# Multiple horizon causality in network analysis: measuring volatility interconnections in financial markets *

Jean-Marie Dufour †
McGill University

Bixi Jian ‡
McGill University

This version: November 14, 2016

* This work was supported by the William Dow Chair in Political Economy (McGill University), the Bank of Canada (Research Fellowship), the Toulouse School of Economics (Pierre-de-Fermat Chair of excellence), the Universitad Carlos III de Madrid (Banco Santander de Madrid Chair of excellence), a Guggenheim Fellowship, a Konrad-Adenauer Fellowship (Alexander-von-Humboldt Foundation, Germany), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds de recherche sur la société et la culture (Québec).

† William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 519, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 4400 ext. 09156; FAX: (1) 514 398 4800; e-mail: jean-marie.dufour@mcgill.ca . Web page: http://www.jeanmariedufour.com

‡ Department of Economics, McGill University, Mailing Address: Department of Economics, McGill University, Leacock Building, 855 Sherbrooke Street West, Montreal, Quebec H3A 2T7, Canada. TEL: (1) 514 998 9885. Email: bixi.jian@mail.mcgill.ca . Web page: www.bixijian.com

**ABSTRACT**

Existing literature cannot provide economic and financial networks with a unified measure to estimate network spillovers for empirical studies. In this paper, we propose a novel time series econometric method to measure high-dimensional directed and weighted market network structures. Direct and spillover effects at different horizons, between nodes and between groups, are measured in a unified framework. We infer causality effects in the network through a causality measure based on flexible VAR models specified by the LASSO approach. (Non-sparse) network structures can be estimated from a sparse set of model parameters. To summarize complex estimated network structures, we also proposed three connectedness measures that fully exploit the flexibility of our network measurement method. We apply our approach to investigate the daily implied volatility interconnections among the S&P 100 stocks over the period of 2000 - 2015 as well as its subperiods. We find that 7 out of the 10 most influential firms in the S&P 100 belong to the financial sector. Top investment banks (Morgan Stanley, Goldman Sachs and Bank of America) have the greatest influence in the financial sector. Market connectedness is especially strong during the recent global financial crisis, and this is mainly due to the high connectedness within the financial sector and the spillovers from the financial sector to other sectors.

**Key words**: Network; Multiple Horizon Causality Measures; LASSO; Financial Systemic Risk; Network Connectedness; Implied Volatility.

**Journal of Economic Literature classification:** C32; C52; C55; G15; G01 .

# Contents

# List of Tables



# List of Figures

# 1. Introduction

Since the financial crisis of 2007-09, academic researchers and financial regulators have a growing interest in investigating interconnections in financial markets. Network models have become increasingly popular to study economic interdependence by looking into the market architecture. Allen and Babus (2008) provide a survey showing a wide range of applications of network analysis in economics and finance. For example, bankruptcy contagion, volatility spillovers, risk propagation and amplification can all be studied in economic and financial network frameworks.[1] As Andersen, Bollerslev, Christoffersen and Diebold (2012) mention, modern network theory can provide a unified framework for systemic risk measures.

In macroeconomics, theoretical literature usually takes market structures as given, and then studies the roles of market architecture in the relationship between idiosyncratic risk and market-wide risk. In finance, economic links between firms may serve as the channel of gradual information diffusion. Individual firm's returns, return volatilities and credit spreads can be predicted via firms' linkages, while these empirical studies require identification of the underlying network structures, such as those from the Input-Output Surveys of the Bureau of Economic Analysis, the reported consumer-supplier relationships by public business enterprises or the international trade flows data from the International Monetary Fund (IMF) Direction of Trade Statistics.[2] In fact, many network structures are latent and not readily available in databases. For instance, the relationships between entities (e.g., detailed information on intra-bank asset and liability exposures) in a financial network are usually unknown. To empirically study a market network from financial data, we need an econometric measurement framework to identify and quantify the underlying network structure. A growing econometric literature is responding to this demand.[3] Perhaps surprisingly, however, very few of them are able to provide a satisfactory tool to measure high-dimensional market networks for general empirical purposes.

In this paper, we propose a novel network econometric measurement framework to better measure directed and weighted network structures using financial time series data in a high-dimensional context. Direct and spillover effects, between nodes and between groups, are measured in a unified framework. Causality at different horizons in the network is measured through a causality measure at different horizons. With this framework at hand, we provide estimated market networks with new econometric connectedness measures. The market systemic risk that is

[1] See Buraschi and Porchia (2012), Elliott, Golub and Jackson (2014), Acemoglu, Ozdaglar and Tahbaz-Salehi (2015*a*), Acemoglu, Akcigit and Kerr (2015) and Acemoglu, Ozdaglar and Tahbaz-Salehi (2015*b*) among others.

[2] See Cohen and Frazzini (2008),Hertzel, Li, Officer and Rodgers (2008), Menzly and Ozbas (2010), Aobdia, Caskey and Ozel (2014), Gençay, Signori, Xue, Yu and Zhang (2015), Albuquerque, Ramadorai and Watugala (2015) and Gençay, Yu and Zhang (2016) among others.

[3] See Billio, Getmansky, Lo and Pelizzon (2012), Hautsch, Schaumburg and Schienle (2014), Diebold and Yilmaz (2014), Demirer, Diebold, Liu and Yilmaz (2015), Bianchi, Billio and Casarin (2015), Barigozzi and Brownlees (2016) and Giudici and Spelta (2016) among others.

quantified by our connectedness measures has an intrinsic network foundation.

More concretely, we apply the short run and long run Granger causality measures[4] as the basic econometric framework to quantify the strengths of directed edges in a market network. We go beyond the simple Granger noncausality testing, i.e. whether an edge exists between two nodes, but explicitly measure the degree of the multiple horizon causality to obtain the strength of interconnections between two sets of nodes. Following Dufour and Taamouti (2010), we estimate the multiple horizon causality in the Vector Autoregressive model (VAR) settings. To overcome high-dimensionality problems in estimation, we use and extend the Least Absolute Shrinkage and Selection Operator (LASSO) techniques in the VAR estimations, which are similar to those developed by Barigozzi and Brownlees (2014) and Barigozzi and Brownlees (2016). Actually, (non-sparse) network structures, which are measured by our causality measures table, can be estimated from a sparse set of autoregressive coefficients and errors concentration matrices. Under mild conditions, we prove the asymptotic consistency of the estimators of our directed and weighted edge measures.

Our network measurement method has the following 7 appealing features:

1. The network edges we measure are directed. Allowing directed network structures provides us with important insights into the direction of network spillovers, since spillovers and relationships in economic and financial networks are generally asymmetric.

2. The network edges we measure are weighted. We do not merely identify the edges between two sets of nodes, but explicitly quantify their economic strengths.

3. In contrast to correlation-based measures, the directed edges we measure have causality implications. This is an important feature for theory verifications, model predictions and policy making.

4. Spillovers at different horizons in an economic network can be identified and measured by analyzing causality measures at different horizons. The multiple horizon causality measures gauge the net effects while simultaneously taking direct and indirect effects into account.

5. Our network measurement method overcomes the high-dimensionality problems in estimations. Note that economic and financial network theories usually study the cases in which the size (number of nodes) of a network is large or even goes to infinity (see, e.g., Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi (2012), Elliott et al. (2014) and Acemoglu, Ozdaglar and Tahbaz-Salehi (2015*b*)).

6. Our network measures provide underlying market network structures with clear graphical representations. Eichler (2007) shows that the multiple horizon causality in Dufour and Renault (1998), the base of the multiple horizon causality measures, is well matched to path

[4] See Dufour and Renault (1998), Dufour, Pelletier and Renault (2006) and Dufour and Taamouti (2010).

diagrams in the multivariate time series context. Thus our network measurement framework is also consistent with the network analysis in graph theory.

7. Point-wise edges, $(i \to j)$, as well as group-wise edges, $([i_1, i_2, ..., i_n] \to [j_1, j_2, ..., j_m])$, can be simultaneously analyzed by our unified network econometric framework. In empirical applications, for example, we can not only measure the relationship between firms, but also measure the relationship between sectors[5] by the same data observations at firm level and the same type of econometric measures.

We argue that a satisfactory econometric framework for studying market networks should at least satisfy Features 1 - 5: the network measurement method should be able to estimate directed and weighted network structures with causality implications, and it can be applied to study network spillover effects in a high-dimensional context. Feature 6 and Feature 7 are the extra advantages of our network measurement method. Moreover, Feature 7 provides us a new angle to study market network connectedness. It is intuitive to decompose market connectedness by the interconnections between different sectors and the connectedness within each sector. This decomposition is straightforward for economic and financial network analysis. However, the group-wise edges measurement method for measuring sectors' interconnections is missing in existing econometric literature. Our network measurement method can exactly fill this blank with our Feature 7.

Considering the economy of interest, which is modelled by a market network, as a $N$-dimensional Euclidean space, we use the causality measures table to provide the coordinates of each firm's location in the multi-dimensional economic space. The interconnectedness of a firm to the network can be characterized by the firm's location in the economic space. Total market connectedness is measured by the mean of the interconnectedness measures of each firm to the economic space. Similar to Billio et al. (2012) and Diebold and Yilmaz (2014), our market connectedness measures are built on underlying market network structures, and thus the market systemic risk quantified by these measures has a market network foundation. Since an economic network can be viewed as a network connected by firms (firm-wise market), whose interconnections are measured by our point-wise edges method $(i \to j)$, or a network connected by sectors (sector-wise market), whose interconnections are measured by our group-wise edges method $([i_1, i_2, ..., i_n] \to [j_1, j_2, ..., j_m])$, we have three types of connectedness measures to gauge network interconnections: i) firm-wise connectedness, which measures the interconnectedness of a firm-wise market; ii) firm-wise connectedness within a sector, which measures the interconnectedness within a given sector in a firm-wise market; and iii) sector-wise connectedness, which measures the interconnectedness of a sector-wise market. These three types of connectedness measures fully take advantage of the flexibility of our network measurement method, so they can be applied to study market network connectedness in more flexible ways than those connectedness measures

[5] A sector can be viewed as a group of firms.

proposed by Billio et al. (2012) and Diebold and Yilmaz (2014).

Our network measurement methods have a wide range of applications and can be applied in a variety of research areas, including identifying and quantifying economic relationships between firms, between sectors and between areas; measuring market connectedness; predicting financial risks; guiding asset allocations in large portfolios; etc. Note that many latent economic and financial network structures can be estimated by our flexible network measurement method with varieties of panel databases, and observing that explicit identified economic network centrality and consumer-supplier linkage have been shown to be new risk factors in asset pricing and new determinants to predict financial variables, e.g., stock return, return volatility, and credit spread[6], we expect more pricing factors and financial and macroeconomic variables drivers are to be discovered by network econometric measurement methods.[7]

To illustrate the usefulness of our method in network analysis, we investigate the S&P 100 implied volatility network in the US stock market. Volatility network in financial markets has been studied in Diebold and Yilmaz (2014), Demirer et al. (2015) and Barigozzi and Brownlees (2016), but they mainly focus on realized volatility. For financial practitioners, the VIX index, calculated from the implied volatilities of S&P 500 index option contracts, is the most popular volatility measure to gauge market turbulences, and it is also known as a "market fear" index. Our implied volatility network among the S&P 100 stocks[8] can thus be naturally viewed as an "individual fear" network. To the best of our knowledge, implied volatility network has not yet been studied in the financial literature.

We first look at the static network with the full sample (2000 - 2015). We identify the most influential firms in the firm-wise market network, the most influential firms in the financial sector, and the most influential sectors in the sector-wise market network. Using rolling subsamples, we estimate the time-varying firm-wise market connectedness before, during and after the recent financial crisis of 2007-09, and compare it with the dynamic patterns of the firm-wise connectedness within each sector and the sector-wise connectedness among different sectors. In particular, we also examine the dynamic interconnections between the financial sector and other sectors.

We find that: i) 7 out of the 10 most influential firms in the S&P 100 belong to the financial sector, and top investment banks (Morgan Stanley, Goldman Sachs and Bank of America) have the greatest influence in the financial sector; ii) market connectedness was especially strong during the recent global financial crisis; iii) the high market connectedness was mainly due to the high connectedness within the financial sector and the spillovers from the financial sector to other sectors; iv) the financial sector had the highest firm-wise connectedness from 2008 to 2010, while the

[6]See Cohen and Frazzini (2008), Hertzel et al. (2008), Menzly and Ozbas (2010), Ahern (2013), Aobdia et al. (2014), Gençay et al. (2015) and Gençay et al. (2016).

[7]For example, Jian (2016) uses a Granger-type method to identify the illiqiudity network in stock markets and finds centralities in illiquidity networks are priced in the cross-section of expected returns.

[8]To be included in the S&P 100, the companies should be among the larger and more stable companies in the S&P 500, and *must have list options*.

connectedness of other sectors also reach relatively high level during this period; v) the causality effects between the financial sector and other sectors were asymmetric and displayed considerable variation over time, which stresses the importance of directed and weighted edges settings in market network analysis.

This paper is motivated by the econometric literatures on the analysis of financial networks and contributes to different strands of literature. The topic of this paper is related to recent econometric literature on financial networks (see Billio et al. (2012), Diebold and Yilmaz (2014), Demirer et al. (2015), Bianchi et al. (2015), Barigozzi and Brownlees (2016), Hautsch et al. (2014), Ahelegbey, Billio and Casarin (2016), and Giudici and Spelta (2016) among others). We differ from the social network econometrics literature, e.g., Bramoulle, Djebbari and Fortin (2009), in the sense that the nodes in our network setting are represented by time series financial variables (e.g., return and volatility). The most closely related econometric literature to this paper includes: Billio et al. (2012), Diebold and Yilmaz (2014), Demirer et al. (2015) and Barigozzi and Brownlees (2016). Billio et al. (2012) detect the edge of a pair of nodes via testing bilateral Granger noncausality without taking into account other nodes in the network, and thus may find misleading "spurious" causality edges and tend to overestimate the number of linkages. Diebold and Yilmaz (2014) and Demirer et al. (2015) overcome the spurious relation problem. They measure the directed and weighted network structure by generalized forecast error variance decompositions in a VAR representation. The generalized forecast error variance decomposition technique is closely related to our multiple horizon causality measures. Unfortunately, Diebold and Yilmaz (2014) neglect the high-dimensionality problem in their study, Demirer et al. (2015) fail to provide the theoretical validity for their estimations and they both require the joint Gaussian innovation assumption in the econometrics model. These drawbacks inevitably limit their applications in market network analysis for general purposes. The time series network estimation settings in Barigozzi and Brownlees (2016) are similar to what we apply in this paper. Yet, their network structure is assumed to be sparse and their edges, measured by long run partial correlations, are basically undirected. Among recent literature[9], only the empirical model proposed in Demirer et al. (2015) is able to study a high-dimensional directed and weighted network structure, and none of them is able to estimate point-wise edges and group-wise edges in a unified framework.

We apply the short run and long run Granger causality measures as our basic network econometric measurement framework. The concept of the noncausality testing introduced by Granger (1969) and Sims (1972) has been widely used to study dynamic relationships between time series in economics and finance. Dufour and Renault (1998) and Dufour et al. (2006) extend this notion to multiple horizon cases to study indirect causality effects. Eichler (2007) connects the short run and long run Granger causality with path diagram in multivariate time series analysis. Based on Geweke (1982), Dufour and Renault (1998) and Dufour et al. (2006), Dufour and Taamouti

[9]Ahelegbey (2015) provides a recent review on the network econometrics in the context of time series analysis.

(2010) propose the multiple causality measures to quantify the causality at any forecast horizon $h \geq 1$. Dufour, Garcia and Taamouti (2012) apply this tool in studying the relationship among returns, realized volatility and implied volatility. Dufour and Zhang (2015) further study the multiple horizons second-order causality. In this paper, we show that market networks, with directed and weighted edges, can be modelled and measured by the well-developed econometrics framework of the multiple horizon causality measures. Moreover, unlike Dufour and Taamouti (2010) and Dufour et al. (2012) who only deal with low-dimensional situations, we estimate the multiple horizon causality measures with the LASSO approach to better fit the multiple horizon causality measure framework into high-dimensional network analysis.

One of the motivations of this paper, identifying and quantifying the degree of interconnections between nodes and between groups in market networks, is to provide a new way to measure market-based systemic risk. Similar to Billio et al. (2012) and Diebold and Yilmaz (2014), our market connectedness measures are also built upon the underlying network structure and contribute to the strand of literature on market-based systemic risk measurement (see Acharya, Pedersen, Philippon and Richardson (2010), Brownlees and Engle (2015), Adrian and Brunnermeier (2011), Billio et al. (2012), Diebold and Yilmaz (2014), Hautsch et al. (2014) and Demirer et al. (2015) among others). Benoit, Colliard, Hurlin and Perignon (2015) provide a comprehensive survey on measurement methods for systemic risk.

Our key contribution is that we propose a novel time series econometrics network measurement framework, which can be applied to measure high-dimensional directed and weighted market network structures, without sparsity assumptions on network structures or the Gaussian assumption on econometric models. We successfully connect the causality literature with the LASSO approach in application to network measurement. Moreover, to the best of our knowledge, our econometric framework is the first one in the network econometric literature to explicitly allow point-wise edges and group-wise edges to be measured in a unified framework.

The rest of this paper is organized as follows. In section 2, we provide a brief description of general directed and weighted network structures and discuss the criteria of a satisfactory network econometric framework in economic and financial network analysis. In section 3, we show that directed and weighted network structures and network spillovers can be measured by the multiple horizon causality measures table. In section 4, we estimate the causality table with the LASSO approach in a high-dimensional context and provide asymptotic consistency results. In section 5, we propose new market network connectedness measures for systemic measurement. In section 6, we investigate the static structure and the time-varying characteristics of the implied volatility network in the US stock market. Finally, in section 7 we provide a short conclusion.

## 2. General economic and financial network

A network is composed by two basic elements: nodes and edges. Financial institutions, for instance, represented by different nodes, are linked through networks of different types of financial contracts, such as derivatives, credits and securities. These contracts or business relationships, between any pair of financial institutions, are represented by their edges in the financial network. While nodes are given and known as they are always referred to some specific institutions, modelling edges is always an elusive part in financial network analysis. Edges represent some implicit economic relationships between nodes. The relationship among financial institutions in many cases are unknown or difficult to be specified. When we study the systemic risk in a financial network, edges could be the position of banks' loans to each other in their balance sheets, or whether they hold a large bilateral position of some securities (e.g., credit default swap (CDS)). Without a prior specific definition of the systemic risk, which financial contract should be selected as the edge to study a financial network is a difficult decision to make, since loan's edges and CDS's edges are both theoretically important but their existences can be independent. Moreover, detail information of the financial contracts that financial institutions are holding and their counterparties is usually unavailable to public. Therefore, what we can measure for the edges from data is at most a proxy of what we are interested in. This provides a board space for econometricians to develop different statistical network measures for different research objectives. One of the main aspects of research papers differing from each other in the financial econometrics network literature is in their rationales of how to construct a statistical measure to quantify the edges in a network.

Despite of it, all networks have basic structures in common. A simple static network has a mathematical notation: $G = \{V,E\}$, where $V = \{1,2,...,N\}$ is the set of nodes and $E = \{e^{ij} : (i, j) \in V \times V\}$ is the set of edges. Usually, the size of the network, $N$, is large. Any pair of nodes in $V$, $(i, j)$, may be linked by an edge in the edge set, $E$, with certain degree of strength $e^{ij}$. When $e^{ij} = e^{ji}$ is assumed, the network is undirected; otherwise, the network is directed. If $e^{ij}$ is assumed to be indexed by $\{0,1\}$, the network is unweighted; if $e^{ij}$ is continuous, the network is weighted.

The directed edges setting is crucial in economic and financial network analysis. Economic relations are usually directed and the directed structures play an important role in network analysis. For instance, the presence of directed intersectoral input-output linkages can explain why single idiosyncratic shocks may lead to market-wide aggregate fluctuations (see Acemoglu et al. (2012)). Economic effects and information flows have directions. We use causal relationships to describe such directed relationships in a economic network. Causality interpretations are required for economic networks because it is the foundation for theory verifications, model predictions and policy makings. Intuitively, if two firms have no business relationship, we do not expect there is a causal relationship between them and vice versa. We notate the directed edges in our network with arrows, $(i \to j)$, which indicates $i$ causes $j$. Figure 1 shows four simple possible relations between node A and node B in a unweighted setting. If the strength of the edges $e^{AB} = e^{BA} = 0$, we say

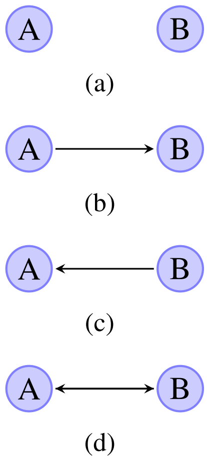

Figure 1. Directions between node A and node B

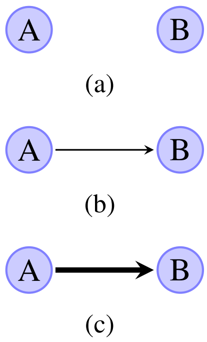

Figure 2. Strengths of the edges from node A to node B

node A and node B are unlinked (fig. 1(a)); if they are linked, then either $e^{AB} = 1$ or $e^{BA} = 1$ or $e^{AB} = e^{BA} = 1$ (fig. 1(b), fig. 1(c) and fig. 1(d)).

The weighted edges setting is also important. Effects in an economic network are weighted. In social networks, knowing how well agents know each others is much more informative than merely knowing whether they know each others, since the probability of information transmissions is highly correlated with their familiarity. In financial networks, when we say a bank is “too big to fail”, it implies that this bank has “big” impacts on others. When studying shock propagations or risk amplifications in a market network, we would be especially interested in quantifying spillover effects. Since spillovers may grow (or disappear) through edges in a network, unweighted edges setting is not able to model the quantitative change in spillover processes. Figure 2 shows three possible strengths of edge from node A to node B. The strength of the edge could be zero, which implies there is no relation from node A to node B (see fig. 2(a)). The strength of the edge could be small and it is represented by a light arrow (see fig. 2(b)); the strength of the edge could also be large and it is represented by a thick arrow (see fig. 2(c)). The thickness of the edge ($i \rightarrow j$) is weighted scaled by $e^{ij}$.

Economic and financial network literature usually reports some graphs of the network they study. The graph representation of a static network does provide us a broad and concise picture of the underlying network structure. A static network, however, only tells us direct effects. The indirect effects, a central part in risk spillover analysis, is nontrivial to be revealed from the direct effects. For instance, suppose there are relations from node A to node B indirectly via two different paths in an unweighted static network, we may naively say that the risk from node A could cascade to node B. However, it is possible that node A has no effect on node B if the indirect effects in those two paths are just cancelled out by each other. Hence, a network graph drew from a static network structure may mislead us to a wrong implication about spillover effects in the true economic network. Surprisingly, econometric literature on financial networks have not yet realized this important issue. Most of them are just focus on estimating static network structures without directly measure spillover effects.

In summary, the size of an economic network is usually large; nodes' relationships in an economic network should have causality interpretations; a directed and weighted edges setting is required to uncover the effects in the underlying economic network structures; network spillover effects need to be measured directly. Therefore, a satisfactory network econometric framework should be able to estimate directed and weighted network structures with causality implications, and it can be applied to study spillover effects in a high-dimensional context.

# 3. Multiple horizon causality and networks

In this section, we model a complex network structure by causality relations, and apply the short run and long run Granger causality measures, introduced by Dufour and Taamouti (2010), to identify and quantify the edges between two sets of nodes in the underlying network structure. We demonstrate that the multiple horizon causality measures satisfies the criteria of a satisfactory network econometric framework. It is able to estimate directed and weighted network structures with causality implications and can be applied to study spillover effects in a high-dimensional context Moreover, our network measurement framework has some other important features.

Suppose we observe a data sample from a jointly strictly stationary process $X = \{X_{1t}, X_{2t}, ..., X_{Nt}\}_{t=1}^{T}$. $N$ is the number of nodes and $T$ is the observable sample size. In context of economic network analysis, the number of nodes, $N$, is large. The process of interest, $X$, can be divided by three sub-processes as $X = \{X_t^W, X_t^Y, X_t^Z\}_{t=1}^{T}$, such that $X_t^W = [X_{1t}, ..., X_{m_1 t}]$, $X_t^Y = [X_{(m_1+1)t}, ..., X_{(m_1+m_2)t}]$ and $X_t^Z = [X_{(m_1+m_2+1)t}, ..., X_{(m_1+m_2+m_3)t}]$, where $m_1, m_2, m_3 \geq 0$ and $m_1 + m_2 + m3 = N$.

**Definition 3.1** MEAN-SQUARE CAUSALITY MEASURE AT FORECAST HORIZON $h$ RELATIVE TO AN INFORMATION SET $I$.

*For* $h \geq 1$, *where by convention* $\ln(0/0) = 0$ *and* $\ln(x/0) = +\infty$ *for* $x > 0$,

$$C_L(X^W \underset{h}{\longrightarrow} X^Y|I) := \ln\left[\frac{det\{\Sigma[X^Y(t+h)|I_{(-W)}(t)]\}}{det\{\Sigma[X^Y(t+h)|I(t)]\}}\right] \tag{3.1}$$

*is the mean-square causality measure from $X^W$ to $X^Y$ ($Y = W$ is allowed) at horizon $h$, given information set $I$. Since we only consider the mean-square measures in this paper, we will just call it as short run and long run Granger causality measures or multiple horizon causality measures hereafter. $I$ denotes the full information set and $I_{-W}$ denotes the full information set without the information generated by $X^W$. If we further assume the full information set is generated only by $X$ itself, $I$ and $I_{-W}$ can be denoted by $I_{WYZ}$ and $I_{YZ}$ respectively, where $I_{WYZ}(t)$ denotes the information set generated by the process $X = \{X^W, X^Y, X^Z\}$ up to time $t$, and $I_{YZ}(t)$ denotes the information set generated by the sub-process $\{X^Y, X^Z\}$ up to time $t$.*

The multiple horizon causality measure, $C_L(X^W \underset{h}{\longrightarrow} X^Y|I)$, gauges the predictive power of $X^W$ to $X^Y$ conditional on $I$. We say $X^W$ causes $X^Y$ at forecast horizon $h$ if and only if $X^W$ helps to predict $X^Y$ at forecast horizon $h$. The value of $C_L(X^W \underset{h}{\longrightarrow} X^Y|I)$ measures the degree of the causal effect from $X^W$ to $X^Y$ at forecast horizon $h$. Consequently, the identified edge, $(W \to Y)$, has causality implications.

There are some important properties of this type of measures.

First, generally speaking, $C_L(X^W \underset{h}{\longrightarrow} X^Y|I) \neq C_L(X^Y \underset{h}{\longrightarrow} X^W|I)$. The effect from $W$ to $Y$ is not presumed to be equal to the effect from $Y$ to $W$. The edges between $W$ and $Y$, $(W \to Y)$ and $(Y \to W)$, are directed.

Second, $C_L(X^W \underset{h}{\longrightarrow} X^Y|I)$ is always nonnegative as $I_{(-W)}(t) \subseteq I(t)$. $C_L(X^W \underset{h}{\longrightarrow} X^Y|I) = 0$ if and only if there is no causal effect from $X^W$ to $X^Y$ at forecast horizon $h$. The value of $C_L(X^W \underset{h}{\longrightarrow} X^Y|I)$ is increasingly monotone to the the predictive power of $X^W$ to $X^Y$. Thus, the strength of the edge, $(W \to Y)$, measured by the value of $C_L(X^W \underset{h}{\longrightarrow} X^Y|I)$, is weighted.

Third, $C_L(X^W \underset{h}{\longrightarrow} X^Y|I)$ measures the indirect effect from $W$ to $Y$ at horizon $h$, while $C_L(X^W \underset{1}{\longrightarrow} X^Y|I)$ measures the direct effect as there is only one step to be considered. For example, suppose $C_L(X^W \underset{1}{\longrightarrow} X^Y|I) = 0$ and $C_L(X^W \underset{h}{\longrightarrow} X^Y|I) > 0$ for a $h > 1$, it implies there is no direct effect from $W$ to $Y$, but there is an indirect effect from $W$ to $Y$ via other node(s) in the network. The spillover effect from $W$ to $Y$ at any step $h$ can thus be directly measured by $C_L(X^W \underset{h}{\longrightarrow} X^Y|I)$.

Fourth, the dimensions of $X^W$ and $X^Y$ are arbitrary. To measure the edge from $W$ to $Y$, we only require the dimensions of the processes $X^W$ and $X^Y$, $m_1$ and $m_2$, such that $m_1, m_2 \geq 1$ and $m_1 + m_2 \leq N$. We can let $W$ and $Y$ represent a single node or a set of nodes. The point-wise edges, where $X^W$ and $X^Y$ are univariate variables ($m_1 = m_2 = 1$), and the group-wise edges, where $X^W$ and $X^Y$ are multivariate variables ($m_1, m_2 > 1$), can be analysed in this unified econometric framework.

For instance, we can not only measure the edges between firms (firm-wise edges), where $X^W$ and $X^Y$ represent firms, but also measure the edges between sectors (sector-wise edges), where $X^W$ and $X^Y$ represent sectors and $m_1$ and $m_2$ are the number of firms in the sectors. Therefore, we can use the same data observations at firm's level and the same type of econometric measures defined in Definition 3.1 to study firm-wise edges and sector-wise edges in a unified framework. In the past, weighted aggregation is usually required if we want to study the sector-wise spillover effect with firm-wise data. However, it would inevitably come to a cost of losing information in firm-wise interconnections. The econometrics approach proposed in this paper overcomes this limitation.

**Remark 3.1** If $X^W$ and $X^Y$ are univariate processes denoted by $X_i$ and $X_j$ respectively, then for $h \geq 1$

$$C_L(X_i \underset{h}{\longrightarrow} X_j|I) := \ln\left[\frac{\sigma^2[X_j(t+h)|I_{(-i)}(t)]}{\sigma^2[X_j(t+h)|I(t)]}\right]. \tag{3.2}$$

The variances of the forecast errors of $X_j(t+h)$, $\sigma^2[X_j(t+h)|I_{(-i)}(t)]$ and $\sigma^2[X_j(t+h)|I(t)]$, are both positive, and $\sigma^2[X_j(t+h)|I_{(-i)}(t)] \geq \sigma^2[X_j(t+h)|I(t)]$. $\sigma^2[X_j(t+h)|I_{(-i)}(t)] = \sigma^2[X_j(t+h)|I(t)]$ if the information generated by node $i$ does not help to decrease the forecast error variance of node $j$. $C_L(X_i \underset{h}{\longrightarrow} X_j|I)$ measures the causality strength from node $i$ to node $j$. For notation convenience, we hereafter let $C_{ij}^h := C_L(X_i \underset{h}{\longrightarrow} X_j|I)$ and $C_{ij} := C_{ij}^1$.

For any given forecast horizon $h \geq 1$, we have the multiple horizon causality measures for each pair of nodes in a network as Table 1 shows. Point-wise edges in the network are measured by the values of $C_{ij}^h$, $i = 1,...,N$ and $j = 1,...,N$. Table 1 is exactly corresponding to a static network structure. The $i$th row and $j$th column element in Table 1 is the strength of the directed edge from node $i$ to node $j$. $C_{ij}$ measures the direct effect from node $i$ to node $j$: $S(i \to j)$, where $S(i \to j)$ denote the effect from node $i$ to node $j$ via the path $(i \to j)$. For $h > 1$, $C_{ij}^h$ measures the total indirect effect from node $i$ to node $j$ via every possible path with length $h$: $S(i \to k_1 \to k_2 \to ... \to k_{h-1} \to j)$ for any $k_i \in V, i = 1,...,h-1$, where $S(i \to k_1 \to k_2 \to ... \to k_{n-1} \to j)$ denote the indirect effect from node $i$ to node $j$ via the path $(i \to k_1 \to k_2 \to ... \to k_{n-1} \to j)$. In other words, $C_{ij}^h$ measures the indirect effect from node $i$ to node $j$ with taking into account all the interconnections in the network. Intuitively, the forecast horizon $h$ can be interpreted as the effect-radius when considering the effect between any pair of nodes. For example, when $h = 1$, we only measure the direct effect (1-step effect); when $h = 100$, the effect between any pair of nodes could "walk" via as many as 99 different other nodes in the network. Another way to understand the difference between $C_{ij}^1$ and $C_{ij}^h$ ($h > 1$) is to consider the difference among standard network centrality measures (e.g., Degree, Closeness, Betweenness and Eigenvector). These centrality measures differ from each others mainly in how to weight the importance of the nodes that a node connected to to measure this node's importance in the network. For instance, the degree centrality only calculate how many nodes that a node directly connected to to characterize this node's importance, while the

Table 1. Causality table (given forecast horizon $h$)

| nodes | 1 | 2 | $\cdots$ | $i$ | $\cdots$ | $j$ | $\cdots$ | $N$ |
|---|---|---|---|---|---|---|---|---|
| 1 | $C_{11}^h$ | $C_{12}^h$ | $\cdots$ | $C_{1i}^h$ | $\cdots$ | $C_{1j}^h$ | $\cdots$ | $C_{1N}^h$ |
| 2 | $C_{21}^h$ | $C_{22}^h$ | $\cdots$ | $C_{2i}^h$ | $\cdots$ | $C_{2j}^h$ | $\cdots$ | $C_{2N}^h$ |
| $\vdots$ | | | | | | | | |
| $i$ | $C_{i1}^h$ | $C_{i2}^h$ | $\cdots$ | $C_{ii}^h$ | $\cdots$ | $C_{ij}^h$ | $\cdots$ | $C_{iN}^h$ |
| $\vdots$ | | | | | | | | |
| $j$ | $C_{j1}^h$ | $C_{j2}^h$ | $\cdots$ | $C_{ji}^h$ | $\cdots$ | $C_{jj}^h$ | $\cdots$ | $C_{jN}^h$ |
| $\vdots$ | | | | | | | | |
| $N$ | $C_{N1}^h$ | $C_{N2}^h$ | $\cdots$ | $C_{Ni}^h$ | $\cdots$ | $C_{Nj}^h$ | $\cdots$ | $C_{NN}^h$ |

eigenvector centrality assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes, thus the degree centrality is a "local" measure, and the eigenvector centrality is a "global" measure. Similarly, $C_{ij}^1$ is 1-step locally measuring the direct effect, and $C_{ij}^h$ is $h$-steps globally measuring the indirect effect.

In terms of mathematical definitions, group-wise edges are the generalization of point-wise edges. They are equivalent when the sizes of the groups equal 1. For any pair of nodes, $i$ and $j$, in a node set $V$, we say $i \underset{C,h}{\nrightarrow} j$ if and only if $C_{ij}^h = 0$. For any pair of groups of nodes $(i_1,...,i_{n_1})$ and $(j_1,...,j_{n_2})$, where $(i_1,...,i_{n_1}) = (j_1,...,j_{n_2})$ or $(i_1,...,i_{n_1}) \cap (j_1,...,j_{n_2}) = \emptyset$ for $(i_1,...,i_{n_1}),(j_1,...,j_{n_2}) \subset V$, we say $(i_1,...,i_{n_1}) \underset{C,h}{\nrightarrow} (j_1,...,j_{n_2})$ if and only if $C_{WY}^h = 0$, where $W = (i_1,...,i_{n_1})$ and $Y = (j_1,...,j_{n_2})$.

**Remark 3.2** Let $V_1 = (i_1,...,i_{n_1})$ and $V_2 = (j_1,...,j_{n_2})$. For any $i \in V_1$ and $j \in V_2$, because of $I_{-V_1} \subset I_{-i}$, $C_{V_1V_2}^h = 0$ $[(i_1,...,i_{n_1}) \underset{C,h}{\nrightarrow} (j_1,...,j_{n_2})]$ implies $C_{iV_2}^h = 0$ $[i \underset{C,h}{\nrightarrow} (j_1,...,j_{n_2})]$, and $C_{V_1j}^h = 0$ $[(i_1,...,i_{n_1}) \underset{C,h}{\nrightarrow} j]$ implies $C_{ij}^h = 0$ $[i \underset{C,h}{\nrightarrow} j]$.

Remark 3.2 says if a set of node(s) has no effect on some other node(s), any element of this set of node(s) also has no effect on those node(s). It is worth to emphasize here that $C_{ij}^h = 0$ $[i \underset{C,h}{\nrightarrow} j]$ for any $i \in V_1$ and for any $j \in V_2$ does NOT necessarily imply $C_{V_1V_2}^h = 0$ $[(i_1,...,i_{n_1}) \underset{C,h}{\nrightarrow} (j_1,...,j_{n_2})]$. In other words, the strength of $[(i,j) \underset{C,h}{\rightarrow} k]$ may be strong even if the strengths of $[i \underset{C,h}{\rightarrow} k]$ and $[j \underset{C,h}{\rightarrow} k]$ are weak. This circumstance is analogy to the difference between pairwise independence and mutually independence. When $X_i$ and $X_j$ are contemporaneously highly correlated, $X_i$'s marginal effect on $X_k$, conditional on $X_j$, will be very small since all relevant information in $X_i$ that helps to predict $X_k$ has been captured by $X_j$.

In fact, studying the role of a group of nodes in a network is an important topic. In social network literature, for instance, just as looking into who is the center node in a network, which can be measured by standard centrality measures (see, e.g., Freeman (1978) and Jackson (2008)), we also may want to find which group of nodes is center in a network, which can be measured by the generalizations of the standard centrality measures (see Everett and Borgatti (1999)). From measurement perspective, the importance of a group of node(s) has to base on the interconnections of this group to other nodes in the network (and all other interconnections in the network). To the best of our knowledge, surprisingly, measuring the effects of a group of nodes on other nodes in a network is still missing in the network econometric literature. Our network measurement method can exactly fill this blank. Moreover, our group-wise edges measurement method is compatible with the classic network literature. Remark 3.2 suggests our generation of pair-wise edges by group-wise edges is in line with the generation of the node centralities in Freeman (1978) by the group centralities in Everett and Borgatti (1999).

From the discussion in this section, we have seen that the multiple horizon measures in Definition 3.1 have causality implications for the edges it measures. It is also very flexible to be applied to study indirect effects in directed and weighted network structures. Properties of network analysis, which can be applied to study complex interconnections in an economic system, have been studied in mathematics and computer science as graph theory. As Eichler (2007) shows, the multiple horizon causality in Dufour and Renault (1998), the the base of the multiple horizon causality measures, is also well matched to path diagrams in the multivariate time series context. Thus our network measurement framework is also in line with the network analysis in graph theory. Lastly, point-wise edges, $(i \to j)$, and group-wise edges, $([i_1, i_2, ..., i_n] \to [j_1, j_2, ..., j_m])$, can be analysed by our multiple horizon causality measures framework.

# 4. LASSO estimation of causality measure

In this section, we estimate the multiple horizon causality measures $C_L(X^W \underset{h}{\longrightarrow} X^Y|I)$ and $C_L(X^W \underset{h}{\longrightarrow} X^W|I)$ in a high-dimensional context. Given the network's nodes processes $X = \{X_{1t}, X_{2t}, ..., X_{Nt}\}_{t=1}^{T}$, following Dufour and Taamouti (2010) we use the VAR framework in our econometric analysis. Network estimation under the VAR representation is desirable since the VAR models are naturally developed to investigate the pairwise effect in a complex linear structure. Unlike Dufour and Taamouti (2010) only deal with low-dimensional situations, we estimate the multiple horizon causality measures with the LASSO approach to better fit the multiple horizon causality measures framework into high-dimensional network analysis.

**Assumption 4.1** PROCESSES VAR REPRESENTATIONS.

*The unrestricted process* $X = \{X_t^W, X_t^Y, X_t^Z\}_{t=1}^{T} = \{X_{1t}, X_{2t}, ..., X_{Nt}\}_{t=1}^{T}$ *is strictly stationary and* *has a VAR($\infty$) representation,*

$$X(t) = \sum_{k=1}^{\infty} A_k X(t-k) + u(t), \tag{4.1}$$

*where $X(t) = [X_{1t}, X_{2t}, ..., X_{Nt}]'$ is a $N \times 1$ vector, $A_k$ is $N \times N$ matrix and $u(t) \sim w.n.(0, \Sigma_u)$. $\Sigma_u$ is a $N \times N$ positive definite matrix.*

*The restricted process $X_0 = \{X_t^Y, X_t^Z\}_{t=1}^T$ is strictly stationary and has a VAR($\infty$) representation,*

$$X_0(t) = \sum_{k=1}^{\infty} \bar{A}_k X_0(t-k) + \varepsilon(t), \tag{4.2}$$

*where $X_0(t) = [X_t^Y, X_t^Z]'$ is a $(N-m_1) \times 1$ vector, $\bar{A}_k$ is $(N-m_1) \times (N-m_1)$ matrix and $\varepsilon(t) \sim w.n.(0, \Sigma_\varepsilon)$. $\Sigma_\varepsilon$ is a $(N-m_1) \times (N-m_1)$ positive definite matrix.*

*The restricted process has the following expanded representation,*

$$X(t) = \sum_{k=1}^{\infty} \bar{A}_k^{\phi} J_2 X(t-k) + \nu(t) \tag{4.3}$$

*where $\bar{A}_k^{\phi} = \begin{bmatrix} \bar{A}_k^W \\ \bar{A}_k \end{bmatrix}_{N \times (N-m_1)}$, $\bar{A}_k$ is defined in (4.2) and $\bar{A}_k^W$ is the expanded coefficients for $X^W$. $J_2 = [0_{(N-m_1)\times m_1}, I_{(N-m_1)\times(N-m_1)}]_{(N-m_1)\times N}$ and $\nu(t) \sim w.n.(0, \Sigma_\nu)$. $\Sigma_\nu$ is a $N \times N$ positive definite matrix.*

**Remark 4.1** Under the Assumption 4.1, the covariance matrix of the forecast error at horizon $h$ for the unrestricted model (4.1) is

$$\Sigma[X(t+h)|\mathscr{F}(t)] \equiv \sum_{q=0}^{h-1} \varphi_q \Sigma_u \varphi_q', \tag{4.4}$$

where $\varphi_q = \sum_{k=1}^{q} A_k \varphi_{q-k}$ and $\varphi_0 = I_N$. The covariance matrix of the forecast error at horizon $h$ for the restricted model (4.2) is

$$\Sigma[X_0(t+h)|\mathscr{F}_{-W}(t)] \equiv \sum_{q=0}^{h-1} \bar{\varphi}_q \Sigma_\varepsilon \bar{\varphi}_q', \tag{4.5}$$

where $\bar{\varphi}_q = \sum_{k=1}^{q} \bar{A}_k \bar{\varphi}_{q-k}$ and $\bar{\varphi}_0 = I_{N-m_1}$.

**Definition 4.1** *Under the Assumption 4.1 and by the Remark 4.1, the multiple horizon causality* *measure, from W to Y, at forecast horizon h is*

$$C_L(X^W \underset{h}{\longrightarrow} X^Y|I) := \ln\left[\frac{det\{J_0\Sigma[X_0(t+h)|\mathscr{F}_{-W}(t)]J_0'\}}{det\{J_1\Sigma[X(t+h)|\mathscr{F}(t)]J_1'\}}\right] \tag{4.6}$$

*where* $J_0 = [I_{m_2},\ 0_{m_2\times m_3}]_{m_2\times(N-m_1)}$ *and* $J_1 = [0_{m_2\times m_1},\ I_{m_2},\ 0_{m_2\times m_3}]_{m_2\times N}$. $\Sigma[X(t+h)|\mathscr{F}(t)]$ *and* $\Sigma[X_0(t+h)|\mathscr{F}_{-W}(t)]$ *are defined in* (4.4) *and* (4.5) *respectively.*

**Remark 4.2** Under the Assumption 4.1, it can be easy to observe that $\Sigma_\varepsilon = J_2\Sigma_\nu J_2'$ and $X_0(t) = J_2X(t)$. Then the forecast error covariance of $X^W$ at horizon $h$, without its past information, is

$$\Sigma_W[X^W(t+h)|\mathscr{F}_{-W}(t)] \equiv J_3\left(\sum_{q=0}^{h-1}\phi_q\Sigma_\nu\phi_q'\right)J_3', \tag{4.7}$$

where $\phi_q = \sum_{k=1}^{q} A_k^{\phi}\phi_{q-k}$, $A_k^{\phi} = \bar{A}_k^{\phi}J_2$, $\phi_0 = I_N$, $J_3 = [I_{m_1\times m_1},\ 0_{m_1\times(N-m_1)}]_{m_1\times N}$.

**Definition 4.2** *Under the Assumption 4.1 and by the Remark 4.2, the multiple horizon causality measure, from W to W, at forecast horizon h is*

$$C_L(X^W \underset{h}{\longrightarrow} X^W|I) := \ln\left[\frac{det\{\Sigma_W[X_0(t+h)|\mathscr{F}_{-W}(t)]\}}{det\{J_3\Sigma[X(t+h)|\mathscr{F}(t)]J_3'\}}\right] \tag{4.8}$$

*where* $\Sigma_W[X_0(t+h)|\mathscr{F}_{-W}(t)]$ *and* $\Sigma[X(t+h)|\mathscr{F}(t)]$ *are defined in* (4.7) *and* (4.4) *respectively.*

In order to obtain $C_L(X^W \underset{h}{\longrightarrow} X^Y|I)$ and $C_L(X^W \underset{h}{\longrightarrow} X^W|I)$, we just need to know the autoregressive matrices, $[A_1, A_2, ..., A_{h-1}]$ and $[\bar{A}_1^{\phi}, \bar{A}_2^{\phi}, ..., \bar{A}_{h-1}^{\phi}]$, and the contemporaneous covariance matrices, $\Sigma_u$ and $\Sigma_\nu$. To estimate these parameters, we consider the truncated models of the unrestricted process (4.1) and the expanded restricted process (4.3) as

$$X(t) = \sum_{k=1}^{p} A_k^p X(t-k) + u^p(t), \tag{4.9}$$

$$X(t) = \sum_{k=1}^{p} \bar{A}_k^p X_0(t-k) + \nu^p(t). \tag{4.10}$$

where $u^p(t) \sim w.n.(0, \Sigma_u^p)$ and $\nu^p(t) \sim w.n.(0, \Sigma_\nu^p)$. $A_k^p$ and $\Sigma_u^p$ are $N$ by $N$ matrices for $k = 1, 2, .., p$. $\bar{A}_k^p$ is a $N$ by $N - m_1$ matrix and $\Sigma_\nu^p$ is a $N$ by $N$ matrix for $k = 1, 2, .., p$.

While the dimensions for two groups in group-wise edge analysis, $m_1$ and $m_2$, are fixed, we assume that the number of nodes, $N$, and the lag $p$ can be functions of $T$ (i.e., $N_T = O(T^{c_1})$ and $p_T = O(T^{c_2})$ for constant $c_1, c_2 > 0$), but for notation simplicity we do not write the subscript $T$ explicitly. Under mild assumptions, the truncated bias is asymptotically negligible such that $||A_k^p - A_k||_\infty = o(1)$ for $k = 1, 2, ..., p$ and $||\Sigma_u^p - \Sigma_u||_\infty = o(1)$ as $T \longrightarrow \infty$. We can therefore estimate the parameters of interest with the truncated models. Similar arguments can be applied to the expanded restricted truncated model. The unrestricted model and the expanded restricted model basically share the same estimation procedure.

The main estimation challenge in a network context is the high-dimensionality problem. We have $N \times N \times p$ unknown parameters in the autoregressive matrices $A_k^p$, $k = 1, ..., p$, as well as $\frac{N(N+1)}{2}$ unknown parameters in the contemporaneous covariance matrix $\Sigma_u^p$, but we only have $N \times T$ observations. For a market network, the number of nodes, $N$, can be large. Traditional estimation methods are not reliable when $N \times p$ is closed to $T$ or even infeasible when $N \times p > T$. One of the popular ways to solve the high-dimensional problem in statistics is by assuming sparsity such that the effective dimension of the parameter space keeps tractable. The statistical intuition is that we can set free of the limited observable data by assuming appropriate sparsity structures to only estimate the nonzero parameters. One thing we need to emphasize here is that we do not assume the network structure, measured by the multiple horizon causality measures table, is sparse. Instead, we only need to assume the autoregressive matrices, and the error concentration matrix are sparse. Since the multiple horizon causality measures are nonlinear functions of the autoregressive matrices and the concentration matrices, the causality table is generally nonsparse. The estimation technique in this section is called the Least Absolute Shrinkage and Selection Operator (LASSO) (see, e.g., Tibshirani (1996)).

Under sparsity assumptions, the autoregressive coefficients and the error concentration matrices could be estimated simultaneously (see Barigozzi and Brownlees (2016)). As the dimension of the unknown parameter space is huge, however, this estimation procedure could be time intensive. Note that the multiple horizon causality measures requires estimating as many as $N + 1$ models (one unrestricted model and $N$ restricted models) to quantify the effects from one to others. The estimation efficiency, in terms of computational time, is also an important issue to be concerned when $N$ is large.

For empirical convenience, we apply a faster two-stage estimation procedure. At stage one, we use the Adaptive LASSO regression (see, e.g., Zou (2006)) to estimate the autoregressive coefficients. At stage two, the error concentration matrix can be estimated by the residuals from the stage one. It comes a cost that the rate of convergence of the estimator in the second stage will depend on the estimator in the first stage (see Barigozzi and Brownlees (2014)), and thus this method is theoretically less desirable than the joint estimation method in Barigozzi and Brownlees (2016). Nonetheless, we believe this tradeoff for empirical convenience is worthwhile because our network measurement method is designed for general empirical applications. Our estimation procedure uses the results provided by Barigozzi and Brownlees (2014).

## 4.1. Autoregressive matrix estimation

Each of the $N$ equation of the unrestricted VAR($p$) model can be written as

$$X_i(t) = \alpha_i' z(t) + u_i^p(t), \tag{4.11}$$

where $X_i(t)$ is the $i$th univariate time series in $X(t)$. $z(t) = (X'(t-1), X'(t-2), ..., X'(t-p))'$ is the $Np \times 1$ vector of lagged observations. $\alpha_i = (\alpha_{1i1}, ..., \alpha_{1iN}, ..., \alpha_{pi1}, ..., \alpha_{piN})'$ is a $Np \times 1$ parameter vector, such that $\text{vec}(\alpha_1', ..., \alpha_N') = \text{vec}([A_1^p, A_2^p, ..., A_p^p]')$. For each of the $N$ equation of the expanded restricted VAR($p$) model, similarly,

$$X_i(t) = \bar{\alpha}_i' z_0(t) + \varepsilon_i^p(t), \tag{4.12}$$

the unknown autoregressive coefficient vector $\bar{\alpha}_i = (\bar{\alpha}_{1i1}, ..., \bar{\alpha}_{1iN-m_1}, ..., \bar{\alpha}_{pi1}, ..., \bar{\alpha}_{piN-m_1})'$ is a $(N-m_1)p \times 1$ vector, such that $\text{vec}(\bar{\alpha}_1', ..., \bar{\alpha}_N') = \text{vec}([\bar{A}_1^p, \bar{A}_2^p, ..., \bar{A}_p^p]')$. $z_0(t) = (X_0'(t-1), X_0'(t-2), ..., X_0'(t-p))'$ is the $(N-m_1)p \times 1$ vector of lagged observations.

The Adaptive LASSO estimators of $\alpha_i$ and $\bar{\alpha}_i$ are defined respectively as

$$\hat{\alpha}_{Ti} = \underset{\alpha}{\text{argmin}} \frac{1}{T} \sum_{t=1}^{T} [X_i(t) - \alpha_i' z(t)]^2 + \frac{\lambda_T}{T} \sum_{j=1}^{Np} w_{Tij} |\alpha_{ij}| \quad \text{for } i = 1, ..., N, \tag{4.13}$$

$$\hat{\bar{\alpha}}_{Ti} = \underset{\bar{\alpha}}{\text{argmin}} \frac{1}{T} \sum_{t=1}^{T} [X_i(t) - \bar{\alpha}_i' z_0(t)]^2 + \frac{\lambda_T}{T} \sum_{j=1}^{(N-m_1)p} \bar{w}_{Tij} |\bar{\alpha}_{ij}| \quad \text{for } i = 1, ..., N \tag{4.14}$$

where $\lambda_T$ is an appropriate pre-selected value controlling the overall estimated sparsity level in the autoregressive models. If $\lambda_T$ equals 0, then the LASSO estimation is simply the OLS estimation and every element in $\alpha_i$ have to be estimated; if $\lambda_T \longrightarrow \infty$, the estimates of the parameter $\alpha_i$ are all zeros, which means the estimated autoregressive coefficients are perfectly sparse. The choice of $\lambda_T$ can be selected by the BIC criterion or by Cross-Validations. $w_{Tij}$ and $\bar{w}_{Tij}$ are pre-estimator weighted penalties to the sparse structures of $\alpha_i$ and $\bar{\alpha}_i$. They help to separate zero coefficients from nonzero coefficients when regressos are highly correlated. Here we use $w_{Tij} = \frac{1}{|\hat{\alpha}_{ij}^{LASSO}|}$ and $\bar{w}_{Tij} = \frac{1}{|\hat{\bar{\alpha}}_{ij}^{LASSO}|}$ as the weighted penalties to $|\alpha_{ij}|$ and $|\bar{\alpha}_{ij}|$ respectively, where $\hat{\alpha}_{ij}^{LASSO}$ and $\hat{\bar{\alpha}}_{ij}^{LASSO}$ are the standard LASSO estimators: $w_{Tij} = 1$ for $\alpha_{ij}$ and $\bar{w}_{Tij} = 1$ for $\bar{\alpha}_{ij}$.

In order to estimate $\alpha_i$ and $\bar{\alpha}_i$ in a high-dimensional context, sparsity assumptions are required. We denote the sets of nonzero entries in $\alpha_i$ and in $\bar{\alpha}_i$ as $\mathscr{A}_i$, which has $q_{Ti}^{\mathscr{A}}$ elements, and $\bar{\mathscr{A}_i}$, which has $q_{Ti}^{\bar{\mathscr{A}}}$ elements. $q_{Ti}^{\mathscr{A}}$ and $q_{Ti}^{\bar{\mathscr{A}}}$ are function of $T$. Since the estimation of $\alpha_i$ is similar to the estimation of $\bar{\alpha}_i$, we here only discuss the sparsity of $\alpha_i$. Following Barigozzi and Brownlees (2014), the key assumptions on the number of nonzero entries in the autoregressive coefficients and the pre-selected penalty constant controlling the overall estimated sparsity level are $q_{Ti}^{\mathscr{A}} = o\left(\sqrt{\frac{T}{\log T}}\right)$, $\frac{\lambda_T}{T}\sqrt{q_{Ti}^{\mathscr{A}}} = o(1)$, $\lim_{T \to \infty} \frac{\lambda_T}{T}\sqrt{\frac{T}{\log T}} = \infty$, $\sqrt{\frac{q_{Ti}^{\mathscr{A}} \log T}{T}} = o\left(\frac{\lambda_T}{T}\right)$ and $\frac{\lambda_T}{T^{1-c_1}}\sqrt{q_{Ti}^{\mathscr{A}}} = O(1)$ for $i = 1, ..., N$. These assumptions provide the restrictions among the the underlying true sparsity level ($q_{Ti}^{\mathscr{A}}$), pre-selected penalty constant controlling the overall estimated sparsity level ($\lambda_T$) and rate of number of nodes $N$ as $T$ grows to infinity ($c_1$). To identify the zero entries in $\alpha_i$, we also need the following assumption on the signal strength: For all $i = 1,...,N$, there exists a sequence of positive real numbers $\{s_{Ti}^{\mathscr{A}}\}$ such that $|\alpha_{ij}| > s_{Ti}^{\mathscr{A}}$ and $\lim_{T\to\infty} \frac{s_{Ti}^{\mathscr{A}}}{\frac{\lambda_T}{T}\sqrt{q_{Ti}^{\mathscr{A}}}} = \infty$ for all $\alpha_{ij} \in \mathscr{A}_i$.

**Proposition 4.1** *Under the Assumption 1 - 6 in Appendix A.1, as $T \longrightarrow \infty$,*

1. *if $\alpha_{ij} \in \mathscr{A}_i^C$, $Prob\{\hat{\alpha}_{Tij} = 0\} \to 1$, $\quad i = 1,...,N$*
2. *if $\bar{\alpha}_{ij} \in \bar{\mathscr{A}}_i^C$, $Prob\{\hat{\bar{\alpha}}_{Tij} = 0\} \to 1$, $\quad i = 1,...,N$*
3. *$\hat{\alpha}_{Ti} \overset{p}{\longrightarrow} \alpha_i$, and thus $\hat{A}_{Tk}^p \overset{p}{\longrightarrow} A_k$ for $k = 1,...,p$*
4. *$\hat{\bar{\alpha}}_{Ti} \overset{p}{\longrightarrow} \bar{\alpha}_i$ and thus $\hat{\bar{A}}_{Tk}^p \overset{p}{\longrightarrow} \bar{A}_k^{\phi}$ for $k = 1,...,p$*

PROOF. See in Appendix A.2. □

Proposition **4.1** states that the Adaptive LASSO estimators in (4.13) and (4.14) correctly select the nonzero coefficients asymptotically, and the estimators are consistent. Even if the dimension of the network is large, this estimation procedure can still safely concentrate on estimating the nonzero coefficients using the limited information from the observable sample, given the sparsity assumption on the true coefficients vector.

## 4.2. Contemporaneous covariance matrix estimation

The contemporaneous covariance matrix can be estimated by the sparse concentration matrix via the sparse errors partial correlations. We use the estimation strategy in Peng, Wang, Zhou and Zhu (2009) and Barigozzi and Brownlees (2014). The errors partial correlations matrix $\rho$ has generic component $\rho_{ij}$. The concentration matrix in the unrestricted model, $S_u^p \equiv [\Sigma_u^p]^{-1}$, and the concentration matrix in the expanded restricted model, $S_v^p \equiv [\Sigma_v^p]^{-1}$, have the following relationship with their respective errors correlations:

$$\rho_{ij}^u \equiv \mathrm{Corr}(u_i^p, u_j^p) = -\frac{s_{ij}^u}{\sqrt{s_{ii}^u s_{jj}^u}} \tag{4.15}$$

$$\rho_{ij}^v \equiv \mathrm{Corr}(v^p, v^p) = -\frac{s_{ij}^v}{\sqrt{s_{ii}^v s_{jj}^v}} \tag{4.16}$$

where $s_{ij}^u$ is the $(i,j)$ component of $S_u^p$ and $s_{ij}^v$ is the $(i,j)$ component of $S_u^v$. Moreover, the errors correlations can be also expressed as the coefficients of the linear regressions (see Lemma 1 in Peng et al. (2009)):

$$u^p_{ti} = \sum_{j\neq i}^{N} \rho^u_{ij} \sqrt{\frac{s^u_{ii}}{s^u_{jj}}} u^p_{tj} + \eta^u_{ti}, \quad i = 1,...,N, \quad t = 1,...,T \tag{4.17}$$

$$v^p_{ti} = \sum_{j\neq i}^{N} \rho^v_{ij} \sqrt{\frac{s^v_{ii}}{s^v_{jj}}} v^p_{tj} + \eta^v_{ti}, \quad i = 1,...,N, \quad t = 1,...,T \tag{4.18}$$

We assume the concentration matrices as well as the correlations matrices are sparse and denote the set of nonzero entries in the unrestricted and restricted errors correlation matrices as $\mathscr{Q}_u$ and $\mathscr{Q}_v$ respectively. The LASSO estimator of the errors partial correlations in the unrestricted model (4.1) and the one in the restricted model (4.3) are defined respectively as

$$\hat{\rho}^u_T = \underset{\rho^u}{\text{argmin}} \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} (\hat{u}_{ti} - \sum_{j\neq i}^{N} \rho^u_{Tij} \sqrt{\frac{\hat{s}^u_{Tii}}{\hat{s}^u_{Tjj}}} \hat{u}_{tj})^2 + \frac{\gamma_T}{T} \sum_{i=2}^{N} \sum_{j=1}^{i-1} |\rho^u_{ij}|, \tag{4.19}$$

$$\hat{\rho}^v_T = \underset{\rho^v}{\text{argmin}} \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} (\hat{v}_{ti} - \sum_{j\neq i}^{N} \rho^v_{Tij} \sqrt{\frac{\hat{s}^v_{Tii}}{\hat{s}^v_{Tjj}}} \hat{v}_{tj})^2 + \frac{\gamma_T}{T} \sum_{i=2}^{N} \sum_{j=1}^{i-1} |\rho^v_{ij}| \tag{4.20}$$

where $\hat{u}_{ti} = X_i(t) - \hat{\alpha}'_i z(t)$ and $\hat{v}_{ti} = X_i(t) - \hat{\tilde{\alpha}}'_i z_0(t)$. $\gamma_T$ is the tuning parameter controlling the model sparsity level as $\lambda_T$ in (4.13) and in (4.14). The estimator of the unrestricted concentration matrix $S^p_u$, denoted as $\hat{S}^u_T$, and the the estimator of the expanded restricted concentration matrix $S^p_v$, denoted as $\hat{S}^v_T$, have entries $\hat{s}^u_{Tij} = -\hat{\rho}^u_{Tij}\sqrt{\hat{s}^u_{Tii}\hat{s}^u_{Tjj}}$ and $\hat{s}^v_{Tij} = -\hat{\rho}^v_{Tij}\sqrt{\hat{s}^v_{Tii}\hat{s}^v_{Tjj}}$. The estimators $\hat{s}^u_{Tii}$ and $\hat{s}^v_{Tii}$ are given respectively by

$$\hat{s}^u_{Tii} = \left[ \frac{1}{T-1} \sum_{t=1}^{T} (\hat{\eta}^u_{ti})^2 \right]^{-1}, \tag{4.21}$$

$$\hat{s}^v_{Tii} = \left[ \frac{1}{T-1} \sum_{t=1}^{T} (\hat{\eta}^v_{ti})^2 \right]^{-1} \tag{4.22}$$

where $\hat{\eta}^u_{ti} = \hat{u}_{ti} - \sum_{j\neq i}^{N} \rho^u_{Tij} \sqrt{\frac{\hat{s}^u_{Tii}}{\hat{s}^u_{Tjj}}} \hat{u}_{tj}$ and $\hat{\eta}^v_{ti} = \hat{v}_{ti} - \sum_{j\neq i}^{N} \rho^v_{Tij} \sqrt{\frac{\hat{s}^v_{Tii}}{\hat{s}^v_{Tjj}}} \hat{v}_{tj}$

The estimator of the unrestricted errors concentration matrix, $\hat{S}^u_T$, can be obtained by iterating between (4.19) and (4.21). The estimator of the expanded restricted errors concentration matrix, $\hat{S}^v_T$, can be obtained by iterating between (4.20) and (4.22). For more discussions on the assumptions to estimate the correlasso matrices, we refer readers to see Peng et al. (2009) and Barigozzi and Brownlees (2014).

**Proposition 4.2** *Under the Assumption 1 - 9 in Appendix A.1, as $T \to \infty$,*

1. *if $\rho^u_{ij} \in \mathscr{Q}^C_u$, $Prob\{\hat{\rho}^u_{Tij} = 0\} \to 1, \quad i, j = 1,...,N$*

2. *if* $\rho^{v}_{ij} \in \mathscr{Q}^{C}_{v}$, $Prob\{\hat{\rho}^{v}_{Tij} = 0\} \to 1, \quad i,j = 1,\ldots,N$

3. $\hat{\rho}^{u}_{Tij} \xrightarrow{p} \rho^{u}_{ij}$, *and thus* $\hat{S}^{u}_{T} \xrightarrow{p} S_{u} \equiv \Sigma^{-1}_{u}$

4. $\hat{\rho}^{v}_{Tij} \xrightarrow{p} \rho^{v}_{ij}$, *and thus* $\hat{S}^{v}_{T} \xrightarrow{p} S_{v} \equiv \Sigma^{-1}_{v}$

PROOF. See in Appendix A.3. □

Proposition **4.2** states that the LASSO estimators in (4.19) and (4.20) correctly select the nonzero coefficients in the errors correlation matrics asymptotically and the estimators are consistent. By the relationships between errors correlations and the concentration matrix in (4.15) and (4.16), we obtain the consistent estimators of the concentration matrices for the unrestricted model (4.1) and the concentration matrices for the expanded unrestricted model (4.3).

## 4.3. Multiple horizons causality measures estimation

Note that the multiple horizon causality measures under the Assumption 4.1 is mainly composed by two parts (see Definition 4.1 and Remark 4.1): i) autoregressive coefficients in the unrestricted model (4.1) and in the expanded restricted model (4.3); ii) contemporaneous covariances in the unrestricted model (4.1) and in the expanded restricted model (4.3). We have already obtained their consistent estimators by Proposition **4.1** and Proposition **4.2**.

Finally, the estimator of the multiple horizon causality measure, from $X^W$ to $X^W$, is defined as

$$\hat{C}^{h}_{TWW} := \ln\left[\frac{\det\{\hat{\Sigma}_W[X^W(t+h)|\mathscr{F}_{-W}(t)]\}}{\det\{J_3\hat{\Sigma}[X(t+h)|\mathscr{F}(t)]J_3'\}}\right], \tag{4.23}$$

where

$$\hat{\Sigma}[X(t+h)|\mathscr{F}(t)] = \sum_{q=0}^{h-1}\hat{\varphi}_q(\hat{S}^{u}_{T})^{-1}\hat{\varphi}_q', \tag{4.24}$$

$$\hat{\Sigma}_W[X^W(t+h)|\mathscr{F}_{-W}(t)] = J_3\left[\sum_{q=0}^{h-1}\hat{\phi}_q(\hat{S}^{v}_{T})^{-1}\hat{\phi}_q'\right]J_3', \tag{4.25}$$

$$\hat{\phi}_q = \sum_{k=1}^{q}(\hat{\bar{A}}^{p}_{Tk}J_2)\hat{\phi}_{q-k}, \tag{4.26}$$

$$\hat{\varphi}_q = \sum_{k=1}^{q}\hat{A}^{p}_{Tk}\hat{\varphi}_{q-k}, \tag{4.27}$$

$$\hat{\varphi}_0 = I_N,\ \hat{\phi}_0 = I_N. \tag{4.28}$$

The estimator of the multiple horizon causality measures, from $X^W$ to $X^Y$, is defined as

$$\hat{C}^h_{TWY} := \ln\left[\frac{\det\{J_0\hat{\Sigma}[X_0(t+h)|\mathscr{F}_{-W}(t)]J_0'\}}{\det\{J_1\hat{\Sigma}[X(t+h)|\mathscr{F}(t)]J_1'\}}\right], \tag{4.29}$$

where

$$\hat{\Sigma}[X(t+h)|\mathscr{F}(t)] = \sum_{q=0}^{h-1}\hat{\varphi}_q(\hat{S}^u_T)^{-1}\hat{\varphi}_q', \tag{4.30}$$

$$\hat{\Sigma}[X_0(t+h)|\mathscr{F}_{-W}(t)] = \sum_{q=0}^{h-1}\hat{\bar{\varphi}}_q(\hat{S}^\varepsilon_T)^{-1}\hat{\bar{\varphi}}_q', \tag{4.31}$$

$$(\hat{S}^\varepsilon_T)^{-1} = J_2(\hat{S}^\nu_T)^{-1}J_2', \tag{4.32}$$

$$\hat{\bar{\varphi}}_q = \sum_{k=1}^{q}(J_2\hat{\bar{A}}^p_{Tk})'\hat{\bar{\varphi}}_{q-k}, \tag{4.33}$$

$$\hat{\varphi}_q = \sum_{k=1}^{q}\hat{A}^p_{Tk}\hat{\varphi}_{q-k}, \tag{4.34}$$

$$\hat{\varphi}_0 = I_N,\ \hat{\bar{\varphi}}_0 = I_{N-m_1}. \tag{4.35}$$

**Theorem 4.3** *Under the Assumption 1 - 9 in Appendix A.1, for any given h,* $h = 1,2,...,$ *as* $T \longrightarrow \infty$,

1. $\hat{\Sigma}[X(t+h)|\mathscr{F}(t)] \xrightarrow{p} \Sigma[X(t+h)|\mathscr{F}(t)]$;
2. $\hat{\Sigma}[X_0(t+h)|\mathscr{F}_{-W}(t) \xrightarrow{p} \Sigma[X_0(t+h)|\mathscr{F}_{-W}(t)]$;
3. $\hat{\Sigma}_W[X^W(t+h)|\mathscr{F}_{-W}(t) \xrightarrow{p} \Sigma_W[X^W(t+h)|\mathscr{F}_{-W}(t)]$;
4. $\hat{C}^h_{TWY} \xrightarrow{p} C_L(X^W \underset{h}{\longrightarrow} X^Y|I)$;
5. $\hat{C}^h_{TWW} \xrightarrow{p} C_L(X^W \underset{h}{\longrightarrow} X^W|I)$.

PROOF. See in Appendix A.4. □

Now, we have the consistent estimators of the multiple horizon causality measures for any given networks. Measuring the point-wise edge strength, $i \underset{C,h}{\rightarrow} j$, and the group-wise edge strength, $(i_1,...,i_{n_1}) \underset{C,h}{\rightarrow} (j_1,...,j_{n_2})$, for arbitrary horizon $h \geq 1$ shares the same estimation procedure suggested in this section.

# 5. Network connectedness measures

The world is not flat. While the relationships of entities in an economy can be modelled by 2-dimensional network representations, the economy itself, however, is multi-dimensionally structured. Different firms play different roles. Some them are alike: insurances companies sell a wide range of insurances; some of them are distinctive: restaurants serve cuisines and the Space X provides space transportation services. We do not assume we have the prior knowledge of their exact roles, but we have their interconnection structures that can be measured by our causality table. If we use one's interconnection relationships to the others as a proxy of a firm's role in an economy, the causality table gives us the firms' coordinates of their roles in the multi-dimensional economy. Once the coordinates in a multi-dimensional space are given, it is easy to define direction and distance measures with middle school's geometry. We consider an economy of interest as a $N$-dimensional Euclidean space. The coordinate of a firm is corresponding to a vector in the multi-dimensional space. The direction of a firm's vector can be interpreted as "what the firm's role is": firm $i$'s vector direction tends to point to the companies that firm $i$ has more relationships to; the norm of a firm's vector can be interpreted as "how strong the firm's role is": the norm of firm $i$'s vector measures the extent of the firm $i$' relationships to all companies in the economy.

Following this logic, we define our new connectedness measures in the market network base on the multi-dimensional economy setting. We hereafter take the estimated causality measures table as given. Note that a network can be divided into several subgroups, the network can be viewed as a combination of its sub-networks. In a stock market, for example, the market index can be viewed as the weighted average of the prices of individual stocks as well as the weighted average of different sector indices. Since an economic network can be viewed as a network among firms (firm-wise market), whose interconnections are measured by our point-wise edges method ($i \rightarrow j$), as well as a network among sectors (sector-wise market), whose interconnections are measured by group-wise edges method ($[i_1, i_2, ..., i_n] \rightarrow [j_1, j_2, ..., j_m]$), we have three types of connectedness measures to gauge network interconnections: i) firm-wise market connectedness, which measures the interconnectedness of a firm-wise market; ii) firm-wise connectedness within a sector, which measures the interconnectedness within a given sector in a firm-wise market; and iii) sector-wise market connectedness, which measures the interconnectedness of a sector-wise market. These three types of connectedness measures fully take advantage of the flexibility of our network measurement method, so they can be applied to study market network connectedness in more flexible ways than Billio et al. (2012) and Diebold and Yilmaz (2014).

## 5.1. Firm-wise market connectedness measures

Market network connectedness can be decomposed by each firm's connectedness to the market. Firms' roles in an economy determine the firms' connectedness to the market network. As firm $i$'s

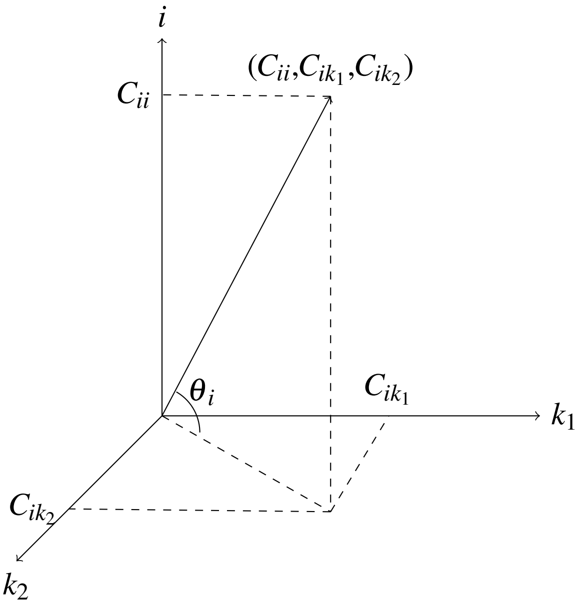

Figure 3. Relative connectedness between $i$ and network

vector direction represents firm $i$'s role in the economy, we will use firm $i$'s vector direction as the foundation to measure the firm $i$'s connectedness to the market network.

The connectedness, in term of economic role, of firm $i$ to the economy can be measured by the angle of firm $i$'s vector to the subspace of the economy composed by all other firms. In Figure 3, we use a simple 3-dimensional economy space to illustrate this idea. The economy has only three firms: $i$, $k_1$ and $k_2$. We want to study firm $i$'s role connectedness to this market. From the causality table, we choose the vector of $i$, $(C_{ii}, C_{ik_1}, C_{ik_2})$. It measures the relationships from $i$ to $i$, $k_1$ and $k_2$. The direction of $(C_{ii}, C_{ik_1}, C_{ik_2})$ in the 3-dimensional space determines firm $i$'s economic role in this market. $\theta_i$ is the angle of $i$'s vector to the subspace of the economy composed by $k_1$ and $k_2$. If we take $k_1$ and $k_2$ as a unit, $\theta_i$ exactly measures the tendency of firm $i$'s economic role to $k_1$ and $k_2$. When $\theta_i = \frac{\pi}{2}$, $i$ plays no role on $k_1$ and $k_2$; when $\theta_i = 0$, $i$ has relationships only to $k_1$ and $k_2$.

Given forecast horizon $h$, in a the $i$th row of the causality table measures the directed and weighted edges from the node $i$ to all nodes in a N-dimensional market network, which has a node set $V = \{1, ..., N\}$. Let $OUT_i^h = [C_{i1}^h, C_{i2}^h, ..., C_{iN}^h]'$, $OUT_i^h$ contains all the directed edges information pointed from $i$. They are the "out" effects from $i$ to all firms in the market. For $i = 1, ..., N \in V$, we define the "out" connectedness angle of the firm $i$ to the economy as $\theta_{i,V}^{out}(h)$,

$$\theta_{i,V}^{out}(h) = \arcsin \frac{C_{ii}^h}{\|OUT_{i,V}^h\|_2} \tag{5.1}$$

where we assume $\|OUT_{i,V}^h\|_2 > 0$, which is equivalent to say there exists $j \in \{1, ..., N\}$ such that $C_{ij}^h \neq 0$. If $\|OUT_{i,V}^h\|_2 = 0$, we let $\theta_{i,V}^{out}(h) = 0$ .

$\theta_{i,V}^{out}(h)$ measures the "out" connectedness from firm $i$ to the economy and is a relative connectedness strength since it has been rescaled in $[0, \pi/2]$. The connectedness angle $\theta_{i,V}^{out}(h) = \pi/2$ if and only if $C_{ii}^h > 0$ and $C_{ij}^h = 0$ for any $j \in \{1, ..., i-1, i+1, ..., N\}$, which implies firm $i$ is isolated with the economy in the sense that it has no impact on other companies. If $C_{ii}^h = 0$ and thus the projection angle $\theta_{i,V}^{out}(h) = 0$, it implies all relationships from firm $i$ to the economy are all from its impacts to other firms in the economy.

The relative connectedness strength of firm $i$ to the economy, measured by $\theta_{i,V}^{out}(h)$, considers the economic role of firm $i$ to the economy. It is a direction measure, and it is more related to relatively economic structures. The extent of how strong the economic roles, however, is not captured by $\theta_{i,V}^{out}(h)$. Besides the connectedness angle, we are also interested in the absolute connectedness strength. We denote a general formula of the absolute connectedness strength of firm $i$ to the economy as $K_{out}\left(\|OUT_{i,V}^h\|, \theta_{i,V}^{out}(h)\right)$. $K_{out}\left(\|OUT_{i,V}^h\|, \theta_{i,V}^{out}(h)\right)$ is a function of firm $i$'s connectedness angle, $\theta_{i,V}^{out}(h)$, and its causation strength to the economy, $\|OUT_{i,V}^h\|$. $K_{out}\left(\|OUT_{i,V}^h\|, \theta_{i,V}^{out}(h)\right)$ should at least satisfy the following properties:

$$K_{out}\left(\|OUT_{i,V}^h\|, \frac{\pi}{2}\right) = 0 \tag{5.2}$$

$$K_{out}\left(0, \theta_{i,V}^{out}(h)\right) = 0 \tag{5.3}$$

$$\frac{\partial K_{out}\left(\|OUT_{i,V}^h\|, \theta_{i,V}^{out}(h)\right)}{\partial \|OUT_{i,V}^h\|} \geq 0 \tag{5.4}$$

$$\frac{\partial K_{out}\left(\|OUT_{i,V}^h\|, \theta_{i,V}^{out}(h)\right)}{\partial \theta_{i,V}^{out}(h)} \leq 0 \tag{5.5}$$

The firm $i$ has no connectedness to the economy if has no impact on all other firms in the economy. The absolute connectedness strength between firm $i$ to the economy, should be an nondecreasing function of its causation strength to the economy and an nonincreasing function of its connectedness angle to the economy.

A simple functional specification of the absolute connectedness strength of firm $i$ to the economy we use in this paper is

$$K_{out}\left(\|OUT_{i,V}^h\|, \theta_{i,V}^{out}(h)\right) = \|OUT_{i,V}^h\| \cos \theta_{i,V}^{out}(h). \tag{5.6}$$

This absolute connectedness strength can be easily decomposed by the causation strength, $\|OUT_{i,V}^h\|$, and the connectedness angle $\theta_{i,V}^{out}(h)$. In geometric terms, $K_{out}\left(\|OUT_{i,V}^h\|, \theta_{i,V}^{out}(h)\right)$ just measures the norm of the projection of $OUT_{i,V}^h$ on the subspace spanned by the all other firms in the economy. We again use a simple 3-dimensional economy space to illustrate this idea. The economy has only three firms: $i$, $k_1$ and $k_2$. The absolute connectedness strength of firm $i$ to this

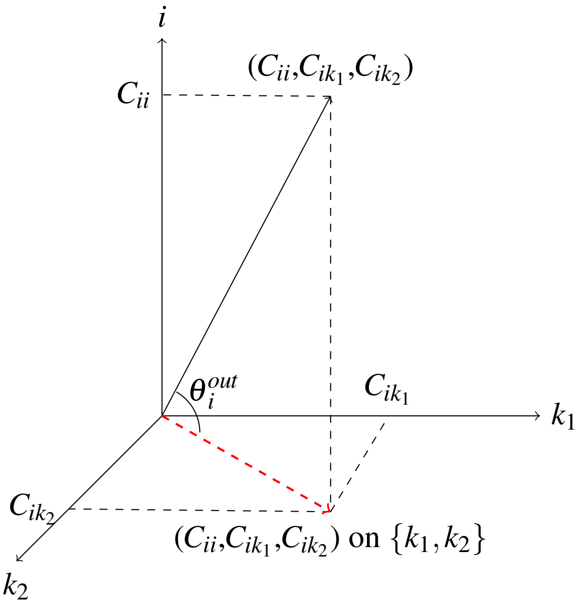


Figure 4. Absolute connectedness between $i$ and network

economy is the projection of the vector $(C_{ii}, C_{ik_1}, C_{ik_2})$ on the subspace spanned by $k_1$ and $k_2$, which is shown in Figure 4.

In summary, our absolute connectedness strength of firm $i$ to the economy simultaneously takes the firm $i$'s economic role structure and the economic role strength into account. In addition, the absolute connectedness strength can be easily decomposed by these two parts. Moreover, it has nice geometric interpretations in a N-dimensional economic space as illustrated by Figure 4. Therefore, our market connectedness measures, the mean of all firms' connectedness measures to the economy, will also enjoy these features.

**Definition 5.1** *Given a market network with node set $V = \{1, 2, ..., N\}$, the Firm-wise Market Network Relative Connectedness Structure Measure of "out" effects at horizon h, $MRC_{V_C}^{out}(h)$, is defined as following:*

$$MRC_{V_C}^{out}(h) = \frac{1}{N} \sum_{i=1}^{N} \cos \theta_{i,V}^{out}(h) \tag{5.7}$$

**Definition 5.2** *Given a market network with node set $V = \{1, 2, ..., N\}$, the Firm wise Market Network Absolute Connectedness Strength Measure of "out" effects at horizon h, $MAC_{V_C}^{out}(h)$, is defined as following:*

$$MAC_{V_C}^{out}(h) = \frac{1}{N} \sum_{i=1}^{N} \|OUT_{i,V}^{h}\| \cos \theta_{i,V}^{out}(h) \tag{5.8}$$

**Remark 5.1** If the edges in a network are unweighted, the connection of node $i$ to the network can be solely characterized by the connectedness angle, $\theta_{i,V}^{out}(h)$, irrespective to its unweighted absolute connectedness magnitude to the network, $\|OUT_{i,V}^{h}\|$. Therefore, the Firm-wise Market Network Relative Connectedness Structure Measure is basically equivalent to the Firm-wise Market Network Absolute Connectedness Strength Measure in the context of unweighted network.

Note that the edges in our network network are directed. Following similar procedures, we can also define market connectedness measures at "in" direction. Given a forecast horizon $h$, the $i$th column of the causality table measures the directed and weighted edges to the firm $i$ from all firms in the N-dimensional market network, which has node set $V = \{1,...,N\}$. Let $IN_i^h = [C_{1i}^h, C_{2i}^h, ..., C_{Ni}^h]'$, $IN_i^h$ contains all the directed edges information pointed from $i$. They are the "in" effects to $i$ from all firms in the market. For $i = 1,...,N \in V$, the "in" connectedness angle of the firm $i$ to the economy, $\theta_{i,V}^{in}(h)$, and the "in" absolute connectedness strength of the the firm $i$ to the economy, $K_{in}\big(\|IN_{i,V}^h\|, \theta_{i,V}^{in}(h)\big)$, are defined respectively as

$$\theta_{i,V}^{in}(h) = \arcsin \frac{C_{ii}^h}{\|IN_{i,V}^h\|_2} \tag{5.9}$$

and

$$K_{in}\big(\|IN_{i,V}^h\|, \theta_{i,V}^{in}(h)\big) = \|IN_{i,V}^h\| \cos \theta_{i,V}^{in}(h) \tag{5.10}$$

**Definition 5.3** *Given a market network with node set $V = \{1,2,...,N\}$, the Firm-wise Market Network Relative Connectedness Structure Measure of "in" effects at horizon h, $MRC_{V_C}^{in}(h)$, is defined as following:*

$$MRC_{V_C}^{in}(h) = \frac{1}{N} \sum_{i=1}^{N} \cos \theta_{i,V}^{in}(h) \tag{5.11}$$

**Definition 5.4** *Given a market network with node set $V = \{1,2,...,N\}$, the Firm-wise Market Network Absolute Connectedness Strength Measure of "in" effects at horizon h, $MAC_{V_C}^{in}(h)$, is defined as following:*

$$MAC_{V_C}^{in}(h) = \frac{1}{N} \sum_{i=1}^{N} \|IN_{i,V}^h\| \cos \theta_{i,V}^{in}(h) \tag{5.12}$$

## 5.2. Firm-wise sector connectedness measures

An economic market can be viewed as a network of sectors. Furthermore, there is also a sub-network for each sector in an economy. In this section, we discuss the firm-wise connectedness within each sector. Without loss of simplicity, we consider an economic network with node set $V$ composed by two sectors, $V_1$ and $V_2$, where $V = \{1,...,N\}$, $V_1 = \{i_1,...,i_{n_1}\}$, $V_2 = \{j_1,...,j_{n_2}\}$,

$V_1 \cap V_2 = \emptyset$, $V_1 \cup V_2 = V$ and $n_1 + n_2 = N$. $V_1$ and $V_2$ are disjoint and complete sub-network elements of $V$.

Within a sector $V_z$, $z = 1$ or 2, our sector connectedness measures are defined in a similar manner as the firm-wise market connectedness measures in section 5.1.

**Definition 5.5** *Given a sector node set $V_z$, $z = 1, 2$, the Firm-wise Sector Relative Connectedness Structure Measure of "out" effects within sector z at horizon h, $MRC^{out}_{V_z}(h)$, is defined as following:*

$$MRC^{out}_{V_z}(h) = \frac{1}{n_z} \sum_{i=1}^{n_z} \cos\theta^{out}_{i,V_z}(h), \tag{5.13}$$

*where $n_z = |V_z|$ is the number of nodes in the sector node set $V_z = \{i_1, ..., i_{n_1}\}$, $i \in V_z$, $\theta^{out}_{i,V_z}(h) = \arcsin \frac{C^h_{ii}}{\|OUT^h_{i,V_z}\|_2}$, and $OUT^h_{i,V_z} = [C^h_{ii_1}, C^h_{ii_2}, ..., C^h_{ii_{n_1}}]'$.*

**Definition 5.6** *Given a sector node set $V_z$, $z = 1, 2$, the Firm-wise Sector Absolute Connectedness Strength Measure of "out" effects within sector z at horizon h, $MAC^{out}_{V_z}(h)$, is defined as following:*

$$MAC^{out}_{V_z}(h) = \frac{1}{n_z} \sum_{i=1}^{n_z} \|OUT^h_{i,V_z}\| \cos\theta^{out}_{i,V_z}(h) \tag{5.14}$$

*where $OUT^h_{i,V_z}$ and $\theta^{out}_{i,V_z}(h)$ are defined as above.*

**Definition 5.7** *Given a sector node set $V_z$, $z = 1, 2$, the Firm-wise Sector Relative Connectedness Structure Measure of "in" effects within sector z at horizon h, $MRC^{in}_{V_z}(h)$, is defined as following:*

$$MRC^{in}_{V_z}(h) = \frac{1}{n_z} \sum_{i=1}^{n_z} \cos\theta^{in}_{i,V_z}(h) \tag{5.15}$$

*where $n_z = |V_z|$ is the number of nodes in the sector node set $V_z = \{i_1, ..., i_{n_1}\}$, $i \in V_z$, $\theta^{in}_{i,V_z}(h) = \arcsin \frac{C^h_{ii}}{\|IN^h_{i,V_z}\|_2}$, and $IN^h_{i,V_z} = [C^h_{i_1 i}, C^h_{i_2 i}, ..., C^h_{i_{n_1} i}]'$.*

**Definition 5.8** *Given a sector node set $V_z$, $z = 1, 2$, the Firm-wise Sector Absolute Connectedness Strength Measure of "in" effects within sector z at horizon h, $MAC^{in}_{V_z}(h)$, is defined as following:*

$$MAC^{in}_{V_z}(h) = \frac{1}{n_z} \sum_{i=1}^{n_z} \|IN^h_{i,V_z}\| \cos\theta^{in}_{i,V_z}(h) \tag{5.16}$$

*where $IN^h_{i,V_z}$ and $\theta^{in}_{i,V_z}(h)$ are defined as above.*

The sector connectedness measures are just the blocked firm-wise market network connectedness measures for each sector in a firm-wise market .

**Remark 5.2** For any given $h$, if we have $C^h_{ij} = C^h_{ji} = 0$ for any $i \in V_1$ and any $j \in V_2$, then

1. $(n_1+n_2)MRC^{out}_{V_C}(h) = n_1 MRC^{out}_{V_1}(h) + n_2 MRC^{out}_{V_2}(h)$;

2. $(n_1+n_2)MAC^{out}_{V_C}(h) = n_1 MAC^{out}_{V_1}(h) + n_2 MAC^{out}_{V_2}(h)$;

3. $(n_1+n_2)MRC^{in}_{V_C}(h) = n_1 MRC^{in}_{V_1}(h) + n_2 MRC^{in}_{V_2}(h)$;

4. $(n_1+n_2)MAC^{in}_{V_C}(h) = n_1 MAC^{in}_{V_1}(h) + n_2 MAC^{in}_{V_2}(h)$.

The market network connectedness can be obtained by the sector connectedness if the sectors are the disjoint and complete decomposition elements of the market network and if there is no point-wise edge between different sectors. Intuitively speaking, the market connectedness is simply the weighted sum of sectors' connectedness when there is no causality edge between firms across different sectors.

## 5.3. Sector-wise market connectedness measures

Similar to the firm-wise market connectedness measures, the sector-wise market connectedness measures also measure market interconnectedness. But the sector-wise market connectedness measures gauge the interconnectedness among different sectors instead of different firms.

In a sector-wise market, nodes are groups of firms. We assume that any firm can only belong an unique sector. Suppose we have $M$ sectors: $V_i$ for $i = 1, 2, ..., M$. Then we have $V = \bigcup_{i=1}^{M} V_i$ and $V_S = \{V_1, V_2, ..., V_M\}$. In this case, the causality table is a $M$ by $M$ matrix. The $i$ row of the causality table, $(C_{V_iV_1}, C_{V_iV_2}, ..., C_{V_iV_M})$, measures the effects from sector $i$ to other sectors. The sector-wise market connectedness measures are defined in a similar manner as the firm-wise market connectedness measures in section 5.1.

**Definition 5.9** *The Sector-wise Market Relative Connectedness Structure Measure of "out" effects at horizon h, $MRC^{out}_{V_S}(h)$, is defined as following:*

$$MRC^{out}_{V_S}(h) = \frac{1}{M}\sum_{i=1}^{M} \cos\theta^{out}_{V_i,V_S}(h), \tag{5.17}$$

*where* $\theta^{out}_{V_i,V_S}(h) = \arcsin\frac{C^h_{V_iV_i}}{\|OUT^h_{V_i,V_S}\|_2}$, *and* $OUT^h_{V_i,V_S} = [C^h_{V_iV_1}, C^h_{V_iV_2}, ..., C^h_{V_iV_M}]'$.

**Definition 5.10** *The Sector-wise Market Absolute Connectedness Strength Measure of "out" effects at horizon h, $MAC^{out}_{V_S}(h)$, is defined as following:*

$$MAC^{out}_{V_S}(h) = \frac{1}{M}\sum_{i=1}^{M} \|OUT^h_{V_i,V_S}\| \cos\theta^{out}_{V_i,V_S}(h), \tag{5.18}$$

*where* $\theta^{out}_{V_i,V_S}(h) = \arcsin\frac{C^h_{V_iV_i}}{\|OUT^h_{V_i,V_S}\|_2}$, *and* $OUT^h_{V_i,V_S} = [C^h_{V_iV_1}, C^h_{V_iV_2}, ..., C^h_{V_iV_M}]'$.

**Definition 5.11** *The Sector-wise Market Relative Connectedness Structure Measure of "in" effects at horizon h, $MRC^{in}_{V_S}(h)$, is defined as following:*

$$MRC^{in}_{V_S}(h) = \frac{1}{M}\sum_{i=1}^{M} \cos\theta^{in}_{V_i,V_S}(h), \tag{5.19}$$

*where* $\theta^{in}_{V_i,V_S}(h) = \arcsin\frac{C^h_{V_iV_i}}{\|IN^h_{V_i,V_S}\|_2}$, *and* $IN^h_{V_i,V_S} = [C^h_{V_1V_i}, C^h_{V_2V_i}, \ldots, C^h_{V_MV_i}]'$.

**Definition 5.12** *The Sector-wise Market Absolute Connectedness Strength Measure of "in" effects at horizon h, $MAC^{in}_{V_S}(h)$, is defined as following:*

$$MAC^{in}_{V_S}(h) = \frac{1}{M}\sum_{i=1}^{M} \|IN^h_{V_i,V_S}\| \cos\theta^{in}_{V_i,V_S}(h), \tag{5.20}$$

*where* $\theta^{in}_{V_i,V_S}(h) = \arcsin\frac{C^h_{V_iV_i}}{\|IN^h_{V_i,V_S}\|_2}$, *and* $IN^h_{V_i,V_S} = [C^h_{V_1V_i}, C^h_{V_2V_i}, \ldots, C^h_{V_MV_i}]'$.

# 6. Application to implied volatility network structures

In previous sections, we have proposed a flexible network econometric measurement framework, a reliable estimation procedure designed for high-dimensional contexts and new market network connectedness measures. In this section, we illustrate the wide range of applications of our market network measurement methods by investigating a high-dimensional volatility network in the US equity market. We would like to study how the volatility network is structured and how it changes over time. Fruitful information extracted from the empirical exercises can be easily visualized by our reporting figures.

More specifically, we study the static volatility network with the full sample from 2000 to 2015 to see how firms and sectors connect to each other. We investigate the dynamics of the network structures to see how the interconnections among firms and the interconnections among sectors varied in the past 15 years. The market connectedness measures proposed in this paper are designed for measuring market systemic risk. It is a common wisdom that the systemic risk played an important role in the 2007-2009 financial crisis. Thus we examine dynamic market connectedness with our measures, and compare it with market indices (i.e. VIX index) before, during and after the crisis period. Our market connectedness measures are constructed based on the directed and weighted edges in the market network, and the superiority of the "directed" and "weighted" edges analysis against the "undirected" and "unweighted" edges analysis is demonstrated by the asymmetric effects between the financial sector and other sectors in the volatility network.

## 6.1. Data

Firms and sectors are connected with trade links or business relationships. It is an impossible mission to collect all qualitative and quantitative business information at firm-level to reveal their interconnections. As Diebold and Yilmaz (2014) argue, however, stock markets, which reflect forward-looking assessments of many thousands of smart, strategic and often privately-informed agents, provide us with feasible information that is close to the true business conditions and interconnections. For instance, there are numerous investment opportunities in the world, and using the S&P 500 index as a benchmark is almost a convention when evaluating excess returns in asset management. Therefore, we will study the crisis-sensitive volatility network in the US stock market. In addition, we are also interested in examining whether our volatility connectedness measures can reflect the underlying market systemic risk that plays an important role in the recent global financial crisis.

The volatility in stock markets is latent, so we need an volatility proxy. The well-known VIX, which has been widely accepted as a market volatility index by financial practitioners, is calculated from implied volatilities of the S&P 500 index options. It is sensitive to market turmoils. For each firm, we also exploit the information in their respective option contracts. We use implied volatility in our volatility network analysis, rather than using realized volatility estimated from stock intraday prices (see Diebold and Yilmaz (2014) and Barigozzi and Brownlees (2016)), for the quantities we are dealing with are more comparable to market indices (e.g., VIX). Similar to the VIX index known as a "market fear" index, our implied volatility network connectedness can also be viewed as "individual fear" connectedness. Volatility or implied volatility is sensitive to "terrifying news" in financial markets. For instance, the 9/11 Attacks terrify people in the stock market and leads implied volatilities to jump up rapidly. Although the 9/11 event had very minor impacts on most firms' real business conditions and their interconnections, its shocks would spill over from firms to firms and from sectors to sectors in stock markets, just because of liquidation concerns and other risk issues faced by investors. The stock implied volatilities are inevitably contaminated by shocks in financial markets since risks are traded on markets. Nevertheless, implied volatility is still an excellent proxy to study the high-dimensional market volatility network. We hope the underlying market network structure can be at least partially uncovered by its implied volatility network.

We estimate the volatility network of the S&P 100 components stocks quoted on 06/30/2015. Similar to the VIX index for the S&P 500 stock composite, in this paper the S&P 100 components[10] implied volatilities are constructed with their respective at-the-money option contracts with 30-day maturity. This implied volatility measures the expected volatility of the underlying stock over the next 30 days. We hereafter only consider the option contracts with 30-day maturity. Generally speaking, an at-the-money call (put) option usually has a delta at approximately 0.5 (-0.5). A

[10]To be included in the S&P 100, the companies should be among the larger and more stable companies in the S&P 500, and *must have list options*.

simple way to get the at-the-money implied volatility is to take the simple arithmetic mean of the interpolated implied volatility of the call option with delta 0.5 and the interpolated implied volatility of the put option with delta -0.5:

$$IV_{i,t} = \frac{1}{2}(IV_{i,t}^{C_{0.5}} + IV_{i,t}^{P_{-0.5}}). \tag{6.21}$$

where $IV_{i,t}^{C_{0.5}}$ is firm $i$'s interpolated implied volatility of the call option with delta 0.5 at time $t$, and $IV_{i,t}^{P_{-0.5}}$ is firm $i$'s interpolated implied volatility of the put option with delta -0.5 at time $t$. The data information of the daily implied volatility with different delta levels are provided in the OptionMetrics - Volatility Surface database. As the firms' implied volatilities measure the expected volatility of their stock prices over the next 30 days, the daily sequence of $\{IV_{i,t}\}_t$ is a highly persistent process. In other words, $IV_{i,t-1}$ would have a strong predictive power to forecast $IV_{i,t}$. To deviate such self-effect that merely comes from the overlapping of measuring periods, we analyze the innovation processes by taking daily first differences on each implied volatility series:

$$\Delta IV_{i,t} = IV_{i,t} - IV_{i,t-1}. \tag{6.22}$$

This manipulation procedure is simple and easy to replicate [11]. We will hereafter use $\Delta IV_{i,t}$ to estimate our implied volatility network.

The date range of the database is from 01/01/1996 to 08/31/2015. The companies whose IPO dates are after 01/01/2000 will be dropped off, such that we can examine the two most important crises in the US stock market (i.e., the IT Bubble Burst and the Financial Crisis of 2007-09). The remaining full sample is from 20/08/1999 to 31/08/2015. There are missing values on some dates for some companies and we take linear interpolations to impute the missing values to get completed time series processes for estimations. We have 90 companies in the final sample, $N = 90$. Appendix B provides the ticker symbol list of nodes and their respective sectors in our implied volatility network. The Industry Group classification for each node is from the North American Industry Groups database from MorningStar, LLC.

As Diebold and Yilmaz (2014) point out, latent market network structures may vary with business circles or may shift abruptly with market environment (e.g., crisis and noncrisis). Whether and how much it varies is ultimately an empirical matter and there is no point to just simply assume it is constant. Hence, we allow network structure to be time-varying, and thus the elements in the causality measures table are also allowed to be time-varying.[12] To capture time variations, we will

[11] Ang, Hodrick, Xing and Zhang (2006) use this manipulation approach to deal with the VIX index to test whether the VIX index is market risk factor

[12] This assumption does not contradict the constant parameters setting we made in estimating the multiple horizon causality measures. The "calendar time" for time-varying measures and the "sampling time" to estimate the measures are conceptually different. We just require the processes to be estimated are locally stationary.

estimate the dynamic implied volatility network structures with rolling samples.

Throughout the empirical exercise, We set the lag $p = 1$ and apply the VAR(1) model to approximate the unconstrained and the constrained models. Setting the same lag makes the conditional covariance to be comparable in the unconstrained models and in the constrained models. We will first estimate the static implied volatility network structure with full sample observations (20/08/1999 - 31/08/2015). As mentioned before, market connectedness can be decomposed by the connectedness within each sector and the interconnections among different sectors. Firm-wise interconnections within each sector and sector-wise interconnections are certainly of interest. To investigate the dynamic patterns of the volatility network structures, there is always a trade-off between estimation accuracies and more current conditional estimates when choosing the width of estimation windows. To examine market connectedness dynamics, we set the width of the moving window to be 2 years and update measures every one month. For example, the estimates on December 2008 are estimated based on the data from January 2007 to December 2008. By moving the estimation windows forward every month, we can obtain the dynamic pictures of the implied volatility network.

In robustness check, we compare our results with those setting the lag $p = 2$ and those using moving estimation windows of 1 year ($T = 252$), to see if our results are robust to different pre-selected modelling settings.

## 6.2. Empirical results

Market network econometric analysis can be worked under two types of network representations: i) firm-wise market structure ($V_C$), under which the nodes in the market are the 90 companies; and ii) sector-wise market structure ($V_S$) under which the nodes in the market are the 8 sectors that the 90 companies belong to. We will apply the point-wise edge analysis technique in the firm-wise market structure and apply the group-wise analysis technique in the sector-wise market structure.

### 6.2.1. Static implied volatility network structures

Firm-wise market network structures give us a broad picture of how firms connect to each other. Sub-market network structures zoom in firms' interconnections within specific sector. Sector-wise market structures merge the firms in the same sector and give us a simple picture of how different sectors connect to each other.

In Figure 5, we show the firm-wise S&P 100 implied volatility network structure. To examine this big network (90 nodes and $90^2$ edges), we only keep the directed and weighted edge ($i \rightarrow j$) if its strength is greater or equal to 90% percentile of the strengths of the edges ($i \rightarrow \cdot$) and 90% percentile of the strengths of the edges ($\cdot \rightarrow j$). In other words, we only keep an edge if and only if this edge is important to the pair of nodes being connected by it. If $i \rightarrow j$ and $j \rightarrow i$ are both kept, we only plot the one with greater strength without confusions. At the first glance, edges are denser around the firms in the financial sector. A majority of the edges being shown in the figure comes from financial firms. Moreover, the financial firms have more interconnections due to the recent financial crisis. It is also documented by Barigozzi and Brownlees (2016) in the S&P 100 realized volatility network. Interestingly, we observe that GE (a major industrial goods company) and SLB (a major supplier to the oil and gas exploration and production industry) have relatively strong interconnections with the financial firms. GE was almost bankrupt in 2009 and 2010. The oil price is very volatile in the past ten years. Figure 5 has reflected some special market situations in the US economy in the past 15 years.

We identify the 10 most influential firms in Table 2. In Table 2, we report the minimum value, the maximum value, the mean value and the quantiles (25%, 50% and 75%) of the entries in each firm's "OUT" vector, $[C_{i1}, C_{i2}, ..., C_{iN}]$. The mean for almost every firms is greater than their median and is closed to the 75% quantile; the discrepancy within the first 25% quantile is very small, but the discrepancy within the last 25% (75% - Max) is much larger. These are strong evidences of the distribution of firms' weighted edges in the "OUT" direction is right skewness. Jackson (2008) also documents right skewness distributions in social networks. We select the median, rather than the mean, to describe the central tendency of the distributions of firms' edges[13]. We sort the firms' tickers by their medians. The most influential firm in the static network is the BAC (Bank of America). BAC helps to increase the forecast precision of the next-day implied volatility by 0.07% for more than a half of the firms in the S&P 100, and by at high as 5.33% for the firm that it affects most. Seven financial firms (BAC, C, BK, AIG, MET, F, JPM and MS) are listed in the top 10 influential firms at Table 2. In Figure 5, we have seen many prominent edges are from financial firms. The firms in the financial sector have great influence in the S&P 100 network. On the other side, Table 3 reports the summary statistics of the entries in the "IN" vector, $[C_{1i}, C_{2i}, ..., C_{Ni}]$. Among the top 10 sensitive firms, only the firm (C) belongs to the financial sector and the other nine firms belong to the basic materials sector or from the Industrial goods sector. Therefore, the influential firms in the S&P 100 network are not those who will easily be affected. The "influential" and "sensitive" we mentioned so far are in the sense of direct effects, in which the causality measures are at forecast horizon $h = 1$. In Table 4, we report the top 10 influential firms at different forecast horizons, $h = 1, 2, 3, 4, 5$, to take spillover effects into account. The firms and their orders in the list of top 10 influential firms are slightly different at different forecast horizons. For instance, in the case of only taking direct effects into account ($h = 1$), the most influential financial firm is BAC and 7 out of 10 most influential firms belong to the financial sector; in the case of taking direct and indirect effects into account ($h = 5$), the most influential financial firm becomes AIG and only 4 out of 10 most influential firms is from the financial sector. The technology

[13]The firm's centrality described by our median measures is in alignment with the "Degree Centrality" in Freeman (1978) and Jackson (2008)

firms are actually influential. In the case of $h = 5$, 4 out 10 most influential firms belong to the technology sector and the top 2 influential firms are from the technology sector, if the Apple Inc. is considered as a technology firm. In short, measuring a static network that only characterizes direct effects in an economic network is far from enough to fully understand all interconnections and indirect effects. In contrast, directly measuring direct and indirect effects with the causality tables at different forecast horizons can provide us "dynamic" pictures of interconnections in the S&P 100 network with different effect-radius. In many cases, what is truly important is firm's total effect (direct effect and indirect effect) rather than just its direct effect.

Next, we zoom in the financial sector and investigate the interconnections within the financial sector. The firm-wise S&P 100 implied volatility network within the financial sector can be visualized by Figure 6. In this figure, we only keep the directed and weighted edges with the strength greater or equal to the 50% percentile of the strengths of edges in this financial network. In other words, only the "big" edges in this financial sector network will be kept. Again, if both $i \to j$ and $j \to i$ are kept, we only show the one with greater strength. We find that the most influential firms in the financial sector, in the sense of the out-degree (number of edges pointing from the firms), are the top investment banks: Morgan Stanley (MS), Goldman Sachs (GS) and Bank of America(BAC). In Table 2, Morgan Stanley and the Bank of America are both listed in the top 10 influential firms and Goldman Sachs is the 16th influential firm. The summary statistics of the entries in the "OUT" vector in the financial network in Table 5 confirms their great influence in the financial sector. Similar to the one in Table 2, the edges distributions in the financial sector are also right skewness. We again use the median to describe the central tendency of these distributions. The top 3 influential firms in the financial sector are in order as: BAC (median = 0.42), MS (median = 0.30) and GS (median = 0.25), compared with the 4th influential firms: BK (median = 0.06). Roughly speaking, we could say that the financial sector is actually controlled by the top investment banks in the past 15 years. It is also interesting to look at who are the most sensitive firms in this financial sector. In Table 6, we sort the firms by their sensitivities. The top 3 sensitive financial firms are in order as: C (median = 0.33), ALL (median = 0.32) and BAC (median = 0.28). C is the only financial firm that is listed in the top 10 sensitive firms in the S&P 100 network, and it is also the most sensitive firms in the financial sector. BAC not only is the most influential firms in the S&P 100 network, but also has strong interconnections with other firms in the financial sector since it is the most influential firm as well as the 3rd most sensitive firm in the financial sector.

Lastly, Figure 7 shows the sector-wise S&P 100 implied volatility network structure. In this network, the nodes are the sectors grouped by their respective firms as $V_i$. We only keep the directed and weighted edges with the strength greater or equal to the 50% percentile of the strengths of edges in this sector-wise network. In other words, only the "big" edges in this network will be kept. Once again, if both $V_i \to V_j$ and $V_j \to V_i$ are kept, we only show the one with greater strength. An important observation is that all sectors are strongly self-affected. It is in line with our common wisdom. Four most influential sectors, in the sense of the out-degree (number of edges pointing from the sectors), are Technology, Consumer Goods, Industrial Goods and Financial. They are also the key industries that support the growth of the US economy in these 15 years. In Table 7, we sort the sectors by their influences and obtain the top 4 influential sectors: Technology (median = 3.27), Industrial Goods (median = 1.55), Consumer Goods (median = 0.90) and Financial (median = 0.48). It is similar to what we have found in Figure 7. Moreover, Technology, Consumer Goods and Financial are also on the list of four least sensitive sectors, as reported in Table 8. Overall, the relationships among different sectors in the S&P 100 network are very asymmetric. There are two groups in this network: the influential sectors (Technology, Industrial Goods, Consumer Goods and Financial) and the sensitive sectors (Services, Basic Materials, Industrial Goods and Healthcare). Interestingly, the most influential sector in the sector-wise network (see Table 7) is the technology sector, rather than the financial sector that has the most influential firms in the frim-wise S&P 100 network found in Table 2. Note that the causality we measure is based on the marginal effect on prediction. When firms' marginal effects are small, their total (sector) margin effect is not necessarily small, especially if the component marginal effects are positive correlated. Even though the technology firms, as single components, are not as influential as the financial firms, the technology sector, as a whole, can be more influential than the financial sector. This circumstance is also discussed theoretically in the Section 3. Therefore, the group-wise network measurement technique is an important complementary for the point-wise network measurement technique to help us understand underlying market network structures.

#### 6.2.2. Connectedness dynamics in firm-wise market

In Figure 8, we show the dynamic patterns of the market relative connectedness structure measures and the market absolute connectedness strength measures in the firm-wise market structure, at forecast horizon 1, $h = 1$, and at forecast horizon 10, $h = 10$. We only report the "out" connectedness measures as the "out" measures and their respective "in" measures are highly correlated. This is not out of surprise, because one's "out" causality measures are just someone's "in" causality measures, and thus their market connectedness measures will have a similar dynamic pattern.

If our market connectedness measures are truly able to measure the market systemic risk in the US stock market, they will vary with market conditions that can be reflected by market indices like the VIX index or the S&P 500 index. The market absolute connectedness strength measures indeed have significant variations across different periods. Prior to 2007, the absolute connectedness strength measures are close to zero, while the VIX index is relatively high before 2003 due to the IT Bubble Burst. Starting from 2007, both the market connectedness strength and the VIX index start to soar and become more fluctuated at relatively high levels until 2011. This is exactly period of the recent global financial crisis. From 2011 to 2015, the market connectedness strength has a new "normal" level that is lower than the level during the crisis but higher than the level before the crisis, while the VIX index decreases to the pre-crisis level. Overall, there is an apparent synchronization between our market connectedness strength measures and the VIX index, expect in the IT Bubble Burst period. It is actually in alignment with our common wisdom that the major difference of the financial crisis of 2007-09 from other crises is the recent global financial crisis is driven and amplified by the systemic risk in financial markets. Our absolute connectedness strength measure ("individual fear connectedness") looks to be positive correlated with the "market fear" level (VIX), but our measures do concentrate more on the systemic risk that comes from the connectedness in financial markets.

Unlike the absolute connectedness strength measure, the relative connectedness structure measure concerns more about the network connectedness structure instead of the connectedness strength. We first look at the relative market connectedness structure at forecast horizon 1 and discuss it in four periods (2000-2003, 2003-2006, 2006-2009, 2009-2015). During 2000-2003, the level of the relative connectedness structure measure is relatively high (0.90-0.95) and the stock market slides due to the IT Bubble Burst. The S&P 500 gets to a bottom in early 2003 and starts to recover, and the VIX index also starts to decrease. In this period (2003-2006), the relative connectedness structure goes down. During the pre-crisis and crisis period (2006-2009), the relative connectedness structure climbs up rapidly, and touches a historical record ($> 0.95$) at the end of 2008 when is the also the most fearful moment in financial markets as shown by the VIX index touching the historical peak and the S&P 500 touching the bottom. During 2009-2015, the VIX index goes down to be normal and the S&P 500 has been fully recovered from the crisis. Interestingly, however, the relative connectedness structure measure still remains at the crisis level ($> 0.95$). Our conjecture is that the financial market is still remaining at a "crisis zone" that can be characterized by the high level of the market connectedness structure.

When comparing the relative connectedness at different forecast horizons, we find the market relative connectedness structure measures at forecast horizon 10 are much closer to the upper bound 1, than at forecast horizon 1. Note that longer forecast horizon allows every node in the network has more steps of paths to connect each other, the relative connectedness structure measure will thus be larger at greater forecast horizons. Hence, we do not expect to find big time variations for the relative connectedness structure measure at long horizons (e.g., $h = 10$), while we still can see the market connectedness structure measures at horizon 10 has a dynamic pattern similar to the one at horizon 1.

In Figure 8, the market relative connectedness structure measures and the market absolute connectedness strength measures have striking different dynamic patterns across our sample period. Absolute connectedness strength measures can be decomposed by relative connectedness structure measures and causation strengths. The difference of the relative measures and the absolute measures is totally accounted by the time-varying causation strengths. By comparing these two types of measures at different periods, we find the causation strengths are relative large during the financial crisis. It again confirms our assertion that our causality measures can capture elements of the market systemic risk.

Also, we provide the 90% bootstrap confidence intervals for the absolute market connectedness strength measures on some specific dates[14] (2004-01-20, 2005-01-20, 2006-01-20, 2007-01-20, 2008-01-20, 2009-01-20, 2010-01-20, 2011-01-20, 2012-01-20, 2013-01-20 and 2014-01-20) in Figure 9. We use the bootstrapping procedure that is similar to the one described in Dufour and Zhang (2015). The raise of the market absolute connectedness strength during the financial crisis period is statically significant.

While our market connectedness measures do show dynamic patterns corresponding to different major market conditions (before crisis, during crisis, and after crisis), it still seems to be counterintuitive that our dynamic connectedness measures are "too volatile". For instance, one may find the market connectedness strength measures jump up and down frequently[15], but the underlying market structures has no way to change at this rate even though the market structures may change abruptly because of crisis. In fact, our estimated implied volatility network not only measures the underlying market structures, but also captures the market effects in the stock market and in the option market. As we have discussed before, firm's implied volatility is sensitive to special events in financial markets. One of the regular important events in the equity market is the quarter earnings announcements. Publicly-traded companies have to release their earning reports every three months regarding their financial conditions, earning forecasts, etc. It means the firm's detail information is only renewed to the public every three months. This kind information is crucial for firm's credit grade and firm's stock price target evaluated by equity analysts in the market. If an earning report beat market expectations, the firm's stock price could jump up overnight and vice versa. As a result, option trading will become much more active during earnings seasons, and thus the implied volatilities are usual more volatile during this period. Moreover, different firms could release their earnings reports on different dates during a earnings season. Some investors would bet on some companies base on others' released performances, especially when these firms are in the same sector where they face a similar business environment. The high leverage and large possible payoffs of the option trading make a large proportion of active investors choose to bet on the option market[16]. Therefore, it is very likely that the connectedness measures of the implied volatility network would become more volatile during earnings seasons. It is mainly due to shocks in the financial market, rather than changes in the underlying market structures. Thus, the dynamics of our implied volatility market connectedness measures can be decomposed by long-run stable market connectedness changes and short-run financial fluctuations, and this is exactly what

[14]We do not report the confidence intervals every month in our sample period because the bootstrapping procedure is time costly.

[15]From 2007 to 2011, for example, we find about 10 spikes in the figure.

[16]Donders, Kouwenberg, Vorst et al. (2000) find firm's implied volatility increases before announcement days and drops afterwards.

we observe in Figure 8.

#### 6.2.3. Connectedness dynamics within single sector in firm-wise market

Taking the diagonal block that contains companies in a sector in the firm-wise market causality measures table, we have the sub-network structure for this sector. We do so for each sector, and then obtain the sector connectedness measures within every single sector in the firm-wise market.

Figure 10 reports the absolute connectedness strength measures within each of the 7 sectors in our implied volatility network[17]. As expected, the financial sector has the highest and the most persistent absolute connectedness strengths during the financial crisis. Other sectors also have higher connectedness strengths in this period, but they are very minor when compared with the financial sector. During the crisis, investors would be more sensitive to news comings, so the implied volatility connectedness could become more fluctuated. Since financial shocks (e.g., quarter earnings releases) to implied volatilities are more easily to spill over within a sector, at most of the times when there are major spikes in the market connectedness strengths in Figure 8, we can find their corresponding ones in one of the sector connectedness strengths in Figure 10.

#### 6.2.4. Connectedness dynamics in sector-wise market

As has been emphasized before, the econometric framework proposed in this paper provides the first unified method to estimate point-wise effects and group-wise effects. The nodes in the sector-wise network structure ($V_S$) in this empirical exercise are the 8 sectors[18] that the S&P 100 components belong to.

In Figure 11, we report the dynamic patterns of the market absolute connectedness strengths and the market relative connectedness structures in the sector-wise market network at forecast horizon 1 ($h = 1$) and at forecast horizon 10 ($h = 10$). The sector-wise market absolute connectedness strength measure at forecast horizon 1 has a sharp peak at the end of 2008. However, it does not persistently remain at a high level compared with the absolute connectedness strength measures in the company-wise market structure during the crisis period shown in Figure 8. In other words, even if there is a high persistent market systemic risk during the financial crisis, it is not due to the connectedness among different sectors. The relative connectedness structure measures and the absolute connectedness strength measures are positively correlated before 2009, while again, the market connectedness structure does not decrease with the market connectedness strength after the crisis. The sector-wise absolute connectedness structure strengths in Figure 11 are generally lower than the firm-wise strengths in Figure 8. It is because sector-wise nodes have weaker interconnections than firm-wise nodes in an economic network. Connected firms usually have closed business relationships and they tend to be in the same sector.

[17]The "Utility" industry is not included as it only contains one company.

[18]Basic Materials, Consumer Goods, Financial, Healthcare, Industrial Goods, Services, Technology and Utilities

### 6.2.5. Directed and weighted edges dynamics in sector-wise market

We now look at the sector-wise network interconnections in more details. In particular, we concentrate on the financial sector, which has the greatest influence on the US stock market in the past 10 years due to the global financial crisis, to see how the interconnections between the financial sector and other sectors look like.

In Figure 12 and Figure 13, we report the time-varying direct effects from the financial sector to other sectors and from other sectors to the financial sectors. As expected, the financial sector have the strongest influence on itself during the financial crisis. At the end of 2008, the magnitude of the financial sector affecting itself soars to a historical peak with the VIX index soaring, and it keep at a relatively high level until 2011. We find the financial sector has an extremely strong effect on itself from 2009 to 2011, which matches the crisis period of financial crisis of 2007-09 as our estimates utilize 2 years rolling samples. The financial crisis is actually not yet over in the global financial market after 2009. Fore instance, the US financial crisis triggers the European debt crisis in early 2010. From 2011 to 2013, the financial sector still has a relative strong effect on itself.

The effect from the financial sector on other sectors also increase at the beginning of the crisis, while the raise only lasts for a few months. In contrast, all other sectors only have negligible effects on the financial sector, compared with the striking magnitude of the financial sector affecting to itself. The effects between the financial sector and other sectors are quite asymmetric: the financial sector has a strong effect on others but the reverse is not. The asymmetry and the time variations in effects between the financial sector and other sectors confirm the importance of directed and weighted edges setting in economic network analysis.

## 6.3. Robustness check

Finally, we conclude this section with checking the robustness of our market connectedness results to the choice of lags $p$ in the VAR($p$) approximation to the causality estimation models and to the width of the estimation sample windows. In fact, different lags, $p$, correspond to different information sets using in causality estimations; different widths of estimation windows correspond to different sample market conditions. The estimates of a given edges will change with different choices of them. Therefore, we do not expect our estimated connectedness measures would be invariant to different lags and to different widths of estimation windows. Instead, if the underlying market systemic risk in the volatility network can truly be measured by our market connectedness measures, the measures, under different pre-selected model settings, should have similar dynamic patterns over time, following the changes in the underlying market systemic risk.

In particular, we compare our estimated results with those estimated with VAR(2) models and with those estimated with 1-year estimation windows. Figure 14 reports the market absolute connectedness strength measures under three different model settings: i) VAR(1) and 2-year estimation windows (Benchmark); ii) VAR(2) and 2-year estimation windows; and iii) VAR(1) and 1-year estimation windows. They are estimated at forecast horizon 1 and at forecast horizon 10 in the firm-wise market network. They all have a similar dynamic pattern (low before 2007, soar up from 2008, resume pre-crisis level in 2011 and has a mild increasing trend from 2012 to present). Figure 15 shows the robustness of the relative connectedness structure measures with the same model settings comparison. All of the three relative connectedness structure measures at forecast horizon 1 have a similar dynamic pattern (relatively high from 2001 to 2003, decline from 2003 to 2006, soar up from 2006 to 2009 and remain at the financial crisis level from 2009 to present). For the connectedness structure measures at forecast horizon 10, they keep at a high level all the time.

To summarize, our robustness check shows that the time-varying characteristics of our market connectedness measures are robust to the choices of $p$ in the VAR($p$) approximation, and also robust to the choices of the widths of the estimation windows.

# 7. Conclusion

Economic and financial network analysis requires a well developed time series econometric framework for empirical studies. Less restrictions on network settings, less assumptions on the time series identification models and more empirical flexibility of the measurement framework would be favoured. In this paper, we propose a novel time series econometric method to measure high-dimensional directed and weighted market network structures. Direct and spillover effects at multiple horizons, between nodes and between groups, are measured in a unified framework. We argue that a satisfactory network econometric framework to study market networks should be able to estimate directed and weighted network structures with causality implications, and it can be applied to study network spillover effects in a high-dimensional context. Indeed, our network estimation method not only satisfies all these criteria, but also enjoys other appealing features.

We measure causality at different horizons in a network through the multiple horizon causality measures based on flexible VAR models specified by the LASSO approach. (Non-sparse) network structures can be estimated from a sparse set of autoregressive coefficients and concentration matrices. Asymptotic consistency results of the estimators of our directed and weighted edges measures are also provided in this paper. We do not require sparsity assumptions on network structures or the Gaussian assumption on econometric models. We successfully connect the causality literature with the LASSO approach in application to economic and financial network measurement. Moreover, to the best of our knowledge, our econometric framework is the first one, in the network econometric literature, to explicitly allow point-wise edges (relationships between firms) and group-wise edges (relationships between sectors) to be measured in a unified framework.

With this framework at hand, we also provide the estimated market network with new connectedness measures that are built upon the underlying network structures. Since an economic network can be viewed as a network among firms as well as a network among sectors, we propose three types of connectedness measures to gauge network interconnections. These types of connectedness measures fully take advantage of the flexibility of our network measurement method, so they can be applied to study market network connectedness in flexible ways

Our network measurement methods have a wide range of applications and can be applied in a variety of research areas, including identifying and quantifying economic relationships between firms, between sectors and between areas; measuring market connectedness; predicting financial risks; guiding asset allocations in large portfolios; etc. Note that many latent economic and financial network structures can be estimated by our flexible network measurement method with varieties of panel databases. Specifically, observing that explicit identified economic network centrality and consumer-supplier linkage have been shown to be new risk factors in asset pricing and new determinants to predict financial variables, we expect more pricing factors and financial and macroeconomic variables drivers are to be discovered by our network econometric measurement methods.

To illustrate the usefulness of our method in network analysis, we investigate the S&P 100 implied volatility network in the US stock market, which can be viewed as a "individual fear" network and has not yet been studied in existing literature. We find that: i) 7 out of the 10 most influential firms in the S&P 100 belong to the financial sector, and top investment banks (Morgan Stanley, Goldman Sachs and Bank of America) have the greatest influence in the financial sector; ii) market connectedness was especially strong during the recent global financial crisis; iii) the high market connectedness was mainly due to the high connectedness within the financial sector and the spillovers from the financial sector to other sectors; iv) the financial sector had the highest firm-wise connectedness from 2008 to 2010, while the connectedness of other sectors also reach relatively high level during this period; v) the causality effects between the financial sector and other sectors were asymmetric and displayed considerable variation over time, which stresses the importance of directed and weighted edges settings in market network analysis.

# Appendix

# A. Proofs

We apply an assumptions set that is similar to the one using in Barigozzi and Brownlees (2014). The proofs of the Proposition **4.1** and the Proposition **4.2** can thus follow their results. The proof of Theorem **4.3** is based on Proposition **4.1** and Proposition **4.2**.

## A.1. Assumptions

1. The N-dimensional random vector process $X(t)$ is non-deterministic, has zero mean, and is covariance stationary. Moreover,

   (a) here exist constants $M_1$ and $M_2$ such that for each $N$, $0 < M_1 < \mu_{\min}(\Gamma_X) \leq \mu_{\max}(\Gamma_X) < M_2 < \infty$, where $\Gamma_X$ is the covariance matrix of $X$ and $\mu_{\min}(\cdot)$ and $\mu_{\max}(\cdot)$ are the smallest and the largest eigenvalues operators respectively.

   (b) there exists constants $M_3(\omega)$ and $M_4(\omega)$ such that, for each $N$ and for any $\omega \in [-\pi, \pi]$, we have $0 < M_3(\omega) \leq \mu_{\min}(s_X(\omega)) \leq \mu_{\max}(s_X(\omega)) \leq M_4(\omega) < \infty$, where $s_X(\omega)$ is the spectral density matrix of (4.1).

   (c) define $\beta = \sup\{c : \sum_{h=1}^{\infty} h^c \sup_{i,j} |E[X(t)_i X(t-h)_j]|\}$, then $\beta > 0$.

   (d) the process has three representation forms (4.1), (4.2) and (4.3) as stated in Assumption 4.1.

2. There exist constants $c_1 > 0$ and $c_2 > 0$ such that $N = O(T^{c_1})$ and $p = O(T^{c_2})$. $\beta > \frac{4c_1}{c_2}$. The dimension of the two parties $W$ and $Y$ of analysis, $m_1$ and $m_2$, is fixed.

3. (a) The set of nonzero entries in $\alpha_i$, $\mathscr{A}_i$, has $q_{Ti}^{\mathscr{A}}$ elements, and $q_{Ti}^{\mathscr{A}}$ satisfies the following conditions:
   $q_{Ti}^{\mathscr{A}} = o\left(\sqrt{\frac{T}{\log T}}\right)$, $\frac{\lambda_T}{T}\sqrt{q_{Ti}^{\mathscr{A}}} = o(1)$, $\lim_{T\to\infty} \frac{\lambda_T}{T}\sqrt{\frac{T}{\log T}} = \infty$, $\sqrt{\frac{q_{Ti}^{\mathscr{A}} \log T}{T}} = o\left(\frac{\lambda_T}{T}\right)$ and $\frac{\lambda_T}{T^{1-c_1}}\sqrt{q_{Ti}^{\mathscr{A}}} = O(1)$ for $i = 1, \ldots, N$.

   (b) The set of nonzero entries in $\bar{\alpha}_i$, $\bar{\mathscr{A}_i}$, has $q_{Ti}^{\bar{\mathscr{A}}}$ elements, and $q_{Ti}^{\bar{\mathscr{A}}}$ satisfies the following conditions:
   $q_{Ti}^{\bar{\mathscr{A}}} = o\left(\sqrt{\frac{T}{\log T}}\right)$, $\frac{\lambda_T}{T}\sqrt{q_{Ti}^{\bar{\mathscr{A}}}} = o(1)$, $\lim_{T\to\infty} \frac{\lambda_T}{T}\sqrt{\frac{T}{\log T}} = \infty$, $\sqrt{\frac{q_{Ti}^{\bar{\mathscr{A}}} \log T}{T}} = o\left(\frac{\lambda_T}{T}\right)$ and $\frac{\lambda_T}{T^{1-c_1}}\sqrt{q_{Ti}^{\bar{\mathscr{A}}}} = O(1)$ for $i = 1, \ldots, N - m_1$.

4. (a) For all $i = 1, \ldots, N$, there exists a sequence of positive real numbers $\{s_{Ti}^{\mathscr{A}}\}$ such that $|\alpha_{ij}| > s_{Ti}^{\mathscr{A}}$ and $\lim_{T\to\infty} \frac{s_{Ti}^{\mathscr{A}}}{\frac{\lambda_T}{T}\sqrt{q_{Ti}^{\mathscr{A}}}} = \infty$ for all $\alpha_{ij} \in \mathscr{A}_i$.

(b) For all $i = 1, ..., N - m_1$, there exists a sequence of positive real numbers $\{s_{Ti}^{\bar{\mathscr{A}}}\}$ such that $|\alpha_{ij}| > s_{Ti}^{\bar{\mathscr{A}}}$ and $\lim_{T\to\infty} \frac{s_{Ti}^{\bar{\mathscr{A}}}}{\frac{\lambda_T}{T}\sqrt{q_{Ti}^{\bar{\mathscr{A}}}}} = \infty$ for all $\bar{\alpha}_{ij} \in \bar{\mathscr{A}_i}$.

5. (a) For each $i = 1, ..., N$, $|\hat{\alpha}_{Tij}^{LASSO} - \alpha_{ij}| = O_p(T^{-\theta})$ with $\theta \in [\frac{1}{4}, \frac{1}{2}]$ for any $j = 1, ..., N$.

   (b) For each $i = 1, ..., N$, $|\hat{\bar{\alpha}}_{Tij}^{LASSO} - \bar{\alpha}_{ij}| = O_p(T^{-\theta})$ with $\theta \in [\frac{1}{4}, \frac{1}{2}]$ for any $j = 1, ..., N$.

6. (a) The set of nonzero entries in $\rho^u$, $\mathscr{Q}_u$, has $q_T^{\mathscr{Q}_u}$ elements, and $q_T^{\mathscr{Q}_u}$ satisfies the following conditions:
$q_T^{\mathscr{Q}_u} = o\left(\sqrt{\frac{T}{\log T}}\right)$, $\frac{\gamma_T}{T}\sqrt{q_T^{\mathscr{Q}_u}} = o(1)$, $\lim_{T\to\infty} \frac{\gamma_T}{T}\sqrt{\frac{T}{\log T}} = \infty$ and $\sqrt{\frac{q_T^{\mathscr{Q}_u} \log T}{T}} = o\left(\frac{\gamma_T}{T}\right)$.

   (b) The set of nonzero entries in $\rho^v$, $\mathscr{Q}_v$, has $q_T^{\mathscr{Q}_v}$ elements, and $q_T^{\mathscr{Q}_v}$ satisfies the following conditions:
$q_T^{\mathscr{Q}_v} = o\left(\sqrt{\frac{T}{\log T}}\right)$, $\frac{\gamma_T}{T}\sqrt{q_T^{\mathscr{Q}_v}} = o(1)$, $\lim_{T\to\infty} \frac{\gamma_T}{T}\sqrt{\frac{T}{\log T}} = \infty$ and $\sqrt{\frac{q_T^{\mathscr{Q}_v} \log T}{T}} = o\left(\frac{\gamma_T}{T}\right)$.

7. (a) For all $\rho_{ij}^u \in \mathscr{Q}_u$, there exists a sequence of positive real numbers $\{s_T^{\mathscr{Q}_u}\}$ such that $|\rho_{ij}^u| > s_T^{\mathscr{Q}_u}$ and $\lim_{T\to\infty} \frac{s_T^{\mathscr{Q}_u}}{\frac{\gamma_T}{T}\sqrt{q_T^{\mathscr{Q}_u}}} = \infty$.

   (b) For all $\rho_{ij}^v \in \mathscr{Q}_v$, there exists a sequence of positive real numbers $\{s_T^{\mathscr{Q}_v}\}$ such that $|\rho_{ij}^v| > s_T^{\mathscr{Q}_v}$ and $\lim_{T\to\infty} \frac{s_T^{\mathscr{Q}_v}}{\frac{\gamma_T}{T}\sqrt{q_T^{\mathscr{Q}_v}}} = \infty$.

8. (a) Let $D_t^u$ be a $\frac{N(N-1)}{2} \times 1$ vector such that it has generic component $d_{tij}^u = \sqrt{\frac{\hat{s}_{Tii}^u}{\hat{s}_{Tjj}^u}}\hat{u}_{ti}$ and let $\Gamma_D = E[(D_t^u)' D_t^u]$, then there exists a constant $M_u < 1$ such that for any $\rho_{ij}^u \in \mathscr{Q}_u^C$, $|\Gamma''_{Dij\mathscr{Q}_u}(\rho^u)[\Gamma''_{D\mathscr{Q}_u\mathscr{Q}_u}(\rho^u)]^{-1} sign(\rho_{\mathscr{Q}_u}^u))| < M_u$, where $\Gamma''_{Dijsq}(\rho^u) := \frac{\partial^2 \Gamma_D}{\partial d_{tij}^u \partial d_{tsq}^u}|_{d_{tij}^u = \rho_{ij}^u, d_{tsq}^u = \rho_{sq}^u}$.

   (b) Let $D_t^v$ be a $\frac{N(N-1)}{2} \times 1$ vector such that it has generic component $d_{tij}^v = \sqrt{\frac{\hat{s}_{Tii}^v}{\hat{s}_{Tjj}^v}}\hat{v}_{ti}$ and let $\Gamma_D = E[(D_t^v)' D_t^v]$, then there exists a constant $M_v < 1$ such that for any $\rho_{ij}^u \in \mathscr{Q}_v^C$, $|\Gamma''_{Dij\mathscr{Q}_v}(\rho^v)[\Gamma''_{D\mathscr{Q}_v\mathscr{Q}_v}(\rho^v)]^{-1} sign(\rho_{\mathscr{Q}_v}^v))| < M_u$, where $\Gamma''_{Dijsq}(\rho^v) := \frac{\partial^2 \Gamma_D}{\partial d_{tij}^v \partial d_{tsq}^v}|_{d_{tij}^v = \rho_{ij}^v, d_{tsq}^v = \rho_{sq}^v}$.

9. (a) For any $\delta > 0$, there exists a constant $K$ such that for $T$ large enough, we have $P\left(\max_{1\le i\le N_T} |\hat{s}_{Tii}^u - s_{ii}^u| \le K\sqrt{\frac{\log T}{T}}\right) \ge 1 - O(T^{-\delta})$.

   (b) For any $\delta > 0$, there exists a constant $K$ such that for $T$ large enough, we have $P\left(\max_{1\le i\le N_T} |\hat{s}_{Tii}^v - s_{ii}^v| \le K\sqrt{\frac{\log T}{T}}\right) \ge 1 - O(T^{-\delta})$.

## A.2. Proof of the Proposition 4.1

Under assumption 5a, the weighted penalty $w_{Tij} = \frac{1}{|\hat{\alpha}_{Tij}^{LASSO}|}$ in (4.13) satisfies the condition 1 for the pre-estimator in Barigozzi and Brownlees (2014). Then under assumptions 1, 2, 3a, 4a and 5a, using the result in the Theorem 1 in Barigozzi and Brownlees (2014), we have

1. for $T$ large enough and for any $\delta > 0$, $\hat{\alpha}_{Ti} = 0$ for $\alpha_i \in \mathscr{A}_i^C$ with at least probability $1 - O(T^{-\delta})$, where $\hat{\alpha}_{Ti}$ is defined in (4.13), and

2. for $T$ large enough and for any $\delta > 0$, there exist a constant $\kappa_u$ such that $\|\hat{\alpha}_{Ti} - \alpha_i\|_2 \leq \kappa_u \frac{\lambda_T}{T}\sqrt{q_{Ti}^{\mathscr{A}}}$ with at least probability $1 - O(T^{-\delta})$.

From assumption 3a, we know $\frac{\lambda_T}{T}\sqrt{q_{Ti}^{\mathscr{A}}} = o(1)$. Thus we have Prob$\{\hat{\alpha}_{Tij} = 0 \text{ if } \alpha_{ij} \in \mathscr{A}_i^C\} \to 1$ and $\hat{\alpha}_{Ti} \xrightarrow{p} \alpha_i$ for $i = 1,...,N$. Note also that $\text{vec}(\alpha_1',...,\alpha_N') = \text{vec}([A_1^p, A_2^p, ..., A_p^p]')$, and by the Lemma 2 in Barigozzi and Brownlees (2014) the truncated bias $\|A_k^p - A_k\|_\infty = o(1)$. Therefore, $\hat{A}_{Tk}^p \xrightarrow{p} A_k$ for $k = 1,...,p$ .

Similarly, for the expanded restricted process, under the assumptions 1, 2, 3b, 4b and 5b, we have Prob$\{\hat{\bar{\alpha}}_{Tij} = 0 \text{ if } \bar{\alpha}_{ij} \in \bar{\mathscr{A}}_i^C\} \to 1$, $\hat{\bar{\alpha}}_{Ti} \xrightarrow{p} \bar{\alpha}_i$ for $i = 1,...,N$ and thus $\hat{\bar{A}}_{Tk}^p \xrightarrow{p} \bar{A}_k^\phi$ for $k = 1,...,p$. □

## A.3. Proof of the Proposition 4.2

For the $\hat{\rho}_T^u$ considered in (4.19), under the assumptions 1, 2, 3a, 4a, 5a, 6a, 7a, 8a and 9a, using the result in Theorem 2 in Barigozzi and Brownlees (2014), we have

1. for $T$ large enough and for any $\delta > 0$, $\hat{\rho}_{Tij}^u = 0$ for $\rho_{ij}^u \in \mathscr{Q}_u^C$ with at least probability $1 - O(T^{-\delta})$, and

2. for $T$ large enough and for any $\delta > 0$, there exists a constant $\kappa_q$ such that $\|\hat{\rho}_T^u - \rho^u\|_2 \leq \kappa_q \frac{\gamma_T}{T}\sqrt{q_T^{\mathscr{Q}_u}}$, or equivalently, $\|\hat{S}_T^u - S^u\| \leq \kappa_q \frac{\gamma_T}{T}\sqrt{q_T^{\mathscr{Q}_u}}$with at least probability $1 - O(T^{-\delta})$

For assumption 6a, we know $\frac{\gamma_T}{T}\sqrt{q_T^{\mathscr{Q}_u}} = o(1)$. Then we have Prob$\{\hat{\rho}_{Tij}^u = 0 \text{ if } \rho_{ij}^u \in \mathscr{Q}_u^C\ \} \to 1$ and $\hat{\rho}_{Tij}^u \xrightarrow{p} \rho_{ij}^u$ for $i, j = 1,...,N$. Therefore, we also have $\hat{S}_T^u \xrightarrow{p} S_u \equiv \Sigma_u^{-1}$.

Similarly, for the $\hat{\rho}_T^v$ considered in (4.20), under the assumptions 1, 2, 3b, 4b, 5b, 6b, 7b, 8b and 9b, we have Prob$\{\hat{\rho}_{Tij}^v = 0 \text{ if } \rho_{ij}^v \in \mathscr{Q}_v^C\ \} \to 1$ and $\hat{\rho}_{Tij}^v \xrightarrow{p} \rho_{ij}^v$ for $i, j = 1,...,N$. Therefore, we also have $\hat{S}_T^v \xrightarrow{p} S_v \equiv \Sigma_v^{-1}$. □

## A.4. Proof of the Theorem 4.3

Under the assumptions 1, 2, 3a, 3b, 4a, 4b, 5a, 5b, 6a, 6b, 7a, 7b, 8a, 8b, 9a, and 9b, which have been used in Proposition **4.1** and **4.2**, and by these propositions, we have the consistent estimators, $\hat{A}^p_{Tk}$, $\hat{\bar{A}}^p_{Tk}$, $\hat{S}^u_T$, $\hat{S}^v_T$ for $A_k$, $\bar{A}^\phi_k$, $S^u$, $S^v$ respectively.

Note that from the Remark 4.1 and from the Remark 4.2, we have

1. The covariance matrix of the forecast error at horizon $h$ for the unrestricted model is

$$\Sigma[X(t+h)|\mathscr{F}(t)] = \sum_{q=0}^{h-1} \varphi_q \Sigma_u \varphi_q', \tag{A.1}$$

   where $\varphi_q = \sum_{k=1}^{q} A_k \varphi_{q-k}$ and $\varphi_0 = I_N$.

2. The covariance matrix of the forecast error at horizon $h$ for the restricted model is

$$\Sigma[X_0(t+h)|\mathscr{F}_{-W}(t)] = \sum_{q=0}^{h-1} \bar{\varphi}_q \Sigma_\varepsilon \bar{\varphi}_q', \tag{A.2}$$

   where $\bar{\varphi}_q = \sum_{k=1}^{q} \bar{A}_k \bar{\varphi}_{q-k}$ and $\bar{\varphi}_0 = I_{N-m_1}$

3. The forecast error covariance of $X^W$, without its past information, at horizon $h$ is

$$\Sigma_W[X^W(t+h)|\mathscr{F}_{-W}(t)] = J_3 \left( \sum_{q=0}^{h-1} \phi_q \Sigma_v \phi_q' \right) J_3', \tag{A.3}$$

   where $\phi_q = \sum_{k=1}^{q} A^\phi_k \phi_{q-k}$, $A^\phi_k = \bar{A}^\phi_k J_2$, $\phi_0 = I_N$, $J_3 = [I_{m_1 \times m_1},\, 0_{m_1 \times (N-m_1)}]_{m_1 \times N}$.

4. $\Sigma_\varepsilon = J_2 \Sigma_v J_2'$ and $\bar{A}_k = (J_2 \bar{A}^\phi_k)'$, where $J_2 = [0_{(N-m_1)\times m_1},\, I_{(N-m_1)\times(N-m_1)}]_{(N-m_1)\times N}$

As $\hat{A}^p_{Tk} \xrightarrow{p} A_k$ and $(\hat{S}^u_T)^{-1} \xrightarrow{p} (S^u)^{-1} = \Sigma_u$, $\hat{\varphi}_q$ is iteratively defined as $\hat{\varphi}_q = \sum_{k=1}^{q} \hat{A}^p_{Tk} \hat{\varphi}_{q-k}$ for $q = 1, \ldots, h-1$, then $\hat{\varphi}_q \xrightarrow{p} \varphi_q$ and thus

$$\hat{\Sigma}[X(t+h)|\mathscr{F}_{-W}(t)] \xrightarrow{p} \Sigma[X(t+h)|\mathscr{F}_{-W}(t)], \tag{A.4}$$

where $\hat{\Sigma}[X(t+h)|\mathscr{F}_{-W}(t)] := \sum_{q=0}^{h-1} \hat{\varphi}_q (\hat{S}^u_T)^{-1} \hat{\varphi}_q'$ and $\hat{\varphi}_0 = I_N$.

As $\hat{\bar{A}}^p_{Tk} \xrightarrow{p} \bar{A}^\phi_k$ and $(\hat{S}^v_T)^{-1} \xrightarrow{p} (S^v)^{-1} = \Sigma_v$, $\hat{\phi}_q$ is iteratively defined as $\hat{\phi}_q = \sum_{k=1}^{q} (\hat{\bar{A}}^p_{Tk} J_2) \hat{\phi}_{q-k}$ for $q = 1, \ldots, h-1$, then $\hat{\phi}_q \xrightarrow{p} \phi_q = \sum_{k=1}^{q} (\bar{A}^\phi_k J_2) \phi_{q-k}$, and thus

$$\hat{\Sigma}_W[X^W(t+h)|\mathscr{F}_{-W}(t)] \xrightarrow{p} \Sigma_W[X^W(t+h)|\mathscr{F}_{-W}(t)], \tag{A.5}$$

where $\hat{\Sigma}_W[X^W(t+h)|\mathscr{F}_{-W}(t)] := J_3 \left( \sum_{q=0}^{h-1} \hat{\phi}_q (\hat{S}^v_T)^{-1} \hat{\phi}_q' \right) J_3'$ and $\hat{\phi}_0 = I_N$.

As $\hat{\bar{A}}^{p}_{Tk} \xrightarrow{p} \bar{A}^{\phi}_{k}$, $(\hat{S}^{\nu}_{T})^{-1} \xrightarrow{p} (S^{\nu})^{-1} = \Sigma_{\nu}$, $\Sigma_{\varepsilon} = J_2 \Sigma_{\nu} J_2'$ and $\bar{A}_k = (J_2 \bar{A}^{\phi}_{k})'$, then $\hat{\Sigma}_{\varepsilon} \xrightarrow{p} \Sigma_{\varepsilon}$ and $(J_2 \hat{\bar{A}}^{p}_{Tk})' \xrightarrow{p} (J_2 \bar{A}^{\phi}_{k})'$, where $\hat{\Sigma}_{\varepsilon} = J_2 (\hat{S}^{\nu}_{T})^{-1} J_2'$. Also $\hat{\bar{\varphi}}_q$ is iteratively defined as $\hat{\bar{\varphi}}_q = \sum_{k=1}^{q} (J_2 \hat{\bar{A}}^{p}_{Tk})' \hat{\bar{\varphi}}_{q-k}$, then $\hat{\bar{\varphi}}_q \xrightarrow{p} \bar{\varphi}_q = \sum_{k=1}^{q} \bar{A}_k \bar{\varphi}_{q-k}$, and thus

$$\hat{\Sigma}[X_0(t+h)|\mathscr{F}_{-W}(t)] \xrightarrow{p} \Sigma[X_0(t+h)|\mathscr{F}_{-W}(t)], \tag{A.6}$$

where $\hat{\Sigma}[X_0(t+h)|\mathscr{F}_{-W}(t)] := \sum_{q=0}^{h-1} \hat{\bar{\varphi}}_q \hat{\Sigma}_{\varepsilon} \hat{\bar{\varphi}}_q'$ and $\hat{\bar{\varphi}}_0 = I_{N-m_1}$.

Finally, we have

$$\begin{aligned} \hat{C}^{h}_{TWY} &= \ln \left[ \frac{\det\{J_0 \hat{\Sigma}[X_0(t+h)|\mathscr{F}_{-W}(t)] J_0'\}}{\det\{J_1 \hat{\Sigma}[X(t+h)|\mathscr{F}(t)] J_1'\}} \right] \\ &\xrightarrow{p} \ln \left[ \frac{\det\{J_0 \Sigma[X_0(t+h)|\mathscr{F}_{-W}(t)] J_0'\}}{\det\{J_1 \Sigma[X(t+h)|\mathscr{F}(t)] J_1'\}} \right] \end{aligned}$$

and

$$\begin{aligned} \hat{C}^{h}_{TWW} &= \ln \left[ \frac{\det\{\hat{\Sigma}_W[X^W(t+h)|\mathscr{F}_{-W}(t)]\}}{\det\{J_1 \hat{\Sigma}[X(t+h)|\mathscr{F}(t)] J_1'\}} \right] \\ &\xrightarrow{p} \ln \left[ \frac{\det\{\Sigma_W[X^W(t+h)|\mathscr{F}_{-W}(t)]\}}{\det\{J_1 \Sigma[X(t+h)|\mathscr{F}(t)] J_1'\}} \right]. \end{aligned}$$

Therefore,

$$\hat{C}^{h}_{TWY} \xrightarrow{p} C_L(X^W \underset{h}{\longrightarrow} X^Y | I), \tag{A.7}$$

$$\hat{C}^{h}_{TWW} \xrightarrow{p} C_L(X^W \underset{h}{\longrightarrow} X^W | I), \tag{A.8}$$

□

# B. S&P 100 components (selected)

| Ticker | Company | Sector | Ticker | Company | Sector |
|---|---|---|---|---|---|
| AAPL | Apple Inc. | Consumer Goods | HPQ | Hewlett-Packard Co | Technology |
| ABT | Abbott Laboratories | Healthcare | IBM | Intl Business Machines Corp | Technology |
| ACN | Accenture plc | Technology | INTC | Intel Corp | Technology |
| AGN | Allergan plc | Healthcare | JNJ | Johnson & Johnson | Healthcare |
| AIG | American Intl Group Inc | Financial | JPM | JP Morgan Chase & Co | Financial |
| ALL | Allstate Corp | Financial | KO | Coca-Cola Co | Consumer Goods |
| AMGN | Amgen Inc | Healthcare | LLY | Lilly Eli & Co | Healthcare |
| AMZN | Amazon.com Inc | Services | LMT | Lockheed Martin | Industrial Goods |
| APC | Anadarko Petroleum Corp | Basic Materials | LOW | Lowe's Cos Inc | Services |
| AXP | American Express Co | Financial | MCD | McDonald's Corp | Services |
| BA | Boeing Co | Industrial Goods | MDT | Medtronic plc | Healthcare |
| BAC | Bank of America Corp | Financial | MET | Metlife Inc | Financial |
| BAX | Baxter Intl Inc | Healthcare | MMM | 3M Co | Industrial Goods |
| BIIB | Biogen Inc | Healthcare | MO | Altria Group Inc | Consumer Goods |
| BK | The Bank of New York Mellon Corp | Financial | MON | Monsanto Co. | Basic Materials |
| BMY | Bristol-Myers Squibb | Healthcare | MRK | Merck & Co Inc | Healthcare |
| C | Citigroup Inc | Financial | MS | Morgan Stanley | Financial |
| CAT | Caterpillar Inc | Industrial Goods | MSFT | Microsoft Corp | Technology |
| CELG | Celgene Corp | Healthcare | NKE | NIKE Inc B | Consumer Goods |
| CL | Colgate-Palmolive Co | Consumer Goods | NSC | Norfolk Southern Corp | Services |
| CMCSA | Comcast Corp | Services | ORCL | Oracle Corp | Technology |
| COF | Capital One Financial | Financial | OXY | Occidental Petroleum | Basic Materials |
| COP | ConocoPhillips | Basic Materials | PEP | PepsiCo Inc | Consumer Goods |
| COST | Costco Wholesale Corp | Services | PFE | Pfizer Inc | Healthcare |
| CSCO | Cisco Systems Inc | Technology | PG | Procter & Gamble | Consumer Goods |
| CVS | CVS Health Corporation | Healthcare | QCOM | QUALCOMM Inc | Technology |
| CVX | Chevron Corp | Basic Materials | RTN | Raytheon Co | Industrial Goods |
| DD | E. I. du Pont de Nemours and Company | Basic Materials | SBUX | Starbucks Corp | Services |
| DIS | Walt Disney Co | Services | SLB | Schlumberger Ltd | Basic Materials |
| DOW | Dow Chemical | Basic Materials | SO | Southern Co | Utilities |
| DVN | Devon Energy Corp | Basic Materials | SPG | Simon Property Group | Financial |
| EBAY | eBay Inc. | Services | T | AT&T Inc. | Technology |
| EMC | EMC Corp | Technology | TGT | Target Corp | Services |
| EMR | Emerson Electric Co | Industrial Goods | TWX | Time Warner Inc | Services |
| EXC | Exelon Corp | Utilities | TXN | Texas Instruments Inc | Technology |
| F | Ford Motor Co | Consumer Goods | UNH | Unitedhealth Group Inc | Healthcare |
| FDX | FedEx Corp | Services | UNP | Union Pacific Corp | Services |
| FOXA | Twenty-First Century Fox, Inc | Services | USB | US Bancorp | Financial |
| GD | General Dynamics | Industrial Goods | UTX | United Technologies Corp | Industrial Goods |
| GE | General Electric Co | Industrial Goods | V | Visa Inc | Services |
| GILD | Gilead Sciences Inc | Healthcare | VZ | Verizon Communications Inc | Technology |
| GS | Goldman Sachs Group Inc | Financial | WBA | Walgreens Boots Alliance Inc | Services |
| HAL | Halliburton Co | Basic Materials | WFC | Wells Fargo & Co | Financial |
| HD | Home Depot Inc | Services | WMT | Wal-Mart Stores | Services |
| HON | Honeywell Intl Inc | Industrial Goods | XOM | Exxon Mobil Corp | Basic Materials |

# References


Acemoglu, D., Akcigit, U. and Kerr, W. (2015), Networks and the macroeconomy: An empirical exploration, Technical report, National Bureau of Economic Research.

Acemoglu, D., Carvalho, V. M., Ozdaglar, A. and Tahbaz-Salehi, A. (2012), 'The network origins of aggregate fluctuations', *Econometrica* **80**(5), 1977–2016.
*http://dx.doi.org/10.3982/ECTA9623

Acemoglu, D., Ozdaglar, A. and Tahbaz-Salehi, A. (2015*a*), Networks, shocks, and systemic risk, Technical report, National Bureau of Economic Research.

Acemoglu, D., Ozdaglar, A. and Tahbaz-Salehi, A. (2015*b*), 'Systemic risk and stability in financial networks', *American Economic Review* **105**(2), 564–608.
*http://www.aeaweb.org/articles.php?doi=10.1257/aer.20130456

Acharya, V. V., Pedersen, L. H., Philippon, T. and Richardson, M. P. (2010), 'Measuring systemic risk', *FRB of Cleveland Working Paper* (10-02).
*http://ssrn.com/abstract=1595075

Adrian, T. and Brunnermeier, M. K. (2011), Covar, Working Paper 17454, National Bureau of Economic Research.
*http://www.nber.org/papers/w17454

Ahelegbey, D. F. (2015), 'The econometrics of networks: A review', *University Ca'Foscari of Venice, Dept. of Economics Research Paper Series No* **13**.

Ahelegbey, D. F., Billio, M. and Casarin, R. (2016), 'Bayesian graphical models for structural vector autoregressive processes', *Journal of Applied Econometrics* **31**(2), 357–386.
*http://dx.doi.org/10.1002/jae.2443

Ahern, K. R. (2013), 'Network centrality and the cross section of stock returns', *Available at SSRN 2197370* .

Albuquerque, R., Ramadorai, T. and Watugala, S. W. (2015), 'Trade credit and cross-country predictable firm returns', *Journal of Financial Economics* **115**(3), 592 – 613.
*http://www.sciencedirect.com/science/article/pii/S0304405X14002293

Allen, F. and Babus, A. (2008), 'Networks in finance', *Wharton Financial Institutions Center Working Paper* (No. 08-07).

Andersen, T. G., Bollerslev, T., Christoffersen, P. F. and Diebold, F. X. (2012), Financial risk measurement for financial risk management, Working Paper 18084, National Bureau of Economic Research.
*http://www.nber.org/papers/w18084

Ang, A., Hodrick, R. J., Xing, Y. and Zhang, X. (2006), 'The cross-section of volatility and expected returns', *The Journal of Finance* **61**(1), 259–299.
*http://dx.doi.org/10.1111/j.1540-6261.2006.00836.x

Aobdia, D., Caskey, J. and Ozel, N. B. (2014), 'Inter-industry network structure and the cross-predictability of earnings and stock returns', *Review of Accounting Studies* **19**(3), 1191–1224.
*http://dx.doi.org/10.1007/s11142-014-9286-7

Barigozzi, M. and Brownlees, C. T. (2014), 'Nets: network estimation for time series', *Manuscript* .

Barigozzi, M. and Brownlees, C. T. (2016), 'Nets: network estimation for time series', *Available at SSRN: http://ssrn.com/abstract=2249909 or http://dx.doi.org/10.2139/ssrn.2249909* .

Benoit, S., Colliard, J.-E., Hurlin, C. and Perignon, C. (2015), 'Where the risks lie: A survey on systemic risk', *HEC Paris Research Paper No. FIN-2015-1088 . Available at SSRN: http://ssrn.com/abstract=2577961 or http://dx.doi.org/10.2139/ssrn.2577961* .

Bianchi, D., Billio, M. and Casarin, R. (2015), 'Modeling contagion and systemic risk', *Available at SSRN: http://ssrn.com/abstract=2537986 or http://dx.doi.org/10.2139/ssrn.2537986* .

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. Market Institutions, Financial Market Risks and Financial Crisis.
*http://www.sciencedirect.com/science/article/pii/S0304405X11002868

Bramoulle, Y., Djebbari, H. and Fortin, B. (2009), 'Identification of peer effects through social networks', *Journal of Econometrics* **150**(1), 41 – 55.
*http://www.sciencedirect.com/science/article/pii/S0304407609000335

Brownlees, C. T. and Engle, R. F. (2015), 'Srisk: A conditional capital shortfall index for systemic risk measurement', *Available at SSRN: http://ssrn.com/abstract=1611229 or http://dx.doi.org/10.2139/ssrn.1611229* .

Buraschi, A. and Porchia, P. (2012), Dynamic networks and asset pricing, *in* 'AFA 2013 San Diego Meetings Paper'.

Cohen, L. and Frazzini, A. (2008), 'Economic links and predictable returns', *The Journal of Finance* **63**(4), 1977–2011.
*http://dx.doi.org/10.1111/j.1540-6261.2008.01379.x

Demirer, M., Diebold, F. X., Liu, L. and Yilmaz, K. (2015), 'Estimating global bank network connectedness', *Available at SSRN 2631479* .

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

Donders, M., Kouwenberg, R., Vorst, T. et al. (2000), 'Options and earnings announcements: an empirical study of volatility, trading volume, open interest and liquidity', *European Financial Management* **6**(2), 149–171.

Dufour, J.-M., Garcia, R. and Taamouti, A. (2012), 'Measuring high-frequency causality between returns, realized volatility, and implied volatility', *Journal of Financial Econometrics* **10**(1), 124–163.
*http://jfec.oxfordjournals.org/content/10/1/124.abstract

Dufour, J.-M., Pelletier, D. and Renault, E. (2006), 'Short run and long run causality in time series: inference', *Journal of Econometrics* **132**(2), 337 – 362.
*http://www.sciencedirect.com/science/article/pii/S0304407605000588

Dufour, J.-M. and Renault, E. (1998), 'Short run and long run causality in time series: Theory', *Econometrica* **66**(5), pp. 1099–1125.
*http://www.jstor.org/stable/2999631

Dufour, J.-M. and Taamouti, A. (2010), 'Short and long run causality measures: Theory and inference', *Journal of Econometrics* **154**(1), 42 – 58.
*http://www.sciencedirect.com/science/article/pii/S0304407609001493

Dufour, J.-M. and Zhang, H. J. (2015), Short and long run second-order causality: Theory, measures and inference, Technical report, Technical report, Department of Economics, McGill University.

Eichler, M. (2007), 'Granger causality and path diagrams for multivariate time series', *Journal of Econometrics* **137**(2), 334 – 353.
*http://www.sciencedirect.com/science/article/pii/S0304407606000480

Elliott, M., Golub, B. and Jackson, M. O. (2014), 'Financial networks and contagion', *American Economic Review* **104**(10), 3115–53.
*http://www.aeaweb.org/articles.php?doi=10.1257/aer.104.10.3115

Everett, M. G. and Borgatti, S. P. (1999), 'The centrality of groups and classes', *The Journal of mathematical sociology* **23**(3), 181–201.

Freeman, L. C. (1978), 'Centrality in social networks conceptual clarification', *Social Networks* **1**(3), 215 – 239.
*http://www.sciencedirect.com/science/article/pii/0378873378900217

Gençay, R., Signori, D., Xue, Y., Yu, X. and Zhang, K. (2015), 'Economic links and credit spreads', *Journal of Banking & Finance* **55**, 157 – 169. Global Governance and Financial Stability.
*http://www.sciencedirect.com/science/article/pii/S0378426615000345

Gençay, R., Yu, X. and Zhang, K. (2016), 'Economic links and return volatility', *Working Paper* .

Geweke, J. (1982), 'Measurement of linear dependence and feedback between multiple time series', *Journal of the American Statistical Association* **77**(378), 304–313.
*http://www.tandfonline.com/doi/abs/10.1080/01621459.1982.10477803

Giudici, P. and Spelta, A. (2016), 'Graphical network models for international financial flows', *Journal of Business & Economic Statistics* **34**(1), 128–138.
*http://dx.doi.org/10.1080/07350015.2015.1017643

Granger, C. W. (1969), 'Investigating causal relations by econometric models and cross-spectral methods', *Econometrica: Journal of the Econometric Society* pp. 424–438.

Hautsch, N., Schaumburg, J. and Schienle, M. (2014), 'Financial network systemic risk contributions', *Review of Finance* .
*http://rof.oxfordjournals.org/content/early/2014/03/24/rof.rfu010.abstract

Hertzel, M. G., Li, Z., Officer, M. S. and Rodgers, K. J. (2008), 'Inter-firm linkages and the wealth effects of financial distress along the supply chain', *Journal of Financial Economics* **87**(2), 374 – 387.
*http://www.sciencedirect.com/science/article/pii/S0304405X07001808

Jackson, M. O. (2008), *Social and economic networks*, Vol. 3, Princeton University Press Princeton.

Jian, B. (2016), 'Centralities in illiquidity transmission networks and the cross-section of expected returns', *Working Paper* .

Menzly, L. and Ozbas, O. (2010), 'Market segmentation and cross-predictability of returns', *The Journal of Finance* **65**(4), 1555–1580.
*http://dx.doi.org/10.1111/j.1540-6261.2010.01578.x

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. PMID: 19881892.
*http://dx.doi.org/10.1198/jasa.2009.0126

Sims, C. A. (1972), 'Money, income, and causality', *The American economic review* pp. 540–552.

Tibshirani, R. (1996), 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society. Series B (Methodological)* **58**(1), pp. 267–288.
*http://www.jstor.org/stable/2346178

Zou, H. (2006), 'The adaptive lasso and its oracle properties', *Journal of the American Statistical Association* **101**(476), 1418–1429.
*http://dx.doi.org/10.1198/016214506000000735

Table 2. Summary statistics of causality measures from each firm to other firms. This table reports the summary statistics of each row of the firm-wise causality table $[C_{i\rightarrow\cdot}]$. The causality table is estimated by the full data sample (20/08/1999 - 31/08/2015). Nodes are the firms of selected S&P 100 components. For each firm $i$, we report the minimum value, the maximum value, the mean value and the quantiles (25%, 50% (median) and 75%) of the entries in its "OUT" vector. The reported values are 100 times of the raw values, and are kept with two digits. We sort the tickers by their median values and identify the top 10 influential firms.

| Sector* | Ticker | Median | Mean | Min | 25% | 75% | Max | Sector* | Ticker | Median | Mean | Min | 25% | 75% | Max |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **F** | **BAC** | 0.07 | 0.43 | 0.00 | 0.00 | 0.51 | 5.33 | I | LMT | 0.00 | 0.03 | 0.00 | 0.00 | 0.01 | 1.80 |
| C | AAPL | 0.07 | 0.16 | 0.00 | 0.02 | 0.20 | 1.03 | C | KO | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 1.34 |
| T | CSCO | 0.07 | 0.27 | 0.00 | 0.00 | 0.31 | 2.63 | H | AGN | 0.00 | 0.03 | 0.00 | 0.00 | 0.01 | 2.27 |
| **F** | **C** | 0.07 | 0.20 | 0.00 | 0.00 | 0.21 | 2.77 | B | CVX | 0.00 | 0.04 | 0.00 | 0.00 | 0.01 | 1.59 |
| **F** | **BK** | 0.06 | 0.33 | 0.00 | 0.00 | 0.33 | 5.75 | T | HPQ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.16 |
| **F** | **AIG** | 0.03 | 0.15 | 0.00 | 0.00 | 0.15 | 2.28 | B | OXY | 0.00 | 0.08 | 0.00 | 0.00 | 0.02 | 3.81 |
| **F** | **MET** | 0.03 | 0.11 | 0.00 | 0.00 | 0.07 | 1.32 | H | BAX | 0.00 | 0.10 | 0.00 | 0.00 | 0.02 | 4.88 |
| C | F | 0.03 | 0.12 | 0.00 | 0.00 | 0.10 | 4.04 | B | DVN | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.10 |
| **F** | **JPM** | 0.03 | 0.11 | 0.00 | 0.00 | 0.08 | 0.83 | H | BMY | 0.00 | 0.01 | 0.00 | 0.00 | 0.02 | 0.08 |
| **F** | **MS** | 0.02 | 0.38 | 0.00 | 0.00 | 0.17 | 4.83 | S | CMCSA | 0.00 | 0.30 | 0.00 | 0.00 | 0.14 | 6.38 |
| H | GILD | 0.02 | 0.09 | 0.00 | 0.00 | 0.10 | 1.39 | F | ALL | 0.00 | 0.20 | 0.00 | 0.00 | 0.07 | 4.67 |
| I | GE | 0.02 | 0.17 | 0.00 | 0.00 | 0.08 | 3.63 | F | USB | 0.00 | 0.16 | 0.00 | 0.00 | 0.04 | 2.62 |
| F | WFC | 0.02 | 0.15 | 0.00 | 0.00 | 0.12 | 3.96 | B | SLB | 0.00 | 0.09 | 0.00 | 0.00 | 0.00 | 3.28 |
| S | TGT | 0.02 | 0.06 | 0.00 | 0.00 | 0.05 | 0.37 | T | TXN | 0.00 | 0.09 | 0.00 | 0.00 | 0.01 | 4.93 |
| T | IBM | 0.02 | 0.07 | 0.00 | 0.00 | 0.06 | 1.02 | S | SBUX | 0.00 | 0.07 | 0.00 | 0.00 | 0.01 | 6.14 |
| F | GS | 0.02 | 0.12 | 0.00 | 0.00 | 0.08 | 1.70 | T | MSFT | 0.00 | 0.06 | 0.00 | 0.00 | 0.02 | 2.96 |
| T | VZ | 0.02 | 0.07 | 0.00 | 0.00 | 0.03 | 4.11 | S | DIS | 0.00 | 0.06 | 0.00 | 0.00 | 0.01 | 5.11 |
| F | SPG | 0.02 | 0.14 | 0.00 | 0.00 | 0.07 | 5.63 | T | ACN | 0.00 | 0.06 | 0.00 | 0.00 | 0.01 | 3.49 |
| S | TWX | 0.01 | 0.07 | 0.00 | 0.00 | 0.05 | 2.47 | H | UNH | 0.00 | 0.06 | 0.00 | 0.00 | 0.01 | 2.78 |
| B | DOW | 0.01 | 0.09 | 0.00 | 0.00 | 0.03 | 2.52 | U | EXC | 0.00 | 0.06 | 0.00 | 0.00 | 0.01 | 4.19 |
| I | BA | 0.01 | 0.11 | 0.00 | 0.00 | 0.03 | 7.36 | H | CVS | 0.00 | 0.05 | 0.00 | 0.00 | 0.01 | 2.97 |
| T | EMC | 0.01 | 0.04 | 0.00 | 0.00 | 0.03 | 0.78 | S | AMZN | 0.00 | 0.05 | 0.00 | 0.00 | 0.03 | 0.97 |
| F | AXP | 0.01 | 0.16 | 0.00 | 0.00 | 0.07 | 2.78 | H | LLY | 0.00 | 0.05 | 0.00 | 0.00 | 0.01 | 3.86 |
| T | ORCL | 0.01 | 0.05 | 0.00 | 0.00 | 0.02 | 3.26 | H | ABT | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 3.48 |
| H | PFE | 0.00 | 0.06 | 0.00 | 0.00 | 0.03 | 2.51 | B | APC | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 1.71 |
| S | COST | 0.00 | 0.02 | 0.00 | 0.00 | 0.02 | 0.98 | T | QCOM | 0.00 | 0.03 | 0.00 | 0.00 | 0.01 | 0.74 |
| H | CELG | 0.00 | 0.03 | 0.00 | 0.00 | 0.01 | 1.14 | B | XOM | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 0.86 |
| T | INTC | 0.00 | 0.04 | 0.00 | 0.00 | 0.03 | 0.68 | I | HON | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.29 |
| U | SO | 0.00 | 0.09 | 0.00 | 0.00 | 0.01 | 7.63 | S | WMT | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.37 |
| F | COF | 0.00 | 0.10 | 0.00 | 0.00 | 0.04 | 4.63 | H | BIIB | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.13 |
| S | UNP | 0.00 | 0.04 | 0.00 | 0.00 | 0.01 | 1.93 | C | NKE | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.31 |
| S | MCD | 0.00 | 0.08 | 0.00 | 0.00 | 0.01 | 6.30 | S | HD | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.14 |
| B | HAL | 0.00 | 0.10 | 0.00 | 0.00 | 0.02 | 6.68 | B | MON | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.11 |
| T | T | 0.00 | 0.05 | 0.00 | 0.00 | 0.02 | 3.70 | C | CL | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.20 |
| H | MRK | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 | 0.69 | I | GD | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.10 |
| S | FOXA | 0.00 | 0.10 | 0.00 | 0.00 | 0.01 | 6.94 | H | AMGN | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.13 |
| I | CAT | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.45 | I | UTX | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 |
| S | V | 0.00 | 0.11 | 0.00 | 0.00 | 0.04 | 7.98 | B | COP | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| I | EMR | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.27 | B | DD | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| S | EBAY | 0.00 | 0.17 | 0.00 | 0.00 | 0.08 | 3.26 | S | FDX | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| S | WBA | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 | 0.70 | C | MO | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| I | MMM | 0.00 | 0.03 | 0.00 | 0.00 | 0.03 | 0.44 | S | LOW | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| H | JNJ | 0.00 | 0.07 | 0.00 | 0.00 | 0.01 | 4.87 | H | MDT | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| I | RTN | 0.00 | 0.07 | 0.00 | 0.00 | 0.01 | 5.79 | C | PEP | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| S | NSC | 0.00 | 0.12 | 0.00 | 0.00 | 0.01 | 8.78 | C | PG | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |

* B: Basic Materials; C: Consumer Goods; F: Financial; H: Healthcare; I: Industrial Goods; S: Services; T: Technology; U: Utilities.

Table 3. Summary statistics of causality measures to each firm from others firms. This table reports the summary statistics of each column of the firm-wise causality table $[C_{\cdot \to i}]$. The causality table is estimated by the full data sample (20/08/1999 - 31/08/2015). Nodes are the firms of selected S&P 100 components. For each firm $i$, we report the minimum value, the maximum value, the mean value and the quantiles (25%, 50% (median) and 75%) of the entries in its "IN" vector. The reported values are 100 times of the raw values, and are kept with two digits. We sort the tickers by their median values and identidy the top 10 sensitive firms.

| Sector* | Ticker | Median | Mean | Min | 25% | 75% | Max | Sector* | Ticker | Median | Mean | Min | 25% | 75% | Max |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| B | OXY | 0.02 | 0.19 | 0.00 | 0.00 | 0.17 | 3.81 | H | BIIB | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.04 |
| I | RTN | 0.02 | 0.09 | 0.00 | 0.01 | 0.03 | 5.79 | S | WMT | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.11 |
| I | LMT | 0.01 | 0.02 | 0.00 | 0.00 | 0.02 | 0.22 | U | SO | 0.00 | 0.09 | 0.00 | 0.00 | 0.01 | 7.63 |
| B | SLB | 0.01 | 0.25 | 0.00 | 0.00 | 0.10 | 3.26 | H | JNJ | 0.00 | 0.20 | 0.00 | 0.00 | 0.01 | 4.88 |
| I | GD | 0.01 | 0.05 | 0.00 | 0.00 | 0.05 | 0.59 | C | NKE | 0.00 | 0.02 | 0.00 | 0.00 | 0.03 | 0.25 |
| F | C | 0.01 | 0.18 | 0.00 | 0.00 | 0.07 | 3.34 | S | EBAY | 0.00 | 0.04 | 0.00 | 0.00 | 0.02 | 1.14 |
| S | DIS | 0.01 | 0.17 | 0.00 | 0.00 | 0.03 | 5.11 | S | AMZN | 0.00 | 0.10 | 0.00 | 0.00 | 0.01 | 4.93 |
| F | WFC | 0.01 | 0.15 | 0.00 | 0.00 | 0.07 | 3.52 | H | CELG | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 1.14 |
| B | COP | 0.00 | 0.09 | 0.00 | 0.00 | 0.09 | 1.32 | B | HAL | 0.00 | 0.11 | 0.00 | 0.00 | 0.02 | 6.68 |
| I | EMR | 0.00 | 0.04 | 0.00 | 0.00 | 0.04 | 0.51 | F | BK | 0.00 | 0.17 | 0.00 | 0.00 | 0.01 | 5.75 |
| I | GE | 0.00 | 0.18 | 0.00 | 0.00 | 0.04 | 2.13 | S | FOXA | 0.00 | 0.15 | 0.00 | 0.00 | 0.01 | 6.94 |
| F | ALL | 0.00 | 0.26 | 0.00 | 0.00 | 0.07 | 3.96 | S | UNP | 0.00 | 0.15 | 0.00 | 0.00 | 0.02 | 2.87 |
| B | DD | 0.00 | 0.05 | 0.00 | 0.00 | 0.03 | 0.98 | S | CMCSA | 0.00 | 0.13 | 0.00 | 0.00 | 0.01 | 6.38 |
| I | BA | 0.00 | 0.14 | 0.00 | 0.00 | 0.04 | 7.36 | S | NSC | 0.00 | 0.13 | 0.00 | 0.00 | 0.02 | 8.78 |
| T | T | 0.00 | 0.10 | 0.00 | 0.00 | 0.03 | 3.70 | F | BAC | 0.00 | 0.13 | 0.00 | 0.00 | 0.04 | 5.33 |
| F | MET | 0.00 | 0.09 | 0.00 | 0.00 | 0.03 | 1.52 | C | F | 0.00 | 0.11 | 0.00 | 0.00 | 0.01 | 4.04 |
| H | MDT | 0.00 | 0.02 | 0.00 | 0.00 | 0.02 | 0.18 | B | CVX | 0.00 | 0.11 | 0.00 | 0.00 | 0.08 | 1.06 |
| S | WBA | 0.00 | 0.02 | 0.00 | 0.00 | 0.02 | 0.20 | B | DVN | 0.00 | 0.09 | 0.00 | 0.00 | 0.05 | 1.84 |
| I | HON | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 | 0.40 | S | V | 0.00 | 0.09 | 0.00 | 0.00 | 0.00 | 7.98 |
| F | SPG | 0.00 | 0.15 | 0.00 | 0.00 | 0.03 | 5.63 | S | MCD | 0.00 | 0.09 | 0.00 | 0.00 | 0.01 | 6.30 |
| F | COF | 0.00 | 0.08 | 0.00 | 0.00 | 0.01 | 4.63 | C | KO | 0.00 | 0.09 | 0.00 | 0.00 | 0.01 | 3.36 |
| H | UNH | 0.00 | 0.12 | 0.00 | 0.00 | 0.03 | 4.22 | H | PFE | 0.00 | 0.07 | 0.00 | 0.00 | 0.02 | 2.63 |
| B | APC | 0.00 | 0.11 | 0.00 | 0.00 | 0.07 | 1.71 | S | TWX | 0.00 | 0.07 | 0.00 | 0.00 | 0.02 | 1.11 |
| T | MSFT | 0.00 | 0.07 | 0.00 | 0.00 | 0.02 | 2.96 | B | XOM | 0.00 | 0.05 | 0.00 | 0.00 | 0.04 | 0.62 |
| F | AXP | 0.00 | 0.12 | 0.00 | 0.00 | 0.04 | 2.75 | B | MON | 0.00 | 0.05 | 0.00 | 0.00 | 0.01 | 0.70 |
| H | ABT | 0.00 | 0.02 | 0.00 | 0.00 | 0.02 | 0.30 | T | EMC | 0.00 | 0.05 | 0.00 | 0.00 | 0.00 | 2.17 |
| H | CVS | 0.00 | 0.06 | 0.00 | 0.00 | 0.01 | 2.97 | T | ORCL | 0.00 | 0.05 | 0.00 | 0.00 | 0.01 | 3.26 |
| F | USB | 0.00 | 0.08 | 0.00 | 0.00 | 0.02 | 2.23 | S | HD | 0.00 | 0.05 | 0.00 | 0.00 | 0.04 | 0.63 |
| B | DOW | 0.00 | 0.05 | 0.00 | 0.00 | 0.01 | 2.52 | T | ACN | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 3.49 |
| F | GS | 0.00 | 0.08 | 0.00 | 0.00 | 0.02 | 1.70 | I | CAT | 0.00 | 0.04 | 0.00 | 0.00 | 0.01 | 0.76 |
| S | SBUX | 0.00 | 0.17 | 0.00 | 0.00 | 0.04 | 6.14 | S | TGT | 0.00 | 0.04 | 0.00 | 0.00 | 0.04 | 0.41 |
| C | CL | 0.00 | 0.02 | 0.00 | 0.00 | 0.02 | 0.39 | S | LOW | 0.00 | 0.04 | 0.00 | 0.00 | 0.02 | 0.50 |
| S | FDX | 0.00 | 0.03 | 0.00 | 0.00 | 0.03 | 0.46 | F | JPM | 0.00 | 0.03 | 0.00 | 0.00 | 0.00 | 1.21 |
| F | MS | 0.00 | 0.17 | 0.00 | 0.00 | 0.05 | 3.63 | H | AGN | 0.00 | 0.03 | 0.00 | 0.00 | 0.00 | 2.27 |
| T | QCOM | 0.00 | 0.04 | 0.00 | 0.00 | 0.02 | 0.93 | F | AIG | 0.00 | 0.03 | 0.00 | 0.00 | 0.00 | 1.86 |
| U | EXC | 0.00 | 0.13 | 0.00 | 0.00 | 0.02 | 4.19 | T | VZ | 0.00 | 0.03 | 0.00 | 0.00 | 0.00 | 0.96 |
| S | COST | 0.00 | 0.13 | 0.00 | 0.00 | 0.03 | 2.59 | C | PEP | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 | 0.64 |
| I | MMM | 0.00 | 0.03 | 0.00 | 0.00 | 0.02 | 0.59 | H | MRK | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 | 0.28 |
| H | AMGN | 0.00 | 0.03 | 0.00 | 0.00 | 0.02 | 0.66 | I | UTX | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 | 0.45 |
| C | AAPL | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 | 0.20 | T | HPQ | 0.00 | 0.02 | 0.00 | 0.00 | 0.02 | 0.18 |
| T | INTC | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.15 | T | TXN | 0.00 | 0.02 | 0.00 | 0.00 | 0.02 | 0.24 |
| H | BAX | 0.00 | 0.09 | 0.00 | 0.00 | 0.01 | 3.48 | T | CSCO | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.20 |
| H | BMY | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.09 | C | PG | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.29 |
| H | LLY | 0.00 | 0.08 | 0.00 | 0.00 | 0.02 | 3.86 | H | GILD | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.13 |
| T | IBM | 0.00 | 0.03 | 0.00 | 0.00 | 0.02 | 1.17 | C | MO | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.08 |

* B: Basic Materials; C: Consumer Goods; F: Financial; H: Healthcare; I: Industrial Goods; S: Services; T: Technology; U: Utilities.

Table 4. Top 10 influential firms at different forecast horizons. This table reports the top 10 influential firms and their respective sector at different forecast horizons, $h = 1, 2, 3, 4, 5$. Given the forecast horizon $h$, we obtain the summary statistics of each row of the firm-wise causality table $[C^h_{i\to\cdot}]$. The causality table is estimated by the full data sample (20/08/1999 - 31/08/2015). Nodes are the firms of selected S&P 100 components. For each firm $i$, we have the median value of the entries in its "OUT" vector. For each given forecast horizon $h$, we sort the tickers by their median values and identify the top 10 influential firms.

| | h=1 | | h=2 | | h=3 | | h=4 | | h=5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| Rank | Sector* | Ticker | Sector* | Ticker | Sector* | Ticker | Sector* | Ticker | Sector* | Ticker |
| 1 | F | BAC | T | CSCO | T | CSCO | T | CSCO | T | CSCO |
| 2 | C | AAPL | C | AAPL | C | AAPL | C | AAPL | C | AAPL |
| 3 | T | CSCO | F | C | F | AIG | F | AIG | F | AIG |
| 4 | F | C | F | AIG | F | C | F | C | F | C |
| 5 | F | BK | F | GS | F | GS | F | GS | F | GS |
| 6 | F | AIG | I | GE | I | GE | I | GE | I | GE |
| 7 | F | MET | F | MS | F | JPM | F | JPM | F | JPM |
| 8 | C | F | F | JPM | C | F | C | F | C | F |
| 9 | F | JPM | F | MET | T | IBM | T | IBM | T | IBM |
| 10 | F | MS | C | F | F | MET | T | EMC | T | EMC |

* B: Basic Materials; C: Consumer Goods; F: Financial; H: Healthcare; I: Industrial Goods; S: Services; T: Technology; U: Utilities.

Table 5. Summary statistics of causality measures from each financial Firm to other financial Firms. This table reports the summary statistics of the firm-wise causality table blocked by the financial sector $[C_{i \to j}]$, where $i, j \in$ Financial Sector. The causality table is estimated by the full data sample (20/08/1999 - 31/08/2015). Nodes are the firms of selected S&P 100 components. For each financial firm $i$, we report the minimum value, the maximum value, the mean value and the quantiles (25%, 50% (median) and 75%) of the entries in its "OUT" vector truncated within the financial sector. The reported values are 100 times of the raw values, and are kept with two digits. We sort the tickers by their median values and identify the top 3 influencial firms in the financial sector.

| Ticker | Median | Mean | Min | 25% | 75% | Max |
|---|---|---|---|---|---|---|
| **BAC** | **0.42** | 1.21 | 0.00 | 0.04 | 1.66 | 5.33 |
| **MS** | **0.30** | 1.19 | 0.00 | 0.04 | 2.21 | 4.83 |
| **GS** | **0.25** | 0.43 | 0.00 | 0.00 | 0.82 | 1.70 |
| BK | 0.06 | 0.91 | 0.00 | 0.00 | 0.98 | 5.75 |
| WFC | 0.03 | 0.49 | 0.00 | 0.00 | 0.38 | 3.96 |
| ALL | 0.01 | 0.30 | 0.00 | 0.00 | 0.37 | 1.46 |
| SPG | 0.01 | 0.54 | 0.00 | 0.00 | 0.08 | 5.63 |
| AXP | 0.00 | 0.43 | 0.00 | 0.00 | 0.24 | 2.78 |
| C | 0.00 | 0.42 | 0.00 | 0.00 | 0.53 | 2.77 |
| AIG | 0.00 | 0.14 | 0.00 | 0.00 | 0.12 | 0.74 |
| COF | 0.00 | 0.39 | 0.00 | 0.00 | 0.12 | 4.63 |
| JPM | 0.00 | 0.10 | 0.00 | 0.00 | 0.15 | 0.51 |
| MET | 0.00 | 0.31 | 0.00 | 0.00 | 0.58 | 1.32 |
| USB | 0.00 | 0.11 | 0.00 | 0.00 | 0.03 | 0.55 |

Table 6. Summary statistics of causality measures to each financial firm from other financial firms. This table reports the summary statistics of the firm-wise causality table blocked by the financial sector $[C_{j\to i}]$, where $i, j \in$ Financial Sector. The causality table is estimated by the full data sample (20/08/1999 - 31/08/2015). Nodes are the firms of selected S&P 100 components. For each firm $i$, we report the minimum value, the maximum value, the mean value and the quantiles (25%, 50% (median) and 75%) of the entries in its "IN" vector truncated within the financial sector. The reported values are 100 times of the raw values, and are kept with two digits. We sort the tickers by their median values and identify the top 3 sensitive firms in the financial sector.

| Ticker | Median | Mean | Min | 25% | 75% | Max |
|---|---|---|---|---|---|---|
| **C** | **0.33** | 0.65 | 0.00 | 0.03 | 0.94 | 3.34 |
| **ALL** | **0.32** | 1.02 | 0.00 | 0.02 | 2.21 | 3.96 |
| **BAC** | **0.28** | 0.64 | 0.00 | 0.00 | 0.63 | 5.33 |
| SPG | 0.22 | 0.85 | 0.00 | 0.13 | 0.64 | 5.63 |
| MET | 0.07 | 0.36 | 0.00 | 0.00 | 0.49 | 1.52 |
| MS | 0.01 | 0.38 | 0.00 | 0.00 | 0.62 | 1.75 |
| WFC | 0.01 | 0.66 | 0.00 | 0.00 | 1.00 | 3.52 |
| BK | 0.00 | 0.96 | 0.00 | 0.00 | 0.75 | 5.75 |
| AXP | 0.00 | 0.44 | 0.00 | 0.00 | 0.31 | 2.75 |
| AIG | 0.00 | 0.04 | 0.00 | 0.00 | 0.01 | 0.30 |
| COF | 0.00 | 0.46 | 0.00 | 0.00 | 0.18 | 4.63 |
| GS | 0.00 | 0.23 | 0.00 | 0.00 | 0.24 | 1.70 |
| JPM | 0.00 | 0.06 | 0.00 | 0.00 | 0.01 | 0.47 |
| USB | 0.00 | 0.23 | 0.00 | 0.00 | 0.00 | 2.23 |

Table 7. Summary statistics of causality measures from each sector to other sectors. This table reports the summary statistics of each row of the sector-wise causality table $[C_{V_i \to V_\cdot}]$. The causality table is estimated by the full data sample (20/08/1999 - 31/08/2015). Nodes are the sectors whose firms are selected in the S&P 100 components. For each sector $V_i$, we report the minimum value, the maximum value, the mean value and the quantiles (25%, 50% (median) and 75%) of the entries in its "OUT" vector. The reported values are 100 times of the raw values, and are kept with two digits. We sort the sectors by their median values and identify the top 4 influential sectors in the economy.

| Sector | Median | Mean | Min | 25% | 75% | Max |
|---|---|---|---|---|---|---|
| **Technology** | **3.27** | 5.75 | 0.00 | 0.00 | 7.40 | 18.84 |
| **Industrial Goods** | **1.55** | 3.28 | 0.00 | 0.01 | 3.16 | 15.04 |
| **Consumer Goods** | **0.90** | 1.07 | 0.00 | 0.43 | 1.47 | 3.02 |
| **Financial** | **0.48** | 9.61 | 0.00 | 0.05 | 5.46 | 57.89 |
| Utilities | 0.08 | 1.59 | 0.00 | 0.00 | 0.27 | 11.83 |
| Services | 0.00 | 7.74 | 0.00 | 0.00 | 2.67 | 52.77 |
| Healthcare | 0.00 | 3.15 | 0.00 | 0.00 | 0.25 | 24.58 |
| Basic Materials | 0.00 | 3.07 | 0.00 | 0.00 | 0.43 | 22.82 |

Table 8. Summary statistics of causality measures to each sector from other sectors. This table reports each column of the summary statistics of the sector-wise causality table $[C_{V_\cdot \to V_i}]$. The causality table is estimated by the full data sample (20/08/1999 - 31/08/2015). Nodes are the sectors whose firms are selected in the S&P 100 components. For each sector $V_i$, we report the minimum value, the maximum value, the mean value and the quantiles (25%, 50% (median) and 75%) of the entries in its "IN" vector. The reported values are 100 times of the raw values, and are kept with two digits. We sort the sectors by their median values and identify the top 4 sensitive sectors in the economy.

| Sector | Median | Mean | Min | 25% | 75% | Max |
|---|---|---|---|---|---|---|
| **Services** | **1.12** | 10.91 | 0.00 | 0.00 | 16.09 | 52.77 |
| **Basic Materials** | **1.11** | 4.43 | 0.00 | 0.00 | 3.63 | 22.82 |
| **Industrial Goods** | **0.69** | 2.72 | 0.00 | 0.06 | 2.28 | 15.04 |
| **Healthcare** | **0.57** | 3.79 | 0.00 | 0.00 | 1.84 | 24.58 |
| Technology | 0.36 | 2.64 | 0.00 | 0.00 | 0.77 | 18.84 |
| Consumer Goods | 0.09 | 0.73 | 0.00 | 0.00 | 0.84 | 3.02 |
| Utilities | 0.06 | 2.08 | 0.00 | 0.00 | 1.25 | 11.83 |
| Financial | 0.00 | 7.96 | 0.00 | 0.00 | 1.45 | 57.89 |

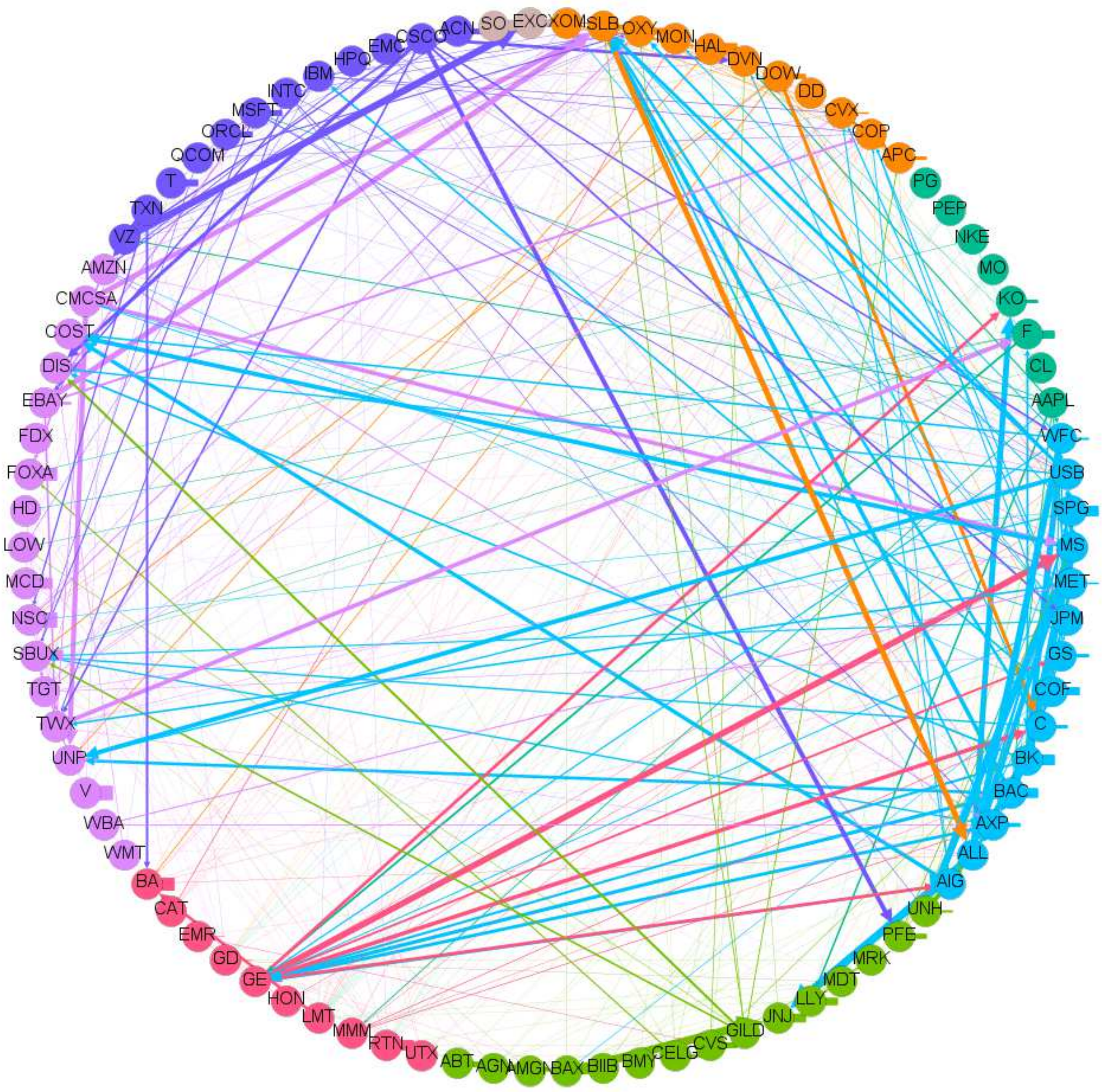


Figure 5. Firm-wise S&P 100 implied volatility network. This is a direct effect network corresponding to the causality table, $[C^1_{ij}]$. The causality measures table is estimated by the full data sample (20/08/1999 - 31/08/2015). Nodes are the firms of selected S&P 100 components. Different colors of the nodes correspond to different sectors that the nodes belong to (skyblue: financial; lawn green: healthcare; pink: industrial goods; purple: services; blue: technology; plum: utilities; orange: basic materials forest green: consumer goods). We only keep the directed and weighted edges $(i \to j)$ if $C_{ij}$ is greater or equal to the 90% percentile element in $OUT^1_i$ $(C_{i\cdot})$ and the 90% percentile element in $IN^1_j$ $(C_{\cdot j})$. When $i \to j$ and $j \to i$ are both kept, only the edge with greater strength will be shown in this figure. The colors of the edges correspond to the colors of the source nodes. The thickness of the edges are weight rescaled.

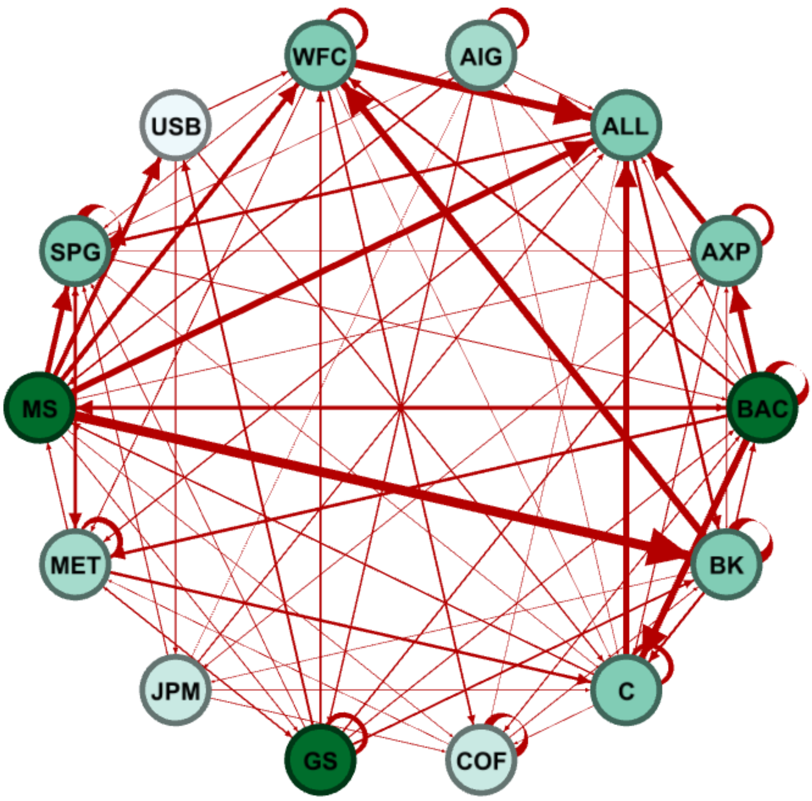


Figure 6. Firm-wise S&P 100 implied volatility network within financial sector. This is a direct effect network corresponding to the blocked causality table, $[C^1_{ij}]$ where $i, j \in$ Financial (both node $i$ and node $j$ are the firms that selected from S&P 100 components and belong the financial sector). The causality measures table is estimated by the full data sample (20/08/1999 - 31/08/2015). Nodes are the firms that selected from S&P 100 components and belong to the financial sector. We only keep the directed and weighted edges ($i \to j$) if $C_{ij}$ is greater or equal to the 50% percentile element in the blocked causality measures table, $[C^1_{ij}]$ where $i, j \in$ Financial. The darkness of the nodes corresponds to the out-degree of the nodes in this filtered network (e.g., MS, BAC and GS have higher out-degree). When $i \to j$ and $j \to i$ are both kept, only the edge with greater strength will be shown in this figure. The thickness of the edges are weight rescaled.

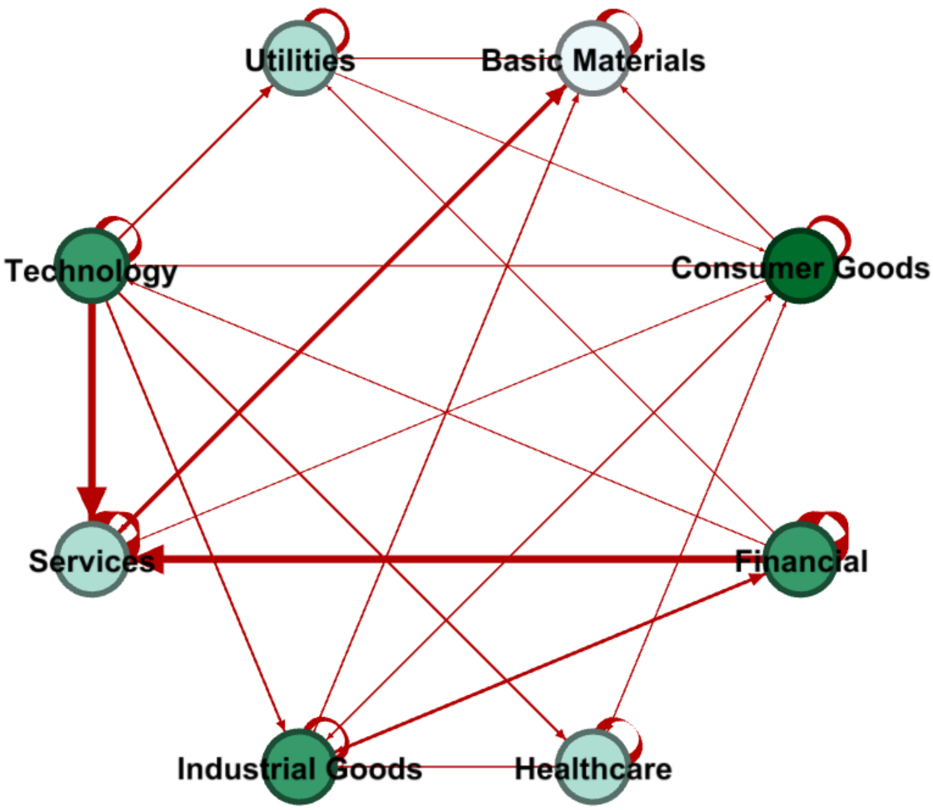


Figure 7. Sector-wise S&P 100 implied volatility network. This is a direct effect network corresponding to the causality table, $[C^1_{V_iV_j}]$. The causality table is estimated by the full data sample (20/08/1999 - 31/08/2015). Nodes are the sectors of the firms selected from S&P 100 components. We only keep the directed and weighted edges ($V_i \to V_j$) if $C_{V_iV_j}$ is greater or equal to the 50% percentile element in the causality measures table, $[C^1_{V_iV_j}]$. The darkness of the nodes corresponds to the out-degree of the nodes in this filtered network (e.g., Consumer Goods, Financial, Industrial Goods and Technology have higher out-degree). When $V_i \to V_j$ and $V_j \to V_i$ are both kept, only the edge with greater strength will be shown in this figure. The thickness of the edges are weight rescaled.

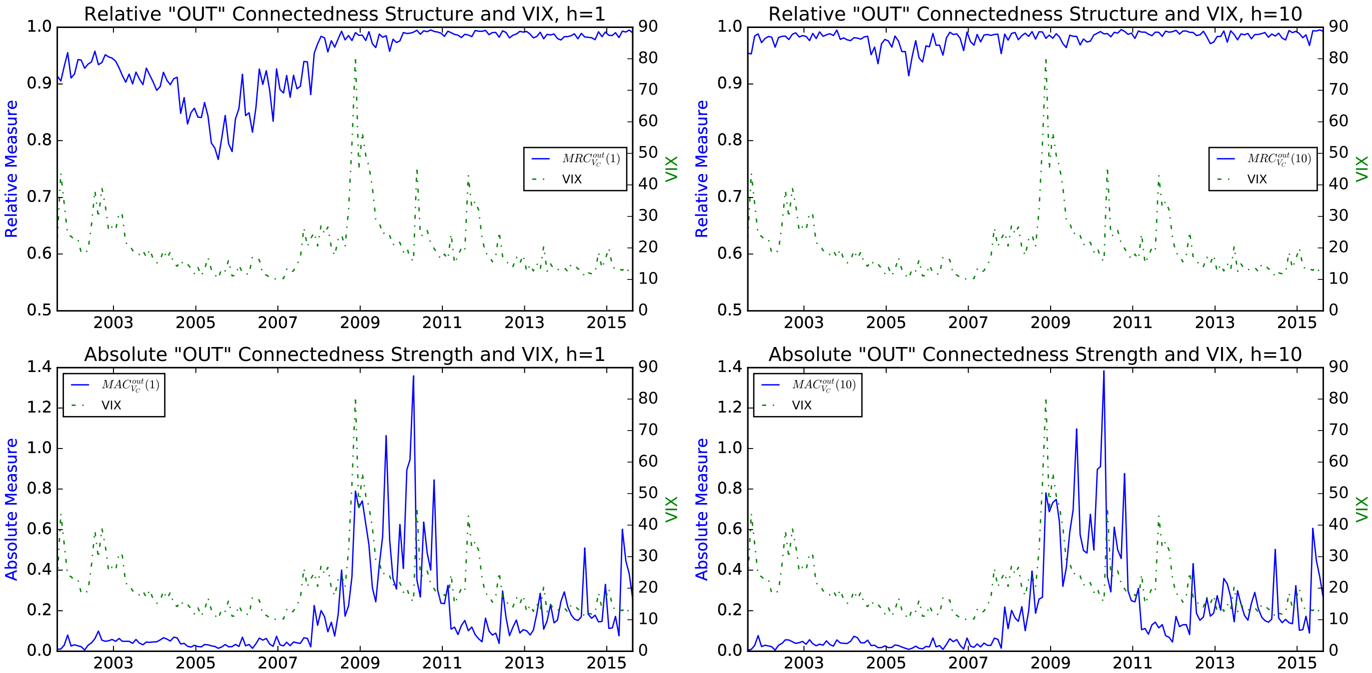


Figure 8. Firm-wise market connectedness "out" measures and the VIX index. The blue solid lines are our "out" connectedness measures (relative connectedness structure: upper row; absolute connectedness strength: bottom row) and the green dash lines are the VIX index. The reported connectedness measures are estimated at forecast horizon 1 (left column) and at forecast horizon 10 (right column). All measures are estimated every 1 month by the VAR(1) models with 2-year rolling estimation windows. The nodes of the underlying market are the companies that selected from the S&P 100 components.

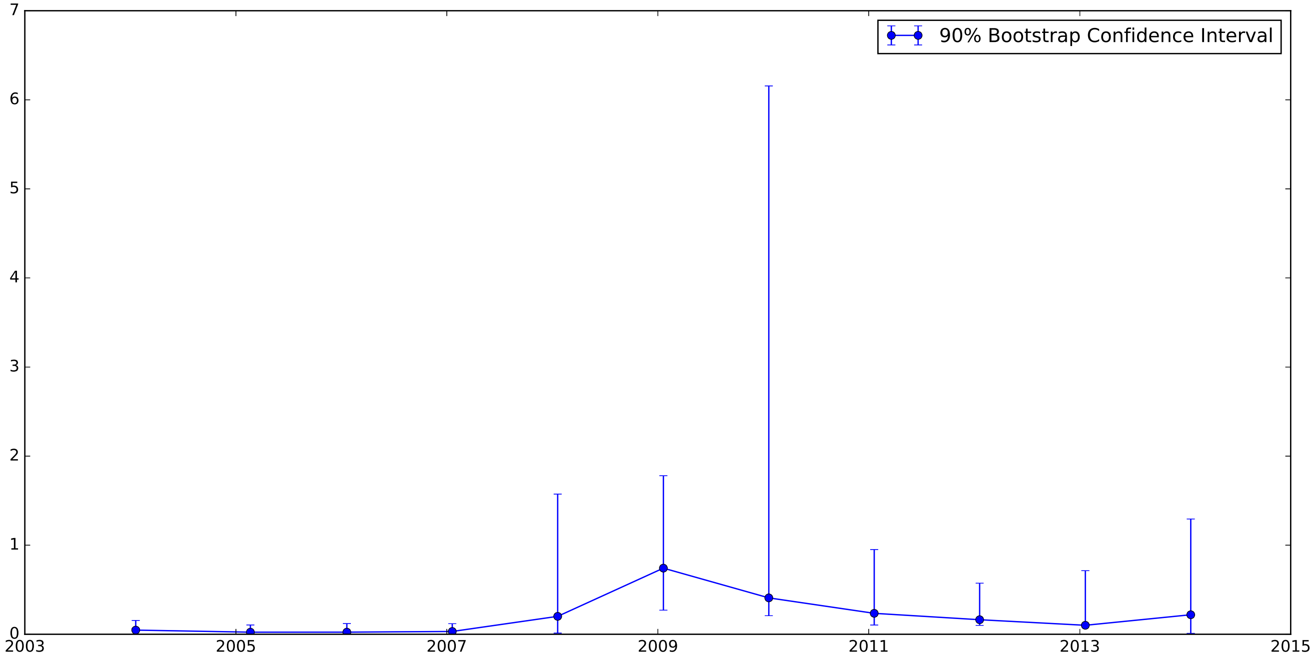


Figure 9. 90% bootstapping confidence intervals of market absolute connectedness strength "out" measures. For each date (20/01/2004, 20/01/2005, 20/01/2006, 20/01/2007, 20/01/2008, 20/01/2009, 20/01/2010, 20/01/2011, 20/01/2012, 20/01/2013, 20/01/2014), we construct a simple resampling bootstrap confidence interval of the market absolute connectedness strength "OUT" measure. The confidence level is set to be 90%. All measures are estimated by the VAR(1) models with 2-year rolling estimation windows. The nodes of the underlying market are the companies selected from the S&P 100 components.

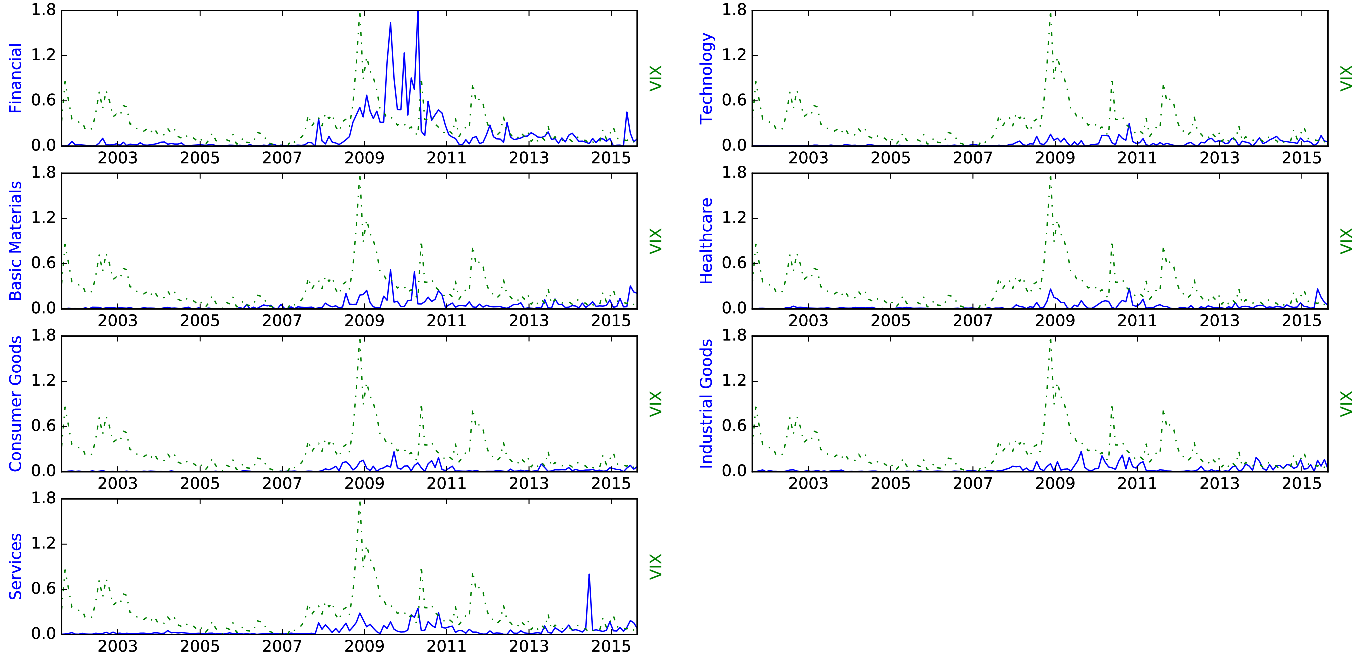


Figure 10. Sector connectedness "out" measures and the VIX index. The sector connectedness measures are defined as the connectedness measures within each of the 7 sectors (Financial, Technology, Basic Materials, Healthcare, Consumer Goods, Industrial Goods and Services). The solid blue lines are our absolute sector connectedness "out" strength measures and the green dash lines are the VIX index. All measures are estimated every 1 month by the VAR(1) models with 2-year rolling estimation windows. The nodes of the underlying market are the companies selected from the S&P 100 components.

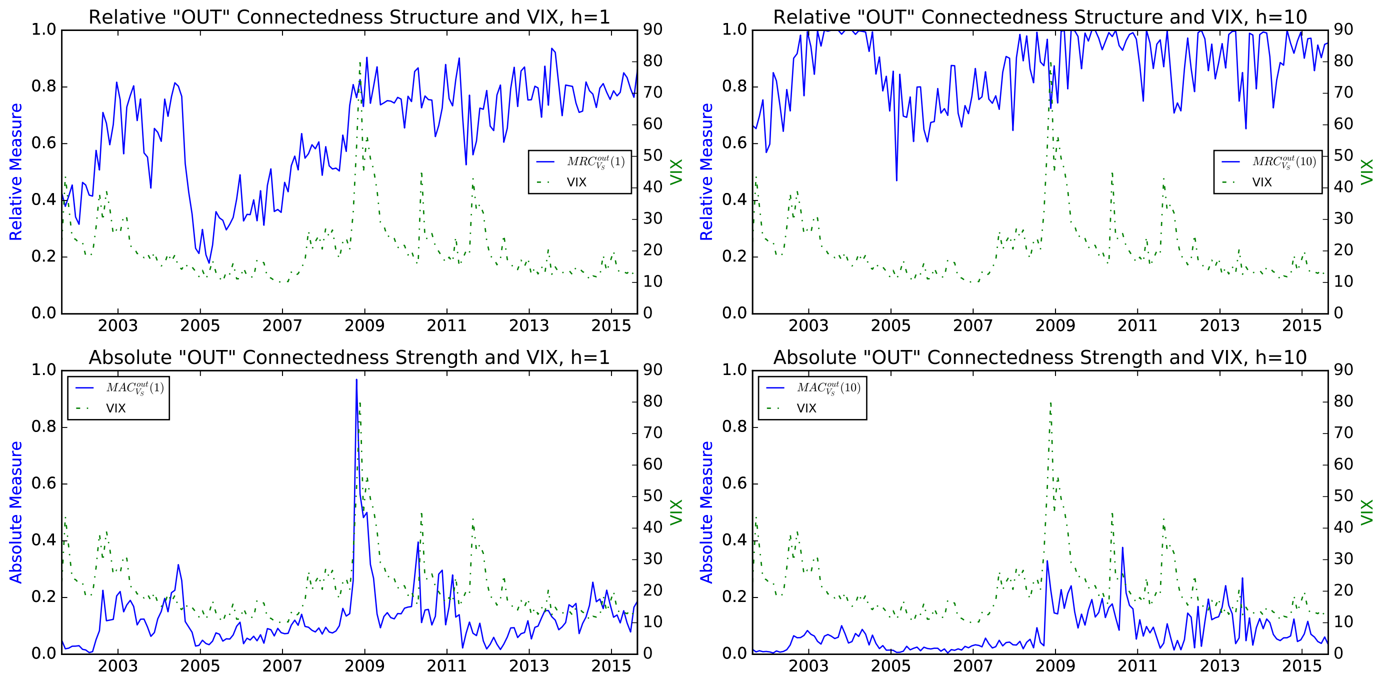


Figure 11. Sector-wise market connectedness “out” measures and the VIX index. The blue solid lines are our “out” connectedness measures (relative connectedness structure: upper row; absolute connectedness strength: bottom row) and the green dash lines are the VIX index. The reported connectedness measures are estimated at forecast horizon 1 (left column) and at forecast horizon 10 (right column). All measures are estimated every 1 month by the VAR(1) models with 2-year rolling estimation windows. The nodes of the underlying market are the sectors whose companies are selected from the S&P 100 components.

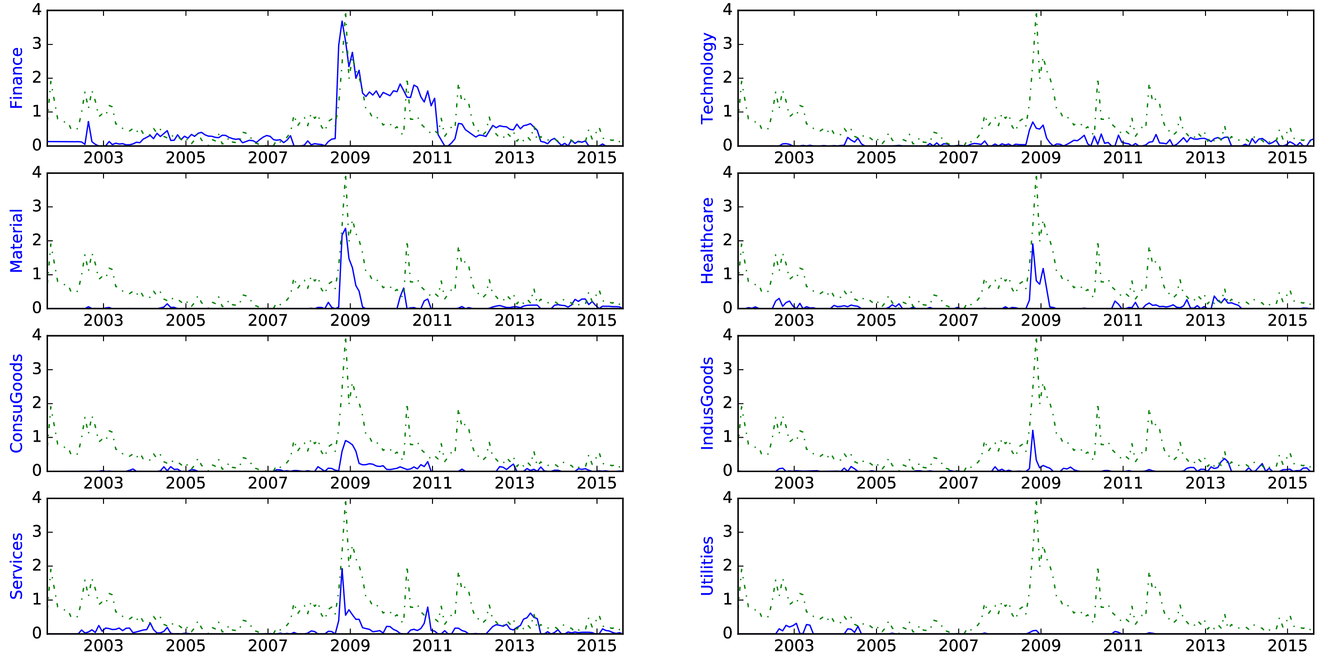


Figure 12. Sector-wise causality measures from the financial sector to other sectors. The blue solid line is our causality measures and the green dash line is the VIX index. This figure reports the direct effects from the financial sector on other sectors. The measures are estimated at forecast horizon $h = 1$ ($C^1_{\text{Fin}\to\cdot}$). All measures are estimated every 1 month by the VAR(1) models with 2-year rolling estimation windows. The nodes of the underlying market are the sectors whose companies are selected from the S&P 100 components.

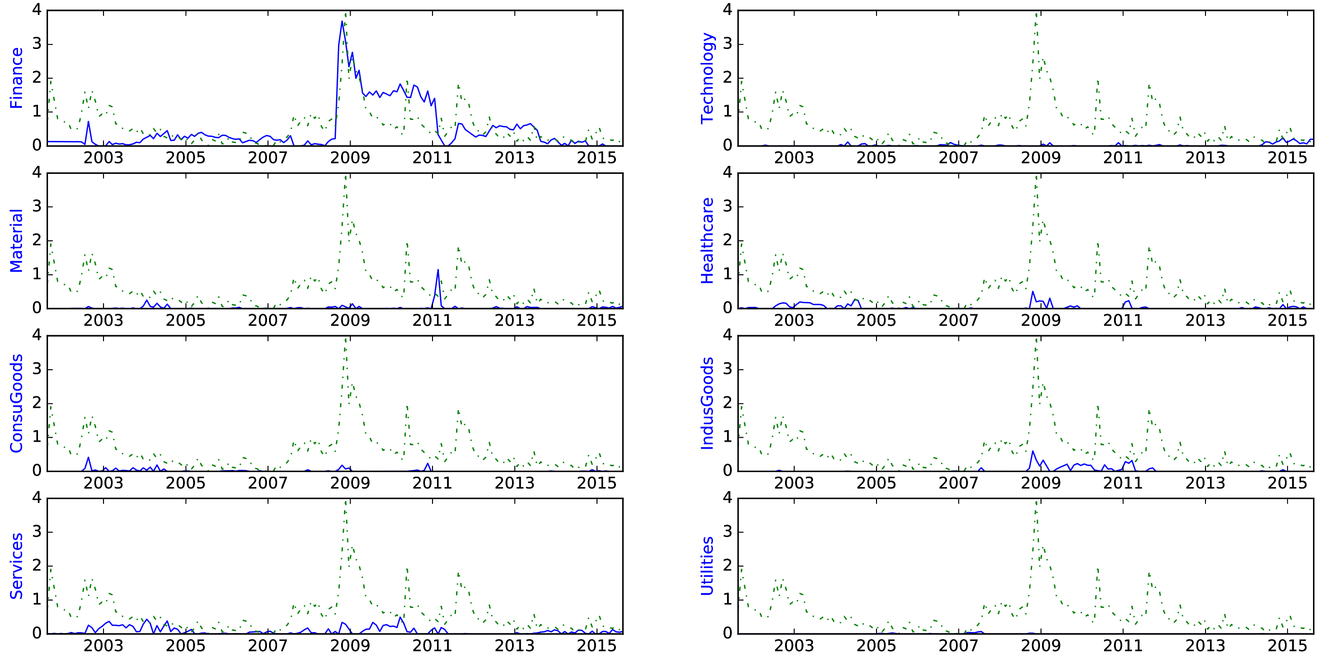


Figure 13. Sector-wise causality measures to the financial sector from other sectors. The blue solid line is our causality measures and the green dash line is the VIX index. This figure reports the direct effects from other sectors to the financial sector. The measures are estimated at forecast horizon $h = 1$ ($C^{1}_{\cdot\to\mathrm{Fin}}$). All measures are estimated every 1 month by the VAR(1) models with 2-year rolling estimation windows. The nodes of the underlying market are the sectors whose companies are selected from the S&P 100 components.

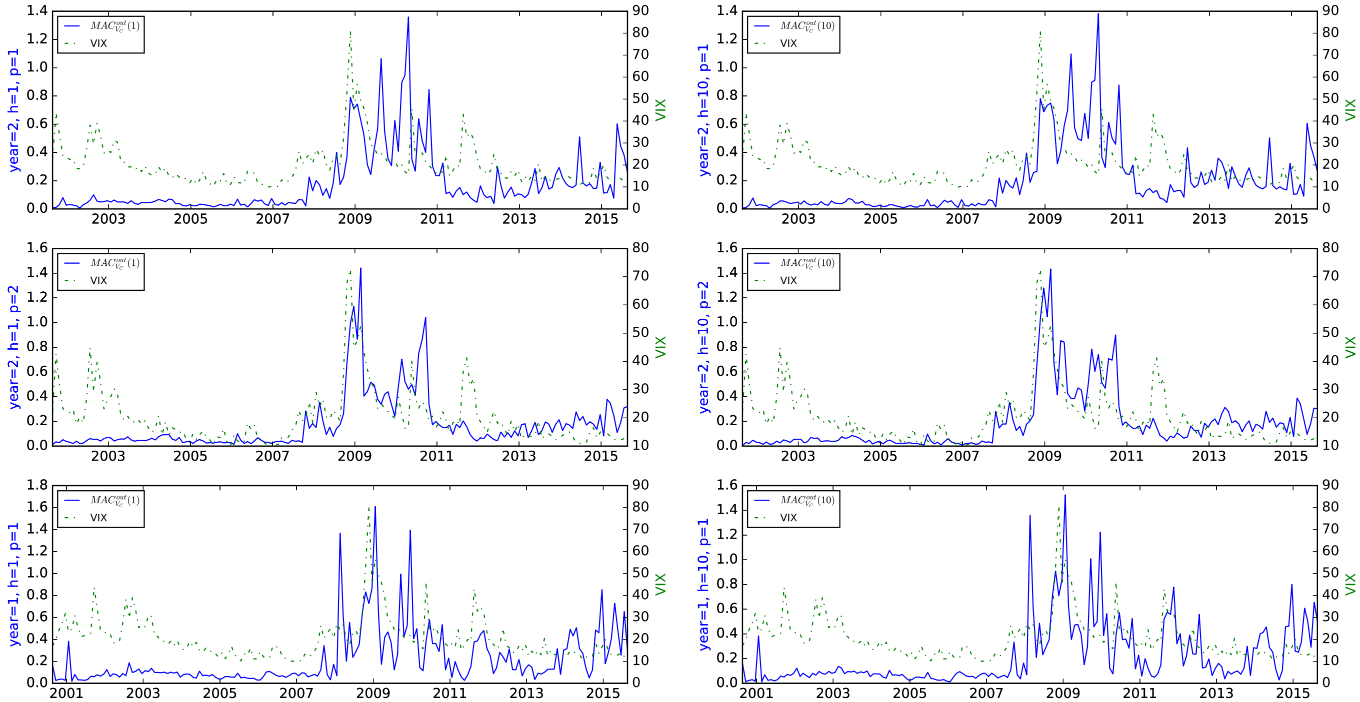


Figure 14. Robustness of firm-wise absolute market connectedness strength "out" measures. The blue solid lines are our absolute market connectedness strength "out" measures and the green dash lines are the VIX index. All measures are estimated every 1 month. The nodes of the underlying market are the companies selected from the S&P 100 components. The upper row reports the measures estimated with the VAR(1) model (p=1) and with 2-year rolling estimation windows (year=2). The middle row reports the measures estimated with the VAR(2) model (p=2) and with 2-year rolling estimation windows (year=2). The bottom row reports the measures estimated with the VAR(1) model (p=1) and with 1-year rolling estimation windows (year=1). The left column reports the measures estimated at forecast horizon 1 (h=1) and the right column reports the measures estimated at forecast horizon 10 (h=10).

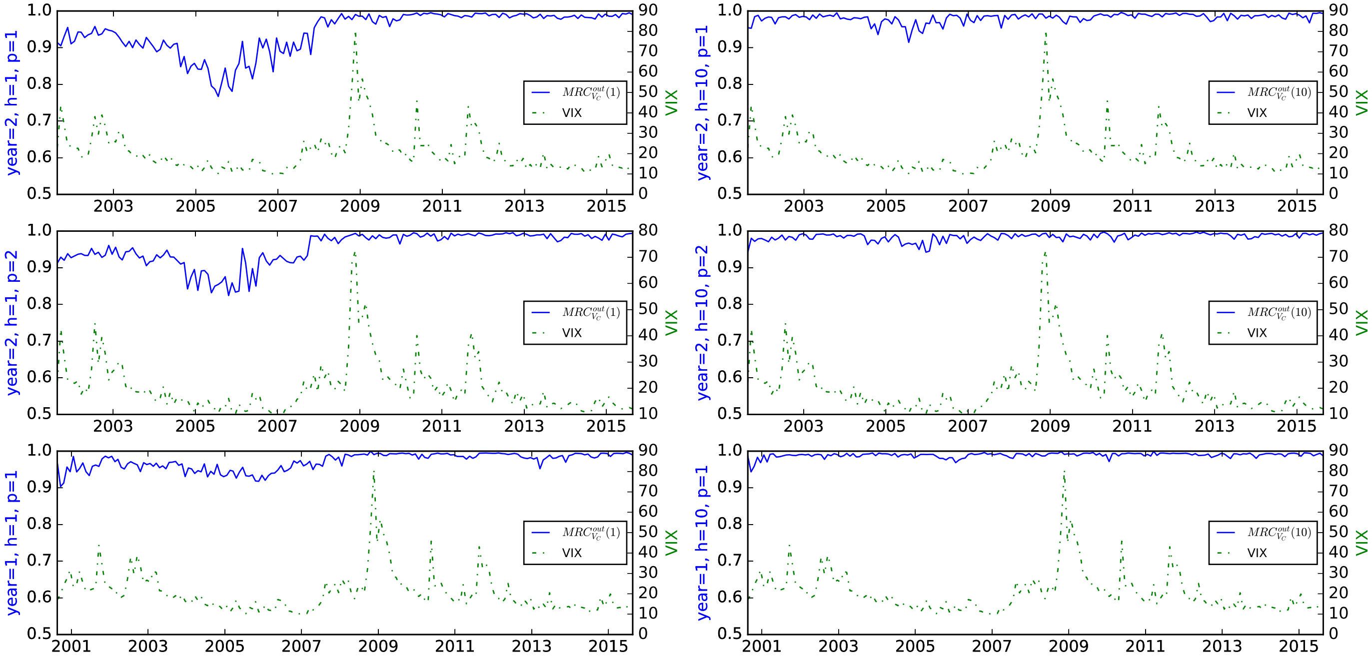


Figure 15. Robustness of firm-wise relative market connectedness structure “out” measures. The blue solid lines are our firm-wise relative market connectedness structure “out” measures and the green dash lines are the VIX index. All measures are estimated every 1 month. The nodes of the underlying market are the companies selected from the S&P 100 components. The upper row reports the measures estimated with the VAR(1) model (p=1) and with 2-year rolling estimation windows (year=2). The middle row reports the measures estimated with the VAR(2) model (p=2) and with 2-year rolling estimation windows (year=2). The bottom row reports the measures estimated with the VAR(1) model (p=1) and with 1-year rolling estimation windows (year=1). The left column reports the measures estimated at forecast horizon 1 (h=1) and the right column reports the measures estimated at forecast horizon 10 (h=10).