

**Matteo Barigozzi and Marc Hallin**

## A network analysis of the volatility of high-dimensional financial series

**Article (Accepted version)  
(Refereed)**

**Original citation:** Barigozzi, Matteo and Hallin, Marc (2016) *A network analysis of the volatility of high-dimensional financial series*. [Journal of the Royal Statistical Society: Series C \(Applied Statistics\)](#) . ISSN 0035-9254  
DOI: [10.1111/rssc.12177](https://doi.org/10.1111/rssc.12177)

© 2016 [Royal Statistical Society](#)

This version available at: <http://eprints.lse.ac.uk/67456/>

Available in LSE Research Online: November 2016

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's final accepted version of the journal article. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

# A Network Analysis of the Volatility of High-Dimensional Financial Series

MATTEO BARIGOZZI<sup>1</sup>      MARC HALLIN<sup>2</sup>

## Abstract

Interconnectedness between stocks and firms plays a crucial role in the volatility contagion phenomena that characterise financial crises, and graphs are a natural tool in their analysis. In this paper, we are proposing graphical methods for an analysis of volatility interconnections in the Standard & Poor's 100 dataset during the period 2000-2013, which contains the 2007-2008 Great Financial Crisis. The challenges are twofold: first, volatilities are not directly observed and have to be extracted from time series of stock returns; second, the observed series, with about 100 stocks, is high-dimensional, and curse of dimensionality problems are to be faced. To overcome this double challenge, we propose a dynamic factor model methodology, decomposing the panel into a factor-driven and an idiosyncratic component modelled as a sparse vector autoregressive model. The inversion of this autoregression, along with suitable identification constraints, produces networks in which, for a given horizon  $h$ , the weight associated with edge  $(i, j)$  represents the  $h$ -step-ahead forecast error variance of variable  $i$  accounted for by variable  $j$ 's innovations. Then, we show how those graphs yield an assessment of how *systemic* each firm is. They also demonstrate the prominent role of financial firms as sources of contagion during the 2007-2008 crisis.

*Keywords:* Dynamic Factor Models, Sparse Vector Autoregression Models, Volatility, Systemic Risk, Standard & Poor's 100 index.

## 1 Introduction

The study of networks as complex systems has been the subject of intensive research in recent years, both in the physics and statistics communities (see, for example, Kolaczyk, 2009, for a review of the main models, methods, and results). Typically, the datasets considered in that literature exhibit a “natural” or pre-specified network structure, as, for example, world trade fluxes (Serrano and Boguñá, 2003; Barigozzi et al., 2010), co-authorship relations (Newman, 2001), power grids (Watts and Strogatz, 1998), social individual relationships (Zachary, 1977), fluxes of migrants (Fagiolo and Santoni, 2015), or political weblog data (Adamic and Glance, 2005). In all those studies, the network structure (as a collection of vertices and edges) is known, and pre-exists the observations. More recently, in the aftermath of the Great Financial Crisis of 2007-2008, networks also have become a popular tool in financial econometrics and, more particularly, in the study of the interconnectedness of financial markets (Diebold and

---

<sup>1</sup>m.barigozzi@lse.ac.uk – London School of Economics and Political Science, Department of Statistics, UK.

<sup>2</sup>mhallin@ulb.ac.be – ECARES, Université libre de Bruxelles CP114/4 B-1050 Brussels, Belgium.

Yilmaz, 2014). In this case, however, the data under study—usually time series of stock returns and volatilities—do not have any particular pre-specified network structure, and the graphical structure (viz., the collection of edges) of interest has to be recovered or estimated from the data.

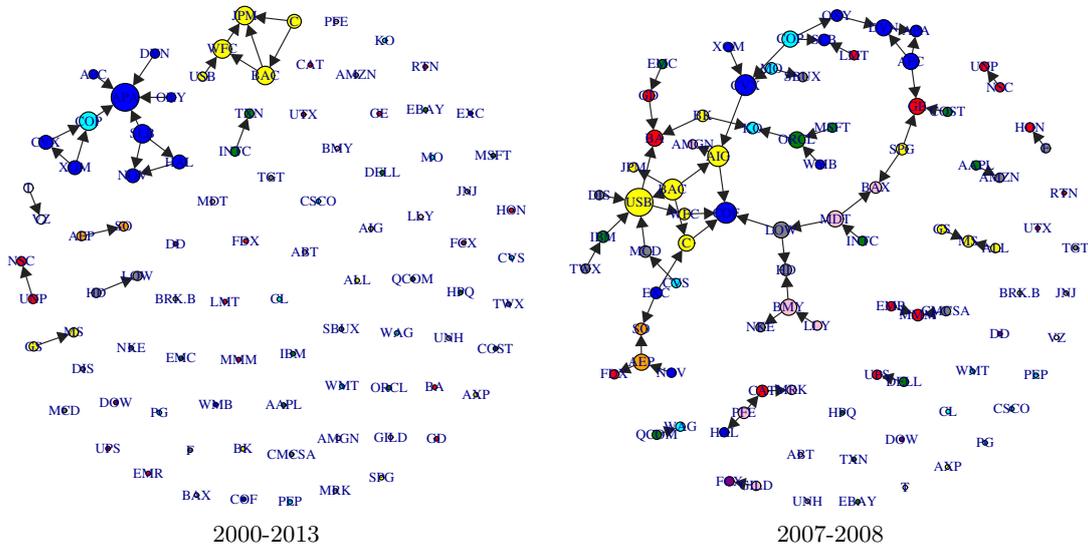
In this paper, we focus on one particular network structure: the Long-Run Variance Decomposition Network (LVDN). Following Diebold and Yilmaz (2014), the LVDN, jointly with appropriate identification assumptions, defines, for a given horizon  $h$ , a weighted and directed graph where the weight associated with edge  $(i, j)$  represents the proportion of  $h$ -step ahead forecast error variance of variable  $i$  which is accounted for by the innovations in variable  $j$ . Therefore, by definition, LVDNs are completely characterised by the infinite vector moving average (VMA) representation given by Wold’s classical representation theorem.

Classical network-related quantities as *in* and *out node strength* or *centrality*, computed for a given LVDN, then admit an immediate economic interpretation in indicating, for instance, which are the stocks or the firms most affected by a global extreme event, and which are the stocks or the firms that, when hit by some extreme shock, are likely to transmit it and spread it over to the whole market. The weights attached to each edge provide a quantitative assessment of the risks attached to such events. That type of network-based analysis is of particular relevance in financial econometrics—see, among others, Billio et al. (2012), Acemoglu et al. (2015), Hautsch et al. (2014, 2015), or Diebold and Yilmaz (2014).

Throughout, we are concentrating on the analysis of the LVDN associated with a panel of daily volatilities of the stocks constituting the Standard & Poor’s 100 (S&P100) index. The observed period is from 3rd January 2000 to 30th September 2013. The stocks considered belong to ten different sectors: Consumer Discretionary, Consumer Staples, Energy, Financial, Health Care, Industrials, Information Technology, Materials, Telecommunication Services, Utilities.

Our main findings are illustrated in Figure 1 which shows, for our dataset, the estimated LVDNs associated (a) with the period 2000-2013, and (b) with the years 2007-2008, which witnessed the so-called “Great Financial Crisis”. Inspection of these graphs reveals a main role of the Energy (blue nodes) and Financial (yellow nodes) sectors. The interconnections within and between those two sectors had a prominent role in the period considered, due to the high energy prices in the years 2005-2007 and the Great Financial Crisis of the years 2007-2008. In particular, when focussing only on the 2007-2008 period, the Financial stocks appear to be the most central ones (in the sense of the eigenvector centrality concept of Bonacich and Lloyd, 2001), and the connectedness of all network structures considered increases quite sizeably, making the whole system considerably more prone to contagion. This increased connectedness is an unsurprising phenomenon, since volatility measures *fear* or *lack of confidence* of investors, which tends to spread during periods of high uncertainty.

The LVDNs in Figure 1 are estimated in two steps. First, we obtain what we call the *idiosyncratic components* of volatilities by removing from the data the pervasive influence of global volatility shocks, which we refer to as *common shocks* or, given the present financial context, as *market shocks*. This is done by applying the general dynamic factor methodology recently proposed by Forni et al. (2015a,b) and adapted by Barigozzi and Hallin (2016) to a study of volatilities. More precisely, the factor model structure we are considering here is the Generalised Dynamic Factor model (GDFM) originally proposed by Forni et al. (2000) and Forni and Lippi (2001). In a second step, the LVDN is obtained by estimating and inverting a sparse vector autoregression (VAR) for the resulting idiosyncratic components, together with suitable identifying constraints. In particular, we consider VAR estimation based on three



**Figure 1:** Graphs of estimated LVDN for the S&P100 idiosyncratic volatilities.

different methods: elastic net (Zou and Hastie, 2005), group lasso (Yuan and Lin, 2006), and adaptive lasso (Zou, 2006). We call this estimation approach the “factor plus sparse VAR” approach.

This paper gives two main contributions to the existing financial literature on networks. First, we show that a combination of dynamic factor analysis and penalised regressions provides an ideal tool in the analysis of volatility and interconnectedness in large panels of financial time series. Second, we generalise to the high-dimensional setting the LVDN estimation originally proposed by Diebold and Yilmaz (2014) for a small number of series.

There are strong reasons in favour of our approach controlling for market volatility shocks— as opposed to a direct sparse VAR analysis that does not control for those shocks. The main motivation is of an economic nature; but forecasting and empirical motivations are important as well.

(1) *Economic motivation.* In the financial context, the pertinence of factor models is a direct consequence of the Arbitrage Pricing Theory (APT) and the related Capital Asset Pricing Model (CAPM) (Ross, 1976; Fama and French, 1993). These models allow us to disentangle and identify the main sources of variation driving large panels of financial time series: (i) a strongly pervasive component typically driven by few common shocks, or factors, affecting the whole market, and (ii) an idiosyncratic weakly connected component driven by local, or sectoral, shocks. In agreement with APT, the factor-driven, or common, component represents the *non-diversifiable* or *systematic* part of risk, while the idiosyncratic component, becomes perfectly *diversifiable* as the dimension of the system grows (Chamberlain and Rothschild, 1983). Many studies, however, provide evidence that connectivity in the idiosyncratic component, although milder than in the common component, still may be quite non-negligible, even in large dimensional systems (see for example the empirical and theoretical results by Jovanovic, 1987; Gabaix, 2011; Acemoglu et al., 2012). This is bound to happen in highly interconnected systems as financial markets. Therefore, when an exceptionally large shock, such as bankruptcy, affects the idiosyncratic component of a particular stock, that shock,

although idiosyncratic, subsequently is likely to spread, and hit, eventually, all idiosyncratic components across the system. Such events are called *systemic*, and diversification strategies against them might be ineffective. Studying the LVDN related to the idiosyncratic volatilities, hence after controlling for market shocks, helps identifying the systemic elements of a panel of times series, and provides the basis for an analysis of contagion mechanisms.

(2) *Forecasting motivation.* It has been shown (see e.g. De Mol et al., 2008) that forecasts obtained via penalised regression are highly unstable in the presence of collinearity. Thus, even though forecasting is not the main goal of this paper, removing the effect of common shocks before turning to sparse VAR estimation methods seems highly advisable.

(3) *Empirical motivation.* When considering partial dependencies, measured by partial spectral coherence, in the idiosyncratic component, that is after common components have been removed, hidden dependencies between and within the Financial and Energy sectors are uncovered. This finding, documented in Section 5, is the empirical justification for preferring a “factor plus sparse VAR” approach rather than a direct application of sparse VAR methods.

In Section 2, we introduce the GDFM for large panels of time series and the definition of LVDN for the idiosyncratic components, allowing for a study of different sources of interdependencies. In Section 3 we discuss estimation. In Section 4, following Barigozzi and Hallin (2016), we show how to extract from financial returns those volatility proxies which will be the object of our analysis. Section 5 presents the results for the S&P100 dataset. A detailed list of the series considered and complementary results are provided in the Web-based supporting materials.

## 2 Factors and networks in large panels of financial time series

We consider large panels of time series data, namely, observed finite realizations, of the form  $\{Y_{it} | i = 1, \dots, n, t = 1, \dots, T\}$ , of some stochastic process  $\mathbf{Y} := \{Y_{it} | i \in \mathbb{N}, t \in \mathbb{Z}\}$ ;  $i$  is a cross-sectional index and  $t$  stands for time. Both the cross-sectional dimension  $n$  and the sample size or series length  $T$  are large and, in asymptotics, we consider sequences of  $n$  and  $T$  values tending to infinity. The notation  $\mathbf{Y}_n := \{\mathbf{Y}_{nt} = (Y_{1t}, Y_{2t}, \dots, Y_{nt})' | t \in \mathbb{Z}\}$  in the sequel is used for the  $n$ -dimensional subprocess of  $\mathbf{Y}$ ; the same notation is used for all  $n$ -dimensional vectors. In this section,  $\mathbf{Y}_n$  stands for a generic panel of time series. Since our interest is to study connections and interdependencies responsible for the contagion phenomena that might lead to financial crises, then in Sections 4 and 5, we will apply the definitions and results presented here to the case of financial volatilities.

In principle, as shown by Diebold and Yilmaz (2014), a LVDN can be estimated via classical VAR estimation. However, when dealing with high-dimensional systems, VAR estimation is badly affected by curse of dimensionality problems, and adequate estimation techniques have to be considered. The most frequent strategy, in presence of large- $n$  datasets, is based on sparsity assumptions allowing for the application of penalised regression techniques (Hsu et al., 2008; Abegaz and Wit, 2013; Nicholson et al., 2014; Basu and Michailidis, 2015; Davis et al., 2015; Kock and Callot, 2015; Barigozzi and Brownlees, 2016; Gelper et al., 2016). The presence of pervasive shocks affecting the large panels of time series considered in macro-econometrics and finance has motivated the development of another dimension reduction technique, the so-called dynamic factor methods. Various versions of those methods have become daily practice in many areas of econometrics and finance; among these the GDFM by Forni et al. (2000) and Forni and Lippi (2001) is the most general, while most other factor models considered in the

time series literature (Stock and Watson, 2002; Bai and Ng, 2002; Lam and Yao, 2012; Fan et al., 2013, to quote only a few) are particular cases.

## 2.1 The Generalised Dynamic Factor Model

In order to introduce the *dynamic factor representation* for  $\mathbf{Y}$ , we make the following assumptions.

**Assumption 1.** *For all  $n \in \mathbb{N}$ , the vector process  $\mathbf{Y}_n$  is second-order stationary, with mean zero and finite variances.*

**Assumption 2.** *For all  $n \in \mathbb{N}$ , the spectral measure of  $\mathbf{Y}_n$  is absolutely continuous with respect to the Lebesgue measure on  $[-\pi, \pi]$ , that is,  $\mathbf{Y}_n$  admits a full-rank (for any  $n$  and  $\theta$ ) spectral density matrix  $\Sigma_{\mathbf{Y};n}(\theta)$ ,  $\theta \in [-\pi, \pi]$  with uniformly (in  $i, j, \theta$ , and  $n$ ) bounded entries  $\sigma_{\mathbf{Y};ij}(\theta)$ .*

Denote by  $\lambda_{\mathbf{Y};n,1}(\theta), \dots, \lambda_{\mathbf{Y};n,n}(\theta)$ ,  $\theta \in [-\pi, \pi]$ , the eigenvalues (in decreasing order of magnitude) of  $\Sigma_{\mathbf{Y};n}(\theta)$ ; the mapping  $\theta \mapsto \lambda_{\mathbf{Y};n,i}(\theta)$  is also called  $\mathbf{Y}_n$ 's  *$i$ th dynamic eigenvalue*. We say that  $\mathbf{Y}$  admits a General Dynamic Factor representation with  $q$  factors if for all  $i$  the process  $\{Y_{it}\}$  decomposes into a *common* component  $\{X_{it}\}$  and an *idiosyncratic* one  $\{Z_{it}\}$ ,

$$Y_{it} = X_{it} + Z_{it} =: \sum_{k=1}^q b_{ik}(L)u_{kt} + Z_{it}, \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \quad (1)$$

such that

- (i) the  $q$ -dimensional vector process of factors  $\mathbf{u} := \{\mathbf{u}_t = (u_{1t} \ u_{2t} \ \dots \ u_{qt})' \mid t \in \mathbb{Z}\}$  is orthonormal zero-mean white noise;
- (ii) the filters  $b_{ik}(L)$  are one-sided and square-summable for all  $i \in \mathbb{N}$  and  $k = 1, \dots, q$ ;
- (iii) the  $q$ th dynamic eigenvalue  $\lambda_{\mathbf{X};n,q}(\theta)$  of  $\mathbf{X}_n$  diverges  $\theta$ -almost everywhere ( $\theta$ -a.e.) in the interval  $[-\pi, \pi]$  as  $n \rightarrow \infty$ ;
- (iv) the first dynamic eigenvalue  $\lambda_{\mathbf{Z};n,1}(\theta)$  of  $\mathbf{Z}_n$  is bounded ( $\theta$ -a.e. in  $[-\pi, \pi]$ ) as  $n \rightarrow \infty$ ;
- (v)  $Z_{k,t_1}$  and  $u_{h,t_2}$  are mutually orthogonal for any  $k, h, t_1$  and  $t_2$ ;
- (vi)  $q$  is minimal with respect to (i)-(v).

For any  $n$ , we can write (1) in vector notation as

$$\mathbf{Y}_{nt} = \mathbf{X}_{nt} + \mathbf{Z}_{nt} =: B_n(L)\mathbf{u}_t + \mathbf{Z}_{nt}, \quad n \in \mathbb{N}, \quad t \in \mathbb{Z}. \quad (2)$$

This actually defines the GDFM. In this model, the common and idiosyncratic components are identified by means of the following assumption on  $\mathbf{Y}_n$ 's dynamic eigenvalues.

**Assumption 3.** *The  $q$ th eigenvalue  $\lambda_{\mathbf{Y};n,q}(\theta)$  of  $\Sigma_{\mathbf{Y};n}(\theta)$  diverges,  $\theta$ -a.e. in  $[-\pi, \pi]$ , while the  $(q+1)$ th one,  $\lambda_{\mathbf{Y};n,q+1}(\theta)$ , is  $\theta$ -a.e. bounded, as  $n \rightarrow \infty$ .*

More precisely, we know from Forni et al. (2000) and Forni and Lippi (2001) that, given Assumptions 1 and 2, Assumption 3 is necessary and sufficient for the process  $\mathbf{Y}$  to admit the dynamic factor representation (1). Hallin and Lippi (2014) moreover provide very weak time-domain primitive conditions under which (1), hence Assumption 3, holds for some  $q < \infty$ .

Finally, the idiosyncratic component  $\mathbf{Z}_n$  always admits a Wold decomposition which, after adequate transformation, yields the vector moving average (VMA) representation

$$\mathbf{Z}_{nt} := D_n(L)\mathbf{e}_{nt}, \quad t \in \mathbb{Z}, \quad \mathbf{e}_{nt} \sim w.n.(\mathbf{0}, I_n), \quad (3)$$

where  $D_n(L) = \sum_{k=0}^{\infty} D_{nk}L^k$  is a square-summable power series in the lag operator  $L$ . Notice that although the magnitude of those coefficients is bounded by condition (iv) above, no sparsity assumption is made.

## 2.2 The Long-Run Variance Decomposition Network

In order to study the interdependencies among series, and following traditional econometric analysis, we focus on the reactions of observed variables to unobserved shocks, i.e. impulse response functions. Large panels of financial time series are affected by market shocks that are, essentially, common to all stocks and represent the non-diversifiable components of risk, and by idiosyncratic shocks that are specific to one or a few stocks in the panel. The GDFM is the ideal tool for disentangling those two sources of variation: (i) the  $q$  market shocks  $\mathbf{u}$  and their impulse responses  $B_n(L)$ , defined in (2), and (ii) the  $n$  idiosyncratic shocks  $\mathbf{e}_n$  and their impulse responses  $D_n(L)$ , defined in (3).

Our focus here is mainly on idiosyncratic shocks. Indeed, once we control for market effects, the study of interdependencies among different stocks is strictly related to systemic risk measures, that is, individual measures of how one given stock is likely to be affected by, and/or is likely to affect, all others.

To this end, representation (3) is what we need, as it characterises all (linear) interdependencies between the components of  $\mathbf{Z}_n$ . In particular,  $D_{n0}$  characterises contemporaneous dependencies, while  $D_{nk}$  for  $k > 0$  characterises the dynamic dependencies with lag  $k$ . Denote by  $d_{k,ij}$  the  $(i, j)$ th entry of  $D_{nk}$ . Then, following Diebold and Yilmaz (2014), we can summarise all dependencies up to lag  $h$  by means of the forecast error variance decomposition and, more particularly, by the ratios

$$w_{ij}^h := 100 \left( \frac{\sum_{k=0}^{h-1} d_{k,ij}^2}{\sum_{\ell=1}^n \sum_{k=0}^{h-1} d_{k,i\ell}^2} \right), \quad i, j = 1, \dots, n. \quad (4)$$

The ratio  $w_{ij}^h$  is the percentage of the  $h$ -step ahead forecast error variance of  $\{Y_{it}\}$  accounted for by the innovations in  $\{Y_{jt}\}$ . Note that, by definition,

$$\frac{1}{100} \sum_{j=1}^n w_{ij}^h = 1 \quad \text{for any } i, \text{ hence } \frac{1}{100} \sum_{i,j=1}^n w_{ij}^h = n.$$

The LVDN is then defined by the set of edges

$$\mathcal{E}_{\text{LVDN}} := \left\{ (i, j) \in \{1 \dots n\}^2 \mid w_{ij}^{\text{LVDN}} := \lim_{h \rightarrow \infty} w_{ij}^h \neq 0 \right\}. \quad (5)$$

In practice an horizon  $h$  has to be chosen to compute those weights, and the operational definition of the LVDN therefore also depends on that  $h$ .

Three measures of connectedness can be based on the quantities defined in (4). First, we define the *from-degree* of component  $i$  and *to-degree* of component  $j$  (also called *in-strength* of node  $i$  and *out-strength* of node  $j$ ) as

$$\delta_i^{\text{From}} := \sum_{j=1, j \neq i}^n w_{ij}^h, \quad i = 1, \dots, n \quad \text{and} \quad \delta_j^{\text{To}} := \sum_{i=1, i \neq j}^n w_{ij}^h, \quad j = 1, \dots, n, \quad (6)$$

respectively. As pointed out by Diebold and Yilmaz (2014), these two measures are closely related to two classical measures of systemic risk considered in the financial literature. The from-degree is directly related with the so-called *marginal expected shortfall* and *expected capital shortfall* (of series  $i$ ), which measure the exposure of component  $i$  to extreme events affecting all other components (see Acharya et al., 2012, for a definition of these measures). As for the to-degree, it is related to *co-Value-at-Risk*, which measures the effect on the whole panel of an extreme event affecting component  $j$  (see Adrian and Brunnermeier, 2016, for a definition). Finally, we can define an overall measure of connectedness by summing all from-degrees (equivalently, all to-degrees):

$$\delta^{\text{Tot}} := \frac{1}{n} \sum_{i=1}^n \delta_i^{\text{From}} = \frac{1}{n} \sum_{j=1}^n \delta_j^{\text{To}}. \quad (7)$$

Given the economic interpretation of these quantities, the LVDN of the idiosyncratic component  $\mathbf{Z}_n$  seems to be an ideal tool for studying systemic risk and, for this reason, in the empirical study of Section 5, we mainly focus on the LVDN of volatilities (see also Section 4 for a motivation).

An LVDN also can be constructed for the common component  $\mathbf{X}_n$  by using a definition analogous to (4), but based on the entries of the matrix polynomial  $B_n(L)$ , defined in (2). However, due to the singularity of  $B_n(L)$ , definition (4) in this case does not measure the proportion of the  $h$ -step ahead forecast error variance of variable  $i$  accounted for by the innovations in variable  $j$ , but rather the proportion of the same forecast error variance explained by the  $j$ th market shock  $\{u_{jt}\}$ .

### 2.3 VAR representations

The LVDN of the idiosyncratic component  $\mathbf{Z}_n$  is defined from the coefficients of the VMA representation (3). That representation can be estimated as an inverted sparse VAR. We accordingly make the following assumption.

**Assumption 4.** *The idiosyncratic component  $\mathbf{Z}_n$  admits, for some  $p$  that does not depend on  $n$ , the VAR( $p$ ) representation*

$$F_n(L)\mathbf{Z}_{nt} = \mathbf{v}_{nt}, \quad t \in \mathbb{Z}, \quad \mathbf{v}_{nt} \sim w.n.(\mathbf{0}, C_n^{-1}), \quad (8)$$

where  $F_n(L) = \sum_{k=0}^p F_{nk}L^k$  with  $F_{n0} = I_n$  and  $\det(F_n(z)) \neq 0$  for any  $z \in \mathbb{C}$  such that  $|z| \leq 1$ , and  $C_n$  has full rank. Moreover, denoting by  $f_{k,ij}$  and  $c_{ij}$  the  $(i, j)$ th entries of  $F_{nk}$  and  $C_n$ ,

$$\max_{j=1 \dots n} \sum_{i=1}^n \mathbb{I}_{(f_{k,ij} \neq 0)} = o(n), \quad k = 1, \dots, p, \quad n \in \mathbb{N}, \quad (9)$$

$$\max_{i=1 \dots n} \sum_{j=1}^n \mathbb{I}_{(f_{k,ij} \neq 0)} = o(n), \quad k = 1, \dots, p, \quad n \in \mathbb{N}, \quad (10)$$

$$\max_{j=1 \dots n} \sum_{i=1}^n \mathbb{I}_{(c_{ij} \neq 0)} = o(n), \quad n \in \mathbb{N}. \quad (11)$$

The first part (8) of this assumption is quite mild, provided that  $p$  can be chosen large enough. The second part requires some further clarification. In (9)-(10), we require the VAR

coefficient matrices in (8) to have only a small number of non-zero entries. In this sense we say that the VAR representation (8) is sparse (see, for example, the definitions of sparsity in Bickel and Levina, 2008; Cai and Liu, 2011). That assumption is needed for the consistent estimation of (8) in the large- $n$  setting, and it is justified by the idea that in a GDFM, once we control for common shocks, most interdependencies among the elements of  $\mathbf{Z}_n$  are quite weak (since the corresponding dynamic eigenvalues are bounded as  $n \rightarrow \infty$ ). However, note that, while a sparse VAR is related to conditional second moments, the GDFM assumptions on the idiosyncratic component are based on unconditional second moments. For this reason, the GDFM assumptions do not imply a sparse VAR representation for  $\mathbf{Z}_n$ , and (9)-(10) are needed. Finally, for convenience, we parametrise the covariance matrix of the VAR innovations by means of its inverse  $C_n$  and in (11) we require this matrix to be sparse too, in accordance with the idea of a sparse global conditional dependence structure of idiosyncratic components.

As a by-product, a Long-Run Granger Causality Network (LGCN) can be defined by the set of edges

$$\mathcal{E}_{\text{LGCN}} := \left\{ (i, j) \in \{1 \dots n\}^2 \mid w_{ij}^{\text{LGCN}} := \sum_{k=0}^p f_{k,ij} \neq 0 \right\}. \quad (12)$$

This network captures the leading/lagging conditional dependencies of a given panel of time series. Such graphical representations of VAR dependencies were initially proposed by Dahlhaus and Eichler (2003) and Eichler (2007), and extend to a time-series context the graphical models for independent data considered by Dempster (1972), Meinshausen and Bühlmann (2006), Friedman et al. (2008), and Peng et al. (2009), to quote only a few.

Two comments are in order here. A network is said to be sparse if its weight matrix has many zero entries. First, notice that, under Assumption 4, the LGCN is likely to be sparse. On the other hand, the GDFM assumptions do not guarantee sparsity of the LVDN but only some weaker restrictions on the magnitude of its entries, as dictated by the boundedness of the eigenvalues of  $\mathbf{Z}_n$ 's spectral density matrix. Second, the economic interpretation of the LGCN is not as straightforward as that of the LVDN, and the LGCN therefore is of lesser interest for the analysis of financial systems: mainly, it will be a convenient tool in the derivation of the LVDN. This is in line with traditional macroeconomic analysis where impulse response functions, i.e. VMA coefficients, rather than VAR ones, are the object of interest for policy makers.

As for the common component  $\mathbf{X}_n$ , Forni et al. (2015a) show that it admits the singular VAR representation

$$A_n(L)\mathbf{X}_{nt} = H_n\mathbf{u}_t, \quad t \in \mathbb{Z}, \quad \mathbf{u}_t \sim w.n.(\mathbf{0}, I_q). \quad (13)$$

Assuming, without loss of generality, that  $n = m(q + 1)$  for some integer  $m$ , the VAR operator  $A_n(L)$  in (13) is block-diagonal, with  $(q + 1) \times (q + 1)$ -dimensional diagonal blocks of the form  $A^{(i)}(L) = \sum_{k=0}^{p_i} A_k^{(i)} L^k$  such that, for any  $i = 1, \dots, m$ ,  $A_0^{(i)} = I_n$  and  $\det(A^{(i)}(z)) \neq 0$  for any  $z \in \mathbb{C}$  such that  $|z| \leq 1$ . Moreover,  $H_n$  is a full-rank  $n \times q$  matrix, and  $\mathbf{u}$  is the  $q$ -dimensional process of common shocks defined in (2).

## 2.4 Identification

Starting from  $\mathbf{Z}_n$ 's VAR representation (8) and comparing it with (3), we have

$$D_n(L) = (F_n(L))^{-1} R_n, \quad (14)$$

where the full-rank matrix  $R_n$  is making the shocks  $R_n^{-1}\mathbf{v}_n =: \mathbf{e}_n$  orthonormal (such matrices under Assumption 4 exist). In other words, the LVDN is obtained from the inversion of the VAR in (8) by selecting a suitable transformation  $R_n$  meeting the required identification constraints. The simplest choice for  $R_n$  follows from a Choleski decomposition of the covariance  $C_n^{-1}$  of the shocks (see Sims, 1980)—namely, selecting  $R_n$  as the lower triangular matrix such that  $C_n^{-1} = R_n R_n'$ . Such a choice is appealing, as it is purely data-driven, but it depends on the ordering of the variables or, equivalently, on the ordering of the components of the shocks vector  $\mathbf{v}_n$ .

Many orderings are possible, and the one we propose is based on  $\mathbf{v}_n$ 's partial correlation structure. More precisely, assuming  $C_n$  to have full rank, the partial correlation between  $\{v_{it}\}$  and  $\{v_{jt}\}$  is

$$\rho^{ij} := \frac{-c_{ij}}{\sqrt{c_{ii}c_{jj}}}, \quad i, j = 1, \dots, n \quad (15)$$

where  $c_{ij}$  is the  $(i, j)$ th entry of  $C_n$ . Associated with this concept of partial correlation is the Partial Correlation Network (PCN), with edges

$$\mathcal{E}_{\text{PCN}} := \{(i, j) \in \{1 \dots n\}^2 \mid w_{ij}^{\text{PCN}} := \rho^{ij} \neq 0\}, \quad (16)$$

which, by Assumption 4, is a sparse network. The PCN can be studied by means of the  $n$  linear regressions

$$v_{it} = \sum_{j=1, j \neq i}^n \beta_{ij} v_{jt} + \nu_{it}, \quad t \in \mathbb{Z}, \quad \nu_{it} \sim w.n.(0, \sigma_i^2), \quad i = 1, \dots, n. \quad (17)$$

Indeed, it can be shown that  $\beta_{ij} = \rho^{ij} \sqrt{\sigma_i^2 / \sigma_j^2}$  (see, for example, Lemma 1 in Peng et al., 2009), so that  $(i, j) \in \mathcal{E}_{\text{PCN}}$  if and only if  $\beta_{ij} \neq 0$ . Thus, the inverse covariance matrix  $C_n$  of the VAR shocks is directly related to the partial correlation matrix of the VAR innovations, and this matrix in turn can be seen as the PCN weight matrix. We then order the shocks by decreasing order of eigenvector centrality (as in Bonacich, 1987) in that PCN, the most central component receiving label one, etc.

The centrality measure considered defines each node's centrality as the sum of the centrality values of its neighbouring nodes. It is easily seen that the mathematical translation of this idea leads to an eigenvector-related concept of centrality. That concept of eigenvector centrality differs from that of degree centrality (the number of neighbours of a given node); indeed, a node receiving many links does not necessarily have high eigenvector centrality, while a node with high eigenvector centrality is not necessarily highly linked (it might have few but important linkers).

It has to be noticed that usually eigenvector centrality is defined for networks in which the sign of the weight associated to a given edge is not taken into account, so negative and positive partial correlations have the same importance. However, based on the idea that contagion is more likely between nodes which are linked through positive weights rather than negative ones, we can also consider eigenvector centrality for a signed network, thus preserving the information about weights' signs. Both approaches are considered in the empirical analysis that follows.

This identification strategy seems well suited to the study of financial contagion. Indeed, in an impulse response exercise, we study the propagation of shocks through the system starting from lag zero up to a given lag  $h > 0$ . Thus, a given order of shocks defines which component

we choose to hit first. By ordering nodes according to their centrality in the PCN of VAR residuals, we are considering the case in which the most contemporaneously interconnected node is firstly affected by an unexpected shock, and then, by means of the subsequent impulse response analysis, we study the propagation of such shock through the whole system. The corresponding row of the LVDN adjacency matrix gives a summary, in terms of explained variance, of the effect of this propagation mechanism after  $h$  lags.

Finally, it is worth mentioning a few other possible methods for identifying the shocks in the LVDN. One could rank the series based on endogenous characteristics of the firms considered, such as market capitalisation. Another data-driven approach is taken in Swanson and Granger (1997), who propose to test the over-identifying restrictions implied by orthogonal shocks. That approach is closely related to ours as those restrictions involve partial correlations of the shocks; however as the dimension  $n$  of the problem increases, the related computational problems rapidly become non-tractable. Diebold and Yilmaz (2014) adopt generalised variance decompositions, a very popular method originally proposed by Koop et al. (1996). This approach, however, is only valid if the shocks have a Gaussian distribution, an assumption which is unlikely to hold true for the S&P100 dataset.

Turning to the common component  $\mathbf{X}_n$ , its LVDN is obtained by inverting the VAR representation (13). Namely, there exists a  $q \times q$  invertible matrix  $K$  such that

$$B_n(L) = (A_n(L))^{-1} H_n K. \quad (18)$$

Here  $K$  is required in order to identify the orthonormal market shocks  $\mathbf{u}$ . However, if  $q = 1$  as in the empirical analysis of Section 5, the choice of  $K$  reduces to that of a sign and a scale.

In the next sections, we first discuss the estimation of GDFMs and LVDNs under the general definition of this section, and then adapt to the particular case of unobserved volatilities.

### 3 Estimation

In this section we review estimation of (14) and (18). The numbers of factors throughout are determined via the information criterion proposed by Hallin and Liška (2007) and based on the behaviour of dynamic eigenvalues as the panel size grows from 1 to  $n$ . In this section we consider a generic observed panel  $\{Y_{it} \mid i = 1, \dots, n, t = 1, \dots, T\}$  of time series, with sample size  $T$  (in the next section these would be either returns or volatilities). Hereafter, we use the superscript  $T$  to denote estimated quantities.

#### 3.1 GDFM estimation

First we recover the common component using its autoregressive representation (13). For a given number of factors  $q$ , the method, described in detail by Forni et al. (2015a,b), is based on the following steps.

- (i) Estimate the spectral density matrix of  $\mathbf{Y}_n$ , denoted as  $\Sigma_{\mathbf{Y};n}^T(\theta)$ , for example using the smoothed periodogram estimator.
- (ii) Use the  $q$  largest dynamic principal components of  $\Sigma_{\mathbf{Y};n}^T(\theta)$  to extract the spectral density matrix of  $\mathbf{X}_n$ , denoted as  $\Sigma_{\mathbf{X};n}^T(\theta)$  (see Brillinger, 1981).
- (iii) Compute the autocovariances of  $\mathbf{X}_n$  by inverse Fourier transform of  $\Sigma_{\mathbf{X};n}^T(\theta)$  and use these to compute the Yule-Walker estimator of the VAR filters  $A_n^T(L)$ ; denote by  $\boldsymbol{\epsilon}_n^T := \{\epsilon_{it}^T \mid i = 1, \dots, n, t = 1, \dots, T\}$  the corresponding residuals.

- (iv) From the sample covariance of  $\boldsymbol{\epsilon}_n^T$ , compute the eigenvectors corresponding to its  $q$  largest eigenvalues, these are the columns of the estimator  $H_n^T$ ; then, by projecting  $\boldsymbol{\epsilon}_n^T$  onto the space spanned by the columns of  $H_n^T$ , obtain the  $q$ -dimensional vector  $\{\mathbf{u}_t^T | t = 1, \dots, T\}$ .
- (v) The estimated LVDN of the common component is given by  $(A_n^T(L))^{-1}H_n^TK$ , where  $K$  is a generic  $q \times q$  invertible matrix such that  $H_n^TK$  consistently estimates  $H_n$ .
- (vi) The estimated common and idiosyncratic components are respectively given by

$$\mathbf{X}_{nt}^T = (A_n^T(L))^{-1}H_n^T\mathbf{u}_t^T \quad \text{and} \quad \mathbf{Z}_{nt}^T = \mathbf{Y}_{nt} - \mathbf{X}_{nt}^T, \quad t = 1, \dots, T.$$

Details on each step and asymptotic properties of the estimators are given in Forni et al. (2015b). In particular, the parameters of the model are estimated consistently as  $n$  and  $T$  tend to infinity, at  $O_P(\max(n^{-1/2}, T^{-1/2}))$  rate.

### 3.2 Sparse VAR estimation

Once we have an estimator of the  $n$  idiosyncratic components, we estimate the VAR( $p$ ) representation (8) by minimising the penalised quadratic loss

$$\mathcal{L}_T = \sum_{t=1}^T \left( Z_{it}^T - \sum_{k=1}^p \mathbf{f}'_{k,i} \mathbf{Z}_{nt-k}^T \right)^2 + \mathcal{P}(\mathbf{f}_{1,i}, \dots, \mathbf{f}_{p,i}), \quad i = 1, \dots, n, \quad (19)$$

where  $\mathbf{f}'_{k,i}$  is the  $i$ -th row of  $F_{kn}$  and  $\mathcal{P}(\cdot)$  is some given penalty which depends on the chosen estimation method. In particular, we consider three alternative strategies.

- (i) Elastic net, as defined by Zou and Hastie (2005), where the penalty is a weighted average of ridge and lasso penalties:

$$\mathcal{P}^{EN}(\mathbf{f}_{1,i}, \dots, \mathbf{f}_{p,i}) = \lambda \|(\mathbf{f}'_{1,i} \dots \mathbf{f}'_{p,i})'\|_1 + (1 - \lambda) \|(\mathbf{f}'_{1,i} \dots \mathbf{f}'_{p,i})'\|_2^2.$$

- (ii) Adaptive lasso, as defined by Zou (2006), where the penalty is a lasso one but conditioned on a pre-estimators  $\tilde{\mathbf{f}}_{k,i}$  of the parameters (typically given by ridge or least squares estimators)

$$\mathcal{P}^{AL}(\mathbf{f}_{1,i}, \dots, \mathbf{f}_{p,i}) = \lambda \frac{\|(\mathbf{f}'_{1,i} \dots \mathbf{f}'_{p,i})'\|_1}{\|(\tilde{\mathbf{f}}'_{1,i} \dots \tilde{\mathbf{f}}'_{p,i})'\|_1}.$$

- (iii) Group lasso, as defined by Yuan and Lin (2006), where the explanatory variables are grouped before penalising, thus in a VAR context the groups are given by the lags of each variable, thus there are  $n$  groups, each of  $p$  elements, and we have

$$\mathcal{P}^{GL}(\mathbf{f}_{1,i}, \dots, \mathbf{f}_{p,i}) = \lambda \sqrt{p} \sum_{j=1}^n \|(f_{1,ij} \dots f_{p,ij})'\|_2.$$

The penalisation constant  $\lambda$ , in all three methods, and the maximum VAR lag  $p$ , are determined by minimising, over a grid of possible values, a BIC-type criterion.

Elastic net and adaptive lasso are particularly useful in a time series context since they are known to stabilise a simple lasso estimator which might suffer on instability due to serial dependence in the data. To the best of our knowledge, elastic net so far has not been considered in the estimation of high-dimensional VARs, while adaptive lasso has been used by Kock and Callot (2015) and Barigozzi and Brownlees (2016), among others. On the other hand, group

lasso in principle is likely to make the LGCN more sparse than the other two methods, and has been used, for VAR estimation, by Nicholson et al. (2014) and Gelper et al. (2016). Other possible penalised VAR estimators, which we do not consider here, include the simple lasso and smoothly clipped absolute deviation (SCAD) (see, for example, Hsu et al., 2008 and Abegaz and Wit, 2013). The consistency of all those methods is proved in the papers referenced when both  $n$  and  $T$  tend to infinity, under a variety of technical assumptions which we do not report here.

### 3.3 Identification of the LVDN

Once an estimator of the VAR coefficients is obtained, we can estimate, in a last step, the PCN of the VAR residuals, as defined in (16), which in turn can be used for LVDN identification. This again can be performed by means of several regularisation methods. Here we chose the approach proposed by Peng et al. (2009), where we refer to for technical details. More precisely, the weights of the PCN are obtained by estimating the regressions in (17) via traditional lasso (Tibshirani, 1996) in order to ensure sparsity as required by Assumption 4. Alternatively, the PCN of the residuals can be estimated jointly with the LGCN as originally proposed by Rothman et al. (2010) for cross-sectional regressions and, in a time series context, by Abegaz and Wit (2013), Gelper et al. (2016) or (with adaptive lasso penalty) Barigozzi and Brownlees (2016). Once we have estimated the PCN of VAR residuals, we can order them according to their centrality in the network. We then estimate the matrix  $R_n^T$ , needed for identification, as the Choleski factor of the sample covariance matrix of the ordered residuals. Finally, from  $R_n^T$  and the estimated VAR operators  $F_n^T(L)$ , we compute the VMA operator  $D_n^T(L) = (F_n^T(L))^{-1}R_n^T$ . The estimated LVDNs weights, denoted as  $w_{ij}^{hT}$ , readily follow from (4).

It has to be noted that, while by estimating a sparse VAR the LGCN by construction is sparse, this is not the case for the LVDN which, being derived from the *inverse* of a sparse VAR, does not necessarily have to be sparse. In other words, the assumption of sparsity of VAR coefficients is made for the purpose of dealing with the curse of dimensionality, and does not entail sparsity of the corresponding LVDN. The matrix  $W^{hT}$  with entries  $w_{ij}^{hT}$  (the LVDN adjacency matrix) is therefore not necessarily sparse. Nevertheless, since most of its entries are quite small, considering a sparse thresholded version  $W_\tau^{hT}$  with threshold  $\tau > 0$  is very natural. Here, we chose the threshold  $\tau$  minimising the sum  $\|(W^{hT}W^{hT'})^{-1/2}(W_\tau^{hT}W_\tau^{hT'})(W^{hT}W^{hT'})^{-1/2} - I_n\|_2$  of squared errors (see for example Fan et al., 2013, for a similar approach, although in a different context).

## 4 Network analysis of financial volatilities

The available datasets, in the study of interdependencies of financial institutions, in general, are (large) panels of stock returns. If our interest is in the systematic, i.e. market-driven, and systemic components of risk, then volatilities are what we need, not returns. Financial crises typically are characterised by unusually high levels of volatility generated by some major systemic events as the bankruptcy of some major institutions. Analysing interdependencies in volatility panels is the first step to a study of financial contagion (see Diebold and Yilmaz, 2014, for a thorough discussion).

Volatilities unfortunately are unobserved, and therefore must be estimated from the panels of returns. Many volatility proxies can be constructed from the series of returns as, for example, the adjusted log-range (Parkinson, 1980) from daily returns, or log-realised volatilities

(Andersen et al., 2003) based on intra-daily returns. Those proxies are generally treated as observed quantities, and nothing is said about the associated estimation error. On the other hand, volatilities can also be estimated from conditionally heteroschedastic models for financial returns such as multivariate GARCH models (see, for instance, the survey by Bauwens et al., 2006). Those are however parametric models which, due to the curse of dimensionality problem, cannot be handled in the present large- $n$  setting. We therefore follow the approach in Barigozzi and Hallin (2016) in adopting a global point of view, with a joint analysis of returns and volatilities in a high-dimensional setting. That analysis is based on a two-step dynamic factor procedure: a first GDFM procedure, applied to the panel of returns, is extracting a (double) panel of volatility proxies, which, in a second step, is analysed via a second GDFM. The LVDNs we are interested in are those of the common and idiosyncratic volatility components resulting from this second step.

More precisely, we consider a panel  $\mathbf{r}_n := \{\mathbf{r}_{nt} = (r_{1t} \ r_{2t} \ \dots \ r_{nt})' \mid t \in \mathbb{Z}\}$  of  $n$  stock returns (such as the constituents of the S&P100 index). We assume that  $\mathbf{r}_n$  satisfies Assumptions 1-3, i.e., admits the GDFM decomposition

$$\mathbf{r}_{nt} = \boldsymbol{\chi}_{nt} + \boldsymbol{\xi}_{nt}, \quad t \in \mathbb{Z}, \quad (20)$$

where  $\boldsymbol{\chi}_n$  is driven by  $q$  common shocks and  $\boldsymbol{\xi}_n$  is idiosyncratic—call them *level-common* and *level-idiosyncratic* components, respectively.

When applied to the real data in Section 5, the Hallin and Liška (2007) criterion very clearly yields  $q = 1$ , that is, a unique level-common shock. Thus the level-common component  $\boldsymbol{\chi}_n$  admits (see (13)) the autoregressive representation

$$\mathcal{A}_n(L)\boldsymbol{\chi}_{nt} =: \boldsymbol{\eta}_{nt} =: (\eta_{1t}, \dots, \eta_{nt})', \quad t \in \mathbb{Z}, \quad (21)$$

where  $\mathcal{A}_n(L)$  is a one-sided square-summable stable block-diagonal autoregressive filter with blocks of size  $2 \times 2$ , and  $\boldsymbol{\eta}_n$  is an  $n$ -dimensional white noise process with a singular covariance matrix of rank one.

We then assume that Assumption 4 holds for the level-idiosyncratic component  $\boldsymbol{\xi}_n$ , which thus admits the sparse VAR representation

$$\mathcal{F}_n(L)\boldsymbol{\xi}_{nt} = \mathbf{v}_{nt} =: (v_{1t}, \dots, v_{nt})', \quad t \in \mathbb{Z} \quad (22)$$

where  $\mathbf{v}_n$  is an  $n$ -dimensional white noise process and  $\mathcal{F}_n(L)$  is a one-sided stable VAR filter with sparse coefficients, the rows of which have a finite number of non-zero terms. Estimators of  $\boldsymbol{\eta}_n$  and  $\mathbf{v}_n$  are obtained by applying the methodology described in the previous section.

Turning to volatilities, define two panels of volatility proxies,

$$\boldsymbol{\sigma}_{nt} := \log(\boldsymbol{\eta}_{nt}^2) \quad \text{and} \quad \boldsymbol{\omega}_{nt} := \log(\mathbf{v}_{nt}^2), \quad t \in \mathbb{Z}, \quad (23)$$

with  $\log(\boldsymbol{\eta}_{nt}^2) := (\log \eta_{1t}^2, \dots, \log \eta_{nt}^2)'$  and  $\log(\mathbf{v}_{nt}^2) := (\log v_{1t}^2, \dots, \log v_{nt}^2)'$ . After due centring, we assume that those two panels of volatility proxies in turn satisfy Assumptions 1-3, and hence admit the GDFM decompositions

$$\hat{\boldsymbol{\sigma}}_{nt} := \boldsymbol{\sigma}_{nt} - \mathbf{E}[\boldsymbol{\sigma}_n] = \boldsymbol{\chi}_{\sigma,nt} + \boldsymbol{\xi}_{\sigma,nt}, \quad t \in \mathbb{Z}, \quad (24)$$

$$\hat{\boldsymbol{\omega}}_{nt} := \boldsymbol{\omega}_{nt} - \mathbf{E}[\boldsymbol{\omega}_n] = \boldsymbol{\chi}_{\omega,nt} + \boldsymbol{\xi}_{\omega,nt}, \quad t \in \mathbb{Z}, \quad (25)$$

where  $\boldsymbol{\chi}_{\sigma,n}$  and  $\boldsymbol{\chi}_{\omega,n}$  are driven by  $q_\sigma$  and  $q_\omega$  common shocks, respectively, and  $\boldsymbol{\xi}_{\sigma,n}$  and  $\boldsymbol{\xi}_{\omega,n}$  are idiosyncratic—call them the volatility common and idiosyncratic components, respectively.

Note that traditional factor models for volatilities, derived from factor models of returns, are assuming that  $\hat{\sigma}_n$  has no idiosyncratic component and, more importantly, that  $\hat{\omega}_n$  has no common component. Such assumption is quite unlikely to hold, as there is no reason for market shocks only affecting the volatility  $\hat{\sigma}_n$  of level-common returns. The empirical results in Barigozzi and Hallin (2016) indeed amply confirm that a non-negligible proportion of the variance of  $\hat{\omega}_n$  is explained by the same market shocks also driving  $\hat{\sigma}_n$ .

The two volatility common components  $\chi_{\sigma,n}$  and  $\chi_{\omega,n}$  jointly define a  $2n$ -dimensional panel made of two blocks. These blocks might be driven by  $q_{\sigma\omega}$  shocks, some of which common in both blocks, and some others common only in one of them (see Hallin and Liška, 2011, for a general theory of factor models with block structure). However, when analysing the real data in Section 5, we find that  $q_{\sigma\omega} = q_\sigma = q_\omega = 1$ , and therefore there is evidence of a unique common shock, denoted as  $\{\varepsilon_t\}$ , driving both blocks, and thus unambiguously qualifying as the market volatility shock. The autoregressive representation of these two common components then reads as (see also (13))

$$\begin{pmatrix} A_{\sigma,n}(L) & 0_n \\ 0_n & A_{\omega,n}(L) \end{pmatrix} \begin{pmatrix} \chi_{\sigma,n} \\ \chi_{\omega,n} \end{pmatrix} = \begin{pmatrix} H_{\sigma,n} \\ H_{\omega,n} \end{pmatrix} \varepsilon_t, \quad t \in \mathbb{Z}, \quad \varepsilon_t \sim w.n.(0, 1), \quad (26)$$

where  $A_{\sigma,n}(L)$  and  $A_{\omega,n}(L)$  are one-sided square-summable block-diagonal stable filters with blocks of size  $2 \times 2$ , and  $H_{\sigma,n}$  and  $H_{\omega,n}$  are  $n$ -dimensional column vectors. All parameters in (26) can be estimated as described in Section 3.1.

The singular LVDNs for the common components of volatilities thus can be built from the VMA filters (see (18))

$$B_{\sigma,n}(L) := (A_{\sigma,n}(L))^{-1} H_{\sigma,n} K_\sigma \quad \text{and} \quad B_{\omega,n}(L) := (A_{\omega,n}(L))^{-1} H_{\omega,n} K_\omega, \quad (27)$$

where  $K_\sigma$  and  $K_\omega$  in this case are just scalars needed to identify the scale and sign of the market shocks. For a given horizon  $h$ , the LVDN weights defined in (4) provide the percentages of  $h$ -step ahead forecast error variance of series  $i$  accounted for by the common market shock.

For the two volatility idiosyncratic components  $\xi_{\sigma,n}$  and  $\xi_{\omega,n}$ , we assume that Assumption 4 holds, yielding the sparse VAR representations

$$F_{\sigma,n}(L) \xi_{\sigma,n} = \nu_{\sigma,nt}, \quad t \in \mathbb{Z}, \quad \nu_{\sigma,nt} \sim w.n.(\mathbf{0}, C_{\sigma,n}^{-1}), \quad (28)$$

$$F_{\omega,n}(L) \xi_{\omega,n} = \nu_{\omega,nt}, \quad t \in \mathbb{Z}, \quad \nu_{\omega,nt} \sim w.n.(\mathbf{0}, C_{\omega,n}^{-1}), \quad (29)$$

where  $F_{\sigma,n}(L)$  and  $F_{\omega,n}(L)$  are one-sided stable filters with sparse coefficients, the rows of which have a finite number of non-zero terms. All parameters in (28) and (29) can be estimated as described in Section 3.2.

Inverting those autoregressive representations, we obtain the VMA filters (see also (14))

$$D_{\sigma,n}(L) := (F_{\sigma,n}(L))^{-1} R_{\sigma,n} \quad \text{and} \quad D_{\omega,n}(L) := (F_{\omega,n}(L))^{-1} R_{\omega,n}. \quad (30)$$

Here,  $R_{\sigma,n}$  and  $R_{\omega,n}$  are  $n \times n$  invertible matrices such that  $R_{\sigma,n}^{-1} \nu_{\sigma,n}$  and  $R_{\omega,n}^{-1} \nu_{\omega,n}$  are orthonormal. As explained in Section 2, we choose those matrices to be the Choleski factors of the covariances of the VAR shocks  $\nu_{\sigma,n}$  and  $\nu_{\omega,n}$ , ordered according to their centrality in the PCN induced by  $C_{\sigma,n}$  and  $C_{\omega,n}$  (see also (15)-(17)). From (30), and for a given horizon  $h$ , the LVDN weights computed from (4) then provide the shares of  $h$ -step ahead forecast error variance for the idiosyncratic volatility of series  $i$  accounted for by innovations in the idiosyncratic volatility of series  $j$ . As already explained in Section 3.3, most of those weights are quite small, and naturally can be thresholded, yielding a sparse network.

## 5 The network of S&P100 volatilities

In this section, we consider the panel of stocks used in the construction of the S&P100 index and, based on the daily adjusted closing prices  $\{p_{it} \mid i = 1, \dots, n, t = 1, \dots, T\}$ , we compute the panel of percentage daily log-returns

$$r_{it} := 100 \log(p_{it}/p_{it-1}), \quad i = 1, \dots, n, \quad t = 1, \dots, T,$$

which are our observed “levels” or “returns”. The observation period is  $T = 3457$  days, from 3rd January 2000 to 30th September 2013 and, since not all 100 constituents of the index were traded during the observation period, we end up with a panel of  $n = 90$  time series. A detailed list of the series considered is provided in the Web-based supporting materials. All results in this section are reported for the whole period 2000-2013, and for the period 2007-2008 corresponding to the Great Financial Crisis. Networks are represented by heat-maps showing the entries of the corresponding adjacency matrices. In all those plots we highlight the Energy and Financial sectors which correspond to the index values 22-33 and 34-46, respectively. Moreover, all LVDNs considered report shocks’ effects over  $h = 20$  periods, corresponding to a one-month horizon (20 days plus the contemporaneous effects). Results for  $h = 5, 10$  are reported in the Web-based supporting materials.

As already explained, the Hallin and Liška (2007) criterion yields  $q^T = 1$ , that is, the level-common component is driven by a one-dimensional common shock. By estimating (20) and (21), we recover the estimated level-common component  $\chi_n^T$ , the rank-one shocks  $\eta_n^T$ , and the level-idiosyncratic component  $\xi_n^T$ . The contribution of the level-common component to the total variation of returns, computed as the ratio between the sum of the (empirical) variances of the estimated level-common components  $\chi_n^T$  to the sum of the (empirical) variances of the observed returns, is 0.36. Turning to the level-idiosyncratic component  $\xi_n^T$ , we recover an  $n$ -dimensional innovation vector  $v_n^T$  by fitting a sparse VAR.

The panels  $\sigma_n^T$  and  $\omega_n^T$  of level-common and level-idiosyncratic volatilities are constructed, as explained in (23), from the  $\eta_n^T$ ’s and  $v_n^T$ ’s; the centred panels  $\hat{\sigma}_n^T$  and  $\hat{\omega}_n^T$  are obtained by subtracting the sample means. Applying the Hallin and Liška (2007) criterion again, we obtain  $q_\sigma^T = q_\omega^T = 1$  and, for the global panel,  $q_{\sigma\omega}^T = 1$ . This implies a unique market shock, common to the two sub-panels. We then compute the estimators  $\chi_{\sigma,n}^T$  and  $\xi_{\sigma,n}^T$  of the two level-common volatility components of the GDFM decomposition (24), and the estimators  $\chi_{\omega,n}^T$  and  $\xi_{\omega,n}^T$  of the two level-idiosyncratic volatility components of the GDFM decomposition (25). The overall contribution of the common components (driven by the unique market shock, hence non-diversifiable) to the total variances of  $\sigma_n^T$  and  $\omega_n^T$  are 0.60 and 0.17 respectively.

### 5.1 The LVDN of volatility common components

We now turn to the LVDNs of the two common components  $\chi_{\sigma,n}^T$  and  $\chi_{\omega,n}^T$ , given by (27). Since both panels are driven by the same unique common shock, the networks are identified once we impose a sign and a scale on the shocks. The sign is set in such a way that the sample correlation between the estimated market shock  $\{\varepsilon_t^T\}$  and the cross-sectional average of all common components is positive. The scale is set in such a way that the following results represent the effect of one-standard-deviation market volatility shock, that is, the sequence of MA loading coefficients is divided by the standard error of the market shock. We then obtain the estimated filters  $B_{\sigma,n}^T(L)$  and  $B_{\omega,n}^T(L)$ .

Sector	2000-2013		2007-2008	
	$w_{\chi_{\sigma},j}^T$	$w_{\chi_{\omega},j}^T$	$w_{\chi_{\sigma},j}^T$	$w_{\chi_{\omega},j}^T$
Consumer Discretionary	10.10	7.63	9.91	7.73
Consumer Staples	10.38	10.70	9.83	10.45
Energy	9.86	13.35	9.83	27.04
Financial	10.12	13.66	10.61	18.19
Health Care	9.92	8.84	9.77	6.24
Industrials	9.41	7.59	9.59	6.34
Information Technology	10.01	10.00	10.11	3.76
Materials	9.92	6.77	10.05	9.53
Telecommunication Services	10.33	9.80	10.29	5.79
Utilities	9.96	11.67	10.01	4.93
Total	100	100	100	100

**Table 1:** Percentages of 20-step ahead forecast error variances due to the market shock.

Since LVDNs in this case are singular, we do not show a graph, but rather report, in Table 1, for each panel  $\chi_{\sigma,n}^T$  and  $\chi_{\omega,n}^T$ , the percentages of sectoral 20-step ahead forecast error variance accounted for by the unique market shock. More precisely, for the ten sectors considered, the figures in Table 1 are the ratios

$$w_{\chi_{\sigma},j}^T := 100 \left( \frac{\sum_{i=1}^{n_j} \sum_{k=0}^{20} (b_{\sigma,k,i}^T)^2}{\sum_{i=1}^n \sum_{k=0}^{20} (b_{\sigma,k,i}^T)^2} \right) \text{ and } w_{\chi_{\omega},j}^T := 100 \left( \frac{\sum_{i=1}^{n_j} \sum_{k=0}^{20} (b_{\omega,k,i}^T)^2}{\sum_{i=1}^n \sum_{k=0}^{20} (b_{\omega,k,i}^T)^2} \right), \quad j = 1, \dots, 10,$$

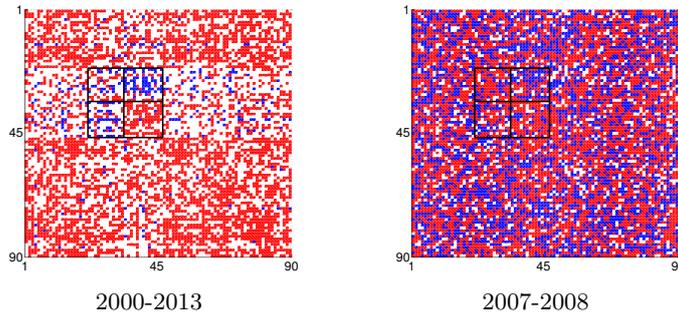
where  $b_{\sigma,k,i}^T$  and  $b_{\omega,k,i}^T$  are the  $i$ -th entries of  $B_{\sigma,nk}^T$  and  $B_{\omega,nk}^T$ , respectively, and  $n_j$  is the number of stocks in sector  $j$ . As the single shock obviously explains all the variance of the common component, we normalise the figures in each column so that their sum equals one hundred.

In both periods considered, the common component of level-common volatility is affected uniformly across all sectors by a market shock (the  $w_{\chi_{\sigma},j}^T$  columns in Table 1), while the common components of level-idiosyncratic volatilities exhibit some interesting inter-sectoral differences (the  $w_{\chi_{\omega},j}^T$  columns in Table 1). In particular, when looking at the  $w_{\chi_{\omega},j}^T$  column, the Energy and Financial sectors are the most affected ones. The impact of a market shock during the crisis is even heavier on those two sectors. We refer to Barigozzi and Hallin (2016) for further results on the volatility of common components.

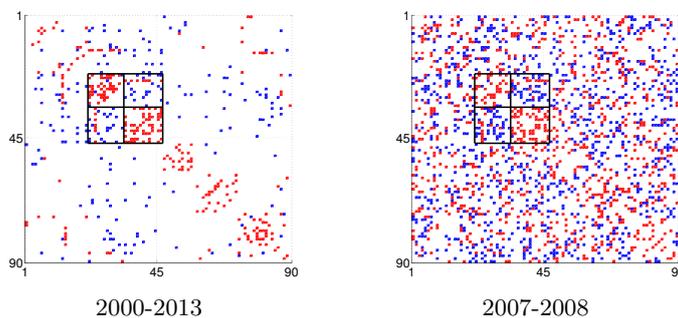
## 5.2 The LVDN of the volatility idiosyncratic components

Turning to the idiosyncratic volatilities  $\xi_{\omega,n}^T$  and  $\xi_{\sigma,n}^T$ , it appears that  $\xi_{\sigma,n}^T$  is essentially uncorrelated, both serially and cross-sectionally. Therefore, we focus only on the idiosyncratic volatility  $\xi_{\omega,n}^T$  of level-idiosyncratic components. Based on a BIC criterion, we estimate a sparse VAR(5) for  $\xi_{\omega,n}$ . The weight associated with the  $(i, j)$  edge of the LGCN of  $\xi_{\omega,nt}^T$  then is the  $(i, j)$  entry of  $F_{\omega,n}^T(1) = \sum_{k=0}^5 F_{\omega,nk}^T$ .

We define the *edge density* of a network with set of edges  $\mathcal{E}$  as the proportion  $\#(\mathcal{E})/(n^2 - n)$  of couples  $(i, j)$  in  $\mathcal{E}$ . When considering the elastic net estimation approach, we obtain, for the period 2000-2013, a LGCN with edge density 53% whereas, for the period 2007-2008, that density is close to 86%. The corresponding LGCNs are shown on Figure 2. Note that, as expected, the LGCNs are much sparser when resulting from group lasso estimation, with densities 14% for the period 2000-2013, and 37% for the period 2007-2008. Nevertheless, since the subsequent results on LVDNs turn out to be qualitatively similar regardless of the way



**Figure 2:** LGCN for  $\xi_{\omega,n}^T$ , negative weights in blue, the positive ones in red.



**Figure 3:** PCN for  $\nu_{\omega,n}^T$ , negative weights in blue, the positive ones in red.

we estimate the VAR, we only show here the elastic net results, deferring group and adaptive lasso results to the Web-based supporting materials.

As explained in the previous sections, identification of the LVDN requires some choices. The Choleski ordering is appealing, since it is data-driven, but requires an ordering of the cross-sectional items (the stocks). Here, we use an identification method based on the partial correlation of the VAR residuals  $\mathbf{v}_{\omega,n}^T$ . As explained in Section 2, we ordered the stocks according to the concept of eigenvector centrality for undirected networks (see Bonacich, 1987) in the estimated PCN. The resulting network is shown in Figure 3 for the two periods under study. The densities of those networks are 6% for the period 2000-2013 and 24% for the period 2007-2008. The ten most central stocks are reported in Table 2. For results using other identification methods we refer to the Web-based supporting material.

Summing up, from the estimation of the VAR model and the analysis of its residuals we see that (i) the Great Financial Crisis has considerably blown up the dynamic interdependencies between stocks, and (ii) the Energy and Financial stocks appear as the most interconnected ones, with more intra-sectoral dependencies rather than inter-sectoral.

The LVDN for  $\xi_{\omega,n}^T$  now can be computed on the basis of the ordering (partially) shown in Table 2. That identification defines (see (30)) the matrix  $R_{\omega,n}^T$  as the Choleski factor of the sample covariance matrix of the ordered shocks. The weight of edge  $(i, j)$  of the  $\xi_{\omega,n}^T$  LVDN is

$$w_{\xi_{\omega,n}^T, ij}^T = 100 \left( \frac{\sum_{k=0}^{20} (d_{\omega,k,ij}^T)^2}{\sum_{j=1}^n \sum_{k=0}^{20} (d_{\omega,k,ij}^T)^2} \right), \quad i, j = 1, \dots, n,$$

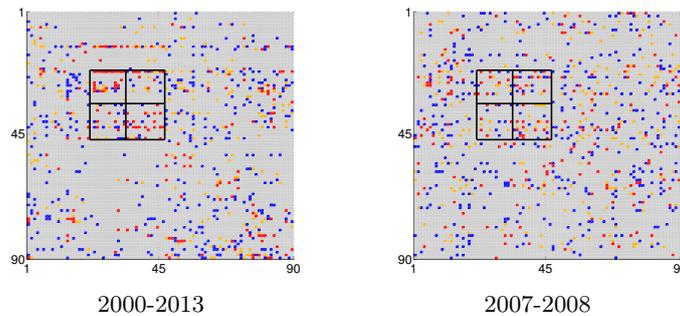
where  $d_{\omega,k,ij}^T$  is the  $(i, j)$  entry of  $D_{\omega,nk}^T$  such that  $D_{\omega,n}^T(L) = \sum_{k=0}^{20} D_{\omega,nk}^T L^k$  as defined in (30). The resulting network is directed and weighted, but it is not sparse, meaning that in principle

2000-2013	2007-2008
JPM JP Morgan Chase & Co.	BAC Bank of America Corp.
C Citigroup Inc.	USB US Bancorp
BAC Bank of America Corp.	JPM JP Morgan Chase & Co.
APA Apache Corp.	MS Morgan Stanley
WFC Wells Fargo	WFC Wells Fargo
COP Conoco Phillips	DVN Devon Energy
OXY Occidental Petroleum Corp.	GS Goldman Sachs
DVN Devon Energy	AXP American Express Inc.
SLB Schlumberger	COF Capital One Financial Corp.
MS Morgan Stanley	UNH United Health Group Inc.

**Table 2:** Eigenvector centrality in the PCN for  $\nu_{\omega,n}^T$ .

percentiles	50 <sup>th</sup>	90 <sup>th</sup>	95 <sup>th</sup>	97.5 <sup>th</sup>	99 <sup>th</sup>	max
2000-2013	0.02	0.13	0.20	0.29	0.48	4.29
2007-2008	0.17	0.71	1.00	1.28	1.76	4.53

**Table 3:** Selected percentiles of  $\xi_{\omega,n}^T$  LVDN weights.

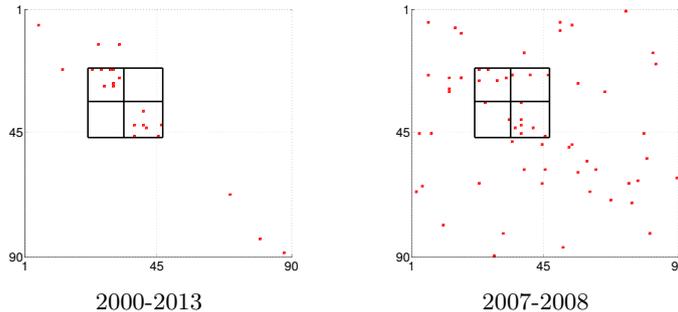


**Figure 4:** LVDN for  $\xi_{\omega,n}^T$ , weights below the 95<sup>th</sup> percentile in grey, between the 95<sup>th</sup> and 97.5<sup>th</sup> percentiles in blue, between the 97.5<sup>th</sup> and 99<sup>th</sup> percentiles in yellow, and above the 99<sup>th</sup> percentile in red.

all its weights can be different from zero. Still, only a small proportion of edges has weights larger than one, corresponding to a proportion of variance larger than 1%. In particular, out of the total number ( $n(n-1) = 8010$ ) of possible edges, only 0.4% have weights larger than 1% over the period 2000-2013, while this number increases considerably, up to 5%, during the period 2007-2008. Selected percentiles are given in Table 3.

Figure 4 shows the  $\xi_{\omega,n}^T$  LVDN weights. Inspection reveals that LVDNs, although not sparse, have many entries close to zero. A thresholded version, as described in Section 3, is reported in the top row of Figure 5. The resulting plots are highly sparse and indeed the optimal value of the thresholding parameter  $\tau$  is found to be 1.17 for the period 2000-2013 and 1.87 for the period 2007-2008. Plots for other values of  $\tau$  are provided in the Web-based supporting materials. Given the few remaining links, this is conveniently visualised, and the corresponding networks are shown in Figure 5. Note that Financial (yellow nodes) and Energy (blue nodes) stocks are the most interconnected ones. When considering the whole sample, there are almost no inter-sectoral links; however, during the Great Financial Crisis, the degree of interconnectedness of the Financial sector with other sectors quite dramatically increases.

Those findings can be quantified by computing from- and to-degrees, as defined in (6). As



**Figure 5:** Thresholded LVDN for  $\xi_{\omega,n}^T$ , non-zero weights in red.

Sector	non-thresholded				thresholded			
	2000-2013		2007-2008		2000-2013		2007-2008	
	from	to	from	to	from	to	from	to
Consumer Discretionary	4.32	2.37	26.31	26.41	0.24	0.24	1.96	3.29
Consumer Staples	3.98	4.65	27.47	22.65	0.00	0.44	2.80	1.36
Energy	5.52	7.92	21.91	33.72	1.90	1.54	4.01	6.10
Financial	4.74	6.22	24.42	35.56	0.77	0.77	5.23	6.69
Health Care	5.00	2.51	28.06	22.36	0.00	0.00	4.27	1.83
Industrials	4.43	3.21	27.26	25.81	0.21	0.21	2.43	3.53
Information Technology	5.03	4.89	29.90	19.98	0.00	0.00	3.35	1.58
Materials	3.24	4.62	26.86	27.01	0.00	0.00	2.99	0.81
Telecommunications Services	6.50	7.26	27.44	16.52	1.91	1.91	0.00	0.00
Utilities	5.15	8.74	29.49	21.54	0.00	0.00	4.94	2.38
Total degree	4.73		26.54		0.47		3.40	

**Table 4:** From- and To-degree sectoral averages in the  $\xi_{\omega,n}^T$  LVDN.

explained in Section 2, the from-degree measures the exposure of a given firm to shocks coming from all other firms, while the to-degree measures the effect of a shock to a given firm on all other ones. In Table 4, we show sectoral averages of from- and to-degrees when computed both for the non-sparse and thresholded LVDNs. We also report sectoral total connectedness and overall total connectedness as defined in (7). Finally, an overall measure of how systemic is an institution is given by its centrality in the network. Table 5 shows, for the two periods considered, the rankings of firms according to a measure of eigenvector centrality adapted to weighted directed networks (see Bonacich and Lloyd, 2001). For the period 2000-2013, only very few stocks are connected and qualify as central.

All results amply demonstrate that the Great Financial Crisis quite spectacularly tightened up the links between firms. In particular, when considering the thresholded LVDN version, the shocks to Financial stocks still account for almost 7% of the total idiosyncratic variance during the crisis, while shocks from other sectors explain about 5% of the Financial intra-sectoral variability (see the last two columns of Table 4). The Financial sector appears to be the most central in the network, the most vulnerable to shocks coming from all other sectors, and the most systemic in the sense that a shock to the Financial stocks is most likely to strongly affect the whole panel.

A few more comments are in order in relation with the results in the Web-based supporting materials. First, when comparing the results for  $h = 20$  with those obtained at shorter horizons ( $h = 5, 10$ ), we see that connectivity increases considerably with the forecast hori-

2000-2013	2007-2008
BAC Bank of America Corp.	USB US Bancorp
JPM JP Morgan Chase & Co.	BAC Bank of America Corp.
WFC Wells Fargo	COF Capital One Financial Corp.
C Citigroup Inc.	AIG American International Group Inc.
USB US Bancorp	C Citigroup Inc.
APA Apache Corp.	WFC Wells Fargo
SLB Schlumberger	BA Boeing Co.
COP Conoco Phillips	CVX Chevron
-	MCD McDonalds Corp.
-	IBM International Business Machines

**Table 5:** Eigenvector centrality in the  $\xi_{\omega,n}^T$  thresholded LVDN.

zon, indicating that the effect of a shock keeps on propagating for a long time. This result is consistent with the fact that volatilities in financial data, although stationary, tend to have long memory (see Barigozzi and Hallin, 2016, for a detailed analysis of this aspect). Second, group lasso and elastic net VAR estimation yield qualitatively similar results. However, despite of sparser LGCNs, group lasso yields estimated LVDNs that are more tightly connected than those resulting from elastic net. This is particularly true during the crisis period when Financial and Energy stocks are the most central ones, but also the Health Care sector seems to have a decisive role.

To conclude, we provide some empirical justification of the “factor plus sparse VAR” approach adopted in this paper by comparing the conditional dependencies in the volatility panel  $\omega_n^T$  with those of its idiosyncratic component  $\xi_{\omega;n}^T$ . To do this, we consider *partial spectral coherence (PSC)*, which is the analogous of partial correlation, but in the spectral domain, and is strictly related with the coefficients of a VAR representation (see Davis et al., 2015, for a definition).

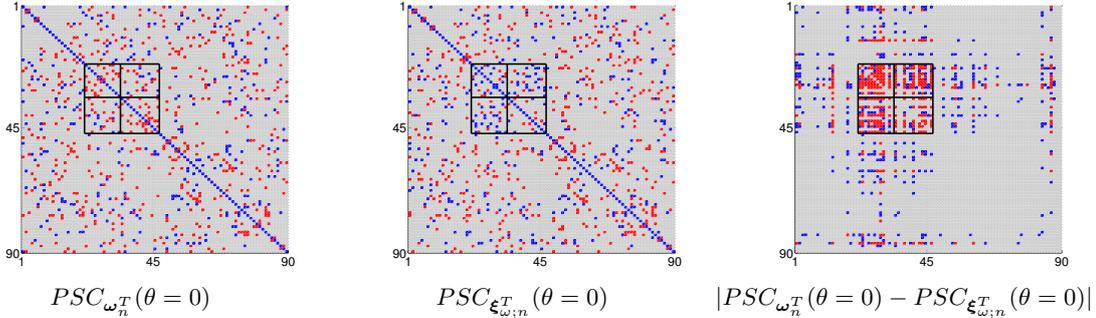
In line with the long-run spirit of the LVDN definition, and since volatilities have strong persistence, we first consider the *PSCs* at frequency  $\theta = 0$ , thus looking at long-run conditional dependencies. Selected percentiles of the distributions of the absolute value of the PSC entries for  $\omega_n^T$  and  $\xi_{\omega;n}^T$  and the distribution of the absolute value of their differences are shown in Table 6. Both *PSCs* have many small (in absolute value) entries, which is consistent with our sparsity assumptions.

The two *PSCs* are shown in the first two panels of Figure 6, while in the rightmost panel we show the absolute values of their differences. Inspection of these results clearly indicates that (i) after removal of the market shock, the idiosyncratic component still contains important dependencies, and (ii) an important benefit of our “factor plus sparse VAR” approach is to uncover the hidden dependencies between and within the Financial and Energy sectors. Moreover, when repeating the same analysis at other frequencies (e.g.,  $\theta = \pi/2, \pi$ ), no significant difference emerges between the two *PSCs*; the benefits of our approach thus are relevant mainly in the long run.

Similar conclusions can be derived by considering spectral densities (at different frequencies) and the squared partial spectral coherence (averaged over all frequencies), a measure which is non-zero if and only if two series are uncorrelated at all leads and lags after taking into account the (linear) effects of all other series in the panel. Results are in the Web-based supporting materials.

	percentiles	50 <sup>th</sup>	90 <sup>th</sup>	95 <sup>th</sup>	97.5 <sup>th</sup>	99 <sup>th</sup>	max
$ PSC_{\omega_n^T}(\theta = 0) $		0.06	0.16	0.18	0.22	0.25	0.35
$ PSC_{\xi_{\omega;n}^T}(\theta = 0) $		0.06	0.16	0.19	0.22	0.25	0.34
$ PSC_{\omega_n^T}(\theta = 0) - PSC_{\xi_{\omega;n}^T}(\theta = 0) $		0.02	0.06	0.08	0.12	0.15	0.24

**Table 6:** Distribution of absolute value of  $PSCs'$  entries.



**Figure 6:**  $PSCs$  at frequency  $\theta = 0$ . Left and middle panels: weights in absolute values below the 90<sup>th</sup> percentile in grey, weights above the 90<sup>th</sup> percentile in red, and below the 10<sup>th</sup> percentile in blue. Right panel: weights below the 90<sup>th</sup> percentile in grey, between the 90<sup>th</sup> and 95<sup>th</sup> percentiles in blue, and above the 95<sup>th</sup> percentile in red.

## 6 Conclusions

In this paper, we extend the study of interconnectedness of volatility panels initiated by Diebold and Yilmaz (2014) to the high-dimensional setting where a factor structure can be assumed for the data. We determine and quantify the different sources of variation driving a panel of volatilities of S&P100 stocks over the period 2000-2013. Our analysis highlights the key role of the Financial sector, which appears to be particularly important during the Great Financial Crisis. Other sectors such as Energy, and in some cases also Health Care, seem to have an important role too. Moreover, we show that, contrary to a direct sparse VAR approach, our “factor plus sparse VAR” method can unveil crucial inter-sectoral dependencies, which can be of tantamount importance for investors’ decisions in the context of risk management.

## References

- Abegaz, F. and E. Wit (2013). Sparse time series chain graphical models for reconstructing genetic networks. *Biostatistics* 14, 586–599.
- Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The network origins of aggregate fluctuations. *Econometrica* 80, 1977–2016.
- Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015). Systemic risk and stability in financial networks. *Economic Journal* 125, 443–463.
- Acharya, V., R. Engle, and M. Richardson (2012). Capital shortfall: a new approach to ranking and regulating systemic risks. *The American Economic Review* 102, 59–64.
- Adamic, L. A. and N. Glance (2005). The political blogosphere and the 2004 US election: divided they blog. In *Proceedings of the 3rd international Workshop on Link Discovery*, pp. 36–43. ACM.
- Adrian, T. and M. K. Brunnermeier (2016). CoVaR. *The American Economic Review*, forthcoming.

- Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrics* 71, 579–625.
- Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica* 70, 191–221.
- Barigozzi, M. and C. Brownlees (2016). NETS: Network estimation for time series. Available at SSRN: <http://ssrn.com/abstract=2249909>.
- Barigozzi, M., G. Fagiolo, and D. Garlaschelli (2010). Multinetwork of international trade: A commodity-specific analysis. *Physical Review E* 81, 046104.
- Barigozzi, M. and M. Hallin (2016). Generalized dynamic factor models and volatilities: Recovering the market volatility shocks. *The Econometrics Journal* 19, 33–60.
- Basu, S. and G. Michailidis (2015). Regularized estimation in sparse high-dimensional time series models. *The Annals of Statistics* 43, 1535–1567.
- Bauwens, L., S. Laurent, and J. V. K. Rombouts (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics* 21, 79–109.
- Bickel, P. and E. Levina (2008). Covariance regularization by thresholding. *Annals of Statistics* 36, 2577–2604.
- Billio, M., M. Getmansky, A. Lo, and L. Pellizzon (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics* 104, 535–559.
- Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 1170–1182.
- Bonacich, P. and P. Lloyd (2001). Eigenvector-like measures of centrality for asymmetric relations. *Social Networks* 23, 191–201.
- Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*. Holden-Day.
- Cai, T. and W. Liu (2011). Adaptive thresholding for sparse covariance matrix estimation. *Journal of the American Statistical Association* 106, 672–684.
- Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure, and mean–variance analysis on large asset markets. *Econometrica* 51, 1281–304.
- Dahlhaus, R. and M. Eichler (2003). Causality and graphical models for time series. In P. Green, N. Hjort, and S. Richardson (Eds.), *Highly Structured Stochastic Systems*, pp. 115–137. Oxford University Press.
- Davis, R. A., P. Zang, and T. Zheng (2015). Sparse vector autoregressive modeling. *Journal of Computational and Graphical Statistics, to appear*.
- De Mol, C., D. Giannone, and L. Reichlin (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics* 146, 318–328.
- Dempster, A. P. (1972). Covariance selection. *Biometrics* 28, 157–175.
- Diebold, F. and K. Yilmaz (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics* 182, 119–134.

- Eichler, M. (2007). Granger causality and path diagrams for multivariate time series. *Journal of Econometrics* 137, 334–353.
- Fagiolo, G. and G. Santoni (2015). Human-mobility networks, country income, and labor productivity. *Network Science* 3, 377–407.
- Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3–56.
- Fan, J., Y. Liao, and M. Mincheva (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society Series B* 75, 603–680.
- Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The Generalized Dynamic Factor Model: Identification and estimation. *The Review of Economics and Statistics* 82, 540–554.
- Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2015a). Dynamic factor models with infinite-dimensional factor spaces: One Sided Representations. *Journal of Econometrics* 185, 359–371.
- Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2015b). Dynamic factor models with infinite dimensional factor space: Asymptotic analysis. EIEF working paper 16/07, einaudi institute for economics and finance, rome.
- Forni, M. and M. Lippi (2001). The Generalized Dynamic Factor Model: Representation theory. *Econometric Theory* 17, 1113–1141.
- Friedman, J., T. Hastie, and R. Tibshirani (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9, 432–441.
- Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica* 79, 733–772.
- Gelper, S., I. Wilms, and C. Croux (2016). Identifying demand effects in a large network of product categories. *Journal of Retailing* 92, 25–39.
- Hallin, M. and M. Lippi (2014). Factor models in high-dimensional time series. A time-domain approach. *Stochastic Processes and their Applications* 123, 2678–2695.
- Hallin, M. and R. Liška (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association* 102, 603–617.
- Hallin, M. and R. Liška (2011). Dynamic factors in the presence of blocks. *Journal of Econometrics* 163, 29–41.
- Hautsch, N., J. Schaumburg, and M. Schienle (2014). Forecasting systemic impact in financial networks. *International Journal of Forecasting* 30, 781–794.
- Hautsch, N., J. Schaumburg, and M. Schienle (2015). Financial network systemic risk contributions. *Review of Finance* 19, 685–738.
- Hsu, N.-J., H.-L. Hung, and Y.-M. Chang (2008). Subset selection for vector autoregressive processes using lasso. *Computational Statistics & Data Analysis* 52, 3645–3657.
- Jovanovic, B. (1987). Micro shocks and aggregate risk. *The Quarterly Journal of Economics* 102, 395–409.
- Kock, A. B. and L. Callot (2015). Oracle inequalities for high dimensional vector autoregressions. *Journal of Econometrics* 186, 325–344.

- Kolaczyk, E. D. (2009). *Statistical Analysis of Network Data: Methods and Models*. Springer Science & Business Media.
- Koop, G., M. H. Pesaran, and S. M. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of econometrics* 74, 119–147.
- Lam, C. and Q. Yao (2012). Factor modeling for high-dimensional time series: inference for the number of factors. *The Annals of Statistics* 40, 694–726.
- Meinshausen, N. and P. Bühlmann (2006). High-dimensional graphs and variable selection with the Lasso. *The Annals of Statistics* 34, 1436–1462.
- Newman, M. E. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences* 98, 404–409.
- Nicholson, W. B., J. Bien, and D. S. Matteson (2014). Hierarchical vector autoregression. arXiv preprint 1412.5250.
- Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *The Journal of Business* 53, 61–65.
- Peng, J., P. Wang, N. Zhou, and J. Zhu (2009). Partial correlation estimation by joint sparse regression models. *Journal of the American Statistical Association* 104, 735–746.
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13, 341–360.
- Rothman, A. J., E. Levina, and J. Zhu (2010). Sparse multivariate regression with covariance estimation. *Journal of Computational and Graphical Statistics* 19, 947–962.
- Serrano, M. Á. and M. Boguñá (2003). Topology of the world trade web. *Physical Review E* 68, 015101.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica* 48, 1–48.
- Stock, J. H. and M. W. Watson (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97, 1167–1179.
- Swanson, N. R. and C. W. Granger (1997). Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association* 92, 357–367.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*, 267–288.
- Watts, D. J. and S. H. Strogatz (1998). Collective dynamics of “small-world” networks. *Nature* 393, 440–442.
- Yuan, M. and Y. Lin (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society Series B* 68, 49–67.
- Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research* 33, 452–473.
- Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association* 101, 1418–1429.
- Zou, H. and T. Hastie (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B* 67, 301–320.