www.uni-corvinus.hu/ind/events/annual-financial-market-liquidity-conference/afml-2020-previous-conferences/?lang=en

[2]2021 Annaul Hungarian MKE Conference: https://www.mktudegy.hu/konferencia/magyar-kozgazdasagtudomanyi-egyesulet-mke-konferenciafelhivas-2021-december-21-22/38282/

linkages can enhance diversification, but such linkages can also lead to cascading defaults and failures in a system. Consequently, identifying and monitoring financial linkages is essential for understanding system vulnerability as a result of large-scale contagions during turmoil.

Besides that, Acemoglu et al. (2015) showed that the connectedness of the financial network enhances stability in the system until the magnitude and the number of shocks hitting the network remain low. However, if the level of financial distress exceeds a certain threshold, the structure of the financial network dictates the extent of contagion. In other words, higher financial connectedness makes the network more sensitive and more prone to contagion of shocks. Thus, not only the analysis of SIFIs but the inclusion of the whole system is necessary. Institutions on the periphery may move to the center of the financial network owing to a significant shock that restructures the linkages.

Recently numerous studies have introduced new empirical frameworks (quantitative methods) to model financial linkages and measure contagion and systemic risk. Bisias et al. (2012) identified more than 30 quantitative measures of systemic risk in the finance literature. In their survey, they classify the measures into six groups. One of them is a network-based approach. On the empirical side of the systemic risk modeling - from network perspective - the widespread methods are Granger causality network (Billio et al., 2012), delta conditional value-at-risk (ΔCoVaR) (Tobias and Brunnermeier, 2016), systemic expected shortfall (SES) (Acharya et al., 2017) and numerous studies appeared based on the VAR based Diebold-Yilmaz framework (Diebold and Yilmaz, 2009, 2012; Diebold and Yılmaz, 2014). Although there are several methods in the literature to measure systemic risk, the Diebold-Yilmaz framework has several favorable properties. First, unlike Granger causality network analysis (Billio et al., 2012), the DY framework estimates weighted connections (Diebold and Yilmaz, 2012). Second, CoVaR (Tobias and Brunnermeier, 2016) and marginal expected shortfall (MES) (Acharya et al., 2012) are related to the directional connectedness indices of Diebold and Yilmaz, so unlike the DY framework, they cannot track the association between individual firms (Diebold and Yilmaz, 2015).

Due to its approving attributes, DY framework has often been used to analyze the spillovers of the financial system through stock prices (Diebold and Yılmaz, 2014; Baruník and Křehlík, 2018) or the sovereigns through sovereign bonds (Alter and Beyer, 2014; Demirer et al., 2018) or sovereign CDS (credit default swaps) (Greenwood-Nimmo et al., 2019; Bratis et al., 2020). In addition, the DY framework has been applied to networks of different asset classes like equities (Baruník et al., 2016), bonds (Claeys and Vašíček, 2014), exchange rates (Bubák et al., 2011), commodity prices (Kang et al., 2017), cryptocurrencies (Moratis, 2021) or across asset classes (Wang et al., 2016; Kurka, 2019).

Focusing on the FIs network, Diebold and Yılmaz (2014) were the first to apply the framework to systemic risk modeling. They used the daily realized volatility time series to examine the daily sensitivity of the connections across major US FIs. They focused on the GFC 4 key events and illustrated the network on specific days with snapshots. Diebold and Yilmaz (2015) extend this analysis by examining the spillover channels of the realized volatility network of major US and European FIs during the GFC and the ESDC. They find the following results related to the two continent's bank system: before the Lehman Brothers' collapse, realized volatility spillovers primarily flowed from U.S. FIs to their European counterparts. However, after Lehman Brothers' bankruptcy in September 2008, the financial crisis evolved into a worldwide phenomenon, causing volatility spillovers and linkages across the Atlantic to become two-way, with a notable decrease in net spillover

from the U.S. to Europe. Demirer et al. (2018) applied a LASSO estimated VAR model to extend the investigated FIs and analyze a global bank network connectedness. They examined a network with the top 150 banks between 2003 and 2014. They conclude that global FI's spillovers have a strong geographic component.

I present the relevant literature of the VAR modelling in Chapter 2 for regularization methods in Chapter 3, and for estimation structural changes in networks in Chapter 6.

2.2 Diebold-Yilmaz framework

Networks based on time series are often represented in graphs, where the nodes (which represent time series) and edges (linkages between the series) are graphically displayed. One type of network is the weighted network, which allows for weights on the edges to represent stronger or weaker linkages between the nodes (time series). Direct networks allow for asymmetric linkages. In the econometrics and finance literature Diebold and Yilmaz (2009) were the first to show in a seminal paper that the forecast error variance decomposition (FEVD) of an estimated VAR(p) model of the network can be interpreted as a weighted directed graph. Diebold and Yilmaz (2009), Diebold and Yilmaz (2012), and Diebold and Yılmaz (2014)[3] suggest a unified framework for measuring linkages or spillovers between the time series. The framework is based on the concept that the shares of own and other time series' shocks to the total variance of a selected times series forecasts can be summarized in a spillover table, which I refer hereafter to as the DY spillover table. The most important merits of the DY framework are the following (Demirer et al., 2018):

1. The method does not require additional restrictions beyond those imposed for VAR(p) model estimation and identification.

2. It provides both the direction and magnitude of the measures. The estimated network will be directed.

3. The VAR(p) estimation step has been used for decades in economics[4], the estimated networks are easily interpretable.

4. The framework allows one to track spillovers between the time series at all levels of the network, from pairwise connections (micro level) to system-wide connections (macro level).

5. It allows for static and dynamic usage, which is relevant in financial applications where events can cause abrupt changes in the network.

6. It's a generalization of the Granger-causality based networks.

2.2.1 VAR model

The examination of multivariate time series is a prevalent statistical challenge across various fields, such as economics, finance (Engle and Granger, 1987; Johansen, 1988; Stock and Watson, 1988; Bai and Ng, 2002; Stock and Watson, 2002b; Bai, 2003; De Gooijer and Hyndman, 2006) and even in genomics, and neuroscience (Seth et al., 2015). In the last

---

[3]the framework is extremely popular in recent years, the three seminal papers citation is more than 2200 just in 2023

[4]impulse response functions (IRF) and FEVDs are commonly used in economics and econometrics

ten years, there has been increasing focus on networks to describe how these multivariate time series connect how they form a system (Eichler, 2007; Billio et al., 2012; Eichler, 2012).

Among the many available models, VAR methods are the most popular - since Sims (1980) introduced - in analyzing the temporal evolution of the network's time series, extracting short and long-term relationships and linkages, or examining the network dynamics through time (Eichler, 2007; Billio et al., 2012; Diebold and Yılmaz, 2014). In the VAR model, the temporal relationship in a multivariate series includes the autoregressive dependence within each series and the interdependence among different series, unlike in the case of univariate time series (Basu and Michailidis, 2015; Kock and Callot, 2015). This method is a powerful tool for analyzing the complex system of time series, and it allows for flexible modeling of the connections between the system's time series. Due to the increasing availability of time series data, vector autoregressive models have captured the attention of numerous theoretical and empirical researchers during the last two decades (Hecq et al., 2023; Krampe and Paparoditis, 2021)[5].

Generally, VAR has been used to identify the lagged linkages between time series. However, the contemporaneous relationships often provide additional information too (Basu and Michailidis, 2015). In the last twenty years, VAR was widely used for forecasting (De Gooijer and Hyndman, 2006)[6]. Besides that, based on the VAR-based tools, impulse response analysis, and forecast error variance decomposition, it can be used for policy analysis (Gertler and Karadi, 2015; Ramey, 2016; Nakamura and Steinsson, 2018). In the last decade, due to the easily computed and commonly applied FEVD, its role in the econometrics methods-based network analysis has increased.

2.2.2 Framework

Following the seminal papers Diebold and Yilmaz (2009) and Diebold and Yılmaz (2014), the network connectedness and spillover measures I derive are based on VAR(p) model coefficient matrix estimation (Sims, 1980) and its forecast error variance decompositions. The framework is based on the concept that, for every time series of the network, we can calculate the forecast error variance based on the estimated VAR(p) model coefficient and covariance matrix. This variance is related to its own and other time series shocks. Due to the VAR(p) model identification, the shares of own and other time series' shocks can be calculated. In the last step of the process, the forecast error variance decompositions can be summarized in a spillover table, which we refer to hereafter as the DY spillover table. In the next paragraphs, I present the Diebold-Yilmaz framework in detail.

The first step of the framework is to specify and estimate a stationary VAR(p) model for the J time series of the network based on the following equation:

$$Y_t = \sum_{i=1}^{p} \beta_i Y_{t-i} + \epsilon_t \tag{1}$$

where Y_t is a $J \times 1$ vector of the time series, β_i is an $J \times J$ autoregressive coefficient matrix, and lastly ϵ_t is an $J \times 1$ vector of error terms. It has a zero mean with a Σ

---

[5]In addition to economics and finance, VAR models have recently become standard approaches in genomics (Shojaie and Michailidis, 2010) and in neuroscience for understanding the complex patterns between brain regions and areas (Friston, 2009; Smith, 2012).

[6]The method can forecast the future of multiple time series dynamically

covariance matrix. No intercept is included in Eq. (1), without loss of generality, I assume, similar to Basu et al. (2019), that all the J time series are mean centered. The VAR(p) process is assumed to be stable and stationary (Lütkepohl, 2013), while the covariance matrix Σ is needed to be positive definite with bounded largest eigenvalue (Melnyk and Banerjee, 2016).

To estimate the DY framework's most important element, the DY spillover table, I need to estimate the coefficient matrices β_1, β_2, ... β_p and the error covariance matrix Σ efficiently. The β_i coefficient matrices reveal the temporal dependence between the time series and Σ reveals the contemporaneous linkages among them (Han et al., 2015; Davis et al., 2016).

The starting point for the DY framework (Diebold and Yilmaz, 2012) is to transform the time series of the VAR(p) in Eq. (1) into its vector moving average (VMA) representation using the Wold representation theorem (Diebold and Yilmaz, 2012; Gabauer et al., 2020) to get Eq. (2):

$$Y_t = \sum_{i=0}^{\infty} A_i \epsilon_{t-i} \tag{2}$$

where A_i is an $J \times J$ moving average coefficient matrix. Based on the Wold's theorem A_i is given by the following recursion $A_i = \sum_{j=1}^{p} \beta_j A_{i-j}$ where $A_j = 0$ for $j < 0$ and A_0 is an identity matrix.

As Diebold and Yilmaz (2012) emphasized, the calculated moving average coefficients and the estimated error covariance matrix (or its nonlinear transformations such as impulse response functions (IRF) or forecast error variance decompositions) are the key to understanding the dynamics of the time series network. With the estimated VAR(p) coefficient and covariance matrix, I can estimate the H-step ahead forecast error variance $(FEV(H))$[7] on the following way:

$$FEV(H) = MSE(Y_t(H)) = E[(Y_{t+H} - Y_t(H)) * (Y_{t+H} - Y_t(H))^{'}] = \sum_{h=0}^{\infty} A_h \Sigma A_h^{'} \tag{3}$$

where $Y_t(H)$ is the optimal predictor for known A_i, A_h is the hth coefficient of the VAR(p) MA representation coefficient in Eq. (2), E() stands for expected value and prime symbol denotes the transposing operation (Diebold and Yilmaz, 2012; Buse and Schienle, 2019).

The above calculated FEV allows me to calculate the fraction of the H step-ahead error variance in forecasting Y_i $(Y_i(H))$ that is due to shocks to other time series such as Y_j (Diebold and Yilmaz, 2012), to which I will hereafter refer as a spillover between Y_i and Y_j. Generally, in the DY framework, the FEVD of the VAR(p) model gives the measures of spillovers between the time series. Unfortunately, the calculation of the FEV requires orthogonal innovations. However, the VAR innovations are generally contemporaneously correlated (Diebold and Yilmaz, 2012; Diebold and Yilmaz, 2014; Basu and Michailidis, 2015).

There are two widely used approaches in the early DY framework-related papers for obtaining the variance decomposition. The first method uses the Cholesky factor orthogonalisation of the covariance matrix Σ which generates orthogonalised innovations. The

---

[7]which is equal to the H-step ahead mean square error of the Y_t (MSE($Y_t(H)$)

weakness of this decomposition is that its results order dependent FEVDs (IRFs and FEVDs are different if we change the order of the network's time series) (Diebold and Yilmaz, 2012; Fengler and Gisler, 2015). The other approach uses the generalized VAR framework, which was introduced in the seminal papers by Koop et al. (1996) and Pesaran and Shin (1998). This framework allows correlated shocks. As a result, this second method produces an order-independent FEVD. However, the results on the individual entries of the DY spillover table will be less interpretable. Applying the second method is more widespread in empirical DY network studies. In my simulation and empirical analysis, I use the generalized FEVD of Koop et al. (1996) and Pesaran and Shin (1998) to avoid the Cholesky factor orthogonalization so that the results will be order independent. The following equation shows the calculation of the generalized FEVD:

$$\theta_{ij}^g(H) = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' A_h \Sigma e_j)^2}{\sum_{h=0}^{H-1} (e_i' A_h \Sigma A_h' e_j)} \tag{4}$$

Where σ_{jj} is the j-th diagonal element of the error term's covariance matrix Σ, A_h is the moving average coefficient matrix multiplying the h-lagged shock vector in the Wold's moving average representation (Eq. (2)) and e_i is a selection vector. That means i-th element of the e_i vector is unity, and others are zeros.

The numerator in Eq. (5) of $\tilde{\theta}_{ij}^g(H)$ represents the contribution of shocks in variable Y_j to the H-step FEVD of time series Y_i. The denominator is the forecast error variance of the time series Y_j. Unfortunately, the sum of the network time series' contribution to the forecast error variance is not necessarily one because, generally, FEVD of the shock terms are not orthogonalized (Diebold and Yilmaz, 2012). Normalization is therefore required, which I calculate in the following way:

$$\tilde{\theta}_{ij}^g(H) = \frac{\theta_{ij}^g(H)}{\Sigma_{k=1}^J \theta_{ik}^g(H)}. \tag{5}$$

In that way, the information in the FEVD can directly be used for the DY spillover matrix (Diebold and Yılmaz, 2014). The Koop et al. (1996) and Pesaran and Shin (1998)'s generalized FEVD approach allows us to gain insight into the direction of spillovers between the time series. Besides that, by aggregating the FEVD on different levels of the network, I can conclude both on macro and micro levels.

The generalized FEVD is used to construct the several systemic/network connectedness measures of the DY framework (Diebold and Yilmaz, 2012; Diebold and Yılmaz, 2014). Firstly, the sum of directional spillovers to time series Y_i from all other time series (FROM spillover index $S_{i \leftarrow \bullet}^g$) is defined with the following equation:

$$S_{i \leftarrow \bullet}^g(H) = \frac{\sum_{k=1, k \neq i}^J \tilde{\theta}_{ik}^g(H)}{\sum_{k=1}^J \tilde{\theta}_{ik}^g(H)} * 100\%. = \sum_{k=1, k \neq i}^J \tilde{\theta}_{ik}^g(H) * 100\%. \tag{6}$$

Second, I am interested in the sum of the shocks transmitted by time series Y_i tO other time series (TO spillover index $S_{\bullet \leftarrow i}^g$):

$$S_{\bullet \leftarrow i}^g(H) = \sum_{k=1, k \neq i}^J \tilde{\theta}_{ki}^g(H) * 100\%. \tag{7}$$

The third relevant measure is the NET spillover index (S_i^g, Eq. (8)), which calculates the difference between the gross transmitted (TO) and received (FROM) shocks from all

other time series:

$$S_i^g(H) = S_{\bullet \leftarrow i}^g(H) - S_{i \leftarrow \bullet}^g(H) \tag{8}$$

Finally, at the macro level of the network analysis, the system-wide spillover index (SUM spillover index S_{total}^g) offers information about the average influence one time series has on all other time series, regardless of the direction on the following way:

$$S_{total}^g(H) = \frac{1}{J} \sum_{i,k=1, i \neq k}^{J} \tilde{\theta}_{ik}^g(H) \tag{9}$$

In other words, the total spillover index is the sum of all the off-diagonal elements of the generalized FEVD matrix relative to the number of time series considered in the VAR(p) model (Diebold and Yilmaz, 2012; Gabauer et al., 2020). It summarises the measure of how much of the FEV of the time series can be explained by spillovers from other time series (Fengler and Gisler, 2015). A large (small) total spillover index means that the average propagation of a shock in one time series to all others in the system is high (low) and, thus, the systemic risk of the network is high (low) (Diebold and Yılmaz, 2014; Gabauer et al., 2020).

I can further decompose the directional spillovers between two-time series into net pairwise directional spillovers. This decomposition allows me to determine the spillover linkages between two specified time series. Pairwise NET spillover index (NETP, S_{ij}^g) between time series Y_i and Y_j is the difference between the gross shocks transmitted from Y_i to Y_j and those transmitted from Y_j to Y_i. I calculate it in the following way:

$$S_{ij}^g(H) = \left(\frac{\tilde{\theta}_{ji}^g(H)}{\Sigma_{i,k=1}^J \tilde{\theta}_{ik}^g(H)} - \frac{\tilde{\theta}_{ij}^g(H)}{\Sigma_{i=1,k}^J \tilde{\theta}_{ik}^g(H)} \right) * 100\%. \tag{10}$$

As Diebold and Yılmaz (2014) and Gabauer et al. (2020) pointed out, a positive (negative) value of the net pairwise directional spillovers implies that time series Y_j is dominate (is dominated) by time series Y_i. To create a full comparison between the indicators, in my empirical analysis, I use both macro (SUM) and micro (FROM, TO, NET, and NETP) level spillover indices.

For a J-variable VAR(p) system and a prespecified forecast horizon, H, the i,j th element of the H steps-ahead FEVD matrix Θ, which is a $J \times J$ matrix and I denoted with $\Theta_{ij}(H)$, express the percentage of the H-steps-ahead forecast uncertainty associated with time series Y_i that can be shocked by time series Y_j (Greenwood-Nimmo and Tarassow, 2022). Hereafter, I refer to this Θ matrix as a DY spillover matrix. During turbulent periods, an increase in the average pairwise spillovers from one time period to another corresponds to an increase in the total spillover index of the system. In the DY empirical literature, abrupt increases in total spillover index are frequently interpreted in relation to systemic shocks (Diebold and Yılmaz, 2014; Greenwood-Nimmo and Tarassow, 2022). Following Diebold and Yılmaz (2014) and Diebold and Yılmaz (2014), almost all the researchers apply a rolling window estimation because it is a simple approach to analyze the dynamics of the linkages between the network of time series.

The DY framework is closely linked to recently proposed measures of systemic risk. However, the MES (Acharya et al., 2012) and CoVaR (Tobias and Brunnermeier, 2016) measures allow only to quantify the linkages between individual time series and the system in one direction or the other, unfortunately not simultaneously in both directions.

An alternative method involves concentrating on the VAR estimated coefficients within a Granger network (GN), as proposed by Billio et al. (2012) in their seminal paper. This approach employs pairwise Granger causality to assess the connections in the network. However, the GN is less granular than the FEVD-based approach because it solely examines whether the VAR coefficients are zero or nonzero without considering the covariance matrix. Hence, the DY framework exhibits two advantageous characteristics compared to the GN. Firstly, DY networks assess both the direction and magnitude of spillovers among time series (while GNs are directed but lack weighting[8]). Secondly, as the DY spillover table originates from a high-dimensional VAR estimation, it remains robust to the issue of omitted variables, a challenge that significantly impacts GNs derived from a series of bivariate regression models (Diebold and Yılmaz, 2014).

To illustrate the Diebold-Yilmaz framework, I calculate the spillover table and the FROM, TO, NET, and SUM spillover indices from a low-dimension system ($J = 5$). I consider that that data-generating process (DGP) is a VAR(1) model, and the coefficient matrix (B) is sparse, with one dominant time series:

$$
B = \begin{bmatrix}
0.40 & 0 & \mathbf{0.08} & 0 & 0 \\
0 & 0.40 & \mathbf{0.07} & 0 & 0 \\
0 & 0 & 0.40 & 0 & 0 \\
0 & 0 & \mathbf{0.08} & 0.40 & 0 \\
0 & 0 & \mathbf{0.09} & 0 & 0.40
\end{bmatrix}
\tag{11}
$$

The diagonal elements in the coefficient matrix refer to the autoregressive dependence within the time series. Based on the third column of the matrix, the third time series leads to the other ones.

To illustrate the role of the covariance matrix (Σ) in the Diebold-Yilmaz framework, I consider two different structures. In the first case, both the covariance matrix is the identity matrix[9]:

$$
\Sigma = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\tag{12}
$$

In the second case, I consider a is a Toeplitz-type structure with a $\rho = 0.3$ parameter[10]:

$$
\Sigma = \begin{bmatrix}
1.00 & 0.20 & 0.04 & 0.01 & 0.00 \\
\mathbf{0.20} & 1.00 & 0.20 & 0.04 & 0.01 \\
\mathbf{0.04} & 0.20 & 1.00 & 0.20 & 0.04 \\
\mathbf{0.01} & 0.04 & 0.20 & 1.00 & 0.20 \\
\mathbf{0.00} & 0.01 & 0.04 & 0.20 & 1.00
\end{bmatrix}
\tag{13}
$$

In this case, the covariance matrix is dense, and the entries exponentially decrease in

---

[8]some newer papers constructed weighted Granger-network based on the F-statistics of the Granger-test (Yun et al., 2019; Zhao et al., 2022), but the common empirical application in finance is still the unweighted GN network (Diebold and Yilmaz, 2023)

[9]it's a common assumption in the econometrics literature, and Diebold and Yilmaz (2009) assumed the same

[10]it's a common assumption in the machine learning literature, Basu and Michailidis (2015) and Hecq et al. (2023) assumed the same

the distance from the diagonal.

Based on the coefficient and covariance matrix, I calculate the theoretical spillover table and the spillover indices by applying the Eqs. (1-9). The following Diebold-Yilmaz spillover table (Table 1) represents the first system where the B is sparse with one dominant time series, and the Σ is the identity matrix:

Table 1: Diebold-Yilmaz spillover table for sparse coefficient matrix B and sparse covariance matrix Σ

	TS1	TS2	TS3	TS4	TS5	FROM
TS1	98.96	0.00	**1.04**	0.00	0.00	1.04
TS2	0.0	99.20	**0.80**	0.00	0.00	0.80
TS3	**0.00**	**0.00**	100.00	**0.00**	**0.00**	**0.00**
TS4	0.00	0.00	**1.04**	98.96	0.00	1.04
TS5	0.00	0.00	**1.31**	0.00	98.69	1.31
TO	0.00	0.00	**4.20**	0.00	0.00	**4.20**
NET	**-1.04**	**-0.80**	**4.20**	**-1.04**	**-1.31**	0

Note: *The coefficient matrix (B) is sparse with one dominant time series, and the covariance matrix (Σ) is the identity matrix.*

The spillover table shows that the third time series transmits the shocks to the other parts of the system (TO = 4.2), and it doesn't receive any shock (FROM = 0). It plays a dominant role in a system as a shock transmitter. There are no connections between the other time series of the network. The total spillover index is very low (SUM = 4.2/5 = 0.8%) due to the sparse coefficient and covariance matrix. In the second case, when the coefficient matrix is the same, but the covariance matrix is dense, the Diebold-Yilmaz spillover table is the following:

Table 2: Diebold-Yilmaz spillover table for sparse coefficient matrix B and dense covariance matrix Σ

	TS1	TS2	TS3	TS4	TS5	FROM
TS1	94.42	4.09	**1.43**	0.06	0.00	5.58
TS2	3.63	90.65	**5.49**	0.22	0.01	9.35
TS3	**0.15**	**3.69**	92.32	**3.69**	**0.15**	**7.68**
TS4	0.01	0.23	**5.85**	90.30	3.61	9.70
TS5	0.00	0.08	**1.72**	4.12	94.08	5.92
TO	3.79	8.09	**14.48**	8.10	3.77	**38.24**
NET	**-1.80**	**-1.25**	**6.80**	**-1.61**	**-2.15**	0

Note: *The coefficient matrix (B) is sparse with one dominant time series, and the covariance matrix (Σ) is a Toeplitz-type matrix with $\rho = 0.3$.*

The third time series dominated the others in this case, too, but the FROM spillover value is not equal to zero (FROM = 7.68) due to the non-diagonal covariance matrix. Besides the connections between the other part of the network, all the time series has positive TO and FROM values. Regarding these network characteristics, the total spillover is higher than in the earlier case (SUM = 38.24/5 = 7.6%).

Most of the articles represent the network with graphs (besides the spillover table) to interpret the connections between the time series and the key participants of the network

more easier[11]. In the following two graphs (Fig. 1), I illustrate the two networks with sparse and dense covariance matrices:

Figure 1: Theoretical Diebold-Yilmaz networks for sparse and dense Σ
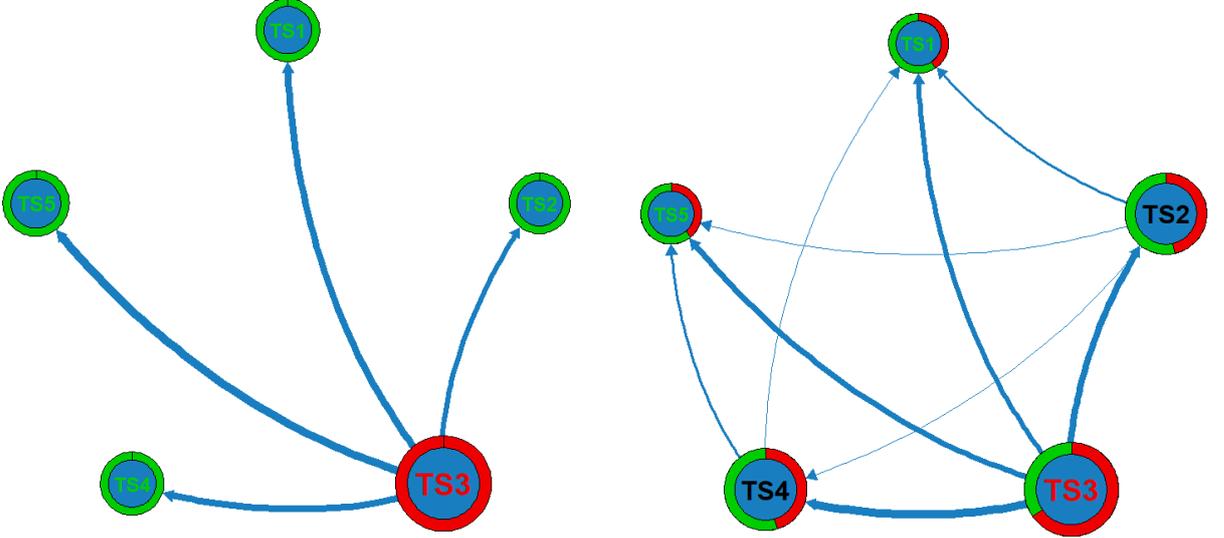
Notes: First (Second) DGP-based network is in the left (right) chart. The coefficient matrix (B) in both cases is sparse with one dominant time series, the covariance matrix (Σ) is the identity matrix in the first case, and a Toeplitz-type matrix with $\rho = 0.3$ in the second case. Nodes represent the time series. Node sizes show the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the time series. Edge directions mark the net spillover indices; their thickness represents the magnitudes. All of the edges are shown on the charts. Nodes with green letters represents time series where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$.

To get a relevant conclusion from these figures, it is essential to clarify the interpretation of the graphs. These graphs encompass a wealth of information. The nodes represent the times series. The pie charts' colors around the nodes represent the proportion of shocks transmitted (TO spillover value marked with red color) and received (FROM spillover value marked with green color) by the time series. The last information related to the nodes is the SUM value, which is indicated by the size of the nodes. The edges represent the NET spillover indices of the time series, and their thickness represents their magnitudes. As in the 5-node network, there are just 10 NET pairwise connections; I present all of them. Nodes with green letters represent time series where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. These graphs aim to identify the key participants of the system (both big shock transmitters and receivers). The SUM indicator measures the centrality of a given time series in the system, regardless of the direction of the shock transmission. Nodes with red (green) letters represent the biggest shock transmitters (receivers).

In the first graph, it's clearly visible the dominance and the shock transmitting role the third time series. The lack of connections between the other time series shows the sparsity of the coefficient and the inverse covariance matrices. In contrast to this, the network on the second graph is dense due to the nonlinear relationship between the B

---

[11]for more than 20 time series, the spillover table is less interpretable

and Σ matrices and the Diebold-Yilmaz spillover table, and the dense structure of the covariance matrix.

These examples with the spillover tables and the network graphs illustrate the non-linear relationship between the B and Σ and Θ and the role of the coefficient and the covariance matrix in the Diebold-Yilmaz framework.

3 Chapter 3. Regularization methods for financial networks

In this chapter, I discuss the machine learning regularization methods, specifically the LASSO and adaptive LASSO methods. In the first part of the chapter, I summarize the advantages of LASSO-based VAR models in high-dimension time series environments. Next, I compare two LASSO-based VAR estimators (row-wise LASSO and penalized maximum likelihood-based LASSO). At the end of this chapter, I explain the connection between the regularization methods and the DY framework.

3.1 VAR estimation for high-dimension time series

In the last two decades, econometrics, economics, and finance have seen a rapid increase in applications involving high-dimensional time series. Different approaches have been explored to address the issue of the dimensionality of the system (large numbers of time series relative to the time series length). The most common approaches are different factor models (Stock and Watson, 2002a; Forni et al., 2000), Bayesian VAR models (Bańbura et al., 2010), reduced-rank VAR models (Carriero et al., 2011), and regularized VAR models (Kock and Callot, 2015; Demirer et al., 2018). Around the turn of the millennium, dynamic factor models were considered to be quite successful in handling and forecasting the large panels of time series data (Forni et al., 2000; Stock and Watson, 2002b; Forni et al., 2005). In a seminal paper, Stock and Watson (2002b) contributed to the earlier literature significantly by employing factors to reduce the dimension of the system and achieve more accurate predictions on macro variables. Factor models assume that the bulk of dynamic connections between the time series can be represented by a few common factors (Bai, 2003). For a deeper survey on high-dimension time series factor models, see Stock and Watson (2006), Bai et al. (2008), and Stock and Watson (2011).

In addition to the factor models, shrinkage methods and sparse modeling are other concepts to handle the high-dimensionality in econometric time series, and it has been popular for a long time in the fields of statistics (Tibshirani, 1996; Zou, 2006; Yuan and Lin, 2006). Shrinkage methods assume a specific structure on the parameter vector instead of common structures between the time series. Compared to factor models, the sparse VAR models have several natural advantages. First, sparse VAR models allow straightforward variable-to-variable relationship (impulse response and forecast error variance decomposition) analysis, which is not feasible in the factor modeling because the variables are represented by the corresponding factors (Kock and Callot, 2015). Second, they are more easily interpretable in many economic applications. Moreover, the third, dynamic factor models estimation is typically a 2-step estimation (the first step is a dimension reduction, and the second step for low dimensional time series forecasting). Sparse VAR models can be estimated in one step, which may lead to greater efficiency (Basu and Michailidis, 2015). Furthermore, it appears natural to assume that a relatively small number of variables can appropriately model the most relevant interactions and linkages among the time series (Kock and Callot, 2015; Hecq et al., 2023).

In the low-dimensional setting (in this case there are sufficient degrees of freedom for efficient VAR estimation), the VAR model in Eq. (1) can be reformulated as a multivariate regression problem Lütkepohl (2013) and estimated directly (equation-by-equation estimation) by the OLS method[12]. However, this estimation has an overfitting issue as

---

[12]It is the most common approach within the DY framework too (Diebold and Yilmaz, 2023)

the system's dimension becomes larger (Basu and Michailidis, 2015; Buse and Schienle, 2019).

Estimating high dimensional VAR models poses practical challenges to researchers: On one hand, a problem of dimensionality is caused by incorporating a huge number of time series and modeling higher-order autoregressive processes (Kock and Callot, 2015; Davis et al., 2016; Basu et al., 2019). Second, the temporal dependence structure in the sparse VAR model gives rise to some theoretical challenges (Basu and Michailidis, 2015). It is a challenging issue to determine which variables and (their) lags are relevant when the sample size is moderate (relative to dimensionality and lag number) (Nicholson et al., 2017; Hecq et al., 2023).

The parameter space of VAR models grows quadratically as the number of variables increases, resulting a rapid exhaustion of the available degrees of freedom (Kock and Callot, 2015). A VAR(p) model, when it is fully parameterized, has J^2p AR (autoregressive) parameters that need to be estimated, where J is the number of series of the system and p is the common lag length. It can be a severe problem if many time series are included in the system. It's commonly undertaken to ensure satisfactory modeling of the network dynamics (to decrease the endogeneity probability). When the VAR lag $p = 3$, and the number of variables (time series) $J = 50$, there are 7500 regression parameters that we need to estimate from the data. As the time series length (T) is typically relatively small for many economic and financial applications (to avoid the problem caused by structural changes), the researcher might find themselves in a scenario where the quantity of parameters significantly exceeds the number of observations (Kock and Callot, 2015).

This heavy parameterization is a serious drawback of the traditional VAR estimation, and this limits its applicability for economics and financial high-dimension time series modeling (Hecq et al., 2023). The existing data lack adequate information for efficient parameter estimation using standard OLS and maximum likelihood methods. This leads to noisy parameter estimates, potential instability in predictions, and challenges in providing clear interpretations of temporal dependence. Besides that, if $T < J^2p$, equation-by-equation least squares is not even feasible (Basu and Michailidis, 2015; Kock and Callot, 2015). Addressing the estimation of VAR model parameters is essential for effectively modeling high-dimensional time series data.

Recent advances in high-dimensional time series modeling have established that estimating a VAR model with relatively few samples is possible even when least squares estimation is not appropriate (Kock and Callot, 2015; Basu and Michailidis, 2015; Davis et al., 2016). It needs to impose a special structure on the coefficient matrix to estimate it consistently. It is generally believed that, for most economic and financial applications (Fan et al., 2011b; Giannone et al., 2021), the true model of the system of time series is sparse, only a small, unknown subset of the variables have significantly non-zero coefficients, and all the other variables have negligible (or even exactly zero) coefficients (Kock and Callot, 2015; Demirer et al., 2018; Hecq et al., 2023). Therefore, it is preferable to fit a sparse VAR model, with many AR parameters set to zero to more accurately approximate the sparse data generating process. An additional advantage of sparse models is that they allow for an ultra-high dimensionality of time series (Fan et al., 2011a, 2014).

Estimators with sparsity properties have demonstrated effective performance across diverse fields, such as economics (Fan et al., 2011b), macroeconomics (Korobilis, 2013), and biostatistics (Friedman, 2012). Over the last ten years, a broad range of studies has been dedicated to the sparse modeling of the VAR process, mostly based on recent advances in variable selection and shrinkage theory from Tibshirani (1996), Zou (2006), and

Yuan and Lin (2006), to effectively handle the above-mentioned dimensionality problem. The academic papers refer to estimators that achieve sparse estimation as regularized methods. VAR model coefficient matrix estimation using regularization methods has become a popular alternative to factor modeling for systems with high-dimensional time series (Bai et al., 2008; Kock and Callot, 2015).

3.2 Regularized estimation of VAR: LASSO and adaptive LASSO

The regularized (penalized) regression approach in cross-section cases leads to sparse estimates of the regression coefficients, achieves both efficient prediction and consistent variable selection, and besides that allows the dimensionality of the high-dimension system to be much larger than the sample size. The two most common methods to address the high-dimension problem are to apply pure shrinkage (L2-penalization) or to estimate with pure selection (as with traditional information criteria: Akaike information criterion (AIC), Bayesian information criterion (BIC) or with L1-penalization).

To understand the equation-by-equation regularized estimator consider the penalized estimation problem Eq. (14) for VAR(p):

$$\underset{\beta ij}{\operatorname{argmin}}\left[\sum_{t=p+1}^{T}(y_{jt}-\sum_{i=1}^{p}\sum_{j=1}^{J}\beta_{ij}y_{j,t-i})^2 + \lambda_j\sum_{i=1}^{p}\sum_{j=1}^{J}\beta_{ij}^q\right] \tag{14}$$

where β_{ij} denotes the i, jth element of the stacked version of VAR(p) coefficient matrix B[13] and λ_j is the regularization parameter[14] for the jth equation of the VAR model (Kock and Callot, 2015; Basu and Michailidis, 2015; Demirer et al., 2018; Basu et al., 2019). The first part of the objective function represents the OLS loss function, and the second is the penalty term for the coefficients. Concave penalty functions produce selection, whereas smooth convex penalties[15] produce shrinkage.

The L2-penalized estimation problem is the following (Eq. (15)) for the ridge estimator:

$$\underset{\beta ij}{\operatorname{argmin}}\left[\sum_{t=p+1}^{T}(y_{jt}-\sum_{i=1}^{p}\sum_{j=1}^{J}\beta_{ij}y_{j,t-i})^2 + \lambda_j\sum_{i=1}^{p}\sum_{j=1}^{J}\beta_{ij}^2\right] \tag{15}$$

Due to the convex penalty, this estimator only shrinks the coefficients towards zero. Unfortunately, it doesn't solve the heavy parameterization problem of VAR(p) coefficient matrix (B) estimation.

The most prominent algorithm among penalized estimators is the LASSO (least absolute shrinkage and selection operator), blending the two core concepts; it shrinks and selects together. The LASSO approach was proposed by Tibshirani (1996). Theoretical aspects of this method have been thoroughly examined since the publication of the seminal paper; see the following references from machine learning literature Zou (2006),

---

[13]I refer Basu et al. (2019) and Krampe and Paparoditis (2021) for a more detailed explanation for the connection between β_{ij} and B

[14]the literature often refers this as a hyperparameter of the method, because during the estimation process, it's necessary to optimize

[15]where $q = 2$, the ridge regression estimator

Meinshausen and Bühlmann (2006), Bickel et al. (2009).

LASSO solves the penalized regression problem with $q = 1$. The LASSO method incorporates a penalty based on the absolute value of the parameters into the objective function of least squares as the estimation problem in Eq. (16) shows:

$$\underset{\beta_{ij}}{\mathrm{argmin}}\left[\sum_{t=p+1}^{T}(y_{jt} - \sum_{i=1}^{p}\sum_{j=1}^{J}\beta_{ij}y_{j,t-i})^2 + \lambda_j\sum_{i=1}^{p}\sum_{j=1}^{J}|\beta_{ij}|\right] \qquad (16)$$

The L1 penalty serves a dual purpose. Initially, by incorporating this penalty into the objective function, it enables the estimation of β even when the number of coefficients in the VAR model surpasses the time series length. Subsequently, it induces sparsity in the estimated autoregressive parameters B by enforcing some elements of β exactly equal to zero. The value of the penalty clearly determines the amount of this selection: the larger of the regularization parameter λ, the sparser the coefficient matrix (B). The concept behind the LASSO estimation is to shrink the OLS estimated parameters (elements of B matrix) towards zero to reduce variance (Kock and Callot, 2015; Demirer et al., 2018).

In recent years, LASSO has become popular in econometrics since it performs automatic variable selection by setting weaker coefficients exactly to zero and parallel with this, making consistent parameter estimation (Kock and Callot, 2015). An additional advantage is that the resulting parameter vector (elements of B matrix) is sparse and easier to interpret. LASSO estimation is computationally feasible (Nicholson et al., 2017, 2020). It is widely applied and most useful in high-dimension VAR(p) settings where the number of coefficients is much larger than the number of observations $(J^2p \geq T)$. As regards the numerical solution of the estimation, the parameters of equation-by-equation LASSO can be efficiently and quickly calculated by coordinate descent algorithms (Lin and Michailidis, 2017). An important limitation of the LASSO arises when the model faces up with highly correlated variables (Zou and Hastie, 2005). In such cases, the LASSO tends to pick only one or just a few coefficients from the B matrix and shrinks the rest of them toward zeros. Hereafter, I refer to the estimation approach as a LASSO-VAR model if the coefficient matrix (B) is estimated with L1-penalization.

In the last two decades, much attention has been devoted to alternative penalized estimators beyond LASSO and Ridge. Other prominent methods are the smoothly clipped absolute deviation algorithm (SCAD) (Fan and Li, 2001), the adaptive LASSO (Zou, 2006), the Dantzig selector (Candes and Tao, 2007), the minimum concave penalty method (Zhang, 2010) and the square root LASSO (Belloni et al., 2011), and the sparse group LASSO (Poignard, 2020). Due to favorable estimation and prediction features, the LASSO and its many extensions are nowadays standard tools in high-dimensional analysis and high-dimensional time series modeling, too (Basu and Michailidis, 2015). Among the alternative penalized regression approach methods, the adaptive LASSO has received the greater attention (Kock and Callot, 2015).

As Eq. (16) shows, LASSO penalizes all parameters equally. If it were feasible to assign a higher penalty to the truly zero parameters (β_{ij}, which are zeros in the data generating parameters) compared to the non-zero ones, one would expect a better estimation and forecast performance. Zou (2006) improved the LASSO regression by introducing an additional weight parameter. The adaptive LASSO method is a two-step algorithm that uses a first-step estimator result (usually the Ridge) to weight the lagged time series. If the first-step estimator classifies one of the estimated parameters as zero $\beta_{ij} = 0$, it is not

included in the second step of the estimation. This concept results in a smaller size of the problem in the second phase. It also has sparse solutions and an even more efficient estimation algorithm than the original LASSO. Shortly, it generalizes the popular LASSO method.

The following equation shows the adaptive LASSO estimation problem:

$$\operatorname*{argmin}_{\beta ij} \left[\sum_{t=p+1}^{T} (y_{jt} - \sum_{i=1}^{p} \sum_{j=1}^{J} \beta_{ij} y_{j,t-i})^2 + \lambda_j \sum_{i=1}^{p} \sum_{j=1}^{J} w_{ij} |\beta_{ij}| \right] \qquad (17)$$

where ($w_{ij} = \frac{1}{|\beta_{ij}^{ridge}|}$) is the adaptive penalty term.

A nice feature of the adaptive LASSO is what Fan and Li (2001) call the oracle property, which means that the adaptive LASSO correctly identifies the zero components of the coefficient matrix with probability tending to 1 (setting all zero parameters in the data generating process exactly equal to zero through the estimation (Zou, 2006)). This property of the algorithm is true both in cross-section regression and high-dimension VAR(p) settings (Kock and Callot, 2015).

3.3   Theoretical properties of LASSO-VAR model

In the early 2000s, the theoretical properties of regularized estimates in cross-section cases under high-dimensional scaling were investigated in numerous studies (Basu and Michailidis, 2015). Considerable effort has been dedicated to establishing error bounds for penalized methods, ensuring consistency for both predictions (Meinshausen and Bühlmann, 2006) and both parameter estimation in high-dimension data sets (Bickel et al., 2009). Even though much progress has been made in this direction, most of these advances are achieved assuming that the samples are independent and identically distributed (IID errors in the linear regression model). Most of the advances in the penalized estimation literature are valid only in this framework and setting (Huang et al., 2008; Meinshausen and Bühlmann, 2006). Theoretical analysis of the penalized estimates in a time series context, characterized by both temporal and cross-sectional dependence in the data, requires further completion (Song and Bickel, 2011; Basu and Michailidis, 2015).

The first two papers that expanded the penalized methods results to high-dimension time series setting were Wang et al. (2007) and Hsu et al. (2008). However, these studies only consider cases where the number of variables is smaller or equal than the sample size. Wang et al. (2007) studied first the L1-penalty efficiency in simultaneously estimating the regression coefficients and selecting the relevant lags in a linear regression model where the error term was allowed to follow an autoregressive process. The paper finds that LASSO consistently selects the relevant variables, and a modified version of LASSO has an oracle property too. In a high-dimension time series setting, Hsu et al. (2008) proposed first a LASSO-VAR model where the coefficient matrices are estimated with L1-penalty, inducing a sparse output for these matrices. Besides that, Chan and Chen (2011) analyzed the oracle properties and model selection consistency for lag selection if the estimated model is from the ARMA "family". The first paper that let the number of variables (the number of regressors or the number of time series) increase with the sample size – to establish the theoretical background of the models – was by Nardi and Rinaldo (2011). However, their focus was limited only to autoregressive (AR) models. They have shown favorable results for model selection consistency (efficient lag selection) and prediction

efficiency for these models.

Recently, the theoretical properties of penalized estimation in high-dimensional time series (VAR setting) have begun to be studied extensively, mostly in machine learning literature. Regularized estimation theory for VAR models is now well established;, see among others, Song and Bickel (2011), Basu and Michailidis (2015), Han et al. (2015), Kock and Callot (2015) and Davis et al. (2016). These papers let the number of variables increase with the sample size (assuming J grows with T, possibly at a faster rate), similar to Nardi and Rinaldo (2011) in the AR setting. Simply, that means if the number of non-zero entries of the B matrix is much less than the dimension of parameters J^2p, then much fewer sample is sufficient for consistent and efficient estimation of the sparse coefficient matrix B.

Song and Bickel (2011) considered first deeply a regularized VAR estimation problem and applied LASSO and group-LASSO for estimating the parameters of the coefficient matrix B. They derived oracle inequalities[16] for the estimation accuracy. However, their analysis focused solely on scenarios where the number of candidate variables was a function of the sample size while keeping the number of relevant variables constant. Moreover, they employed a strong assumption on the dependency structure in the VAR (Basu and Michailidis, 2015; Davis et al., 2016). Basu and Michailidis (2015) and Kock and Callot (2015) were the first to study the LASSO's properties and its variants in high-dimensional Gaussian VAR models from a theoretical perspective. Basu and Michailidis (2015) established the model selection consistency (chooses the relevant variables as the number of observations increases) for the LASSO estimator in high-dimensional time series settings with weak dependence (non-diagonal covariance matrix), extending Song and Bickel (2011). They established consistency for the L1-penalized OLS loss function and the maximum likelihood estimator of the coefficients in a high-dimension time series setting. They calculate the error bounds of parameter estimation by building on the prior work (Loh and Wainwright, 2011). The approach exploits the spectral properties of a general VAR model, providing insights into the dependency structure of the process. Kock and Callot (2015) derived the oracle property of the adaptive LASSO for high-dimension VAR models. Han et al. (2015) extended the Dantzig selector of Candes and Tao (2007) to be applicable to weakly sparse[17] high-dimension time series (VAR) models. Davis et al. (2016) proposed a special two-step estimator based on partial coherence and the t-test statistic.

These papers required quite restrictive assumptions on the innovation covariance matrix (Gaussianity in the error term) when investigating the theoretical properties of L1-penalized estimators in high-dimension time series models. More recently, researchers relaxed this assumption in several ways. Medeiros and Mendes (2016a) considered conditionally heteroskedastic errors for the VAR(p) model. They showed that the adaptive LASSO is model selection consistent and retains oracle property even when the errors are non-Gaussian. Xue and Taniguchi (2020) derived model selection consistency for a modified version of the LASSO in VAR(p) models with long-memory innovations. Masini et al. (2022) focused on weakly sparse high-dimensional VAR models under very general conditions, with heavy-tailed, weakly dependent innovations. They derived theoretical error bounds for the LASSO estimator in these settings. Adamek et al. (2023) developed theoretical results for desparsified LASSO in near-epoch dependent VARs.

---

[16]assumptions about the data generating process to earn the estimation method the oracle property

[17]only weakly sparse matrix means, that it contains no zeroes, but its entries follow a geometrically decaying pattern (Adamek et al., 2023)

As the above references show, the L1–penalized high-dimension VAR modeling approach has the advantage of performing model selection and parameter estimation simultaneously, and it can also be applied under the "large-J-small-T" settings. The use of penalized methods for estimating high-dimension vector autoregressive (HD-VAR) models in economics and finance has increased significantly in recent years, see e.g. Wilms and Croux (2016), Nicholson et al. (2017), Basu and Michailidis (2015), Nicholson et al. (2020), Hecq et al. (2023) and Camehl (2023).

3.4 Two ways to estimate the LASSO-VAR model

Compared to linear regression (cross-sectional regression), analysis of high-dimension VAR models requires critical consideration from a specific point. Since the response variable is multivariate in the VAR systems, choosing the loss function in the estimation step (sum of squared residuals or negative log-likelihood) is challenging (Basu and Michailidis, 2015). In a multivariate VAR setting, the applied loss function plays a critical role as it influences the LASSO-VAR model parameter estimation and accuracy, even if the multivariate time series have Gaussian errors[18]. This phenomenon becomes more serious when the multivariate error process has highly correlated components (Basu and Michailidis, 2015). The reason is that the negative log-likelihood function considers the inverse covariance matrix of the error term, but the OLS (based on the sum of squared residuals) estimation is not. This distinction will generally lead to different estimations of high-dimensional VAR models unless the unknown covariance matrix equals of the identity matrix[19] (Davis et al., 2016; Basu and Michailidis, 2015). To compare the different estimation methods' strengths and weaknesses, in the two sections, I summarize both the sum of squared residuals and the penalized maximum likelihood-based LASSO estimator.

3.4.1 Row-wise LASSO-VAR

In estimating the LASSO-VAR method, a standard, and most common approach is to apply the uni-variate OLS regression estimation separately on each response variable (Kock and Callot, 2015)). This essentially involves restructuring the VAR model into a linear regression model, where the current values of the time series are regarded as the response variable, and lagged values are considered as the explanatory variables. After that, researchers apply penalized estimation equation-by-equation separately, often with different tuning parameters for different equations (Lee and Liu, 2012; Basu and Michailidis, 2015; Kock and Callot, 2015; Davis et al., 2016; Deshpande et al., 2019). Hereafter I refer this estimation as a row-wise LASSO (RW LASSO), and the adaptive version of this estimation as an adaptive row-wise Lasso (ARW LASSO)[20].

Although simple and popular, this estimation strategy ignores the joint information among the response variables. It does not consider the dependence between the time series (the contemporaneous correlation), which may lead to poor predictive performance under certain circumstances. A research of Lee and Liu (2012) points out that the correlation between error terms of the VAR model has a significant impact on the estimated parameters in a penalized regression. Disregarding the serial correlation in the regularization step can be dangerous because the theoretical risk bounds of the estimation depend on

---

[18]which is a common assumption in econometrics

[19]or a scalar multiple of identity matrix

[20]Eq. (16) is related to row-wise LASSO, and Eq. (21) to the adaptive row-wise LASSO estimation

the degree of contemporaneous cross-correlation between the time series (Song and Bickel, 2011; Basu and Michailidis, 2015).

Interestingly, the question of optimally choosing the loss function has little attention yet in the econometrics literature on LASSO-VAR models. Most of the econometrics-related research papers (Kock and Callot, 2015; Medeiros and Mendes, 2016b; Hecq et al., 2023; Adamek et al., 2023) employed the sum of squared residuals as the loss function and did not explore the option of choosing the negative log-likelihood estimation instead of that. Basu and Michailidis (2015) shows that the equation-by-equation estimation imposes strict assumptions on the dependence structure between the cross-sectional units. While the above-mentioned econometrics-related research studies yield highly interpretable and valuable models, the lack of this research is that they do not explicitly model the correlation structure of the residuals, essentially assuming that the covariance matrix is diagonal. Of course, if its assumption is true, the row-wise LASSO-VAR coincides with the penalized maximum likelihood-based LASSO-VAR estimation (Davis et al., 2016; Basu and Michailidis, 2015).

Several machine learning-related papers have considered this LASSO-VAR estimation question in the last ten years (Song and Bickel, 2011; Basu and Michailidis, 2015; Davis et al., 2016). Song and Bickel (2011) were the first to give a theoretical explanation of the consequences of applying LASSO estimation directly (equation-by-equation estimation) to the time series without taking into account the covariance structure (contemporaneous correlation) between the time series. A few years later, Davis et al. (2016) provides numerical evidence based on an extensive simulation that the VAR forecasting performance can be improved by using a log-likelihood-based loss function. This estimation method incorporates information on the data-generating process correlation matrix into the parameter estimation.

Besides the LASSO-VAR estimation problem, in the last two decades, significant work has been dedicated in the machine learning literature to the penalized estimation of a positive-definite sparse precision matrix. Yuan and Lin (2007) assumed an element-wise sparse precision matrix (inverse covariance matrix) and proposed a LASSO estimator to minimize the negative log-likelihood loss function. Besides that, Bickel and Levina (2008) introduced a threshold operator for the inverse covariance matrix, which constrains the parameters toward zero. Rothman et al. (2010) showed that this estimator satisfies the oracle property.

3.4.2 PML based LASSO-VAR

In order to ease of notation, consider the same VAR model appearing in Eq. (1). I rewrite the VAR model - J dimensional vector of time series with T length and P lag – in the following matrix[21] form:

$$Y = XB + \Sigma \tag{18}$$

where Y is a $K \times J$ matrix, with $K = T - p$, $y = [y_p, ..., y_T]$. Let $X = [X_1, X_2, X_3..., X_p]$ be a lagged time series the matrix of with size $K \times Jp$. Besides that let B be the $Jp \times J$ matrix $B = [B_1, B_2, B_3..., B_p]$ containing the autoregressive coefficients β_i for $i = 1, ..., p$ and Σ is the $K \times J$ innovation matrix (Rothman et al., 2010; Barbaglia et al., 2020).

As I discussed in the earlier section, one possible way to estimate B is to build J single equations separately and solve all of them with the least square loss function. However, this method ignores all the information in the Σ matrix. This separate modeling approach

---

[21]it's a common representation of the VAR model in a machine learning literature

can work well only if Σ is diagonal because it does not use the joint information among the response variables.

Let's denote the inverse covariance matrix: $\Sigma^{-1} = \Omega$. If Σ is known, the negative log-likelihood function for (B,Ω) is the following:

$$g(\hat{B}, \hat{\Omega}) = \frac{1}{2K}trace\left[(Y - XB)\Omega(Y - XB)'\right] - \frac{1}{2}\log|\Omega| \qquad (19)$$

The maximum likelihood function involves the inverse covariance matrix (Ω); however, the estimation of B is equal to separate ordinary least squares estimation for each of the response variables. In this case, the maximum likelihood estimation does not take advantage of the known Ω. However, when we add L1 or L2 penalties on the log-likelihood function, the penalized maximum likelihood estimator can bring a significant advantage in the estimation process (Rothman et al., 2010; Lee and Liu, 2012; Ma and Michailidis, 2016). In addition, in case $J \times p \geq T$, the maximum likelihood estimator does not exist. The penalized maximum likelihood estimator in contrast, is still computable.

While variable selection and covariance estimation have long, rich histories in the machine learning and econometrics regularization literature, a joint solution has only recently gotten more attention. To the best of my knowledge, Rothman et al. (2010) were the first who suggested this solution in the context of cross-sectional regressions. They considered simultaneous sparse estimation of the regression parameters and the inverse covariance matrix of errors, and introduced a new estimation method solving a penalized maximum likelihood problem as an L1-regularized log-likelihood objective function, accounting for correlated errors in the estimation process. Rothman et al. (2010) coined their new approach to multivariate regression with covariance estimation (MRCE) During the estimation process, they simultaneously optimized an objective function with L1 penalties on both the coefficient matrix B and the inverse covariance matrix Ω. This method differs on one key point from the earlier penalized covariance estimation approaches like Yuan and Lin (2007) and Bickel and Levina (2008). Here, the covariance matrix is estimated with penalization to improve the estimation of the coefficients and the prediction rather than as a stand-alone parameter.

Following Rothman et al. (2010), several authors have proposed similar methods for solving the joint problem of coefficient (variable selection) and covariance matrix estimation. Lee and Liu (2012) proposed a weighted version of Rothman et al. (2010)'s method. They considered nearly the same objective function but with adaptive L1 penalties. Besides that, they also provided theoretical results for the application of the method in a low-dimensional setting. Ma and Michailidis (2016) introduced a convergent estimation procedure and established the theoretical properties of this algorithm. They pointed out this joint optimization problem's weakness: it can be numerically unstable in high dimensions. However, all of these authors report relatively good performance in estimating both the response variables and the inverse covariance matrix (Lee and Liu, 2012; Ma and Michailidis, 2016; Deshpande et al., 2019).

In the last decade, two econometrics-related studies applied the MRCE method in a times series context. First, Wilms and Croux (2016) introduced a penalized maximum likelihood approach for estimating the cointegrating vectors and a rank in a vector error correction model (VECM). The estimation assumed sparse coefficients, and they analyzed the model's forecasting performance on cointegrated times series. They showed in an extensive simulation that the penalized maximum likelihood-based estimation significantly

outperforms the benchmark Johansen method (Johansen, 1988) significantly when the space cointegrating vectors are sparse. A few years later, Barbaglia et al. (2020) proposed a PML method to estimate a t-LASSO-VAR. They were the first who assumed the fat-tailed distribution of the VAR model errors. Besides that, after the parameter estimation – autoregressive matrix and the inverse covariance matrix – they calculated the DY spillover network too. They showed that the t-LASSO-VAR attains better forecast accuracy than standard estimators (row-wise LASSO) if the error terms follow a fat-tailed distribution.

Both Wilms and Croux (2016) and Barbaglia et al. (2020) estimate the VARs under the assumption that the autoregressive matrix and the inverse covariance matrix is band. It is important to note that this assumption is different from that used in other econometrics-related papers such as Kock and Callot (2015) and Medeiros and Mendes (2016b), where sparsity assumptions are formulated only for the autoregressive matrices. It is closer to the machine learning papers like Basu and Michailidis (2015) and Davis et al. (2016). In the econometrics literature, as far as I know, only Hecq et al. (2023) has the same assumptions for the data-generating process.

Following Rothman et al. (2010), Barbaglia et al. (2020) was the first to adapt the penalized maximum likelihood estimation (based on the MRCE concept) in a high-dimension VAR context. In this case, the penalized maximum likelihood estimator jointly estimates the coefficient (B) and inverse covariance matrix (Ω) by solving the following problem:

$$
(\hat{B}, \hat{\Omega}) = \underset{B,\Omega}{\operatorname{argmin}} \left[\frac{1}{2J} trace \left[((Y - XB)\Omega(Y - XB))' \right] - \frac{1}{2} \log |\Omega| \right.
$$
$$
\left. + \lambda \sum_{s=1}^{p} \sum_{i,k=1}^{J} |\beta_{s,ik}| + \gamma \sum_{i=k}^{J} |\omega_{ik}| \right] \tag{20}
$$

where $\lambda > 0$ and $\gamma > 0$ are the hyperparameters and ω_{ij} is the i-th row j-th column element of the inverse covariance matrix Ω. This proposed method simultaneously estimates the coefficients and the elements of the inverse covariance matrix of the VAR(p) model.

This L1-penalty on B ensures that the estimation of the coefficient matrix is feasible even if the number of estimated parameters exceeds the length of the time series. Besides that, it estimated some elements of $\hat{B}$ exactly zero. Larger λ indicates sparser $\hat{B}$: more zero estimated elements in the coefficient matrix. Barbaglia et al. (2020) select the L1 penalty on the off-diagonal entries of the inverse error covariance Ω for two different reasons. First, this penalty ensures that the estimation of the inverse covariance matrix is feasible even if the number of parameters exceeds the length of the time series. Secondly, similar to the coefficient matrix, it sets some elements of $\hat{\Omega}$ exactly to zero. Again, a larger regularization parameter indicates a sparser covariance matrix. If J is large, estimating a non-sparse inverse covariance matrix Ω means that the method has $O((J)^2)$ additional parameters in compared with row-wise LASSO method (Rothman et al., 2010). Sparse inverse covariance matrix estimation has significantly lower variability; for this reason, L1-penalty is preferred against L2 (ridge) penalty (Rothman et al., 2010; Barbaglia et al., 2020). Hereafter, I refer to this estimator as a penalized maximum likelihood LASSO (PML LASSO).

3.5 Regularized DY framework

The sparsity assumptions are essential in the Diebold-Yilmaz network estimation, too. The variance decomposition matrix that drives DY connectedness measures and spillover indices is a nonlinear transformation of the VAR coefficients and the inverse covariance matrix. Mostly, the underlying systems are based on high-dimension time series, and we need to assume the coefficients and the inverse covariance matrix are sparse to estimate the true network efficiently. Using models for the VAR estimation with the best forecast accuracy is essential to achieve a good network and spillover index approximation (Demirer et al., 2018).

To address the issue of high dimensionality financial network estimation, Demirer et al. (2018) were the first to introduce the regularized VAR(p) estimation techniques – adaptive elastic net (Zou and Hastie, 2005) – into the Diebold-Yilmaz literature when they estimated the network across 96 global international banks. This seminal paper does not formulate sparsity assumptions on the spillover matrix Θ; instead, they assume sparsity on the approximating VAR model coefficient matrix B. Their approach made an equation-by-equation sparse penalized estimation of the VAR model coefficient matrix B.

Based on Demirer et al. (2018) seminal paper, in recent years, different regularized methods for estimating high-dimension VAR have been extensively used in financial system modeling, especially in the DY framework empirical literature. While Yang et al. (2021), Just and Echaust (2022), and Tiwari et al. (2022) used equation-by-equation LASSO, Yi et al. (2018), Wen and Wang (2020), and Cheng et al. (2022) applied group LASSO, Greenwood-Nimmo and Tarassow (2022) and Greenwood-Nimmo et al. (2017) estimated the VAR with adaptive LASSO. Besides that Greenwood-Nimmo et al. (2019), Bostanci and Yilmaz (2020) and Fang and Shao (2022) applied elastic net and Grant and Yung (2021) and Jena et al. (2022) used adaptive elastic net for the VAR(p) model estimation.

4 Chapter 4. Research questions and thesis contribution

In this chapter, I highlight the limitations of the Diebold-Yilmaz framework and the room for improvement. I summarize the shortcomings of the empirical network literature and present the contribution of my thesis in detail.

4.1 Limitations of the DY framework

In recent years, the DY framework has already been extended by several methodological innovations (frequency decomposition by Baruník and Křehlík (2018) regime-switching by Lee and Chang (2013); BenSaïda et al. (2018)). However, there are still some relevant limitations of the framework.

DY framework is based on a VAR estimation, resulting in a closed network. Omitted relevant factors from the financial markets can lead to misleading results. Some researchers (Kock and Callot, 2015) tried to handle this shortcoming of the original approach with a FAVAR model estimation ((Claeys and Vašíček, 2014; Barigozzi et al., 2021, 2023), but most articles do not address this issue.

The simplicity of VAR supporting only liner relations between the time series is a major shortcoming in the original DY approach. While researchers, assuming nonlinear linkages between the financial institutions, can apply the time-varying parameter VAR (TVP-VAR) version of the framework (Antonakakis et al., 2020), the method has its own problems, especially with cointegrated time-series. However, it's questionable that this method captures all the nonlinear relations between the time series.

Most empirical applications of the DY framework focusing on relatively small models estimating simple VAR models have used traditional estimators such as ordinary least squares (OLS) (Demirer et al. (2018)). Regularization methods must be used in larger systems or when estimation is based on a small time series sample because traditional estimation methods are challenging to perform well in a high dimensional time series network context. However, applying these machine-learning methods raises certain questions. Which method is the most effective in the DY framework? How should we optimize the regularization parameters of the methods? In the case of which dimension is it worth using? Is it worth regularizing the covariance matrix of the error term?

Some articles have already examined the changes in DY networks in the empirical literature. However, no formal statistical test has been spread in the empirical literature to measure the structural changes in the network. Only visualization and some simple measures were primarily used in the DY literature. Nevertheless, formal statistical tests must be included to create a more complex analysis of the financial network and systemic risk.

4.2 Shortcomings of the empirical network literature

Almost all empirical studies analyze the realized volatility network of the financial system. Although some works have examined skewness (Greenwood-Nimmo et al., 2017) and liquidity (Smimou and Khallouli, 2016; Grillini et al., 2022) spillovers, their behavior in the financial networks is still questionable. Besides that, only a few studies (Wang et al., 2021, 2023) focus on comparing different networks.

Despite the many recent connectedness-related articles and the widespread methods, the deeper structure of the networks (analyzing on both macro and micro levels) has been investigated by far fewer. In addition to that, only a few studies (Hautsch et al., 2015; Nucera et al., 2016; Hué et al., 2019) tried to identify the key participants of the financial networks (and analyze the dynamics of their linkages), and less one analyze these players' role in crisis periods (Salim et al., 2023).

As far as I know, only a few studies analyzed the effect of structural changes and economic events on the networks (Claeys and Vašíček, 2014; Buse et al., 2022; Greenwood-Nimmo and Tarassow, 2022). Both the econometric and economic background are incomplete in this crucial question. Regulators need to monitor the rapid changes in the network, understand the system dynamics on different levels, and identify the key participants and their role in the financial networks.

Most of the empirical studies that apply the Diebold-Yilmaz framework analyze only the point estimates of the SUM, FROM, TO, NET, and pairwise NET spillover indices (Diebold and Yilmaz, 2012; Diebold and Yılmaz, 2014; Greenwood-Nimmo and Tarassow, 2022). The large drawback of this limitation is that the structural changes in the network are quite complex, and it is often not explored with a deep analysis. The analyses that measure the significant structural changes in the macro (total spillover index) and micro level (net pairwise connections) of the system allow us to extract more complex information from the analyzed network. Besides that, if we narrow down our analysis to the point estimation of the spillover indices, this prevents us from conveying the information surrounding the uncertainty of our estimation.

4.3 Thesis contribution

My thesis contributes to the existing literature in three points. Firstly, I contribute to high-dimension time series modeling and network analysis with a methodological innovation. I propose an extension of the adaptive row-wise LASSO method, which accounts for possible sparsity both in the coefficient (B) and the inverse covariance matrix (Ω) of the estimated (VAR) model.

Secondly, I contribute to that part of network analysis literature, which analyses the structural changes in the systems with a methodological innovation. I extend the Diebold-Yilmaz framework with an event study tool to estimate network structural changes and provide more in-depth insights about contagion channels in the systems during structural changes. Integrating a moving-block bootstrap method (MBB) into the framework, I can investigate how the observed shocks transform the networks.

Thirdly, I contribute to the empirical network-based systemic risk modeling literature with a conceptual innovation. I characterize and compare FIs' illiquidity and volatility networks. I show that the DY framework extended with the MBB method is a powerful tool to identify the key participants (troubled financial institutions) and contagion channels during crisis periods.

My contributions have not yet been published, but I would like to highlight my related article (Badics et al. (2023)), which focuses on modeling high-dimension time series and network analysis, but is not included in my thesis.

4.3.1 Research method innovations

Regularisation in the estimation step of the VAR model may be useful for high-dimension time series modeling. In Chapter 5, I extend the commonly used row-wise

LASSO method. I propose a new regularization method that is useful in the DY framework, as it is able to shrink and select the essential coefficients of the VAR and simultaneously account for possible sparsity in the distribution of the VAR model errors. In Chapter 5.2, I perform an extensive Monte Carlo simulation study and show that my new method estimates the elements of both the VAR(p) model's coefficient (B) and inverse covariance matrix (Ω) and the DY spillover table (Θ) more effectively than the commonly used methods. Besides that, I investigate the robustness of my method of the network parameters (dimension of the system, time series lengths). In general, the advantage in the estimation accuracy becomes much greater in shorter time series. Relative estimation accuracy is sensitive to the dimension of the system, the highest in high-dimension settings.

Combining the DY framework for network modeling with event study methodology may be useful at turbulent times. In Chapter 6.1.3, I extend the original Diebold-Yilmaz framework with an event study tool to provide more insights into the contagion channels appearing in the networks during structural changes.

I introduce a formal test to compare the distribution of the SUM, FROM, TO, NET, and NETP (Eqs. (6-10)) spillover indices over time. I integrate the bootstrap algorithm into the DY framework's estimation step to determine whether the DY spillover indices' empirical distribution on different periods is equal. The gain of this method is that we can use it not just for analyzing the significant structural changes in the network on a macro level (total spillover index) but on a micro level (pairwise spillover indices) too. Combining the residual-based MBB Brüggemann et al. (2016) method with the DY framework (MBB-based DY framework), I can investigate how the observed shocks transform the financial networks. By interpreting the system as a network, I can monitor the network on a daily basis and analyze structural changes with the combined method.

In the earlier empirical literature, only two papers integrate these methods (the DY framework and MBB algorithm). Buse et al. (2022) applied only for the total spillover index, and Greenwood-Nimmo and Tarassow (2022) used to create probabilistic analysis of spillover scenarios. To the best of my knowledge, I am the first to apply the combined method to analyze the significant changes in both macro (total spillover index) and micro (net pairwise spillover indices) levels of the network.

4.3.2 Conceptual innovations

To address the gap in the empirical literature, in Chapter 6.2, I present a characterization of the illiquidity connectedness of a US financial network with my proposed AJ LASSO method in the Diebold-Yilmaz framework. I investigate the significant similarities and differences between the volatility and illiquidity network on the macro and micro levels with the combined MBB-based DY framework.

In Chapter 6.3, I consider the critical financial events of the GFC, as Diebold and Yılmaz (2014), with the analytical purpose of testing the performance of my new event study framework. Applying the MBB-based DY framework, I analyze the volatility and liquidity shocks, focusing on four key events of the GFC. Besides that, I provide daily snapshots of the financial network to illustrate that illiquidity spillovers act as contagion channels during turmoil time.

The analysis in Chapter 6 contributes to the empirical network analysis literature at three related points. First, as far as I know, I am the first to compare the volatility and illiquidity network with the MBB-based DY framework. My empirical results suggest that the volatility and illiquidity network differs not just on the macro level (total spillover

index) but on the micro (pairwise net spillover indices) level too. I conclude that micro and macro-level illiquidity spillover indices display important information for the financial network.

Second, I apply the MBB-based DY framework to analyze the effect of key events in the network. An important limitation of the original DY framework is the absence of a formal statistic to test whether the changes in the spillover matrix are significant. Applying the MBB-based DY framework, I prove that the observed shocks temporarily transform the financial network and significantly alter the spillover matrix.

Lastly, I show that DY illiquidity spillover better tracks the dominant shock transmitters in the system and signals during financial turmoil than volatility spillover both in the macro and micro level of the network. My empirical results strengthen the findings of Gai and Kapadia (2010) and Acemoglu et al. (2015). First, I found that the impact of a shock depends on which node of the network it hits. Second, the network transformation's persistence depends on the shock size.

To the best of my knowledge, I am the first to apply the MBB-based DY framework for a network event study and to examine the key participants of the US FIs network during the GFC.

5   Chapter 5. Adaptive joint LASSO based network estimation

In this chapter, I propose a new regularization method that is useful in the DY framework, as it is able to shrink and select the essential coefficients of the VAR and simultaneously account for possible sparsity in the distribution of the VAR model errors. This method is an extended penalized maximum likelihood estimator for the VAR coefficient and inverse covariance matrix, and I refer to this method as the adaptive joint LASSO algorithm (AJ LASSO). In the second part of the chapter, I perform an extensive simulation study to present my proposed estimator's statistical properties. I compare my new approach with the most widely used VAR model estimation methods (OLS, RW LASSO, ARW LASSO)[22].

5.1   Adaptive Joint LASSO

The main contribution of this chapter is to propose a new methodology to estimate large, sparse VARs. I introduce a novel adaptive joint LASSO-based algorithm. The highlight of the procedure is that it simultaneously estimates the autoregressive matrices and the entries of the residual inverse covariance matrix. To understand the strength of my new method I formally introduce my estimator, in the following subsection.

In order to achieve more efficient estimation, I propose an extended penalized maximum likelihood method for VAR estimation. To develop a better estimator for B and Ω, I consider incorporating the inverse covariance matrix of the error term (Ω) in the estimation process, and regularize both the coefficient and the inverse covariance matrix, following Rothman et al. (2010), Lee and Liu (2012) and Barbaglia et al. (2020). Unlike previous studies, I estimate the model with adaptive-LASSO, minimizing the following weighted negative log-likelihood function:

$$(\hat{B}, \hat{\Omega}) = \operatorname*{argmin}_{B,\Omega} \left[\frac{1}{2J} trace\left[((Y - XB)\Omega(Y - XB))' \right] - \frac{1}{2}\log|\Omega| \right.$$
$$\left. + \lambda \sum_{s=1}^{p} \sum_{i,k=1}^{J} w_{s,ik}|\beta_{s,ik}| + \gamma \sum_{i=k}^{J} w_{ik}|\omega_{ik}| \right] \tag{21}$$

where λ and γ are the two hyperparameters of the method.

I apply the adaptive version of LASSO for two reasons. Firstly, as Kock and Callot (2015) stands, only the row-wise adaptive LASSO has an oracle property for the high-dimension VAR models. Second, as Lee and Liu (2012) pointed out, the adaptive version of the MRCE approach has better parameter estimation accuracy for cross-sectional data.

I take the weights of the adaptive LASSO as a reciprocate of the L2 estimator, where the ridge estimator is based on the following minimization:

---

[22]these regularization approaches are the most commonly applied methods in the DY literature

$$(\hat{B}, \hat{\Omega}) = \underset{B,\Omega}{\mathrm{argmin}} \left[\frac{1}{2J} trace \left[(Y - XB)\Omega(Y - XB)' \right] - \frac{1}{2} \log |\Omega| \right.$$
$$\left. + \lambda_2 \sum_{s=1}^{p} \sum_{i,k=1}^{J} \beta_{s,ik}^2 + \gamma_2 \sum_{i=k}^{J} \omega_{ik}^2 \right] \tag{22}$$

I refer to my new algorithm as the adaptive joint LASSO (AJ LASSO) method. The essence of the model is the maximum likelihood estimation with joint penalization on both B and Ω matrices. The adaptive joint LASSO extends the MRCE method (Rothman et al., 2010) and PML LASSO method (Barbaglia et al., 2020) with the adaptive estimation step. It has two substantial advantages compared to the commonly used adaptive row-wise LASSO method. Firstly, it incorporates the inverse covariance matrix (Ω) into the coefficient matrix (B) estimation. Based on the earlier results of the MRCE literature (Rothman et al., 2010; Lee and Liu, 2012) this feature of the estimation can be important if the inverse covariance matrix has a dense or band structure ($\Omega \neq I_J$). Second, besides the coefficient matrix (B) regularization, this approach penalizes the elements of the inverse covariance matrix (Ω) too[23]. It can be effective if the inverse covariance matrix is sparse ($\Omega = I_J$). Besides these advantages, I highlight that the Diebold-Yilmaz spillover table depends on both B and Ω, and the strength of my method is that it jointly estimates these matrices. Mention the disadvantages of my model as well; it uses only one hyperparameter for the coefficient and inverse covariance matrix estimation, while the row-wise LASSO optimizes the selection equation-by-equation. If some time series dominate the system, the row-wise LASSO can be a better choice for estimating B.

I don't proof the oracle property of my new method for VAR(p) model coefficient (B) and inverse covariance matrix (Ω) estimation, but in the next section, I conduct an extensive set of Monte Carlo simulation and represent that my method achieves better estimation and forecast accuracy than the row-wise, and adaptive row-wise LASSO estimators [24].

5.1.1 Details of the algorithm

The optimization problem in Eq. (21) is not convex concerning the coefficient and covariance matrices (B, Ω), however, solving for either B or Ω with keeping the other fix is the objective function[25] become convex (Rothman et al., 2010; Ma and Michailidis, 2016).

To solve this, I present an algorithm in line with the estimation of Rothman et al. (2010)'s MRCE method. Solving Eq. (21) for Ω with fixed B_0 yields the following estimation problem:

$$\hat{\Omega}(B_0) = \underset{\Omega}{\mathrm{argmin}} \left[\frac{1}{2K} tr \left[(Y - XB_0)\Omega(Y - XB_0)' \right] - \frac{1}{2} \log |\Omega| \right.$$
$$\left. + \gamma \sum_{i=k}^{J} w_{ik} |\omega_{ik}| \right] \tag{23}$$

---

[23]row-wise LASSO and adaptive row-wise LASSO only penalize B

[24]we can call this method as a heuristic approach

[25]the objective function is bi-convex

This is exactly a weighted (adaptive) L1-regularized covariance estimation problem considered by several machine learning studies like Yuan and Lin (2007); Bickel and Levina (2008) and Friedman et al. (2010). Similar to Rothman et al. (2010) and Barbaglia et al. (2020), I use Friedman et al. (2010)'s graphical LASSO (gLASSO) algorithm for optimization. The advantage of the gLASSO is well known: it is fast and the most commonly used algorithm for solving inverse covariance matrix penalization. It needs to be mentioned that following Friedman et al. (2010), I left out the diagonal elements from the objective function; therefore, only the off-diagonal elements are penalized.

Solving Eq. (21) for B with fixed Ω_0 yields the following estimation problem:

$$\hat{B}(\Omega_0) = \underset{B}{\operatorname{argmin}} \left[\frac{1}{2J} tr \left[(Y - XB)\Omega_0(Y - XB)' \right] + \lambda \sum_{s=1}^{p} \sum_{i,k=1}^{J} w_{s,ik}|\beta_{s,ik}| \right] \qquad (24)$$

A solution of Eq. (24) can be efficiently computed using cyclical-coordinate descent algorithm (Lee and Liu, 2012; Ma and Michailidis, 2016). In line with Rothman et al. (2010), Lee and Liu (2012) and Barbaglia et al. (2020), I iterate the optimization of the two objective functions (Eqs. (23-24)) until convergence.

As Lee and Liu (2012) pointed out, the weakness of this iterative estimation is that it may fail to converge to the global optimum. It can be unstable, especially in high-dimensional settings where $K \times J \geq T$. If some diagonal element of the first term in Eqs. (23-24) are zeros, then it dominates the others during the estimation process. Similar to most non-convex problems, for fast convergence, good initial parameters are beneficial (Ma and Michailidis, 2016).

To create a stopping rule for the estimation, I iterate the algorithm minimizing the objective function Eq. (21) until the relative change in it is less than tolerance level ϵ. This parameter of the algorithm needs to be specified following Wilms and Croux (2016) and Barbaglia et al. (2020). I choose to be $\epsilon = 10^{-2}$. While there is no formal proof of the convergence of this algorithm, based on my knowledge, I have consistently observed convergence in all of my simulation scenarios, typically occurring within an average of two iterations. To summarize the convergence of the iterative method, based on my experiment, the coordinate-descent algorithm works very efficiently. Since the adaptive joint LASSO algorithm method involves estimation of both B and Ω, with an additional ridge estimation Eq. (22), unfortunately, the computation can be intensive when the dimension of the VAR is high.

To start the iterative estimation process, we need starting values for both B and Ω. I take the identity matrix of Ω for starting values. It's a common choice, and I follow Rothman et al. (2010) and Barbaglia et al. (2020). For the starting value of B, I use the estimated coefficients of the row-wise LASSO method in line with Lee and Liu (2012). In my simulations, I performed numerous numerical experiments to assess how these choices influence the algorithm's accuracy. For both matrices, the results are robust in low and high-dimensional VAR settings, too.

The regularization parameters λ and γ in Eqs. (21-24) are not known, we need to optimize these hyper-parameters. Besides that, due to the adaptive estimation, we have to fine-tune two other parameters, the penalties of the ridge objective function Eq. (22). That means estimating adaptive joint LASSO algorithm-based VAR models involves choosing the optimal values of four tuning parameters. Achieving good performance in terms of coefficient and covariance matrix estimation requires the appropriate selection

of the penalty parameters (Kock and Callot, 2015; Hecq et al., 2023).

While some theoretical results on the choice of these hyperparameters are available in the machine learning literature (Basu and Michailidis, 2015), in the empirical literature, many different data-driven methods exist to fine-tune these parameters, giving varying results. The choice of λ and γ are usually based on certain information criteria (AIC, BIC) or cross-validation, which is challenging in time series context (Hecq et al., 2023).

I employ a computationally efficient approach[26], and in line with Rothman et al. (2010), Wilms and Croux (2016) and Barbaglia et al. (2020) I use the Bayesian information criterion introduced by Schwarz (1978) for both λ, γ and the ridge penalty optimization. When the iterative method solves the Eq. (24), I use a specified grid of the λ hyperparameter and search for the optimal value minimizing the BIC based on the following equation:

$$BIC_\lambda = -2logL_\lambda + df_\lambda log(N) \tag{25}$$

where $logL_\lambda$ is the log-likelihood of the estimated VAR model with a λ penalty and df_λ is the degree of freedom of the model which can be well approximated by the number of nonzero coefficients of the estimated coefficient matrix B (Zou et al., 2007; Rothman et al., 2010).

Similarly, when the iterative method solves the Eq. (23), I use a specified grid of the γ hyperparameter and search the optimal value minimizing the BIC based on the following equation:

$$BIC_\gamma = -2logL_\gamma + df_\gamma log(N) \tag{26}$$

In this case, df_λ is the nonzero lower-diagonal elements of the estimated inverse covariance matrix Ω (Barbaglia et al., 2020).

An alternative way to optimize the hyperparameters with information criteria is the AICc, a corrected version of AIC. In a VAR estimation, Hecq et al. (2023) showed it has a better finite sample performance for small sample sizes. My simulations indicate no significant difference in terms of estimation accuracy of the adaptive joint LASSO-based VAR estimation when either using BIC or the AICc to select the regularization parameters λ and γ.

Perhaps the most popular way in cross-section regression is to choose the tuning parameter, cross-validation (CV). However, as Bergmeir et al. (2018) showed, the CV-based LASSO is not always appropriate in high-dimension time series settings without serious modifications. I conducted simulations involving cross-validation, but this did not substantially improve results despite being notably slower.

---

[26]my approach has computation intensive two-step iteration process

5.2 Monte Carlo simulation

In this section, I undertake a simulation study to assess the performance and illustrate the beneficial qualities of the previously proposed adaptive joint LASSO algorithm. I simulate different large sparse VAR systems and then use my proposed algorithm to estimate the parameters of the VAR model and the DY network. The aim of this simulation is to analyze the accuracy of my new method in the coefficient (B), covariance matrix (Ω), and the Diebold-Yilmaz spillover table (Θ) estimation.

To explore the relative performance of the new method, I compare it with the most commonly applied regularization methods in the empirical Diebold-Yilmaz literature (Diebold and Yilmaz, 2023). The first method is the OLS, where the VAR parameters are estimated equation-by-equation without penalty. The other two methods are the row-wise LASSO (RW LASSO, Eq. (16)) and adaptive row-wise LASSO (ARW LASSO, Eq. (17)). These are similar to the OLS; both are estimated equation-by-equation, but the objective function has a penalty term. For the first model (RW LASSO), the estimation is a one-step method with L1-penalty. The parameters are estimated in a two-step for the second model (ARW LASSO). Specifically, in the first stage, I use L2-penalty (ridge penalty) for the parameter estimation, and in the second stage, I use the weighted variables to estimate the objective function with L1-penalty. In all regularization methods, I use BIC and the same grid to estimate the hyperparameters. Without this, the results are less comparable. These methods are the most common ways of obtaining VAR estimates and obtaining the DY spillover matrix.

My Monte Carlo simulation in the next section relates to three literature stands. First, my work relates to the literature on the sparse estimation of high-dimension VAR models, such as Song and Bickel (2011), Basu and Michailidis (2015), Kock and Callot (2015) and Davis et al. (2016). It is also connected to the econometric literature on networks for time series models, which includes research by Billio et al. (2012), Diebold and Yılmaz (2014), Demirer et al. (2018), Barigozzi and Brownlees (2019) and Barbaglia et al. (2020). Lastly, this research also linked to the literature on regularized covariance estimation and sparse precision matrices, such as Meinshausen and Bühlmann (2006), Yuan and Lin (2007), Bickel and Levina (2008), Friedman et al. (2010) and Peng et al. (2009).

5.2.1 Simulation design

In addition to comparing the different methods, I also investigate the sensitivity of the simulation parameters on the performance of the four estimators. I test the performance of the 4 methods with different sample sizes and time series numbers, using 1000 simulation runs, and also test the sensitivity of the covariance and auto-correlation structure. I consider both sparse and dense coefficient and covariance matrices to investigate the effect of the autoregressive and the covariance structure on the methods. The notations are similar to the earlier B represents the coefficient matrix, Σ and Ω marks covariance and the inverse covariance matrix and I note the DY spillover matrix with Θ. J represents the number and T is the length of the time series.

Assuming various data-generating processes, represented by various VAR coefficient structures can have different theoretical and empirical motivations. In order to assess the robustness I investigate the effect of both sparse and dense B.

I consider two alternative coefficient matrices for the VAR model adapted from Kock and Callot (2015) and Hecq et al. (2023). In the first case (Eq. (27)), the coefficient matrix is as sparse as possible:

$$B = \begin{bmatrix} 0.4 & 0 & \dots & 0 \\ 0 & 0.4 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 0.4 \end{bmatrix} \quad (27)$$

This structure was investigated in Kock and Callot (2015) (Experiment A in the simulation design) and Hecq et al. (2023) (DGP 1 in the simulation design)[27]. In this truly sparse model, each variable depends only on its own past, but if the covariance structure is not an identity matrix, the DY spillover table can be dense (due to the nonlinear relationship between B, Ω and Θ). In each equation, we have J-1 redundant variables.

In the second case, the coefficient matrix (Eq. (28)) is given by the following way[28]:

$$B = \begin{bmatrix} (-1)^{|i-j|}0.4^{|i-j|+1} & \dots & (-1)^{|i-j|}0.4^{|i-j|+1} \\ \vdots & \ddots & \vdots \\ (-1)^{|i-j|}0.4^{|i-j|+1} & \dots & (-1)^{|i-j|}0.4^{|i-j|+1} \end{bmatrix} \quad (28)$$

where i and j represent the rows and columns of the matrix. This matrix is equal with Kock and Callot (2015) (Experiment D in the simulation design) and Hecq et al. (2023) (DGP 2 in the simulation design). To understand the concept behind this choice, I illustrate this dense matrix with an example. For $J = 5$ time series the coefficient matrix is the following:

$$B = \begin{bmatrix} 0.40 & -0.16 & 0.06 & -0.03 & 0.01 \\ -0.16 & 0.40 & -0.16 & 0.06 & -0.03 \\ 0.06 & -0.16 & 0.40 & -0.16 & 0.06 \\ -0.03 & 0.06 & -0.16 & 0.40 & -0.16 \\ 0.01 & -0.03 & 0.06 & -0.16 & 0.40 \end{bmatrix} \quad (29)$$

The choice can be motivated by financial networks, with both negative and positive connections between the assets and FIs. The distress experienced by one FI is transmitted at most to the closer institution (Hecq et al., 2023). I refer to this case as dense B, because it illustrates the violation of the sparsity assumption of the estimation methods. Many of the parameters are small, but none of them is zero. The coefficients decrease exponentially in the distance from the diagonal. The stationarity of the VAR is ensured in both designs.

Similar to the coefficient matrix (B) the results of the different estimation methods are reported for different types of covariance structures (Σ). I consider 3 different structures for Σ and its inverse Ω. In line with the earlier papers (Rothman et al., 2010; Basu and Michailidis, 2015; Barbaglia et al., 2020; Hecq et al., 2023), my first choice for the innovation covariance matrix (Σ) is a Toeplitz-type structure with a ρ parameter (Eq. (30)). Larger values indicate the error process of the systems is more correlated.

$$\Sigma = \begin{bmatrix} \rho^{|i-j|} & \dots & \rho^{|i-j|} \\ \vdots & \ddots & \vdots \\ \rho^{|i-j|} & \dots & \rho^{|i-j|} \end{bmatrix} \quad (30)$$

---

[27]The results are robust for the coefficient value in the diagonal

[28]adapted from Kock and Callot (2015) and Hecq et al. (2023)

That means the inverse error covariance matrix (Ω) is a band matrix[29]. To illustrate this, for $J = 5$ and $\rho = 0.3$ the inverse error covariance matrix is the following:

$$\Omega = \begin{bmatrix} 1.1 & -0.3 & 0.0 & 0.0 & 0.0 \\ -0.3 & 1.2 & -0.3 & 0.0 & 0.0 \\ 0.0 & -0.3 & 1.2 & -0.3 & 0.0 \\ 0.0 & 0.0 & -0.3 & 1.2 & -0.3 \\ 0.0 & 0.0 & 0.0 & -0.3 & 1.1 \end{bmatrix} \tag{31}$$

The second case is related to the first one. In case of $\rho = 0$, $\Sigma = I_J$, and both Σ and its inverse (Ω) are sparse. This choice is in line with Kock and Callot (2015). Lastly, to investigate the violation of the sparsity assumption, I consider a dense covariance matrix too (Eq. (32)). I follow Wilms and Croux (2018) and simulate the error term from the following structure:

$$\Sigma = \begin{bmatrix} 0.5 * ((|i - j| + 1)^{2*0.9} - 2 * |i - j|^{2*0.9} + (|i - j| - 1)^{2*0.9}) & \cdots & \cdots \\ \vdots & \ddots & \vdots \\ \vdots & \cdots & \vdots \end{bmatrix} \tag{32}$$

At this case both Σ and Ω has a dense structure, for $J = 5$ the inverse error covariance matrix is the following:

$$\Omega = \begin{bmatrix} 2.4 & -1.3 & -0.2 & -0.2 & -0.3 \\ -1.3 & 3.1 & -1.3 & -0.1 & -0.2 \\ -0.2 & -1.3 & 3.1 & -1.3 & -0.2 \\ -0.2 & -0.1 & -1.3 & 3.1 & -1.3 \\ -0.3 & -0.2 & -0.2 & -1.3 & 2.4 \end{bmatrix} \tag{33}$$

I simulate all runs from VAR(1) and assume p is known through the estimation process for every method. I perform the experiment for different sample sizes $T = 100, 250, 500$. I also conduct an extensive robustness evaluation of my results to the choice of the dimension of the system, and the contemporaneous correlation between the time series. I vary the number of time series from $J = 10$ to $J = 25$. I considered $J = 10$ low-dimension setting, and $J = 25$ high-dimension setting. The dimensions of the VAR are in line with the empirical study made in the next chapter. To investigate the contemporaneous correlation sensitivity of the results I vary the Toeplitz-type covariance matrix parameter (ρ) from 0.1 to 0.9. I replicate the simulation 1000 times for each parameter set. In the simulation study, I focus on the horizon of ten days ($H = 10$). This value of the DY spillover table parameter is a common choice in an empirical setting.

The parameter choices of the simulation are in line with the earlier papers (Basu and Michailidis, 2015; Kock and Callot, 2015; Hecq et al., 2023), but the investigation is more deep. My Monte Carlo simulation takes into consideration the time series length, the dimension (low and high setting), and the correlation at the same time.

In this simulation, I aim to present the performance gain of my new method in coefficient (B) and inverse covariance matrix (Ω) parameter estimation and investigate the robustness of the results. Besides that, I also investigate the efficiency of the DY spillover

---

[29]A band matrix is characterized by having non-zero elements exclusively along diagonal bands situated either above or below the main diagonal

matrix (Θ) estimation. Lastly, I analyze different settings to explore in which circumstances perform the adaptive joint LASSO best compared to the other methods.

5.2.2 Performance measures

Through the Monte Carlo simulation, I measure the performance of the applied methods from two perspectives. First is the sparsity recognition accuracy for both B and Ω. Secondly, I calculate the estimation accuracy of the parameters for B, Ω, and Θ in an absolute and quadratic way.

Similar to the earlier papers that investigate the effectiveness of regularization methods (Rothman et al., 2010; Kock and Callot, 2015; Wilms and Croux, 2018), I measure the sparsity recognition by the true positive rate (TPR), true negative rate (TNR), and the weighted average of this for both B and Ω. I calculate the true positive and true negative rates in the following way:

$$TPR(B, \hat{B}) = \frac{\#(i,j) : \hat{b}_{ij} \neq 0, b_{ij} \neq 0}{\#(i,j) : b_{ij} \neq 0} \tag{34}$$

$$TPR(\Omega, \hat{\Omega}) = \frac{\#(i,j) : \hat{\omega}_{ij} \neq 0, \omega_{ij} \neq 0}{\#(i,j) : \omega_{ij} \neq 0} \tag{35}$$

$$TNR(B, \hat{B}) = \frac{\#(i,j) : \hat{b}_{ij} = 0, b_{ij} = 0}{\#(i,j) : b_{ij} = 0} \tag{36}$$

$$TNR(\Omega, \hat{\Omega}) = \frac{\#(i,j) : \hat{\omega}_{ij} = 0, \omega_{ij} = 0}{\#(i,j) : \omega_{ij} = 0} \tag{37}$$

The true positive rate gives the estimation accuracy of important variables, whereas the true negative rate gives the accuracy of excluding the unimportant variables. Sparsity recognition weighed these measures with the relative number of non-zero and zero elements in the B and Ω.

To evaluate the parameter estimation accuracy, I use the mean absolute estimation error (MAEE) in line with Barbaglia et al. (2020) and Wilms and Croux (2018) and the mean Frobenius norm (MFBN) followed by Basu and Michailidis (2015).

The first one is given by

$$MAEE(B, \hat{B}) = \frac{1}{J^2} \sum_{i,j=1}^{J} |\beta_{ij} - \hat{\beta}_{ij}| \tag{38}$$

the second one is the following

$$MFBN(B, \hat{B}) = ||\hat{B} - B||_F / ||B||_F \tag{39}$$

The calculation is similar for Ω and Θ. In every different DGP and parameter setting I calculate the mean of these performance measures based on 1000 experiments.

5.3 Simulation results

In the next four subsections, I report the estimation accuracy of the four methods in different settings. In the first subsection, I focus on the base setting where the coefficients matrix (B) is sparse and the inverse covariance matrix (Ω) is band. I analyze the effect of the time series (T) length on the sparsity recognition and the parameter estimation for the methods. In the next subsection, to estimate the robustness of the results, I vary the time series dimension and the contemporaneous correlation between the series. After that, I investigated the effect of the covariance structure. I analyze the robustness of the results if the Ω is sparse and dense. In the last subsection, I examine the case of non-sparse B. I report the results for dense B with two different settings of Ω (band and dense).

5.3.1 Sparse coefficient and band inverse covariance matrix

Simulation results on the sparsity recognition are given in Table 3, which reports the average true positive and negative rates and the sparsity recognition for B and Ω matrices with various time series length. Panel A reports an extreme case of $T = 100$, and the time series length increases over the panels until $T = 500$ (Panel C). I only report the results for $J = 15$ and $\rho = 0.3$ in this table.

It is clear from Table 3 Panel A that regularization methods significantly outperform OLS in sparsity recognition in the case of B estimation. As expected, sparsity recognition accuracy increases with the increase of T. These patterns are similar to the true negative rate. These observations align with Basu and Michailidis (2015) and Kock and Callot (2015). In the case of inverse covariance estimation, the picture is less clear. LASSO estimators are less powerful; only the adaptive joint LASSO algorithm identifies zero elements in a band Ω matrix. The largest forecast accuracy gains of the adaptive joint LASSO estimator relative to the row-wise LASSO estimators are obtained for short time series. The average sparsity recognition gain is 20% for B and 100% for Ω matrix.

Table 3: Summary of simulation results for sparse coefficient matrix B and band inverse covariance matrix Ω for OLS, RW LASSO, ARW LASSO, and AJ LASSO methods

Panel A: T = 100	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	97%	97%	93%
True negative rate (B)	0%	79%	82%	95%
Sparsity recognition (B)	7%	80%	83%	95%
True positive rate (Ω)	100%	100%	100%	99%
True negative rate (Ω)	0%	0%	0%	26%
Sparsity recognition (Ω)	19%	19%	19%	40%
Panel B: T = 250	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	80%	85%	99%
Sparsity recognition (B)	7%	82%	86%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	22%
Sparsity recognition (Ω)	19%	19%	19%	37%

Panel C: T = 500	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	81%	87%	99%
Sparsity recognition (B)	7%	83%	88%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	14%
Sparsity recognition (Ω)	19%	19%	19%	30%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0.3$. In the table, in the first column, the rows indicate the mean of sparsity recognition performance measures across simulated paths, and columns 2 through 5 represent the alternative methods: OLS, RW LASSO, ARW LASSO, and AJ LASSO. The structures of B and Ω, along with J = 15, are fixed throughout the simulation. Time series length (T) varies across the Panels A, B, and C in the Table.*

Table 4 reports the accuracy gain of the regularization methods in B, Ω, and Θ estimation, compared to the benchmark OLS model (parameters are the same as in Table 3). The estimation accuracy of the regularization methods is significantly better than that of the OLS estimator for all time series lengths. In case of the coefficient matrix, the accuracy gain is stable independently of the time series length and the error form. As time series length increases, this gain decreases for Ω and Θ estimation. This pattern is similar for both error forms (absolute error and Frobenius norm), except B, where the accuracy gain is higher in absolute terms. Interestingly, row-wise LASSO and adaptive row-wise LASSO have similar performance - sometimes RW LASSO outperforms ARW but not significantly - for every time series length. These findings extend Kock and Callot (2015) who investigate the oracle properties of LASSO and adaptive LASSO. In all simulation settings, adaptive LASSO outperformed the other regularization methods. My finding shades this.

Table 4: Accuracy comparison of regularization methods with a simulation of sparse coefficient matrix B and band inverse covariance matrix Ω

Panel A: T = 100	RW LASSO	ARW LASSO	AJ LASSO
MAEE (B,$\hat{B}$)	27%	27%	16%
MAEE (Ω, $\hat{\Omega}$)	75%	76%	46%
MAEE (Θ, $\hat{\Theta}$)	45%	50%	34%
MFBN (B,$\hat{B}$)	54%	53%	47%
MFBN (Ω, $\hat{\Omega}$)	75%	76%	50%
MFBN (Θ, $\hat{\Theta}$)	44%	49%	33%
Panel B: T = 250	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
MAEE (B,$\hat{B}$)	25%	24%	11%
MAEE (Ω, $\hat{\Omega}$)	89%	90%	63%
MAEE (Θ, $\hat{\Theta}$)	48%	51%	39%
MFBN (B,$\hat{B}$)	51%	49%	40%
MFBN (Ω, $\hat{\Omega}$)	88%	89%	68%
MFBN (Θ, $\hat{\Theta}$)	49%	52%	41%
Panel C: T = 500	*RW LASSO*	*ARW LASSO*	*AJ LASSO*

MAEE (B,$\hat{B}$)	27%	24%	10%
MAEE (Ω, $\hat{\Omega}$)	94%	94%	75%
MAEE (Θ, $\hat{\Theta}$)	56%	58%	48%
MFBN (B,$\hat{B}$)	55%	51%	36%
MFBN (Ω, $\hat{\Omega}$)	93%	93%	78%
MFBN (Θ, $\hat{\Theta}$)	57%	60%	51%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0.3$. In the table, in the first column, the rows indicate the mean of sparsity error measures (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm), and columns 2 through 4 represent the three alternative methods: RW LASSO, ARW LASSO, and AJ LASSO. Results are in percentage of the benchmark OLS model. The structures of B and Ω, along with J = 15, are fixed throughout the simulation. Time series length (T) varies across the Panels A, B, and C in the Table.*

To understand the efficiency of the adaptive joint LASSO algorithm, I report in Table 5 the simulation results based on the accuracy of the model in a percentage of the best row-wise LASSO method. It is clear from Table 5 that the adaptive joint LASSO algorithm significantly outperforms the row-wise LASSO methods. The estimated matrices' absolute error and the Frobenius norm are significantly lower. The mean absolute estimation error is 40-60% lower in B estimation, and it increases as the time series length is longer. Average Ω estimation error is also found to be more than 20% better. Interestingly, the accuracy gain is higher for longer time series. In the case of Θ estimation, the error is 15-25% lower than the row-wise LASSO's two versions, and it's growing as time series are shorter. These patterns are similar to Forbenius norm-based relative error, but the accuracy gains are smaller.

Table 5: Summary of the accuracy gains of the AJ LASSO method relative to the better RW LASSO with sparse coefficient (B) and band inverse covariance (Ω) matrix.

Relative Error	*T = 100*	*T = 250*	*T = 500*
MAEE (B,$\hat{B}$)	59%	45%	41%
MAEE (Ω, $\hat{\Omega}$)	61%	71%	79%
MAEE (Θ, $\hat{\Theta}$)	75%	81%	85%
MFBN (B,$\hat{B}$)	88%	81%	69%
MFBN (Ω, $\hat{\Omega}$)	67%	77%	84%
MFBN (Θ, $\hat{\Theta}$)	76%	84%	90%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0.3$. Rows contain the mean of accuracy measures across simulated paths, and columns represent the explored methods. In the table, the first column, the rows indicate the mean of sparsity error measures for (B) and (Ω) matrix (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm). The relative performance of the AJ LASSO in comparison with the better RW LASSO with time series lengths (T) varying from 100 to 500 are presented in Columns 2 through 4. Results are in percentage of the better RW LASSO model. The structures of B and Ω, along with J*

= 15, are fixed throughout the simulation. Relative error calculation is based on Panel A, B, and C of Table 4.

To summarize the results for the basic parameter set in sparse B and band Ω VAR setting, the adaptive joint LASSO estimator outperforms OLS and row-wise LASSO estimators significantly in all time series lengths. The advantage in the estimation accuracy of the algorithm becomes much greater in shorter time series (except B). These results extend the earlier penalized maximum likelihood studies (Rothman et al., 2010; Basu and Michailidis, 2015; Wilms and Croux, 2016; Barbaglia et al., 2020). In the following subsection, I demonstrate that these results are robust to a wide range of simulation - number of time series and correlation between the series - parameters.

5.3.2 Sparse B and band Ω robustness analysis

Robustness analysis regarding the number of time series in the VAR system is important for two reasons. First, in network modeling, the dimension is often high. Several recent papers applied the Diebold-Yilmaz spillover method in a high-dimension setting (Demirer et al., 2018; Greenwood-Nimmo et al., 2019), using RW or ARW LASSO methods. Secondly, considering multiple numbers of time series in the VAR is able to magnify the differences between the methods explored (Basu and Michailidis, 2015; Kock and Callot, 2015).

Simulation results for sparsity recognition in low and high-dimension cases are given in Table 6. I only report the results for $\rho = 0.3$. The sparsity pattern recognition is computed for four different VAR dimensions.

Table 6: Summary of simulation results with sparse coefficient B and band inverse covariance Ω matrix

Panel A: J = 10	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	78%	88%	99%
Sparsity recognition (B)	10%	81%	90%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	14%
Sparsity recognition (Ω)	28%	28%	28%	38%
Panel B: J = 15	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	80%	85%	99%
Sparsity recognition (B)	7%	82%	86%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	22%
Sparsity recognition (Ω)	19%	19%	19%	37%
Panel C: J = 20	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	82%	85%	99%
Sparsity recognition (B)	5%	83%	85%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	24%

Sparsity recognition (Ω)	15%	15%	15%	35%
Panel D: J = 25	*OLS*	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	86%	87%	99%
Sparsity recognition (B)	4%	86%	88%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	22%
Sparsity recognition (Ω)	12%	12%	12%	31%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0.3$. In the table, in the first column, the rows indicate the mean of sparsity recognition performance measures across simulated paths, and columns 2 through 5 represent the alternative methods: OLS, RW LASSO, ARW LASSO, and AJ LASSO. The structures of B and Ω are fixed throughout the simulation. The time series (T) length is fixed at 250, while J (dimension of the matrices) varies across Panels A through D in the Table. Note: Panel B of the table is the same as Panel B of Table 3.*

The general pattern observed in Table 3 can also be recognized in Table 6. In the case of B, as the dimension increases, the sparsity recognition will be better for regularized methods. This pattern is only true for Ω estimation if we use the adaptive joint LASSO algorithm. I also observe that the accuracy gain for Ω sparsity recognition is higher with the increase of N. This gain is between 30-150%. Overall, the adaptive joint LASSO algorithm outperforms all other methods in sparsity pattern recognition, and the largest gain is high-dimension for Ω estimation.

The simulation results for the estimation accuracy of B, Ω, and Θ estimation in the low and high-dimensional designs are given in Table 7. The estimation accuracy of the regularization methods is significantly better than that of the OLS estimator for all VAR dimensions. The accuracy gain increase for both B, Ω, and Θ as the dimension increase. These patterns are similar in both error terms, and the gain is always smaller with Frobenius norm-based error. Similar to Table 4, the difference between the two row-wise LASSO methods is not significant.

Table 7: Accuracy comparison of regularization methods with a simulation of sparse coefficient matrix B and band inverse covariance matrix Ω using a range of the dimension (J) of the system

Panel A: J = 10	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
MAEE ($B,\hat{B}$)	30%	24%	13%
MAEE ($\Omega, \hat{\Omega}$)	93%	94%	75%
MAEE ($\Theta, \hat{\Theta}$)	55%	56%	48%
MFBN ($B,\hat{B}$)	55%	50%	37%
MFBN ($\Omega, \hat{\Omega}$)	93%	93%	79%
MFBN ($\Theta, \hat{\Theta}$)	56%	58%	52%
Panel B: J = 15	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
MAEE ($B,\hat{B}$)	25%	24%	11%
MAEE ($\Omega, \hat{\Omega}$)	89%	90%	63%
MAEE ($\Theta, \hat{\Theta}$)	48%	51%	39%

MFBN (B,$\hat{B}$)	51%	49%	40%
MFBN (Ω, $\hat{Ω}$)	88%	89%	68%
MFBN (Θ, $\hat{Θ}$)	49%	52%	41%

Panel C: J = 20	RW LASSO	ARW LASSO	AJ LASSO
MAEE (B,$\hat{B}$)	22%	23%	8%
MAEE (Ω, $\hat{Ω}$)	86%	87%	52%
MAEE (Θ, $\hat{Θ}$)	44%	49%	32%
MFBN (B,$\hat{B}$)	48%	49%	35%
MFBN (Ω, $\hat{Ω}$)	85%	86%	57%
MFBN (Θ, $\hat{Θ}$)	45%	50%	34%

Panel C: J = 25	RW LASSO	ARW LASSO	AJ LASSO
MAEE (B,$\hat{B}$)	17%	19%	7%
MAEE (Ω, $\hat{Ω}$)	82%	84%	48%
MAEE (Θ, $\hat{Θ}$)	39%	44%	29%
MFBN (B,$\hat{B}$)	44%	43%	32%
MFBN (Ω, $\hat{Ω}$)	79%	82%	52%
MFBN (Θ, $\hat{Θ}$)	38%	43%	29%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0.3$. In the table, in the first column, the rows indicate the mean of sparsity error measures (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm), and columns 2 through 4 represent the three alternative methods: RW LASSO, ARW LASSO, and AJ LASSO. Results are in percentage of the benchmark OLS model. The length of the time series (T) is fixed at 250, while J (dimension of the matrices) varies from 10 to 25 across the Panels A through D in the Table. Panel B of the table is the same as Panel B of Table 4.*

Table 8 reports the results relative accuracy of the adaptive joint LASSO algorithm in the low ($J = 10$) and high-dimensional ($J = 25$) settings (in a percentage of the best row-wise LASSO method). The estimation accuracy of the adaptive joint LASSO algorithm is significantly better than the row-wise LASSO methods in low and high-dimension settings, too. The largest estimation accuracy gains (MAEE) of the ADL estimator relative to the other regularization methods estimator are obtained in the case of B (50-60%). Average Ω and Θ estimation error is also found to be more than 20% and 15% better. Accuracy gain increases in both matrices as the number of time series increases. These patterns are generally in line with the results presented in Table 5, further confirming that regularization on both B and Ω is extremely useful.

Table 8: Summary of the accuracy gains of the AJ LASSO method relative to the better RW LASSO with sparse coefficient (B) and band inverse covariance (Ω) matrix using a range of the dimension (J) of the system

Relative Error	*J = 10*	*J = 15*	*J = 20*	*J = 25*
MAEE (B,$\hat{B}$)	53%	45%	39%	42%
MAEE (Ω, $\hat{Ω}$)	80%	71%	60%	59%
MAEE (Θ, $\hat{Θ}$)	88%	81%	73%	74%
MFBN (B,$\hat{B}$)	74%	81%	72%	74%

MFBN (Ω, Ω̂)	85%	77%	67%	66%
MFBN (Θ, Θ̂)	92%	84%	75%	75%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0.3$. Rows contain the mean of accuracy measures across simulated paths, and columns represent the explored methods. In the table, the first column, the rows indicate the mean of sparsity error measures for (B) and (Ω) matrix (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm). The relative performance of the AJ LASSO in comparison with the better RW LASSO with dimension of the time series (J) - varies from 10 to 25 -, are presented in Columns 2 through 5. Results are in percentage of the better RW LASSO model. The structures of B and Ω, along with $T = 250$, are fixed throughout the simulation. Relative error calculation is based on Panel A, B, and C of Table 7.*

To summarize the simulation result in relation to the VAR system dimension, the adaptive joint LASSO algorithm significantly outperforms the B, Ω and Θ estimation of the row-wise LASSO methods both in low and high-dimension settings, and accuracy gain increasing as the dimension increasing. These findings extend the earlier penalized maximum likelihood applications (Wilms and Croux, 2016; Barbaglia et al., 2020).

Figure 2: Relative estimation accuracy of the AJ LASSO algorithm as a function of the covariance matrix (Σ) Toeplitz-parameter (ρ).
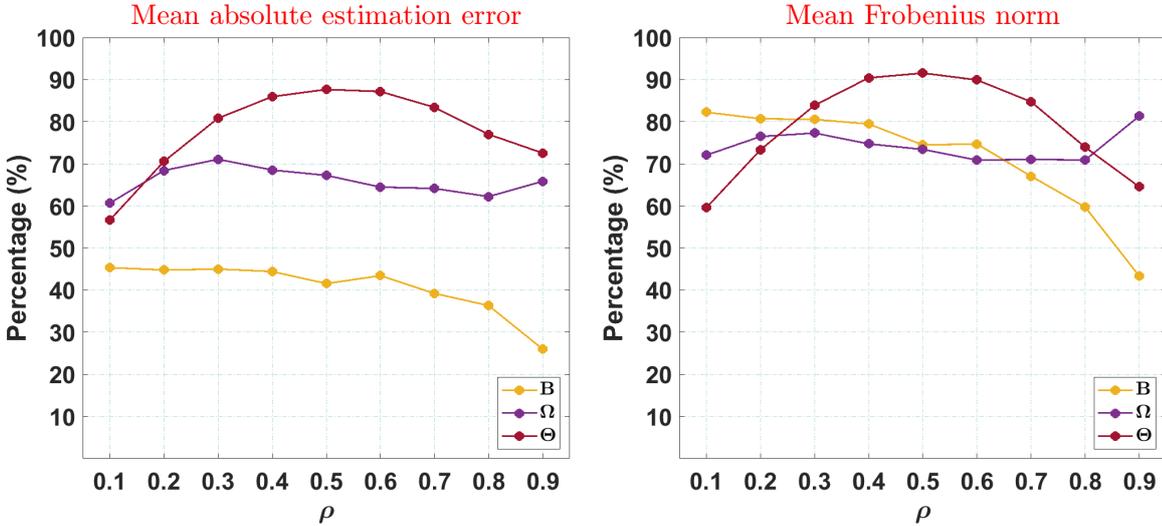
Notes: Results are in percentage of better row-wise LASSO model based on 1000 simulation runs. The structure of B, Ω and Θ, along with $T = 250$ and $J = 15$ is fixed throughout the simulation. The setting corresponding to $\rho = 0.3$ is the same as the second column in Table 5 and in Table 8.

My robustness analysis also considers the variation of the covariance matrix Σ. Different degrees of correlation between the time series error terms (ρ) can influence the estimator performance, especially for regularization methods (Zou, 2006). Several studies, including Basu and Michailidis (2015) and Hecq et al. (2023) confirm this for VAR systems in both low and high-dimension settings, and Rothman et al. (2010) for penalized maximum likelihood estimation. In line with these studies, I vary the ρ parameter of the

Toeplitz error matrix to assess robustness analysis. I consider several values of ρ between $\rho = 0.1$ and $\rho = 0.9$ to obtain a precise evaluation of the methods. I only report the results for $T = 250$ and $J = 15$. Sparsity recognition results are similar to Table 3 and Table 6. I report only the estimation accuracy results. The evolution of the relative estimation accuracy compared to the two row-wise LASSO in a ρ-based setting is plotted in Figure 2. Detailed simulation results are plotted in Figure A1 and Figure A2.

It is clear from Figure 2 that the proposed AJ LASSO significantly outperforms the RW LASSO methods for all alternative ρ values. Both mean absolute estimation error (MAEE) and mean Frobenius norm (MFBN) measures of the B, Ω, and Θ matrices are significantly lower in case of AJ LASSO method in comparison with other methods. However, it is important to note the differences in patterns in the change of MAEE and MFBN with alternative ρ values. In the case of B, as the correlation ρ increases, the algorithm's accuracy gain also increases. The gain is stable for Ω between 30-40%. In parallel with this, the accuracy gain has a hump-shaped pattern for Θ, primarily due to the improvement in the estimation of the autoregressive coefficient matrix (B). Simulation results indicate that in highly correlated systems, the relative performance of the AJ LASSO method performs remarkably well, especially for the coefficient matrix (B) estimation.

To summarize the simulation results for the sparse B and band Ω VAR setting, I conclude that the adaptive joint LASSO estimator outperforms row-wise LASSO estimators significantly in the B, Ω and Θ estimation for all parameter sets. The advantage in the estimation accuracy becomes much greater in shorter time series and in high-dimension settings. In the following subsection, I demonstrate that these results are robust to different covariance structures.

5.3.3 Sparse coefficient and different (sparse and dense) inverse covariance matrices

Assuming various data-generating processes represented by various covariance structures can have different theoretical and empirical motivations. To assess the robustness of my results related to the covariance structure in this subsection, I investigate the effect of sparse and dense Ω on the relative estimation performance of the adaptive joint LASSO algorithm. In line with Kock and Callot (2015) and Hecq et al. (2023) I represent sparse Ω with a Toeplitz-type covariance matrix with $\rho = 0$. This covariance structure is a common choice, assuming the contemporaneous correlation between time series is zero. All regularized method outperforms the OLS model, and sparsity recognition pattern is similar to the earlier results (Table A1, Table A3, Table A5 and A7), so I focus on the estimation accuracy gain of my novel approach compared to the two row-wise LASSO model. In the case of sparse and dense Ω, the correlation parameter is fixed through the simulation, while T and J vary. Detailed simulation results are represented in Table A2, Table A4, Table A6, and Table A8 which are in the similar structure like tables in the earlier two subsections.

Simulation results based on error terms drawn from a sparse Ω (Toeplitz-type Σ with $\rho = 0$) are presented in Tables 9-10. These tables are structured similarly to Table 5 and Table 8.

Table 9: Summary of the accuracy gains of the AJ LASSO method relative to the better RW LASSO with sparse coefficient (B) and band inverse covariance (Ω) matrix.

Relative Error	$T = 100$	$T = 250$	$T = 500$
MAEE ($B,\hat{B}$)	60%	44%	39%
MAEE ($\Omega, \hat{\Omega}$)	27%	23%	16%
MAEE ($\Theta, \hat{\Theta}$)	26%	17%	9%
MFBN ($B,\hat{B}$)	92%	80%	68%
MFBN ($\Omega, \hat{\Omega}$)	40%	39%	37%
MFBN ($\Theta, \hat{\Theta}$)	31%	19%	10%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0$. Rows contain the mean of accuracy measures across simulated paths, and columns represent the explored methods. In the table, the first column, the rows indicate the mean of sparsity error measures for (B) and (Ω) matrix (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm). The relative performance of the AJ LASSO, in comparison with the better RW LASSO with time series lengths (T), varies from 100 to 500, are presented in Columns 2 through 4. Results are in percentage of the better RW LASSO model. The structures of B and Ω, along with J = 15, are fixed throughout the simulation. Relative error calculation is based on Panel A, B, and C of Table A2.*

The adaptive joint LASSO estimator always obtains lower MAEE values than the row-wise LASSO estimators. The difference in the accuracy is significant in every setting independently of T and J. The smallest estimation accuracy gains of the AJ LASSO relative to the other estimators are obtained for the coefficient matrix B. This is generally the opposite conclusion of the results presented in the earlier substructures, further confirming that the robustness analysis for the covariance structure is necessary. On the other hand, the accuracy gain of the new method for the Ω and Θ is much higher than in the band Ω setting. This gain is between 65-85% for Ω and a bit higher than 75-90% for Θ estimation. These results are almost double and triple, as we observed in Table 4 and Table 7. The results further confirm the relevancy of my new method.

As we can see in Table A1 and Table A3, the new method tends to capture better sparsity of the inverse covariance matrix (Ω). This is represented by higher values and the true negative rate of (Ω). While in the band (Ω) setting (Table 3 and Table 6), these values ranged between 10-30%, now they are between 75-90%, and almost independent from parameters T and J.

Table 10: Summary of the accuracy gains of the AJ LASSO method relative to the better RW LASSO with sparse coefficient (B) and sparse inverse covariance (Ω) matrix using a range of the dimension (J) of the system

Relative Error	$J = 10$	$J = 15$	$J = 20$	$J = 25$
$MAEE\ (B,\hat{B})$	49%	44%	40%	40%
$MAEE\ (\Omega,\ \hat{\Omega})$	33%	23%	23%	21%
$MAEE\ (\Theta,\ \hat{\Theta})$	18%	17%	17%	17%
$MFBN\ (B,\hat{B})$	74%	80%	76%	74%
$MFBN\ (\Omega,\ \hat{\Omega})$	51%	39%	37%	35%
$MFBN\ (\Theta,\ \hat{\Theta})$	21%	19%	20%	20%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0$. Rows contain the mean of accuracy measures across simulated paths, and columns represent the explored methods. In the table, the first column, the rows indicate the mean of sparsity error measures for (B) and (Ω) matrix (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm). The relative performance of the AJ LASSO in comparison with the better RW LASSO with dimension of the time series (J) - varies from 10 to 25 -, are presented in Columns 2 through 5. Results are in percentage of the better RW LASSO model. The structures of B and Ω, along with T = 250, are fixed throughout the simulation. Relative error calculation is based on Panel A, B, and C of Table A4.*

In a sparse inverse covariance matrix setting, the new method of accounting for the characteristics of the DGP for the estimation process proves extremely useful. The pattern of results was found to be similar for the Frobenius norm error. As the dimension of the time series increases, the relative error diminishes in the case of all matrices (B, Ω) except Θ. This means that the new method is useful in low-dimension systems, too, if the Ω) is sparse. The pattern for B and Ω is similar to the case of band Ω covariance structure. However, from the view of the time series length parameter, as it decreases, the difference decreases as well. That means the new method can be useful for longer window sizes of the DY network analysis. In general, the enhancements in forecast accuracy are more significant for these sparse (Ω) designs than for the sparse (Ω) designs in Table 4 and Table 7. Besides that, in the sparse Ω setting, the accuracy gain is significantly high for low-dimension systems.

Simulation results based on error terms drawn from a dense Ω are presented in Tables 11-12. These tables are structured similarly to Table 5 and Table 8. Similar to the earlier covariance structures, the newb AJ LASSO method always obtains lower MAEE values than the row-wise LASSO estimators. The difference in the accuracy is significant in every setting independently T and J, but smaller than the case of sparse Ω. The results are closer to the band Ω setting (the similarity can be observed with Table 4 and Table 7).

Table 11: Summary of the accuracy gains of the AJ LASSO method relative to the better RW LASSO with sparse coefficient (B) and dense inverse covariance (Ω) matrix.

Relative Error	$T = 100$	$T = 250$	$T = 500$
MAEE (B,$\hat{B}$)	46%	34%	30%
MAEE (Ω, $\hat{\Omega}$)	62%	70%	83%
MAEE (Θ, $\hat{\Theta}$)	89%	85%	82%
MFBN (B,$\hat{B}$)	68%	60%	51%
MFBN (Ω, $\hat{\Omega}$)	66%	73%	85%
MFBN (Θ, $\hat{\Theta}$)	88%	81%	74%

Notes: *Results are based on 1000 simulation runs. Rows contain the mean of accuracy measures across simulated paths, and columns represent the explored methods. In the table, the first column, the rows indicate the mean of sparsity error measures for (B) and (Ω) matrix (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm). The relative performance of the AJ LASSO in comparison with the better RW LASSO with time series lengths (T) varying from 100 to 500 are presented in Columns 2 through 4. Results are in percentage of the better RW LASSO model. The structures of B and Ω, along with J = 15, are fixed throughout the simulation. Relative error calculation is based on Panel A, B, and C of Table A6.*

The largest estimation accuracy gains of the new estimator estimator relative to the others are obtained for the coefficient matrix (B). The accuracy gain is between 55-70%. It's almost 50% higher than the case of band Ω. The new method tends to capture better sparsity of the coefficient matrix (B) in this case.

Table 12: Summary of the accuracy gains of the AJ LASSO method relative to the better RW LASSO with sparse coefficient (B) and dense inverse covariance (Ω) matrix using a range of the dimension (J) of the system

Relative Error	$J = 10$	$J = 15$	$J = 20$	$J = 25$
MAEE (B,$\hat{B}$)	32%	34%	32%	30%
MAEE (Ω, $\hat{\Omega}$)	79%	70%	63%	63%
MAEE (Θ, $\hat{\Theta}$)	87%	85%	83%	83%
MFBN (B,$\hat{B}$)	48%	60%	59%	57%
MFBN (Ω, $\hat{\Omega}$)	82%	73%	67%	67%
MFBN (Θ, $\hat{\Theta}$)	82%	81%	75%	75%

Notes: *Results are based on 1000 simulation runs. Rows contain the mean of accuracy measures across simulated paths, and columns represent the explored methods. In the table, the first column, the rows indicate the mean of sparsity error measures for (B) and (Ω) matrix (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm). The relative performance of the AJ LASSO in comparison with the better RW LASSO with dimension of the time series (J) - varies from 10 to 25 -, are presented in Columns 2 through 5. Results are in percentage of the better RW LASSO model. The structures of B and Ω, along with T = 250, are fixed throughout the simulation. Relative error calculation is based on Panel A, B, and C of Table A8.*

In a dense inverse covariance matrix setting, the new method, which accounts for this characteristic of the DGP for the two-step estimation process, proves extremely useful. The information of dense Ω in the (B) estimation process is the key point of the method efficiency (Rothman et al., 2010). The sensitivity of accuracy gain for different T and J parameters are similar to the sparse Ω setting. As the time series length decreases, the difference decreases as well. Besides that, if the dimension of the series increases, the relative error diminishes in the case of all matrices (B, Ω) except Θ. This finding confirms the results of Tables 9-10, and I conclude that the new method is useful in low-dimension systems too.

To summarize the results of the covariance structure robustness analysis, I can define three new findings. First, the accuracy gain is sensitive to the covariance structure; in the case of sparse Ω setting it is remarkable (between 75-90% for Θ estimation). The results prove that the new method efficiently accounts for the characteristics of the sparse Ω in the estimation via the penalized maximum likelihood objective function. Secondly, this gain is higher in the case of the dense Ω settings, too (compared to the basic band Ω set), especially for the estimation of coefficient matrix (B). The information of dense Ω in the (B) estimation process is the key point of the new method efficiency. Thirdly, in both cases, the gain is also significant in the low-dimension systems $(J = 10)$. That means the new method can be useful for DY network analysis in low-dimension systems. In the following subsection, I demonstrate that the main results of the base set are robust to the dense coefficient matrix (B) too.

5.3.4 Dense coefficient matrix

Assuming various data-generating processes, represented by various VAR coefficient structures, can have different theoretical and empirical motivations. To assess the robustness of the basic set related to the autoregressive structure (sparse B) in this subsection, I investigate the effect of dense B on the relative estimation performance of the adaptive joint LASSO algorithm. The choice of the dense B is in line with Kock and Callot (2015) and Hecq et al. (2023).

Similar to the earlier cases, the regularized methods outperform the OLS model, and the sparsity recognition pattern is similar to the sparse B set. Therefore, I focus on the estimation accuracy gain of my novel approach compared to the two row-wise LASSO model. Besides that, I compare the results to the sparse B case. I analyze two cases for the covariance structure: band and dense Ω. The reason behind the decision is the same in the sparse B case. In the case of band and dense Ω, the correlation parameter is fixed through the simulation (equal with sparse B setting), while T and J vary. Detailed simulation results are represented in Tables A9-A10 and Tables A12-A13 which are in the similar structure like tables in the earlier two subsections.

Simulation results based on error terms drawn from a band Ω (Toeplitz-type Σ with $\rho = 0.3$) and dense B are presented in Fig 3 and Fig 4. The figures report the accuracy gain of the new AJ LASSO estimation method in both sparse and dense B settings, while the covariance structure is band.

Figure 3: Relative estimation accuracy of the AJ LASSO algorithm as a function of the length of time series (T).
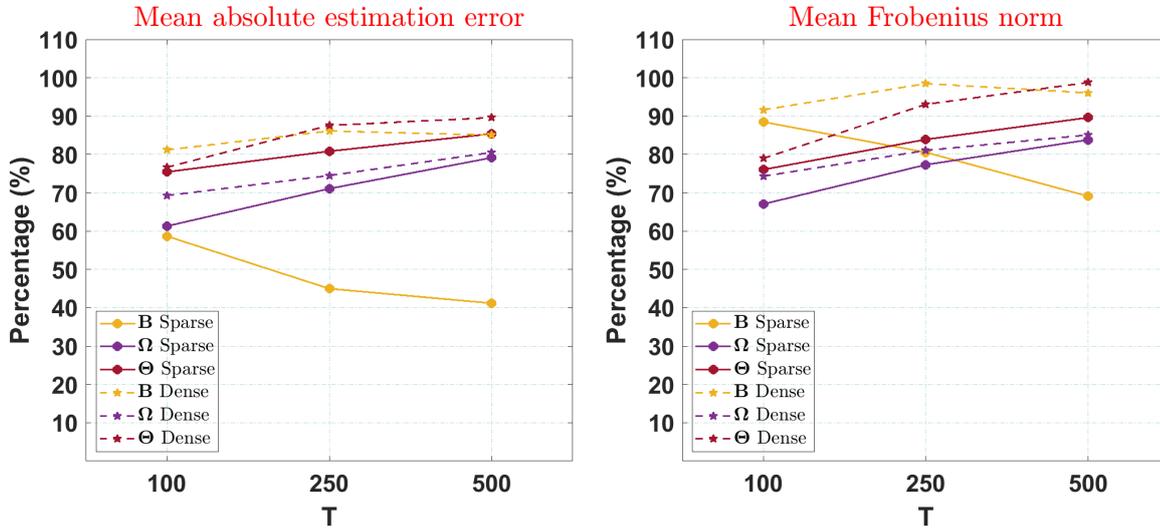
Notes: Results are in percentage of better row-wise LASSO model and are based on 1000 simulation runs. The structure of B, Ω and Θ, along with $J = 15$ and $\rho = 0.3$ is fixed throughout the simulation. Solid line represents sparse (B), dashed line represents dense (B), the setting corresponding to $T = 250$ is the same as the second column in Table 5, Table 8 and in Tables A9-A10.

The pattern of the results is very similar to the dense B case. However, the AJ LASSO seems to have poor estimation accuracy for B in every T and J setting compared to the

Figure 4: Relative estimation accuracy of the AJ LASSO algorithm as a function of the number of time series (J).
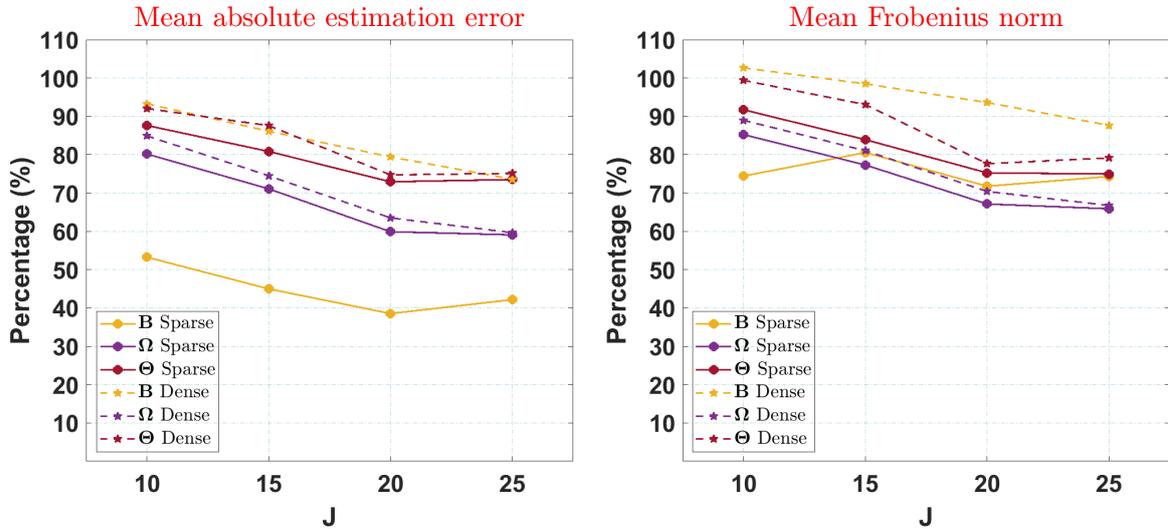
Notes: Results are in percentage of better row-wise LASSO model based on 1000 simulation runs. The structure of B, Ω and Θ, along with $T = 250$ and $\rho = 0.3$ is fixed throughout the simulation. Solid line represents sparse (B), dashed line represents dense (B), the setting corresponding to $J = 15$ is the same as the second column in Table 5, Table 8, Table A9 and in Table A10.

earlier findings. This result is based on a significant under-performance in sparsity pattern recognition for the inverse covariance matrix. In parallel with this, the estimation of the DY spillover matrix (Θ) is worse, but the difference is generally not significant.

The accuracy gains are lower, they are between 5-25% for B, 15-40% Ω and 10-23% for Θ. However, these numbers are smaller than the earlier setting, and the method still significantly outperforms the row-wise LASSO estimation. The patterns for increasing the system dimension and length of time series are similar to the sparse (B) case. Overall, the accuracy gain is higher for high-dimension and short time series. These results extend the earlier LASSO-based VAR modeling studies, focusing on simulation (Kock and Callot, 2015; Hecq et al., 2023).

Besides the estimation accuracy of B, Ω, and Θ, for the dense B, I calculated the number of large inaccuracies in the parameter estimation. Based on my definition, the parameter estimation has large inaccuracy if the estimation error is more than 20%. This measure can capture the outliers in the estimation. Lower measures mean better estimation.

Table 13 reports the results of the relatively large inaccuracy (relative number of parameters where the estimation error is more than 20%) for the regularization methods compared to OLS. Time series length (T) varies across the panels of the table. The structure of B, Ω and Θ, along with $J = 15$ is fixed throughout the simulation. Similar information is reported in Table A11, where T is fixed and J varies.

Table 13: Summary of large inaccuracies in the estimation with a dense coefficient (B) and band inverse covariance (Ω) matrix

Panel A: T = 100	RW LASSO	ARW LASSO	AJ LASSO
RLIA (B,$\hat{B}$)	41%	32%	17%
RLIA (Ω, $\hat{\Omega}$)	75%	77%	63%
RLIA (Θ, $\hat{\Theta}$)	94%	94%	92%
Panel B: T = 250	RW LASSO	ARW LASSO	AJ LASSO
RLIA (B,$\hat{B}$)	49%	35%	17%
RLIA (Ω, $\hat{\Omega}$)	70%	81%	72%
RLIA (Θ, $\hat{\Theta}$)	51%	33%	17%
Panel C: T = 500	RW LASSO	ARW LASSO	AJ LASSO
RLIA (B,$\hat{B}$)	51%	33%	17%
RLIA (Ω, $\hat{\Omega}$)	91%	91%	82%
RLIA (Θ, $\hat{\Theta}$)	99%	99%	100%

Notes: *Error terms are drawn from a Toeplitz-type covariance matrix with $\rho = 0.3$. Results are based on 1000 simulation runs. Rows contain the mean of wrongly estimated parameters across simulated paths, and columns represent the explored methods. RLIA = relative large inaccuracy (relative number of parameters where the estimation error is more than 20%). Results are in percentage of the benchmark OLS model. Time series length (T) varies across the panels of the table. The structure of B, Ω and Θ, along with $J = 15$ is fixed throughout the simulation.*

Based on Table 13 and Table A11, the regularization methods outperform the OLS estimator significantly in all settings in RLIA (relative large inaccuracy) measure. Besides that, for B estimation, the adaptive joint LASSO outperforms the row-wise LASSO methods, and the parameters where the estimation error is more than 20% is almost 50% lower. Interestingly, from this perspective, the ARW LASSO almost always outperforms RW LASSO.

It is important to note that I have shown earlier in Eq. (20) that the structure of the coefficient matrix influences the accuracy gain of the AJ LASSO method. To address the robustness of my methods, I analyze this sensitivity for dense B in the following subsections. As I concluded earlier, the structure of the coefficient matrix has large influence on the accuracy gain of the AJ LASSO method. In the next part of this subsection, I analyze this sensitivity for dense B.

Simulation results based on error terms drawn from a dense Ω and dense B are presented in Fig 5 and Fig 6. The structure of plots is very similar to Fig 3 and Fig 4. The pattern of the results (difference between the sparse and band B settings) are very similar to the band Ω case; however, the AJ LASSO seems to have relatively poor estimation accuracy nut only for B but Ω too, in every T and J setting. In parallel with this, the estimation of the DY spillover matrix (Θ) is worse. However, as we conclude in sparse B and dense Ω settings, if the inverse covariance matrix is non-sparse, it can be useful for the AJ LASSO estimator. The results are worse than in the sparse B setting but still outperform significantly the best row-wise LASSO estimation.

The accuracy gain is between 30-50% for B, 0-25% for Ω and 0-20% for Θ. These results are significant even though the coefficient matrix is non-sparse. The accuracy gain

Figure 5: Relative estimation accuracy of the AJ LASSO algorithm as a function of the length of time series (T).
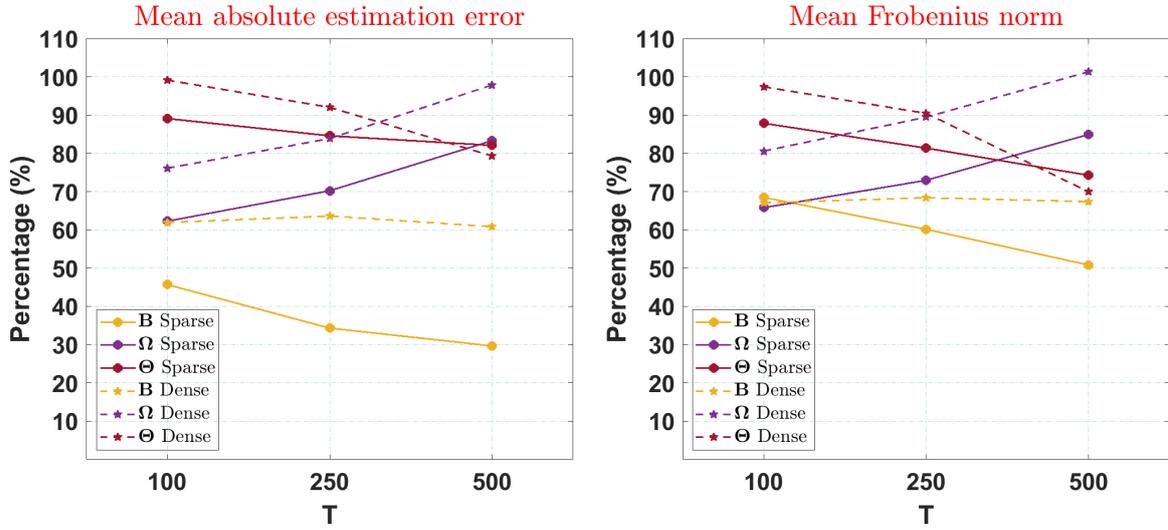
Notes: Results are in percentage of better row-wise LASSO model and are based on 1000 simulation runs. The structure of B, Ω and Θ, along with $J = 15$ and $\rho = 0.3$ is fixed throughout the simulation. Solid line represents sparse (B), dashed line represents dense (B), the setting corresponding to $T = 250$ is the same as the second column in Tables 11-12, Table A12 and in Table A10.

Figure 6: Relative estimation accuracy of the adaptive joint LASSO algorithm as a function of the number of time series (J).
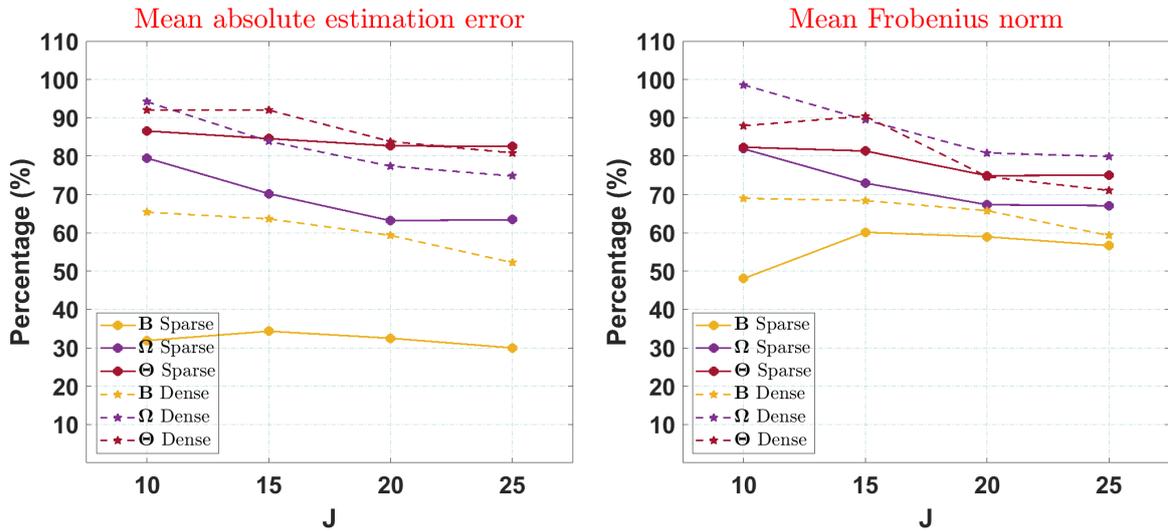
Notes: Results are in percentage of better row-wise LASSO model and are based on 1000 simulation runs. The structure of B, Ω and Θ, along with $T = 250$ and $\rho = 0.3$ is fixed throughout the simulation. Solid line represents sparse (B), dashed line represents dense (B), the setting corresponding to $J = 15$ is the same as the second column in Tables 11-12, Table A12 and in Table A13.

is higher if the time series length (T) is longer and stable as the dimension (J) varies. These results are similar to the sparse B and dense Ω settings. Similar to the dense Ω case in this subsection, I calculated the large inaccuracies in the parameter estimation.

Table 14 reports the results of the relative large inaccuracy (relative number of parameters where the estimation error is more than 20%) for the regularization methods compared to OLS if the Ω is dense. Time series length (T) varies across the panels of the table. Similar information is reported in Table A14, where T is fixed and J varies.

Table 14: Summary of large inaccuracies in the estimation with dense coefficient (B) and dense inverse covariance (Ω) matrix

Panel A: T = 100	RW LASSO	ARW LASSO	AJ LASSO
RLIA ($B,\hat{B}$)	31%	28%	17%
RLIA ($\Omega, \hat{\Omega}$)	94%	92%	80%
RLIA ($\Theta, \hat{\Theta}$)	61%	63%	61%
Panel B: T = 250	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
RLIA ($B,\hat{B}$)	39%	29%	17%
RLIA ($\Omega, \hat{\Omega}$)	98%	97%	83%
RLIA ($\Theta, \hat{\Theta}$)	48%	33%	17%
Panel C: T = 500	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
RLIA ($B,\hat{B}$)	48%	33%	17%
RLIA ($\Omega, \hat{\Omega}$)	99%	96%	87%
RLIA ($\Theta, \hat{\Theta}$)	27%	27%	12%

Notes: *Rows contain the mean of wrongly estimated parameters across simulated paths, and columns represent the explored methods. RLIA = relative large inaccuracy (relative number of parameters where the estimation error is more than 20%). Results are in percentage of the benchmark OLS model. Time series length (T) varies across the panels of the table. The structure of B, Ω, and Θ, along with $J = 15$, is fixed throughout the simulation.*

Based on Table 14 and Table A14, the regularization methods similar to the band Ω setting outperform OLS estimator significantly in all settings in RLIA measure. However, in this case, besides in B in Ω, estimation outperforms the adaptive joint LASSO the other methods. The parameters where the estimation error is more than 20% are almost 50% lower, and the accuracy gain increases as T and J increase.

To summarize the results for the dense coefficient matrix (B), I can define two new findings. First, the accuracy gain of the AJ LASSO method is smaller for non-sparse B but still significant. The smallest gains of the new estimator relative to the others are obtained for non-sparse B and band Ω. Secondly, even in these cases, the AJ LASSO can be useful for high-dimension VAR system (and DY network) modeling because the number of the estimated parameters where the error is more than 20% is almost 50% lower than for the other methods.

5.3.5 Summary of the simulation results

In this subsection, I summarize my extensive Monte Carlo simulation. To illustrate the results, I compared the AJ LASSO performance with the better RW LASSO method in all simulated cases in Figure 7. The figure represents the relative MAEE accuracy for my method in B, Ω, and Θ estimation for the selected DGPs.

Figure 7: Summary of relative estimation accuracy of the AJ LASSO algorithm compared to the better RW LASSO for different coefficient (B) and inverse covariance (Ω) matrices
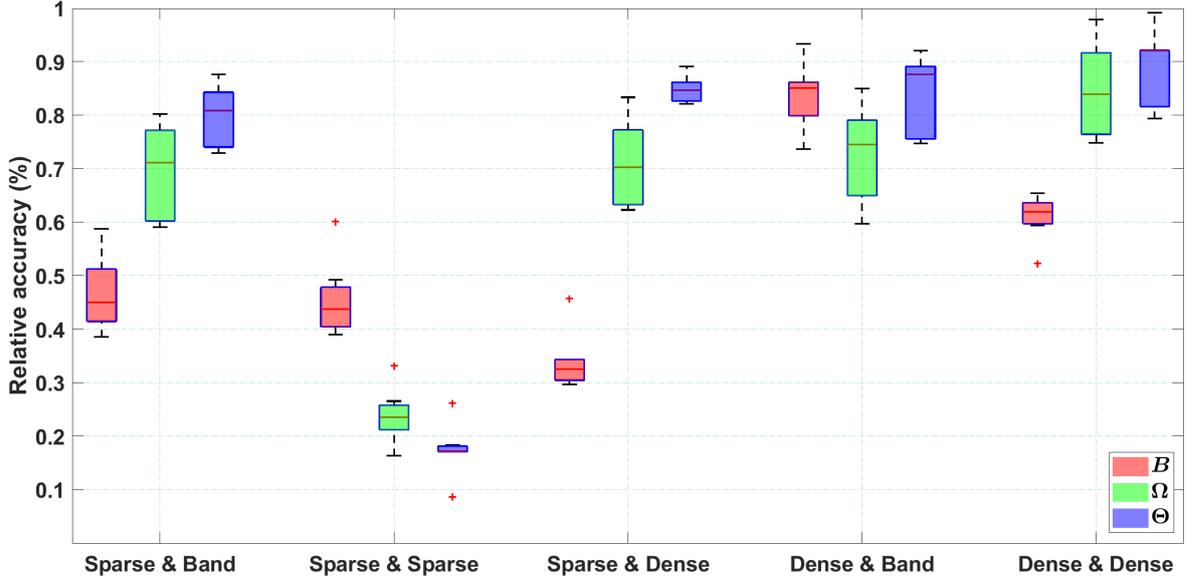
Notes: Red, green, and blue boxplots represent the relative estimation accuracy of the AJ LASSO algorithm compared to the better RW LASSO for the coefficient (B), inverse covariance (Ω), and DY spillover (Θ) matrices. The estimation error is measured by MAEE. The cases on the x-axes are the following: sparse B and band Ω; sparse B and sparse Ω; sparse B and dense Ω; dense B and band Ω; dense B and dense Ω. T, J, and ρ are varies through the cases.

The main findings are the following observations:

1. Based on Figure 7, in all the five cases (sparse and dense B; band, sparse and dense Ω), the AJ LASSO estimator outperforms RW LASSO estimators significantly in both B, Ω and Θ estimation.

2. For the basic setup (sparse coefficient B and band inverse covariance Ω matrices), simulation results indicate that in highly correlated systems, the relative performance of the AJ LASSO method performs remarkably well, especially for the coefficient matrix (B) estimation.

3. The accuracy gain is sensitive to the covariance structure, in case of sparse Ω – the common assumption in theoretical econometrics – setting it is the highest (between 75-90% for Θ estimation). The results demonstrate that the new method efficiently accounts for the characteristics of the sparse Ω in the estimation via the jointly penalized (regularization on both B and Ω) objective function.

4. The relative accuracy is high in the case of the dense Ω settings, too (compared to the basic band Ω set), especially for the estimation of coefficient matrix (B).

Incorporating dense Ω into the B estimation process is the critical point of the new method efficiency. The superior accuracy for the DY spillover matrix estimation (Θ) is still present with dense Ω.

5. In both sparse and dense Ω cases, the gain is also significant in the low-dimension systems ($J = 10$). That means the new method can be useful for DY network analysis in low-dimension systems too.

6. The accuracy gain of the AJ LASSO method is smaller for dense B but still significant. The smallest gains of the new estimator relative to the others are obtained for the dense B and band Ω case.

7. Even when AJ LASSO is only marginally superior (dense B), there are still benefits because the number of extreme outliers where the error is more than 20% in the estimation is significantly lower than for the other methods.

8. In general, the advantage in the estimation accuracy becomes much greater in shorter time series. Relative estimation accuracy is sensitive to the dimension of the system, the highest in high-dimension settings.

Based on these results, AJ LASSO can be useful for high-dimension VAR system (and DY network) modeling. In Chapter 6, I will use my proposed method to estimate a VAR model for the DY volatility and illiquidity network.

6 Chapter 6. Combination of DY framework with event study methodology

This chapter is based on two working papers as a result of my collaboration with Áron Dénes Hartvig. Earlier versions of the research were presented at the 11th and 12th Annual Financial Market Liquidity (AFML) Conference[30] and the 2021 Annual Hungarian MKE Conference[31]. Major modifications are made to align the papers' content with the dissertation format.

6.1 Event study framework for financial networks

Financial linkages and contagion channels often go unrevealed due to the complexity of the financial network. Financial interconnectedness is diverse and heterogeneous; furthermore, it is constantly evolving; thus, it is difficult to observe it empirically. Event study analysis tools are capable of uncovering these connections.

Various techniques have been referenced as event study analysis in economic and sociological studies (Altman et al., 1977) over the years. The main difference between cross-sectional analysis and event study methods is that the latter incorporates and considers changes in independent variables over time (Hill et al., 1996). Among others, MacKinlay (1997), Corrado (2011) and Miller (2023) provide comprehensive overview of the methodology of event studies.

The influence of financial events like mergers or acquisitions on stock prices has already been widely investigated in the literature (Malatesta and Thompson, 1985; Duso et al., 2010; Dilshad, 2013; Adnan et al., 2016). Although studies with event study analysis tend to focus on acquiring and target firms, there are a number of financial events suitable for the analysis. Specifically, a failure of a financial institution or large currency devaluation in the context of financial system is a suitable candidate for event study analysis. events might alter the financial system. As discussed in previous chapters, understanding the transition of volatility spillover in financial networks from the initially impacted institution is of regulatory and academic interest. To provide the necessary "zoomed-in" view, combining the event study method with the DY spillover index might be a suitable approach. In the following sections, I present the importance and the possible methods to execute event study analysis of systemic risk.

6.1.1 Regulatory importance of financial event history analysis

From a regulatory perspective, event study analysis from a financial network perspective serves as a powerful tool in various fields. First, balance sheet-based methods or macro-level analysis can only observe a slower, general evolution of systemic risk. During crises, financial connections change rapidly. Therefore, daily network monitoring can help us understand the progression at each stage of the prevailing financial crisis.

Second, the GFC has revealed, beyond the unimaginable, the magnitude of potential economic fallout caused by the failure of financial institutions. Even a single bankruptcy can lead to a cascade of issues in country, as we have seen recently with the failure of

---

[30]AFML Previous Conferences: https://www.uni-corvinus.hu/ind/events/annual-financial-market-liquidity-conference/afml-2020-previous-conferences/?lang=en

[31]2021 Annaul Hungarian MKE Conference: https://www.mktudegy.hu/konferencia/magyar-kozgazdasagtudomanyi-egyesulet-mke-konferenciafelhivas-2021-december-21-22/38282/

SVB, or a group of bank failures in one country may have international ramifications. Consequently, the linkages of a troubled institution need to be mapped to other network participants to estimate the potential consequences. Regulators must consider all important network linkages of a troubled institution when they are considering a possible bailout and its extent.

Third, as Acemoglu et al. (2015) showed, the connectedness of the financial network facilitates system stability up until a certain threshold of shock levels and number of shocks. However, if the level of financial distress exceeds a certain threshold, local structural properties of the financial network dictate the extent of contagion of shocks. In other words, higher financial interconnectedness makes the system more sensitive and more prone to contagion. Thus not only the analysis of SIFIs, but the inclusion of the whole system is necessary. Institutions on the periphery may move to the center of the financial network owing to a significant shock that restructures the linkages.

Finally, by dissecting each step (e.g., each day or week) leading up to a major financial crisis allows us to have a "slow-motion" view of the changes in the network and can provide important data for future stress testing. A stress test, introduced after the GFC by the US federal reserve, is an analysis or simulation designed to determine the resistance of a given financial institution to deal with an economic shock. To design relevant and insightful scenarios, trigger events causing material changes in the financial network structural could be valuable. In the absence of such valuable information, countries major banks used to stress test for lower-than-expected GDP increases.

6.1.2 Estimation of structural changes in networks

Leveraging on the extensive financial network literature (Elliott et al., 2014; Acemoglu et al., 2015), I discuss three methods for providing insights about the network status, specifically to differentiate across tranquil and turmoil periods. Specifically, in order of complexity I discuss three methods: visualization, measure and similarity statistics.

A practical way to measure the importance of a node is to count its edges in the tenth percentile of all edges (Diebold and Yılmaz, 2014). It is a simple and useful measure; however, it can only be applied to individual nodes and has a subjective parameter, the percentile. Otherwise, Demirer et al. (2018) used the relative weight of cross- and within-country connections to explore the dynamics of global bank connectedness. Moreover, other simple measures describe the system from a network perspective. It is common to describe the correlation matrix of the financial system with the mean, the standard deviation, or the kurtosis of the edges (Onnela et al., 2003; Coelho et al., 2007; Fenn et al., 2011), and it has also been applied to systemic risk analysis to MES and ΔCoVaR connections (Billio et al., 2016).

In addition, there are more sophisticated metrics called similarity statistics to quantify the structural evolution of the financial network. These measures estimate the similarity of different adjacency matrices. The Jaccard index, developed by Jaccard (1912), is a statistic used for gauging the similarity and diversity of sample sets. The index provides information on the number of links between sample periods, thus only applicable to networks with unweighted edges. Jaccard index has been used to observe correlation in financial systems (Nobi et al., 2014; Zhao et al., 2018), and systemic risk with Granger causality networks (Li et al., 2019; Dungey et al., 2019) or with DY networks (Chowdhury et al., 2019). Besides the Jaccard index, the Frobenius norm is often applied to measure changes in correlation networks (Barnett and Onnela, 2016; Cabrieto et al., 2018) and systemic risk (Torri et al., 2018). Finally, the ordinal association between two measured

quantities can be measured by Kendall rank correlation coefficient or Kendall's τ coefficient. Kenett et al. (2015) and Junior et al. (2015) used Kendall's τ to explore the rank correlation of stocks over time.

Alternatively, status change can be identified with more complex statistical tests. In the $\Delta CoVaR$ framework Castro and Ferrari (2014) and Bernal et al. (2014) developed a test relying on the Bootstrap Kolmogorov-Smirnov (KS) test (Abadie, 2002)) to assess whether a sector/institution is statistically significantly risky for the system, and also to measure the difference between the sectors/institutions. KS statistic has been applied to test the significance of the $\Delta CoVaR$ measure to identify SIFIs (Bernal et al., 2014; Castro and Ferrari, 2014; Drakos and Kouretas, 2015; Liu, 2017; de Mendonça and da Silva, 2018), but Hurlin et al. (2017) also used it to compare the riskiness of FIs by testing ΔCoVaR, MES, and %SRISK (Systemic Risk Measure) measures.

In Table 15, I summarise the methods that can be used to estimate structural changes in networks.

Table 15: Event study approaches to analyse financial networks

	Complexity	Level	Method
Visualization	Low	Individual and system	Number of node's edges, the weight of the pairwise connections, the density of the connections.
Measure	Medium	Individual and system	Number of edges, the relative weight of cross- and within-group connections, the mean, the standard deviation, or the kurtosis of the edges.
Similarity statistics	High	System	Statistics: Jaccard index, Frobenius norm, Kendall's τ. Statistical test: Bootstrap Kolmogorov-Smirnov test.

6.1.3 Moving Block Bootstrap method

By constructing bootstrap confidence intervals for spillover measures on macro (total spillover index) and micro levels (net pairwise spillover indices) of the estimated network, I can distinguish significant spillover changes from others. As a result, we can evaluate the magnitude and statistical significance of the shocks. Following Buse et al. (2022) and Greenwood-Nimmo and Tarassow (2022), a bootstrap-based analysis is taken to explore the uncertainty of the estimated spillover measures. The choice of the bootstrap method needs to reflect the dependency structure inherent in the analyzed time series.

Various bootstrapping techniques have been designed in the last two decades to preserve the asymptotically valid inference for VAR estimation (Lahiri and Lahiri, 2003). The most common methods in the VAR context are recursive-design wild bootstrap, fixed-design wild bootstrap, and pairwise bootstrap (Brüggemann et al. (2016)). In my analysis, I apply the residual-based moving-block bootstrap method developed by (Brüggemann et al., 2016). Both Buse et al. (2022) and Greenwood-Nimmo and Tarassow (2022) applied this method in the DY framework due to its desirable features. First, it doesn't require knowledge about the distribution of the DY spillover indices, it is suit-

able for small samples, and third, it remains valid under conditional heteroscedasticity (Greenwood-Nimmo and Tarassow, 2022). For the validity of the residual moving-block bootstrap algorithm, I refer to Brüggemann et al. (2016).

The logic behind the MBB approach is to simulate $s = 1, 2, \ldots, S$ bootstrap samples of the residuals from the estimated VAR model and re-estimate the model on the bootstrap sample too. After re-estimating the VAR model for every bootstrap sample (MBB-VAR), I calculate the DY spillover indices, and the resulting bootstrap spillover metrics (SUM, FROM, TO, and NET indices) can be used for inference. For the VAR estimation, I apply the adaptive joint LASSO algorithm. A detailed description of the algorithm is in Appendix B.1.

The MBB-VAR method efficiency depends on the sample size and bootstrap sample. We can estimate the confidence intervals more accurately if the sample size and the bootstrap sample are higher. Through my empirical analysis, I conducted 1000 bootstrap samples. I use 68% confidence interval (it's a common choice in VAR bootstrap modeling (Sims and Zha, 1999; Wright, 2012; Caggiano et al., 2020)) for the network snapshots (Fig. 12 and Fig. 19-32), and 95%[32] when I analyze the structural changes in the total spillover index (Table 18-19).

The Diebold-Yilmaz matrix is estimated on historical data, often on a window of 100-200 days (Diebold and Yilmaz, 2009, 2012; Baruník et al., 2016). So, the network snapshots represent an overview of the system over a short period. The shorter the window size, the more pronounced the current changes in the network. Nevertheless, small window size also diminishes the degrees of freedom, and the number estimated in the VAR framework is considerable. Moving the historical data window, DY spillover matrices can be estimated for each observed day. Therefore, I can provide a series of snapshots and observe the financial network before and after a severe shock (to analyze the structural change in the network).

To illustrate my bootstrap method's usability in structural shock analysis, I estimate the Diebold-Yilmaz framework on two networks and estimate the significant differences with the Moving Block Bootstrap Method. Let's consider a low-dimension system (J=5) where the data-generating process is the same as the second case in Chapter 1. The coefficient matrix B is sparse, with one dominant time series (Eq. (11)), and the covariance matrix (Σ) is a Toeplitz-type matrix with $\rho = 0.3$ (Eq. (13)). To analyze the structural change in a network, I add a shock to the system. The following coefficient matrix (B) represents the data-generating process after the shock[33]:

$$B = \begin{bmatrix} 0.40 & \mathbf{\color{red}0.12} & 0.08 & 0 & 0 \\ 0 & 0.40 & 0.07 & 0 & 0 \\ 0 & \mathbf{\color{red}0.12} & 0.40 & 0 & 0 \\ 0 & \mathbf{\color{red}0.11} & 0.08 & 0.40 & 0 \\ 0 & \mathbf{\color{red}0.13} & 0.09 & 0 & 0.40 \end{bmatrix} \tag{40}$$

The NET values changed for every time series in the following way: -5.31, 15.35 -0.79, -4.11, - 5.25. These numbers illustrate the change in the network. During the structural change, the shock-transmitting role of the second time series increased. Besides that, the shock-receiving role strengthened for the first, fourth, and fifth time series. To illustrate a structural change more clearly, I show this change in a Fig. 8, which is similar to Fig.

---

[32]Estimating the significant changes is easier for the total spillover index than for net pairwise values.
[33]the covariance matrix is the same after the shock

, but more complex.

Figure 8: Structural change in the Diebold-Yilmaz network
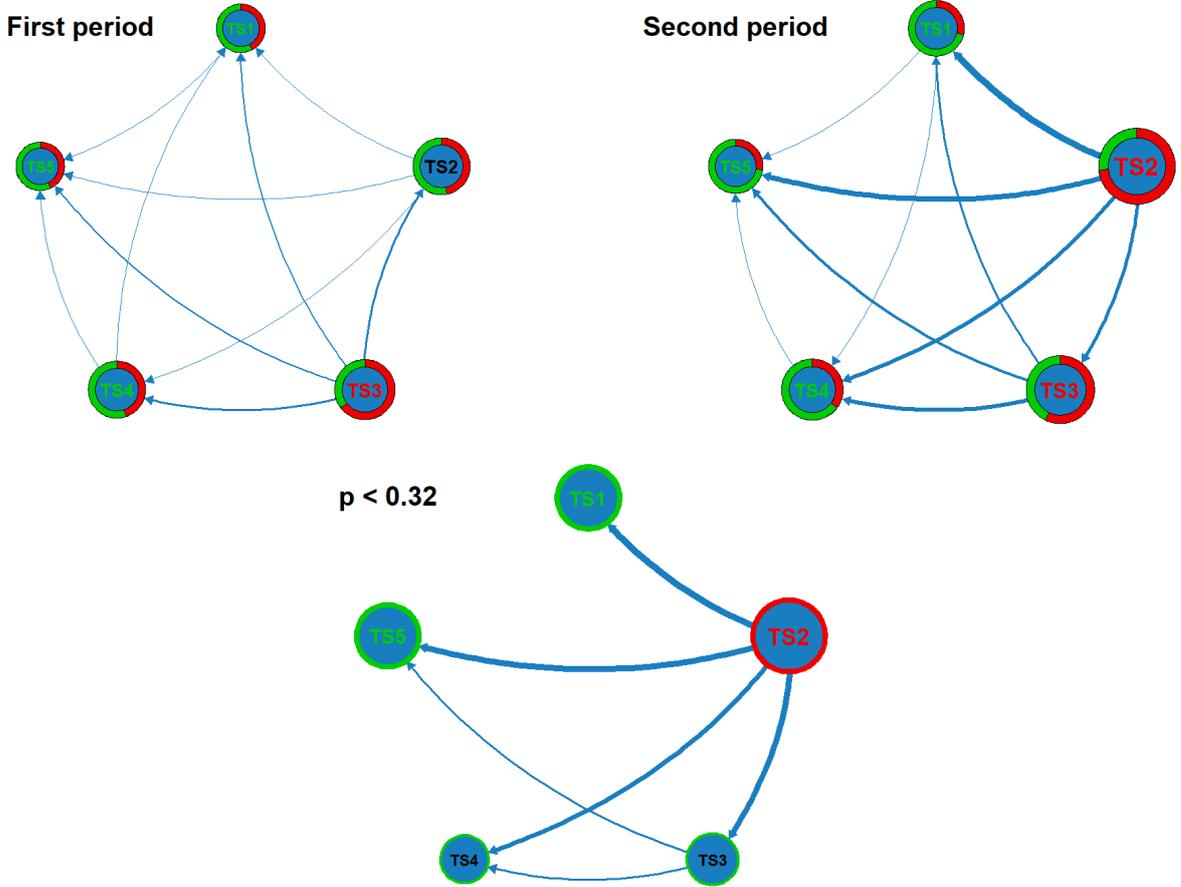
Notes: The network before the shock (after the shock) is in the upper left (right) chart. In the first case, the coefficient matrix (B) is sparse with one dominant time series (TS1). In the second case, it's also sparse with two dominant time series (TS1, TS2). The covariance matrix (Σ) is a Toeplitz-type matrix with $\rho = 0.3$ in both cases. The bottom chart shows their difference at 32% significance levels. Nodes represent the time series. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the time series. Edge directions mark the net spillover indices; their thickness represents the magnitudes. All of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Nodes in the upper charts with green letters represent time series where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent time series where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of the Diebold-Yilmaz network is calculated from a $T = 1000$, adaptive joint LASSO-VAR(2) estimation.

To get a relevant conclusion from these snapshots, it is important to clarify the interpretation of the network graphs. These graphs encompass a wealth of information. At this case, each snapshot figure combines three different charts. The top two are obtained

from the original time series-based VAR estimations, while the bottom one displays the Moving Block Bootstrap-based significant differences between the first and second day at 32% levels[34]. Nodes represent the time series. The pie charts' colors around the nodes represent the proportion of shocks transmitted (TO spillover value marked with red color) and received (FROM spillover value marked with green color) by the time series. The last information related to the nodes is the SUM value, which is indicated by the size of the nodes. On both charts in the top panel, the edges represent the NET spillover indices of the time series, and their thickness represents their magnitudes. As in the 5-node network, there are 10 NET pairwise connections, and I present all of them. Nodes with green letters represent time series where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. The node sizes and the edge thickness are rescaled based on the first and second periods; the biggest node size in the bottom charts represents the maximum SUM value in the two periods, and the smaller one represents the minimum SUM value. The thickest edge in the bottom charts represents the maximum NET value, and the least thicker represents the minimum NET value.

The bottom graphs indicate the significant differences between the SUM, FROM, TO, and NET measures shown by the charts of the two periods at the top. In this graph, only the significantly changed edges are displayed. Their direction and thickness indicate the direction and magnitude of the impact that the structural change generated during the periods. Here, the red (green) color of the node borders represents if there was a significant change based on the MBB method in the transmitted (received) shocks. That means the border color is red (green) if the net spillover index increased (decreased) between the selected days. If the node's size is above (below) the average size, a significant increase (decrease) happens in the SUM spillover value. Nodes with green letters represent time series where the FROM/SUM ratio increased more than 10%, and reds where TO/SUM increased more than 10%. The SUM indicator roughly translates to the centrality of a given time series in the system.

These snapshots aim to identify the key participants of the system (both big shock transmitters and receivers) and analyze how these roles change during specified days. Besides that, I can separate the significant and insignificant changes in the network with the Moving Block Bootstrap method. The SUM indicator measures the centrality of a given time series in the system, regardless of the direction of the shock transmission. Nodes with red (green) letters represent the biggest shock transmitters (receivers). Increased size in the top panel marks the elevated role of the time series in the system, while the smaller size means a less important participant in the network. I interpret the significant changes in the transmitter role with significant red pie slice growth on the top panel (red border on the bottom). Similarly, the receiver role changes significantly with significant green pie slice growth on the top panel (green border on the bottom panel).

Figure 8 illustrates the structural change in the network. The upper charts show the increased total spillover in the system: the edges are thicker, especially that are related to the second time series. Besides that, the node sizes are higher for most of the time series. The node of the second time series shows the changed role of this network's participant. However, the two upper charts only show the increased total spillover and whose role changed in the system. Based on these graphs, we can't decide which change was significant and which was not. The bottom part of the figure solves this problem. This chart illustrates that the shock-transmitting role significantly increased for the second time series, and for the other four participants in the system became more shock-transmitting

---

[34]it is a common choice in VAR bootstrap estimation

due to the structural change

Based on this extension of the original DY framework (MBB-based DY framework), In the next section, I can analyze the differences between the illiquidity and volatility networks and analyze the statistically significant structural changes during the key events for the system.

6.2 Comparing illiquidity and volatility network

Volatility is often used as a measure of risk; for this reason, volatility spillover is widely applied to analyze systemic risk. Volatility is implicitly associated with risk, so measures of spillovers can provide an 'early warning system' for emergent turmoil (Billio et al., 2012). It is also called "fear connectedness" (Diebold and Yılmaz, 2014). However, it does not capture all the relevant information of the system.

According to Amihud (2002), expected market illiquidity positively affects ex-ante stock excess return but is negatively related over time to contemporaneous unexpected illiquidity. Analogous to volatility spillover as "fear connectedness", illiquidity spillover is "fire sales connectedness" expressed by market participants as they trade. In recent years, illiquidity spillovers have also been investigated in several articles, like Smimou and Khallouli (2016) or Schneider et al. (2018); however, DY illiquidity spillovers have only been estimated outside the financial system (Andrikopoulos et al., 2014; Inekwe, 2020).

To address this situation, in this section, I characterize static and dynamic volatility and illiquidity spillover of 16 U.S. FIs via the DY framework. Through these spillover indices besides investors' fear, the spread of illiquidity can be followed too. Besides the dynamic of the total spillover index, I analyze the differences between the distribution of the TO, FROM, NET, and pairwise NET indices of the volatility and illiquidity network. Combining the MBB approach with the DY framework, I examine the differences between the daily snapshots of the two networks.

6.2.1 Data

The data cover the period January 2, 2004, to June 25, 2019, with a total of 3897 daily observations. According to analytical purposes, I use the subsample of Demirer et al. (2018) focus on 16 U.S. FIs. The institutions are listed in Table B1 in the Appendix. To analyze a longer time span in this subsection, I skip the "troubled" FIs (Bear Sterns, Lehman Brothers). I use low and high-price data from Bloomberg in order to measure volatility following Diebold and Yilmaz (2012). Accordingly, illiquidity time series based on Amihud (2002). A description of volatility and illiquidity calculation is provided in Table 16.

Table 16: Summary about the applied realized volatility and illiquidity measures

Variable name	Definition	Source	Formula
Volatility	I calculated the log annualized daily percent standard deviation from daily high/low prices.	Diebold and Yilmaz (2012)	$\sigma_{it} = log(100\sqrt{365 \cdot 0.361[ln(P_{i,t}^{high}) - ln(P_{i,t}^{low})]^2})$
Illiquidity	The absolute (percentage) price change per dollar of daily trading volume.	Amihud (2002)	$ILLIQ_{i,t} = \dfrac{\mid ln(P_{i,t}/P_{i,t-1}) \mid}{VOL_{i,t} \cdot P_{i,t}}$

I use the 200-day rolling window, $H = 10$ forecast horizon, and $p = 2$ lags for the analysis. These are the most commonly used parameters in the empirical literature (Diebold and Yilmaz, 2009, 2012; Baruník et al., 2016).

6.2.2 Empirical results of volatility and illiquidity network comparison

To investigate the dynamics of the total spillover indices, I perform a rolling-sample analysis on the volatility, and illiquidity network. Figure 9 represents the result with a 200-period rolling window, $H = 10$ forecast horizon, and $p = 2$ lags.

Figure 9: Illiquidity and volatility total spillover indices for the network from 2004 to 2019.
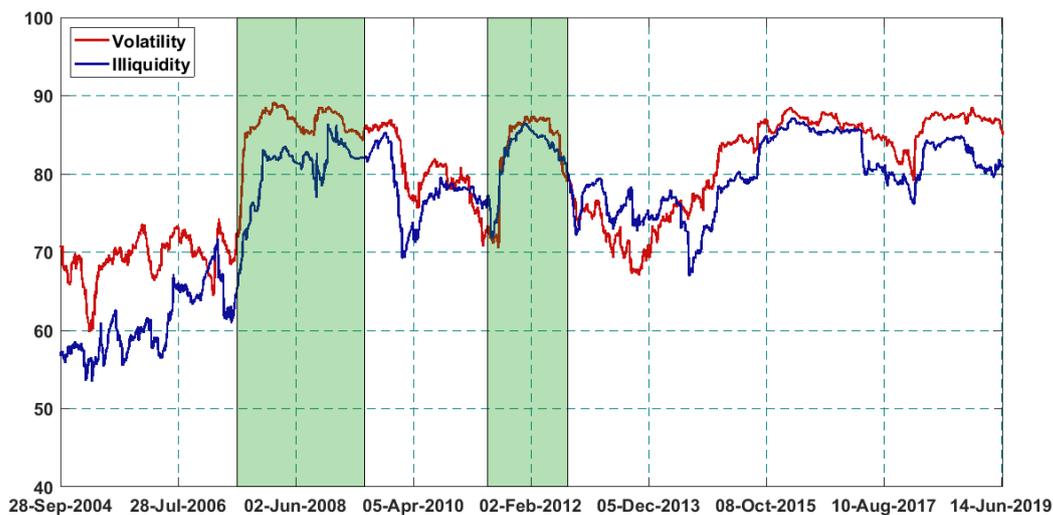
Notes: The information of Diebold-Yilmaz network is calculated from a rolling window analysis with $T = 200$, adaptive joint LASSO-VAR(2) estimation. Dates correspond to the end date of the windows. Vertical (green) shaded highlight crisis periods (GFC and ESDC).

In general, the total volatility spillover index is higher than the illiquidity index (on 69% of observed days), but several highly connected periods can be identified for the processes. The correlations between the indices are between 0.54 and 0.84, so co-movement reflects a strong linear stochastic relationship. Co-movement is spectacular when large shocks hit the economy. First, starting from mid-2007, both of the indices sharply increased by at least 10 percentage points due to the beginning of the Financial Crisis. Then, the European Sovereign Debt Crisis, in August 2011, came up with a large common jump. Consistently, but not identically, each reacts to shocks in the economy. Illiquidity spillover index, similar to volatility was peaking during the GFC, during the ESDC, and in a high range in early 2016. Briefly, I assume that the total illiquidity spillover index contains additional information to volatility spillover indices on the macro level.

Intriguingly, both total spillover indices showed several abrupt change points. These may have occurred due to jumps/noise in the time series, sudden changes in the structure of the financial system (a common shock hits the system), or asymmetry of the pairwise connectedness causing distortion in the total spillover index. As mentioned, the total spillover index represents the average connectedness, so the distribution of the pairwise

Figure 10: Distribution of illiquidity and volatility TO spillover indices for the network from 2004 to 2019.

Notes: The information of Diebold-Yilmaz network is calculated from a rolling window analysis with $T = 200$, adaptive joint LASSO-VAR(2) estimation. Dates correspond to the end date of the windows. Horizontal shaded areas around the mean of the TO spillover indices (red for the volatility index and blue for the illiquidity index) represent the distribution with one standard deviation. Vertical (green) shaded highlight crisis periods (GFC and ESDC).

spillover indices can provide additional information on the network, especially if the distribution is asymmetric. Figure 10 and Figure 11 illustrate the dynamics of the distribution of the TO and FROM spillover indices for both measures. Besides the mean value of the 16 spillover index Figures displays the one standard deviation too.

Temporal changes in the shape or skew of the distribution may contain useful information for a deeper analysis of the financial network. Similar to the results of Diebold and Yılmaz (2014), the distribution of the TO and FROM measures is quite different. The range for the TO spillover index is almost three times wider. During the crisis periods, the FROM index is left-skewed; parallel with this, the other one is right-skewed. Economically that means during turbulent times, only a few firms transmit the shock to the system.

Figure B3 and Figure B4 display the distribution dynamics of the NET and pairwise NET spillover indices for the volatility and illiquidity network. Figures show the distribution of the indices for the volatility network is wider in both cases.

To illustrate the differences between the networks more profoundly, I provide snapshots of volatility and illiquidity networks for the same day. I picked three distinct time periods to show how the indicators' distribution differs from each other during stressed and tranquil times. Figure B6, Figure B5 and Figure 12 represents the volatility and illiquidity networks on the selected days.

In the top panels, larger node sizes indicate institutions that play a more significant role in the bank network, while smaller nodes are less important for the overall system. In the visualization, nodes labeled in red (green) letters represent assets that are major shock transmitters (receivers). To interpret the changes in the receiver role (between the selected days), I mark with an increase in the size of the green pie slice in the top panel and a green

65

Figure 11: Distribution of illiquidity and volatility FROM spillover indices for the network from 2004 to 2019.
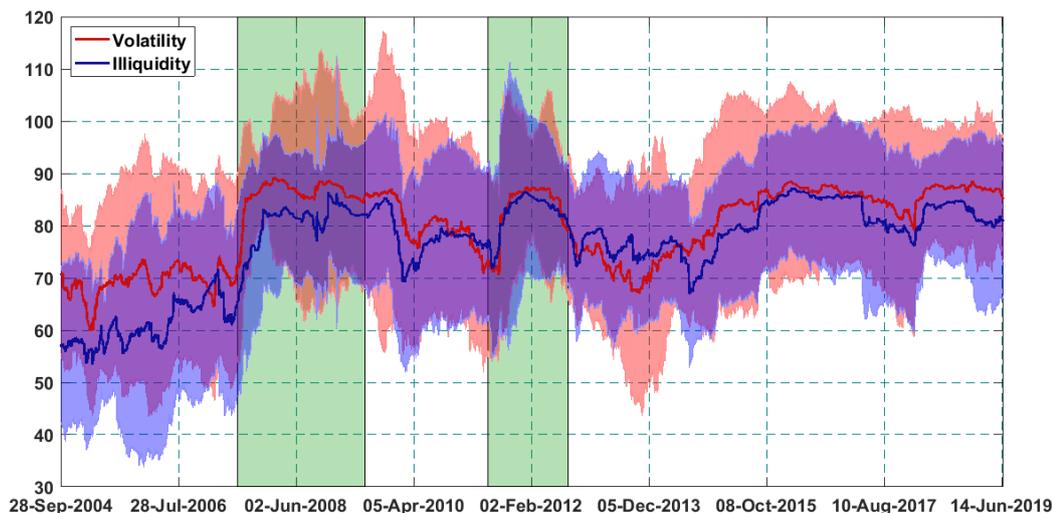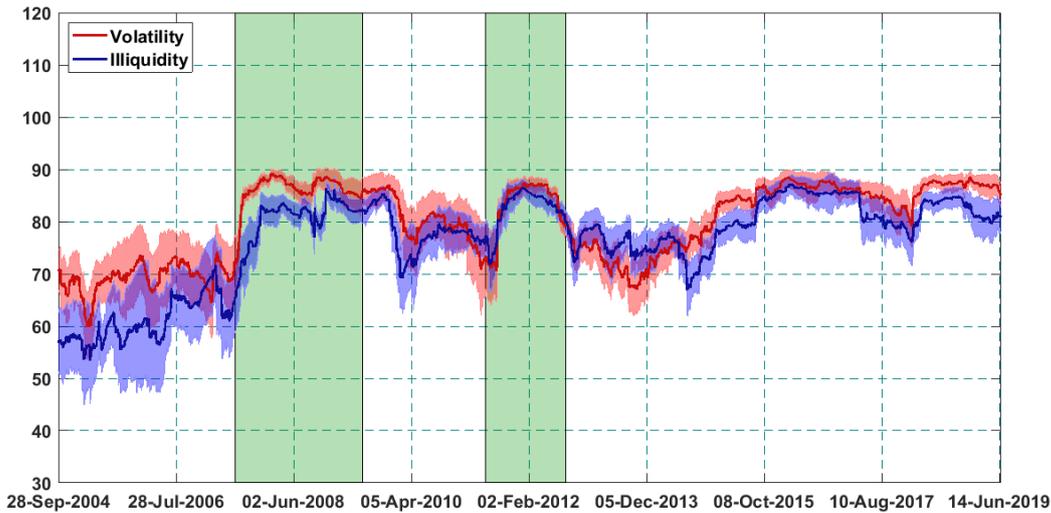

Notes: The information of Diebold-Yilmaz network is calculated from a rolling window analysis with $T = 200$, adaptive joint LASSO-VAR(2) estimation. Dates correspond to the end date of the windows. Horizontal shaded areas around the mean of the FROM spillover indices (red for the volatility index and blue for the illiquidity index) represent the distribution with one standard deviation. Vertical (green) shaded highlight crisis periods (GFC and ESDC).

border in the bottom panel if the net spillover index diminished. This suggests that certain institutions are becoming more influential in receiving shocks. Conversely, to understand the changes in the transmitter role (between the selected days), observe a significant expansion in the red pie slice in the top panel or a red border in the bottom panel. This indicates that particular institutions are gaining importance as shock transmitters in the system, which I measure with increasing net spillover values.

At these periods, the values of total volatility and illiquidity spillovers are almost identical[35], while the shape of the networks differs. Dominant volatility and illiquidity shock transmitters and receivers can be identified; however, they do not match the two networks. For example, in the sample from 2011-12-22, we can see JPM acting as a dominant transmitter in the illiquidity network; however, the role of this FI is on opposite in the volatility network. Some other institutions play different roles in the two systems, too (BAC, C, BBT, and RF). The bottom panel of Figure 12 shows the significant differences between the networks. This chart shows the strength of the moving-block bootstrap method. On the top panels, almost half of the nodes have different characteristics in the volatility and illiquidity network, but for only two (JPM and BBT) is the difference significant. Figure B5 and B6 represent the similar shape of the networks.

The empirical results suggest that illiquidity total spillover indices are also relevant in analyzing systemic risk as they behave differently than total volatility connectedness. On the macro level, both indices react to shocks in the financial system, but even in tranquil periods, the dynamics might vary. Furthermore, if I analyze the pairwise connections between FIs periodically, dominant illiquidity shock transmitters and receivers appear

---

[35]I select the days with the smallest differences between the total spillover values

Figure 12: Estimated volatility and illiquidity network on 2011-12-22

Notes: Volatility (Illiquidity) network are in the upper left (right) chart, and the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "small" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 200$, adaptive joint LASSO-VAR(2) estimation.

in the system. The results of this subsection suggest that both micro and macro-level illiquidity spillover indices display important information for the financial network.

6.3 Analysing the key events of the Global Financial Crisis

In this subsection, I extend the Diebold-Yilmaz (DY) framework with event study analysis tools to shed more light on the contagion channels appearing in the financial network during turmoil times. With the components of my framework, like visualization and the Moving Block Bootstrap approach, we are able to observe the financial connections described by Elliott et al. (2014) and Acemoglu et al. (2015). I provide daily snapshots of the financial network focusing on four key events of the GFC as Diebold and Yılmaz (2014) of 2007-08 with the analytical purpose to illustrate the performance of my combined framework and show that illiquidity spillovers act as contagion channels during turmoil.

Illiquidity contagion has already been analyzed several times in the financial system (Kamara et al., 2008; Karolyi et al., 2012; Cespa and Foucault, 2014). Kamara et al. (2008) demonstrated that the cross-sectional variation of liquidity commonality has increased over the period of 1963–2005, while Karolyi et al. (2012) showed that commonality in liquidity is greater in countries with and during times of high market volatility. Financial markets dried up during the crisis, and liquidity contagion spread across the system. I presume that illiquidity shocks had a greater impact on the observed FIs than volatility and that pairwise illiquidity linkages indicate the route of cascading failures.

Diebold and Yılmaz (2014) examine troubled banks between 2007 and 2008, paying special attention to the influence of Lehman Brothers' bankruptcy on the financial network. They base their measurement on the NET directional spillover and visualization. Their results suggest that Lehman Brothers became the main volatility shock transmitter after the bankruptcy. Nevertheless, the presented NET pairwise directional connections in their graphs are chosen somewhat arbitrarily. Furthermore, they do not display the financial network during the shock of other troubled FIs like Bear Stearns or Wachovia Bank. Finally, based on NET volatility directional spillover indices, the acquisition of Wachovia Bank had an even greater impact on the network than did the Lehman crisis. To overcome these drawbacks, I present an extensive analysis of all events via my new event study framework. For the sake of comparability, I use the stock prices of the 17 U.S. FIs that are also included in the analysis of Diebold and Yılmaz (2014).

6.3.1 Total spillover indices during Global Financial Crisis

The data cover the period from April 17, 2006, to December 22, 2008, with a total of 450 daily observations of the same 17 U.S. FIs that were also analyzed by Diebold and Yılmaz (2014). The institutions are listed in Table 17. Similar to the earlier section, I divide the FIs into three subgroups: big, medium, and troubled institutions. Similar to the earlier subsection, I use realized volatility following Diebold and Yilmaz (2012), and the ILLIQ measure of Amihud (2002). A detailed summary of descriptive statistics of individual time series is provided in Table B2.

Table 17: Financial Institutions included in event study analysis

Panel A: Big Financial Institutions	Ticker	Market Cap.
J.P. Morgan	JPM	156.2
Bank of America	BAC	171.6
Citigroup	C	218.7
Wells Fargo	WFC	57.5
Goldman Sachs	GS	112.0
Panel B: Medium Financial Institutions	**Ticker**	**Market Cap.**
Morgan Stanley	MS	104.5
Bank of New York Mellon	BK	19.8
U.S. Bancorp	USB	23.8
PNC Group	PNC	13.9
American Express	AXP	16.1
Panel C: Troubled Financial Institutions	**Ticker**	**Market Cap.**
Fannie Mae	FNMA	87.9
Freddie Mac	FMCC	79.4
AIG Group	AIG	104.8
Bear Stearns	BS	39.5
Merrill Lynch	MER	102.0
Wachovia Bank	WB	78.3
Lehman Brothers	LEHMQ	69.1

Notes: *Market capitalization is in billion $ at 31/12/2007*

Figure 13: Illiquidity and volatility total spillover indices from January 1, 2008 to December 22, 2008.

Notes: The information of Diebold-Yilmaz network is calculated from a rolling window analysis with $T = 200$, adaptive joint LASSO-VAR(2) estimation. Dates correspond to the end date of the windows.

In my analysis, I dynamically calculate the total DY spillover index to explore the average connectedness in the system during the GFC. Similar to the earlier section, I use a 200-day rolling window, $H = 10$ forecast horizon, and $p = 2$ lags for the analysis[36].

In Figure 13, I show the total spillover index in 2008. Focusing on the crisis, we can see that the total illiquidity spillover index was mainly stagnating apart from the several jumps. The first jump happened in March, and three more jumps occurred in September. After the first jump, the level of the illiquidity spillover index dropped under 80% and stayed there until the next jumps in September. However, the sudden moves in September significantly altered the behavior of the index, and its volatility substantially increased.

To explore the underlying factors of the jumps in the connectedness, I walk through the key banking events of the GFC and observe the behavior of NET spillover indices.

6.3.2 Troubled Financial Institutions

During the GFC, various FIs that got involved in mortgage-backed security trades experienced liquidity troubles (Langley, 2014). In this section, I present four key crisis events that were also considered by Diebold and Yılmaz (2014). I use the NET spillover indices to perform event history analysis of the affected FIs from an individual perspective.

The first major FI that was hit drastically by the crisis was Bear Stearns. The global investment bank narrowly avoided bankruptcy when J.P. Morgan acquired it on March 16, 2008. In Figure 14, I present the NET spillover indices of Bear Stearns during 2008.

Figure 14: Illiquidity and volatility NET spillover indices of Bear Stearns from February 1, 2008 to April 29, 2008.
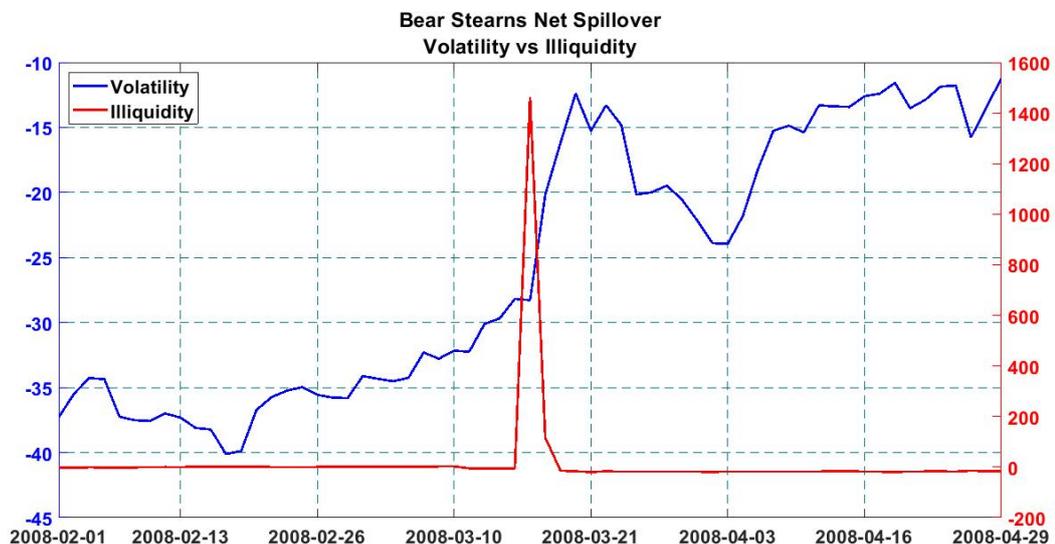
Notes: The information of Diebold-Yilmaz network is calculated from a rolling window analysis with $T = 200$, adaptive joint LASSO-VAR(2) estimation. Dates correspond to the end date of the windows.

On Monday, March 17, 2008, the illiquidity NET spillover index experienced an exceptional jump and achieved a level higher than 1400% points. This extreme value was due to the acquisition that happened the day before. A high positive NET spillover value

---

[36]I use the same parameters as Diebold and Yılmaz (2014)

indicates that Bear Stearns acted as a dominant shock transmitter in the financial system. Nevertheless, I cannot identify which FIs received the transmitted shock in the system.

The mortgage crisis continued in the summer of 2008, and Fannie Mae and Freddie Mac suffered great losses. Fannie Mae, the Federal National Mortgage Association, and Freddie Mac, the Federal Home Loan Mortgage Corporation, bought mortgages on the secondary market, pooled them, and sold them as a mortgage-backed security to investors on the open market. They made sizable profits for more than two decades as housing prices increased. During the housing market's meltdown by mid-2008, Fannie Mae and Freddie Mac began to experience large losses on their retained portfolios.

Figure 15: Illiquidity and volatility NET spillover indices of Fannie Mae from July 25, 2008 to October 22, 2008.
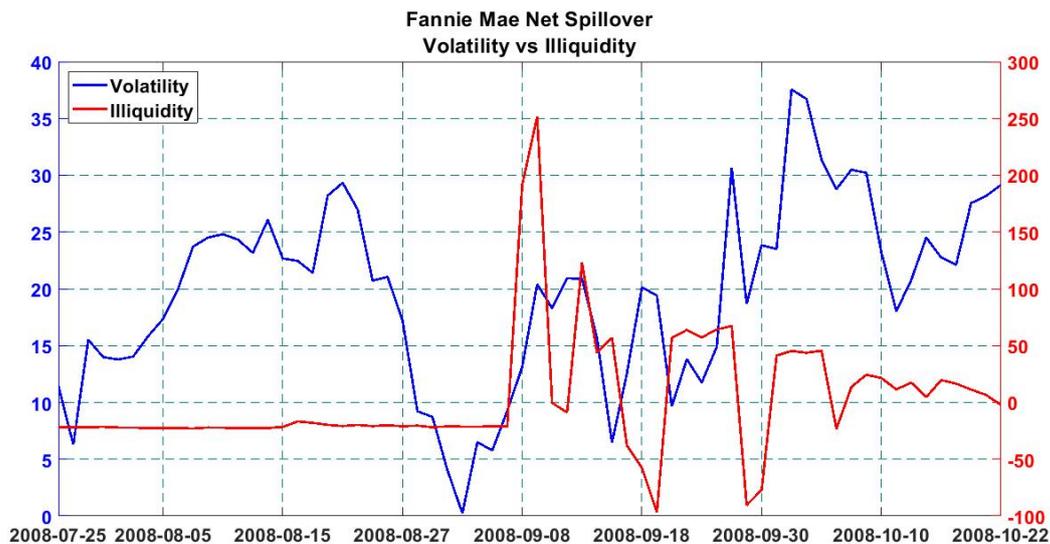
Notes: The information of Diebold-Yilmaz network is calculated from a rolling window analysis with $T = 200$, adaptive joint LASSO-VAR(2) estimation. Dates correspond to the end date of the windows.

With the outbreak of the mortgage crisis, their products became hazardous, spreading contagion in the whole financial system. As the crisis evolved during the summer of 2008, tension around the two FIs increased. Finally, on Saturday, September 6, 2008, governmental intervention was required in the form of $190 billion in the bailout, and on Monday, September 8, the NET illiquidity spillover index jumped above 200% for both Fannie Mae and Freddie Mac as shown in Figure 15 and 16.

Figure 16: Illiquidity and volatility NET spillover indices of Freddie Mac from July 25, 2008 to October 22, 2008.
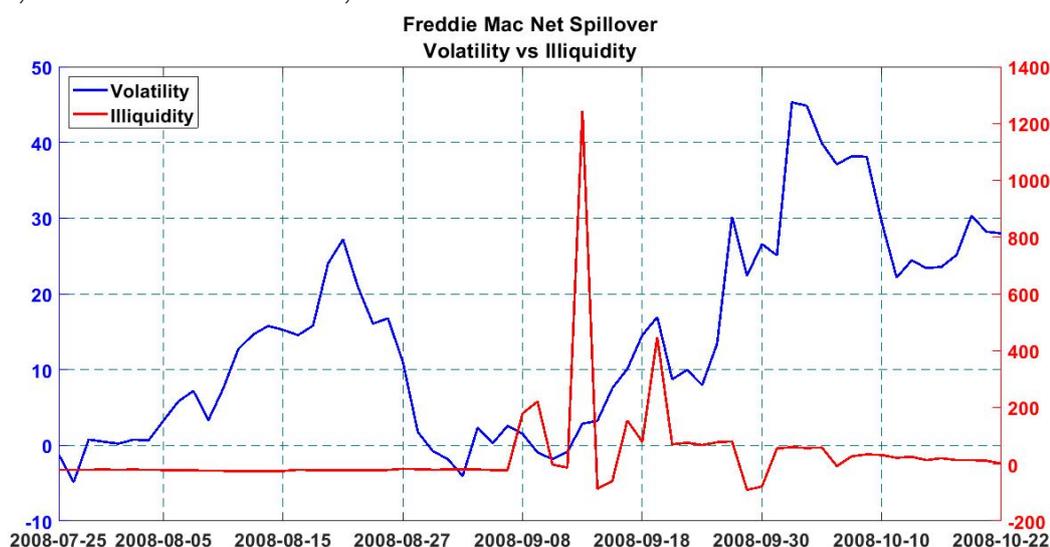
Notes: The information of Diebold-Yilmaz network is calculated from a rolling window analysis with $T = 200$, adaptive joint LASSO-VAR(2) estimation. Dates correspond to the end date of the windows.

Intriguingly, on September 10, the NET illiquidity spillover index dropped to its normal level, then jumped again on September 12, when Freddie Mac's NET spillover index peaked. Illiquidity spillover already signaled the turmoil around the mortgage-backed securities that spilled over various FIs.

Then, on September 15, the infamous bankruptcy of Lehman Brothers held onto large positions in subprime and other lower-rated mortgage tranches when securitizing the underlying mortgages. As described by Elliott et al. (2014), the assets of Fannie Mae and Freddie Mac devalued, and although they managed to avoid bankruptcy through cross-holdings, it imposed losses on their counterparties. In the first two weeks of September, the stock price of Lehman Brothers plunged massively several times. The government summoned several banks to negotiate financing for its reorganization unsuccessfully, and Lehman Brothers filed a Chapter 11 petition[37].

On Figure 17, I can observe that Lehman Brothers' NET illiquidity spillover reached a similar level as Bear Stearns' and Freddie Mac's, but contrary to them, the index remained high for a longer period, five days. Furthermore, on September 16, the Federal Reserve provided a two-year loan to AIG to prevent its bankruptcy and further stress on the global economy. Although the maximum level of NET illiquidity spillover of Lehman Brothers is lower than Bear Stearns', it must be considered that while in March, Bear Stearns was the only troubled institution, in September, several shock transmitters appeared in the system.

Finally, the financial distress around Wachovia Bank started on June 2, 2008, when the chief executive officer was replaced. Still, on September 26 stock price plunged 27%, moreover, several large depositors withdrew money from their accounts in order to drop their balances below the insurance threshold of the Federal Deposit Insurance Corporation. As a result, Wachovia Bank lost a total of $5 billion in deposits that day, and Citigroup had

---

[37]Chapter 11 is a Bankruptcy Code in the US, which is frequently referred to as a "reorganization" bankruptcy

Figure 17: Illiquidity and volatility NET spillover indices of Lehman Brothers from August 15, 2008 to November 12, 2008.
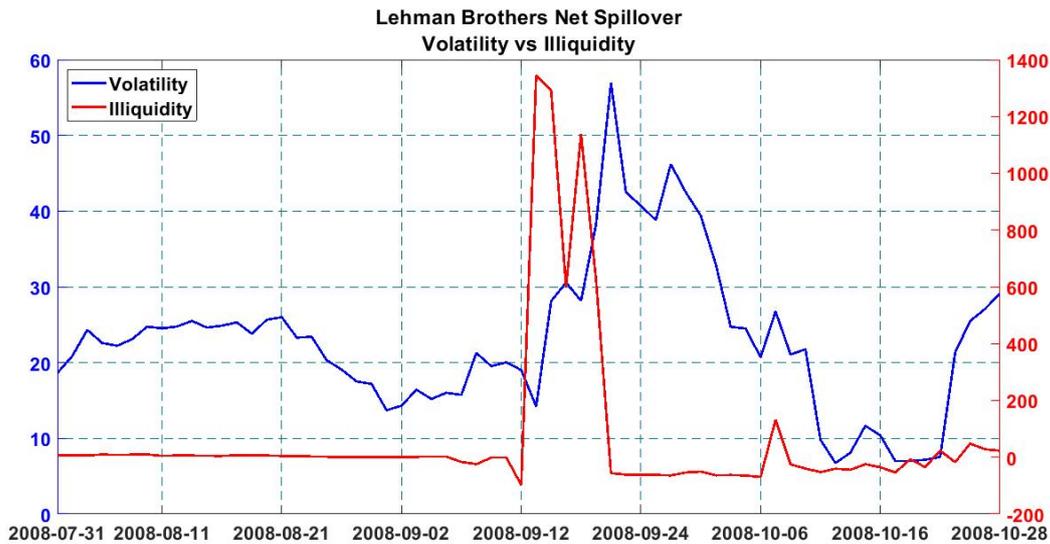
Lehman Brothers Net Spillover
Volatility vs Illiquidity

Notes: The information of Diebold-Yilmaz network is calculated from a rolling window analysis with $T = 200$, adaptive joint LASSO-VAR(2) estimation. Dates correspond to the end date of the windows.

to provide liquidity to continue its operations. On this occasion, the stock price plunge was enough for the NET illiquidity spillover to signal, as presented in Figure 18.

From the results, I assume that the market already priced the merger of Wachovia Bank with Wells Fargo, which only happened a few days later, on October 3.

Figure 18: Illiquidity and volatility NET spillover indices of Wachovia Bank from August 15, 2008 to November 12, 2008.
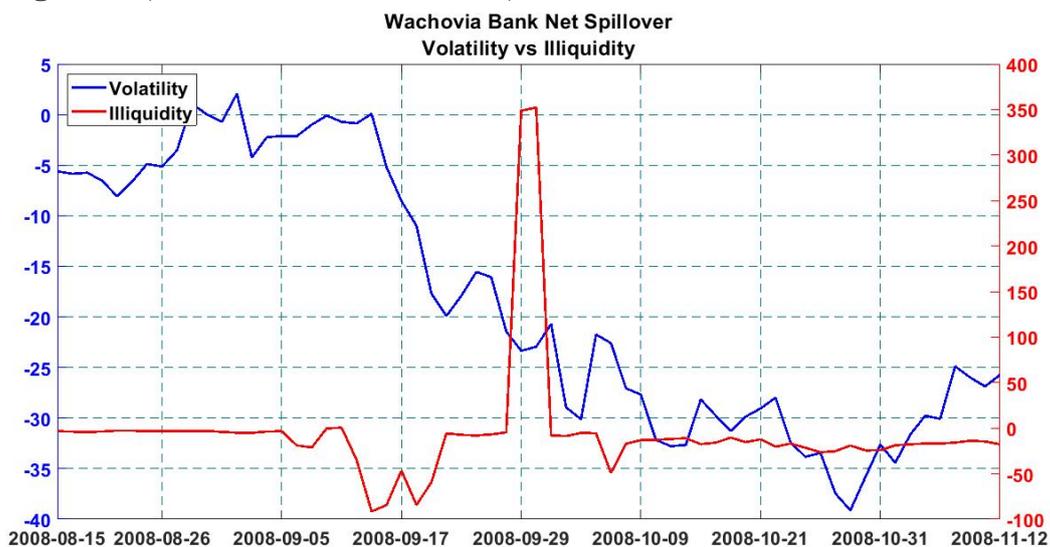
Wachovia Bank Net Spillover
Volatility vs Illiquidity

Notes: The information of Diebold-Yilmaz network is calculated from a rolling window analysis with $T = 200$, adaptive joint LASSO-VAR(2) estimation. Dates correspond to the end date of the windows.

Based on the above-described four events, I can conclude that the NET illiquidity spillover index effectively signals or even forecasts financial turmoil. The behavior of the index clearly differs during calm and distressed periods. Nevertheless, NET spillover indices can only track the association between an individual institution and the overall market. We must observe the system more extensively to explore which other network members were hit by these shocks. Consequently, in the following sections, I present snapshots of the financial system during the days of distress to overcome this drawback of NET spillover indices.

6.3.3 Event analysis from network perspective

I use the NET pairwise spillover indices to investigate the origin of the spikes in the illiquidity spillover index to execute an event study analysis. Similar to the earlier section, I use MBB-based DY networks to provide a sequence of 200-day snapshots of illiquidity connectedness. By examining the financial network daily, we can better understand the ongoing events. I take a close look at the financial network during the days of the financial events examined in the previous section.

Similar to the earlier section, the SUM indicator measures the centrality of a specific FI within a system. In the top panels, larger node sizes indicate institutions that play a more significant role in the bank network, while smaller nodes are less important for the overall system. In the visualization, nodes labeled in red (green) letters represent assets that are major shock transmitters (receivers). To interpret the changes in the receiver role (between the selected days), I mark with an increase in the size of the green pie slice in the top panel and a green border in the bottom panel if the net spillover index diminished. This suggests that certain institutions are becoming more influential in receiving shocks. Conversely, to understand the changes in the transmitter role (between the selected days), observe a significant expansion in the red pie slice in the top panel or a red border in the bottom panel. This indicates that particular institutions are gaining importance as shock transmitters in the system, which I measure with increasing net spillover values.

Bear Stearns' acquisition

The first identified event was the acquisition of Bear Stearns on March 16 (Sunday), 2008. In Figure 19, Figure 20 and Figure 21, I present the illiquidity network snapshots with the significantly changed connections and nodes during the business days between March 14 and 18. Figure 19 plots the DY spillover network on two following days (2008-03-14 and 2008-03-17)[38].

Figure 19: Estimated illiquidity network on 2008-03-14 and 2008-03-17
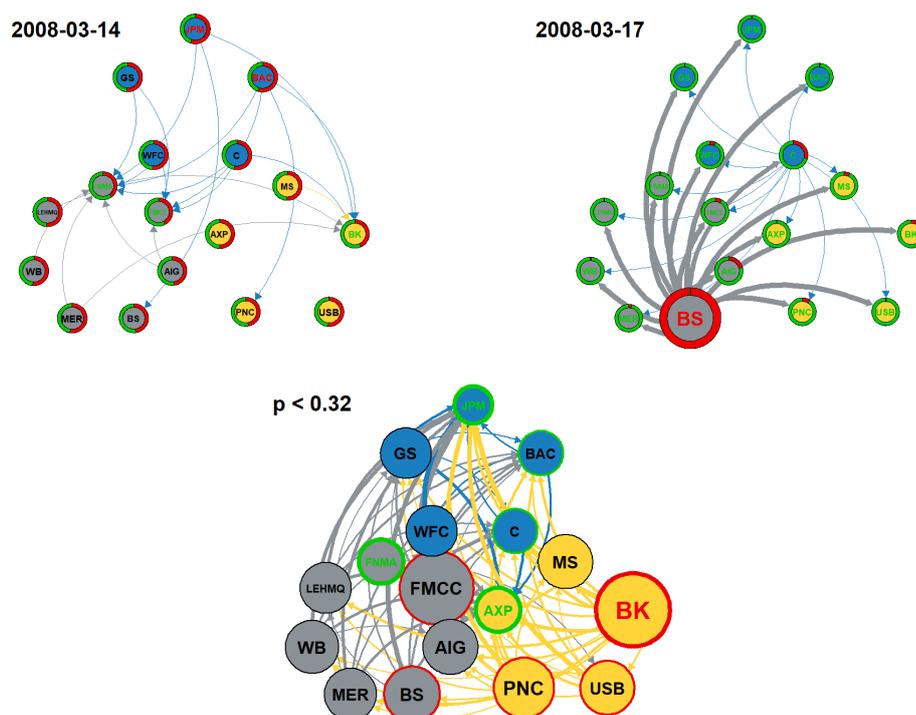
Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FIs. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.

On March 14, Bear Stearns' stock price plunged by almost 50%; however, before the acquisition, there is no sign of the shock in the network; Bear Stearns is even connected to AIG. Then, on March 17, the business day after the acquisition, the illiquidity TO

---

[38]Friday and Monday

connectedness index of Bear Stearns drastically increased. The network also reveals how central Bear Stearns became, as all NET pairwise connectedness indices were strongly positive. It was connected to all other FIs. As shown on the bottom panel of Figure 19, the net illiquidity spillover index changed significantly for other FIs besides Bear Stearns. The net spillover index significantly increased for FMCC, PNC, USB, and BK and decreased for BAC, JPM, C, AXP, and FNMA. This means that the FIs from groups "troubled" and "medium" became shock transmitter participants (relative to the earlier day) of the system.

Figure 20: Estimated illiquidity network on 2008-03-17 and 2008-03-18

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.

On the next day, the connections weakened, but still, Bear Stearn was the key participant in the system. During the day, the role of some participants in the financial network changed significantly. Opposite to Figure 19, the bottom panel of Figure 20 shows how FIs from the group "big" became the shock transmitter participants of the system, rela-

tive to the earlier day. The net spillover index significantly increased for JPM, BAC and C.

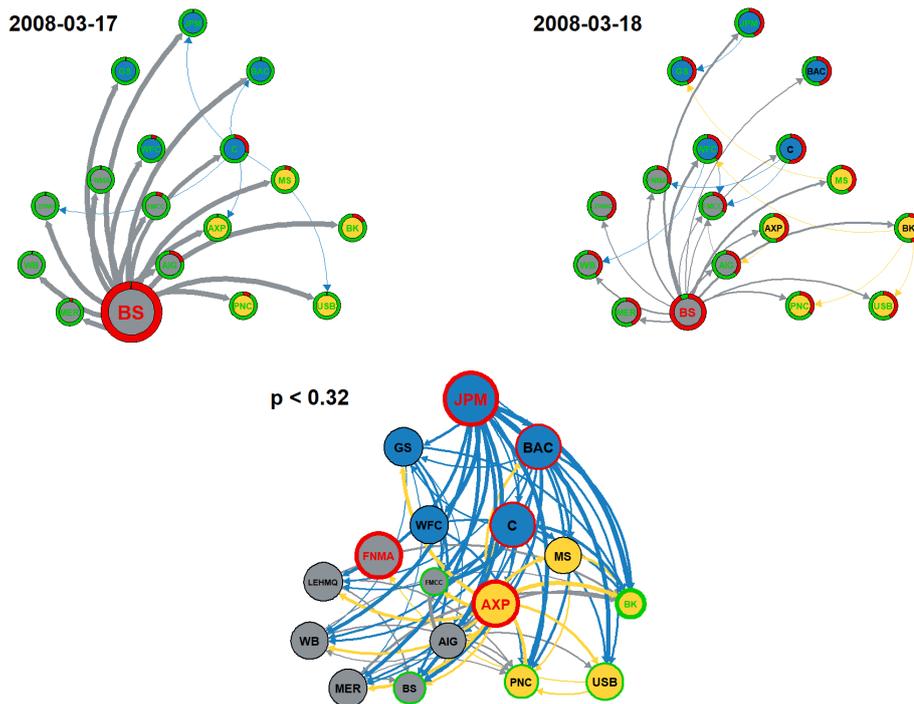**Figure 21: Estimated illiquidity network on 2008-03-18 and 2008-03-19**
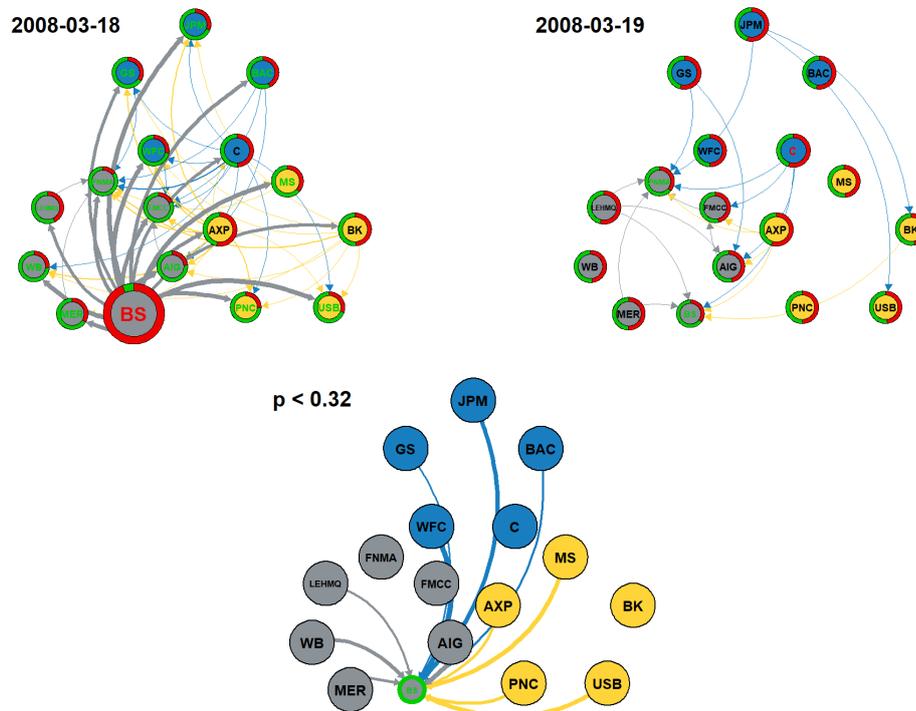

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.

Based on Figure 21 a day later, the connections weakened, and on March 19, the shock disappeared from the system. Bear Stearns acted as a shock receiver – in a relative term – between March 18 and March 19. The bottom panel indicates that all significant changes in the net spillover are related to Bear Stearns.

The example of Bear Stearns showed that the turmoil of an isolated node does not lead to a cascade of failures. Although the shock was severe in the case of Bearn Stearns, it could not spill over to other institutions due to the lack of channels. Because of the shock, the troubled bank temporarily moved into the center of the network, and all its connections strengthened, but the shock disappeared from the system rapidly.

Fannie Mae's and Freddie Mac's bailout

The next event leading to intensified illiquidity connectedness was the bailout of Fannie Mae and Freddie Mac on September 6, 2008. Again, the shock clearly appears in the illiquidity network. On Figure 22 Figure 23 and Figure 24. Based on the Figures, I can see that after the bailout, Fannie Mae and Freddie Mac became large illiquidity shock transmitters; however, this time, both Fannie Mae and Freddie Mac had linkages (instead of the case of Bear Stearns acquisition).

Figure 22: Estimated illiquidity network on 2008-09-05 and 2008-09-08

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

The troubled institutions were linked to AIG, Citigroup, Merrill Lynch, and Lehman Brothers on September 5, the day before the disturbance. The shock lasted for two days, and then the network recovered. As the bottom panels of the figures show, the role of the two institutes changed during these days. Between 2008-09-05 and 2008-09-06[39], FNMA and FMCC became the key participants of the system, as the shock transmission (net illiquidity spillover index) significantly increased. Between these days, JPM, C, and USB were the only institutes whose role changed significantly for both the receiver role strength, which means the net illiquidity index decreased for these institutions.

Figure 23: Estimated illiquidity network on 2008-09-08 and 2008-09-09

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **_Top panels:_** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **_Bottom panel:_** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **_Nodes:_** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **_MBB-VAR:_** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.
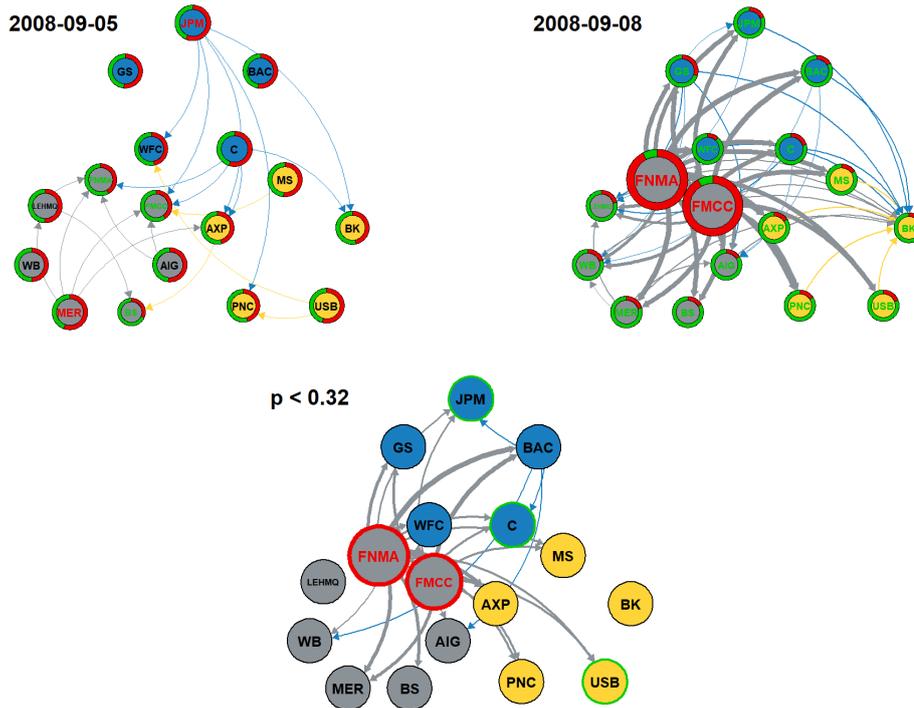
---

[39]Friday and Monday

The shock lasted for two days, and then the network recovered. The top panel of Figure 24 is almost the opposite of Figure 22. Between 2009-09-09 and 2009-09-10, the FNMA and FMCC's role in the system decreased. The bottom panel of the figures strengthens this observation because the change was significant.
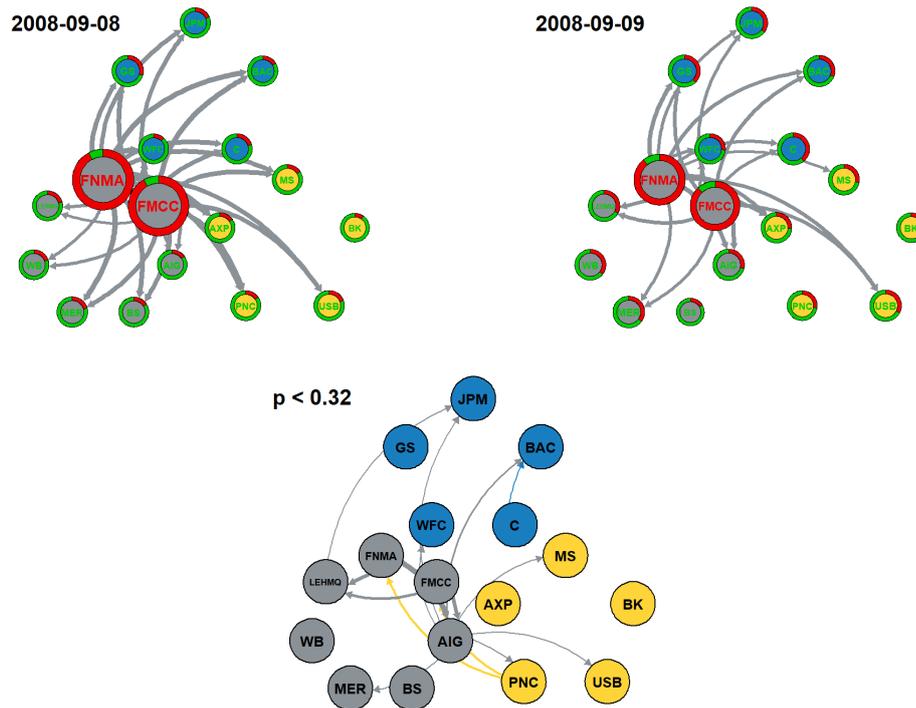
Figure 24: Estimated illiquidity network on 2008-09-09 and 2008-09-10

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

To analyze the robustness of the estimation window, I examined the first network pair (2008-09-05 and 2005-09-08) of the Fannie Mae's and Freddie Mac's bailout case with an estimated VAR model on $T = 100$ and $T = 150$ data point. The results are presented in Figure B7 and B8, and I conclude that the illustrated illiquidity contagion is robust - at least in this case - on the window size.

Lehman Brothers' bankruptcy

In the middle of September, several events happened. Almost all FIs connected to Fannie Mae and Freddie Mac failed. Merrill Lynch was acquired by Bank of America (September 14), Lehman Brothers collapsed (September 15), and AIG was bailed out (September 16). Only Citigroup survived this period, which may be due to its large asset size, which made it more resistant to external shocks (Citigroup had the largest market capitalization at the end of 2007, see in Table 17 in the Appendix). The evolution of illiquidity spillover presented in Figures 25 - 30 clearly reflects the events.

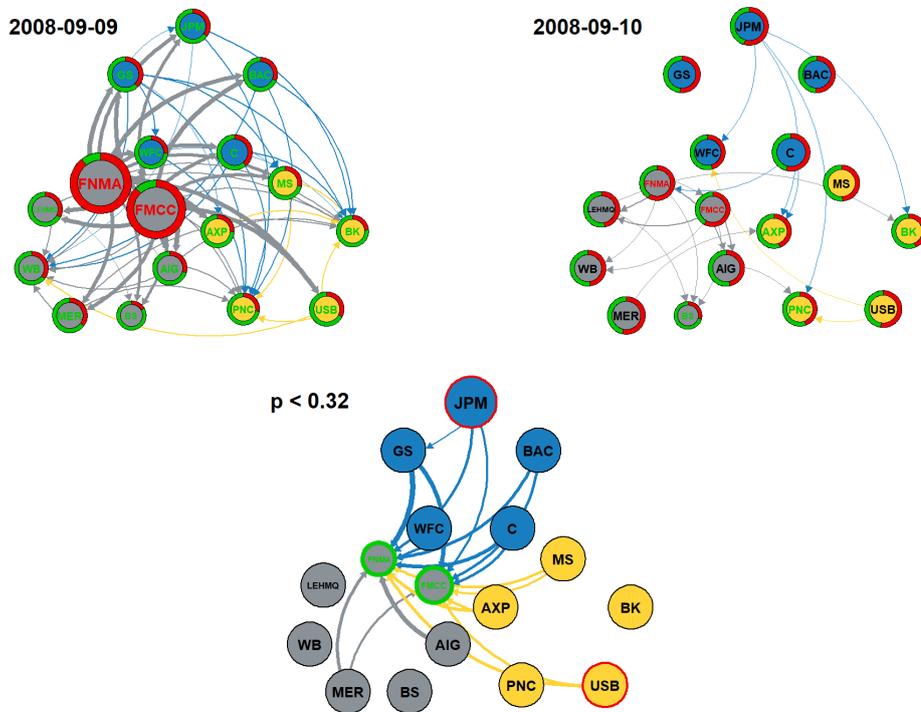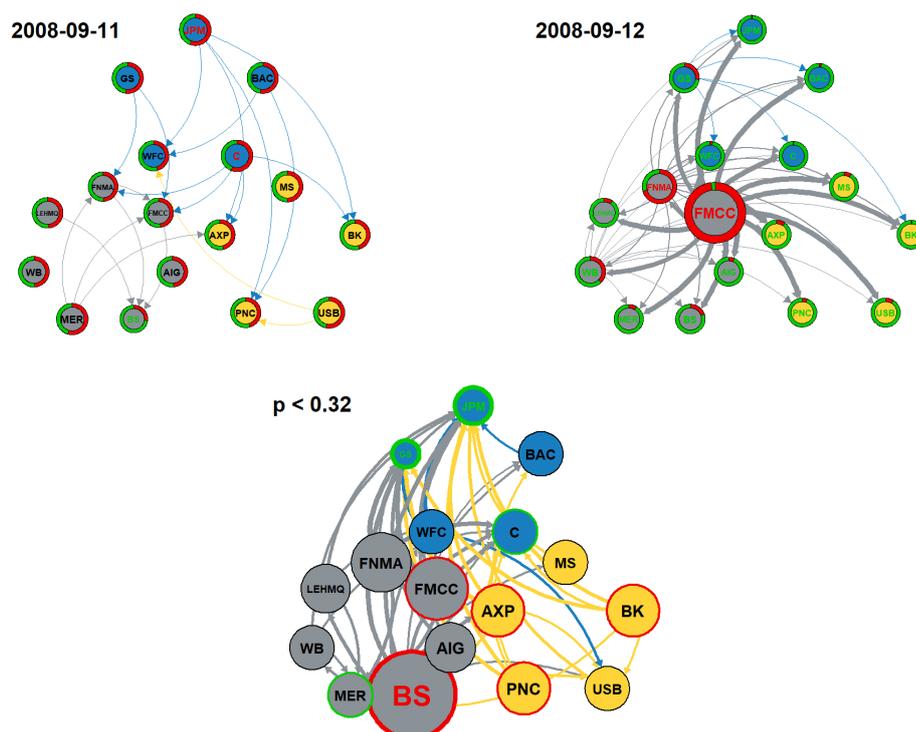**Figure 25: Estimated illiquidity network on 2008-09-11 and 2008-09-12**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

On Friday, September 12, Freddie Mac, the issuer of mortgage-backed securities, stepped up as a dominant shock transmitter as it did a few days earlier, the Fannie Mae. The bottom panel of Figure 25 indicates that the net illiquidity spillover index

changed significantly, not just for FMCC but for some other financial institutes, too. FIs from groups "troubled" and "medium" became shock transmitter participants (relative to the earlier day) of the system, while some institutions from the group "big" transformed in the opposite way. The net spillover index significantly increased for JPM, C, and GS.

Figure 26: Estimated illiquidity network on 2008-09-12 and 2008-09-15

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

Then, on September 12, when the total illiquidity spillover index peaked, the network drastically changed, and Lehman Brothers replaced Freddie Mac. We can see a similar pattern as in the case of Bear Stearns, but strong connections for Lehman Brothers remained even until September 19. Interestingly, this change does not seem significant on the bottom panel of Figure 26, but it can be related to extreme shock (extreme outlier in the time series) on this day what the moving-block bootstrap method can't handle.

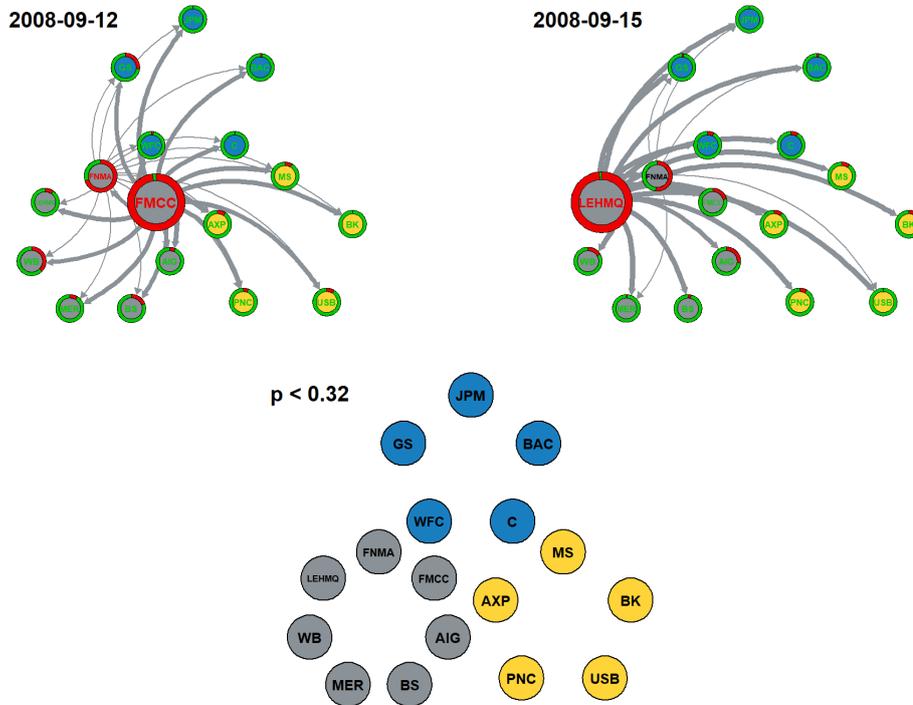**Figure 27: Estimated illiquidity network on 2008-09-15 and 2008-09-16**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

Between the days of 2008-09-15 and 2008-09-17, the network was stable, Lehman Brothers was the key participant - as a shock transmitter - as Figure 27 and Figure 28 show. The bottom panels indicate no significant structural change during these days.
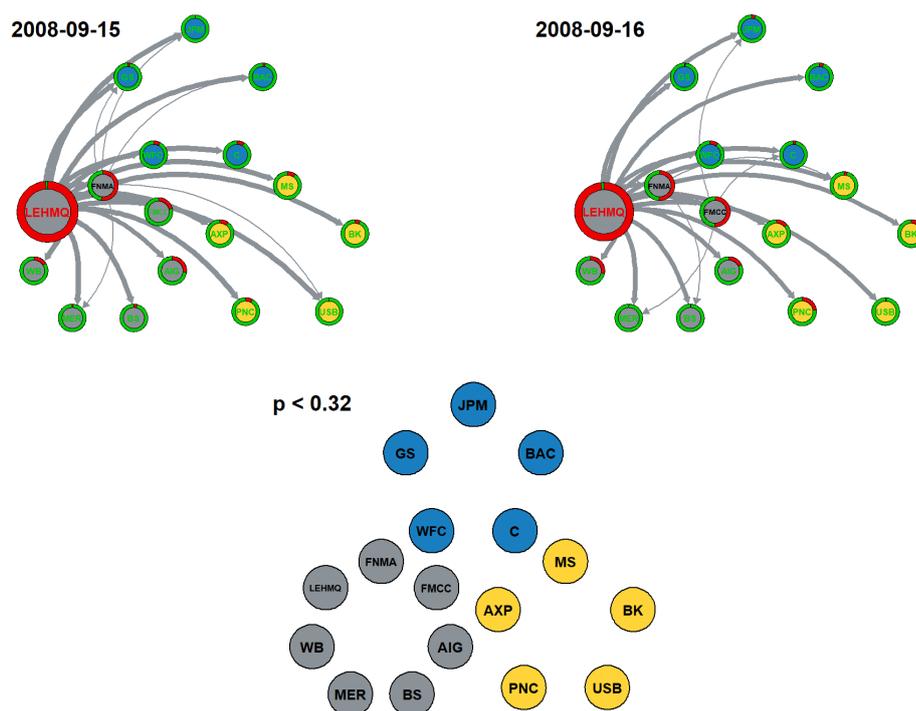
Figure 28: Estimated illiquidity network on 2008-09-16 and 2008-09-17

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

On the next day, the only strong connections that remained besides the Lehman Brothers spillovers were between the most troubled banks and the mortgage-backed security sellers, clearly signaling the spread of the contagion. Furthermore, new connections of Fannie Mae and Freddie Mac appeared with Bank of New York Mellon, J.P. Morgan, and Wachovia Bank. Figure 29 indicates several new significant connections in the system.
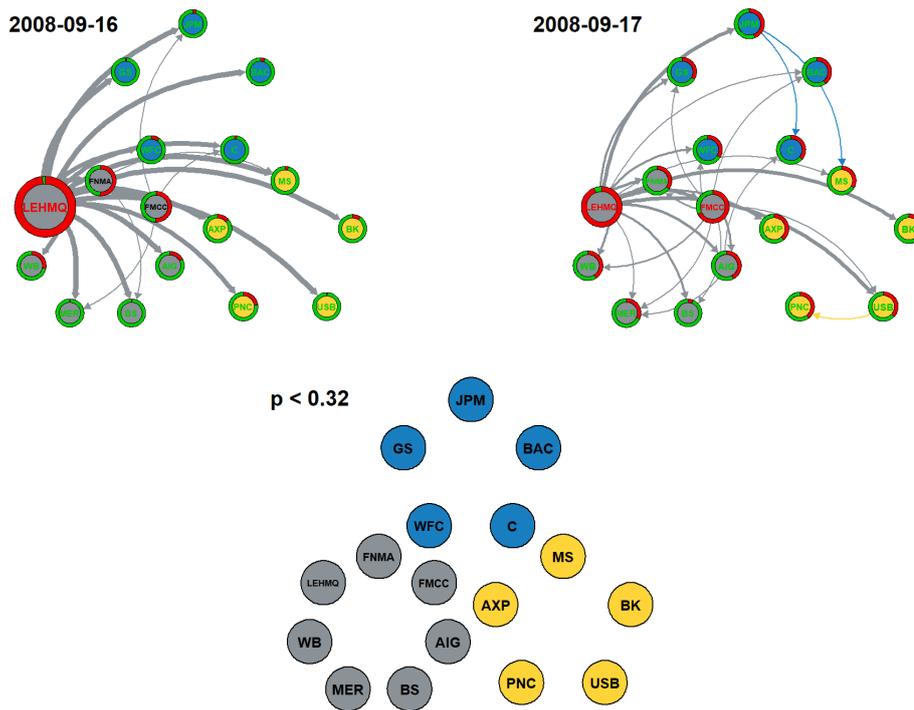
Figure 29: Estimated illiquidity network on 2008-09-18 and 2008-09-19

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 90$, adaptive joint LASSO-VAR(2) estimation.

Between 2009-09-19 (Friday) and 2008-09-22 (Monday), the system changed significantly, as Figure 30 indicates. The role of the Lehman Brothers diminished during the weekend; however, parallel with this, FMCC became a shock transmitter relative to the earlier days. As the MBB-based bottom Panel indicates, these changes are significant.
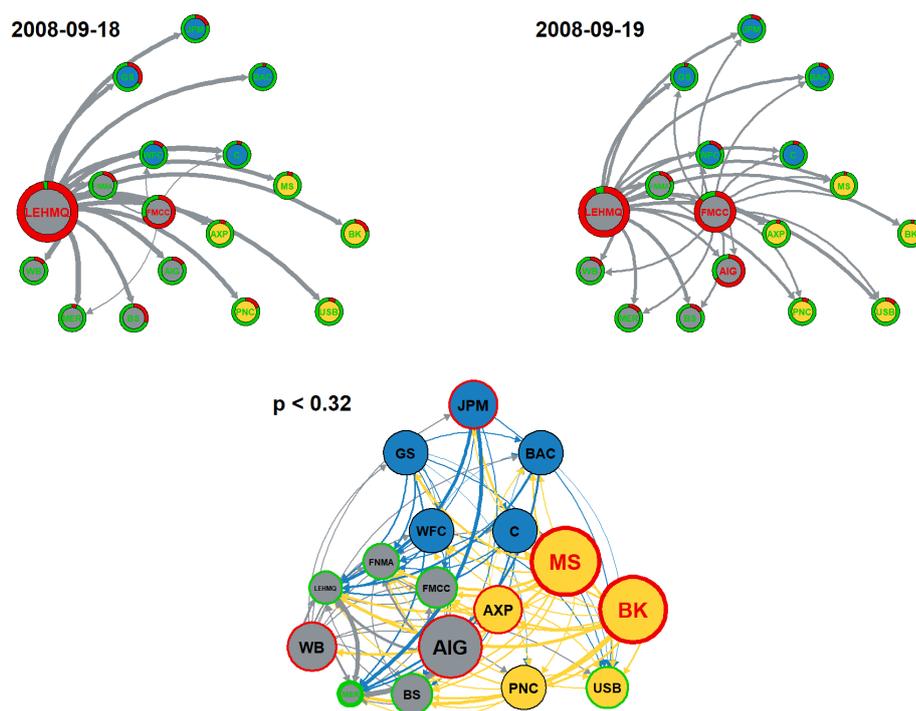
Figure 30: Estimated illiquidity network on 2008-09-19 and 2008-09-22

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red) borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 90$, adaptive joint LASSO-VAR(2) estimation.

Based on Figures B9 - 30 in contrast to the earlier cases during the shock of Lehman Brothers, the effect was more persistent. The failure of Lehman Brothers caused a longer-term transformation of the illiquidity network, highlighting the severeness of its bankruptcy.

Wachovia Bank's acquisition

Finally, the last key event I identified is the merger of Wachovia Bank and Wells Fargo. Although the merger happened on October 3, I present the network from September 26 to October 1 in Figures 31 and 32, since it seems that after the stock price plunge on September 26 the market quickly priced the merger and by the time it actually happened it did not shock the system, as it is also shown in the next section.
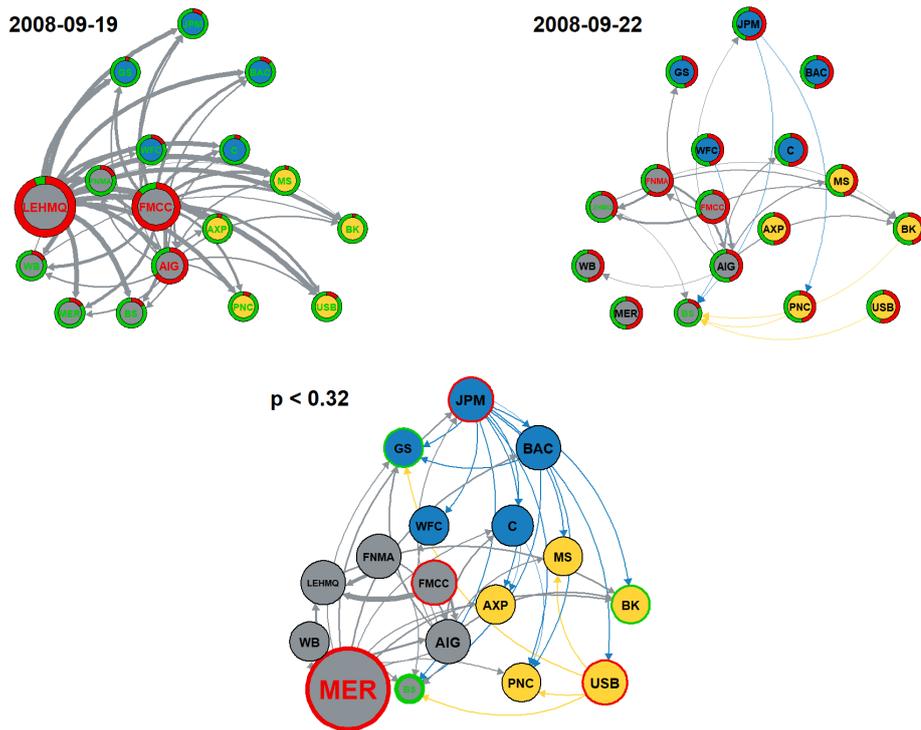
Figure 31: Estimated illiquidity network on 2008-09-26 and 2008-09-29

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.

Stock prices of AIG and Wachovia Bank plunged on September 26 (Friday) and appeared in the illiquidity network on September 29 (Monday). Intriguingly, on this occasion, the stock price plunge was enough for AIG and Bank of New York Mellon to act as dominant variance transmitters, as 31 shows. Before the shock, all three FIs were connected to Fannie Mae and Freddie Mac. The only institution connected to Fannie Mae

and Freddie Mac that avoided the severe shock was JPM, an institution with large assets again. The only significant change is related to the WB, as the bottom panel of 31 shows.

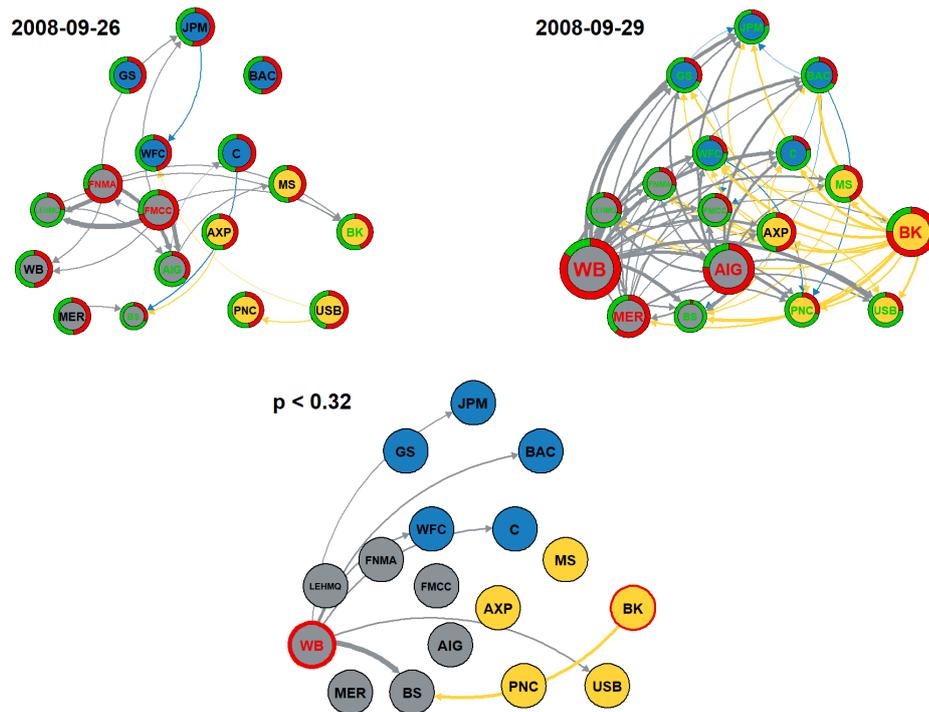**Figure 32: Estimated illiquidity network on 2008-09-30 and 2008-10-01**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.

The next day, on September 30, Wachovia Bank's strong spillovers remained in the system. As described earlier, the merger of Wachovia Bank with Wells Fargo only happened on October 3, and until then, the bank could only continue its operations because Citigroup provided liquidity to it. Since the negotiations about the merger started right after the stock price plunge, it seems that the merger itself did not shock the market; the investors already incorporated it to their expectations.

A day later, the role of the Wachovia Bank diminished, and the system returned to its original state. The top panels of Figure 32 illustrate the shocked and the calm networks, and the bottom panel shows that the only significant change is related to the Wachovia Bank.

Through these network snapshots, I can conclude that NET pairwise illiquidity spillovers act as contagion channels during the crisis described by Elliott et al. (2014) and Acemoglu et al. (2015). During tranquil periods, connectedness remains low in the system, and NET pairwise illiquidity spillovers reflect genuine linkages. Then, if a shock greater than a certain threshold hits a node, the financial network becomes overheated, and strong, short-term connections appear in the system temporarily. However, only the original genuine financial linkages act as contagion channels propagating the distress. If a FI becomes troubled, its value decreases and imposes losses on the other nodes linked to it, and these losses then propagate to others, leading to a cascade of failures (Elliott et al., 2014). As Bearn Stearns was isolated, the losses could not propagate. Nevertheless, the devaluation of Fannie Mae and Freddie Mac's assets initially led to the trouble of AIG, Merrill Lynch, and Lehman Brothers, and later the Bank of New York Mellon and Wachovia Bank. These connections are only visible on a pairwise level thus, visualization indeed revealed additional information to the NET spillover indices. My empirical results are in line with Gai and Kapadia (2010) findings who stands from a theoretical model, that the impact of a shock depends on which node of the network it hits.

6.3.4 Comparison with volatility network

In my analysis, I apply the DY framework to an illiquidity measure, however, this framework has mostly been applied to volatility (Diebold and Yilmaz, 2012; Diebold and Yılmaz, 2014; Demirer et al., 2018). Diebold and Yılmaz (2014) interpret volatility spillovers as "fear spillovers", while illiquidity spillovers can be described as "fire sales spillovers". During the GFC, FIs suffered from liquidity and insolvency difficulties, therefore, I assume that illiquidity spillover is more adequate to assess our research question. Nevertheless, I also estimate the volatility network of the involved FIs to explore whether it reveals any additional information.

Similar to the illiquidity network snapshots in Section 6.3.3, I present the evolution of the volatility network during the observed shocks in Figure B11 - Figure B24. The failure of Lehman Brothers had the greatest impact on the volatility network. Nevertheless, from the snapshots, it is clear that the bankruptcy did not affect the volatility network as severely as the illiquidity network.

In general, volatility spillover networks react much slower. During the Lehman crisis on September 12, U.S. Bancorp was the largest NET shock transmitter. Fannie Mae and Freddie Mac were connected to AIG and Merrill Lynch but not Lehman Brothers. The next day, Lehman Brothers' connections strengthened; nevertheless, U.S. Bancorp still had the highest NET connectedness value. Finally, on September 16, Lehman Brothers stepped up as the main NET shock transmitter. Contrary to the illiquidity network, the shock increased the number of connections of all the FIs. Lehman Brothers could not act as a dominant shock transmitter.

Besides that, the bottom panels of Figure B11 - Figure B24 highlights another important observation. None of these charts show a significant change in the role of the nodes during the four crisis period. It is in stark contrast to the illiquidity network snapshots.

Based on these Figures, I conclude that while illiquidity connections are capable of identifying contagion channels between FIs, volatility linkages do not reflect the spread of contagion accurately. Although a few important connections are outlined in the volatility network, like the link between Lehman Brothers and AIG, the connections between the troubled institutions are considerably less comprehensible.

6.3.5 Structural changes in the networks

In the previous subsection, I analyzed the network changes during financial distress through visualization. We could see that the illiquidity network experienced several sudden structural transformations. Besides that, with the MBB method, I measured on a node level which changes are significant, who are the key participants on the system transformation, and who plays a critical role in a contagion. This approach can help us to identify contagion channels in the financial system. Now, I examine whether these differences are significant on the whole network level. Similar to the node level analysis (micro level analysis), we can apply the MBB method to identify significant changes in the illiquidity network, too. In the analysis, I compare the bootstrapped distributions of total spillover indices on two distinct days. I tested the consecutive days presented on the snapshots. The results of this approach for the illiquidity network are presented in Table 18. Red letters indicate when the total spillover increases significantly between the sample dates, and green letters indicate when it diminishes.

The test results confirm the conclusions drawn in the previous sections. During the days of financial distress, the distribution of the total spillover index significantly changes from one day to the next. As the earlier results suggest, the shock of Bear Stearns' acquisition hit the market on March 17, 2008, significantly altering the illiquidity network. Panel A of Table 18 shows that the total spillover index significantly increased on this day. Then, the next day, as the shock started to disappear, the distribution of the edges returned to its original form, and the total spillover index diminished.

The outcome of the tests is different for the shocks of Fannie Mae, Freddie Mac, and Wachovia Bank. As Panel B and D of Table 18 shows, we cannot observe a significant change in the spillover index in these crisis-related days. I assume these severe shocks only caused idiosyncratic changes in the illiquidity network.

Table 18: Moving Block Bootstrap based Diebold-Yilmaz total spillover index confidence interval for illiquidity network during financial distress

Panel A: Acquisition of Bear Stearns			
Sample Dates		**First date**	**Second date**
2008-03-14	2008-03-17	**82.3 - 85.0**	**86.6 - 87.7**
2008-03-17	2008-03-18	**86.6 - 87.7**	**81.6 - 84.4**

Panel B: Bailout of Fannie Mae and Freddie Mac			
Sample Dates		**First date**	**Second date**
2008-09-05	2008-09-08	79.0 - 82.7	77.6 - 81.2
2008-09-08	2008-09-09	78.2 - 81,9	79.0 - 84.2
2008-09-09	2008-09-10	78.8 - 86.4	77.3 - 83.3

Panel C: Bankruptcy of Lehman Brothers			
Sample Dates		**First date**	**Second date**
2008-09-11	2008-09-12	**78.4 - 81.6**	**81.9 - 89.0**
2008-09-12	2008-09-15	80.5 - 88.5	81.7 - 93.3
2008-09-15	2008-09-16	81.3 - 93.0	82.9 - 92.5
2008-09-16	2008-09-17	**83.4 - 92.5**	**78.0 - 83.0**
2008-09-18	2008-09-19	83.3 - 88.2	82.7 - 88.8
2008-09-19	2008-09-22	**83.1 - 88.2**	**76.8 - 80.2**

Panel D: Merger of Wachovia Bank			
Sample Dates		**First date**	**Second date**
2008-09-26	2008-09-29	77.3 - 80.5	79.3 - 82.4
2008-09-30	2008-10-01	78.4 - 83.6	76.0 - 79.3

Notes: Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of the Diebold-Yilmaz network is calculated from a $T = 200$ adaptive joint LASSO-VAR(2) estimation. The range of the confidence interval is based on 5% significance level. Red letters indicate when the total spillover increases significantly between the sample dates, and green letters indicate when it diminishes.

On the contrary, in the case of Lehman Brothers, the shock was more persistent. From Panel C of Table 18, we can see that the total spillover index significantly increased on September 11, and even on September 16, it was still significantly different from its earlier status. The strong connection between the two federally backed home mortgage companies, Lehman Brothers and AIG, remained in the system for several days and the distribution of the connections did not return to its previous state. The failure of Lehman Brothers caused a longer-term transformation of the illiquidity network, highlighting the severeness of its bankruptcy. Besides that, Table 18 reports that the spillover index decreased in two steps, first on 2008-09-17 and second on 2008-09-22.

Based on the Moving Block Bootstrap approach, I can assume that an illiquidity network has tranquil and distressed statuses. During financial turmoil, the illiquidity connections strengthen, and the total spillover index significantly changes. Still, they revert to their tranquil state at most times. However, if the shock is grave enough, the network structure might transform for longer. This assumption is in line with the results of Acemoglu et al. (2015). If the financial system is hit by a shock with a magnitude higher than a certain threshold, financial interconnections make the system more sensitive and more prone to contagion. With the findings based on these acquisition and bankruptcy cases, I extend the theoretical results Acemoglu et al. (2015) with an empirical one. Besides that my empirical results are in line with Gai and Kapadia (2010) findings, I showed, that the impact of a shock depends on which node of the network it hits.

As shown before, the volatility network is more resistant to shocks, and volatility connections are less revealing regarding genuine connections. In Table 19, I present the MBB method for the same days as for the illiquidity network in Table 18.

Table 19: Moving Block Bootstrap based Diebold-Yilmaz total spillover index confidence interval for volatility network during financial distress

Panel A: Acquisition of Bear Stearns			
Sample Dates		**First date**	**Second date**
2008-03-14	2008-03-17	82.9 - 87.0	83.1 - 86.9
2008-03-17	2008-03-18	82.9 - 87.1	83.3 - 87.5
Panel B: Bailout of Fannie Mae and Freddie Mac			
Sample Dates		**First date**	**Second date**
2008-09-05	2008-09-08	78.5 - 83.1	79.0 - 82.9
2008-09-08	2008-09-09	79.0 - 83.5	78.8 - 82.7
2008-09-09	2008-09-10	79.2 - 83.3	78.9 - 82.9
Panel C: Bankruptcy of Lehman Brothers			
Sample Dates		**First date**	**Second date**
2008-09-11	2008-09-12	79.6 - 83.9	79.6 - 83.6
2008-09-12	2008-09-15	80.3 - 84.0	79.6 - 84.4
2008-09-15	2008-09-16	80.8 - 84.5	80.0 - 83.9
2008-09-16	2008-09-17	80.8 - 84.3	80.3 - 84.6
2008-09-18	2008-09-19	84.9 - 87.8	84.5 - 88.3
2008-09-19	2008-09-22	85.1 - 88.2	83.7 - 87.8
Panel D: Merger of Wachovia Bank			
Sample Dates		**First date**	**Second date**
2008-09-26	2008-09-29	79.5 - 83.9	79.6 - 83.4
2008-09-30	2008-10-01	80.3 - 83.7	79.7 - 83.1

Notes: Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of the Diebold-Yilmaz network is calculated from a $T = 200$ adaptive joint LASSO-VAR(2) estimation. The range of the confidence interval is based on 5% significance level. Red letters indicate when the total spillover increases significantly between the sample dates, and green letters indicate when it diminishes.

The volatility network has not been altered by the severe shocks that hit the financial system during the crisis. All the observed events had sudden and substantial effects on the financial market, which is clearly reflected in the stock prices. However, I cannot identify any of these high-impact events that are reflected in the volatility network. From these results, I can conclude that shocks enter the volatility network considerably slower, and it is not capable of tracking sudden changes in the system.

6.3.6 Robustness analysis

To check the sensitivity of the results to the window size, I calculated the financial network with 150-day and 500-day windows as of September 15, 2008. Figure 33 shows that illiquidity connectedness is not sensitive to window size, however, the volatility network is becoming a bit denser by reducing the window size.

Figure 33: Robustness check of illiquidity and volatility network for September 15, 2008, with 150-day, 200-day, and 500-day windows.
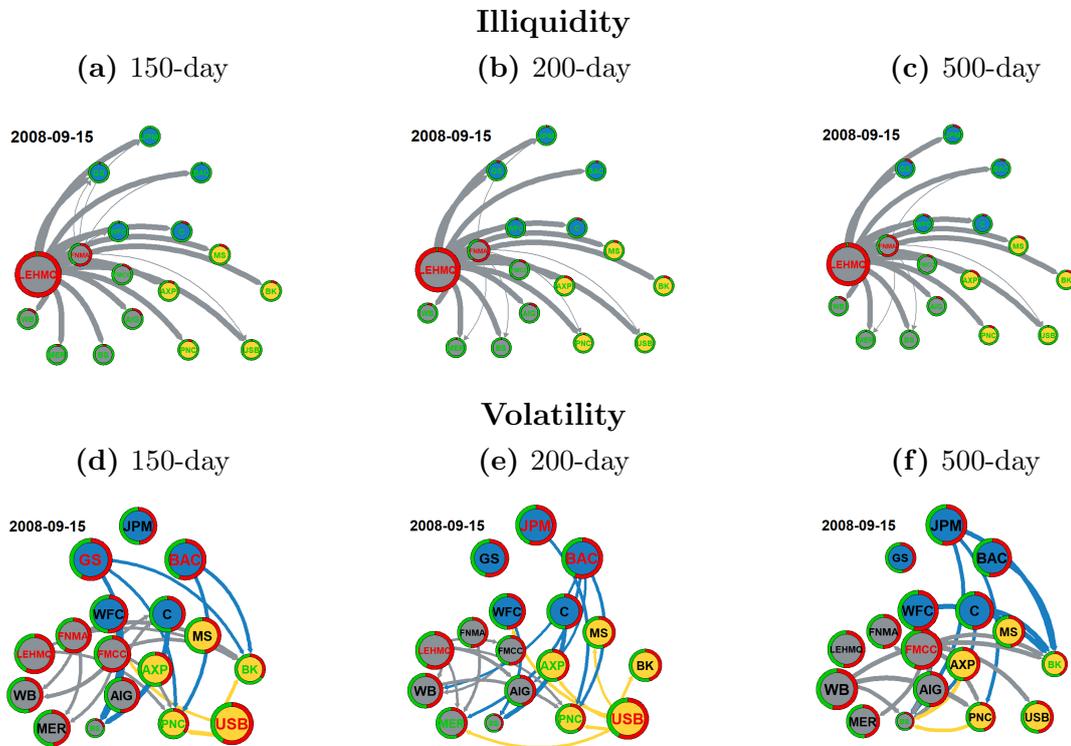
Notes: Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **_Nodes:_** Blue nodes represent FI from the group "big", yellow represents FIs from the group "medium", and grey nodes are related to the "troubled" FIs. Nodes with green letters represents FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$.

With a 150-day window, six institutes have a higher TO/SUM rate than 55%; when the estimation window has 500 days, that number is only one. Furthermore, in the volatility network, even the dominant role of Lehman Brothers is questionable, as with a 500-day window, Freddie Mac is the largest transmitter. While in the illiquidity network, it is unequivocal; thus, it is more robust to the window size than volatility.

6.3.7 Summary of the Global Financial Crisis event study

In this section, I have shown that event study analysis tools like visualization, dynamic measures, and the new formal statistical test help to identify key events, FIs, and contagion channels in the financial system. Furthermore, I have found that on both macro and micro level analysis, the illiquidity networks of FIs during the GFC display relevant information.

I showed that while volatility networks do not always react to financial turmoil, troubled FIs become the main shock transmitters in illiquidity networks. Illiquidity connections act as financial linkages described by Acemoglu et al. (2015), thus, they indicate the spread of contagion in the system during turmoil. Severe financial shocks temporarily alter the illiquidity network, and dominant shock transmitters appear in the system.

I conclude that the DY framework with the MBB approach is a powerful tool for regulators to identify potential failure cascades and SIFIs. Through event study analysis, they can observe shocks that lead to significant changes in the financial network.

7 Chapter 7. Conclusion

In this chapter, I summarize my contribution to the applied finance literature in three parts. In the first two sections, I summarize the methodological contribution by summarizing the statistical benefits of my proposed adaptive joint LASSO method (AJ LASSO) in high-dimension network modeling, and my proposed event study integrated DY framework for focused crisis tracking. Last, I discuss my conceptual contribution, a shift from traditional volatility and return network analysis towards illiquidity network modelling which is consistent with the BASEL IV guidelines emphasizing liquidity management.

7.1 Methodological contribution

My first theoretical contribution to the existing network literature is that I proposed an extension of the row-wise LASSO-VAR method. I referred to this method as the adaptive joint LASSO algorithm. The strength of my adaptive joint LASSO method accounts for possible sparsity both in the coefficient and the covariance matrix of the estimated VAR model.

I compared my new algorithm via simulated data with the most commonly applied regularization methods in the Diebold-Yilmaz literature (row-wise LASSO and adaptive row-wise LASSO). I performed an extensive Monte Carlo simulation study to analyze my proposed estimator's statistical properties. Besides that, I compared the accuracy of my new method to the most commonly used regularization methods' performance. I simulated different large sparse VAR systems and used my proposed algorithm to estimate the parameters of the coefficient and covariance matrix and the DY network.

For the basic setup, sparse coefficient matrix, and band inverse covariance matrix, I concluded that the adaptive joint LASSO estimator outperforms row-wise LASSO estimators significantly in the case of the coefficient, covariance, and the Diebold-Yilmaz spillover matrix estimation for all parameter sets. Simulation results indicated that in highly correlated systems, the relative performance of the AJ LASSO method performs remarkably well, especially for the coefficient matrix estimation.

The accuracy gain is sensitive to the covariance structure. In the case of sparse inverse covariance matrix setting, which is the common assumption in theoretical econometrics, the accuracy gain was 90% for the DY spillover matrix estimation. The result demonstrated that the new method efficiently accounted for the characteristics of the sparse inverse covariance matrix in the estimation via the penalized maximum likelihood objective function.

The accuracy gain of the adaptive joint LASSO method was smaller for the non-sparse coefficient matrix but still significant. The smallest gains of the new estimator relative to the others were obtained for the non-sparse coefficient and inverse covariance matrix. Even when AJ LASSO was only marginally superior, there were still benefits because the number of extreme outliers where the error was more than 20% in the estimation was significantly lower than for the other methods.

Based on these results, I concluded that AJ LASSO can be useful for high-dimension VAR system (and DY network) modeling. In the empirical part of my thesis, I used my proposed method to estimate a VAR model for the DY volatility and illiquidity network.

Combining the DY framework for network modeling with event study methodology may be useful at turbulent times. My second theoretical contribution to the existing network literature is that I extended the original Diebold-Yilmaz framework with an

event study tool to provide more insights into the contagion channels appearing in the networks during structural changes.

I introduced a formal test to compare the distribution of the commonly used spillover indices over time. I integrated the residual-based MBB algorithm into the DY framework's estimation step to determine whether the DY spillover indices' empirical distribution on different periods is equal. The gain of this method is that we can use it not just for analyzing the significant structural changes in the network on a macro level (total spillover index) but on a micro level (pairwise spillover indices) too.

7.2 Conceptual contribution

In the empirical part of my thesis, I presented a characterization of the illiquidity connectedness of a US financial network with my proposed AJ LASSO method in the Diebold-Yilmaz framework. I investigated the significant similarities and differences between the volatility and illiquidity network on the macro and micro levels with the combined MBB-based DY framework.

I showed that illiquidity total spillover indices are also relevant in analyzing systemic risk as they behave differently than total volatility connectedness. On the macro level, both indices react to shocks in the financial system, but even in tranquil periods, the dynamics might vary. My empirical results suggest that the volatility and illiquidity network differs not just on the macro level (total spillover index) but on the micro (pairwise net spillover indices) level too. I concluded that micro and macro-level illiquidity spillover indices display important information for the financial network.

Besides that, I extended the Diebold-Yilmaz (DY) framework with event study analysis tools. With the components of my combined framework, like visualization and the MBB approach, I was able to observe the financial connections during crisis periods. I provided daily snapshots of the financial network, focusing on four key events of the GFC with the analytical purpose of illustrating the performance of my framework and showing that illiquidity spillovers act as contagion channels during turmoil.

In my empirical analysis, I have shown that event study analysis tools like visualization, dynamic measures, and the new formal statistical test help to identify key events, FIs, and contagion channels in the financial system. Furthermore, I have found that on both macro and micro level analysis, the illiquidity networks of FIs during the GFC display relevant information. I showed that while volatility networks do not always react to financial turmoil, troubled FIs become the main shock transmitters in illiquidity networks. Illiquidity connections act as financial linkages described by Acemoglu et al. (2015); thus, they indicate the spread of contagion in the system during turmoil. Severe financial shocks temporarily altered the illiquidity network, and dominant shock transmitters appeared in the system. My empirical results are in line with Gai and Kapadia (2010) findings, too. I found that the impact of a shock depends on which node of the network it hits.

Domestic and international financial networks are ever evolving, and increasingly complex with the emergence of new forms of systemic risks (e.g., climate risk) and re-emergence of to-big-to fail concerns as a result of major bank mergers (UBS-Credit Suisse). While no network analysis can be complete, I believe that my methodological and conceptual contribution to closed network analysis can aid future work to provide valuable insights for regulators and policymakers.

References

Abadie, A. (2002). Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American statistical Association*, 97(457):284–292.

Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608.

Acharya, V., Engle, R., and Richardson, M. (2012). Capital shortfall: A new approach to ranking and regulating systemic risks. *American Economic Review*, 102(3):59–64.

Acharya, V. V., Pedersen, L. H., Philippon, T., and Richardson, M. (2017). Measuring systemic risk. *The review of financial studies*, 30(1):2–47.

Adamek, R., Smeekes, S., and Wilms, I. (2023). Lasso inference for high-dimensional time series. *Journal of Econometrics*, 235(2):1114–1143.

Adnan, A., Hossain, A., Adnan, A., and Hossain, A. (2016). Impact of m&a announcement on acquiring and target firm's stock price: An event analysis approach. *International Journal of Finance and Accounting*, 5(5):228–232.

Alter, A. and Beyer, A. (2014). The dynamics of spillover effects during the european sovereign debt turmoil. *Journal of Banking & Finance*, 42:134–153.

Altman, E. I., Haldeman, R. G., and Narayanan, P. (1977). Zetatm analysis a new model to identify bankruptcy risk of corporations. *Journal of banking & finance*, 1(1):29–54.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets*, 5(1):31–56.

Andrikopoulos, A., Angelidis, T., and Skintzi, V. (2014). Illiquidity, return and risk in g7 stock markets: interdependencies and spillovers. *International Review of Financial Analysis*, 35:118–127.

Antonakakis, N., Chatziantoniou, I., and Gabauer, D. (2020). Refined measures of dynamic connectedness based on time-varying parameter vector autoregressions. *Journal of Risk and Financial Management*, 13(4):84.

Badics, M. C., Huszar, Z. R., and Kotro, B. B. (2023). The impact of crisis periods and monetary decisions of the fed and the ecb on the sovereign yield curve network. *Journal of International Financial Markets, Institutions and Money*, page 101837.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Bai, J., Ng, S., et al. (2008). Large dimensional factor analysis. *Foundations and Trends® in Econometrics*, 3(2):89–163.

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large bayesian vector auto regressions. *Journal of applied Econometrics*, 25(1):71–92.

Barbaglia, L., Croux, C., and Wilms, I. (2020). Volatility spillovers in commodity markets: A large t-vector autoregressive approach. *Energy Economics*, 85:104555.

Barigozzi, M. and Brownlees, C. (2019). Nets: Network estimation for time series. *Journal of Applied Econometrics*, 34(3):347–364.

Barigozzi, M., Cho, H., and Owens, D. (2023). Fnets: Factor-adjusted network estimation and forecasting for high-dimensional time series. *Journal of Business & Economic Statistics*, pages 1–13.

Barigozzi, M., Lippi, M., and Luciani, M. (2021). Large-dimensional dynamic factor models: Estimation of impulse–response functions with i (1) cointegrated factors. *Journal of Econometrics*, 221(2):455–482.

Barnett, I. and Onnela, J.-P. (2016). Change point detection in correlation networks. *Scientific reports*, 6(1):1–11.

Baruník, J., Kočenda, E., and Vácha, L. (2016). Asymmetric connectedness on the us stock market: Bad and good volatility spillovers. *Journal of Financial Markets*, 27:55–78.

Baruník, J. and Křehlík, T. (2018). Measuring the frequency dynamics of financial connectedness and systemic risk. *Journal of Financial Econometrics*, 16(2):271–296.

Basu, S., Das, S., Michailidis, G., and Purnanandam, A. (2019). A system-wide approach to measure connectivity in the financial sector. *Available at SSRN 2816137*.

Basu, S. and Michailidis, G. (2015). Regularized estimation in sparse high-dimensional time series models.

Belloni, A., Chernozhukov, V., and Wang, L. (2011). Square-root lasso: pivotal recovery of sparse signals via conic programming. *Biometrika*, 98(4):791–806.

BenSaïda, A., Litimi, H., and Abdallah, O. (2018). Volatility spillover shifts in global financial markets. *Economic Modelling*, 73:343–353.

Bergmeir, C., Hyndman, R. J., and Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis*, 120:70–83.

Bernal, O., Gnabo, J.-Y., and Guilmin, G. (2014). Assessing the contribution of banks, insurance and other financial services to systemic risk. *Journal of Banking & Finance*, 47:270–287.

Bickel, P. J. and Levina, E. (2008). Covariance regularization by thresholding.

Bickel, P. J., Ritov, Y., and Tsybakov, A. B. (2009). Simultaneous analysis of lasso and dantzig selector.

Billio, M., Casarin, R., Costola, M., and Pasqualini, A. (2016). An entropy-based early warning indicator for systemic risk. *Journal of International Financial Markets, Institutions and Money*, 45:42–59.

Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of financial economics*, 104(3):535–559.

Bisias, D., Flood, M., Lo, A. W., and Valavanis, S. (2012). A survey of systemic risk analytics. *Annu. Rev. Financ. Econ.*, 4(1):255–296.

Bostanci, G. and Yilmaz, K. (2020). How connected is the global sovereign credit risk network? *Journal of Banking & Finance*, 113:105761.

Bratis, T., Laopodis, N. T., and Kouretas, G. P. (2020). Systemic risk and financial stability dynamics during the eurozone debt crisis. *Journal of Financial Stability*, 47:100723.

Brüggemann, R., Jentsch, C., and Trenkler, C. (2016). Inference in vars with conditional heteroskedasticity of unknown form. *Journal of econometrics*, 191(1):69–85.

Bubák, V., Kočenda, E., and Žikeš, F. (2011). Volatility transmission in emerging european foreign exchange markets. *Journal of Banking & Finance*, 35(11):2829–2841.

Bühlmann, P. and Van De Geer, S. (2011). *Statistics for high-dimensional data: methods, theory and applications.* Springer Science & Business Media.

Buse, R. and Schienle, M. (2019). Measuring connectedness of euro area sovereign risk. *International Journal of Forecasting*, 35(1):25–44.

Buse, R., Schienle, M., and Urban, J. (2022). Assessing the impact of policy and regulation interventions in european sovereign credit risk networks: What worked best? *Journal of International Economics*, 139:103673.

Cabrieto, J., Tuerlinckx, F., Kuppens, P., Hunyadi, B., and Ceulemans, E. (2018). Testing for the presence of correlation changes in a multivariate time series: A permutation based approach. *Scientific reports*, 8(1):1–20.

Caggiano, G., Castelnuovo, E., and Figueres, J. M. (2020). Economic policy uncertainty spillovers in booms and busts. *Oxford Bulletin of Economics and Statistics*, 82(1):125–155.

Camehl, A. (2023). Penalized estimation of panel vector autoregressive models: A panel lasso approach. *International Journal of Forecasting*, 39(3):1185–1204.

Candes, E. and Tao, T. (2007). The dantzig selector: Statistical estimation when p is much larger than n.

Carriero, A., Kapetanios, G., and Marcellino, M. (2011). Forecasting large datasets with bayesian reduced rank multivariate models. *Journal of Applied Econometrics*, 26(5):735–761.

Castro, C. and Ferrari, S. (2014). Measuring and testing for the systemically important financial institutions. *Journal of Empirical Finance*, 25:1–14.

Cespa, G. and Foucault, T. (2014). Illiquidity contagion and liquidity crashes. *The Review of Financial Studies*, 27(6):1615–1660.

Chan, K.-S. and Chen, K. (2011). Subset arma selection via the adaptive lasso. *Statistics and its Interface*, 4(2):197–205.

Cheng, T., Liu, J., Yao, W., and Zhao, A. B. (2022). The impact of covid-19 pandemic on the volatility connectedness network of global stock market. *Pacific-Basin Finance Journal*, 71:101678.

Chowdhury, B., Dungey, M., Kangogo, M., Sayeed, M. A., and Volkov, V. (2019). The changing network of financial market linkages: The asian experience. *International Review of Financial Analysis*, 64:71–92.

Claeys, P. and Vašíček, B. (2014). Measuring bilateral spillover and testing contagion on sovereign bond markets in europe. *Journal of Banking & Finance*, 46:151–165.

Coelho, R., Gilmore, C. G., Lucey, B., Richmond, P., and Hutzler, S. (2007). The evolution of interdependence in world equity markets—evidence from minimum spanning trees. *Physica A: Statistical Mechanics and its Applications*, 376:455–466.

Corrado, C. J. (2011). Event studies: A methodology review. *Accounting & Finance*, 51(1):207–234.

Davis, R. A., Zang, P., and Zheng, T. (2016). Sparse vector autoregressive modeling. *Journal of Computational and Graphical Statistics*, 25(4):1077–1096.

De Gooijer, J. G. and Hyndman, R. J. (2006). 25 years of time series forecasting. *International journal of forecasting*, 22(3):443–473.

de Mendonça, H. F. and da Silva, R. B. (2018). Effect of banking and macroeconomic variables on systemic risk: An application of δcovar for an emerging economy. *The North American Journal of Economics and Finance*, 43:141–157.

Demirer, M., Diebold, F. X., Liu, L., and Yilmaz, K. (2018). Estimating global bank network connectedness. *Journal of Applied Econometrics*, 33(1):1–15.

Deshpande, S. K., Ročková, V., and George, E. I. (2019). Simultaneous variable and covariance selection with the multivariate spike-and-slab lasso. *Journal of Computational and Graphical Statistics*, 28(4):921–931.

Diebold, F. X. and Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal*, 119(534):158–171.

Diebold, F. X. and Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1):57–66.

Diebold, F. X. and Yılmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134.

Diebold, F. X. and Yilmaz, K. (2015). Trans-atlantic equity volatility connectedness: Us and european financial institutions, 2004–2014. *Journal of Financial Econometrics*, 14(1):81–127.

Diebold, F. X. and Yilmaz, K. (2023). On the past, present, and future of the diebold–yilmaz approach to dynamic network connectedness. *Journal of Econometrics*, 234:115–120.

Dilshad, M. N. (2013). Profitability analysis of mergers and acquisitions: an event study approach. *Business and economic research*, 3(1).

Drakos, A. A. and Kouretas, G. P. (2015). Bank ownership, financial segments and the measurement of systemic risk: An application of covar. *International Review of Economics & Finance*, 40:127–140.

Dungey, M., Harvey, J., and Volkov, V. (2019). The changing international network of sovereign debt and financial institutions. *Journal of International Financial Markets, Institutions and Money*, 60:149–168.

Duso, T., Gugler, K., and Yurtoglu, B. (2010). Is the event study methodology useful for merger analysis? a comparison of stock market and accounting data. *International Review of Law and Economics*, 30(2):186–192.

Eichler, M. (2007). Granger causality and path diagrams for multivariate time series. *Journal of Econometrics*, 137(2):334–353.

Eichler, M. (2012). Graphical modelling of multivariate time series. *Probability Theory and Related Fields*, 153:233–268.

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

Engle, R. F. and Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276.

Fan, J., Feng, Y., and Song, R. (2011a). Nonparametric independence screening in sparse ultra-high-dimensional additive models. *Journal of the American Statistical Association*, 106(494):544–557.

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, 96(456):1348–1360.

Fan, J., Lv, J., and Qi, L. (2011b). Sparse high-dimensional models in economics. *Annu. Rev. Econ.*, 3(1):291–317.

Fan, J., Ma, Y., and Dai, W. (2014). Nonparametric independence screening in sparse ultra-high-dimensional varying coefficient models. *Journal of the American Statistical Association*, 109(507):1270–1284.

Fang, Y. and Shao, Z. (2022). The russia-ukraine conflict and volatility risk of commodity markets. *Finance Research Letters*, 50:103264.

Fengler, M. R. and Gisler, K. I. (2015). A variance spillover analysis without covariances: What do we miss? *Journal of International Money and Finance*, 51:174–195.

Fenn, D. J., Porter, M. A., Williams, S., McDonald, M., Johnson, N. F., and Jones, N. S. (2011). Temporal evolution of financial-market correlations. *Physical review E*, 84(2):026109.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *Review of Economics and statistics*, 82(4):540–554.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2005). The generalized dynamic factor model: one-sided estimation and forecasting. *Journal of the American statistical association*, 100(471):830–840.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1):1.

Friedman, J. H. (2012). Fast sparse regression and classification. *International Journal of Forecasting*, 28(3):722–738.

Friston, K. (2009). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS biology*, 7(2):e1000033.

Gabauer, D., Gupta, R., Marfatia, H., and Miller, S. M. (2020). Estimating us housing price network connectedness: Evidence from dynamic elastic net, lasso, and ridge vector autoregressive models. *Lasso, and Ridge Vector Autoregressive Models (July 26, 2020)*.

Gai, P. and Kapadia, S. (2010). Contagion in financial networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466(2120):2401–2423.

Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Giannone, D., Lenza, M., and Primiceri, G. E. (2021). Economic predictions with big data: The illusion of sparsity. *Econometrica*, 89(5):2409–2437.

Grant, E. and Yung, J. (2021). The double-edged sword of global integration: Robustness, fragility, and contagion in the international firm network. *Journal of Applied Econometrics*, 36(6):760–783.

Greenwood-Nimmo, M., Huang, J., and Nguyen, V. H. (2019). Financial sector bailouts, sovereign bailouts, and the transfer of credit risk. *Journal of Financial Markets*, 42:121–142.

Greenwood-Nimmo, M., Nguyen, V. H., and Shin, Y. (2017). What's mine is yours: Sovereign risk transmission during the european debt crisis.

Greenwood-Nimmo, M. and Tarassow, A. (2022). Bootstrap-based probabilistic analysis of spillover scenarios in economic and financial networks. *Journal of Financial Markets*, 59:100661.

Grillini, S., Ozkan, A., and Sharma, A. (2022). Static and dynamic liquidity spillovers in the eurozone: The role of financial contagion and the covid-19 pandemic. *International Review of Financial Analysis*, 83:102273.

Han, F., Lu, H., and Liu, H. (2015). A direct estimation of high dimensional stationary vector autoregressions. *Journal of Machine Learning Research*.

Hautsch, N., Schaumburg, J., and Schienle, M. (2015). Financial network systemic risk contributions. *Review of Finance*, 19(2):685–738.

Hecq, A., Margaritella, L., and Smeekes, S. (2023). Granger causality testing in high-dimensional vars: a post-double-selection procedure. *Journal of Financial Econometrics*, 21(3):915–958.

Hill, N. T., Perry, S. E., and Andes, S. (1996). Evaluating firms in financial distress: An event history analysis. *Journal of Applied Business Research (JABR)*, 12(3):60–71.

Hsu, N.-J., Hung, H.-L., and Chang, Y.-M. (2008). Subset selection for vector autoregressive processes using lasso. *Computational Statistics & Data Analysis*, 52(7):3645–3657.

Huang, J., Ma, S., and Zhang, C.-H. (2008). Adaptive lasso for sparse high-dimensional regression models. *Statistica Sinica*, pages 1603–1618.

Hué, S., Lucotte, Y., and Tokpavi, S. (2019). Measuring network systemic risk contributions: A leave-one-out approach. *Journal of Economic Dynamics and Control*, 100:86–114.

Hurlin, C., Laurent, S., Quaedvlieg, R., and Smeekes, S. (2017). Risk measure inference. *Journal of Business & Economic Statistics*, 35(4):499–512.

Inekwe, J. N. (2020). Liquidity connectedness and output synchronisation. *Journal of International Financial Markets, Institutions and Money*, 66:101208.

Jaccard, P. (1912). The distribution of the flora in the alpine zone. 1. *New phytologist*, 11(2):37–50.

Jena, S. K., Tiwari, A. K., Abakah, E. J. A., and Hammoudeh, S. (2022). The connectedness in the world petroleum futures markets using a quantile var approach. *Journal of Commodity Markets*, 27:100222.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, 12(2-3):231–254.

Junior, L. S., Mullokandov, A., and Kenett, D. Y. (2015). Dependency relations among international stock market indices. *Journal of Risk and Financial Management*, 8(2):227–265.

Just, M. and Echaust, K. (2022). Dynamic spillover transmission in agricultural commodity markets: What has changed after the covid-19 threat? *Economics Letters*, 217:110671.

Kamara, A., Lou, X., and Sadka, R. (2008). The divergence of liquidity commonality in the cross-section of stocks. *Journal of Financial Economics*, 89(3):444–466.

Kang, S. H., McIver, R., and Yoon, S.-M. (2017). Dynamic spillover effects among crude oil, precious metal, and agricultural commodity futures markets. *Energy Economics*, 62:19–32.

Karolyi, G. A., Lee, K.-H., and Van Dijk, M. A. (2012). Understanding commonality in liquidity around the world. *Journal of financial economics*, 105(1):82–112.

Kenett, D. Y., Huang, X., Vodenska, I., Havlin, S., and Stanley, H. E. (2015). Partial correlation analysis: Applications for financial markets. *Quantitative Finance*, 15(4):569–578.

Kock, A. B. and Callot, L. (2015). Oracle inequalities for high dimensional vector autoregressions. *Journal of Econometrics*, 186(2):325–344.

Koop, G., Pesaran, M. H., and Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, 74(1):119–147.

Korobilis, D. (2013). Var forecasting using bayesian variable selection. *Journal of Applied Econometrics*, 28(2):204–230.

Krampe, J. and Paparoditis, E. (2021). Sparsity concepts and estimation procedures for high-dimensional vector autoregressive models. *Journal of Time Series Analysis*, 42(5-6):554–579.

Kurka, J. (2019). Do cryptocurrencies and traditional asset classes influence each other? *Finance Research Letters*, 31:38–46.

Lahiri, S. and Lahiri, S. (2003). *Resampling methods for dependent data*. Springer Science & Business Media.

Langley, P. (2014). *Liquidity lost: The governance of the global financial crisis*. OUP Oxford.

Lee, H.-C. and Chang, S.-L. (2013). Spillovers of currency carry trade returns, market risk sentiment, and us market returns. *The North American Journal of Economics and Finance*, 26:197–216.

Lee, W. and Liu, Y. (2012). Simultaneous multiple response regression and inverse covariance matrix estimation via penalized gaussian maximum likelihood. *Journal of multivariate analysis*, 111:241–255.

Li, J., Yao, Y., Li, J., and Zhu, X. (2019). Network-based estimation of systematic and idiosyncratic contagion: The case of chinese financial institutions. *Emerging Markets Review*, 40:100624.

Lin, J. and Michailidis, G. (2017). Regularized estimation and testing for high-dimensional multi-block vector-autoregressive models. *Journal of Machine Learning Research*, 18.

Liu, X. (2017). Measuring systemic risk with regime switching in tails. *Economic Modelling*, 67:55–72.

Loh, P.-L. and Wainwright, M. J. (2011). High-dimensional regression with noisy and missing data: Provable guarantees with non-convexity. *Advances in neural information processing systems*, 24.

Lütkepohl, H. (2013). Vector autoregressive models. *Handbook of research methods and applications in empirical macroeconomics*, 30.

Ma, J. and Michailidis, G. (2016). Joint structural estimation of multiple graphical models. *The Journal of Machine Learning Research*, 17(1):5777–5824.

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of economic literature*, 35(1):13–39.

Malatesta, P. H. and Thompson, R. (1985). Partially anticipated events: A model of stock price reactions with an application to corporate acquisitions. *Journal of Financial Economics*, 14(2):237–250.

Masini, R. P., Medeiros, M. C., and Mendes, E. F. (2022). Regularized estimation of high-dimensional vector autoregressions with weakly dependent innovations. *Journal of Time Series Analysis*, 43(4):532–557.

Medeiros, M. C. and Mendes, E. F. (2016a). l (1)-regularization of high-dimensional time-series models with non-gaussian and heteroskedastic errors.

Medeiros, M. C. and Mendes, E. F. (2016b). 1-regularization of high-dimensional time-series models with non-gaussian and heteroskedastic errors. *Journal of Econometrics*, 191(1):255–271.

Meinshausen, N. and Bühlmann, P. (2006). High-dimensional graphs and variable selection with the lasso.

Melnyk, I. and Banerjee, A. (2016). Estimating structured vector autoregressive models. In *International Conference on Machine Learning*, pages 830–839. PMLR.

Miller, D. L. (2023). An introductory guide to event study models. *Journal of Economic Perspectives*, 37(2):203–230.

Moratis, G. (2021). Quantifying the spillover effect in the cryptocurrency market. *Finance Research Letters*, 38:101534.

Nakamura, E. and Steinsson, J. (2018). High-frequency identification of monetary non-neutrality: the information effect. *The Quarterly Journal of Economics*, 133(3):1283–1330.

Nardi, Y. and Rinaldo, A. (2011). Autoregressive process modeling via the lasso procedure. *Journal of Multivariate Analysis*, 102(3):528–549.

Nicholson, W. B., Matteson, D. S., and Bien, J. (2017). Varx-l: Structured regularization for large vector autoregressions with exogenous variables. *International Journal of Forecasting*, 33(3):627–651.

Nicholson, W. B., Wilms, I., Bien, J., and Matteson, D. S. (2020). High dimensional forecasting via interpretable vector autoregression. *The Journal of Machine Learning Research*, 21(1):6690–6741.

Nobi, A., Lee, S., Kim, D. H., and Lee, J. W. (2014). Correlation and network topologies in global and local stock indices. *Physics Letters A*, 378(34):2482–2489.

Nucera, F., Schwaab, B., Koopman, S. J., and Lucas, A. (2016). The information in systemic risk rankings. *Journal of Empirical Finance*, 38:461–475.

Onnela, J.-P., Chakraborti, A., Kaski, K., Kertesz, J., and Kanto, A. (2003). Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, 68(5):056110.

Peng, J., Wang, P., Zhou, N., and Zhu, J. (2009). Partial correlation estimation by joint sparse regression models. *Journal of the American Statistical Association*, 104(486):735–746.

Pesaran, H. H. and Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, 58(1):17–29.

Poignard, B. (2020). Asymptotic theory of the adaptive sparse group lasso. *Annals of the Institute of Statistical Mathematics*, 72:297–328.

Ramey, V. A. (2016). Macroeconomic shocks and their propagation. *Handbook of macroeconomics*, 2:71–162.

Rothman, A. J., Levina, E., and Zhu, J. (2010). Sparse multivariate regression with covariance estimation. *Journal of Computational and Graphical Statistics*, 19(4):947–962.

Salim, M. Z., Ramdhan, D., and Daly, K. (2023). Centrality measures of financial system interconnectedness: A multiple crises study. *Heliyon*, 9(4).

Schneider, M., Lillo, F., and Pelizzon, L. (2018). Modelling illiquidity spillovers with hawkes processes: an application to the sovereign bond market. *Quantitative Finance*, 18(2):283–293.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, pages 461–464.

Seth, A. K., Barrett, A. B., and Barnett, L. (2015). Granger causality analysis in neuroscience and neuroimaging. *Journal of Neuroscience*, 35(8):3293–3297.

Shojaie, A. and Michailidis, G. (2010). Discovering graphical granger causality using the truncating lasso penalty. *Bioinformatics*, 26(18):i517–i523.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, pages 1–48.

Sims, C. A. and Zha, T. (1999). Error bands for impulse responses. *Econometrica*, 67(5):1113–1155.

Smimou, K. and Khallouli, W. (2016). On the intensity of liquidity spillovers in the eurozone. *International Review of Financial Analysis*, 48:388–405.

Smith, S. M. (2012). The future of fmri connectivity. *Neuroimage*, 62(2):1257–1266.

Song, S. and Bickel, P. J. (2011). Large vector auto regressions. *arXiv preprint arXiv:1106.3915*.

Stock, J. H. and Watson, M. W. (1988). Testing for common trends. *Journal of the American statistical Association*, 83(404):1097–1107.

Stock, J. H. and Watson, M. W. (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, 97(460):1167–1179.

Stock, J. H. and Watson, M. W. (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162.

Stock, J. H. and Watson, M. W. (2006). Forecasting with many predictors. *Handbook of economic forecasting*, 1:515–554.

Stock, J. H. and Watson, M. W. (2011). Dynamic factor models.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288.

Tiwari, A. K., Abakah, E. J. A., Gabauer, D., and Dwumfour, R. A. (2022). Dynamic spillover effects among green bond, renewable energy stocks and carbon markets during covid-19 pandemic: Implications for hedging and investments strategies. *Global Finance Journal*, 51:100692.

Tobias, A. and Brunnermeier, M. K. (2016). Covar. *The American Economic Review*, 106(7):1705.

Torri, G., Giacometti, R., and Paterlini, S. (2018). Robust and sparse banking network estimation. *European Journal of Operational Research*, 270(1):51–65.

Wang, G.-J., Wan, L., Feng, Y., Xie, C., Uddin, G. S., and Zhu, Y. (2023). Interconnected multilayer networks: Quantifying connectedness among global stock and foreign exchange markets. *International Review of Financial Analysis*, 86:102518.

Wang, G.-J., Xie, C., Jiang, Z.-Q., and Stanley, H. E. (2016). Who are the net senders and recipients of volatility spillovers in china's financial markets? *Finance research letters*, 18:255–262.

Wang, G.-J., Yi, S., Xie, C., and Stanley, H. E. (2021). Multilayer information spillover networks: measuring interconnectedness of financial institutions. *Quantitative Finance*, 21(7):1163–1185.

Wang, H., Li, G., and Tsai, C.-L. (2007). Regression coefficient and autoregressive order shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 69(1):63–78.

Wen, T. and Wang, G.-J. (2020). Volatility connectedness in global foreign exchange markets. *Journal of Multinational Financial Management*, 54:100617.

Wilms, I. and Croux, C. (2016). Forecasting using sparse cointegration. *International Journal of Forecasting*, 32(4):1256–1267.

Wilms, I. and Croux, C. (2018). An algorithm for the multivariate group lasso with covariance estimation. *Journal of Applied Statistics*, 45(4):668–681.

Wright, J. H. (2012). What does monetary policy do to long-term interest rates at the zero lower bound? *The Economic Journal*, 122(564):F447–F466.

Xue, Y. and Taniguchi, M. (2020). Modified lasso estimators for time series regression models with dependent disturbances. *Statistical Methods & Applications*, 29(4):845–869.

Yang, J., Li, Z., and Miao, H. (2021). Volatility spillovers in commodity futures markets: A network approach. *Journal of Futures Markets*, 41(12):1959–1987.

Yi, S., Xu, Z., and Wang, G.-J. (2018). Volatility connectedness in the cryptocurrency market: Is bitcoin a dominant cryptocurrency? *International Review of Financial Analysis*, 60:98–114.

Yuan, M. and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 68(1):49–67.

Yuan, M. and Lin, Y. (2007). Model selection and estimation in the gaussian graphical model. *Biometrika*, 94(1):19–35.

Yun, T.-S., Jeong, D., and Park, S. (2019). "too central to fail" systemic risk measure using pagerank algorithm. *Journal of Economic Behavior & Organization*, 162:251–272.

Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty.

Zhao, L., Li, Y., and Wu, Y. J. (2022). An identification algorithm of systemically important financial institutions based on adjacency information entropy. *Computational Economics*, 59(4):1735–1753.

Zhao, L., Wang, G.-J., Wang, M., Bao, W., Li, W., and Stanley, H. E. (2018). Stock market as temporal network. *Physica A: Statistical Mechanics and its Applications*, 506:1104–1112.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, 101(476):1418–1429.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2):301–320.

Zou, H., Hastie, T., and Tibshirani, R. (2007). On the "degrees of freedom" of the lasso.

Appendix

A Simulation Appendix

A.1 Sparse B and band Ω

Figure A1: Relative estimation accuracy (MAEE) of the regularization methods as a function of the covariance matrix (Σ) Toeplitz-parameter (ρ).
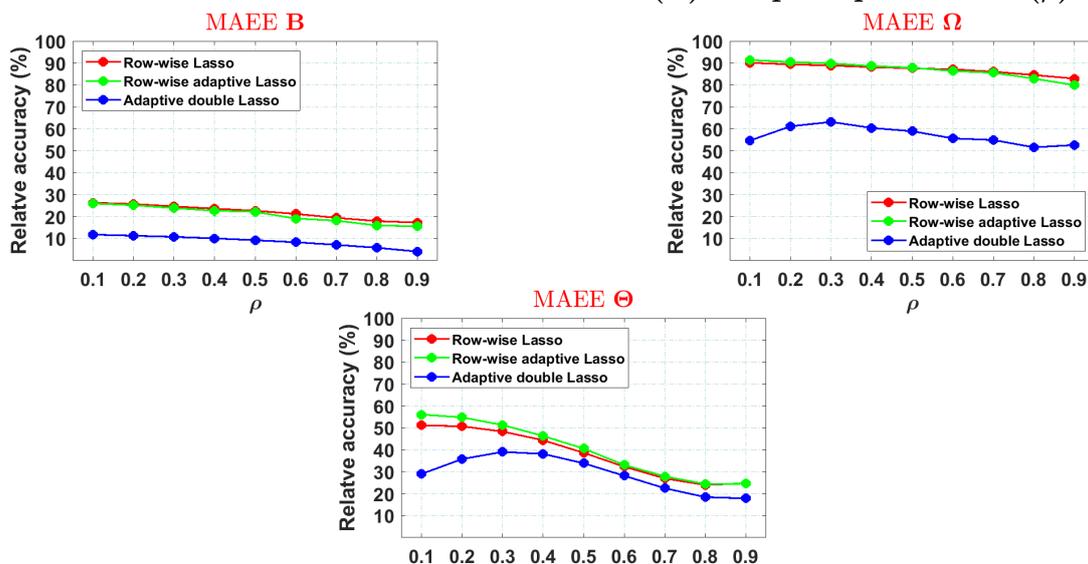
Notes: Results are in percentage of the benchmark OLS model and are based on 1000 simulation runs. The structure of B, Ω and Θ, along with $T = 250$ and $J = 15$ is fixed throughout the simulation. The setting corresponding to $\rho = 0.3$ is the same as the second column in Table 4 and in Table 7.

Figure A2: Relative estimation accuracy (MFBN) of the regularization methods as a function of the covariance matrix (Σ) Toeplitz-parameter (ρ).
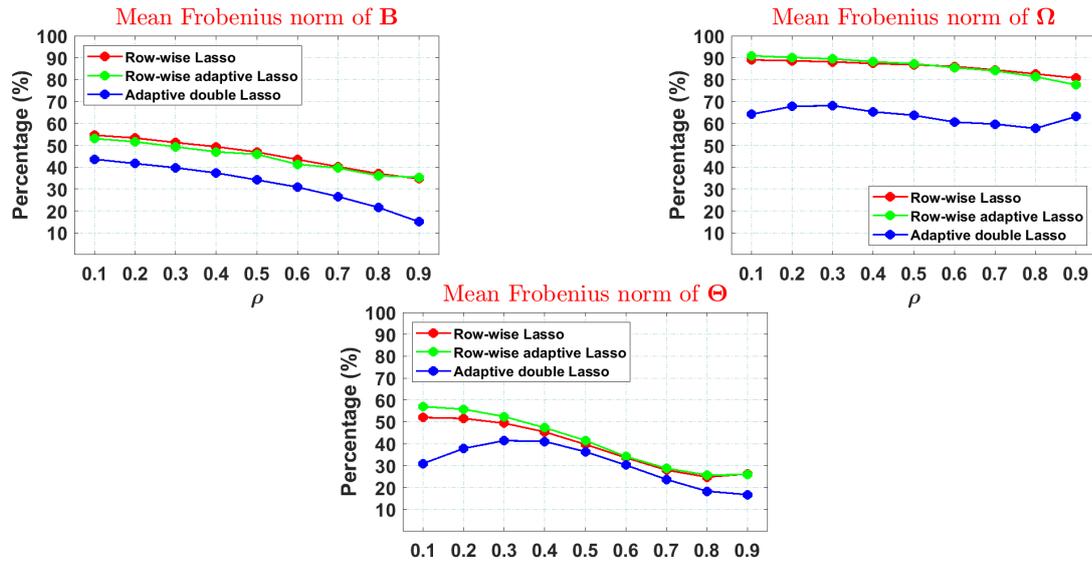
Notes: Results are in percentage of the benchmark OLS model. Results are based on 1000 simulation runs. The structure of B, Ω and Θ, along with $T = 250$ and $J = 15$ is fixed throughout the simulation. The setting corresponding to $\rho = 0.3$ is the same as the second column in Table 4 and in Table 7.

Table A1: Summary of simulation results for sparse coefficient matrix B and sparse inverse covariance matrix Ω for OLS, RW LASSO, ARW LASSO, and AJ LASSO methods

Panel A: T = 100	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	97%	97%	93%
True negative rate (B)	0%	79%	84%	96%
Sparsity recognition (B)	7%	80%	85%	96%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	76%
Sparsity recognition (Ω)	7%	7%	7%	77%
Panel B: T = 250	*OLS*	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	81%	84%	99%
Sparsity recognition (B)	7%	82%	85%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	81%
Sparsity recognition (Ω)	7%	7%	7%	83%
Panel C: T = 500	*OLS*	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
True positive rate (B)	100%	100%	100%	100%
True negative rate (B	0%	82%	86%	99%
Sparsity recognition (B)	7%	84%	87%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	90%
Sparsity recognition (Ω)	7%	7%	7%	91%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0$. In the table, in the first column, the rows indicate the mean of sparsity recognition performance measures across simulated paths, and columns 2 through 5 represent the alternative methods: OLS, RW LASSO, ARW LASSO, and AJ LASSO. The structures of B and Ω, along with $J = 15$, are fixed throughout the simulation. Time series length (T) varies across Panels A, B, and C in the table.*

Table A2: Accuracy comparison of regularization methods with a simulation of sparse coefficient matrix B and sparse inverse covariance matrix Ω

Panel A: T = 100	RW LASSO	ARW LASSO	AJ LASSO
MAEE (B,$\hat{B}$)	28%	28%	17%
MAEE (Ω, $\hat{\Omega}$)	77%	78%	20%
MAEE (Θ, $\hat{\Theta}$)	49%	54%	13%
MFBN (B,$\hat{B}$)	58%	56%	51%
MFBN (Ω, $\hat{\Omega}$)	76%	78%	31%
MFBN (Θ, $\hat{\Theta}$)	51%	56%	16%
Panel B: T = 250	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
MAEE (B,$\hat{B}$)	27%	27%	12%
MAEE (Ω, $\hat{\Omega}$)	91%	92%	21%
MAEE (Θ, $\hat{\Theta}$)	50%	56%	9%
MFBN (B,$\hat{B}$)	56%	54%	43%
MFBN (Ω, $\hat{\Omega}$)	89%	91%	35%
MFBN (Θ, $\hat{\Theta}$)	52%	57%	10%
Panel C: T = 500	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
MAEE (B,$\hat{B}$)	28%	28%	11%
MAEE (Ω, $\hat{\Omega}$)	95%	96%	15%
MAEE (Θ, $\hat{\Theta}$)	52%	58%	4%
MFBN (B,$\hat{B}$)	59%	56%	39%
MFBN (Ω, $\hat{\Omega}$)	94%	95%	34%
MFBN (Θ, $\hat{\Theta}$)	53%	60%	6%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0$. In the table, in the first column, the rows indicate the mean of sparsity error measures (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm), and columns 2 through 4 represent the three alternative methods: RW LASSO, ARW LASSO, and AJ LASSO. Results are in percentage of the benchmark OLS model. The structures of B and Ω, along with $J = 15$, are fixed throughout the simulation. Time series length (T) varies across the Panels A, B, and C in the table.*

Table A3: Summary of simulation results with sparse coefficient B and sparse inverse covariance Ω matrix

Panel A: J = 10	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	80%	85%	99%
Sparsity recognition (B)	10%	82%	87%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	74%
Sparsity recognition (Ω)	10%	10%	10%	77%
Panel B: J = 15	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	81%	84%	99%
Sparsity recognition (B)	7%	82%	85%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	81%
Sparsity recognition (Ω)	7%	7%	7%	83%
Panel C: J = 20	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	82%	86%	99%
Sparsity recognition (B)	5%	83%	86%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	79%
Sparsity recognition (Ω)	5%	5%	5%	80%
Panel D: J = 25	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	84%	85%	99%
Sparsity recognition (B)	4%	85%	85%	99%
True positive rate (Ω)	100%	100%	100%	100%
True negative rate (Ω)	0%	0%	0%	78%
Sparsity recognition (Ω)	4%	4%	4%	79%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0$. In the table, in the first column, the rows indicate the mean of sparsity recognition performance measures across simulated paths, and columns 2 through 5 represent the alternative methods: OLS, RW LASSO, ARW LASSO, and AJ LASSO. The structures of B and Ω are fixed throughout the simulation. The length of the time series (T) is fixed at 250, while J (dimension of the matrices) varies between 10 and 25 across the Panels A through D in the table. Panel B of the table is the same as Panel B of Table A1.*

Table A4: Accuracy comparison of regularization methods with a simulation of sparse coefficient matrix B and sparse inverse covariance matrix Ω using a range of the dimension (J) of the system

Panel A: J = 10	RW LASSO	ARW LASSO	AJ LASSO
MAEE (B,$\hat{B}$)	32%	28%	14%
MAEE (Ω, $\hat{\Omega}$)	94%	95%	31%
MAEE (Θ, $\hat{\Theta}$)	52%	58%	10%
MFBN (B,$\hat{B}$)	61%	55%	41%
MFBN (Ω, $\hat{\Omega}$)	93%	94%	47%
MFBN (Θ, $\hat{\Theta}$)	54%	61%	11%
Panel B: J = 15	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
MAEE (B,$\hat{B}$)	27%	27%	12%
MAEE (Ω, $\hat{\Omega}$)	91%	92%	21%
MAEE (Θ, $\hat{\Theta}$)	50%	56%	9%
MFBN (B,$\hat{B}$)	56%	54%	43%
MFBN (Ω, $\hat{\Omega}$)	89%	91%	35%
MFBN (Θ, $\hat{\Theta}$)	52%	57%	10%
Panel C: J = 20	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
MAEE (B,$\hat{B}$)	23%	23%	9%
MAEE (Ω, $\hat{\Omega}$)	88%	89%	19%
MAEE (Ψ, $\hat{\Psi}$)	47%	52%	8%
MFBN (B,$\hat{B}$)	51%	50%	38%
MFBN (Ω, $\hat{\Omega}$)	86%	87%	32%
MFBN (Θ, $\hat{\Theta}$)	49%	54%	10%
Panel C: J = 25	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
MAEE (B,$\hat{B}$)	20%	22%	8%
MAEE (Ω, $\hat{\Omega}$)	84%	86%	18%
MAEE (Θ, $\hat{\Theta}$)	45%	51%	8%
MFBN (B,$\hat{B}$)	47%	48%	35%
MFBN (Ω, $\hat{\Omega}$)	81%	84%	29%
MFBN (Θ, $\hat{\Theta}$)	46%	52%	9%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with $\rho = 0$. In the table, in the first column, the rows indicate the mean of sparsity error measures (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm), and columns 2 through 4 represent the three alternative methods: RW LASSO, ARW LASSO, and AJ LASSO. Results are in percentage of the benchmark OLS model. The time series (T) length is fixed at 250, while J (dimension of the matrices) varies across Panels A through D in the table. Panel B of the table is the same as Panel B of Table A2.*

A.2 Sparse B and sparse & dense Ω

Table A5: Summary of simulation results for sparse coefficient matrix B and dense inverse covariance matrix Ω for OLS, RW LASSO, ARW LASSO and AJ LASSO methods

Panel A: T = 100	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	95%	93%	93%
True negative rate (B)	0%	78%	83%	95%
Sparsity recognition (B)	7%	79%	84%	95%
Panel B: T = 250	*OLS*	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	77%	86%	99%
Sparsity recognition (B)	7%	78%	87%	99%
Panel C: T = 500	*OLS*	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	79%	86%	99%
Sparsity recognition (B)	7%	80%	87%	99%

Notes: *Results are based on 1000 simulation runs. In the table, in the first column, the rows indicate the mean of sparsity recognition performance measures across simulated paths, and columns 2 through 5 represent the alternative methods: OLS, RW LASSO, ARW LASSO, and AJ LASSO. The structures of B and Ω, along with J = 15, are fixed throughout the simulation. Time series length (T) varies across Panels A, B, and C in the table. The table contains only the sparsity recognition performance for B because Ω is dense.*

Table A6: Accuracy comparison of regularization methods with a simulation of sparse coefficient matrix B and dense inverse covariance matrix Ω

Panel A: T = 100	RW LASSO	ARW LASSO	AJ LASSO
MAEE ($B,\hat{B}$)	22%	23%	10%
MAEE ($\Omega, \hat{\Omega}$)	70%	68%	42%
MAEE ($\Theta, \hat{\Theta}$)	52%	56%	47%
MFBN ($B,\hat{B}$)	43%	46%	29%
MFBN ($\Omega, \hat{\Omega}$)	69%	67%	44%
MFBN ($\Theta, \hat{\Theta}$)	49%	54%	43%
Panel B: T = 250	RW LASSO	ARW LASSO	AJ LASSO
MAEE ($B,\hat{B}$)	23%	20%	7%
MAEE ($\Omega, \hat{\Omega}$)	86%	85%	59%
MAEE ($\Theta, \hat{\Theta}$)	49%	50%	41%
MFBN ($B,\hat{B}$)	44%	43%	26%
MFBN ($\Omega, \hat{\Omega}$)	85%	84%	61%
MFBN ($\Theta, \hat{\Theta}$)	43%	45%	35%
Panel C: T = 500	RW LASSO	ARW LASSO	AJ LASSO
MAEE ($B,\hat{B}$)	23%	22%	6%
MAEE ($\Omega, \hat{\Omega}$)	93%	93%	77%
MAEE ($\Theta, \hat{\Theta}$)	43%	45%	36%
MFBN ($B,\hat{B}$)	46%	45%	23%
MFBN ($\Omega, \hat{\Omega}$)	91%	91%	78%
MFBN ($\Theta, \hat{\Theta}$)	39%	41%	29%

Notes: *Results are based on 1000 simulation runs. In the table, in the first column, the rows indicate the mean of sparsity error measures (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm), and columns 2 through 4 represent the three alternative methods: RW LASSO, ARW LASSO, and AJ LASSO. Results are in percentage of the benchmark OLS model. The structures of B and Ω, along with $J = 15$, are fixed throughout the simulation. Time series length (T) varies across Panels A, B, and C in the table.*

Table A7: Summary of simulation results with sparse coefficient B and dense inverse covariance Ω matrix

Panel A: J = 10	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	78%	86%	99%
Sparsity recognition (B)	10%	81%	88%	99%
Panel B: J = 15	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	77%	86%	99%
Sparsity recognition (B)	7%	78%	87%	99%
Panel C: J = 20	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	83%	88%	99%
Sparsity recognition (B)	5%	84%	89%	99%
Panel D: J = 25	OLS	RW LASSO	ARW LASSO	AJ LASSO
True positive rate (B)	100%	100%	100%	100%
True negative rate (B)	0%	83%	88%	99%
Sparsity recognition (B)	5%	84%	89%	99%

Notes: *Results are based on 1000 simulation runs. In the table, in the first column, the rows indicate the mean of sparsity recognition performance measures across simulated paths, and columns 2 through 5 represent the alternative methods: OLS, RW LASSO, ARW LASSO, and AJ LASSO. The structures of B and Ω are fixed throughout the simulation. The length of the time series (T) is fixed at 250, while J (dimension of the matrices) varies between 10 and 25 across the Panels A through D in the table. Panel B of the table is the same as Panel B of Table A5.*

Table A8: Accuracy comparison of regularization methods with a simulation of sparse coefficient matrix B and dense inverse covariance matrix Ω using a range of the dimension (J) of the system

Panel A: J = 10	RW LASSO	ARW LASSO	AJ LASSO
MAEE ($B,\hat{B}$)	27%	27%	8%
MAEE ($\Omega, \hat{\Omega}$)	93%	91%	72%
MAEE ($\Theta, \hat{\Theta}$)	54%	57%	47%
MFBN ($B,\hat{B}$)	50%	50%	24%
MFBN ($\Omega, \hat{\Omega}$)	91%	91%	74%
MFBN ($\Theta, \hat{\Theta}$)	49%	52%	40%
Panel B: J = 15	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
MAEE ($B,\hat{B}$)	23%	20%	7%
MAEE ($\Omega, \hat{\Omega}$)	86%	85%	59%
MAEE ($\Theta, \hat{\Theta}$)	49%	50%	41%
MFBN ($B,\hat{B}$)	44%	43%	26%
MFBN ($\Omega, \hat{\Omega}$)	85%	84%	61%
MFBN ($\Theta, \hat{\Theta}$)	43%	45%	35%
Panel C: J = 20	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
MAEE ($B,\hat{B}$)	17%	16%	5%
MAEE ($\Omega, \hat{\Omega}$)	83%	81%	51%
MAEE ($\Theta, \hat{\Theta}$)	40%	41%	33%
MFBN ($B,\hat{B}$)	37%	38%	22%
MFBN ($\Omega, \hat{\Omega}$)	80%	79%	53%
MFBN ($\Theta, \hat{\Theta}$)	37%	39%	27%
Panel C: J = 25	*RW LASSO*	*ARW LASSO*	*AJ LASSO*
MAEE ($B,\hat{B}$)	15%	17%	5%
MAEE ($\Omega, \hat{\Omega}$)	80%	79%	50%
MAEE ($\Theta, \hat{\Theta}$)	37%	40%	30%
MFBN ($B,\hat{B}$)	36%	39%	20%
MFBN ($\Omega, \hat{\Omega}$)	77%	76%	51%
MFBN ($\Theta, \hat{\Theta}$)	34%	38%	26%

Notes: *Results are based on 1000 simulation runs. In the table, in the first column, the rows indicate the mean of sparsity error measures (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm), and columns 2 through 4 represent the three alternative methods: RW LASSO, ARW LASSO, and AJ LASSO. Results are in percentage of the benchmark OLS model. The time series (T) length is fixed at 250, while J (dimension of the matrices) varies between 10 and 25 across Panels A through D in the Table. Panel B of the table is the same as Panel B of table A6.*

A.3 Dense B and band, sparse & dense Ω

Table A9: Summary of the accuracy gains of the AJ LASSO method relative to the better RW LASSO with dense coefficient (B) and band inverse covariance (Ω) matrix.

Relative Error	$T = 100$	$T = 250$	$T = 500$
$MAEE\ (B,\hat{B})$	81%	86%	85%
$MAEE\ (\Omega,\ \hat{\Omega})$	69%	74%	81%
$MAEE\ (\Theta,\ \hat{\Theta})$	77%	88%	90%
$MFBN\ (B,\hat{B})$	92%	98%	96%
$MFBN\ (\Omega,\ \hat{\Omega})$	74%	81%	85%
$MFBN\ (\Theta,\ \hat{\Theta})$	79%	93%	99%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with ρ = 0.3. Rows contain the mean of accuracy measures across simulated paths, and columns represent the explored methods. In the table, the first column, the rows indicate the mean of sparsity error measures for (B) and (Ω) matrix (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm). The relative performance of the AJ LASSO in comparison with the better RW LASSO with time series lengths (T) varying from 100 to 500 are presented in Columns 2 through 4. Results are in percentage of the better RW LASSO model. The structures of B and Ω, along with J = 15, are fixed throughout the simulation.*

Table A10: Summary of the accuracy gains of the AJ LASSO method relative to the better RW LASSO with dense coefficient (B) and band inverse covariance (Ω) matrix using a range of the dimension (J) of the system

Relative Error	$J = 10$	$J = 15$	$J = 20$	$J = 25$
$MAEE\ (B,\hat{B})$	93%	86%	79%	74%
$MAEE\ (\Omega,\ \hat{\Omega})$	85%	74%	64%	60%
$MAEE\ (\Theta,\ \hat{\Theta})$	92%	88%	75%	75%
$MFBN\ (B,\hat{B})$	103%	98%	94%	88%
$MFBN\ (\Omega,\ \hat{\Omega})$	89%	81%	70%	67%
$MFBN\ (\Theta,\ \hat{\Theta})$	99%	93%	78%	79%

Notes: *Results are based on 1000 simulation runs. Error terms are drawn from a Toeplitz covariance matrix with ρ = 0.3. Rows contain the mean of accuracy measures across simulated paths, and columns represent the explored methods. In the table, the first column, the rows indicate the mean of sparsity error measures for (B) and (Ω) matrix (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm). The relative performance of the AJ LASSO in comparison with the better RW LASSO with dimension of the time series (J) - varies from 10 to 25 -, are presented in Columns 2 through 5. Results are in percentage of the better RW LASSO model. The structures of B and Ω, along with T = 250, are fixed throughout the simulation.*

Table A11: Summary of large inaccuracies in the estimation with a dense coefficient (B) and band inverse covariance (Ω) matrix II.

Panel A: J = 10	RW LASSO	ARW LASSO	AJ LASSO
RLIA (B,$\hat{B}$)	61%	37%	22%
RLIA (Ω, $\hat{\Omega}$)	93%	100%	82%
RLIA (Θ, $\hat{\Theta}$)	98%	99%	101%
Panel B: T = 15	RW LASSO	ARW LASSO	AJ LASSO
RLIA (B,$\hat{B}$)	49%	35%	17%
RLIA (Ω, $\hat{\Omega}$)	70%	81%	72%
RLIA (Θ, $\hat{\Theta}$)	96%	97%	96%
Panel C: T = 20	RW LASSO	ARW LASSO	AJ LASSO
RLIA (B,$\hat{B}$)	41%	29%	13%
RLIA (Ω, $\hat{\Omega}$)	73%	75%	64%
RLIA (Θ, $\hat{\Theta}$)	96%	96%	96%
Panel D: T = 25	RW LASSO	ARW LASSO	AJ LASSO
RLIA (B,$\hat{B}$)	36%	27%	10%
RLIA (Ω, $\hat{\Omega}$)	58%	69%	46%
RLIA (Θ, $\hat{\Theta}$)	96%	96%	96%

Notes: *Error terms are drawn from a Toeplitz-type covariance matrix with $\rho = 0.7$. Rows contain the mean of wrongly estimated parameters across simulated paths, and columns represent the explored methods. RLIA = relative large inaccuracy (relative number of parameters where the estimation error is more than 20%). Results are in percentage of the benchmark OLS model. The number of time series (J) varies across the panels of the table. The structure of B, Ω and Θ, along with $T = 250$, is fixed throughout the simulation. Panel B of the table is the same as Panel B of Table 13*

Table A12: Summary of the accuracy gains of the AJ LASSO method relative to the better RW LASSO with dense coefficient (B) and dense inverse covariance (Ω) matrix.

Relative Error	$T = 100$	$T = 250$	$T = 500$
MAEE (B,$\hat{B}$)	62%	64%	61%
MAEE (Ω, $\hat{\Omega}$)	76%	84%	98%
MAEE (Θ, $\hat{\Theta}$)	99%	92%	79%
MFBN (B,$\hat{B}$)	67%	68%	67%
MFBN (Ω, $\hat{\Omega}$)	81%	89%	101%
MFBN (Θ, $\hat{\Theta}$)	97%	90%	70%

Notes: *Results are based on 1000 simulation runs. Rows contain the mean of accuracy measures across simulated paths, and columns represent the explored methods. In the table, the first column, the rows indicate the mean of sparsity error measures for (B) and (Ω) matrix (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm). The relative performance of the AJ LASSO, in comparison with the better RW LASSO with time series lengths (T) varying from 100 to 500, are presented in Columns 2 through 4. Results are in percentage of the better RW LASSO model. The structures of B and Ω, along with J = 15, are fixed throughout the simulation.*

Table A13: Summary of the accuracy gains of the AJ LASSO method relative to the better row-wise LASSO with dense coefficient (B) and dense inverse covariance (Ω) matrix using a range of the dimension (J) of the system

Relative Error	$J = 10$	$J = 15$	$J = 20$	$J = 25$
MAEE (B,$\hat{B}$)	65%	64%	59%	52%
MAEE (Ω, $\hat{\Omega}$)	94%	84%	77%	75%
MAEE (Θ, $\hat{\Theta}$)	92%	92%	84%	81%
MFBN (B,$\hat{B}$)	69%	68%	66%	59%
MFBN (Ω, $\hat{\Omega}$)	99%	89%	81%	80%
MFBN (Θ, $\hat{\Theta}$)	88%	90%	75%	71%

Notes: *Results are based on 1000 simulation runs. Rows contain the mean of accuracy measures across simulated paths, and columns represent the explored methods. In the table, the first column, the rows indicate the mean of sparsity error measures for (B) and (Ω) matrix (either MAEE = mean absolute estimation error or MFBN = mean Frobenius norm). The relative performance of the AJ LASSO in comparison with the better RW LASSO with dimension of the time series (J) - varies from 10 to 25 -, are presented in Columns 2 through 5. Results are in percentage of the better RW LASSO model. The structures of B and Ω, along with T = 250, are fixed throughout the simulation.*

Table A14: Simulation results with dense coefficient (B) and dense inverse covariance (Ω) matrix II.

Panel A: J = 10	RW LASSO	ARW LASSO	AJ LASSO
RLIA ($B,\hat{B}$)	50%	36%	21%
RLIA ($\Omega, \hat{\Omega}$)	99%	94%	89%
RLIA ($\Theta, \hat{\Theta}$)	45%	47%	29%
Panel B: T = 15	RW LASSO	ARW LASSO	AJ LASSO
RLIA ($B,\hat{B}$)	39%	29%	17%
RLIA ($\Omega, \hat{\Omega}$)	98%	97%	83%
RLIA ($\Theta, \hat{\Theta}$)	38%	40%	31%
Panel C: T = 20	RW LASSO	ARW LASSO	AJ LASSO
RLIA ($B,\hat{B}$)	34%	25%	13%
RLIA ($\Omega, \hat{\Omega}$)	98%	96%	83%
RLIA ($\Theta, \hat{\Theta}$)	37%	39%	29%
Panel D: T = 25	RW LASSO	ARW LASSO	AJ LASSO
RLIA ($B,\hat{B}$)	30%	25%	10%
RLIA ($\Omega, \hat{\Omega}$)	96%	95%	81%
RLIA ($\Theta, \hat{\Theta}$)	40%	42%	30%

Notes: *Rows contain the mean of wrongly estimated parameters across simulated paths, and columns represent the explored methods. RLIA = relative large inaccuracy (relative number of parameters where the estimation error is more than 20%). Results are in percentage of the benchmark OLS model. The number of time series (J) varies across the panels of the table. The structure of B, Ω and Θ, along with T = 250, is fixed throughout the simulation. Panel B of the table is the same as Panel B of Table 14.*

# B Event Study Appendix

B.1 Moving Block Bootstrap algorithm

Step 1. Obtain the residuals $\hat{u}_t = y_t - \hat{B}_1 y_{t-1} - \hat{B}_2 y_{t-2} - \cdots - \hat{B}_p y_{t-p}$, $t = 1, ..., T$ using $\hat{B}_1, \hat{B}_2 \ldots, \hat{B}_p$ from an estimated the VAR(p) model. I use the adaptive joint LASSO algorithm for estimating the VAR(p) model.

Step 2. Generate $s = 1, 2, \ldots, S$ bootstrap samples of the residuals in the next way. Choose a block length $l < T$, and let $N = T/l$ be the number of blocks. In this case, we achieve that $lN \leq T$. Get $(J \times l)$-dimensional block $S_{i,l} = (\hat{u}_{i+1}, \ldots, \hat{u}_{i+l})$. After that I generate i.i.d. random variables uniformly distributed on the set $\{0, 1, 2, \ldots, T - l\}$ where $i = 0, \ldots, T - l$, and $i_0, \ldots, i_{N-1}$. Discard the last $Nl - T$ values from the blocks $S_{i0,l}, \ldots S_{iN-1,l}$, and get bootstrap residuals $\hat{u}^*_1, \ldots, \hat{u}^*_T$.

Step 3. Center the residuals for all $t = 1, \ldots, T$ as $u^*_{jl+s} = \hat{u}^*_{jl+s} - E^*(\hat{u}^*_{jl+s}) = \hat{u}^*_{jl+s} - \frac{1}{T-l+1} \sum_{r=0}^{T-l} \hat{u}_{s+r}$ to get $E^*(u^*_t) = 0$.

Step 4. Set the initial values of $y^*_{p+1} \ldots, y^*_0$ equal to zero and generate bootstrap sample $y^*_1, \ldots, y^*_T$ by fitting the following model $y^*_t = \hat{B}_1 y^*_{t-1} \hat{B}_2 y^*_{t-2} + \cdots + \hat{B}_p y^*_{t-p} + u^*_t$.

Step 5. Re-estimate the adaptive joint LASSO-VAR(p) model on sample $y^*_{p+1}, y^*_p \ldots, y^*_0$ and calculate the Diebold-Yilmaz spillover measures.

Step 6. Repeat Step 2 to Step 5 several times.

B.2   Illiquidity and Volatility network

**Table B1: U.S. FIs, tickers, and total assets as of the end of 2018, trillion USD**

Panel A: Big Financial Institutions	Ticker	Total assets
JPMorgan Chase & Co	JPM	2.6225
Bank of America	BAC	2.3545
Citigroup	C	1.9174
Wells Fargo	WFC	1.8959
Goldman Sachs Group	GS	0.9318
Panel B: Medium Financial Institutions	Ticker	Total assets
Moran Stanley	MS	0.8535
Bank of New York Mellon	BK	0.3629
U.S. Bancorp	USB	0.4674
PNC Financial Services Group	PNC	0.3823
Capital One Financial	COF	0.3725
Panel C: Small Financial Institutions	Ticker	Total assets
State Street Corporation	STT	0.2446
BB&T Corp	BBT	0.2257
SunTrust Banks	STI	0.2155
American Express	AXP	0.1886
Fifth Third Bancorp	FITB	0.1461
Regions Financial Corporation	RF	0.1257

Figure B3: Distribution of illiquidity and volatility NET spillover indices for the network from 2004 to 2019.
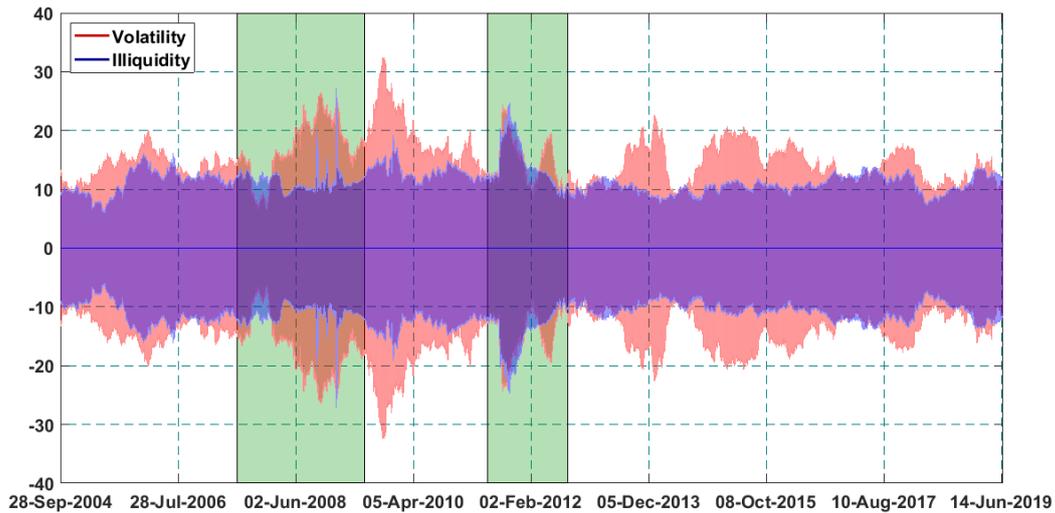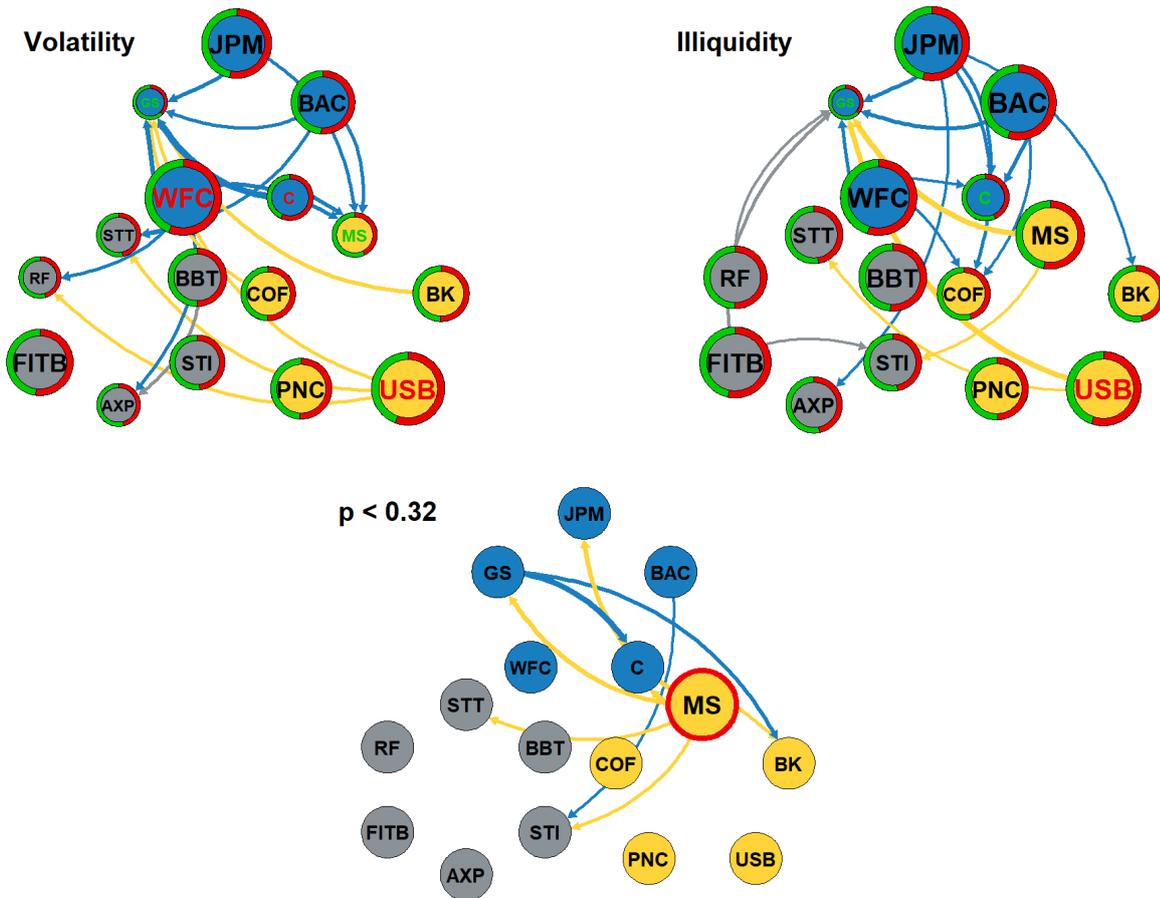

Notes: The information of Diebold-Yilmaz network is calculated from a rolling window analysis with $T = 200$, adaptive joint LASSO-VAR(2) estimation. Dates correspond to the end date of the windows. Horizontal shaded areas around the mean of the NET spillover indices (red for the volatility index and blue for the illiquidity index) represent the distribution with one standard deviation. Vertical (green) shaded highlight crisis periods (GFC and ESDC).

Figure B4: Distribution of illiquidity and volatility pairwise NET spillover indices for the network from 2004 to 2019.
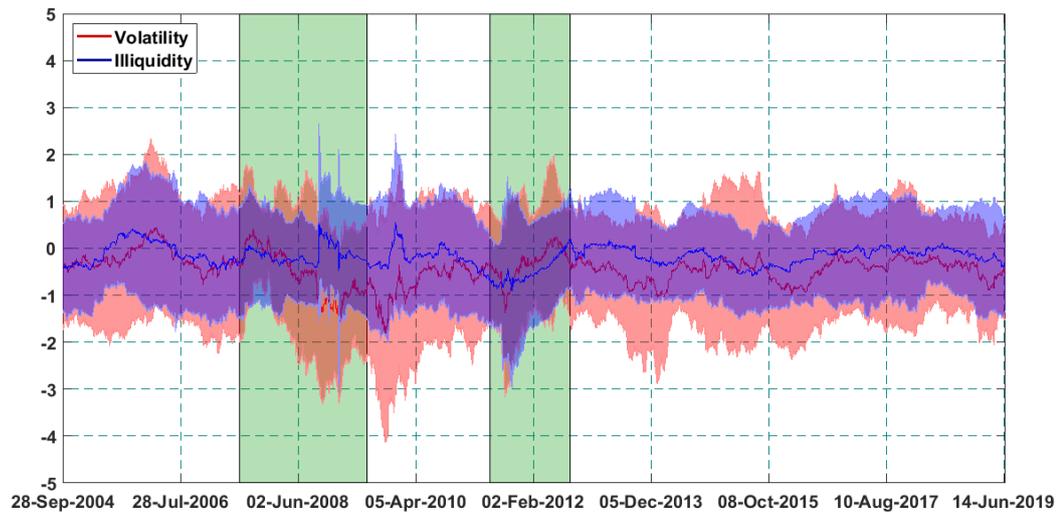
Notes: The information of Diebold-Yilmaz network is calculated from a rolling window analysis with $T = 200$, adaptive joint LASSO-VAR(2) estimation. Dates correspond to the end date of the windows. Horizontal shaded areas around the mean of the pairwise NET spillover indices (red for the volatility index and blue for the illiquidity index) represent the distribution with one standard deviation. Vertical (green) shaded highlight crisis periods (GFC and ESDC).

Figure B5: Estimated volatility and illiquidity network on 2010-10-27

Notes: Volatility (Illiquidity) network are in the upper left (right) chart, and the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "small" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 200$, adaptive joint LASSO-VAR(2) estimation.

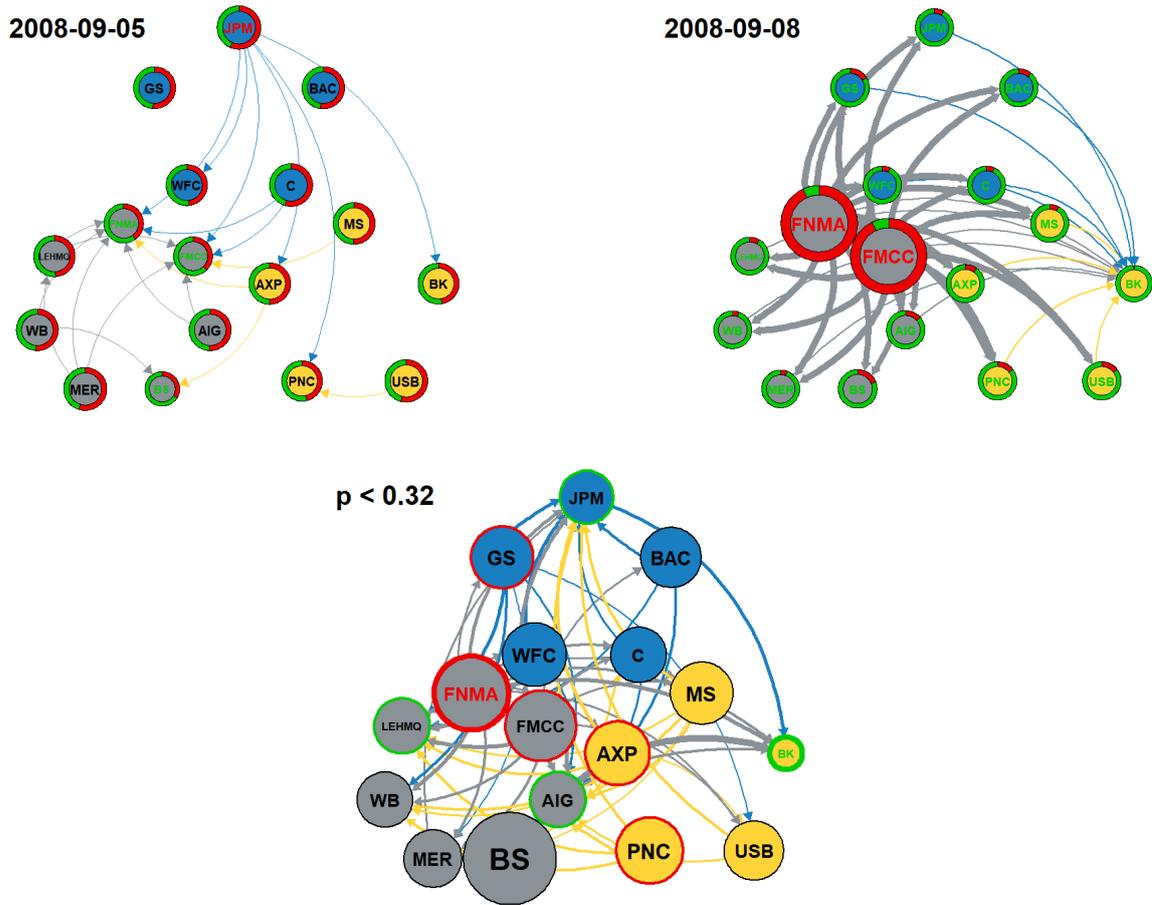**Figure B6: Estimated volatility and illiquidity network on 2014-05-23**
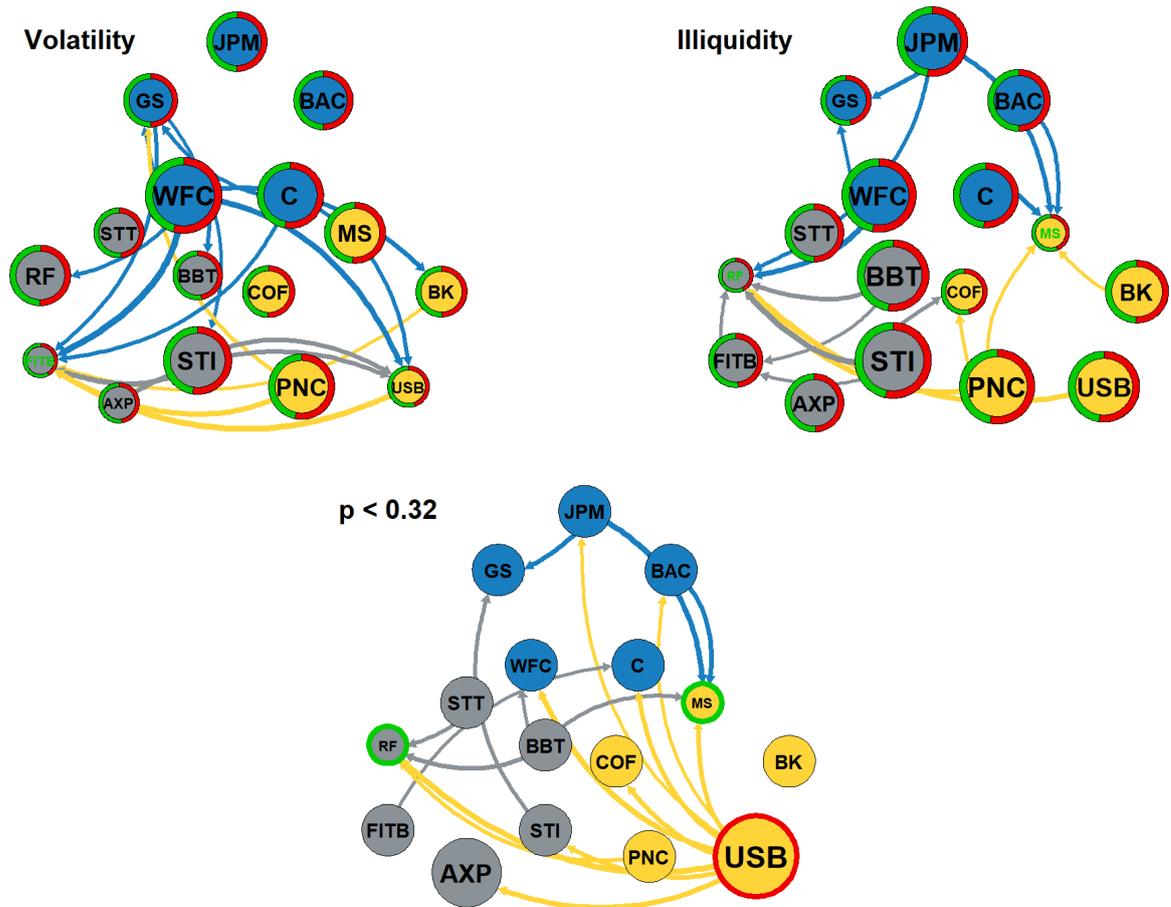
Notes: Volatility (Illiquidity) network are in the upper left (right) chart, and the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) of SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "small" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 200$, adaptive joint LASSO-VAR(2) estimation.

B.3 Key events of the GFC

Table B2: Summary statistics of volatility and illiquidity of stock prices between January 5, 2000, and December 22, 2008
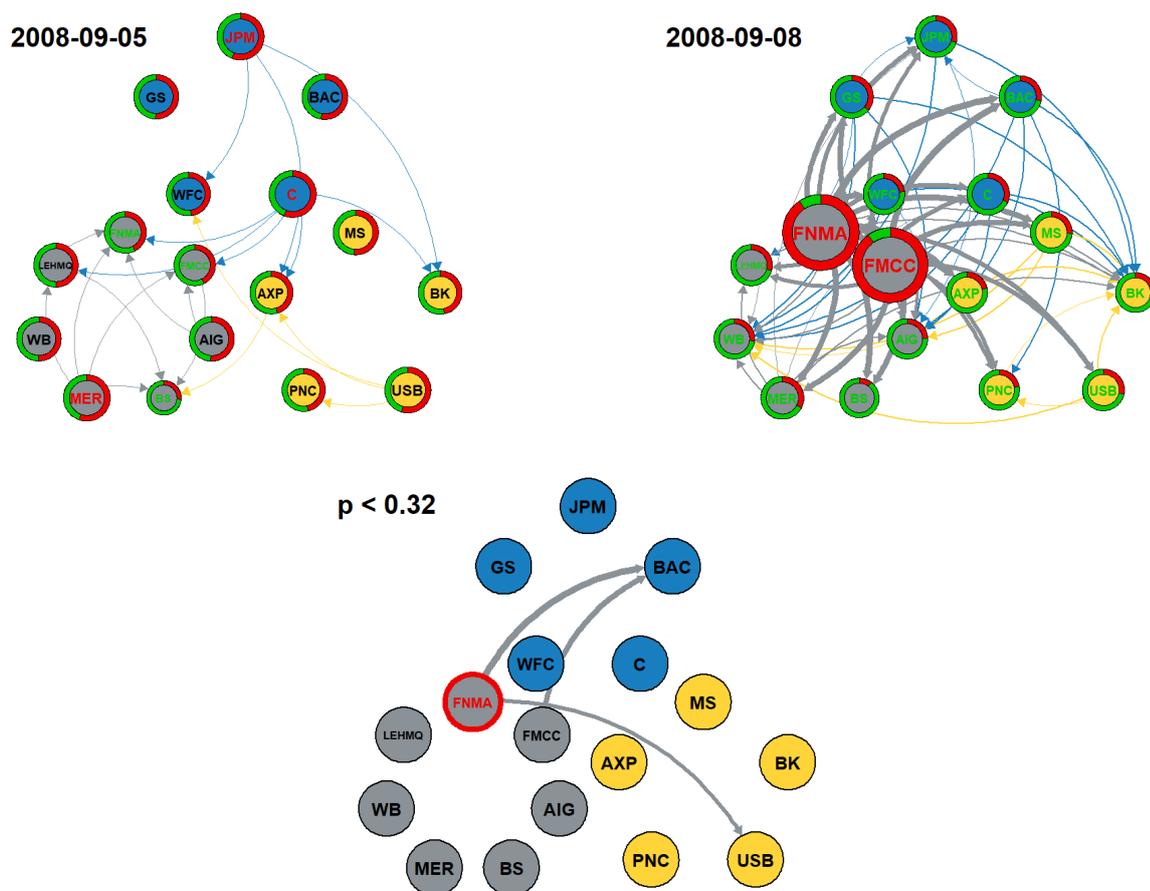
Ticker	Volatility					Illiquidity (10^{-10})				
	Mean	Median	Std	Min	Max	Mean	Median	Std	Min	Max
Panel A: Big Financial Institutions										
JPM	3.299	3.295	0.649	1.590	5.710	0.362	0.219	0.417	0	3.435
BAC	3.065	3.003	0.661	1.536	5.946	0.256	0.146	0.316	0	3.369
C	3.220	3.206	0.666	1.405	5.852	0.185	0.121	0.194	0	1.872
WFC	3.035	2.979	0.658	1.424	5.592	0.402	0.273	0.484	0	6.546
GS	3.357	3.313	0.568	1.562	5.951	0.478	0.222	0.788	0	9.071
Panel A: Medium Financial Institutions										
MS	3.502	3.472	0.607	2.009	6.755	0.668	0.473	0.662	0	6.292
BK	3.298	3.266	0.612	1.557	6.343	1.314	0.852	1.440	0	13.916
USB	3.161	3.148	0.663	1.487	5.305	1.237	0.571	1.870	0	17.771
PNC	3.091	3.046	0.606	1.273	6.027	1.729	1.081	2.090	0	21.461
AXP	3.225	3.205	0.666	1.057	5.444	0.646	0.420	0.695	0	5.056
Panel C: Troubled Financial Institutions										
FNMA	3.350	3.235	0.719	1.759	7.184	0.870	0.408	3.944	0	104.607
FMCC	3.297	3.158	0.752	1.620	7.583	1.343	0.534	9.071	0	357.380
AIG	3.165	3.098	0.673	1.291	7.406	0.380	0.231	0.755	0	19.406
BS	3.335	3.313	0.562	1.867	6.708	1.953	0.868	2.922	0	36.823
MER	3.423	3.395	0.640	1.684	6.225	0.489	0.335	0.484	0	3.320
WB	3.211	3.124	0.700	1.366	7.123	0.899	0.426	1.273	0	23.857
LEHMQ	3.526	3.442	0.710	1.736	7.995	7.866	0.550	82.362	0	2151.951

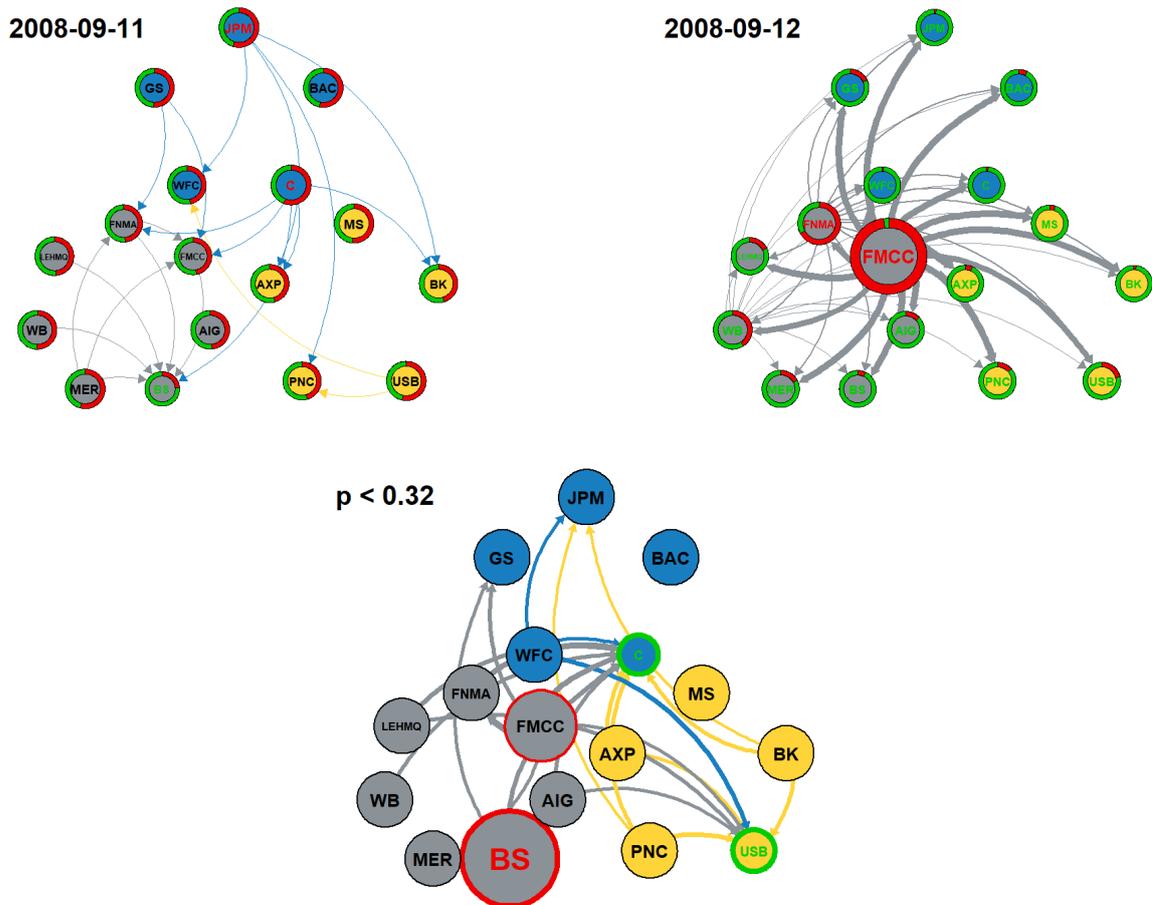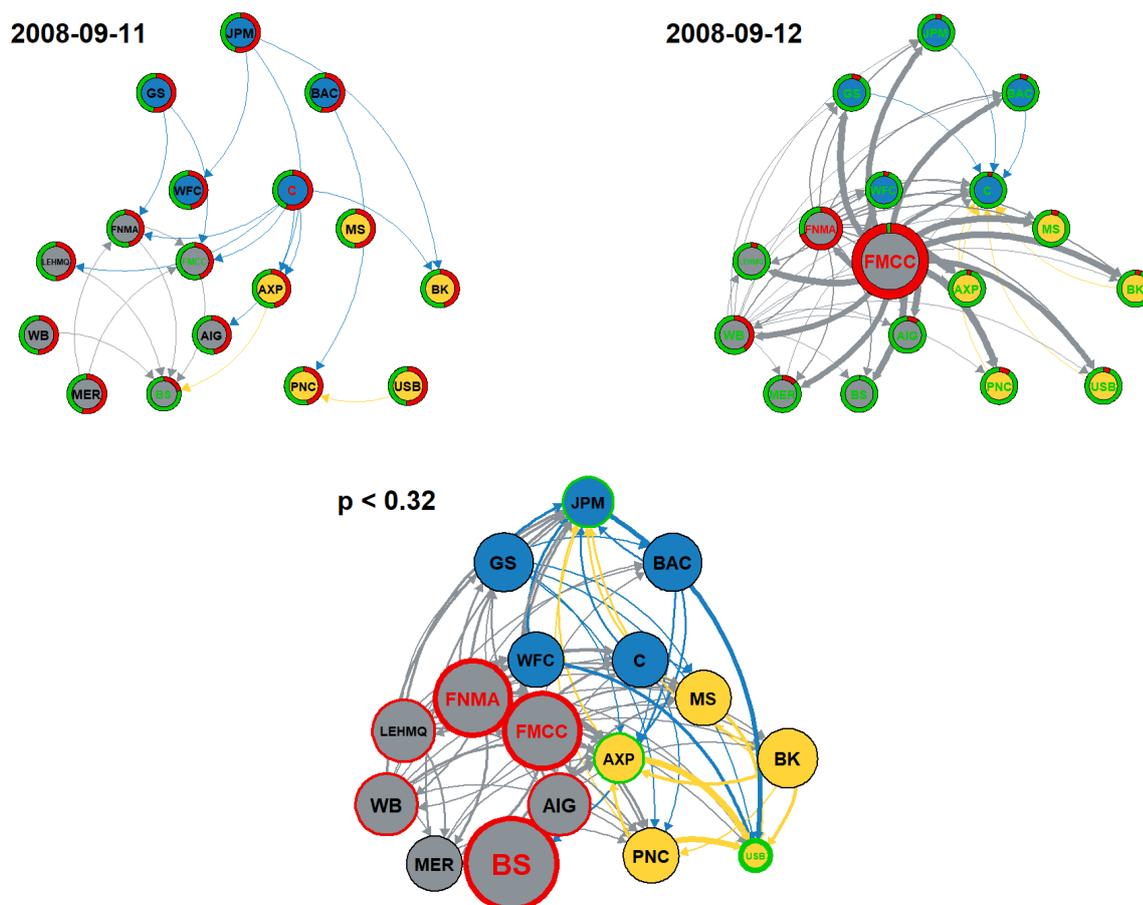**Figure B7: Estimated illiquidity network on 2008-09-05 and 2008-09-08**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 100$, adaptive joint LASSO-VAR(2) estimation.

Figure B8: Estimated illiquidity network on 2008-09-05 and 2008-09-08

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.

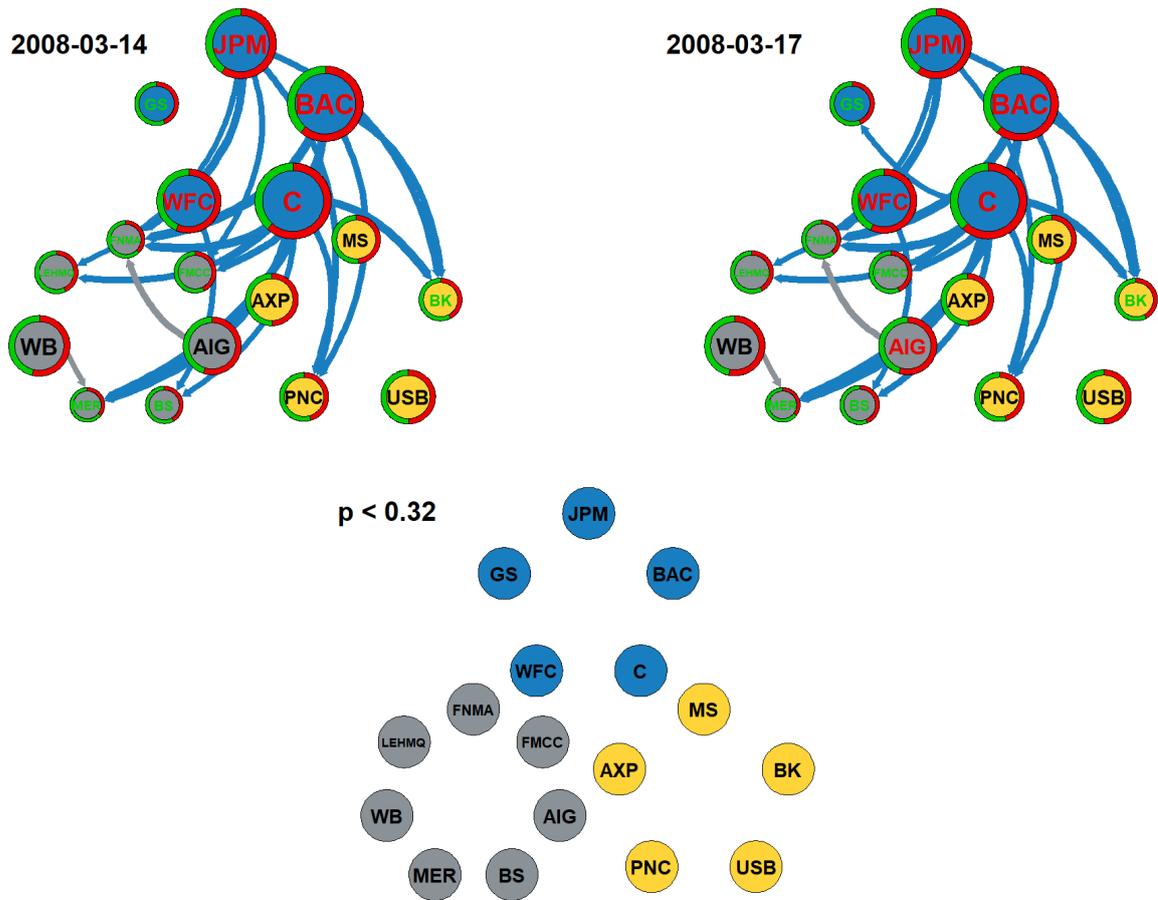**Figure B9: Estimated illiquidity network on 2008-09-11 and 2008-09-12**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.

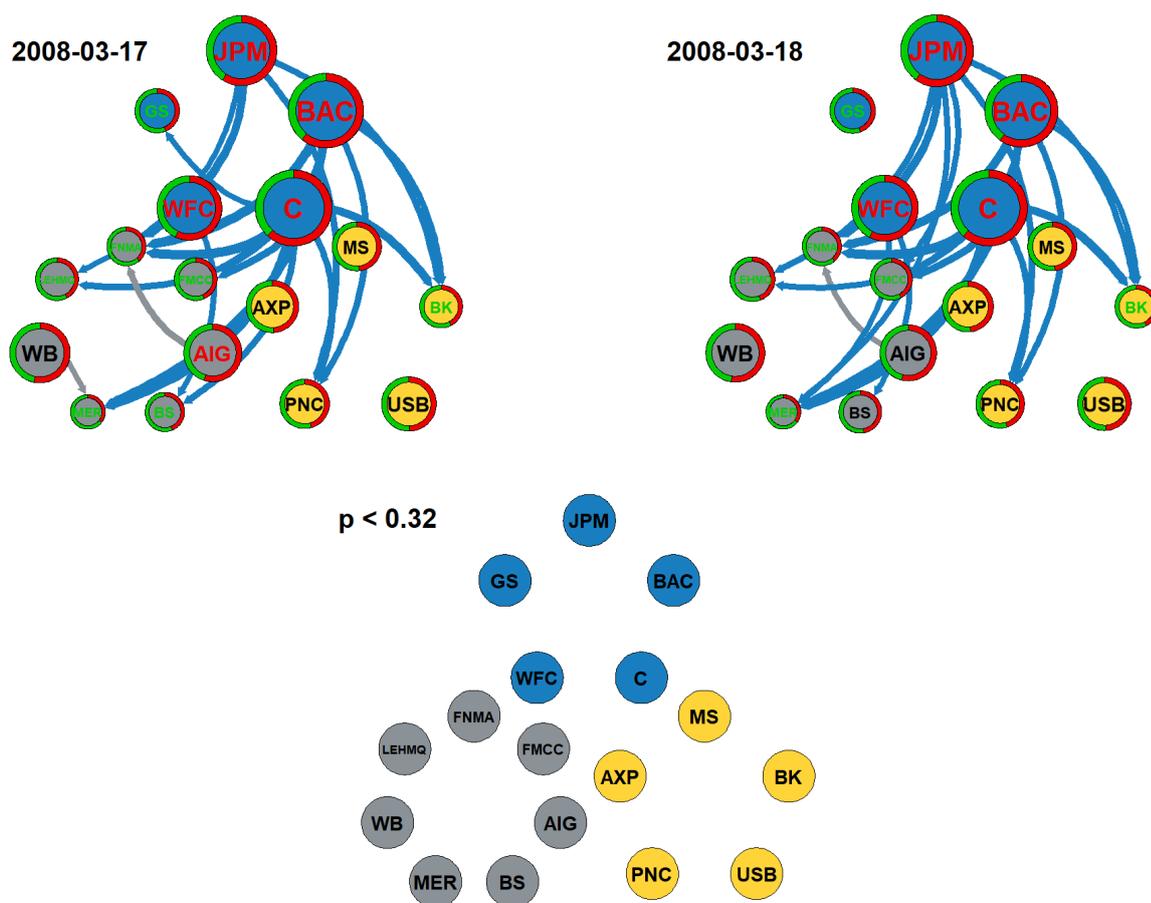**Figure B10: Estimated illiquidity network on 2008-09-11 and 2008-09-12**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 200$, adaptive joint LASSO-VAR(2) estimation.

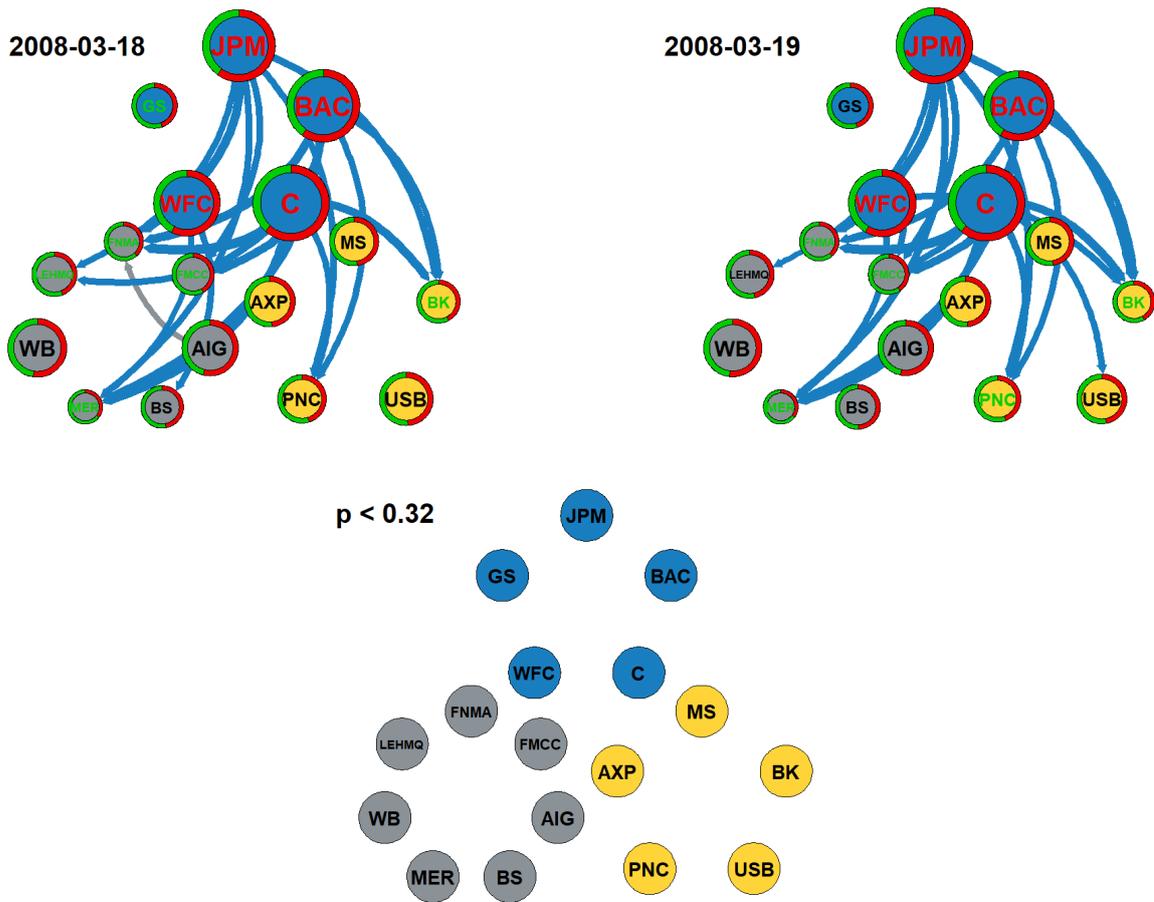**Figure B11: Estimated volatility network on 2008-03-14 and 2008-03-17**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.

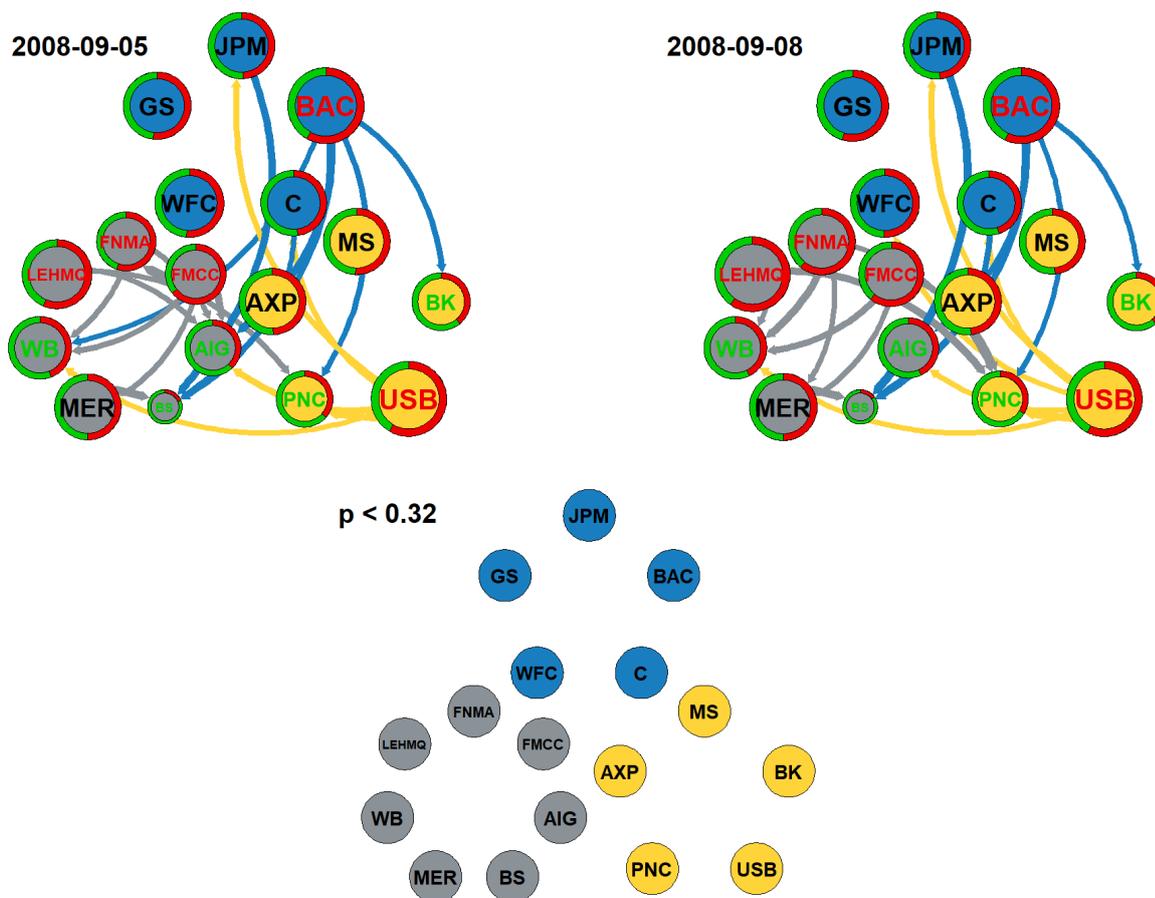Figure B12: Estimated volatility network on 2008-03-17 and 2008-03-18

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.

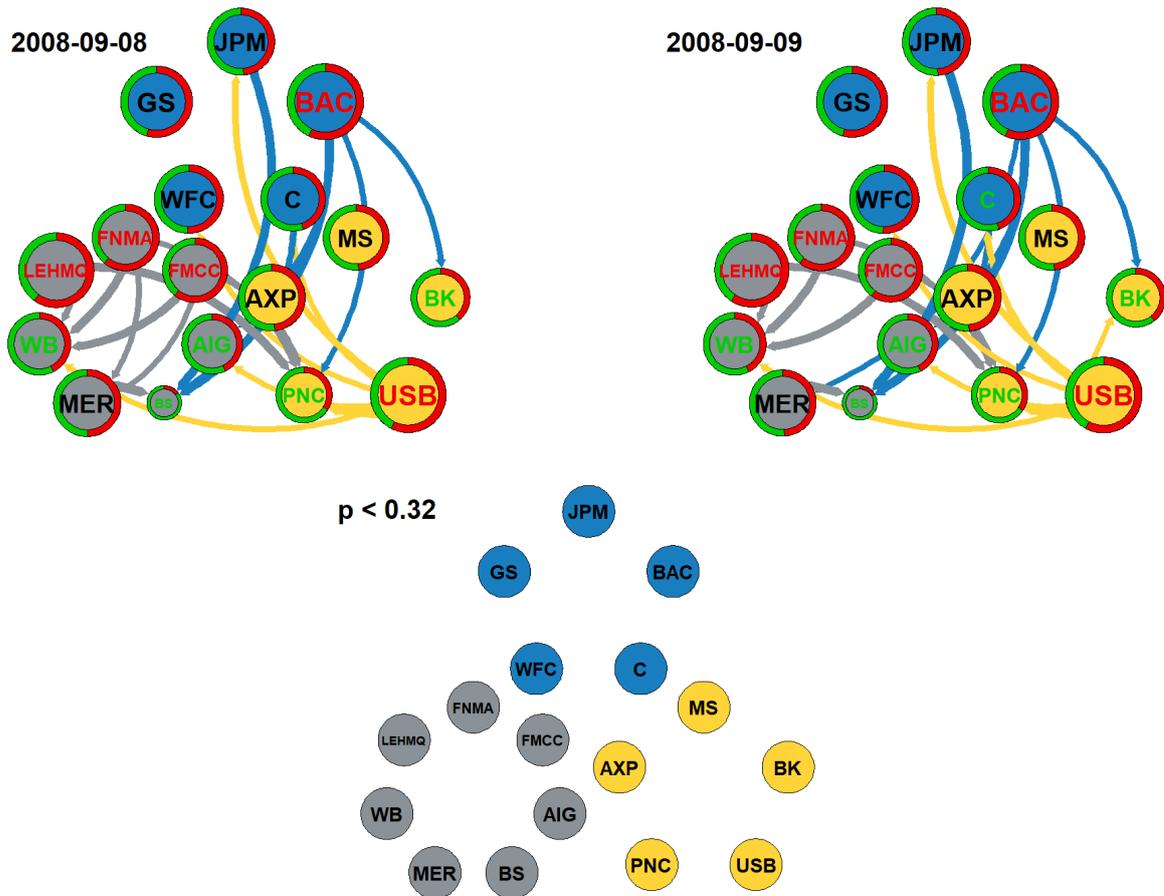**Figure B13: Estimated volatility network on 2008-03-18 and 2008-03-19**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.

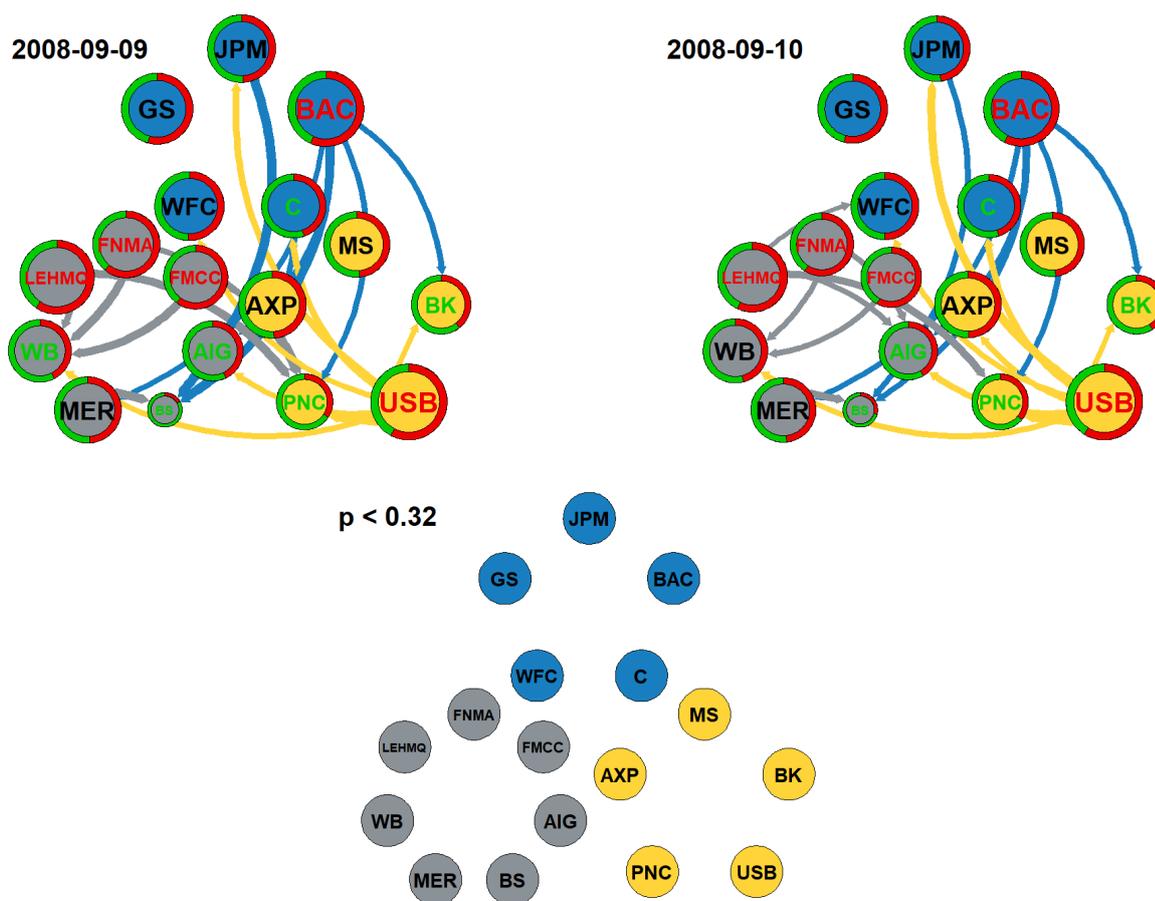**Figure B14: Estimated volatility network on 2008-09-05 and 2008-09-08**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

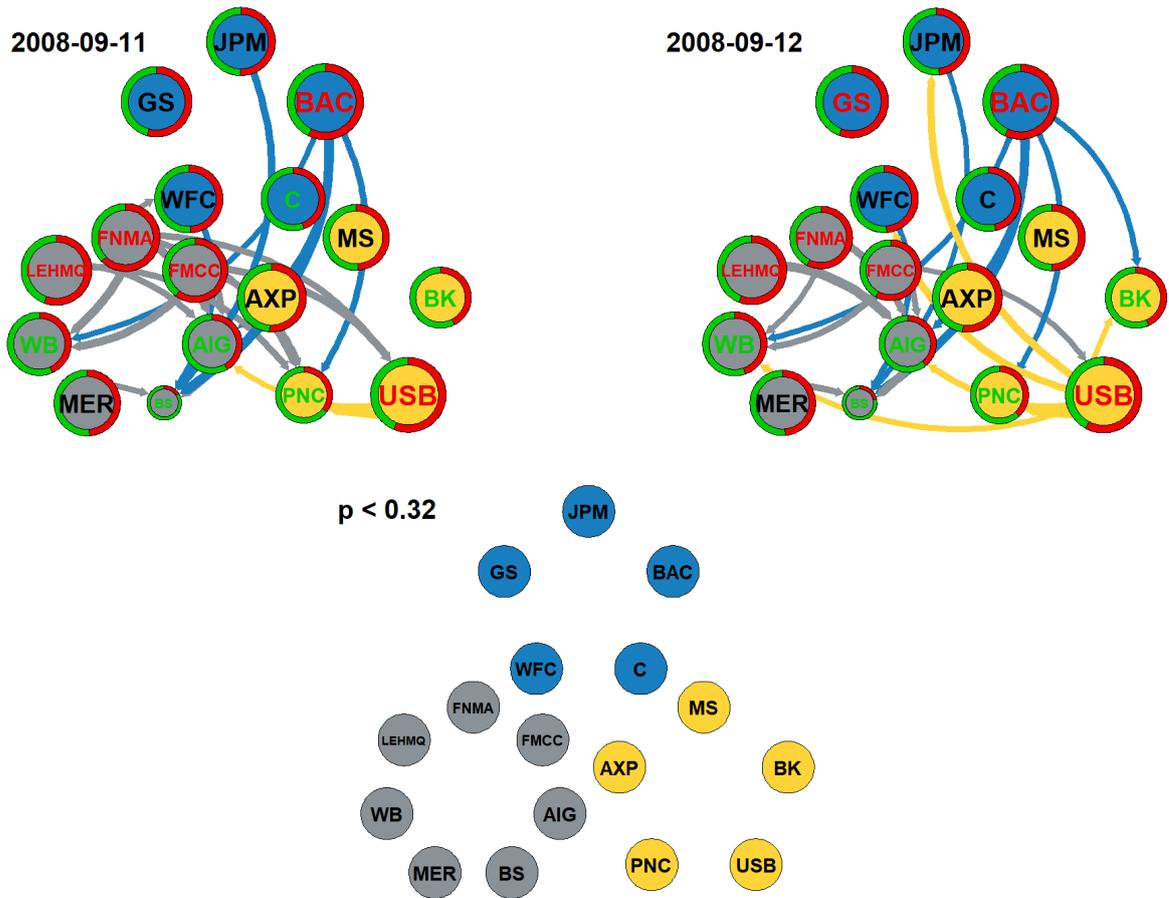**Figure B15: Estimated volatility network on 2008-09-08 and 2008-09-09**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

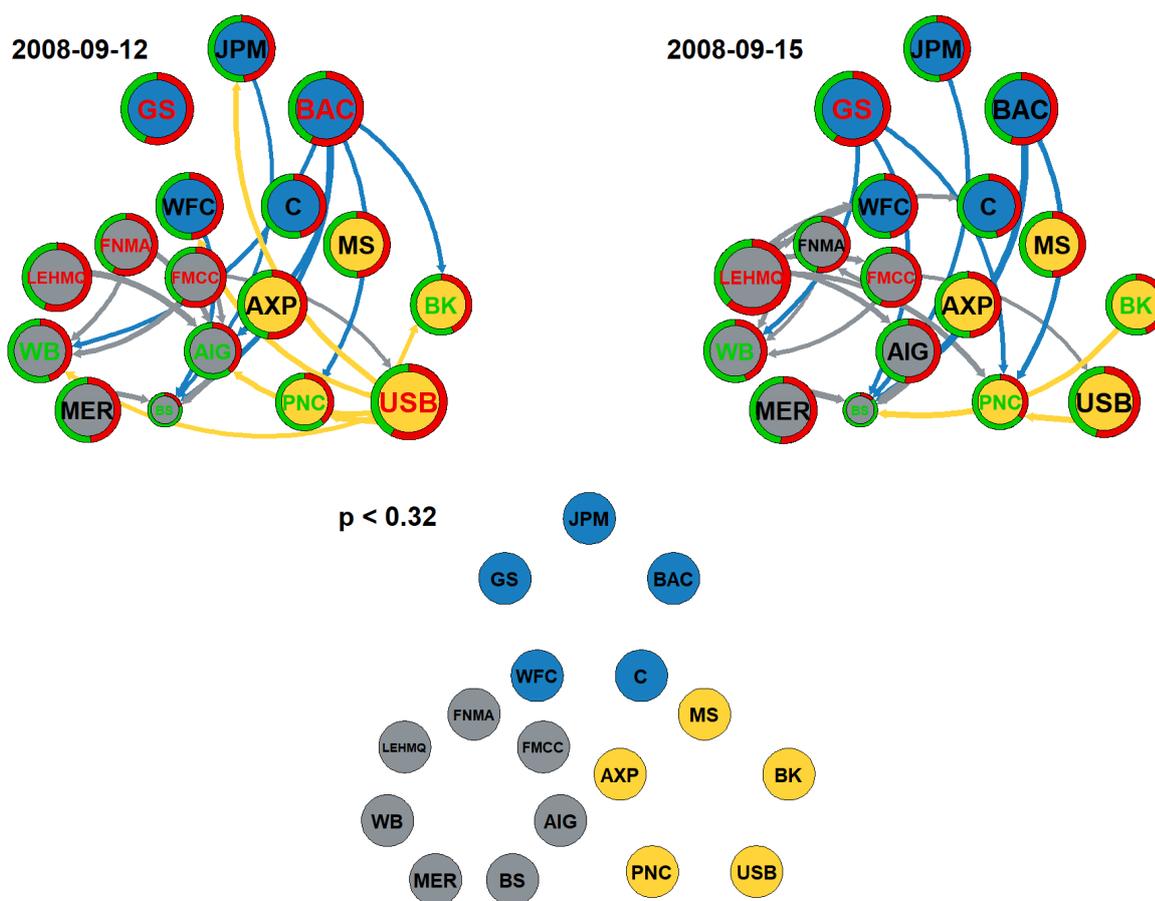# Figure B16: Estimated volatility network on 2008-09-09 and 2008-09-10

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

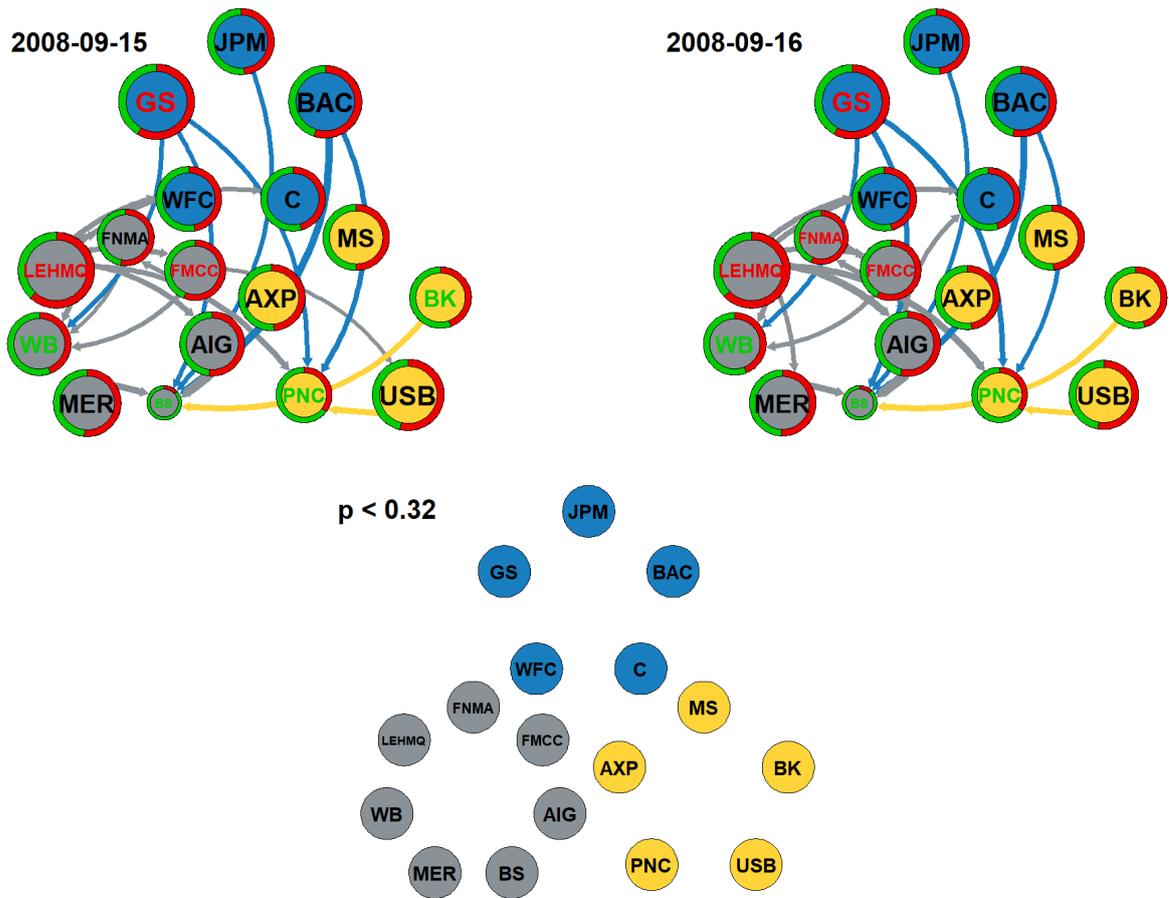**Figure B17: Estimated volatility network on 2008-09-11 and 2008-09-12**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

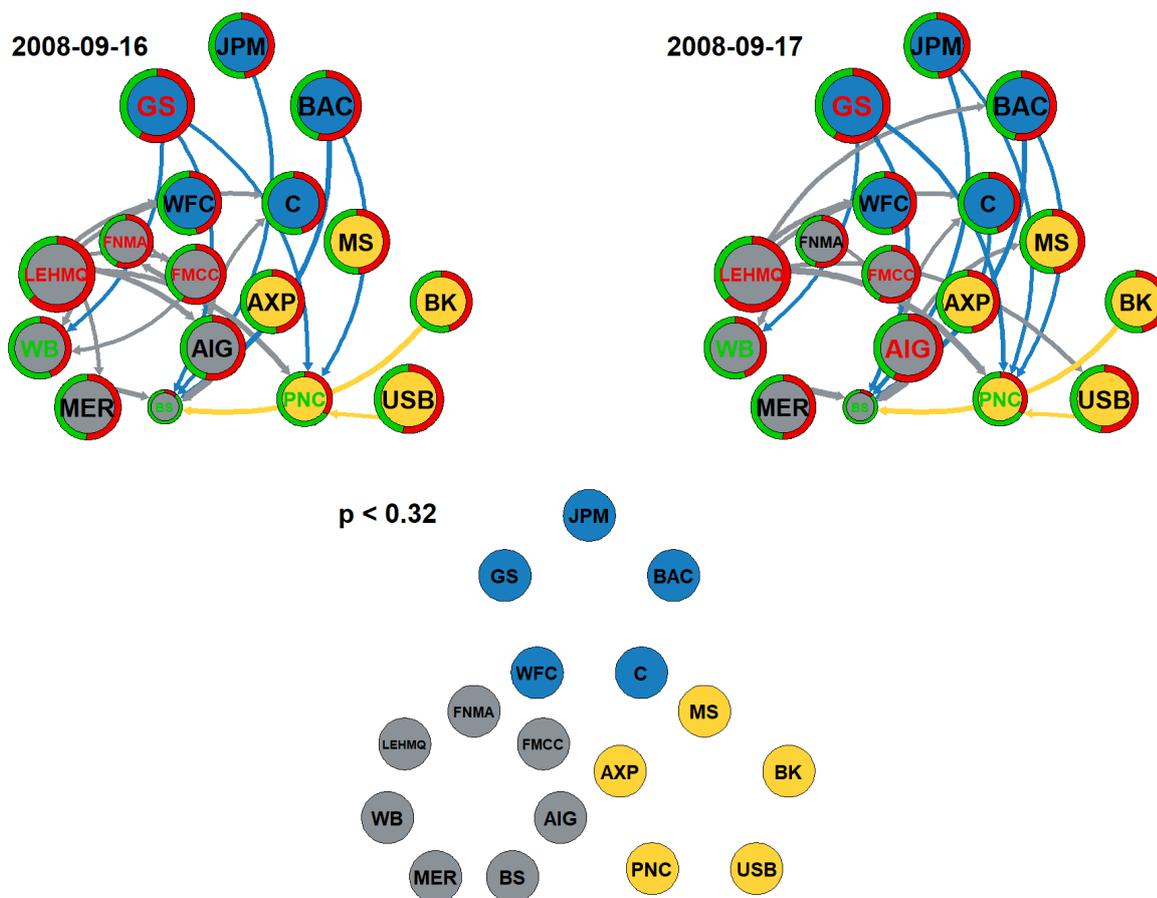**Figure B18: Estimated volatility network on 2008-09-12 and 2008-09-15**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

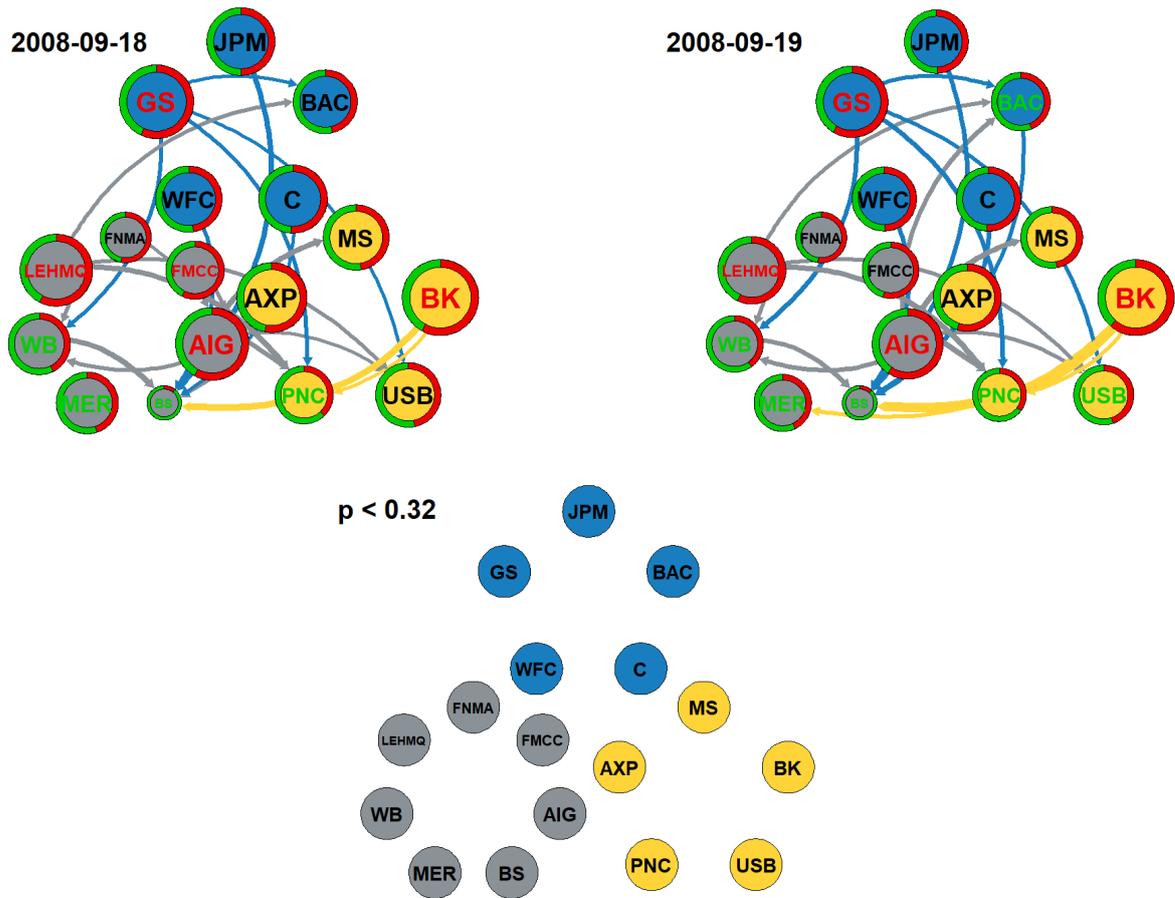**Figure B19: Estimated volatility network on 2008-09-15 and 2008-09-16**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

145

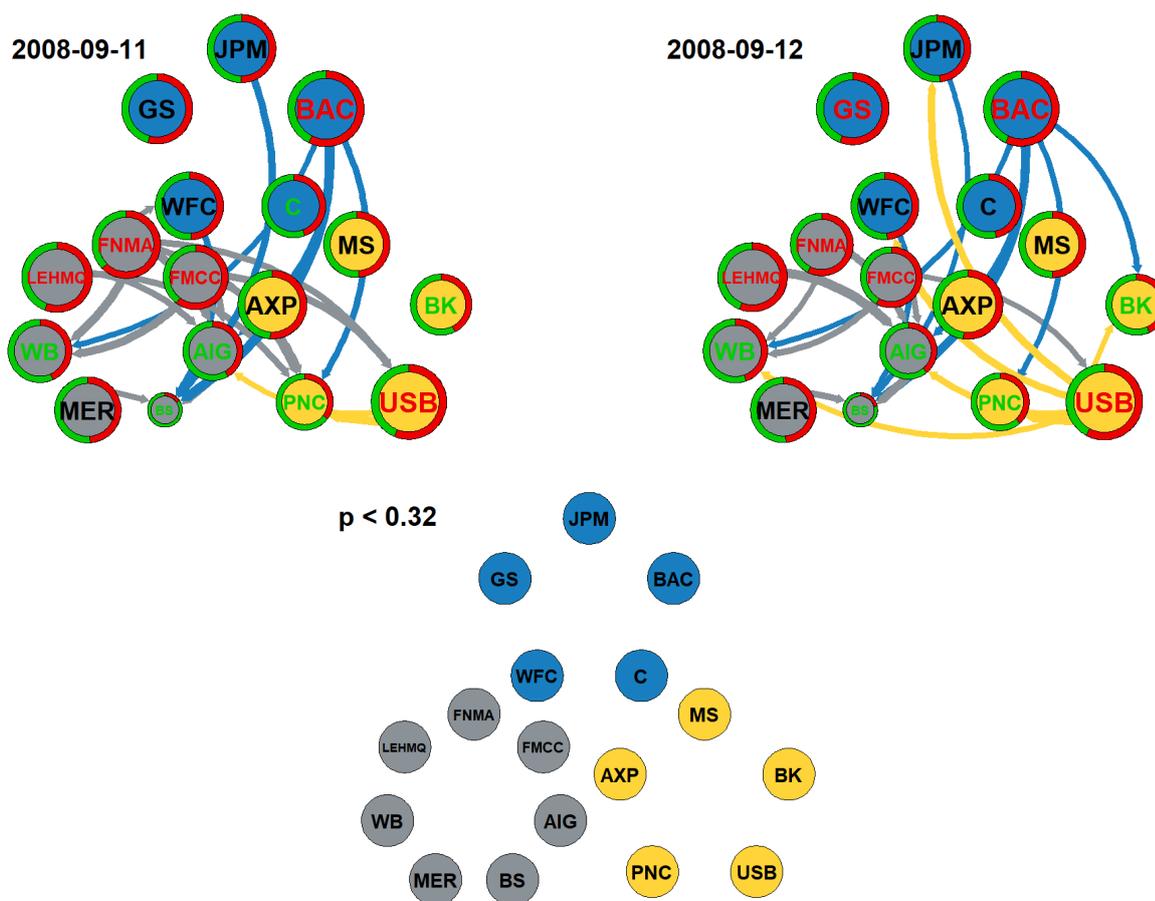**Figure B20: Estimated volatility network on 2008-09-16 and 2008-09-17**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 125$, adaptive joint LASSO-VAR(2) estimation.

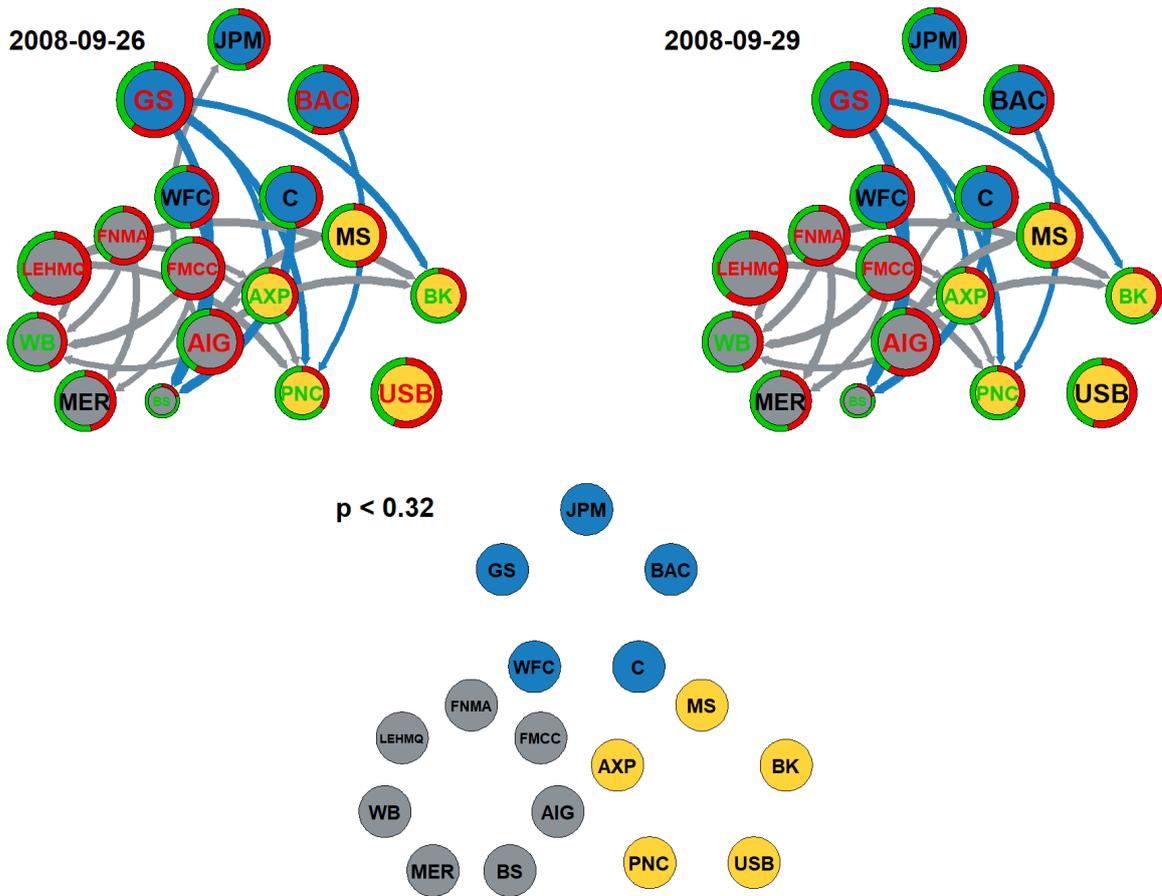**Figure B21: Estimated volatility network on 2008-09-18 and 2008-09-19**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 90$, adaptive joint LASSO-VAR(2) estimation.

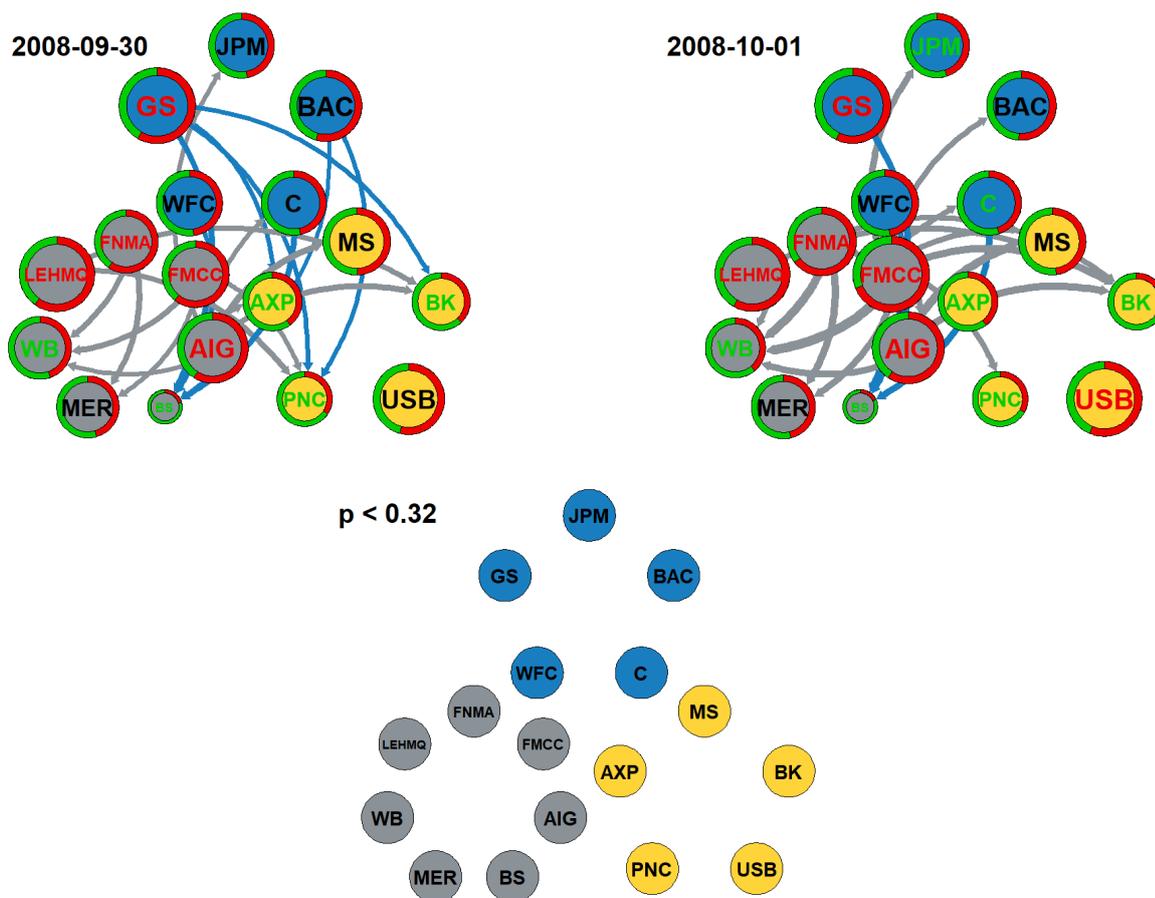**Figure B22: Estimated volatility network on 2008-09-19 and 2008-09-22**

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 90$, adaptive joint LASSO-VAR(2) estimation.

Figure B23: Estimated volatility network on 2008-09-26 and 2008-09-29

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. **Top panels:** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. **Bottom panel:** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. **Nodes:** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. **MBB-VAR:** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.

Figure B24: Estimated volatility network on 2008-09-30 and 2008-10-01

Notes: First day (second day) networks are in the upper left (right) chart; the bottom chart shows their difference at 32% significance levels. ***Top panels:*** Node sizes shows the SUM values. The green (red) slice of the pie charts represents the share of the FROM (TO) spillover indices of the FI. Edge directions mark the net spillover indices; their thickness represents the magnitudes. Only the top 15% of the edges are shown on the charts. ***Bottom panel:*** Green (red borders represent a significant decrease (increase) of the node's net spillover index. Smaller (Larger) sizes than the average size represent a significant decrease (increase) in SUM values. ***Nodes:*** Blue nodes represent FIs from the group "big," yellow represents FIs from the group "medium," and grey nodes are related to the "troubled" FIs. Nodes in the upper charts with green letters represent FIs where $FROM/SUM \geq 55\%$, reds where $TO/SUM \geq 55\%$. In the bottom chart, green letters represent FIs where the FROM/SUM ratio increased more than 10% and reds where TO/SUM increased more than 10%. ***MBB-VAR:*** Significance levels are estimated with moving-block bootstrap method (block size = 15, simulation run = 1000). The information of Diebold-Yilmaz network is calculated from a $T = 150$, adaptive joint LASSO-VAR(2) estimation.