Billio, Monica; Caporin, Massimiliano; Frattarolo, Lorenzo; Pelizzon, Loriana

Working Paper

Networks in risk spillovers: A multivariate GARCH perspective

SAFE Working Paper, No. 225

Provided in Cooperation with:
Leibniz Institute for Financial Research SAFE

Suggested Citation: Billio, Monica; Caporin, Massimiliano; Frattarolo, Lorenzo; Pelizzon, Loriana (2018) : Networks in risk spillovers: A multivariate GARCH perspective, SAFE Working Paper, No. 225, Goethe University Frankfurt, SAFE - Sustainable Architecture for Finance in Europe, Frankfurt a. M., https://doi.org/10.2139/ssrn.3239369

This Version is available at:
https://hdl.handle.net/10419/181755

HOUSE OF FINANCE
WORKING PAPER SERIES

Monica Billio – Massimiliano Caporin – Lorenzo Frattarolo – Loriana Pelizzon

Networks in Risk Spillovers: A Multivariate GARCH Perspective

SAFE Working Paper No. 225

SAFE | Sustainable Architecture for Finance in Europe
A cooperation of the Center for Financial Studies and Goethe University Frankfurt

House of Finance | Goethe University
Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main

Tel. +49 69 798 30080 | Fax +49 69 798 33910
info@safe-frankfurt.de | www.safe-frankfurt.de

Networks in risk spillovers: A multivariate GARCH perspective*

Monica Billio[†] **Massimiliano Caporin**[‡] **Lorenzo Frattarolo**[§] **Loriana Pelizzon**[¶]

Abstract

We propose a spatiotemporal approach for modeling risk spillovers using time-varying proximity matrices based on observable financial networks and introduce a new bilateral specification. We study covariance stationarity and identification of the model, and analyze consistency and asymptotic normality of the quasi-maximum-likelihood estimator. We show how to isolate risk channels and we discuss how to compute target exposure able to reduce system variance. An empirical analysis on Euro-area cross-country holdings shows that Italy and Ireland are key players in spreading risk, France and Portugal are the major risk receivers, and we uncover Spain's non-trivial role as risk middleman.

Keywords: spatial GARCH; network; risk spillover; financial spillover.

JEL Classification: C58, G10

*We thank the participants at the International Symposium on Forecasting 2015, Riverside, USA; and the Computational and Financial Econometrics Conference 2015 in Pisa, Italy; the SYRTO Conference held in Amsterdam in 2015; the 79th International Atlantic Economic Conference held in Milan in 2015; the 9th Financial Risk International Forum of the Bachelier Institute held in Paris in 2016; the 3rd Annual Conference of the International Association for Applied Econometrics held in Milan in 2016; the SYRTO Conferece held in Paris in 2016; the Econometrics and Statistics conference held in Hong Kong in 2017; and seminar presentations at the Complutense University of Madrid and Tansmanian School of Business and Economics for the comments received from them. The usual disclaimers apply. The authors acknowledge financial support from the The European Union, Seventh Framework Program FP7/2007-2013 under grant agreement SYRTO-SSH-2012-320270, the MIUR PRIN project MISURA - Multivariate Statistical Models for Risk Assessment, the SAFE Research Center, funded by the State of Hessen initiative for research, LOEWE, and the Department of Statistical Sciences of the University of Padova through the MIUR Project of Excellence "Statistical methods and models for complex data."

[†]University Ca' Foscari Venezia (Italy)

[‡]University of Padova (Italy) - Corresponding author, Department of Statistical Sciences, Via C. Battisti 241, 35121, Padova, massimiliano.caporin@unipd.it, +39-0498274199.

[§]University Ca' Foscari Venezia (Italy)

[¶]SAFE-Goethe University Frankfurt (Germany) and University Ca' Foscari Venezia (Italy)

1 Introduction

The US subprime and European sovereign bond crises sparked a renaissance in the research related to contagion and risk spillovers (Corsetti et al. (2011); Forbes (2012)). In line with the literature (Bekaert and Harvey (1997); Ng (2000); Billio and Pelizzon (2003)), we define a variance spillover as the contribution to the time t variance (risk) of a recipient entity due to the variance or shocks impacting on a source entity before time t, or due to the covariance between the two entities before time t. This definition is signed and directional; it includes the time dimension and can account for feedback effects. On the other hand, it excludes a systematic shift of variances due to a common factor affecting both entities.[1] We also focus on covariance spillovers, which we define as the contribution to the time-t covariance (thus dependence) between two entities due to the variance (risk) of the two entities, or due to the shocks affecting the variance evolution of the two entities, always before time t.[2] The covariance spillover becomes relevant in a framework where the correlations are not directly modeled, and they can be indirectly retrieved by means of variances and covariances. The definitions we adopt are strict and target only one symptom, of the many phenomena usually related to contagion and, more broadly, to systemic risk (see Silva et al. (2017) and Benoit et al. (2017) and reference therein). Being focused on covariance allows us to fill a void in the current literature that rarely addresses the quantitative investigation of economic transmission channels as potential pathways for risk spillovers.

We aim to introduce an economically grounded medium through which variance and covariance spillovers could flow. We consider the simplest model for this medium by focusing on pairwise directed linkages as summarized by a network and introduce and exploit a parallel between the network approach and the tools commonly used in spatial econometrics. This parallel allows us to take advantage of spatial econometrics literature

[1] Our approach is meant either for a case in which returns are not affected by a common factor or for modeling the residuals of a reduced-form system of equations.

[2] Where we assume that the covariance shocks are a function of the shocks affecting the variances of the two entities.

of the past decades (see Anselin (2013), LeSage and Pace (2009), Elhorst (2003)). The intersection and interaction of different research fields, such as network science, spatial econometrics, economics, and finance, give rise to new developments and tools. In particular, one of the most advanced strands of the spatial econometrics literature (see Keiler and Eder (2013); Blasques et al. (2016); Caporin and Paruolo (2015); Tonzer (2015); Billio et al. (2017)) focuses on statistical, economic, and financial relationships. Our work belongs to this strand of the literature and aims at building a bridge between contagion, volatility spillover, and networks.

The starting point of our paper is the contribution of Caporin and Paruolo (2015), which introduces spatiotemporal econometrics tools for the analysis of conditional volatility models, with the objective of estimating and measuring risk channels, but our work departs in a significant way. Even if we focus only on their BEKK class of models, our network relationships are intended to be derived from specific and potentially granular financial variables (in our application, cross-border exposures of national banking systems). The variables we suggest for consideration are those commonly perceived as potential time-varying transmission channels of shocks, and are thus characterized by asymmetry and time variation. The asymmetry in the weight matrix allows us to introduce network dependence in the model in two different ways and to focus on both risk spillover sources and risk spillover receivers. This distinction is possible only using jointly a vectorial specification for spatial parameters and an asymmetric weight matrix, and was not considered anywhere before. In addition, this new specification makes straightforward the inference on the inferred network links (by allowing simple significance tests for each link) and lets us understand if a particular risk channel between two entities is mediated by the variable included in the selected weight matrix.

More precisely, from a methodological point of view, the asymmetric weight matrix requires a generalization of the identification conditions of Caporin and Paruolo (2015) that we derive and, motivated by the time variation of the network, we are the first to discuss covariance stationarity using the joint spectral radius of Rota and Strang (1960)

and Jungers (2009). Both identification and covariance stationarity conditions are more complex in our framework. We obtain the identification condition by deriving a matrix that linearly links our parameters with those of a full BEKK specification and demonstrate that the rank of that matrix is linked to the algebraic connectivity characteristics of an auxiliary undirected network, resulting in identification conditions more refined than the ones derived in Caporin and Paruolo (2015).[3] The identification restriction does not imply a sign restriction on parameters associated with the network, and this leads to the possibility of variance reducing network effects.

Understanding the role of joint spectral characteristics also allows us to justify a more general weight matrix normalization with respect to the one considered in the literature (Lee and Yu (2012), Qu et al. (2017)). We complete the methodological discussion by adapting to the time-varying coefficients case the proof of consistency and asymptotic normality of the quasi-maximum-likelihood estimator (QMLE) of the BEKK model obtained in Hafner and Preminger (2009). Again, the joint spectral radius is a convenient tool to generalize the existing proof, in which an essential part is played by bounds on the usual spectral radius.

All methodological advances are driven and motivated by the development of tools able to investigate and measure the amount of spillover that could be explained by the transmission channels, as summarized by the selected network. This is particularly important in applications since, as already remarked, we are not only able to obtain the statistical significance of a given channel, as in Tonzer (2015), but also to distinguish the importance of different nodes by using diagonal matrix parameters. Moreover, the use of directed networks (i.e., asymmetric proximity) allows us to weight the transmission channel according to their direction, and we recover information on the statistically relevant channels of contagion, resulting in a better understanding of risk circulation.

We present an empirical analysis, considering a sample period covering the collapse of Lehman Brothers, Greece's bailout, and Brexit. We focus on the role of the network

[3]Nevertheless, we show how in the symmetric case we recover their results.

of cross-country banking system exposures in explaining the European sovereign CDS spillovers and provide a counterfactual analysis evaluating the risk-reducing exposures on the period of the Brexit referendum.

The paper is organized as follows. In Section 2, we introduce the econometric model for the spatial interpretation of risk and discuss identification, covariance stationarity, and QMLE estimation. In Section 3, we discuss inference-based networks, system variance decomposition, and the counterfactual analysis. In Section 4, we apply the methodology to CDS differences for the major countries in the Euro area during the subprime, sovereign debt crises, and Brexit. Finally, Section 5 summarizes our findings and concludes. The Appendices A,B, and C include definitions, corollaries, and proofs of the main results. The online supplementary material, S, reports several additional and complementary tables and figures.

2 A Model for Spatial Econometrics of Networked Risk

In this section, we start to describe our framework for a spatiotemporal econometric treatment of risk and dependence relationships across entities. Two parts compose the framework. First, a multivariate GARCH model, in which lagged shocks propagate through fluctuating relationships. Second, an exclusive set of diagnostic tools able to investigate risk circulation. Both parts have relevance in building the framework. In particular, the model chosen is one of the possible alternatives for studying risk spillovers in a network of entities. It is a workable example on which we can test the framework. About the second part, we designed a set of diagnostic tools, adaptable to other modeling proposals, but already able to showcase the benefits of mapping risk flows. Here, we study our Spatial Bilateral BEKK (SB-BEKK) specification in all its identification and estimation details and postpone to the next section the description of the risk tools.

Our approach focuses on the risk dimension as Caporin and Paruolo (2015), but ex-

tends the historical spatial econometrics definition of proximity. In fact, usually, neighboring relations are constant, based on geographical measures of distance, as in Anselin (2013) and Elhorst (2003), or fixed economic properties, such as the industry sectors membership in Caporin and Paruolo (2015). We choose, instead, the language of network theory to describe not only the static similarities among entities already encoded in the usual weight matrices, but the actual complex relationships among the nodes. This different interpretation of the weight matrix as a network can characterize those relationships in all their weighted, directed, and fluctuating generality. This new translation of old ideas will have unanticipated outcomes both in methodology (see our identification conditions in Theorem 1), and applications (see our inferred networks in subsection 3.2).

In this enterprise, we take inspiration from a recent strand of spatial econometrics literature that includes finance-based weight matrices in their analysis (Keiler and Eder (2013), Blasques et al. (2016), Tonzer (2015), and Billio et al. (2017)), and starts to introduce the network terminology in the discussion. However, building on Caporin and Paruolo (2015) differentiates us from the latter papers. First, we give an alternative perspective in our focus on the risk dimension, as opposed to their aim to explain expected returns. Second, they include only contemporaneous (spatial) transmission effect of the linkages, as opposed to our lagged cross-sectional (spatiotemporal) effect.

In addition, one of the novel features of our framework, with respect to the latter efforts, is a proper treatment of time dependence in the weight matrix. Previous applications of spatial methods to financial markets average these time-varying relationships, reducing to a static framework (see, as an example, Tonzer (2015)). This contribution is not completely new; however, we are the first to embed the time-varying nature of financial proximity within a conditional covariance model, going beyond the static weight matrix used in Caporin and Paruolo (2015). In fact, recent advances in the different modeling approach of spatial panels address the point and study its statistical consequences Lee and Yu (2012), Qu et al. (2017). Unfortunately, their assumption on weight-matrix normalization is restrictive in the financial framework. Our use of joint spectral characteristics

leads instead, to less constraining normalizations (see subsections 2.3 and 3.1).

The other major contribution of the paper is our bilateral specification of proximity. The specification has two main ingredients. The first one is the use of variable-specific proximity parameters, which allows heterogeneity in the network mediated effects. Heterogeneity was already present in the classical literature and Caporin and Paruolo (2015) reintroduced it in a volatility framework. The second ingredient is our consideration of the weight matrix as a directed network. Directedness is not present in classical approaches, but is already considered in Billio et al. (2017); Tonzer (2015) and Blasques et al. (2016). In fact, in classical spatial econometrics, if A is a neighbor of B with a given distance between them, the reverse is also true. In a financial framework, instead, symmetry is not usual. As an example, we can consider the relationship induced by the level of financial claims. It is improbable that the amount of claim that A has on B would be the same amount that B has on A. In a financial setting, then, the relationship is likely to be directed. Moreover, transforming the relationship to an undirected one by considering net claims, would result in an information loss. In fact, investors could perceive the riskiness of the connection to be dependent on its direction. Heterogeneity and directedness were already used, but no one understood the consequences of their combination. In fact, as detailed in the following subsection, the joint use of the two leads to different multiplication possibilities in proximity matrices. We will show how this little improvement in the design of proximity will have vast consequences in the econometric and empirical analysis.

2.1 Directed Networks and Bilateral Proximity

In our opinion, network theory is the correct language to describe general relationships of the type needed in our analysis. This subsection integrates the concepts of the adjacency matrix of a graph with the weight and proximity matrices discussed in the spatial econometrics literature. It also introduces our new bilateral specification for network-based proximity matrices. We can define a weighted network (graph) as an ordered pair of sets and a function $G = (V, E, w)$ where $V = \{1, \ldots, n\}$ is the set of vertices (or nodes),

$E \subset V \times V$ the set of edges (or arcs), and $w(e) : E \mapsto \mathbb{R}^+$ is the weight function attributing strength to the edges. An edge between two nodes exists if there is a relationship between them and it can be identified as the (ordered) pair $\{i, j\}$ with $i, j \in V$. If there is no direction in the connection between nodes, then an edge $\{u, v\}$ is an unordered pair of nodes and the graph G is said to be undirected, whereas if a direction exists, then each edge $\{i, j\}$ is defined as an ordered pair of nodes and the graph G is said to be directed graph (or digraph). Different edges could have different strength as summarized by the weight function. The vertex adjacency structure of a n-order graph $G = (V, E, w)$ can be represented through a n-dimensional matrix W called an adjacency matrix. Each element $\omega_{i,j}$ of the adjacency matrix is equal to $w(\{i, j\})$ if there is an edge $\{i, j\} \in E$ (i.e., an edge from institution u to institution v with $i, j \in V$), and $\omega_{i,j} = 0$ otherwise. If the graph is undirected, then $\omega_{i,j} = \omega_{i,j}$, that is, the adjacency matrix is symmetric.

In the literature (see Anselin (2013), LeSage and Pace (2009), Elhorst (2003)), a proximity matrix is a matrix whose entries quantify the distance between entities. The prototypical example is the real geographical distance. These are generally summarized into a weight matrix W, whose entries $\omega_{i,j}$ correspond to the physical distance from i to j. The matrix W is, then, static and symmetric. In the language of networks, W corresponds to the adjacency matrix of a weighted undirected network. We relax this implicit symmetry hypothesis by considering W as the adjacency matrix of a weighted directed network. Usually (Elhorst (2003)), W is row normalized to maintain reasonable magnitudes for the parameters. In addition, often, the spatial impact is measured by a single coefficient ρ, which pre-multiplies the weight matrix W. This implies a unique impact across the entities involved in the analyses.[4]

Following Caporin and Paruolo (2015), we adopt a wider viewpoint, introducing the proximity matrix P. P is a linear combination of an adjacency matrix W of a weighted

[4]The two most common specifications are the spatial autoregressive model (SAR), where a vector (a cross-section) of observations Y obeys the linear model $Y = \rho WY + \varepsilon$, and the spatial error model (SEM), where, for the same observation, we have $Y = \varepsilon + \rho W\varepsilon$. In both cases, the coefficient ρ monitors the spatial impact, that is, the response of Y to the neighbors' values (in the SAR model) or to the neighbors' shocks (in the SEM model).

directed network and an identity matrix I_n,

$$P = \rho_0 I_n + \rho_1 W,$$

where ρ_1 and ρ_0 are scalars. In particular, ρ_1 represents the global impact of a network on the nodes, and ρ_0 is a constant common to all the nodes. The previous authors, among others, include heterogeneus impacts by transforming the scalar coefficients into diagonal parameter matrices. In their specification, the proximity is an affine function of the network:

$$P = \text{diag}\left(\rho_0\right) I_n + \text{diag}\left(\rho_1\right) W.$$

Moreover, we realized that commutativity is not anymore guaranteed for the product of a diagonal parameter matrix and a non-symmetric W, as it was in the case of symmetric W. Consquently, pre- and post-multiplication of W by the coefficient matrix represent different effects. Accordingly, we introduce our bilateral representation of the proximity matrix, which extends and includes all the previous formulations:

$$P\left(W_t\right) \quad = \quad \text{diag}\left(\rho_0\right) I_n + \text{diag}\left(\rho_{1,L}\right) W_t + W_t \text{diag}\left(\rho_{1,R}\right), \tag{1}$$

where n is the number of series, I_n is the $n \times n$ identity matrix, and we introduce two parameter vectors associated with left multiplication, $\rho_{1,L}$, and right multiplication $\rho_{1,R}$. This new form of proximity could be of independent interest even were it abstracted from our risk framework and applied to models for conditional mean. This specification, as we show in the following, allows for, in our model characterization of sources and receivers of risk, the derivation of a linear restriction test for the significance of network edges, and also leads to model-identification conditions that do not constrain parameter signs and can describe variance-reducing spillovers, and several other features.

2.2 Spatial Bilateral BEKK

A popular specification adopted for the estimation of conditional variance matrices is the BEKK model of (Engle and Kroner, 1995). Unfortunately, even in its most parsimonious specification (the BEKK(1,1)), this model is computationally infeasible, even for moderate values of n, due to its large number of parameters $(2n^2 + 0.5n\,(n+1))$. For this reason, the standard practice is to restrict A and B to be either scalar or diagonal. Despite being feasible, these restricted specifications impose strong limitations on the interpretability of the model outcomes because they exclude or sensibly limit the presence of risk spillovers, included in A, and variance feedbacks, coming from B.

To overcome these critical aspects, Caporin and Paruolo (2015) introduce the Spatial-BEKK GARCH model, in which the full parameter matrices A and B are replaced by proximity matrices. The spatial version of the BEKK model has the advantage of being more parsimonious than the full BEKK case, as, at the same time, it is more flexible than the diagonal specification, and it includes spillovers and feedback effects. We further extend their model considering a time-varying weighted directed network $G_t = (V, E_t, w_t)$, with fixed nodes and adjacency matrices W_t, and allowing a bilateral specification for proximity.

Given a vector y_t of n cross-sectional observations at time t, we define $u_t = y_t - \bar{y}$, where $\bar{y}$ is the vector of sample means. Our Spatial Bilateral BEKK GARCH (SB-BEKK) has the following structure:

$$
\begin{aligned}
u_t &= \Sigma_t^{1/2}\epsilon_t \quad \epsilon_t \sim \mathcal{N}\left(0, I_n\right), \quad t = 1, \ldots, T \\
\Sigma_t &= CC' + A\left(W_t\right) u_{t-1}u_{t-1}' A\left(W_t\right)' + B\left(W_t\right)\Sigma_{t-1}B\left(W_t\right)'
\end{aligned}
$$

where C is a lower triangular matrix, $\Sigma_t^{1/2}$ is the Cholesky decomposition of Σ_t,[5] and the parameter matrices have a proximity specification described in equation (1) that includes

[5] Alternatively to the Cholesky, we can compute the square root by resorting to the spectral decomposition and set $\Sigma_t^{1/2} = D_t P_t^{1/2} D_t'$, where D_t is the matrix of eigenvectors and P_t is the diagonal matrix of eigenvalues.

both left multiplication and right multiplication:

$$\begin{aligned} A(W_t) &= A_0 + A_{1,L}W_t + W_tA_{1,R} = \text{diag}(a_0)\, I_n + \text{diag}(a_{1,L})\, W_t + W_t\text{diag}(a_{1,R}) \\ B(W_t) &= B_0 + B_{1,L}W_t + W_tB_{1,R} = \text{diag}(b_0)\, I_n + \text{diag}(b_{1,L})\, W_t + W_t\text{diag}(b_{1,R}), \end{aligned}$$

where a_0, $a_{1,M}$, b_0 $b_{1,M}$, with $M = L, R$, are $n \times 1$ vectors. Within the SB-BEKK framework, left and right multiplication allow researchers to focus on different aspects of risk propagation. To better understand this model attitude, it is advisable to recall the notions of direct and indirect effects of shock diffusions, previously introduced in the spatial econometrics literature (see LeSage and Pace (2014)) and generalized here for the SB-BEKK model. The starting point is the Spatial Error Model (SEM), where the $n-$variate dependent variable v_t depends on an $n-$dimensional vector of shocks u_t, on a weight matrix W, and on a scalar parameter θ

$$v_t = (I_n + \theta W)\, u_t.$$

LeSage and Pace (2014) decompose the error term in the direct effect v_t^0 and the local indirect effect v_t^1 as follows:

$$v_t = v_t^0 + v_t^1 \quad , \quad v_{i,t}^0 = [I_n u_t]_i = u_{i,t} \quad , \quad v_{i,t}^1 = [\theta W u_t]_i = [W\theta u_t]_i = \theta \sum_{j=1}^{n} \omega_{i,j} u_{j,t}$$

where $[X]_{i,j}$ identifies the element of position i, j of the argument matrix X with one single index if X is a vector, $\omega_{i,j}$ represents the "distance"between subject i and subject j coming from the spatial weight matrix W (time invariant, for simplicity), and by definition $\omega_{i,i} = 0$.

This means that the target variable $v_{i,t}$ depends on its own shock, as monitored by v_t^0, the direct impact. Further, it is also affected by the indirect impact v_t^1. The latter captures the effect coming from neighboring elements $v_{j,t}$ with $i \neq j$ and with an impact only from those j such that $\omega_{i,j} \neq 0$. We note that in the SEM model, left and right

multiplication are identical due to the presence of a scalar parameter θ. We translate these elements into the SB-BEKK model and provide a novel decomposition. We focus on the ARCH part of the model because we want to highlight the role of innovations. We note that

$$
\begin{aligned}
v_t = A(W)\, u_t &= (A_0 + A_{1,L}W + WA_{1,R})\, u_t = v_t^0 + v_{L,t}^1 + v_{R,t}^1 && (2)\\
v_{i,t}^0 &= [A_0 u_t]_i = a_{0,i} u_{i,t} && (3)\\
v_{L,i,t}^1 &= [A_{1,L} W u_t]_i = a_{1,L,i} \sum_{j=1}^{n} \omega_{i,j} u_{j,t}. && (4)\\
v_{R,i,t}^1 &= [W A_{1,R} u_t]_i = \sum_{j=1}^{n} \omega_{i,j} a_{1,R,j} u_{j,t}. && (5)
\end{aligned}
$$

The $i-$th element of v_t depends on its own past shock, weighted by the coefficient $a_{0,j}$ (direct effect), on the past shocks of its neighbors weighted by the distance, loaded with the sum of the same coefficient, $a_{1,L,j}$ (indirect left effect), and a coefficient different from each source $a_{1,R,i}$ (indirect right effect). For the GARCH part, similarly to the ARCH case, we introduce:

$$
\begin{aligned}
m_{i,t}^0 &= [B_0 u_t]_i = b_{0,i} u_{i,t} && (6)\\
m_{L,i,t}^1 &= [B_{1,L} W u_t]_i = b_{1,L,i} \sum_{j=1}^{n} \omega_{i,j} u_{j,t} && (7)\\
m_{R,i,t}^1 &= [W B_{1,R} u_t]_i = \sum_{j=1}^{n} \omega_{i,j} b_{1,R,j} u_{j,t}. && (8)
\end{aligned}
$$

Consequently, bearing in mind that we are discussing properties of a conditional covariance model, the left multiplication term allows us to investigate which are the risk receivers. In the right multiplication term, distinct from the left multiplication case, the coefficients in the indirect effect are not pointing at the subject we are monitoring (subject j) but at the subject originating the shock (subject i). With the right multiplication term, the parameters magnify the effect of the sources of risk, allowing us to focus on risk spreaders.

2.3 Model estimation

We suggest to estimate the parameters of our SB-BEKK, conditional on the availability of the full sequence W_t for $t = 1, 2, \ldots, T$, by means of QMLE methods. If we denote by $\theta \equiv (\text{vec}(C), a_{0,M}, a_{1,M}, b_{0,M}, b_{1,M})$ the vector of model parameters, the log-likelihood $L_T(\theta)$ is :

$$\ell_t(\theta) = \frac{n}{2}\log(2\pi) + \frac{1}{2}\det(\Sigma_t) + \frac{1}{2}u_t(\Sigma_t)^{-1}u_t' \quad , \quad L_T(\theta) = -\frac{1}{T}\sum_{t=1}^{T}\ell_t(\theta) \ .$$

In the following proposition and lemma, we extend results included in Caporin and Paruolo (2015), with a detailed study of identification for a directed network sequence. Moreover, as the network matrices W_t might evolve on a time scale lower than that adopted for entities, we focus here on the collection of the K distinct matrices W_k, $k = 1, 2, \ldots K$ with $K \leq T$. Our focus does not mean that we need for the estimation of the model a lower frequency for the weight matrices.

The study of identification is possible by making explicit the expression of $A(W_t)$ and $B(W_t)$ as a linear function of a vector of parameters. Considering $A(W_t)$ as an example, we could write:

$$\text{vec}(A(W_k)) = M(W_t)\begin{bmatrix} a_0 \\ a_{1,L} \\ a_{1,R} \end{bmatrix} \quad , \quad M(W_t) = \begin{bmatrix} I_n \otimes I_n & W_t' \otimes I_n & I_n \otimes W_t \end{bmatrix}(I_3 \otimes H)$$

where $H = \sum_{i=1}^{n} e_i \otimes e_i e_i'$ and e_i is the i-th column of I_n.

This explicit expression is novel with respect of Caporin and Paruolo (2015), even in the symmetric case. Moreover, an anonymous referee to whom we are particularly grateful pointed out how this linear map, in the asymmetric case, cannot be injective due to the rank deficiency of $M(W_t)$. A proper identification of the model is, then, impossible, without imposing further restrictions on the parameters. The following important refinement of the identification conditions elucidates the hidden role of the connectivity of the net-

work sequence unexpected, given the simpler conditions for the symmetric case in Caporin and Paruolo (2015).[6]

Before stating the identification conditions in the next proposition and lemma, let us define an auxiliary network $\mathfrak{G}_k = (V, \mathfrak{E}_k, \mathfrak{w}_k)$, defined on the same nodes, but with undirected edges and different weights, according to the adjacency matrix $\mathfrak{W}_k$:

$$\mathfrak{W}_k = \left(W_k' \odot W_k'\right)\left(D_{W_k \odot W_k}\right)^{-1}\left(W_k \odot W_k\right) \quad , \quad D_{W_k \odot W_k} = \operatorname{diag}\left(\left(W_k \odot W_k\right)\mathbf{1}\right).$$

We define as c_k the number of connected components in $\mathfrak{G}_k$.

Proposition 1 (Identification)**.** *Assume that at least one of the matrices W_t is not symmetric. Let π_K correspond to the vectorized collection of either $A(W_k)$ or $B(W_k)$ for $k = 1, 2, \ldots K$, that is $\pi_K = \operatorname{vec}\left(\left[A(W_1)' : A(W_2)' : \ldots : A(W_K)'\right]'\right)$, where $A(W_k)$ or $B(W_k)$ matrices are those coming from a full BEKK model fitted on each subsample, and are thus globally identified. Let ψ be the parameter vector in the SB-BEKK representation corresponding to π_K. Additionally, let $c = \min_{1,\ldots,K} c_k$ be the minimum number of connected components of the induced networks $\mathfrak{G}_k$ in the sample. Then, a necessary and sufficient condition for the identification of ψ is that we place c linear restriction on it.*

Lemma 1 (Single Restriction)**.** *A sufficient condition for having $c = 1$ is that at least one of the original weight matrices W_k is fully indecomposable, inducing an ultrastrong graph in the sense of Brualdi (1967).*

The condition of Lemma 1 is not restrictive for small dense networks such as the one we use in the empirical analysis. In those cases, the condition is trivially violated when one node is isolated (i.e., if all the matrices W_k have the same column (row) full of zeros). In that case, the corresponding right (left) parameter must be set to zero. For bigger and sparser networks, sequences one should check the number of components of the induced networks sequence $\mathfrak{G}_k$.

[6] We reproduce conditions in Caporin and Paruolo (2015) for the symmetric case; see corollary 1 in Appendix subsection B.1

In the full BEKK(1,1), it is sufficient to set $A_{1,1}, B_{1,1} > 0$ to have global identification, and in our case, this is equivalent to the conditions $a_{0,1} > 0$ and $b_{0,1} > 0$. Then, under Lemma 1, we choose to achieve identification by imposing an equal sum of left and right parameters:

$$a_{1,R,n} = -\sum_{i=1}^{n} a_{1,L,i} + \sum_{i=1}^{n-1} a_{1,R,i} \quad , \quad b_{1,R,n} = -\sum_{i=1}^{n} b_{1,L,i} + \sum_{i=1}^{n-1} b_{1,R,i}$$

We stress that this identification strategy allows for the presence of coefficients of both signs in $a_0, b_0, a_{1,L}, a_{1,R}, b_{1,L}$ and $b_{1,R}$. In subsection 3.3, we show how this allows contributions that reduce the variance.

Subsection B.1 in the Appendix reports the proof of the previous proposition and lemma, alongside corollary 1 for symmetric W_t matrices.

A necessary and sufficient condition for stationarity and geometric ergodicity, and thus a sufficient condition to ensure the ergodicity and strict stationarity of the process implied by the model, is covariance stationarity, Boussama et al. (2011). According to the discussion in Avarucci et al. (2013), this condition, distinct from the univariate case, appears necessary for consistency and asymptotic normality of the QMLE estimator. To show the covariance stationarity, we introduce a VARMA representation of the process. In addition, the condition for covariance stationarity will be expressed using the joint spectral radius of Rota and Strang (1960) and Jungers (2009). It is well known that a BEKK model admits a VARMA representation (see, for example, Hafner and Rombouts (2007)). In our case, the VARMA coefficients will depend on the network sequence, and ensuring covariance stationarity requires the study of convergence of a geometric series with heterogeneus terms, analogous to the one studied in the generalized autoregressive model; see Brandt (1986) and Bougerol and Picard (1992).

The VARMA representation of the SB-BEKK(1,1) has the following structure:

$$X_t = \tilde{C} + \left(\tilde{A}(W_t) + \tilde{B}(W_t)\right) X_{t-1} - \tilde{B}(W_t)\,\eta_{t-1} + \eta_t \tag{9}$$

where

$$X_t = \text{vech}(u_t u_t'), \quad \xi_t = \text{vech}(\Sigma_t), \quad \eta_t = X_t - \xi_t, \quad \tilde{C} = \text{vech}(CC')$$
$$\tilde{A}(W_t) = L_n (A(W_t) \otimes A(W_t)) D_n, \quad \tilde{B}(W_t) = L_n (B(W_t) \otimes B(W_t)) D_n,$$

and L_n is the elimination matrix, while D_n is the duplication matrix (see Magnus and Neudecker (1999)). Given the distribution of ϵ_t in (2), adopting the terminology of Hafner and Rombouts (2007), we have that u_t is a strong GARCH process and η_t is a martingale difference sequence. This remains true if the innovation distribution is misspecified, but remains i.i.d. Equation (9) puts the model outside of the general specifications of dynamic spatial panel models given in Elhorst (2001), because of the presence of a spatiotemporal moving average term. The specification closer to ours is the STARMA model introduced in Pfeifer and Deutsch (1980); for a recent overview, see Subba Rao and Costa Antunes (2004). However, such a model adopts a static network and scalar coefficients. Due to the presence of a time-varying network, our model is also close to the generalized autoregressive model of Brandt (1986) and Bougerol and Picard (1992). In order to derive stationarity conditions for our case, we need a condition similar to those based on the Lyapunov exponent in Brandt (1986) and Bougerol and Picard (1992). Accordingly, to introduce it, we first need the following definitions concerning the joint spectral characteristic of a set of matrices.

Let us consider an infinite countable set of $n \times n$ matrices $\mathcal{A} = \{A_i\}_{i=0}^{\infty}$ with the convention that $A_0 = I_n$. A generic product of t elements from $\mathcal{A}$ could be obtained by extracting uniformly, with replacement, t elements from $\mathcal{A}$ and matrix-multiplying them. For example, suppose that the elements sampled have indexes $\sigma_1 = 44, \sigma_2 = 44, \sigma_3 = 20, \ldots, \sigma_t = 1$; the product will be $A_{44}^2 A_{20} \cdots A_1$. Let us define the set of all those possible products $\mathcal{A}^t = \left\{ M \in \mathcal{A}^t | M = \prod_{i=1}^{t} A_{\sigma_i}, s.t. A_{\sigma_1} \in \mathcal{A}, \ldots A_{\sigma_t} \in \mathcal{A} \right\}$.
We have:

Definition 2.1. *Joint Spectral Radius. Given a proper norm $\|\cdot\|$ on $\mathbb{R}^n \times \mathbb{R}^n$, we define*

on the set $\mathcal{A}$ *the joint spectral radius* $\varrho(\mathcal{A})$ *by*

$$\varrho(\mathcal{A}) = \lim_{t\to\infty} \hat{\rho}_t(\mathcal{A}) \tag{10}$$

$$\hat{\rho}_t(\mathcal{A}) = \sup_{A_{\sigma_1},\ldots,A_{\sigma_t}\in\mathcal{A}} \left(\left\| \prod_{i=1}^{t} A_{\sigma_i} \right\| \right)^{1/t} \tag{11}$$

Definition 2.2. *Lyapunov Exponent. Given a proper norm* $\|\cdot\|$ *on* $\mathbb{R}^n \times \mathbb{R}^n$*, assuming that we have a probability measure on* $\mathcal{A}_t$*, we define on the set* $\mathcal{A}$ *the Lyapunov Exponent* $\gamma(\mathcal{A})$ *by*

$$\gamma(\mathcal{A}) = \lim_{t\to\infty} \frac{1}{t}\mathbb{E}\left[\log\left(\left\| \prod_{i=1}^{t} A_{\sigma_i} \right\| \right) \right] \tag{12}$$

Using the Jensen inequality, it is possible to show that to $e^{\gamma(\mathcal{A})} \leq \varrho(\mathcal{A})$; see, for instance, Tsitsiklis and Blondel (1997).

We are now ready to discuss the covariance stationarity of our model.

Theorem 1 (Covariance Stationarity)**.** *If the joint spectral radius*

$$\varrho\left(\left\{ \tilde{A}(W_t) + \tilde{B}(W_t) \right\}_{t=-\infty}^{\infty} \right) < 1,$$

the SB-BEKK(1,1) process is covariance stationary conditionally on the network sequence.

We stress that the use of bounds on the joint spectral radius is a condition on dynamic stability strictly weaker than the ones already present in the literature. For example, the uniform boundedness assumption size in Lee and Yu (2012), being based on row and column sum norm, implies our joint-spectral-radius condition. This is relevant for our empirical application where the choice of normalization is different from the standard row sum, and motivated by the economics of the data used (see section 3.1).

We now discuss the consistency and asymptotic normality of the QMLE estimator along the lines of Hafner and Preminger (2009). In detail, we modify their proof to generalize them to the case of time-varying $\tilde{A}(W_t)$ and $\tilde{B}(W_t)$. In the following, $\|\cdot\|$

represents the norm operator, with different norms being specified when needed. Denote by θ the vector of stacked parameters that implicitly satisfy the identification condition of Theorem 1, and denote the true parameter vector as θ_0. Define the QMLE as $\theta_T = \arg\max_{\theta\in\theta} L_T(\theta)$. Additionally, let $\tilde{\Sigma}_t$ be the process where the starting values are drawn from their stationary distribution, and let $\tilde{\xi}_t$, $\tilde{L}_T$ and $\tilde{\ell}_t$ be defined analogously.

We begin by discussing the assumptions needed for consistency.

Assumption 1. *The parameter space* Θ *is compact and* $\varrho\left(\left\{\tilde{A}(W_t) + \tilde{B}(W_t)\right\}_{t=-\infty}^{\infty}\right) < 1$.

Assumption 2. $\{u_t\}$ *is strictly stationary and ergodic, and* $\exists s > 0$ *s. t.* $\mathbb{E}\left[\|u_t\|^s\right] < \infty$.

Assumption 3. $\mathbb{E}\left[\|\epsilon_t\|^s\right] < \infty$, $\mathrm{Var}\left[\epsilon_t\right] = I_n$.

Assumption 4. *The model is identified: i.e., conditions in Theorem 1 are satisfied.*

Our assumptions parallel those in Hafner and Preminger (2009). The only conceptual difference, once we take into account our use of time-varying matrices, is in Assumption 1, where we bound the joint spectral radius of the sum of $\tilde{A}(W_t)$ and $\tilde{B}(W_t)$, instead of a condition that points only at the joint spectral radius of $\tilde{B}(W_t)$ matrices. In Lemma 4 in subsection B.2 of the Appendix, we show that one condition implies the other. This was done for the standard BEKK in Boussama et al. (2011), where they also show that bounding the spectral radius of the sum was a sufficient condition for strictly stationarity and ergodicity. Then, in the standard framework, the first part of assumption 2 would be redundant. Showing that this also applies to our case is outside the scope of the paper.

Theorem 2 (Consistency)**.** *Under Assumptions 1-4* $\hat{\theta}_T \to_{a.s.} \theta_0$.

Proof of Theorem 2 is in section B of the appendix.

To establish asymptotic normality, the following additional assumptions are needed.

Assumption 5. *The parameter* θ_0 *is an interior point of* Θ

Assumption 6. $\mathbb{E}\left[\|u_t\|^6\right] < \infty$

Assumption 7. $\sup \max_{i=1,\ldots,n} \sum_{j=1}^{n} W_{t,ij} \leq d^* < \infty$ *a.s.*

Assumptions 5 and 6 are identical to assumptions adopted in Hafner and Preminger (2009). Assumption 7, pointing at the network structure, is not particularly restrictive. In fact, it is trivially verified for the row-normalization case, but it also justified for different normalization schemes. For example, the network used in our empirical analysis has $d^* = 1$ (c.f. section 4.1.1).

Further, let us also define the following matrices

$$V = \mathbb{E}\left[\frac{\partial \tilde{\ell}\left(\theta_0\right)}{\partial \theta}\frac{\partial \tilde{\ell}\left(\theta_0\right)}{\partial \theta'}\right], \qquad J = \mathbb{E}\left[\frac{\partial^2 \tilde{\ell}\left(\theta_0\right)}{\partial \theta \partial \theta'}\right].$$

Theorem 3 (Asymptotic Normality)**.** *Under Assumptions 1-7*

$$\sqrt{T}\left(\hat{\theta}_T - \theta_0\right) \to_D N\left(0, J^{-1}VJ^{-1}\right).$$

Proof of Theorem 3 is in section B of the Appendix.

In the empirical part, we compute the covariance analytically without using the second derivatives, following the suggestions in Hafner and Herwartz (2008). Parameters constrained to be positive are handled with an exponential transformation and the usual modification of the gradient. Finally, under the asymptotic normality it is straightforward to derive the distribution of the coefficients $a_{1,R,n}$ and $b_{1,R,n}$ linearly constrained by the identification restrictions.

3 Tools for Spatial Econometrics of Networked Risk

In this section, we conclude the description our framework for a spatiotemporal econometric treatment of risk and dependence relationships across entities, introducing a risk toolbox able to investigate the flow of risk spillovers. In the following subsections, we discuss normalization issues for the network sequence, the insight provided by inference-based networks, the decomposition of the system variance, and we introduce a counterfactual covariance-reduction analysis. The practical impacts of these elements will be highlighted

in the empirical analysis presented in Section 4.

3.1 Normalization and Robustness

Taking into account the time variation for the spatial proximity matrices W_t obliges us to pay particular attention to the way in which we normalize these matrices. In fact, a simple row normalization at each time would make the comparison of the proximity matrices over time very difficult. Furthermore, a time-specific or matrix-specific normalization would lead to a loss of information, as both disregard the evolution over time of the network structure.

In order to obtain parameters of a reasonable magnitude, but also to retain differences in matrix norms across time (which could be an important driver of dependence), we divide each row of W_t by an (economic) measure of the magnitude of the entities, which we denote as $M_{i,t}$. Note that this quantity is time varying to account for changing conditions, states, or entities. As examples of possible measures, we mention the use of gross domestic products or public debt for networks of countries, and the use of revenues or balance-sheet-based indicators for networks of companies.

If we adopt this approach, we stress that the spatial matrices W_t will not have the maximum eigenvalue equal to 1, as is standard in spatial econometrics. This new possibility in choosing normalization schemes is granted by the use of the joint spectral radius in our covariance stationarity condition in Theorem 1. In turn, this allows heterogeneity in the spectral radii of single terms W_t in the network sequence and, consequently, to use normalizations as devices to include additional economic and financial effects.

3.2 Inference-Based Networks

Within our SB-BEKK model, time-varying parameter matrices are composed of two elements: the series W_t, which we assume to be observed without errors, and the parameters estimates, which, following Theorem 3, are characterized by an asymptotic normal distribution. By combining these two components, we can revise our knowledge of the un-

derlying network by building several "inference-based"networks, whose characterization derives from the dependence on specific inferential aspects of the BEKK model parameters. We focus here on the A matrices and note that the same line of reasoning can be applied to the B matrices.

Using a simple linear restriction, we can test whether an off-diagonal element of $A(W_t)$ is statistically significant:

$$H_{0,i,j} : [A(W_t)]_{i,j} = (a_{1,L,i} + a_{1,R,j})\,\omega_{t,i,j} = 0 \qquad i \neq j$$

Then, under the asymptotic normality of the QMLE estimator, the test statistic for $H_{0,i,j}$ is asymptotically distributed as a standard normal. The simplicity in this kind of inference is one of the main features of our left-right specification not accessible to any previous spatial models in the literature.

We thus define $\mathcal{W}_t^A$ as a filtered ARCH network, or inference-based ARCH network (i.e., the network whose signed weight matrix are the off-diagonal statistically significant elements of $A(W_t)$). We also remark here that the sign of $[A(W_t)]_{i,j}$ depends on the relative sign and magnitude of the parameters involved, having the same sign of $(a_{1,L,i} + a_{1,R,j})$.

Using a similar approach, we derive filtered networks from the GARCH parameters. The network that can be filtered from the ARCH matrices represents the response to a shock in the previous period, while that associated with the GARCH matrices represents the covariance persistence, that is, the response to the whole history of past shocks.

3.3 Covariance Decomposition

The introduction of proximity matrices in the dynamic of BEKK models allows the estimation $0.5(n+1)n$ series of filtered conditional covariance elements. For $n \geq 3$, it is difficult to interpret directly all the recovered series, and it is therefore desirable to have summary measures backed by some theoretical line of reasoning. This is a classical issue

in spatial econometrics, where we observe the same difficulty in interpreting the impact of explanatory variables or innovations. The complexity stems from the large cross-sectional dimension of the analyzed data (or series), as in our case. The traditional solution is to resort to summary measures of the direct and indirect effects of explanatory variables and shocks; see LeSage and Pace (2009) and LeSage and Pace (2014). We follow a similar approach and introduce a decomposition of the sequence of conditional covariances provided by the SB-BEKK model. Nevertheless, there are two important distinctions: first, focusing on conditional covariance matrices, we deal with quadratic forms where spatial spillovers appear twice with an increase in the terms appearing in the decomposition; second, being that the SB-BEKK is a spatiotemporal model, we have a breakdown conditional to the past. We propose a four-term decomposition of the system-conditional covariance:

1. *Constant Contribution*: This represents the part of the covariance that is unrelated to the model dynamic and independent from both the network and the time;

2. *Direct Contribution*: This represents the covariance contribution from each entity's own past; it is the variance due to past direct effects, and therefore has no dependence on the network;

3. *Indirect contribution*: This represents the covariance contribution due to indirect effects that are due to the assets' network exposures;

4. *Mixed contribution*: This represents the covariance contribution originating from the quadratic form of the model, and is due to the interaction of both direct and indirect elements.

To introduce the algebra of our decomposition, we take as a working example a case where we have non-null values in the time-invariant matrices W. From equation (2), the conditional covariance at time t is given by the sum of three elements: the constant; a quadratic term associated with the shocks; and a quadratic term associated with the past

conditional covariance. In our decomposition, the *constant term* is simply given by the constant of the conditional covariance, thus CC'. We now focus on the shock-response term, the ARCH part of the model. We remind the reader that we introduce in equations (3) and (4) a definition of direct $v^0_{i,t}$ and indirect effects $v^1_{L,i,t}, v^1_{R,i,t}$ within the ARCH part. We now decompose the entire shock response term as follows:

$$\begin{aligned} A(W)\, u_{t-1} u'_{t-1} A(W)' &= v^0_{t-1} \left(v^0_{t-1}\right)' \\ &+ v^1_{L,t-1} \left(v^0_{t-1}\right)' + v^0_{t-1} \left(v^1_{L,t-1}\right)' + v^1_{R,t-1} \left(v^0_{t-1}\right)' + v^0_{t-1} \left(v^1_{R,t-1}\right)' \\ &+ v^1_{L,t-1} \left(v^1_{L,t-1}\right)' + v^1_{R,t-1} \left(v^1_{R,t-1}\right)' + v^1_{R,t-1} \left(v^1_{L,t-1}\right)' + v^1_{R,t-1} \left(v^1_{L,t-1}\right)'. \end{aligned}$$

The first element in the ARCH term decomposition refers to the variance (or covariance) own shock. This is comparable to a direct shock contribution. The third term represents the contribution to the covariance due to network weights. This corresponds to an indirect effect, that is, the shocks' impact due to the network. The second term can be interpreted as mixed effects because they combine both direct and indirect elements.

Moving to the GARCH part of the model, the direct, indirect, and mixed contributions correspond to covariances between the terms in equations equation (6), equation (7) and equation (8). In fact, for the indirect contribution originating from the GARCH part of the model, using the bilinearity of the conditional covariance operator and conditionally on the network W, we have:

$$\left[\Omega^{1,1}_{L,L,t-1}\right]_{i,j} = \mathbb{C}\mathrm{ov}\left(m^1_{L,i,t-1}, m^1_{L,j,t-1} \middle| I_{t-2}, W\right) = \left[B_{1,L} W \Sigma_{t-1} W' B'_{1,L}\right]_{i,j}. \quad (13)$$

We can recover similar identities for the direct and mixed contributions. In Table (1), we summarize the elements appearing in the conditional covariance decomposition.

Further, we highlight that the decomposition is time varying by construction and might also be affected by the dynamic in the network structure. Since the model specifications allow for positive and negative signs on both ARCH and GARCH coefficients, in principle, diversification benefits could arise from all four contributions. The variance

decompositions outlined above are specific to a single element of the covariance matrix. Of particular interest in the empirical section will be the decomposition of the variance of single series, a breakdown able to highlight which entity risk is the most affected by the presence of the exposure network (i.e., which are the more fragile nodes in the network). However, we might be interested in recovering a synthetic measure of the decomposition at the entire covariance level. We propose to define this synthetic (and time-varying) measure starting from a portfolio representation of the system, with portfolio weights given by the vector $\mathbf{z}$, such that $\sum_{i=i}^{n} z_i = 1$, and therefore leading to

$$\begin{aligned} \sigma_t^2 &= \mathbb{V}\mathrm{ar}\left(\mathbf{z}' y_t | \mathcal{I}_{t-1}\right) \\ &= \left(\sigma_t^{Constant}\right)^2 + \left(\sigma_t^{Direct}\right)^2 + \left(\sigma_t^{Indirect}\right)^2 + \left(\sigma_t^{Mixed}\right)^2. \end{aligned} \tag{14}$$

The previous decomposition is able to report, globally, the importance of the network sequence through time. Another possibility is instead to compute the marginal spillover contributions $MSC_{k,t}$ that we could attribute to entity k:

$$\begin{aligned} MSC_{k,t} &= \frac{1}{2}\frac{\partial}{\partial z_k}\mathbb{V}\mathrm{ar}\left(\mathbf{z}' y_t | \mathcal{I}_{t-1}\right) \\ &= MSC_{k,t}^{Constant} + MSC_{k,t}^{Direct} + MSC_{k,t}^{Indirect} + MSC_{k,t}^{Mixed} \end{aligned} \tag{15}$$

The marginal spillover contribution is normalized, guaranteeing that, if we define the vector $\mathbf{MSC}_t' = [MSC_{1,t}, \ldots, MSC_{n,t}]'$, we have $\mathbf{z}'\mathbf{MSC}_t = \mathbb{V}\mathrm{ar}\left(\mathbf{z}' y_t | \mathcal{I}_{t-1}\right)$. $MSC_{k,t}$ is, then, a measure of the importance of node k as a source of spillovers in time t. In particular, our decomposition allows us to disentangle the direct sources of spillovers from the ones that are mediated by the network sequence. Among the many possible choices, in the empirical application, we chose the simplest one, and thus consider a portfolio characterized by equal weights for each entity. This is also equivalent to considering the behavior of an average element of the covariance (i.e., the average linear dependence in the system). The use of different weighting schemes, with potentially better economic explanations, is left for further empirical research. The explicit expressions for variance

decompositions and marginal spillover contributions are reported in section A of the Appendix.

3.4 Simulated Counterfactual Network and Covariance Reduction

Finally, we rely on a simulated counterfactual analysis (Mccallum (1988), Rotemberg and Woodford (1997)) to investigate how changes in the network impact the variance of an equally weighted index. In particular, we change the network in order to minimize the multistep variance forecast based on ex-ante data. This choice of the objective function is motivated by its equivalence with the minimization of the average forecasted covariance coefficent (i.e., on the average linear dependence and risk in the system). The counterfactual innovation paths are bootstrapped from estimated innovations (see section C of the Appendix for the details of the circular bootstrap Politis and Romano (1992) we used). In a similar way, we recover a realized innovation path on the out-of-sample period using the previously estimated parameters and the out-of-sample observed network. By construction, using the realized innovation path in the model with the realized weights will reproduce the multivariate time series observed in the out-of-sample period. If, instead, we feed the model with the previously obtained optimal weights and the realized innovation path, we obtain a counterfactual multivariate time series that takes into account the effect of the optimal change in the network. We then compare, as a proxy for system variance (i.e., the average covariance in the system), the squared observations of the equally weighted indexes obtained from the realized and optimal counterfactual time series. If we observe a reduction of the optimal counterfactual proxy with respect to the realized one, we can conclude that, indeed, we can reduce the variance by optimally changing the network based only on ex-ante data. In the following, we describe the details of the procedure.

Conditional on the bootstrapped innovations $\tilde{\epsilon}_{T+l}^{[b]}$ with $b \in [1, \dots, N_B]$ and $l \in [1, \dots, h]$, and assuming that the network is constant over the forecast horizon, the

forecasted covariance path is a function of the network at time T, $W_T \mapsto \hat{\Sigma}^F_{T+l}(W_T)$ $l \in [1, \ldots, h]$. This raises the interesting possibility of obtaining a target network that can reduce the future risk and dependence in the system. To define the optimal target network, we require that it minimizes, at least locally, the average covariance coefficient of the system, which we compute as the variance of an equally weighted portfolio of all the series. We underline here that the methodology can easily accommodate different objective functions, being based on numerical optimization. A detailed investigation of more refined and economically motivated targets is left for future research. With our approach, we exploit the frequency mismatch between the data used to estimate the network and the series for which the risk is evaluated. Such situations are not rare, as financial networks might be built from lower-frequency data (using, for instance, balance-sheet data), while financial market data are available at a daily or even higher frequency. In particular, we assume that the network changes every q observations. That is, in the full sample T, we have $[T/q] = Q$ networks, or, alternatively, we have Q sub-periods in which the network is stable. In the forecast exercise, we assume that $W_{T+l} = W_Q$ for each $l \in [1, \ldots, h]$, such that $T+1$ and $T+h$ are the beginning and end of the period $Q+1$. We require that the average forecasted variance of the equally weighted index over period $Q+1$; i.e., the first sub-period following the estimation sample, conditional on the bootstrapped innovations, is minimized by numerically solving the following constrained optimization problem:

$$\min_{\mathrm{vec}W^\star} \left\{ \frac{1}{h} \sum_{l=1}^{h} \frac{1}{n^2} \mathbf{1}' \hat{\Sigma}^F_{T+l}(W^\star) \mathbf{1} \right\}$$

$$s.t. \quad 0 \leq [W^\star]_{i,j} \leq 1 \quad for \quad i,j = 1 \ldots n, \quad \text{and} \quad \mathrm{Tr}(W^\star) = 0$$

where $\mathbf{1}$ is the $n \times 1$ column vector whose elements are all equal to 1 and $\mathrm{Tr}(.)$ is the trace operator. It is important to note that the estimated network $W^\star$ is weighted and directed, but is totally unrelated to the last available network. We thus also consider a more realistic constraint in which the out (in) strengths of the nodes, defined as the row (column) sums of the optimal network, are set to be the same as the out (in) strengths

of the nodes of the last network W_Q. For the row-sum case, we impose

$$\sum_{j=1}^{n} [W^\star]_{i,j} = \sum_{j=1}^{n} [W_Q]_{i,j}\,, \tag{16}$$

and we can write a similar constraint for the column sum. These constraints avoid a change in the strengths of the nodes and correspond to a simple redistribution of the weights across the system. We refer to the constraint in equation (16) as a redistribution constraint. Moreover, we introduce an additional alternative constraint imposing that the resulting optimal network differs from the previous one in an ordinary way, and does not represent an exceptional change. We implement this by computing the Frobenius norm on the historical changes in the network

$$\|\Delta W_q\|_F = \|W_q - W_{q-1}\|_F = \sqrt{\mathrm{vec}\,(W_q - W_{q-1})'\,\mathrm{vec}\,(W_q - W_{q-1})} \qquad q = 1, \dots Q-1, \tag{17}$$

and imposing that the norm of optimal change $\|\Delta W^\star\|_F = \|W^\star - W_Q\|_F$ is less than or equal to the empirical 0.95 quantile of historical norm changes $q_{0.95}^{\|\Delta W\|_F}$

$$\|\Delta W^\star\|_F \leq q_{0.95}^{\|\Delta W\|_F}. \tag{18}$$

We refer to the constraint in equation (18) as a Frobenius norm constraint.

To evaluate the performance of the proposed out-of-sample methodology, we suggest comparing two estimates of the model, one excluding the out-of-sample data, and the second including the forecasted data. This enables us to compute the filtered innovations for the forecasted periods, conditional on the true, observed $Q+1$ network:

$$\hat{\epsilon}_{T+l} = \hat{\Sigma}_{T+l}^{-\frac{1}{2}}\,(W_{Q+1})\,u_{T+l} \qquad l \in [1, \dots, h]\,. \tag{19}$$

Then we can reconstruct the us and the counterfactual proxy for the equally weighted index's conditional variance, as if the realized network for the period of interest is the

optimal one, $W^{\star}$:

$$\tilde{u}^{\star}_{T+l} = \hat{\Sigma}^{\frac{1}{2}}_{T+l}\left(W^{\star}\right)\hat{\epsilon}_{T+l} \tag{20}$$

$$\mathbb{V}\mathrm{ar}\left(\frac{1}{n}\mathbf{1}'y^{\star}_{T+l}\middle|\mathcal{I}_{T+l-1}\right) = \mathbb{V}\mathrm{ar}\left(\frac{1}{n}\mathbf{1}'u^{\star}_{T+l}\middle|\mathcal{I}_{T+l-1}\right) \simeq \left(\frac{1}{n}\mathbf{1}'u^{\star}_{T+l}\right)^{2} \tag{21}$$

In this way, we can compare the obtained optimal counterfactual proxy with the realized proxy, the latter being robust against model misspecification. Even if this simulated counterfactual analysis is similar to the one used in monetary-policy evaluation exercises (see Rotemberg and Woodford (1997), Primiceri (2005) and more recently, Bikbov and Chernov (2013)), we are aware that, concerning our empirical application, regulators have, at the moment, no possibility of intervention in incentivizing the redistribution of banking exposures. An additional issue in the use of this procedure in policy evaluation, as noted in the literature, is that such a methodology is prone to the critique in Lucas (1976), as we are not considering the market reaction to the network change. Regarding those issues, we recall that endogenizing the network sequence is beyond the scope of the present investigation and that in Tonzer (2015), in a framework similar to ours, exogeneity failed to be rejected by several tests. In our case, in addition, the time-varying nature of the network during the estimation period mitigates this issue. In fact, the market reaction to changes in the network should have been at least partially encoded in estimated parameters. Following this line of reasoning, we propose a procedure to test the constancy of parameter estimates when we optimally change the network. In particular, we estimate again the model using the in-sample Q periods and a $Q+1$ period in which we use $W^{\star}$ and the $\tilde{u}^{\star}$ of equation (20). We call the estimates for the network parameters obtained in this way $a^{\star}_{1,L}, a^{\star}_{1,R}, b^{\star}_{1,L}, b^{\star}_{1,R}$. We also estimate the model using the in-sample Q periods and the $Q+1$ realized out-of-sample period. Then, considering as given the estimates implied by the optimal network, we test for the joint hypothesis

$$H_0 : a_{1,L} = a^{\star}_{1,L}, a_{1,R} = a^{\star}_{1,R}, b_{1,L} = b^{\star}_{1,L}, b_{1,R} = b^{\star}_{1,R}, \tag{22}$$

using the estimates and covariance obtained by the realized sample in a Wald test. Failing to reject the null hypothesis would support parameter constancy and limit relevance of the Lucas critique, making possible the policy exercise Engle and Hendry (1993). In particular, we explicity designed the Frobenius norm constraint equation (18) to increase the possibilities of obtaining parameter constancy. With those limitations in mind, we highlight that the product of our counterfactual exercise is a first step toward the determination of the target exposures that can be helpful for policy makers from a monitoring perspective.

4 Risk Spillovers among European Sovereign CDS

To better clarify the advantages and potential benefits of our methodology, we consider an application to publicly available data, and in particular, we consider the Euro area sovereign CDS premia. We consider only the Euro area because there is a documented currency firewall effect both for sovereign, Groba et al. (2013), and banks CDS, Alemany et al. (2015). Inside the EMU, we pick the two major economies and the peripheral countries excluding Greece, whose bailout make the CDS series unmanageable from the second semester of 2010. We use two different data sources: (i) the changes in the five-years sovereign CDS spreads for a selection of European countries and (ii) the matrices of foreign claims collected by the BIS. As we detail in the following subsection, these data refer to the claims that the banking sector of a country A has with respect to the banking (public and private) sector of another country B. There is clearly an asymmetry between the dependent variable and the data source for the weighting matrices. Nevertheless, several theoretical models show tight linkages and feedback loops between sovereign and banking risk both in single and multi-country economies; see Bolton and Jeanne (2011); Acharya et al. (2014); Gennaioli et al. (2014); Farhi and Tirole (2016). In Dungey et al. (2017) this theoretical link is also empirically investigated. The aim of our analysis is to characterize, identify and evaluate the sovereign risk of the system, considering the

total sovereign risk of the Euro area as the volatility of a weighted-average portfolio of European sovereign bonds. Risk spillovers are driven by the weight matrices, based on cross-country cross-credit exposures.

4.1 Data Description

4.1.1 BIS Banking Statistics

We use data at a quarterly frequency to describe the network of foreign claims among Portugal, Italy, Ireland, Spain, France, and Germany from Q4 2008 to Q4 2016, as they are produced by BIS in the consolidated banking statistics (ultimate risk basis). The quarterly claims are converted to a daily basis by repeating them for each day in the quarter, thus obtaining the sequence of daily matrices. This interpolating choice has an implicit assumption: foreign claims variation is much slower than the changes in CDS spreads.[7] BIS-consolidated banking statistics provide internationally comparable measures of national banking systems' exposures to country risk (see McGuire and Wooldridge (2005)). Country risk refers to countrywide events, which can lead to systemic instability that prevents obligors (whether direct debtors or guarantors of claims on other borrowers) from fulfilling their obligations. Banks contributing to the consolidated statistics report a full-country breakdown of claims booked by their offices worldwide. Only assets are reported. The residence of the ultimate obligor, or the country of ultimate risk, is defined as the country in which the guarantor of a financial claim resides, or the place in which the head office of a legally dependent branch is located. Foreign claims, in the ultimate risk basis, reported by country A with country B as a counterpart, are all on-balance-sheet financial assets, with the exclusion of derivative contracts guaranteed by public or private entities of country B and owned by the banking system of country A. Our choice differs from that in Tonzer (2015) who, mainly motivated by sample-length consideration and exchange-rates-adjustment reasons, uses locational banking statistics, also from BIS. We motivate this by the mixed nature of the data and the importance of the local banking

[7] We also used a linear interpolation scheme, and the estimation results were unaltered.

system in international financial intermediation (McCauley et al. (2010)). Further, the motivations in Tonzer (2015) are not relevant in our framework because we focus on a smaller and more homogeneous group of countries, with a common currency, for which the more reliable consolidated-ultimate-risk-basis statistics are available. We expect that, if A reports a claim with B as a counterpart (i.e., if A's banking system owns a certain amount of B's public and private debt, investors will perceive A's sovereign risk to be dependent on B's sovereign risk in terms proportional to the claim amount. In the next subsection, 3.1, we discuss the impact of normalization schemes on the estimations. In particular, we stress how our choice of normalization removes a clear deleveraging trend and gives a higher importance to peripheral countries. Here, instead, we underline how the use of an economic normalization scheme is possible only with our careful examination of identification and covariance stationarity conditions.

Our choice for the normalization M_{jt} of the j-th reporting country is its quarterly time series of total ultimate-risk-basis claims, which includes claims from the selected countries, but also from the rest of the world.[8] In addition, we stress that using this normalization allows us also to control for claims outside the chosen countries and to evaluate the exposure channel importance at the country level. The comparison between un-normalized series (Figure (S.1) in the supplementary material) and normalized series (Figure S.2 in the supplementary material) show relevant differences. With respect to the un-normalized series, the deleveraging trend is less evident and the peripheral countries are more important. We note, in particular, the relevance of Spain as a counterpart of Portugal and a big bump in Q4 2010 - Q3 2011 in Ireland's claims with Germany as a counterpart (a fact that was hidden when considering the dollar amount).

[8]We also investigated other choices for normalizations: the absence of normalization, row normalization, the GDP of the reporting country, and the public debt of the reporting country. In the full sample estimation, total claims outperform, in likelihood terms, the alternative normalization schemes in the vast majority of models, and when this is not the case, the difference in likelihoods is negligible. Estimation results for these alternatives are available upon request.

4.1.2 Sovereign CDS spreads

We use the daily changes in the five-year sovereign CDS spreads, from 9/10/2008 to 30/12/2016, for France, Germany, Ireland, Italy, Portugal, and Spain, as downloaded from Thomson Reuters Eikon. The choice of the five-year maturity is due to the more volatile nature of shorter maturities. As can be seen from Table S.1 in the supplementary material, mean and median are negligible, justifying our volatility approach. The Kurtosis of some series — in particular, that of Portugal, but also those of Germany and Ireland — is striking, adding a further motivation for a GARCH-type modeling.

The correlation is high between specific pairs — namely, Spain and Portugal, Spain and Italy, and Italy and Portugal — highlighting the closeness between those economies. Despite all being positive, several correlations display relatively small values, even if they are all significantly different from zero with a p-value below 0.01. Most interestingly, the smallest correlations are those between Germany and the other European countries (France excluded).

4.2 Parameter Estimation

In Table 2, we report QMLE results for the relevant parameters of the model. Table S.2 in the supplementary material includes some specification tests. We estimate the model by using a numerical-constrained optimization with bounds for the parameters function of the joint spectral radius of the network sequence. We checked covariance stationarity after the estimation by verifying the joint-spectral-radius condition of Theorem 2. In turn, we do that by computing a numerical bound obtained with the conic ellipsoid algorithm of the JSR matlab toolbox (see Vankeerberghen et al. (2014)).

As the first panel of Table S.2 shows, the SB-BEKK models outperform the diagonal model. We check the joint significance of parameters by likelihood ratio test statistics and a Wald-type statistic. Notably, both tests strongly reject the null, thus supporting the relevance of networks in variance-spillover analysis. Other tests on the significance of different parts of the model are also rejected. Table 2 shows the significance of the vast

majority of coefficients of the proposed model. In addition, the presence of significant coefficients of both signs induces variance-reducing contributions mediated by the claim matrix, but at this stage, due to the nonlinearity of the model, is not easy to understand which ones they are, and we are mainly commenting on their magnitude. At the 10% significance level, all the spatial coefficients for France, Ireland, Spain, and Portugal are statistically different from zero. Italy, instead, appears highly relevant in short-term risk spreading, having an ARCH right coefficient that is highly significant and the second-highest in magnitude. Germany, on the opposite, has a risk-receiving propensity in the long run, having only a GARCH left coefficient statistically different from zero, but its magnitude is the second-to-last in the GARCH left part. This is in line with the low correlation with other series other than Germany, as displayed in the summary statistics, and it is also coherent with Germany's claims that have Italy and Ireland as counterpart.

Considering the magnitudes of coefficients, we expect Ireland to play a big role as a risk spreader, both in the short run and in the long run, because its right coefficients are highly significant and have the highest magnitude, even if they differ in sign. Given the interesting dynamics uncovered in the normalized claims owned by Germany, this is not surprising. Instead, France has the biggest risk-receiving coefficients in both the short run and long run. Again, this can be understood looking at the series of its claims with Italy as counterpart. Portugal appears to have a risk-receiving propensity; the risk channel should be the claims it has with Spain as a counterpart. Spain, in this first rough analysis, shows a role both in spreading and receiving risk.

4.3 Inferred Networks

We use the methodology outlined in Section 3.2 to build a graphical representation that allows us to monitor the edges of the network — that is, the level of spillover between two specific countries as measured by the time-varying off-diagonal element of matrices A and B. In particular, one of the main features of our model is the ability to test the significance of those implied edges with a linear restriction (c.f. 3.2). Due to the high number of

coefficients and their time variation, it is difficult to conceive a tabular representation for our results. Instead, in order to display the topology of spillover flows, we propose a signed weighted directed network representation of the sequences of off diagonal elements in $A(W_t)$ and $B(W_t)$, with the weight proportional to the edge width and color representing the sign. In particular, in Figure 1, we report the relevant networks for three crisis periods: Q4 2008 (i.e., the collapse of Lehman Brothers and the bank bailout in Ireland), Q2 2010 (i.e., Greece's bailout), Q2 2016 (i.e., the Brexit referendum).(Section S in the supplementary material includes the complete representation.) Since we are conducting $2 * (n^2 - n) = 60$ test, it could be argued that we need to adopt a multiple-hypothesis testing framework; for this reason, we apply a Bonferroni correction to the nominal 5% significance level and show in Figure 1 only edges with p-values smaller than $\frac{0.05}{60} = 8.3e-04$.[9] The model imposes a stable pattern on the network topology and signature. The ARCH spillover topology shows how Spain, in the short term, is the central node receiving risk from Italy and Portugal and giving it to Ireland and France. Further, while France is giving risk back to Spain, Ireland is transferring risk to France and fueling a feedback loop with Germany. The overall picture is more involved in the long run (GARCH part), with several feedback loops, the strongest of which is between Portugal and Spain. We register, also, the appearance of a feedback loop between the two major economies: Germany and France. Here, Ireland has the same incoming interaction, but has two new targets, Italy and Portugal. Regarding the signature, a negative sign implies a sign change of the incoming shock, but we remind the reader that covariance contribution is a quadratic form in the shocks, and therefore negative contribution does not directly imply a covariance reduction. Again, the presence of edges of differing signs in both series of matrices favors the presence of covariance reducing terms. However, given the nonlinearity of the model, it is again difficult to understand where covariance reduction is located. In the next subsection, we will have a clearer description using covariance

[9]We report in Table S.3 of the supplementary material the p-values relevant for each off-diagonal element of the sequences $A(W_t)$ and $B(W_t)$, which under our hypothesis of no measurement errors in weights is time invariant.

decompositions.

The time variation we detect is in the edge weight (width). This is visible, for instance, in the ARCH edge from Italy to Spain; it is less effective in the Greek crisis, in the GARCH spillover from France to Germany, in the collapse of Lehman Brothers, in the sign-changing ARCH transmission from Ireland to Germany, and during the Brexit referendum. The big effect in 2011 from Germany to Ireland can be also appreciated by the pictures in the supplementary material S. The previous results point out the role of Spain and, to a lesser extent, of Ireland as risk middlemen able to transfer risk from peripheral countries to major economies.

4.4 Covariance Decompositions

The use of a Bonferroni correction could hide the role of less-significant but relevant links. In addition, as already noted, in the previous analyses, it is difficult to understand where and when diversification benefits are at work. For these reasons, we use herein the methodology of Subsection 3.3 on the covariance decomposition of the SB-BEKK model. This tool will allow us to understand who benefited the most from variance-reducing contributions mediated by the network, who is most impacted by network exposures, and who are the major risk spreaders.

Figure 2 reports the percentage of the system-variance, constant, mixed, and indirect contributions.[10] The representation of the system is an equally weighted portfolio.[11] The mixed contribution has both variance-increasing and variance-reducing contributions, with only a few large peaks of diversifying effect, mostly around turmoil. The indirect contribution is always positive and present during turbulent periods. In the second part of the sample, the relevance of network contribution is reduced in favor of the constant,

[10] We do not include the direct contribution since, although it is the biggest, it does not depend on the network links.

[11] It could be argued that an equally weighted portfolio does not represent a well-diversified portfolio. Indeed, this is the point of view put forward in Brunnermeier et al. (2017), which proposed weights proportional to the GDP to properly capture the sovereign credit risk of the Euro area. This possibility and the relevance of our model in evaluating the proposed European Safe Bonds will be investigated elsewhere.

with the notable exception of some important episodes, such as the Brexit referendum. This change in behavior could be understood by the developments in the second half of 2012, with Mario Draghi's "whatever it takes" speech in July, which led to the Outright Monetary Transactions framework — but also with the July establishment and September implementation of the European Stability Mechanism. These surrogates of a centralized lender of last resort dampened the relevance of cross-border claims as a contagion channel. Considering also variance decomposition for the other series, variance-reducing contributions come from the mixed part only in some periods, and are prevalent only for some countries. We summarize our findings with the following table, in which we report the cumulative percentage contribution on the whole estimation period.

Those results testify that most of the variance contribution is not mediated by the banking system exposures. This is not surprising since this is only one of several contagion channels. We also note how the indirect contribution is always superseded, in absolute value, by the mixed one, but the order of magnitude is, in the majority of cases, the same. In addition, the table shows that only Spain, Germany, and, to a lesser extent, Italy benefited from network effects. For Portugal and France in particular, and Ireland to a lesser extent, participation in the network of claims was dangerous. Looking again at the previously derived topology of spillover, we could now argue that those are the countries receiving risk from Spain and, in particular, that the countries most severely hit are the ones involved in an unbalanced feedback loop that favors Spain.

In Table 4. we report the downside. In fact, cumulative marginal spillover contributions, in percentage of the system variance, are able to identify the major sources of spillovers. Here we see that the most important countries are Portugal and Spain, but they contribute mostly in the amplification of their own variance. Instead, if we consider the contribution mediated through the banking exposures, the major sources of risk are Italy and Ireland, and, to a lesser extent, Spain. From the topology displayed before, we could argue that Ireland has a worrisome relationship with Germany in the short run and long run, and that it is contributing to the risk of France, Portugal, and Italy. It

seems that Italy's role as one of the major risk contributor to Spain has a big secondary impact to other countries through the previously described feedback loops. In addition, if we look at the P-values (Table S.3 in the supplementary material) we see that Italy impacts directly, but in a less significant way, on Portugal and France. France's shocks seem to have little cumulative diversifying effect. This could be due to the slightly prevalent covariance-reducing effect on Germany, with respect to secondary effects triggered by its interaction with Spain. Section S in the supplementary material reports the complete series of decompositions for each country, excluding direct contributions.

4.5 Estimated Counterfactual Exposures

In the previous section, the bank exposure network is characterized as significant but minor in the sample spillover channel. In this subsection, we evaluate, out of sample, by an ex-ante counterfactual analysis, to what extent changing exposures could be beneficial. In fact, the model allows for an estimation of the networks' impact on the system variance. Therefore, it is possible to evaluate the estimated model by focusing on the identification of the optimal (in terms of the conditional linear dependence in the system) network design. In particular, we propose to minimize the predicted path of the conditional average covariance in the system, looking for the optimal network structure, by following the methodology outlined in Subsection 3.4. We choose to optimize the forecast path in the quarter of the Brexit referendum (Q2 2016), while also including part of the subprime crisis and Greek bailout in the estimation sample (from Q4 2008 to Q1 2016). Our results convince us that the spillovers are channeled through foreign claims in the same way on the three occasions. We start by analyzing the counterfactual effect on the variance proxy of the equally weighted index. We perform the analysis by replacing the network of Q2 2016 with one of the following choices: a network with all exposures equal to zero, a network resulting from unconstrained optimization, a network coming from a redistribution-constrained optimization, and, finally, a network coming from a Frobenius-norm-constrained optimization. For the redistribution-constrained model, equation (16)

implies that there is only a redistribution of the claims among the considered countries; for the Frobenius-constrained model, equation (18) implies that the change in the network is comparable with the historically registered changes, but the total amount of claims can change; for the unconstrained model, the total amount of claims changes for each country, and extreme changes with respect to the historical ones are allowed.

Figure 3 shows the realized and counterfactual variance proxy of the equally weighted index during the sovereign debt crisis. We compute the variance according to equations (20) and (21). In addition, we plot the proxy coming from a network with zero exposures. The prevalent diversifying effect of exposures is evident in comparing the realized proxy with the zeroed one. The percentage change of the cumulative zeroed proxy with respect to the realized one is an increment of 16.55%. This also clarifies the fact that the naïve complete deleveraging solution, in our example, amplifies risk. This points out the necessity of a nontrivial exposure change, if we want to achieve risk mitigation. Our optimal counterfactual exercise is the first step in this direction. In particular, we think that several additional insights could be obtained by looking at the optimal proxy in 3. First of all, we are indeed able to reduce the out-of-sample counterfactual variance of the system by optimizing the network, based only on ex-ante information. This remains true also if we are imposing that the amount of claims for each country remain the same, or that the change is aligned with historical ones. As we were expecting, the optimization without constraints yielded the best percentage of cumulative reduction (-22.35%) and, in this case, the Brexit peak is mitigated and delayed. The drawback of this approach is that the variance is at several points in time above the realized one. On the opposite side, the reduction is halved, with the Frobenius-norm constraint (-11.18%), reduced by a lesser but comparable magnitude (-9.66%) with the redistribution constraint. The Brexit peak drop is less perceptible, but is indeed present and again higher for the Frobenius-norm constraint. Nevertheless, risk mitigation is much more stable in time and the variance is always less than or equal to the realized one in both constrained cases. Additionally, our procedure reduces an average correlation proxy (see Figure S.3 in the supplementary

material). Instead, a procedure relying directly upon the minimization of conditional average correlation has the drawback of increasing the variance (i.e., the risk) of some countries.[12]

The remarkable performance of the unconstrained procedure is obtained through drastic changes in the exposures (see Table S.4 in the supplementary material). We cannot be sure that extraordinary changes do not trigger a reaction from economic agents, which in turn could lead to a structural break in network parameters. As can be seen from our test on parameter constancy (see Table 5), we cannot reject the null hypothesis (i.e., parameter constancy) in this extreme case. In any case, we have stronger evidence considering the constrained instances. In particular, it is reassuring that the Frobenius-norm constraint is the best-performing optimization procedure, because it was designed as a solution to this particular issue. It is also interesting that an economically motivated constraint such as the redistribution one has comparable evidence in terms of parameter constancy. Considering the amount of variance reduction, its path through time, and the evidence for parameter constancy, the best compromise appears to be the optimal networks obtained by imposing the Frobenius-norm constraint.

To evaluate the feasibility of its changes, in Table 6, we report, for the Frobenius case only, the differences in billions of US dollars in the amounts needed to achieve the optimal network. Looking at Table 6, we rediscover some previous results. In particular, the results show in a clear way the necessity of deleveraging from Italy and Ireland, the countries where the major sources of risk exist in the mixed and indirect cases. Germany should invest more in other countries, as the evidence from Table 3 suggests it is the less-fragile country, and indeed it is asked to invest highly in all but Ireland and Italy, and in particular in Portugal. Portugal, in general, should be the target of most new investments and at the same time should reduce all its exposures, again in line with previous results in Table 3 and its small importance in 4. France's role appears in a less-clear way, and we

[12]Result are available upon request.

could argue that it is expected to strengthen its feedback loop with Germany in order to support Portugal and, to a lesser extent, Spain, whose nontrivial contribution was already pointed out. Here, the target network proposes a better balance in the feedback loop with Portugal, and a deleverage in all the others. Furthermore, it is suggested that Germany, and to a lesser extent France, support the country, with other peripheral countries that divert their exposures from it. In general, the variance reduction we obtain is implied by exposures changes that would be hard to enforce in a single quarter. In our opinion, a lower but still meaningful variance reduction can be obtained by considering stricter and economically sound maximum redistribution constraints, leading to an implementable enforcement of redistribution. This is already possible with a minor modification of our methodology that enables us to account for any kind of constraint by simply changing the equation (16). In addition, the procedure is based on conditional covariance and could be regularly updated, closely following actual market evolution after the changes.

5 Conclusions

This paper illustrates how financial networks can be efficiently integrated within a multivariate GARCH model for risk analyses. We refer to the proposed framework as spatiotemporal econometrics of networked risk, since it exploits spatial econometrics methodology in the investigation of a network of risk relationships. The spatiotemporal econometrics of risk enables a number of evaluations and analyses aimed at disentangling and understanding the role of asset interconnection in the evolution of the risk of a system. In particular, our parsimonious econometric model directly measures how variance and covariance spillovers flow through an exposure channel. Moreover, it can investigate long-run and short-run spillovers, through the ARCH and the GARCH parts, and it entails the possibility of diversifying or enhancing spillover contributions through coefficients of both signs. Our work builds on the introduction of spatial methods into volatility models in Caporin and Paruolo (2015). The model considers proximity matrices conveying the

economic distances among assets and capturing the interdependence across the modeled variables. This simple network model of an economic medium, based on data, through which risk can flow allows a better investigation of the determinants of dependence and spillover effects, with respect to purely statistical devices. We make a number of contributions that go beyond the framework of Caporin and Paruolo (2015). We take advantage of the non-commutativity of matrices in modeling, and focus on both the risk-receiving propensity and risk-spreading effectiveness of spillovers. We use graphtheoretic algebraic connectivity results in the study of identification. We make use of joint spectral characteristics for a set of matrices and derive the convergence of a heterogeneous geometric series to obtain covariance stationarity; we obtain this result through a VARMA representation of the model with time-varying coefficients. Moreover, we derive the asymptotic normality of the QMLE under conditions that are standard in the literature. Again, the usefulness of our joint-spectral-radius approach is evident in adapting proof of consistency and asymptotic normality from the standard framework in Hafner and Preminger (2009). The study of the identification, covariance stationarity, and asymptotic normality in the general case allows us to include a general normalization procedure for proximity matrices, driven by economic insight, whose payoff is made clear in the empirical application.

In addition, the empirical application shows the ability of our model to give a reasonable description of European spillovers during the sovereign crisis, uncovering the fundamental role of France and Portugal as risk receivers and the risk-spreading effectiveness of Italy and Ireland. Our inferred network methodology also points out Spain's middleman role, which is not understandable from the single-parameter significance. We also derive a covariance decomposition that allows us to understand the network-mediated contribution to variance, pointing out in particular how several network configurations can reduce the covariance. This effect is possible only with our complete study of identification conditions, which leaves unrestricted the sign of most of the model coefficients. In our empirical analysis, this diversifying effect was produced only by the mixed contribution in specific time periods. In addition, we document a change of relevance of

exposures after regulatory interventions in the second half of 2012. Spillover spreading and receiving effectiveness are heterogeneous across countries, and we study them by our variance decomposition and the marginal spillover contribution. The picture that emerges globally from our in-sample analysis points to Ireland and Italy as the major sources of risk and Portugal and France as the recipients of most of it. However, while Portugal is also a lesser source of risk, shocks coming from France appear mostly innocuous, if not beneficial, having a small risk-reducing and stabilizing effect. Germany transforms most risks it receives into diversification benefits, while contributing negligibly as a risk spreader. Spain's less-trivial role exploits network exposures for diversification benefits, while at the same time, it is the third source of network-mediated risk in the system. This middleman behavior in transferring risk from peripheral to major economies is direct with respect to France, and aided and mediated through Ireland with respect to Germany.

Finally, inspired by the monetary policy literature, we devise a counterfactual analysis able to obtain target exposures for risk mitigation based only on ex-ante information. The empirical results on the quarter during which the Brexit referendum took place are an additional soundness check of our model. They confirm the narrative implied by the in-sample analysis and the essential diversifying effect of banking-system exposures. Since the seminal works of Allen and Gale (2000), Eisenberg and Noe (2001) Freixas et al. (2000), network of bank exposures were considered an important theoretical channel of spillover effects. The focus on covariance allows, in our opinion, a better description of the phenomenon with respect to the use of plain spatial models that were already fruitful in this respect (see Tonzer (2015)). According to the recent review by Toniolo and White (2016) that focuses on the financial-stability mandate across countries and across history, the principal interventions that central banks took to maintain financial stability were the liquidity provision and the monitoring of systemically important financial institutions. In our paper, we propose a new econometric framework with the potential ability to help the regulator to fulfill the financial system's monitoring requirement in an empirically measurable way. To achieve this ambitious goal, two major obstacles remain. The first is

data disclosure: the bilateral exposures are made available only at the country level. The second one is methodological. In fact, the application of our methodology to a greater number of players requires the study of a covariance-targeting estimator of the model in order to obtain a number of parameters growing linearly with the number of nodes. Since rigorous covariance-targeting results for the standard multivariate volatility models were obtained only recently in the literature (Pedersen and Rahbek (2014),Francq et al. (2016)), this aspect is left for future research.

Although we are aware of those limitations, we feel that the major steps toward a quantification of the importance of a particular economic spillover channel have been completed through an intriguing example of cross-fertilization between application and econometric innovation.

References

Viral Acharya, Itamar Drechsler, and Philipp Schnabl. A pyrrhic victory? bank bailouts and sovereign credit risk. *The Journal of Finance*, 69(6):2689–2739, 2014.

Aida Alemany, Laura Ballester, and Ana Gonzalez-Urteaga. Volatility spillovers in the european bank cds market. *Finance Research Letters*, 13:137 – 147, 2015.

Franklin Allen and Douglas Gale. Financial contagion. *Journal of Political Economy*, 108(1): pp. 1–33, 2000.

Torben G Andersen, Tim Bollerslev, Peter F Christoffersen, and Francis X Diebold. Volatility and correlation forecasting. *Handbook of economic forecasting*, 1:777–878, 2006.

L. Anselin. *Spatial Econometrics: Methods and Models*. Studies in Operational Regional Science. Springer Netherlands, 2013.

Marco Avarucci, Eric Beutner, and Paolo Zaffaroni. On moment conditions for quasi-maximum likelihood estimation of multivariate arch models. *Econometric Theory*, 29(3):545–566, 2013.

Geert Bekaert and Campbell R Harvey. Emerging equity market volatility. *Journal of Financial economics*, 43(1):29–77, 1997.

Sylvain Benoit, Jean-Edouard Colliard, Christophe Hurlin, and Christophe Perignon. Where the risks lie: A survey on systemic risk*. *Review of Finance*, 21(1):109–152, 2017.

Ruslan Bikbov and Mikhail Chernov. Monetary policy regimes and the term structure of interest rates. *Journal of Econometrics*, 174(1):27 – 43, 2013.

Monica Billio and Loriana Pelizzon. Volatility and shocks spillover before and after emu in european stock markets. *Journal of Multinational Financial Management*, 13(4):323–340, 2003.

Monica Billio, Massimiliano Caporin, Roberto Calogero Panzica, and Loriana Pelizzon. The impact of network connectivity on factor exposures, asset pricing and portfolio diversification. SAFE Working Paper Series 166, Research Center SAFE - Sustainable Architecture for Finance in Europe, Goethe University Frankfurt, 2017.

Francisco Blasques, Siem Jan Koopman, Andre Lucas, and Julia Schaumburg. Spillover dy-

namics for systemic risk measurement using spatial financial time series models. *Journal of Econometrics*, 195(2):211 – 223, 2016.

Vincent D. Blondel and Yurii Nesterov. Computationally efficient approximations of the joint spectral radius. *SIAM Journal on Matrix Analysis and Applications*, 27(1):256–272, 2005.

Patrick Bolton and Olivier Jeanne. Sovereign default risk and bank fragility in financially integrated economies. *IMF Economic Review*, 59(2):162–194, Jun 2011.

Philippe Bougerol and Nico Picard. Strict stationarity of generalized autoregressive processes. *Ann. Probab.*, 20(4):1714–1730, 10 1992.

Farid Boussama, Florian Fuchs, and Robert Stelzer. Stationarity and geometric ergodicity of bekk multivariate garch models. *Stochastic Processes and their Applications*, 121(10):2331–2360, 2011.

Andreas Brandt. The stochastic equation yn 1=anyn bn with stationary coefficients. *Advances in Applied Probability*, 18(1):211,Äì220, 1986.

Richard A. Brualdi. Kronecker products of fully indecomposable matrices and of ultrastrong digraphs. *Journal of Combinatorial Theory*, 2(2):135 – 139, 1967.

Markus K. Brunnermeier, Sam Langfield, Marco Pagano, Ricardo Reis, Stijn Van Nieuwerburgh, and Dimitri Vayanos. Esbies: safety in the tranches. *Economic Policy*, 32(90):175–219, 2017.

Massimiliano Caporin and Paolo Paruolo. Proximity-structured multivariate volatility models. *Econometric Reviews*, 34(5):559–593, 2015.

Giancarlo Corsetti, Marcello Pericoli, and Massimo Sbracia. Correlation analysis of financial contagion. *Financial contagion: The viral threat to the wealth of nations*, pages 11–20, 2011.

Mardi Dungey, John Harvey, and Vladimir Volkov. The changing international network of sovereign debt and financial institutions. Working Papers 2017-04, University of Tasmania, Tasmanian School of Business and Economics, 2017.

Larry Eisenberg and Thomas H Noe. Systemic risk in financial systems. *Management Science*, 47(2):236–249, 2001.

J. Paul Elhorst. Dynamic models in space and time. *Geographical Analysis*, 33(2):119–140, 2001.

J Paul Elhorst. Specification and estimation of spatial panel data models. *International regional science review*, 26(3):244–268, 2003.

Robert F. Engle and David F. Hendry. Testing superexogeneity and invariance in regression models. *Journal of Econometrics*, 56(1):119 – 139, 1993.

Robert F Engle and Kenneth F Kroner. Multivariate simultaneous generalized arch. *Econometric theory*, 11(01):122–150, 1995.

Emmanuel Farhi and Jean Tirole. Deadly embrace: Sovereign and financial balance sheets doom loops. Working Paper 21843, National Bureau of Economic Research, January 2016.

Trevor I. Fenner and Georghios Loizou. On fully indecomposable matrices. *Journal of Computer and System Sciences*, 5(6):607 – 622, 1971.

Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak mathematical journal*, 23(2): 298–305, 1973.

Miroslav Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Mathematical Journal*, 25(4):619–633, 1975.

Kristin Forbes. The" big c": Identifying contagion. Technical report, National Bureau of Economic Research, 2012.

Christian Francq, Lajos Horvath, and Jean-Michel Zakoian. Variance targeting estimation of multivariate garch models. *Journal of Financial Econometrics*, 14(2):353–382, 2016.

Xavier Freixas, Bruno M Parigi, and Jean-Charles Rochet. Systemic risk, interbank relations, and liquidity provision by the central bank. *Journal of money, credit and banking*, pages 611–638, 2000.

Nicola Gennaioli, Alberto Martin, and Stefano Rossi. Sovereign default, domestic banks, and

financial institutions. *The Journal of Finance*, 69(2):819–866, 2014.
Jonatan Groba, Juan A. Lafuente, and Pedro Serrano. The impact of distressed economies on the eu sovereign market. *Journal of Banking & Finance*, 37(7):2520 – 2532, 2013.
Leonid Gurvits. Stability of discrete linear inclusion. *Linear algebra and its applications*, 231: 47–85, 1995.
Christian M Hafner and Helmut Herwartz. Analytical quasi maximum likelihood inference in multivariate volatility models. *Metrika*, 67(2):219–239, 2008.
Christian M Hafner and Arie Preminger. On asymptotic theory for multivariate garch models. *Journal of Multivariate Analysis*, 100(9):2044–2054, 2009.
Christian M. Hafner and Jeroen V. K. Rombouts. Estimation of temporally aggregated multivariate garch models. *Journal of Statistical Computation and Simulation*, 77(8):629–650, 2007.
Raphaël Jungers. *The Joint Spectral Radius: Theory and Applications*. Lecture Notes in Control and Information Sciences. Springer Berlin Heidelberg, 2009. ISBN 9783540959793.
Sebastian Keiler and Armin Eder. Cds spreads and systemic risk: A spatial econometric approach. *Discussion Paper N.01/2013 Deutsche Bundesbank*, 2013.
Lung-fei Lee and Jihai Yu. Qml estimation of spatial dynamic panel data models with time varying spatial weights matrices. *Spatial Economic Analysis*, 7(1):31–74, 2012.
James LeSage and Robert Kelley Pace. *Introduction to spatial econometrics*. CRC press, 2009.
James P LeSage and R Kelley Pace. Interpreting spatial econometric models. In *Handbook of Regional Science*, pages 1535–1552. Springer, 2014.
Robert E. Lucas. Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, 1(Supplement C):19 – 46, 1976.
J.R. Magnus and H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley Series in Probability and Statistics: Texts and References Section. Wiley, 1999. ISBN 9780471986331.
Bennet T. Mccallum. Robustness properties of a rule for monetary policy. *Carnegie-Rochester Conference Series on Public Policy*, 29(Supplement C):173 – 203, 1988.
Robert McCauley, Patrick McGuire, and Goetz von Peter. The architecture of global banking: from international to multinational? *BIS Quarterly Review*, page 25, 2010.
Patrick McGuire and Philip Wooldridge. The bis consolidated banking statistics: structure, uses and recent enhancements. *BIS Quarterly Review*, page 73, 2005.
Angela Ng. Volatility spillover effects from japan and the us to the pacific–basin. *Journal of international money and finance*, 19(2):207–233, 2000.
Rasmus S. Pedersen and Anders Rahbek. Multivariate variance targeting in the bekk-garch model. *The Econometrics Journal*, 17(1):24–55, 2014.
Phillip E. Pfeifer and Stuart Jay Deutsch. Identification and interpretation of first order space-time arma models. *Technometrics*, 22(3):397–408, 1980. doi: 10.1080/00401706.1980.10486172.
Dimitris N Politis and Joseph P Romano. A circular block-resampling procedure for stationary data. *Exploring the limits of bootstrap*, pages 263–270, 1992.
Dimitris N Politis and Halbert White. Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 23(1):53–70, 2004.
Giorgio E. Primiceri. Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, 72(3):821–852, 2005.
Xi Qu, Lung fei Lee, and Jihai Yu. Qml estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices. *Journal of Econometrics*, 197(2):173 – 201, 2017.
Oscar Rojo. A nontrivial upper bound on the largest laplacian eigenvalue of weighted graphs.

Linear Algebra and its Applications, 420(2):625 – 633, 2007.

Gian-Carlo Rota and W. Gilbert Strang. A note on the joint spectral radius. *Indagationes Mathematicae (Proceedings)*, 63:379 – 381, 1960.

Julio J. Rotemberg and Michael Woodford. An optimization-based econometric framework for the evaluation of monetary policy. *NBER Macroeconomics Annual*, 12:297–346, 1997.

Walmir Silva, Herbert Kimura, and Vinicius Amorim Sobreiro. An analysis of the literature on systemic financial risk: A survey. *Journal of Financial Stability*, 28:91 – 114, 2017.

T. Subba Rao and Ana Monica Costa Antunes. *Spatio-temporal Modelling of Temperature Time Series: A Comparative Study*, pages 123–150. Springer New York, New York, NY, 2004.

Gianni Toniolo and Eugene N White. The evolution of the financial stability mandate. In M.D. Bordo, Ø. Eitrheim, M. Flandreau, and J.F. Qvigstad, editors, *Central Banks at a Crossroads: What Can We Learn from History?*, Studies in Macroeconomic History, chapter 11, pages 424–492. Cambridge University Press, 2016.

Lena Tonzer. Cross-border interbank networks, banking risk and contagion. *Journal of Financial Stability*, 18:19 – 32, 2015.

John N. Tsitsiklis and Vincent D. Blondel. The lyapunov exponent and joint spectral radius of pairs of matrices are hard—when not impossible—to compute and to approximate. *Mathematics of Control, Signals and Systems*, 10(1):31–40, Mar 1997.

Guillaume Vankeerberghen, Julien Hendrickx, and Raphaël M. Jungers. Jsr: A toolbox to compute the joint spectral radius. In *Proceedings of the 17th International Conference on Hybrid Systems: Computation and Control*, HSCC '14, pages 151–156, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2732-9.

A Portfolio Variance Decomposition and Marginal Spillovers contributions

$$\left(\sigma_t^{Constant}\right)^2 = \sum_{i,j=1}^{n} z_i \left[CC'\right]_{i,j} z_j \tag{23}$$

$$\left(\sigma_t^{Direct}\right)^2 = \sum_{i,j=1}^{n} z_i \left(v_{i,t-1}^0 v_{j,t-1}^0 + \left[\Omega_{t-1}^{0,0}\right]_{ij}\right) z_j \tag{24}$$

$$\begin{aligned}\left(\sigma_t^{Indirect}\right)^2 &= \sum_{i,j=1}^{n} z_i \Big(v_{L,i,t-1}^1 v_{L,j,t-1}^1 + v_{R,i,t-1}^1 v_{R,j,t-1}^1 + \left[\Omega_{L,L,t-1}^{1,1}\right]_{i,j} + \left[\Omega_{R,R,t-1}^{1,1}\right]_{i,j} \\ &+ v_{L,i,t-1}^1 v_{R,j,t-1}^1 + v_{R,i,t-1}^1 v_{L,j,t-1}^1 + \left[\Omega_{L,R,t-1}^{1,1}\right]_{i,j} + \left[\Omega_{R,L,t-1}^{1,1}\right]_{i,j}\Big) z_j\end{aligned} \tag{25}$$

$$\begin{aligned}\left(\sigma_t^{Mixed}\right)^2 &= \sum_{i,j=1}^{n} z_i \Big(v_{L,i,t-1}^1 v_{j,t-1}^0 + v_{i,t-1}^0 v_{L,j,t-1}^1 + \left[\Omega_{L,t-1}^{1,0}\right]_{i,j} + \left[\Omega_{L,t-1}^{0,1}\right]_{i,j} \\ &+ v_{R,i,t-1}^1 v_{j,t-1}^0 + v_{i,t-1}^0 v_{R,j,t-1}^1 + \left[\Omega_{R,t-1}^{1,0}\right]_{i,j} + \left[\Omega_{R,t-1}^{0,1}\right]_{i,j}\Big) z_j\end{aligned} \tag{26}$$

$$MSC_{k,t}^{Direct} = \sum_{j=1}^{n} \left(v_{k,t-1}^0 v_{j,t-1}^0 + \left[\Omega_{t-1}^{0,0}\right]_{kj}\right) z_j \tag{27}$$

$$\tag{28}$$

$$\begin{aligned}MSC_{k,t}^{Indirect} &= \sum_{j=1}^{n} \Big(v_{L,k,t-1}^1 v_{L,j,t-1}^1 + v_{R,k,t-1}^1 v_{R,j,t-1}^1 + \left[\Omega_{L,L,t-1}^{1,1}\right]_{k,j} + \left[\Omega_{R,R,t-1}^{1,1}\right]_{k,j} \\ &+ v_{L,k,t-1}^1 v_{R,j,t-1}^1 + v_{R,k,t-1}^1 v_{L,j,t-1}^1 + \left[\Omega_{L,R,t-1}^{1,1}\right]_{k,j} + \left[\Omega_{R,L,t-1}^{1,1}\right]_{k,j}\Big) z_j\end{aligned} \tag{29}$$

$$\begin{aligned}MSC_{k,t}^{Mixed} &= \sum_{j=1}^{n} \Big(v_{L,k,t-1}^1 v_{j,t-1}^0 + v_{k,t-1}^0 v_{L,j,t-1}^1 + \left[\Omega_{L,t-1}^{1,0}\right]_{k,j} + \left[\Omega_{L,t-1}^{0,1}\right]_{k,j} \\ &+ v_{R,k,t-1}^1 v_{j,t-1}^0 + v_{k,t-1}^0 v_{R,j,t-1}^1 + \left[\Omega_{R,t-1}^{1,0}\right]_{k,j} + \left[\Omega_{R,t-1}^{0,1}\right]_{k,j}\Big) z_j\end{aligned} \tag{30}$$

B Proofs and Auxiliary Results

B.1 Identification

B.1.1 Proof of Proposition 1 on page 13

Proof. Without loss of generality, let us consider $A\left(W_k\right)$, we have:

$$\begin{aligned}\operatorname{vec}\left(A\left(W_k\right)\right) &= \operatorname{vec}\left(\operatorname{diag}\left(a_0\right)\right) + W_k' \otimes I_n \operatorname{vec}\left(\operatorname{diag}\left(a_{1L}\right)\right) + I_n \otimes W_k \operatorname{vec}\left(\operatorname{diag}\left(a_{1R}\right)\right) \\ &= \begin{bmatrix} I_n \otimes I_n & W_k' \otimes I_n & I_n \otimes W_k \end{bmatrix} \begin{bmatrix} \operatorname{vec}\left(\operatorname{diag}\left(a_0\right)\right) \\ \operatorname{vec}\left(\operatorname{diag}\left(a_{1,L}\right)\right) \\ \operatorname{vec}\left(\operatorname{diag}\left(a_{1,R}\right)\right) \end{bmatrix} = M\left(W_k\right) \begin{bmatrix} a_0 \\ a_{1,L} \\ a_{1,R} \end{bmatrix}.\end{aligned}$$

and analogously

$$\text{vec}\left(\left[A\left(W_1\right)' : A\left(W_2\right)' : \ldots : A\left(W_K\right)'\right]'\right) = \left[M\left(W_1\right)' : M\left(W_2\right)' \ldots : M\left(W_K\right)'\right]' \begin{bmatrix} a_0 \\ a_{1,L} \\ a_{1,R} \end{bmatrix}$$

The $[Kn^2 \times 3n]$ matrix $\bar{M} = \left[\left(W_1\right)' : M\left(W_2\right)' \ldots : M\left(W_K\right)'\right]'$, is of column rank equal to $\max_{1\ldots k} \text{rank}\left(M\left(W_k\right)\right)$. We will show in lemma 2 that $M\left(W_k\right)$ is of column rank $3n - c_k$. Then, if we impose $c = \min_{1\ldots k} c_k$ linear restriction on the parameter vector, variation in the parameters induces a unique variation in π_K. □

Lemma 2 (Rank of M)**.** *We have:*

$$\text{rank}\left(M\left(W_k\right)\right) = 2n + \text{rank}\left(L\left(\mathfrak{G}_k\right)\right), \tag{31}$$

where $L\left(\mathfrak{G}_k\right)$ is the Laplacian of the undirected graph $\mathfrak{G}_k$ whose rank is equal to $n - c_k$.

Proof.

$$\begin{aligned} M\left(W_t\right) &= \begin{bmatrix} I_n \otimes I_n & W_k' \otimes I_n & I_n \otimes W_k \end{bmatrix} \left(I_3 \otimes H\right) \\ &= \sum_{i=1}^{n} \left[\left(e_i \otimes e_i\right) e_i', \left(W_k' e_i \otimes e_i\right) e_i', \left(e_i \otimes W_k e_i\right) e_i'\right] \\ M\left(W_k\right)' M\left(W_k\right) &= \begin{bmatrix} I_n & \\ & \mathcal{M}\left(W_k\right) \end{bmatrix} \quad \mathcal{M}\left(W_k\right) = \begin{bmatrix} D_{W_k \odot W_k} & W_k \odot W_k \\ W_k' \odot W_k' & D_{W_k' \odot W_k'} \end{bmatrix} \\ D_{W_k' \odot W_k'} &= \left(W_k' \odot W_k'\right) \mathbf{1}_n \quad , \quad D_{W_k \odot W_k} = \left(W_k \odot W_k\right) \mathbf{1}_n \ . \end{aligned}$$

Given that the rank of a generic matrix T is equal to the rank of $T'T$, it suffices to compute the rank of $M\left(W_k\right)' M\left(W_k\right)$. Having the identity I_n as one of the diagonal blocks, $M\left(W_k\right)' M\left(W_k\right)$ has at least rank n. In addition, using the property of the determinant for block matrices, the characteristic polynomial of $\mathcal{M}\left(W_k\right)$, is:

$$\begin{aligned} \det\left(\mathcal{M}\left(W_k\right) - \lambda I_{n^2}\right) &= \det\left(D_{W_k^{\odot} W_k} - \lambda I_n\right) \det\left(L\left(\mathfrak{G}_k\right) - \lambda I_n\right) = 0 \\ L\left(\mathfrak{G}_k\right) &= D_{W_k' \odot W_k'} - \mathfrak{W}_k \end{aligned}$$

Since $D_{W_k^{\odot} W_k}$ is diagonal, $\det\left(D_{W_k^{\odot} W_k} - \lambda I_n\right) = 0$ has exactly n solutions and the rank of $M\left(W_k\right)' M\left(W_t\right)$ is at least $2n$. We obtain the final result for the rank of $M\left(W_k\right)' M\left(W_k\right)$ by discussing the dimension of the kernel of $L\left(\mathfrak{G}_k\right)$. Since $\mathfrak{W}_k \mathbf{1}_n = D_{W_k' \odot W_k'}$, $L\left(\mathfrak{G}_k\right)$ can be interpreted as the Laplacian of the graph $\mathfrak{G}_k$ with weight matrix $\mathfrak{W}_k$, and it is well known from the work of Fiedler (Fiedler, 1973, 1975) that the dimension of the kernel of $L\left(\mathfrak{G}_k\right)$ is the number of connected components c_k of $\mathfrak{G}_k$. Then, we have $\text{rank}\left(M\left(W_k\right)\right) = \text{rank}\left(M\left(W_k\right)' M\left(W_k\right)\right) = 3n - c_k$. □

B.1.2 Proof of Lemma 1 on page 13

Proof. The result follow from the fact that the non zero elements of $\mathfrak{W}_k$ are the same of $\left(W_k' \odot W_k'\right)\left(W_k \odot W_k\right)$ and consequently of $W_k' W_k$. Then, since the multiplication of

two fully indecomposable matrices is fully indecomposable, and fully indecomposability implies irreducibility, Fenner and Loizou (1971), under the hypothesis of the lemma there is a at least one $k = 1 \dots K$ for which $\mathfrak{W}_k$ has only one component. Then, the result follows from Proposition 1. □

B.1.3 Identification for Symmetric Matrices

Corollary 1. *If the matrices W_k are all symmetric, to achieve identification the model must include either the left multiplication or the right multiplication elements. The previous result holds with a corresponding simplification in the structure of $M(W_k)$ and the model is identified if the corresponding $\bar{M}$ is full column rank.*

Proof. The corollary is a consequence of symmetry of matrices W_k. Suppose we focus on the shock component and assume a constant W. We have $A_L u_{t-1} u'_{t-1} A'_L = (A_0 + A_{1,L}W)\, u_{t-1} u'_{t-1} (A_0 + W' A_{1,L})$ thanks to the diagonal form of the parameter matrices. Moreover, by symmetry, $(A_0 + A_{1,L}W)\, u_{t-1} u'_{t-1} (A_0 + W' A_{1,L})$ and the latter equals $(A_0 + W' A_{1,L})\, u_{t-1} u'_{t-1} (A_0 + A_{1,L}W)$. The latter is equal to the right multiplication case if $W = W'$.

Then, with only left(right) parameters, we have :

$$M(W_k)' M(W_k) = \begin{bmatrix} I_n & \\ & D_{W_k \odot W_k} \end{bmatrix}.$$

This matrix is full rank unless some of the row(column) sums are zero. Since the W_k are non negative this can happen only if the row(column) elements are all zero and we reduce to the conditions given in (Caporin and Paruolo, 2015) □

B.2 Covariance Stationarity

The following lemma considers a deterministic network sequence and gives a sufficient condition on the joint spectral radius such that it probably collapses to the one given in terms of the top Lyapunov exponent in Brandt (1986) for scalar coefficients. Bougerol and Picard (1992) extended the last result to coefficient matrices.[13] Note that results in Bougerol and Picard (1992) are valid under the assumption that the network sequence has a stationary distribution. A formal proof of the equivalence is left for future research.

Lemma 3 (Convergence of Heterogeneous Geometric Series)**.** *If $\varrho(\mathcal{A}) < 1$ then*

$$S = \lim_{T \to \infty} S_n = \lim_{T \to \infty} \sum_{t=0}^{T} \prod_{i=0}^{t} A_i < \infty \tag{32}$$

Proof. We can use here the Cauchy convergence criterion. Given $m > n \in \mathcal{N}$ with fixed but arbitrary $r = m - n$, we have to show that $\|S_m - S_n\| \to 0$ when $n \to \infty$.

[13] A similar condition is also used by Hafner and Preminger (2009) in showing stationarity and ergodicity of the VEC model.

We have :

$$0 \leq \|S_m - S_n\| = \left\| \sum_{k=n}^{m} B_k \right\| \leq \sum_{k=n}^{m} \sup_{A_{\sigma_1},\ldots,A_{\sigma_k} \in \mathcal{A}} \left(\left\| \prod_{i=1}^{k} A_{\sigma_i} \right\| \right) \tag{33}$$

and we can write

$$\begin{aligned} \lim_{n\to\infty} \sum_{k=n}^{m} \sup_{\sup_{A_{\sigma_1},\ldots,A_{\sigma_k} \in \mathcal{A}}} \left(\left\| \prod_{i=1}^{k} A_{\sigma_i} \right\| \right) &= \lim_{n\to\infty} \sum_{k=n}^{m} \left(\hat{\rho}_k (\mathcal{A}) \right)^k \\ = \lim_{n\to\infty} \varrho(\mathcal{A})^n \lim_{n\to\infty} \sum_{k=n}^{m} \frac{\left(\hat{\rho}_k (\mathcal{A}) \right)^k}{\varrho(\mathcal{A})^n} &= \lim_{n\to\infty} \frac{1 - \varrho(\mathcal{A})^{r+1}}{1 - \varrho(\mathcal{A})} \varrho(\mathcal{A})^n \to 0 \end{aligned}$$

□

The use of the joint spectral radius leads, also, to an interesting property for matrix sets that leave a proper cone invariant.

Definition B.1. *Cones and Proper Cones. A cone in $\mathbb{R}^n$ is a subset $K \subseteq \mathbb{R}^n$ such that $\lambda v \in K$ for all $\lambda \geq 0$ and $v \in K$. K is proper if it is closed, convex, has non empty interior, and contains no straight lines.*

Definition B.2. *Cone Invariance. $A_i \in \mathcal{A}$ leave a proper cone invariant if there exist a proper cone $K \subseteq \mathbb{R}^n$ such that if $v \in K$ $u_i = A_i v \in J_i$ with $J_i \subseteq K$ for each i.*

According to Blondel and Nesterov (2005) it is possible to define a norm $\|\cdot\|_K$ associated with the cone K such that, if A and B leave the proper cone K invariant, then $\|A\|_K \leq \|A + B\|_K$. This norm will be the essential in the proof of the following lemma.

Lemma 4. *Consider two infinite set of $n \times n$ matrices $\mathcal{A} = \{A_i\}_{i=0}^{\infty}$ and $\mathcal{B} = \{B_i\}_{i=0}^{\infty}$ and their sum $\mathcal{A} + \mathcal{B} = \{M \in \mathcal{A} + \mathcal{B} | M = A_i + B_i, A_i \in A, B_i \in \mathcal{B}\}$ with $A_0, B_0 = I_n$. Suppose that $\mathcal{A}$ and $\mathcal{B}$ leave the cone K invariant. We have $\varrho(\mathcal{B}) \leq \varrho(\mathcal{A} + \mathcal{B})$.*

Proof. $\prod_{i=1}^{t} A_{\sigma_i} + B_{\sigma_i}$ is a sum of 2^t products of t terms. By the assumptions the matrices C_j $j = 1, \ldots, 2^t$ obtained by each of this products leave K invariant. In addition we note that $C_{2^t} = \prod_{i=1}^{t} B_{\sigma_i}$. We obtain

$$\left\| \prod_{i=1}^{t} (A_{\sigma_i} + B_{\sigma_i}) \right\|_K = \left\| \prod_{i=1}^{t} B_{\sigma_i} + \sum_{j=1}^{2^t - 1} C_j \right\|_K \geq \left\| \prod_{i=1}^{t} B_{\sigma_i} \right\|_K$$

so

$$\begin{aligned} \varrho(\mathcal{B}) &= \lim_{t\to\infty} \sup_{B_{\sigma_1},\ldots,B_{\sigma_t} \in \mathcal{B}} \left(\left\| \prod_{i=1}^{t} B_{\sigma_i} \right\|_K \right)^{1/t} = \lim_{t\to\infty} \sup_{\left(A_{\sigma_1} + B_{\sigma_1}\right),\ldots,\left(A_{\sigma_t} + B_{\sigma_t}\right) \in (\mathcal{A}+\mathcal{B})} \left(\left\| \prod_{i=1}^{t} B_{\sigma_i} \right\|_K \right)^{1/t} \\ &\leq \lim_{t\to\infty} \sup_{\left(A_{\sigma_1} + B_{\sigma_1}\right),\ldots,\left(A_{\sigma_t} + B_{\sigma_t}\right) \in (\mathcal{A}+\mathcal{B})} \left(\left\| \prod_{i=1}^{t} (A_{\sigma_i} + B_{\sigma_i}) \right\|_K \right)^{1/t} = \varrho(\mathcal{A} + \mathcal{B}) \end{aligned}$$

□

B.2.1 Proof of theorem 1 on page 16

Proof. Let us consider the filtration $\mathcal{F}_{t-1} = \{\mathcal{U}_{t-1}, \mathcal{W}_{t-1}\}$ where $\mathcal{U}_{t-1}$ is the information set given by the past u_t and $\mathcal{W}_{t-1}$ is the information set of the past network W_t. We remark that, by definition of the processes, $\mathcal{W}_{t-2} \subset \mathcal{U}_{t-1}$.

Since η_t is a martingale difference sequence $\mathbb{E}\left[\eta_t \left|\mathcal{U}_{t-1}, \mathcal{W}_{t-1}\right.\right] = 0$, and we also have

$$\mathbb{E}\left[X_t \left|\mathcal{U}_{t-1}, \mathcal{W}_{t-1}\right.\right] = \tilde{C} + \left(\tilde{A}\left(W_t\right) + \tilde{B}\left(W_t\right)\right) X_{t-1} - \tilde{B}\left(W_t\right)\eta_{t-1}.$$

Using the iterated expectation theorem and the recursion for X_t

$$\begin{aligned}
\mathbb{E}\left[X_t \left|\mathcal{U}_{t-2}, \mathcal{W}_{t-1}\right.\right] &= \mathbb{E}\left[\mathbb{E}\left[X_t \left|\mathcal{U}_{t-1}, \mathcal{W}_{t-1}\right.\right] \left|\mathcal{U}_{t-1}, \mathcal{W}_{t-1}\right.\right] \\
&= \mathbb{E}\left[\tilde{C} + \left(\tilde{A}\left(W_t\right) + \tilde{B}\left(W_t\right)\right) X_{t-1} - \tilde{B}\left(W_t\right)\eta_{t-1} \left|\mathcal{U}_{t-2}, \mathcal{W}_{t-1}\right.\right] \\
&= \tilde{C} + \left(\tilde{A}\left(W_{t-1}\right) + \tilde{B}\left(W_{t-1}\right)\right) \mathbb{E}\left[X_{t-1} \left|\mathcal{U}_{t-2}, \mathcal{W}_{t-1}\right.\right].
\end{aligned}$$

Analogously,

$$\begin{aligned}
\mathbb{E}\left[X_t \left|\mathcal{U}_{t-3}, \mathcal{W}_{t-1}\right.\right] &= \left[I + \left[\tilde{A}\left(W_{t-1}\right) + \tilde{B}\left(W_{t-1}\right)\right]\right] \tilde{C} \\
&+ \left[\tilde{A}\left(W_{t-1}\right) + \tilde{B}\left(W_{t-1}\right)\right] \left[\tilde{A}\left(W_{t-2}\right) + \tilde{B}\left(W_{t-2}\right)\right] \mathbb{E}\left[X_{t-2} \left|\mathcal{U}_{t-3}, \mathcal{W}_{t-1}\right.\right],
\end{aligned}$$

Further, by induction

$$\begin{aligned}
\mathbb{E}\left[X_t \left|\mathcal{U}_{t-\tau}, \mathcal{W}_{t-1}\right.\right] &= \left\{I + \sum_{k=1}^{\tau-2} \prod_{l=1}^{k} \left[\tilde{A}\left(W_{t-l}\right) + \tilde{B}\left(W_{t-l}\right)\right]\right\} \tilde{C} \\
&+ \left\{\prod_{l=1}^{\tau-1} \left[\tilde{A}\left(W_{t-l}\right) + \tilde{B}\left(W_{t-l}\right)\right]\right\} \mathbb{E}\left[X_{t-\tau+1} \left|\mathcal{U}_{t-\tau}, \mathcal{W}_{t-1}\right.\right].
\end{aligned}$$

When $\tau \to \infty$, if $\varrho\left(\left\{\tilde{A}\left(W_t\right) + \tilde{B}\left(W_t\right)\right\}_{t=-\infty}^{\infty}\right) < 1$, by Lemma 3, conditionally on the network sequence, the first term converges to a finite limit and the remainder goes to zero according to Gurvits (1995) and Jungers (2009). □

B.3 Consistency and Asymptotic Normality

In this subsection we denote as $\|\cdot\|$ the operator norm. Different norms are specified when needed. D_n is the duplication matrix and C_{nm} is the commutation matrix. Under assumption 1 and by Lemma 4, since the set of half vectorized symmetric positive semidefinite matrices is a proper cone of the set of half vectorized symmetric matrices, i.e of the vector space $\mathbb{R}^{(n(n+1)/2)}$, and $\tilde{A}\left(W_t\right)$ and $\tilde{B}\left(W_t\right)$ leave this cone invariant (c.f subsection 4.2 in Boussama et al. (2011)). we have

$$\varrho_B = \varrho\left(\left\{\tilde{B}\left(W_t\right)\right\}_{t=-\infty}^{\infty}\right) \leq \varrho\left(\left\{\tilde{A}\left(W_t\right) + \tilde{B}\left(W_t\right)\right\}_{t=-\infty}^{\infty}\right) < 1 \quad , \quad \bar{\varrho}_B = \sup_{\theta\in\Theta} \varrho_B < 1.$$

Lemma 5. *Under Assumptions 1-4* $\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\Sigma_t-\tilde{\Sigma}_t\right\|^{s/2}\right]=O\left(\bar{\varrho}_B\right)$

Proof. By solving equation (9) recursively, we get

$$\begin{aligned}\xi_t &= \operatorname{vech}\left(\Sigma_t\right)=\tilde{C}+\tilde{A}\left(W_t\right)X_{t-1}+\tilde{B}\left(W_t\right)\xi_{t-1}\\ &= \tilde{C}+\tilde{A}\left(W_t\right)y_{t-1}+\sum_{\tau=1}^{t-1}\left\{\prod_{l=1}^{\tau-1}\tilde{B}\left(W_{t-l}\right)\right\}\left(\tilde{C}+\tilde{A}\left(W_{t-\tau}\right)X_{t-\tau-1}\right)+\left\{\prod_{l=1}^{t-1}\tilde{B}\left(W_{t-l}\right)\right\}\xi_0.\end{aligned}$$

In addition

$$\left\|\Sigma_t-\tilde{\Sigma}_t\right\|\leq\left\|\Sigma_t-\tilde{\Sigma}_t\right\|_2\leq\left\|D_n^+\right\|\left\|D_n\right\|\left\|\xi_t-\tilde{\xi}_t\right\|\leq\left\|D_n^+\right\|\left\|D_n\right\|\varrho_B^t\left\|\xi_0-\tilde{\xi}_0\right\|$$

By Assumption 2 and 3, and the c_r inequality we have $\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\Sigma_t-\tilde{\Sigma}_t\right\|^{s/2}\right]=O\left(\bar{\varrho}_B\right)$ □

Lemma 6. *Under Assumptions 1-7*

$$i)\quad \mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\partial\tilde{\xi}_t}{\partial a'}\right\|^3\right]<\infty\qquad\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\partial\tilde{\xi}_t}{\partial b'}\right\|^3\right]<\infty$$

$$ii)\quad \mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\partial^2\tilde{\xi}_t}{\partial a'\partial a'}\right\|^3\right]<\infty\quad\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\partial^2\tilde{\xi}_t}{\partial b'\partial a'}\right\|^3\right]<\infty\quad\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\partial^2\tilde{\xi}_t}{\partial b'\partial b'}\right\|^3\right]<\infty$$

Proof. Using the linear map that links $A\left(W_t\right)$ and $B\left(W_t\right)$ to a and b, it is possible to obtain the derivatives of $\tilde{A}\left(W_t\right)$ and $\tilde{B}\left(W_t\right)$. Consider, as an example, $\tilde{B}\left(W_t\right)$:

$$\begin{aligned}\operatorname{vec}\left(\tilde{B}\left(W_t\right)\right) &= \left(D_n'\otimes D_n^+\right)\left(I_n\otimes C_{nn}\otimes I_n\right)\operatorname{vec}\left(M\left(W_t\right)b\right)\otimes\operatorname{vec}\left(M\left(W_t\right)b\right)\\ &= \left(D_n'\otimes D_n^+\right)\left(I_n\otimes C_{nn}\otimes I_n\right)\left(M\left(W_t\right)\otimes M\left(W_t\right)\right)\left(b\otimes b\right)\\ \frac{\partial}{\partial b'}\operatorname{vec}\left(\tilde{B}\left(W_t\right)\right) &= \left(D_n'\otimes D_n^+\right)\left(I_n\otimes C_{nn}\otimes I_n\right)\left(M\left(W_t\right)\otimes M\left(W_t\right)\right)2D_{3n}D_{3n}^+\left(b\otimes I_{3n}\right)\\ \frac{\partial}{\partial b'}\operatorname{vec}\left(\frac{\partial}{\partial b'}\operatorname{vec}\left(\tilde{B}\left(W_t\right)\right)\right) &= \left[I_{3n}\otimes\left[\left(D_n'\otimes D_n^+\right)\left(I_n\otimes C_{nn}\otimes I_n\right)\left(M\left(W_t\right)\otimes M\left(W_t\right)\right)2D_{3n}D_{3n}^+\right]\right]\frac{\partial}{\partial b'}\operatorname{vec}\left(b\otimes I_{3n}\right)\\ &= \left[I_{3n}\otimes\left[\left(D_n'\otimes D_n^+\right)\left(I_n\otimes C_{nn}\otimes I_n\right)\left(M\left(W_t\right)\otimes M\left(W_t\right)\right)2D_{3n}D_{3n}^+\right]\right]\times\\ &\times\ \left(1\otimes C_{3n3n}\otimes I_{3n}\right)\left(I_{3n}\otimes\operatorname{vec}\left(I_{3n}\right)\right)\end{aligned}$$

Let us define the following quantities

$$\begin{array}{ccc}\|C_{nn}\|=K_{1,n} & \|(D_n'\otimes D_n^+)\|=K_{2,n} & \|(D_nD_n^+)\|=K_{3,n}\\ \sup_{\theta\in\Theta}\|a\|=K_a & \sup_{\theta\in\Theta}\|b\|=K_b & \sup_{\theta\in\Theta}\|C\|=K_c.\end{array}$$

Then, let us bound

$$\begin{aligned}\|(M(W_t)\otimes M(W_t))\| &= \sqrt{\rho\left((M(W_t)\otimes M(W_t))'(M(W_t)\otimes M(W_t))\right)}\\ &= \rho\left(M(W_t)'M(W_t)\right)=\rho_M.\end{aligned}$$

We already studied the characteristic polynomial of $M(W_t)'M(W_t)$ in the proof of Proposition 1; it has n eigenvalues equal to 1, eigenvalues equal to $\mu_i=\sum_{j=1}^n W_{t,ij}^2$ for $i=1,\ldots,n$ and the last n eigenvalues are the eigenvalues of the Laplacian $L(\mathfrak{G}_t)$. In particular, we are interested in $\rho(L(\mathfrak{G}_t))$. Bounds could be found in Rojo (2007), the most trivial one being the column sum over row maxima of the weight matrix.

$$\rho(L(\mathfrak{G}_t))<\sum_{j=1}^n \max_{i=1,\ldots,n}\mathfrak{W}_{t,ij}$$

By Assumption 7 we can bound $W_{t,ij}\leq d^*$ and $\mu_i\leq n(d^*)^2$. We have

$$\rho_M=\max\left(1,n(d^*)^2,\sum_{j=1}^n\max_{i=1,\ldots,n}\mathfrak{W}_{t,ij}\right) \leq \max\left(1,n(d^*)^2\right) \tag{34}$$

$$\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\mathrm{vec}\left(\tilde{B}(W_t)\right)\right\| \leq K_{2,n}K_{1,n}\rho_M K_b^2 \tag{35}$$

$$\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\mathrm{vec}\left(\tilde{B}(W_t)\right)\right\| \leq K_{2,n}K_{1,n}\rho_M 2K_{3,3n}K_b \tag{36}$$

$$\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\mathrm{vec}\left(\frac{\partial}{\partial b'}\mathrm{vec}\left(\tilde{B}(W_t)\right)\right)\right\| \leq K_{2,n}K_{1,n}\rho_M 2K_{3,3n}K_{1,3n} \tag{37}$$

$$\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\mathrm{vec}\left(\tilde{A}(W_t)\right)\right\| \leq K_{2,n}K_{1,n}\rho_M K_b^2 \tag{38}$$

$$\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial a'}\mathrm{vec}\left(\tilde{A}(W_t)\right)\right\| \leq K_{2,n}K_{1,n}\rho_M 2K_{3,3n}K_a \tag{39}$$

$$\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial}{\partial a'}\mathrm{vec}\left(\tilde{A}(W_t)\right)\right)\right\| \leq K_{2,n}K_{1,n}\rho_M 2K_{3,3n}K_{1,3n} \tag{40}$$

In addition, in the following we will need bounds for the first and second derivatives of the products. For the first derivative we have:

$$\begin{aligned}&\frac{\partial}{\partial b'}\mathrm{vec}\left(\left\{\prod_{l=r}^s\tilde{B}(W_{t-l})\right\}\right)\\ =\ &\sum_{l=r}^s\left\{\prod_{m'=s-l+1}^s\tilde{B}(W_{t-m'})\right\}'\otimes\left\{\prod_{m=r}^{s-l-1}\tilde{B}(W_{t-m})\right\}\frac{\partial}{\partial b'}\mathrm{vec}\left(\tilde{B}\left(W_{t-(s-l)}\right)\right)\end{aligned}$$

$$
\begin{aligned}
&\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\mathrm{vec}\left(\left\{\prod_{l=r}^{s}\tilde{B}\left(W_{t-l}\right)\right\}\right)\right\| \\
\leq\ & \sum_{l=r}^{s}\left\{\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\left\{\prod_{m'=s-l+1}^{s}\tilde{B}\left(W_{t-m'}\right)\right\}\right\|\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\left\{\prod_{m=r}^{s-l-1}\tilde{B}\left(W_{t-m}\right)\right\}\right\|\times\right. \\
\times\ & \left.\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\mathrm{vec}\left(\tilde{B}\left(W_t\right)\right)\right\|\right\} \\
\leq\ & \sum_{l=r}^{s}\bar{\varrho}_B^{l}\bar{\varrho}_B^{s-l-r}\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\mathrm{vec}\left(\tilde{B}\left(W_t\right)\right)\right\| \\
\leq\ & \left(s-r+1\right)\bar{\varrho}_B^{s-r}\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\mathrm{vec}\left(\tilde{B}\left(W_t\right)\right)\right\| \qquad (41)
\end{aligned}
$$

For the second derivative we need the following intermediate result:

$$
\begin{aligned}
\frac{\partial}{\partial b'}\mathrm{vec}\left(\left\{\prod_{m'=s-l+1}^{s}\tilde{B}\left(W_{t-m'}\right)\right\}'\otimes\left\{\prod_{m=r}^{s-l-1}\tilde{B}\left(W_{t-m}\right)\right\}\right) &= \left[I_{n(n+1)/2}\otimes C_{n(n+1)/2}\otimes I_{n(n+1)/2}\right]\times \\
\times\left\{\frac{\partial}{\partial b'}\mathrm{vec}\left(\left\{\prod_{m'=s-l+1}^{s}\tilde{B}\left(W_{t-m'}\right)\right\}'\right)\otimes\mathrm{vec}\left(\left\{\prod_{m=r}^{s-l-1}\tilde{B}\left(W_{t-m}\right)\right\}\right)+\right. \\
\left.+\ \mathrm{vec}\left(\left\{\prod_{m'=s-l+1}^{s}\tilde{B}\left(W_{t-m'}\right)\right\}'\right)\otimes\frac{\partial}{\partial b'}\mathrm{vec}\left(\left\{\prod_{m=r}^{s-l-1}\tilde{B}\left(W_{t-m}\right)\right\}\right)\right\}
\end{aligned}
$$

by which, we get the bound:

$$
\begin{aligned}
&\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\mathrm{vec}\left(\left\{\prod_{m'=s-l+1}^{s}\tilde{B}\left(W_{t-m'}\right)\right\}'\otimes\left\{\prod_{m=r}^{s-l-1}\tilde{B}\left(W_{t-m}\right)\right\}\right)\right\| \\
&\qquad\leq K_{1,n(n+1)/2}\left(l+1\right)\bar{\varrho}_B^{l}\bar{\varrho}_B^{s-l-r}\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\mathrm{vec}\left(\tilde{B}\left(W_t\right)\right)\right\| \\
&+K_{1,n(n+1)/2}\left(s-l-r+1\right)\bar{\varrho}_B^{s-l-r}\bar{\varrho}_B^{l}\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\mathrm{vec}\left(\tilde{B}\left(W_t\right)\right)\right\| \\
&\qquad= K_{1,n(n+1)/2}\left(s-r+1\right)\bar{\varrho}_B^{s-r}\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\mathrm{vec}\left(\tilde{B}\left(W_t\right)\right)\right\| \qquad (42)
\end{aligned}
$$

For the second derivative we have:

$$\begin{aligned}
&\frac{\partial}{\partial b'}\text{vec}\left(\frac{\partial}{\partial b'}\text{vec}\left(\left\{\prod_{l=r}^{s}\tilde{B}\left(W_{t-l}\right)\right\}\right)\right) \quad = \\
&= \sum_{l=r}^{s}\frac{\partial}{\partial b'}\text{vec}\left(\left\{\prod_{m'=s-l+1}^{s}\tilde{B}\left(W_{t-m'}\right)\right\}'\otimes\left\{\prod_{m=r}^{s-l-1}\tilde{B}\left(W_{t-m}\right)\right\}\frac{\partial}{\partial b'}\text{vec}\left(\tilde{B}\left(W_{t-(s-l)}\right)\right)\right) \\
&\quad = \sum_{l=r}^{s}\left[\left[\frac{\partial}{\partial b'}\text{vec}\left(\tilde{B}\left(W_{t-(s-l)}\right)\right)\right]'\otimes I_{n(n+1)/2}\right]\times \\
&\quad \times\frac{\partial}{\partial b'}\text{vec}\left(\left\{\prod_{m'=s-l+1}^{s}\tilde{B}\left(W_{t-m'}\right)\right\}'\otimes\left\{\prod_{m=r}^{s-l-1}\tilde{B}\left(W_{t-m}\right)\right\}\right) \\
&+\sum_{l=r}^{s}\left\{I_{3n}\otimes\left\{\left\{\prod_{m'=s-l+1}^{s}\tilde{B}\left(W_{t-m'}\right)\right\}'\otimes\left\{\prod_{m=r}^{s-l-1}\tilde{B}\left(W_{t-m}\right)\right\}\right\}\right\}\times \\
&\qquad\qquad \times\ \frac{\partial}{\partial b'}\text{vec}\left(\frac{\partial}{\partial b'}\text{vec}\left(\tilde{B}\left(W_{t-(s-l)}\right)\right)\right)
\end{aligned}$$

and finally

$$\begin{aligned}
&\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\text{vec}\left(\frac{\partial}{\partial b'}\text{vec}\left(\left\{\prod_{l=r}^{s}\tilde{B}\left(W_{t-l}\right)\right\}\right)\right)\right\| \\
&\leq\ (s-r+1)^2\,\bar{\varrho}_B^{s-r}K_{1,n(n+1)/2}\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\text{vec}\left(\tilde{B}\left(W_t\right)\right)\right\|^2 \\
&+\ (s-r+1)^2\,\bar{\varrho}_B^{s-r}\sup_{\theta\in\Theta,t\in(-\infty,\infty)}\left\|\frac{\partial}{\partial b'}\text{vec}\left(\frac{\partial}{\partial b'}\text{vec}\left(\tilde{B}\left(W_t\right)\right)\right)\right\| \qquad (43)
\end{aligned}$$

For part (i)

$$\begin{aligned}
\xi_t &= \tilde{C}+\tilde{A}\left(W_t\right)y_{t-1}+\sum_{\tau=1}^{\infty}\left\{\prod_{l=1}^{\tau-1}\tilde{B}\left(W_{t-l}\right)\right\}\left(\tilde{C}+\tilde{A}\left(W_{t-\tau}\right)X_{t-\tau-1}\right) \\
\frac{\partial\xi_t}{\partial a'} &= X_{t-1}'\otimes I_{n(n+1)/2}\frac{\partial\text{vec}\left(\tilde{A}\left(W_t\right)\right)}{\partial a'} \\
&+ \sum_{\tau=1}^{\infty}\left\{\prod_{l=1}^{\tau-1}\tilde{B}\left(W_{t-l}\right)\right\}\left(\left(X_{t-\tau-1}'\otimes I_{n(n+1)/2}\right)\frac{\partial\text{vec}\left(\tilde{A}\left(W_{t-\tau}\right)\right)}{\partial a'}\right) \qquad (44) \\
\frac{\partial\xi_t}{\partial b'} &= \sum_{\tau=1}^{\infty}\left[\left(\tilde{C}+\tilde{A}\left(W_{t-\tau}\right)X_{t-\tau-1}\right)'\otimes I_{n(n+1)/2}\right]\frac{\partial}{\partial b'}\text{vec}\left(\left\{\prod_{l=1}^{\tau-1}\tilde{B}\left(W_{t-l}\right)\right\}\right) \qquad (45)
\end{aligned}$$

By using equation (44), equation (45), 39,36 and 41 and applying the Hölder and Minkowski inequalities, we get

$$\begin{aligned}
\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\partial\tilde{\xi}_t}{\partial a'}\right\|^3\right] &\leq \left\{\sum_{\tau=0}^{\infty}\varrho_B^{\tau-1}\left[\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\left(\left(X'_{t-\tau-1}\otimes I_{n(n+1)/2}\right)\frac{\partial \mathrm{vec}\left(\tilde{A}\left(W_{t-\tau}\right)\right)}{\partial a'}\right)\right\|\right]^3\right]^{1/3}\right\}^3 \\
&\leq \left\{\sum_{\tau=0}^{\infty}\varrho_B^{\tau-1}\left[\mathbb{E}\left[\|u_t\|^6\right]^{1/6}K_{2,n}K_{1,n}\rho_M 2K_{3,3n}K_a\right]\right\}^3 < \infty \\
\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\partial\tilde{\xi}_t}{\partial b'}\right\|^3\right] &\leq \left\{\sum_{\tau=0}^{\infty}(\tau-1)\varrho_B^{\tau-2}\left[K_c^2+\mathbb{E}\left[\|u_t\|^6\right]^{1/6}K_{2,n}K_{1,n}\rho_M 2K_{3,3n}K_b\right]\right\}^3 < \infty
\end{aligned}$$

.

For part (ii)

$$\begin{aligned}
\frac{\partial^2\xi_t}{\partial a'\partial a'} &= \left(X'_{t-1}\otimes I_{n(n+1)/2}\right)\frac{\partial}{\partial a'}\left(\mathrm{vec}\left(\frac{\partial \mathrm{vec}\left(\tilde{A}\left(W_t\right)\right)}{\partial a'}\right)\right) \\
&+ \sum_{\tau=1}^{\infty}\left\{\prod_{l=1}^{\tau-1}\tilde{B}\left(W_{t-l}\right)\right\}\left(\left(X'_{t-\tau-1}\otimes I_{n(n+1)/2}\right)\frac{\partial}{\partial a'}\left(\mathrm{vec}\left(\frac{\partial \mathrm{vec}\left(\tilde{A}\left(W_t\right)\right)}{\partial a'}\right)\right)\right) \quad (46)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2\xi_t}{\partial a'\partial b'} &= \sum_{\tau=1}^{\infty}\left[\left(\left(X'_{t-\tau-1}\otimes I_{n(n+1)/2}\right)\frac{\partial \mathrm{vec}\left(\tilde{A}\left(W_{t-\tau}\right)\right)}{\partial a'}\right)'\otimes I_{n(n+1)}\right]\times \\
&\times \left\{\sum_{j=1}^{\tau-1}\left\{\prod_{l=1}^{\tau-1-j}\tilde{B}\left(W_{t-l}\right)'\right\}\otimes\left\{\prod_{l=\tau-1-(j-1)}^{\tau-1}\tilde{B}\left(W_{t-l}\right)\right\}\frac{\partial}{\partial b'}\mathrm{vec}\left(\tilde{B}\left(W_{t-\tau+j}\right)\right)\right\} \quad (47)
\end{aligned}$$

$$\frac{\partial^2\xi_t}{\partial b'\partial b'} = \sum_{\tau=1}^{\infty}\left[\left(\tilde{C}+\tilde{A}\left(W_{t-\tau}\right)X_{t-\tau-1}\right)'\otimes I_{n(n+1)/2}\right]\frac{\partial}{\partial b'}\mathrm{vec}\left(\frac{\partial}{\partial b'}\mathrm{vec}\left(\left\{\prod_{l=r}^{s}\tilde{B}\left(W_{t-l}\right)\right\}\right)\right). \quad (48)$$

Then, from equation (46),equation (47), equation (48), 40,37 and 43 and the Hölder and Minkowski inequalities, we have

$$
\begin{aligned}
\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\partial^2\tilde{\xi}_t}{\partial a'\partial a'}\right\|^3\right] &\leq \left\{\sum_{\tau=0}^{\infty}\bar{\varrho}_B^{\tau-1}K_{2,n}K_{1,n}\rho_M 2K_{3,3n}K_{1,3n}\left[\mathbb{E}\left[\|u_t\|^6\right]^{1/6}\right]\right\}^3 < \infty \\
\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\partial^2\tilde{\xi}_t}{\partial b'\partial a'}\right\|^3\right] &\leq \left\{\sum_{\tau=0}^{\infty}(\tau-1)\,\bar{\varrho}_B^{\tau-2}K_{2,n}^2K_{1,n}^2\rho_M^2 2K_{3,3n}^2K_aK_b\left[\mathbb{E}\left[\|u_t\|^6\right]^{1/6}\right]\right\}^3 < \infty \\
\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\partial^2\tilde{\xi}_t}{\partial b'\partial b'}\right\|^3\right] &\leq \left\{\left[K_c^2+\mathbb{E}\left[\|u_t\|^6\right]^{1/6}K_{2,n}K_{1,n}\rho_M 2K_{3,3n}K_b\right]\times\right. \\
&\quad\times\left[\sum_{\tau=0}^{\infty}(\tau-1)^2\,\bar{\varrho}_B^{\tau-3}K_{2,n}^2K_{1,n}^2\rho_M^2 2K_{3,3n}^2K_b^2\right. \\
&\quad+\left.\left.\sum_{\tau=0}^{\infty}(\tau-1)^2\,\bar{\varrho}_B^{\tau-2}K_{2,n}K_{1,n}\rho_M 2K_{3,3n}K_{1,3n}\right]\right\}^3 < \infty
\end{aligned}
$$

□

B.3.1 Proof of theorem 2 on page 17

Proof. The proof is analogous to the proof of Theorem 2 in Hafner and Preminger (2009) if we substitute in part (ii) of the proof of their Lemma 2 their bound of $\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\Sigma_t-\tilde{\Sigma}_t\right\|\right]$ (equation 17 in their paper) with our bound in Lemma 5 . □

B.3.2 Proof of theorem 3 on page 18

Proof. The proof is again almost identical to the proof of the equivalent Theorem 3 in Hafner and Preminger (2009) if we replace in part (i) and (ii) of their proof of Lemma 3 the bound on the derivatives of $\tilde{\xi}_t$ with respect to parameters (equations 28,29 and 33 in their paper) with the bound we derive in Lemma 6. □

C Multistep Forecast

Since analytical expressions for the multistep volatility forecast are not available in closed form, the most common way to obtain a robust multistep forecast involves the use of bootstrapping techniques, Andersen et al. (2006). In particular, we proceed with the following methodology. Consider an estimation window $t \in [1,\ldots,T]$ from which estimates for $\hat{C}$, $\hat{A}(W_t)$, $\hat{B}(W_t)$ and the series of filtered conditional covariances $\hat{\Sigma}_t$ can be obtained, and that we are interested in computing the path of the forecasted covariance matrix from, assume, $T+1$ to $T+h$. The first step computes the $n\times 1$ vector of filtered innovations (standardized residuals) $\hat{\epsilon}_t$ for each time in the estimation period $t\in[1,\ldots,T]$. We obtain the innovations by multiplying the vector u_t and the inverse of the Cholesky decomposition of the estimated conditional covariance $\hat{\Sigma}_t$:

$$\hat{\epsilon}_t = \hat{\Sigma}_t^{-\frac{1}{2}}u_t, \qquad t\in[1,\ldots,T]\,.$$

The second step builds N_B bootstrap samples of length h from the $n \times T$ matrix of filtered innovations $[\hat{\epsilon}_1, \ldots, \hat{\epsilon}_T]$. We use a bootstrap procedure that preserves as much as possible the residual longitudinal dependence in the data, so that it is robust to misspecification in the model. We use a circular block bootstrap, Politis and Romano (1992), with automatic block length selection, Politis and White (2004).[14] In this way, we obtain the bootstrapped innovations $\tilde{\epsilon}^{[b]}_{T+l}$ with $b \in [1, \ldots, N_B]$ and $l \in [1, \ldots, h]$. In turn, these allow computing the bootstrapped mean innovations u_{T+l} and the bootstrapped covariances for each l and b:

$$\begin{aligned}
\tilde{u}^{[b]}_{T+l} &= \hat{\Sigma}^{\frac{1}{2}}_{T+l} \tilde{\epsilon}^{[b]}_{T+l} \\
\tilde{\Sigma}^{[b]}_{T+l} &= \hat{C}\hat{C}' + \hat{A}\left(W_{T+l}\right) \tilde{u}^{[b]}_{T+l} \tilde{u}^{[b\prime]}_{T+l} \hat{A}\left(W_{T+l}\right)' + \hat{B}\left(W_{T+l}\right) \tilde{\Sigma}^{[b]}_{T+l-1} \hat{B}\left(W_{T+l}\right)'
\end{aligned}$$

Finally, we set the covariance matrix forecast equal to the average across the N_B paths:

$$\hat{\Sigma}^{F}_{T+l} = \frac{1}{N_B} \sum_{b=1}^{N_B} \tilde{\Sigma}^{[b]}_{T+l}.$$

Note that even quantiles could be considered in place of the mean, thus focusing on low/high states for volatility forecasts and that the previous approach is valid for any parametrization of the covariance dynamics, thus including the case of the Spatial-BEKK model. However, we stress that when the parameter matrices are function of a time-varying network, the forecast are conditional to the last observed network. Alternatively, if there exists a model to forecast the network evolution, this can be integrated with the previous covariance forecast approach, allowing the computation of forecasts accounting for the network variability.

[14] In particular, we apply the procedure for selecting the block length to each univariate series and then take the maximum of the obtained lengths.

Table 1: Decomposition of $[\Sigma_t]_{i,j}$

	shock response (ARCH)	persistence (GARCH)
Constant	$[CC']_{i,j}$	
direct	$v^0_{i,t-1}v^0_{j,t-1}$	$\left[\Omega^{0,0}_{t-1}\right]_{i,j}$
indirect	$v^1_{L,i,t-1}v^1_{L,j,t-1} + v^1_{R,i,t-1}v^1_{R,j,t-1}$ $+v^1_{L,i,t-1}v^1_{R,j,t-1} + v^1_{R,i,t-1}v^1_{L,j,t-1}$	$\left[\Omega^{1,1}_{L,L,t-1}\right]_{i,j} + \left[\Omega^{1,1}_{R,R,t-1}\right]_{i,j}$ $+ \left[\Omega^{1,1}_{L,R,t-1}\right]_{i,j} + \left[\Omega^{1,1}_{R,L,t-1}\right]_{i,j}$
mixed	$v^1_{L,i,t-1}v^0_{j,t-1} + v^0_{i,t-1}v^1_{L,j,t-1}$ $+ v^1_{R,i,t-1}v^0_{j,t-1} + v^0_{i,t-1}v^1_{R,j,t-1}$	$\left[\Omega^{1,0}_{L,t-1}\right]_{i,j} + \left[\Omega^{0,1}_{L,t-1}\right]_{i,j}$ $+ \left[\Omega^{1,0}_{R,t-1}\right]_{i,j} + \left[\Omega^{0,1}_{R,t-1}\right]_{i,j}$

Table 2: Estimated Relevant Parameters of SB-BEKK on Daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. Standard deviation in parenthesis. * parameters significant at the 10% level. ** parameters significant at the 5% level. *** parameters significant at the 1% level

	$a_{1,L}$	$a_{1,R}$	$b_{1,L}$	$b_{1,R}$	a_0	b_0
DE	-0.056(0.11)	-0.022(0.10)	-0.088(0.03)***	-0.014(0.03)	0.258(0.03)***	0.965(0.01)***
IT	0.089(0.09)	0.849(0.27)***	0.001(0.03)	-0.058(0.07)	0.192(0.02)***	0.972(0.00)***
FR	-0.636(0.19)***	0.549(0.22)**	0.209(0.05)***	-0.292(0.10)***	0.376(0.03)***	0.887(0.01)***
IE	-0.388(0.11)***	-1.719(0.53)***	0.149(0.03)***	0.757(0.16)***	0.352(0.04)***	0.929(0.01)***
ES	0.618(0.18)***	-0.568(0.22)***	-0.138(0.04)***	0.076(0.04)*	0.107(0.02)***	0.987(0.00)***
PT	0.206(0.12)*	-0.742(0.10)***	0.086(0.04)**	0.252(0.01)***	0.290(0.03)***	0.912(0.01)***

Table 3: The percentage of cumulative variance decomposition was obtained from data on daily changes in the five-year EMU sovereign CDS spreads from 9/10/2008 to 30/12/2016.

	Constant	Mixed	Indirect	Direct
σ^2_{DE} %	0.0170	-0.0104	0.0021	0.9913
σ^2_{IT} %	0.0170	-0.0034	0.0019	0.9845
σ^2_{FR} %	0.0563	0.0137	0.0104	0.9196
σ^2_{IE} %	0.0167	0.0063	0.0037	0.9732
σ^2_{ES} %	0.0209	-0.0282	0.0192	0.9881
σ^2_{PT} %	0.0344	0.0436	0.0059	0.9161

Table 4: The percentage of cumulative marginal spillover contribution (MSC) was obtained from data on daily changes in the five-year EMU sovereign CDS spreads from 9/10/2008 to 30/12/2016.

	Total	Costant	Mixed	Indirect	Direct
MSC_{DE} %	0.0893	0.0015	0.0011	0.0000	0.0866
MSC_{IT} %	0.0984	0.0022	0.0063	0.0005	0.0894
MSC_{FR} %	0.1442	0.0058	-0.0011	0.0010	0.1385
MSC_{IE} %	0.1413	0.0023	0.0071	-0.0005	0.1323
MSC_{ES} %	0.2566	0.0059	0.0042	0.0002	0.2463
MSC_{PT} %	0.2692	0.0037	0.0019	-0.0002	0.2638

Table 5: Parameter-constancy test for the different optimal networks

	No Constraints	Redistribution Constraint	Frobenius Constraint
Wald Stat	27.1	8.2	7.4
P-value	0.3012	0.9988	0.9995

Table 6: Investment Needed to Reach Target Exposures (Billions of USD), obtained from daily changes in five-year EMU sovereign CDS spreads.

	Optimal Network with Frobenius Constraint (Delta wrt true) Billions of USD					
Location	DE	IT	FR	IE	ES	PT
Reporting						
DE		-41.2	32.6	-5.6	41.5	109.7
IT	8.1		-6.5	-7.7	-11.4	15.3
FR	44.9	-121.9		-10.8	9.0	85.3
IE	-0.8	-1.3	-0.7		-1.9	2.4
ES	-10.9	-56.3	-19.8	-2.8		24.1
PT	-0.7	-3.6	-1.7	-0.4	-0.9	

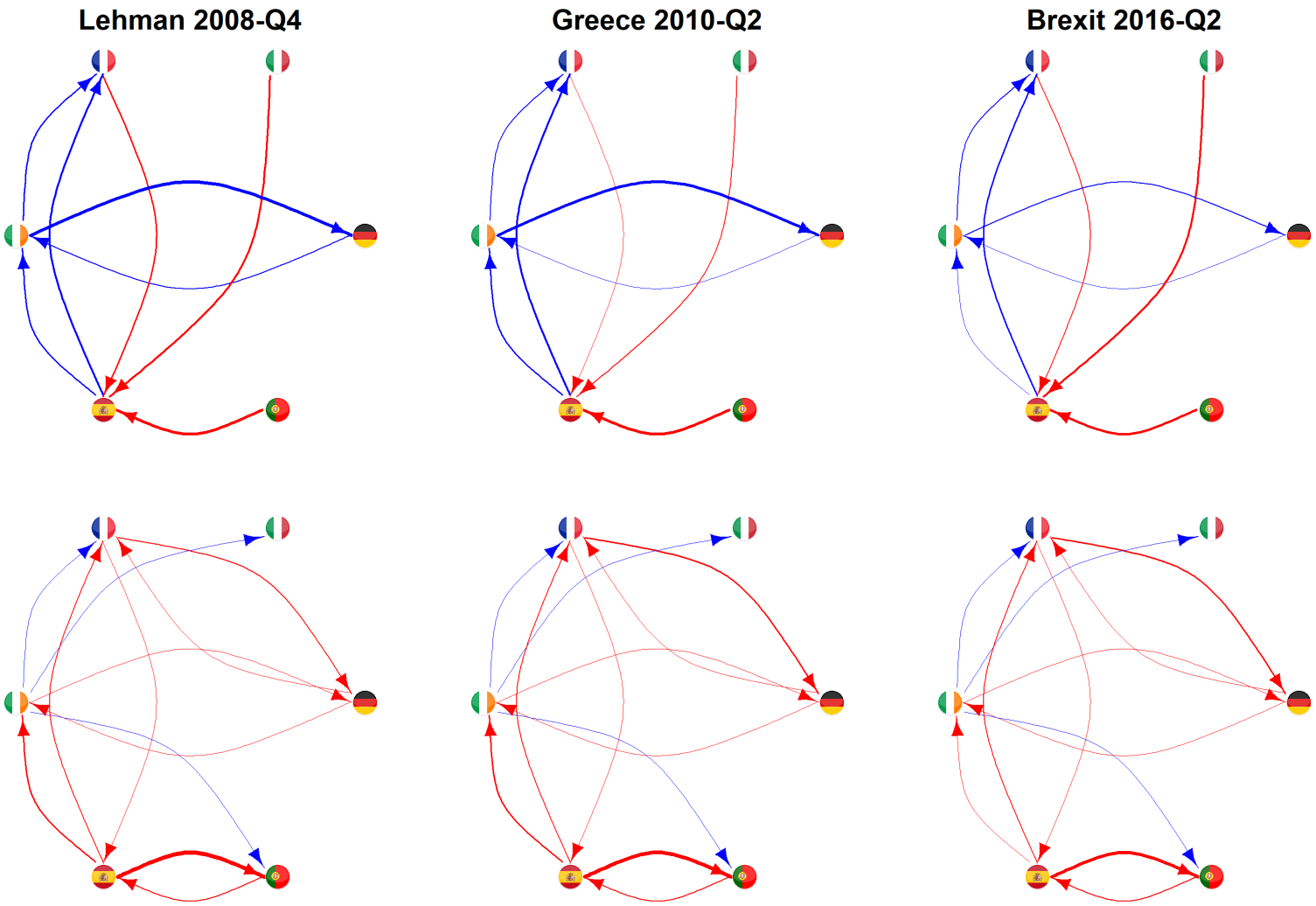

Figure 1: Selected graphical representations of the shock-response networks $\mathcal{W}_t^A$ in the first row and the persistence networks $\mathcal{W}_t^B$ in the second row were all obtained from daily changes in the five-year EMU sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows go from counterparty to reporting, and positive weights are in red while negative are in blue. Line thickness is proportional to coefficent size. We only report edges that are significant at the 5% level, taking into consideration a familywise Bonferroni correction.

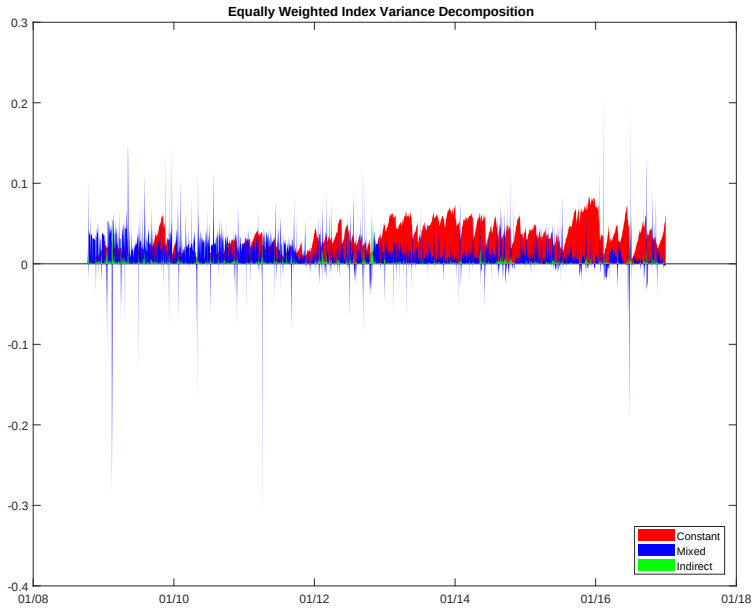

Figure 2: The relative variance decomposition of the equally weighted index was obtained from data on daily changes in the five-year EMU sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted.

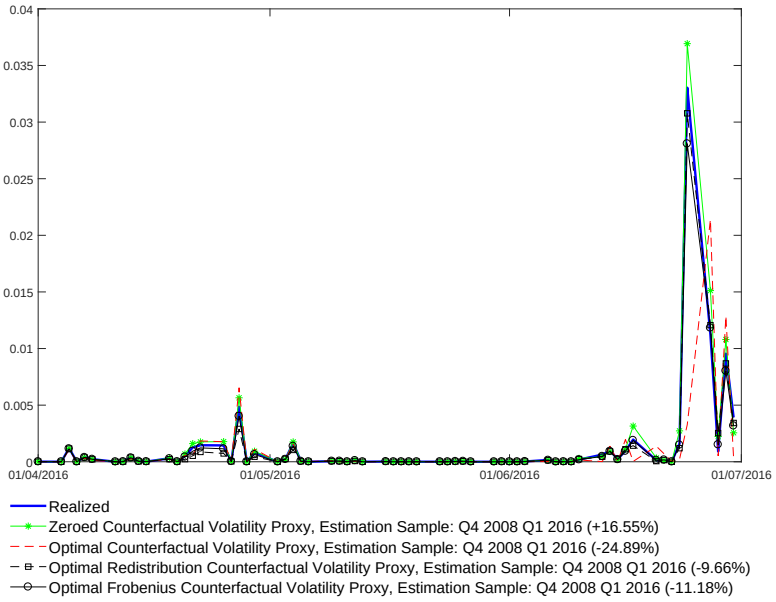

Figure 3: The average covariance proxy during the Brexit referendum of Q2 2016 was obtained from data on daily changes in five-year EMU sovereign CDS spreads. The percentage of cumulative proxy change with regard to the realized one is reported in parenthesis

Appendix to *Networks in risk spillovers: A multivariate GARCH perspective*

June 20, 2018

S Supplementary Tables and Figures

Table S.1: Daily Changes in Five-Year Sovereign CDS spreads from 9/10/2008 to 30/12/2016. Summary Statistics

	Germany	Italy	France	Ireland	Spain	Portugal
mean	-0.0003	0.0003	-0.0001	-0.0002	-0.0001	0.0006
s.d.	0.063	0.046	0.049	0.036	0.047	0.045
Skewness	0.16	-0.18	0.39	0.10	0.01	-0.68
Kurtosis	17.90	13.50	9.92	15.33	11.46	21.33
min	-0.51	-0.45	-0.27	-0.33	-0.42	-0.59
max	0.59	0.34	0.36	0.25	0.34	0.31
median	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Correlations

	Italy	France	Ireland	Spain	Portugal
Germany	0.36	0.36	0.31	0.34	0.31
Italy		0.51	0.62	0.83	0.68
France			0.43	0.50	0.45
Ireland				0.61	0.60
Spain					0.71

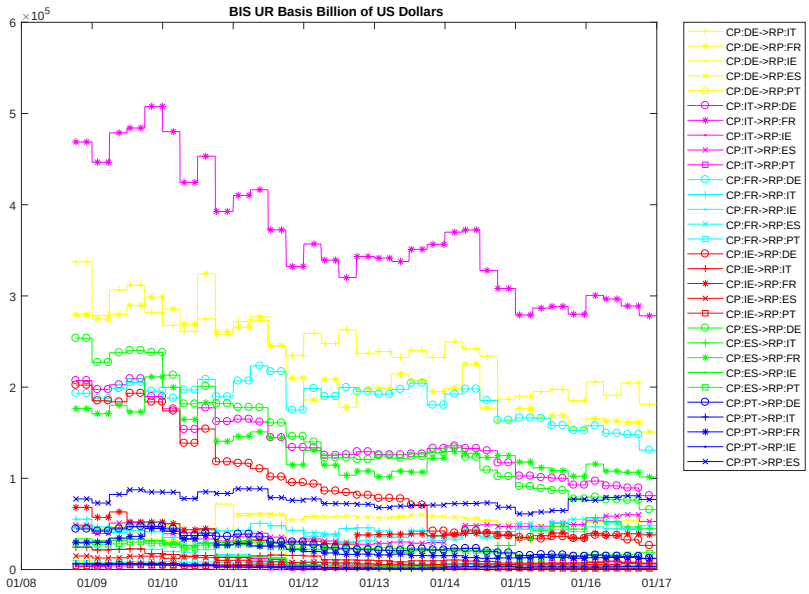

Figure S.1: BIS claims in billions of US Dollars by Counterparty (CP) and Reporting country(RP)

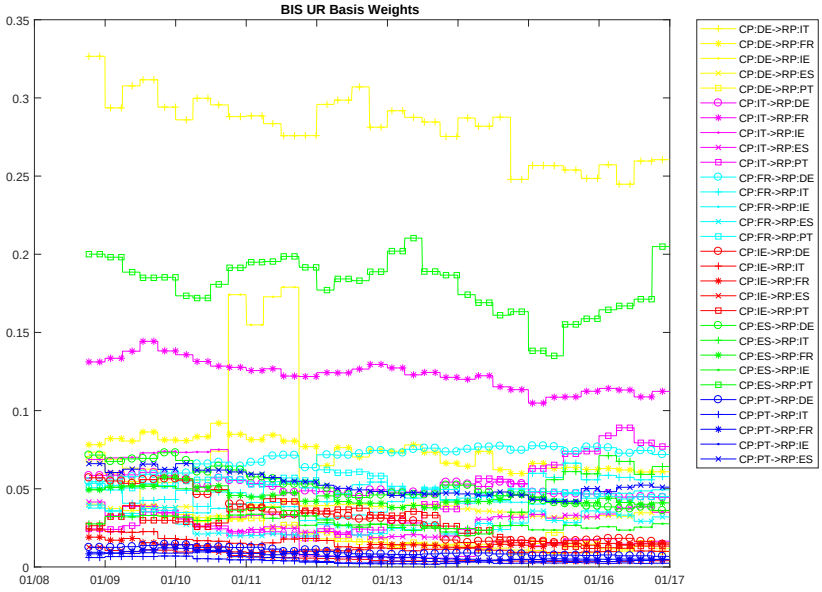

Figure S.2: BIS claims normalized by the worldwide amount declared by reporting country, by Counterparty (CP) and Reporting country (RP)

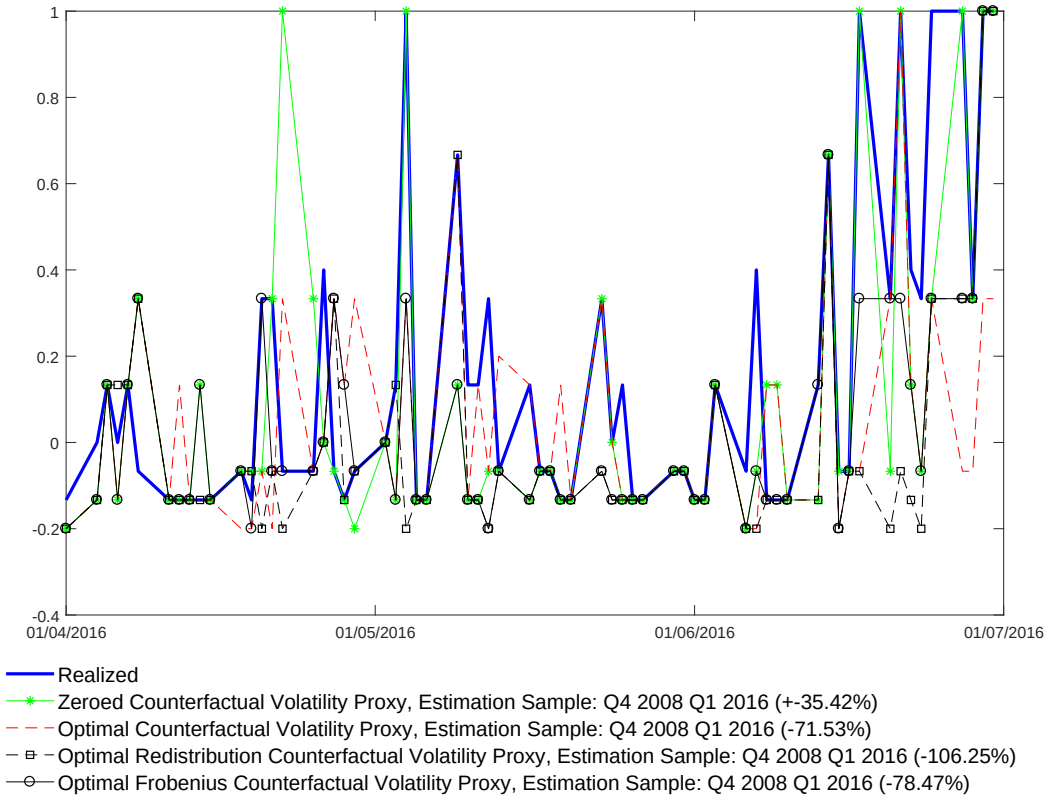

Figure S.3: Average correlation proxy during the Brexit Referendum of Q2 2016, obtained from data on Daily Changes in Five-Year EMU Sovereign CDS spreads. Percentage cumulative proxy change wrt realized is reported in parenthesis

Table S.2: Specification tests for the Spatial-BEKK model on Daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016.

Joint Significance of Network Parameters				
	LR	Wald	Wald L	Wald R
Stat	268	176	97	55
P-value	0.0000	0.0000	0.0000	0.0000
	Wald $a_{1,L}$	Wald $a_{1,R}$	Wald $b_{1,L}$	Wald $b_{1,R}$
Stat	38	27	80	32
P-value	0.0000	0.0002	0.0000	0.0000

Table S.3: P-values of off-diagonal element of the sequences $A(W_t)$ and $B(W_t)$ for Daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016.

$A(W_t)$							$B(W_t)$						
	DE	IT	FR	IE	ES	PT		DE	IT	FR	IE	ES	PT
DE		0.007	0.045	0.001	0.008	0.007	DE		0.042	0.000	0.000	0.797	0.145
IT	0.033		0.004	0.004	0.050	0.036	IT	0.064		0.004	0.000	0.075	0.026
FR	0.002	0.292		0.000	0.000	0.000	FR	0.000	0.003		0.000	0.000	0.000
IE	0.000	0.086	0.476		0.000	0.000	IE	0.000	0.163	0.148		0.000	0.000
ES	0.007	0.000	0.000	0.077		0.783	ES	0.006	0.016	0.000	0.001		0.449
PT	0.253	0.002	0.004	0.005	0.010		PT	0.085	0.707	0.048	0.000	0.000	

Table S.4: Observed and Optimal Exposures, obtained from Data on Daily Changes in Five-Year EMU Sovereign CDS spreads

	True Network of Banking Exposures					
Location	DE	IT	FR	IE	ES	PT
Reporting						
DE		0.045	0.073	0.018	0.037	0.007
IT	0.245		0.057	0.012	0.067	0.004
FR	0.062	0.113		0.014	0.041	0.005
IE	0.010	0.015	0.047		0.024	0.003
ES	0.032	0.038	0.033	0.004		0.051
PT	0.011	0.089	0.043	0.014	0.167	
	Optimal Network (Delta wrt true)					
Location	DE	IT	FR	IE	ES	PT
Reporting						
DE		0.269	0.109	0.018	0.957	0.059
IT	0.736		-0.056	-0.008	-0.062	-0.003
FR	0.723	-0.110		-0.007	0.152	0.515
IE	-0.008	-0.013	-0.044		-0.021	0.002
ES	-0.030	-0.034	-0.015	0.000		0.042
PT	0.274	-0.051	-0.034	-0.003	0.207	
	Optimal Redistribution Network (Delta wrt true)					
Location	DE	IT	FR	IE	ES	PT
Reporting						
DE		-0.044	-0.071	-0.017	-0.032	0.171
IT	0.113		-0.029	-0.009	-0.064	0.003
FR	0.166	-0.113		-0.013	-0.040	0.005
IE	-0.009	-0.014	-0.045		-0.022	0.096
ES	0.024	-0.037	-0.013	0.028		-0.006
PT	0.087	-0.085	-0.040	-0.005	0.035	
	Optimal Frobenius Network (Delta wrt true)					
Location	DE	IT	FR	IE	ES	PT
Reporting						
DE		-0.020	0.016	-0.003	0.020	0.053
IT	0.010		-0.008	-0.010	-0.015	0.020
FR	0.017	-0.047		-0.004	0.003	0.033
IE	-0.009	-0.014	-0.008		-0.020	0.025
ES	-0.007	-0.036	-0.013	-0.002		0.016
PT	-0.008	-0.038	-0.018	-0.004	-0.010	

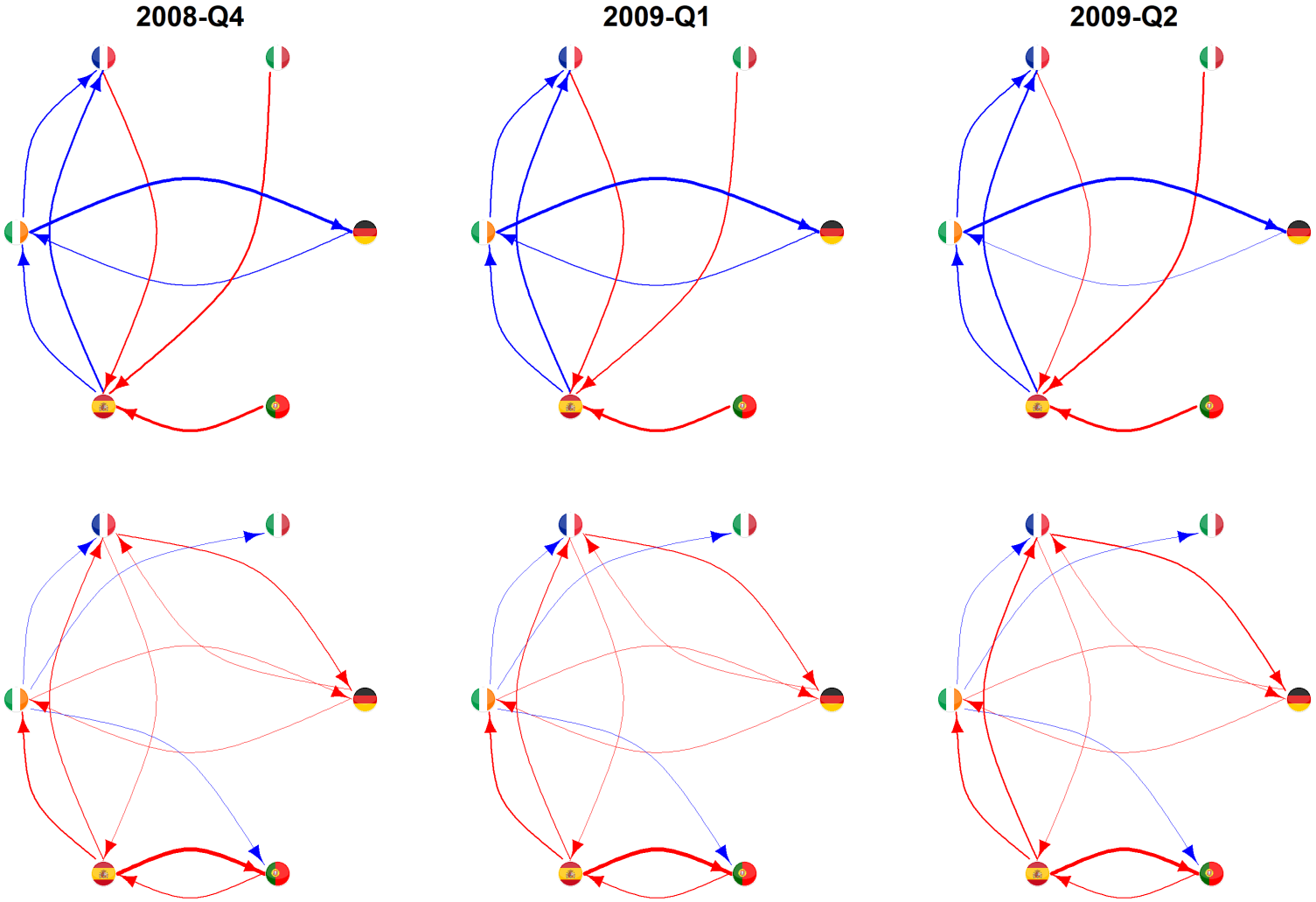

Figure S.4: Selected graphical representation of the shock response networks $\mathcal{W}_t^A$ in the first row, the persistence networks $\mathcal{W}_t^B$ in the second row, all obtained from daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows goes from counterparty to reporting and positive weights are inred while negative are in blue.

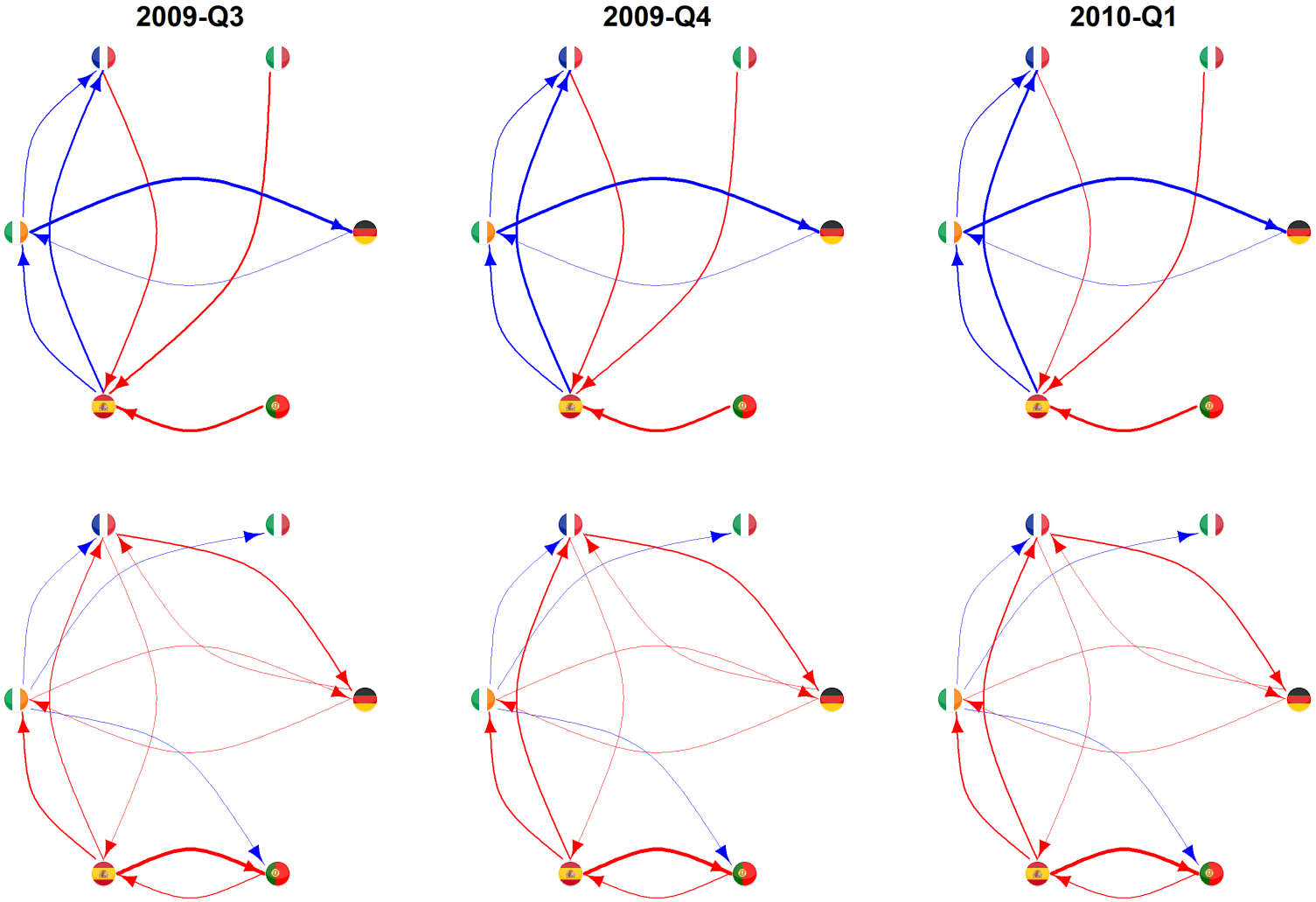

Figure S.5: Selected graphical representation of the shock response networks $\mathcal{W}_t^A$ in the first row, the persistence networks $\mathcal{W}_t^B$ in the second row, all obtained from daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows goes from counterparty to reporting and positive weights are inred while negative are in blue.

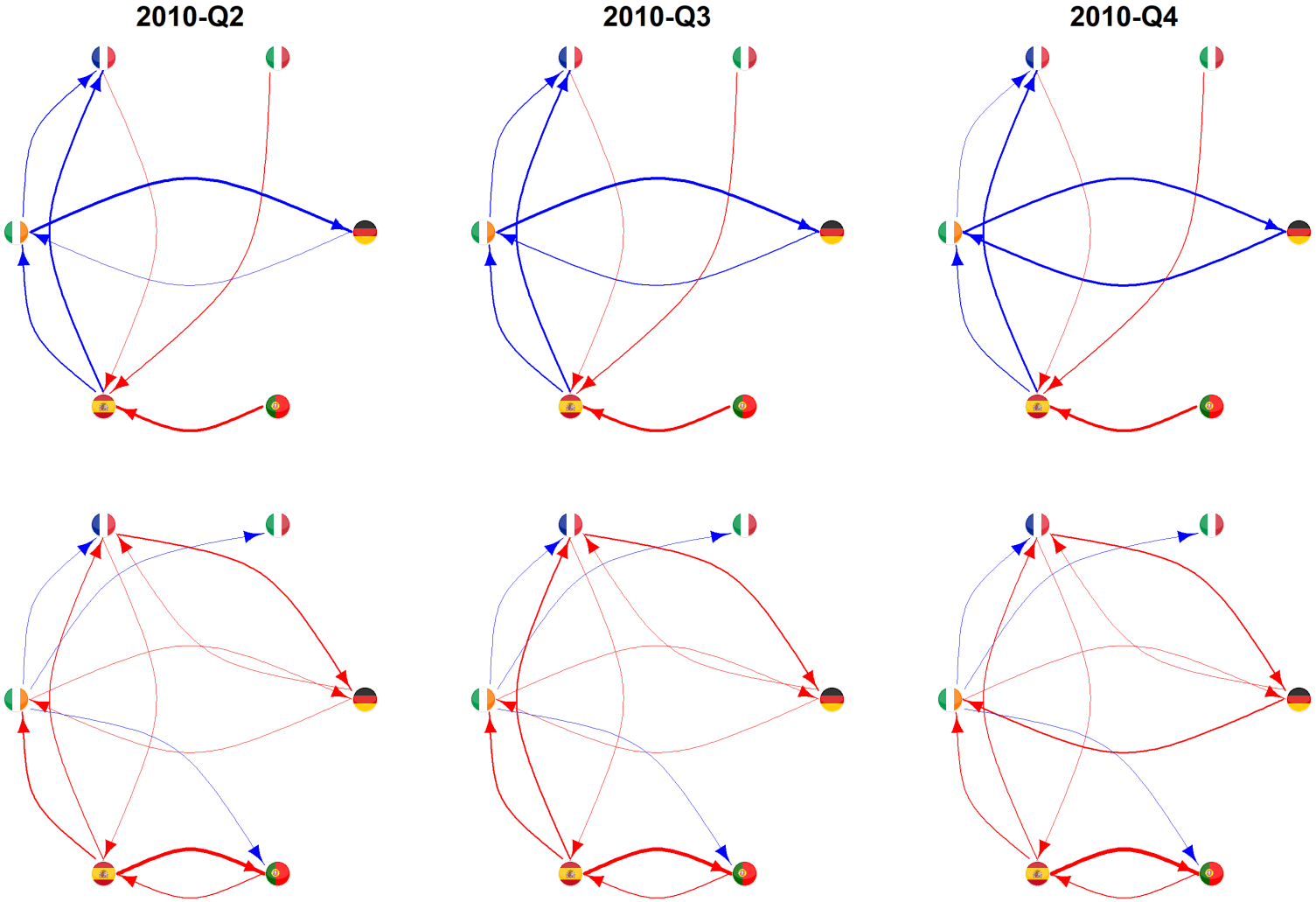

Figure S.6: Selected graphical representation of the shock response networks $\mathcal{W}_t^A$ in the first row, the persistence networks $\mathcal{W}_t^B$ in the second row, all obtained from daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows goes from counterparty to reporting and positive weights are inred while negative are in blue.

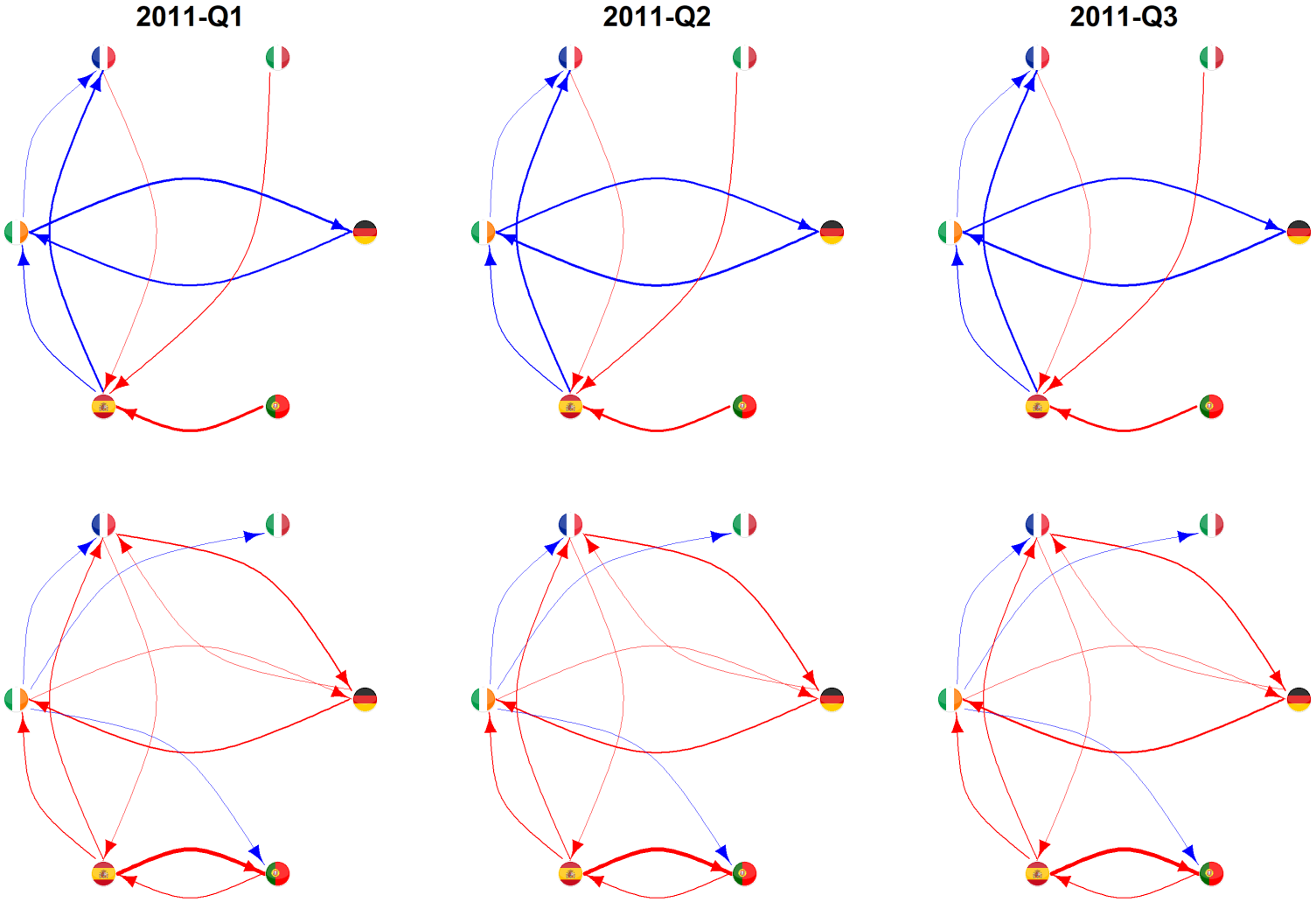

Figure S.7: Selected graphical representation of the shock response networks $\mathcal{W}_t^A$ in the first row, the persistence networks $\mathcal{W}_t^B$ in the second row, all obtained from daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows goes from counterparty to reporting and positive weights are inred while negative are in blue.

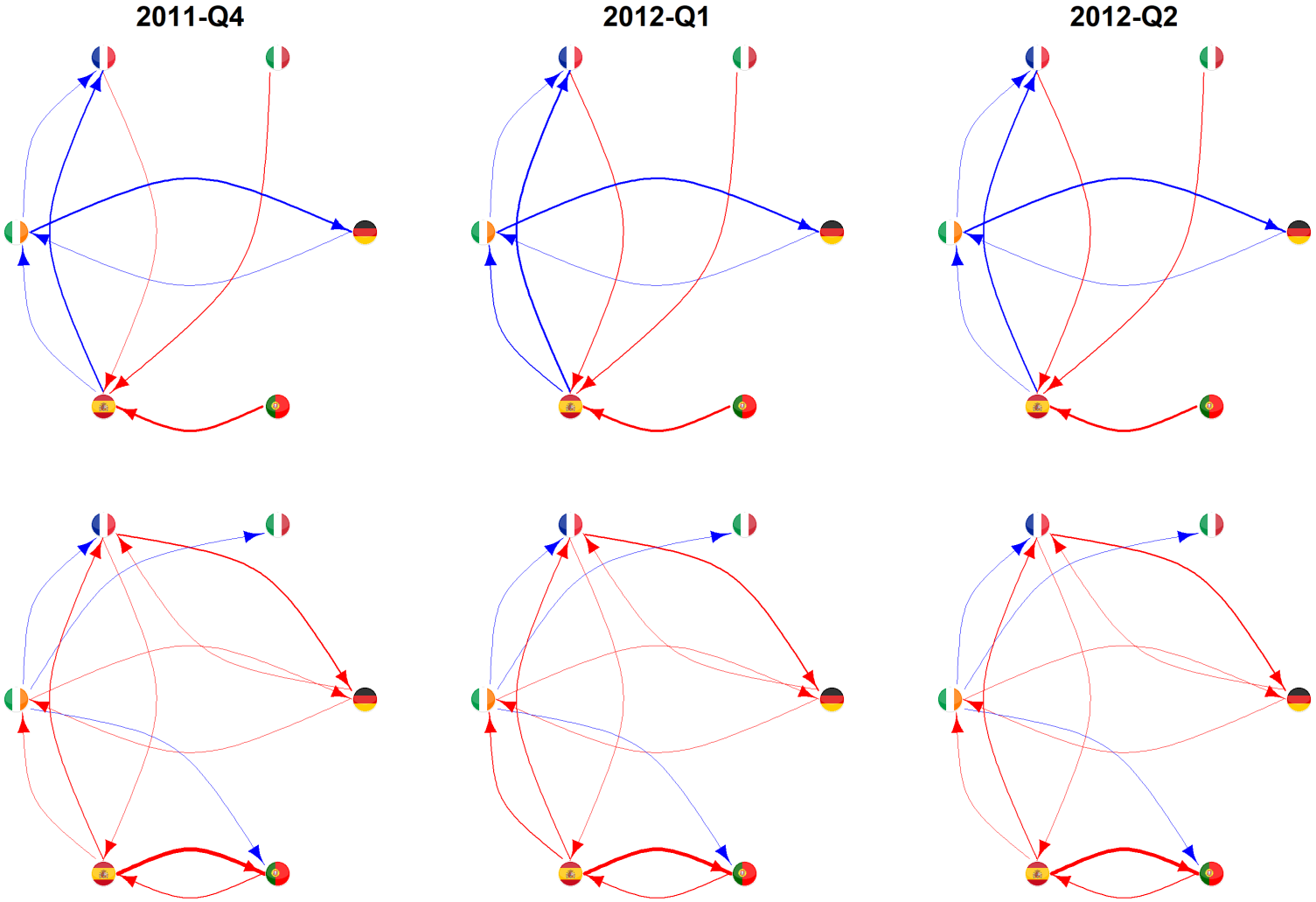

Figure S.8: Selected graphical representation of the shock response networks $\mathcal{W}_t^A$ in the first row, the persistence networks $\mathcal{W}_t^B$ in the second row, all obtained from daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows goes from counterparty to reporting and positive weights are inred while negative are in blue.

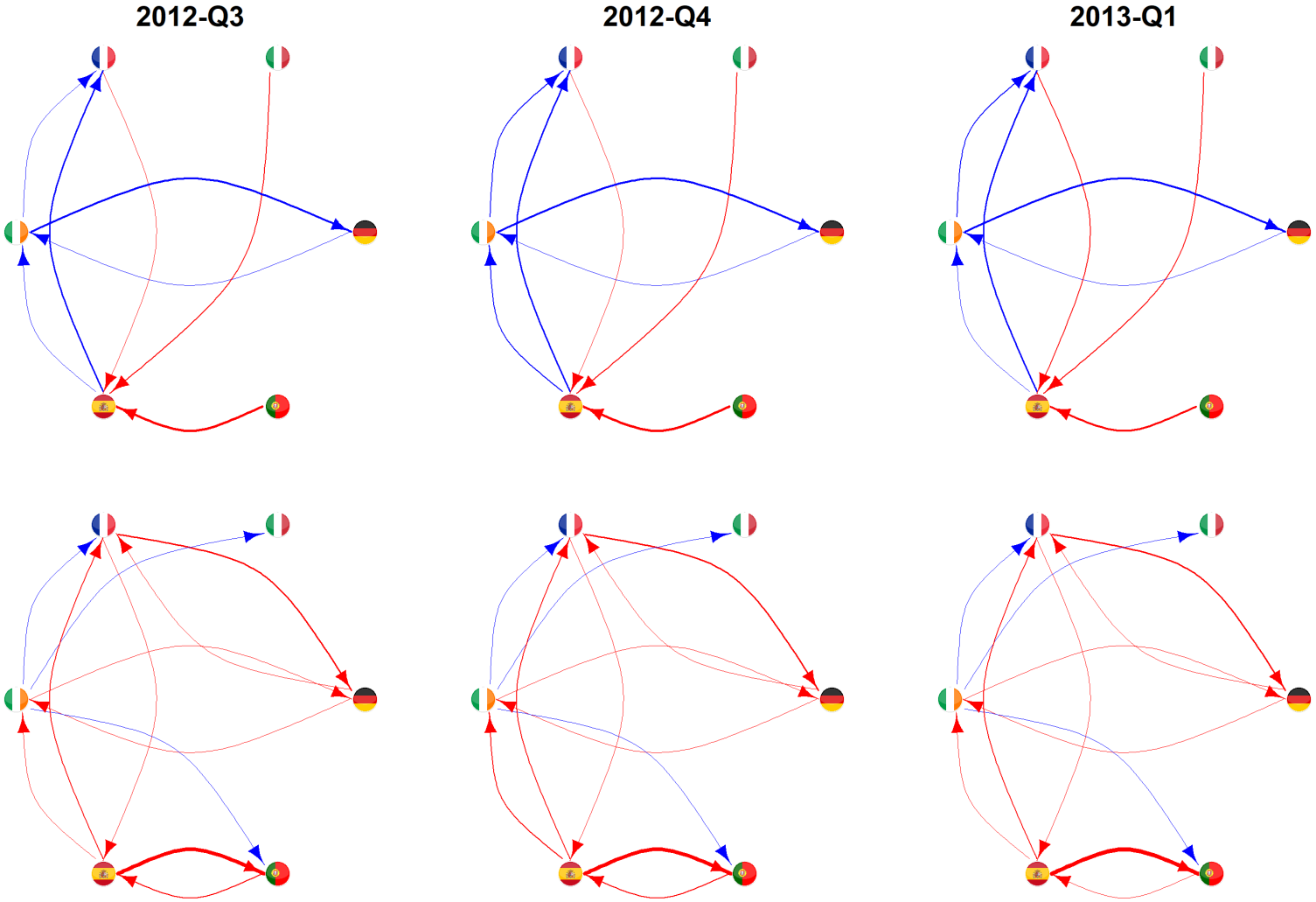

Figure S.9: Selected graphical representation of the shock response networks $\mathcal{W}_t^A$ in the first row, the persistence networks $\mathcal{W}_t^B$ in the second row, all obtained from daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows goes from counterparty to reporting and positive weights are inred while negative are in blue.

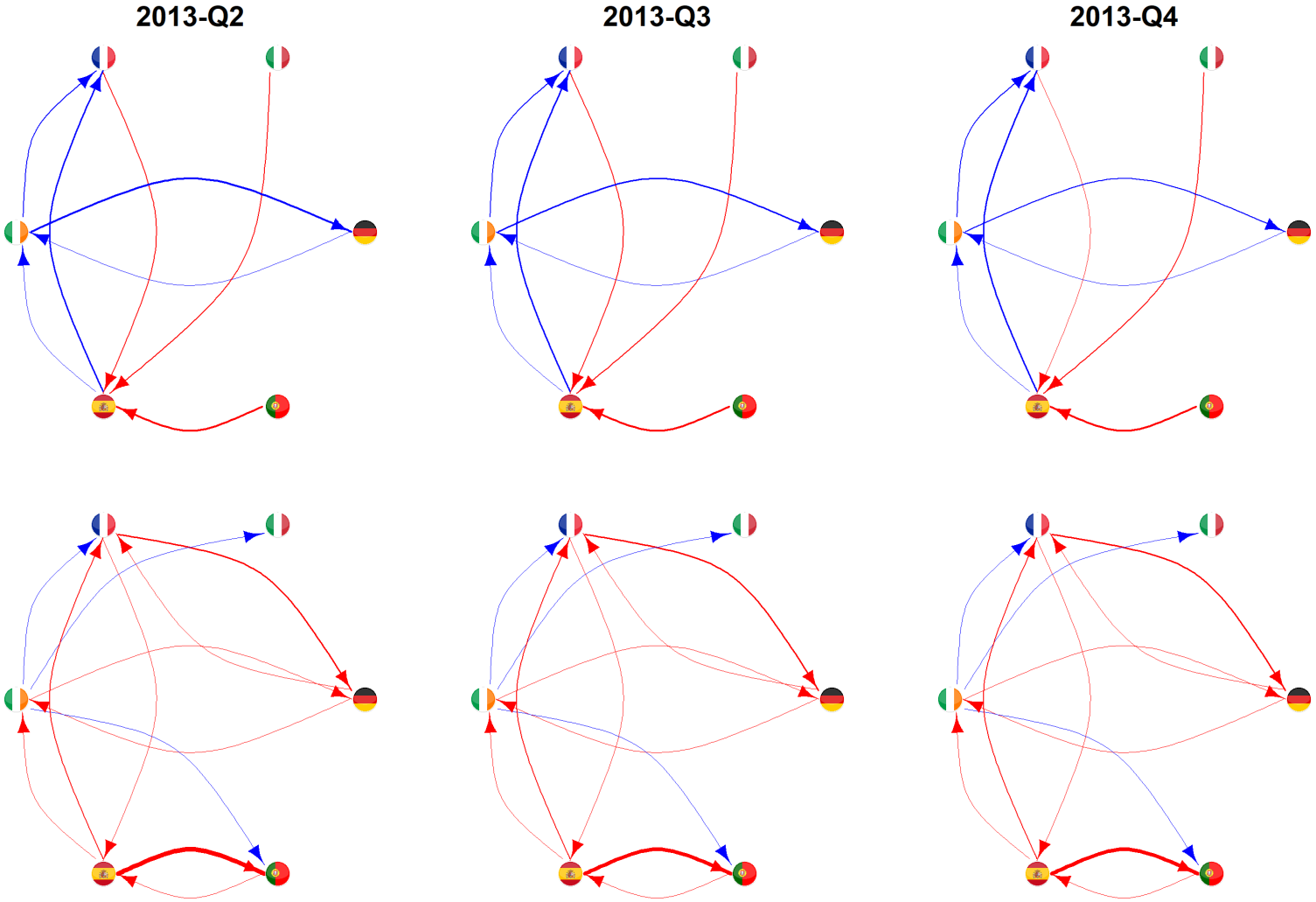

Figure S.10: Selected graphical representation of the shock response networks $\mathcal{W}_t^A$ in the first row, the persistence networks $\mathcal{W}_t^B$ in the second row, all obtained from daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows goes from counterparty to reporting and positive weights are inred while negative are in blue.

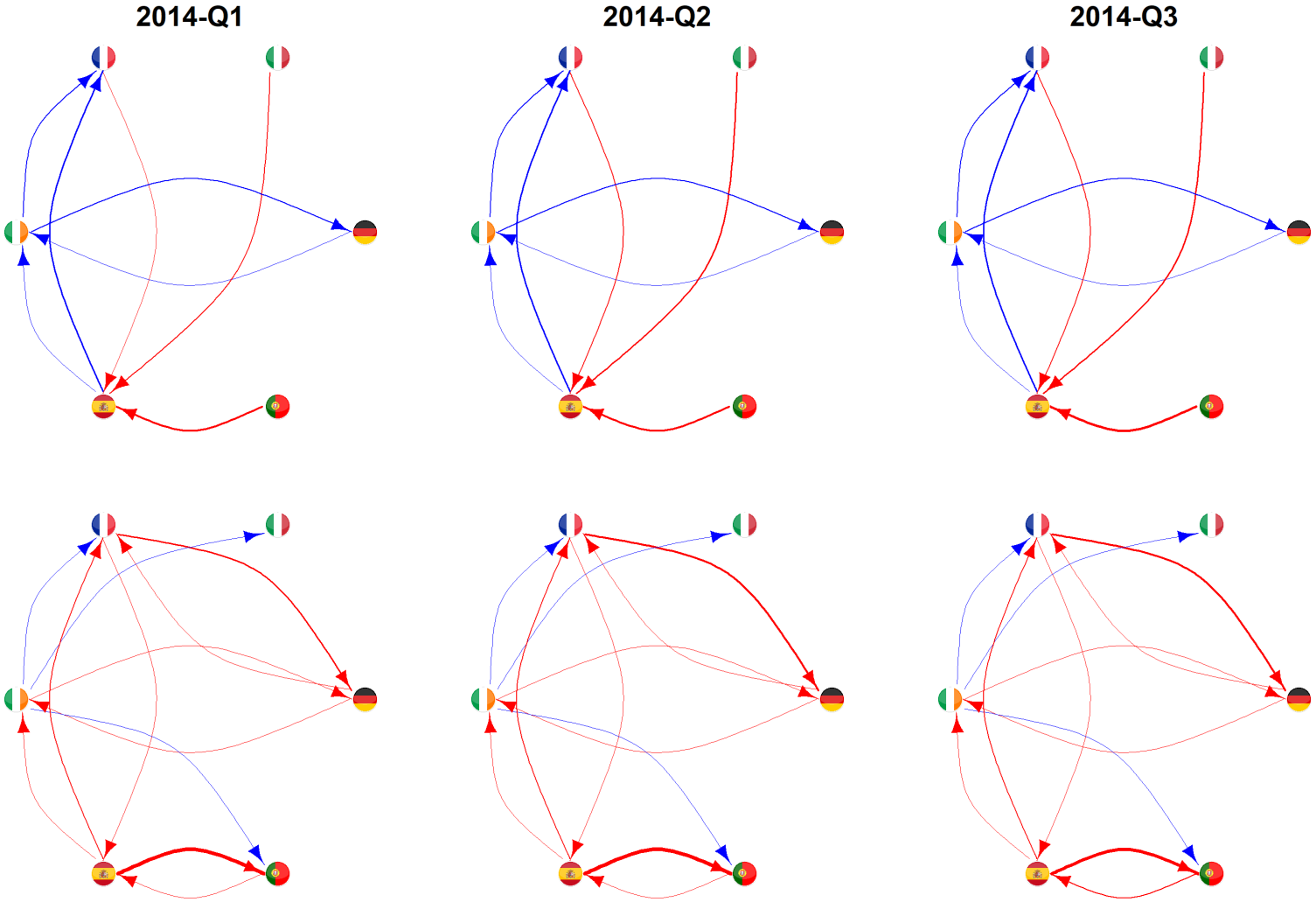

Figure S.11: Selected graphical representation of the shock response networks $\mathcal{W}_t^A$ in the first row, the persistence networks $\mathcal{W}_t^B$ in the second row, all obtained from daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows goes from counterparty to reporting and positive weights are inred while negative are in blue.

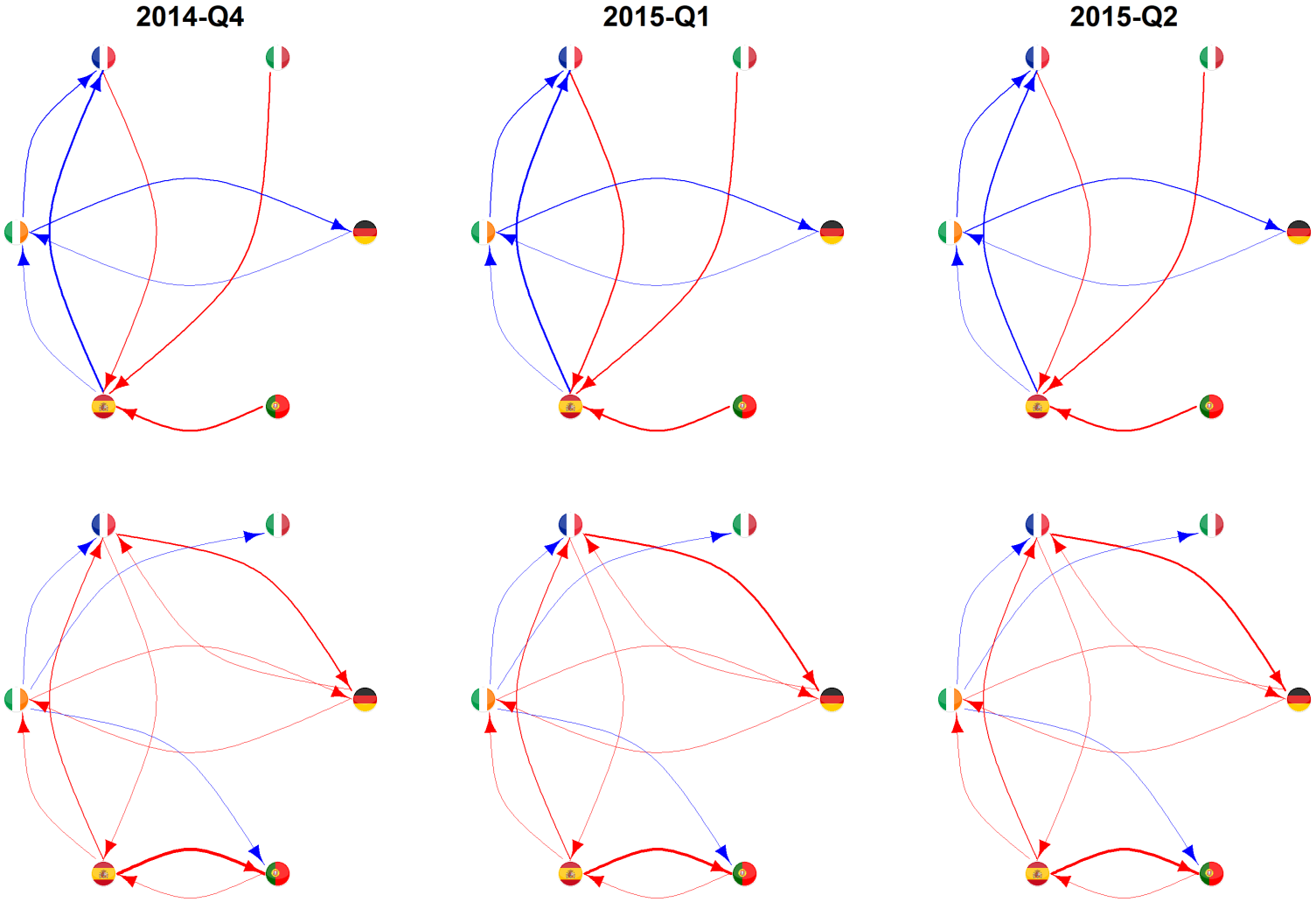

Figure S.12: Selected graphical representation of the shock response networks $\mathcal{W}_t^A$ in the first row, the persistence networks $\mathcal{W}_t^B$ in the second row, all obtained from daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows goes from counterparty to reporting and positive weights are inred while negative are in blue.

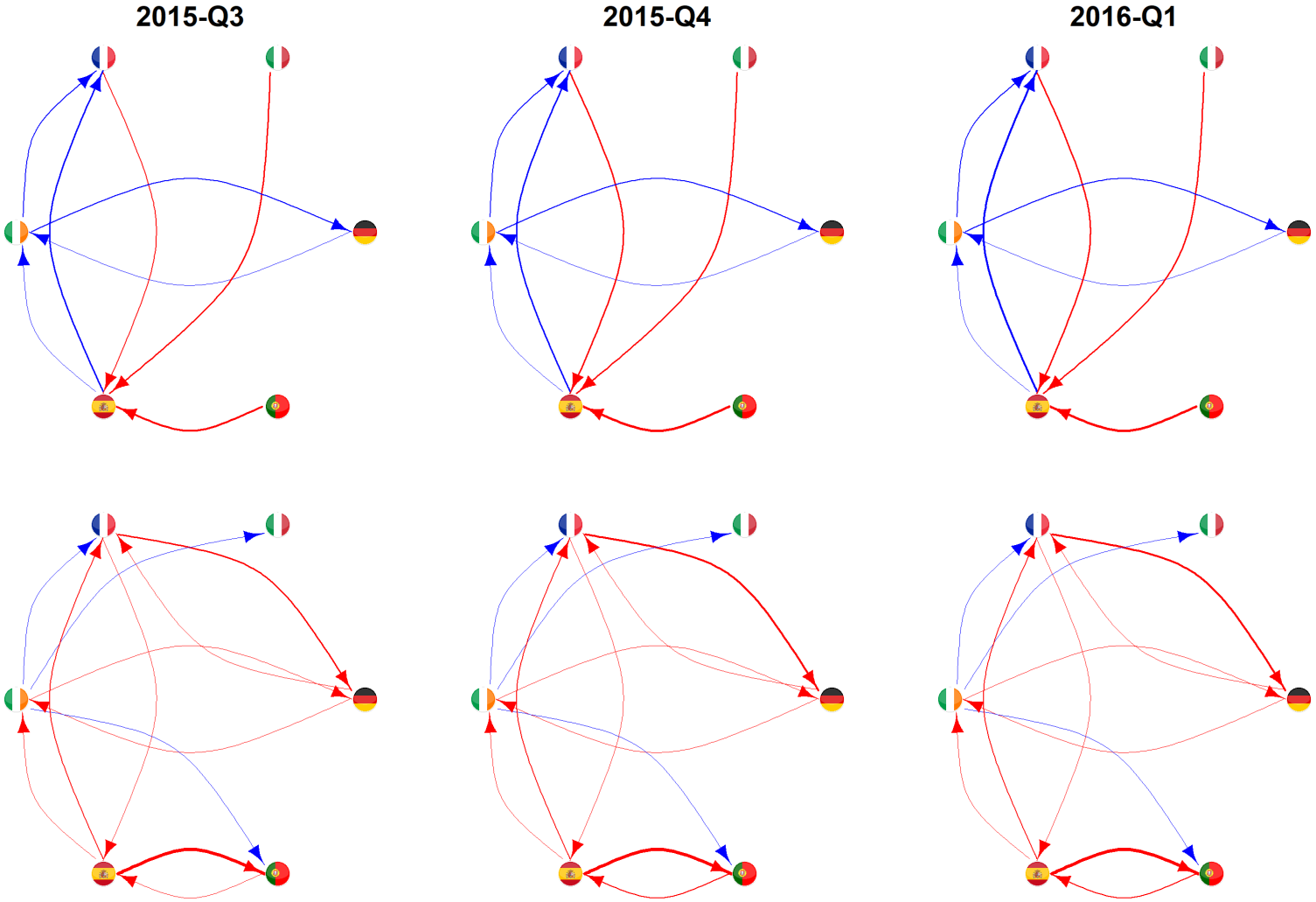

Figure S.13: Selected graphical representation of the shock response networks $\mathcal{W}_t^A$ in the first row, the persistence networks $\mathcal{W}_t^B$ in the second row, all obtained from daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows goes from counterparty to reporting and positive weights are inred while negative are in blue.

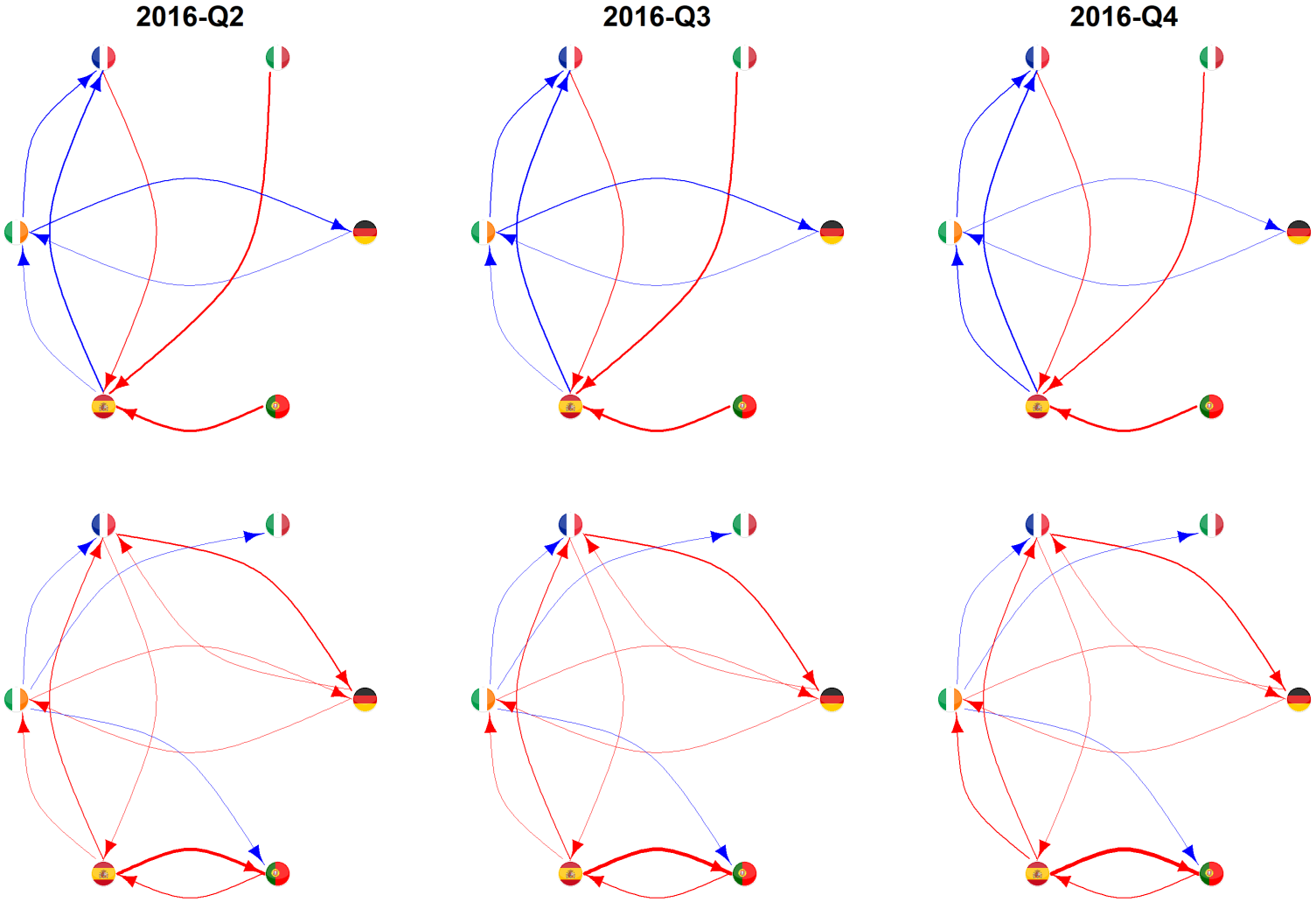

Figure S.14: Selected graphical representation of the shock response networks $\mathcal{W}_t^A$ in the first row, the persistence networks $\mathcal{W}_t^B$ in the second row, all obtained from daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016. The arrows goes from counterparty to reporting and positive weights are inred while negative are in blue.

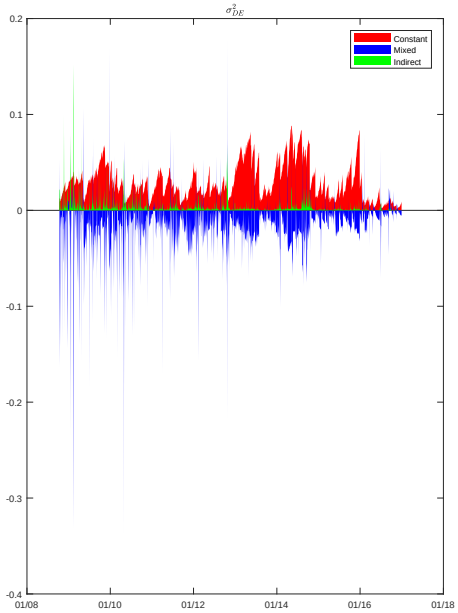

Figure S.15: Relative variance decomposition, of Germany, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

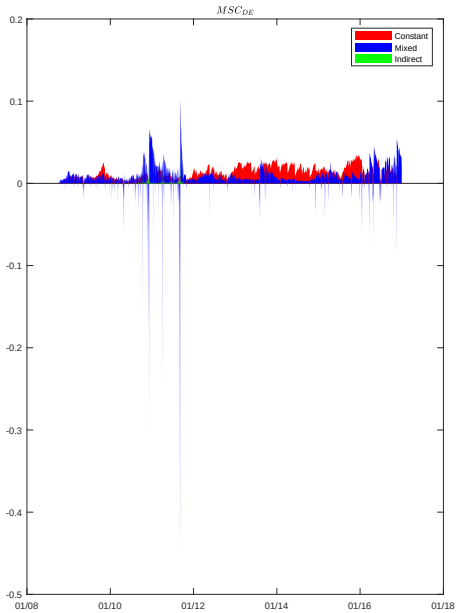

Figure S.16: Relative marginal spillover contribution, of Germany, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

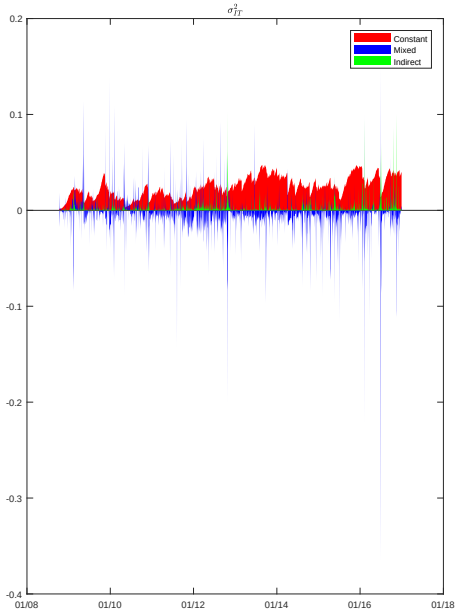

Figure S.17: Relative variance decomposition, of Italy, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

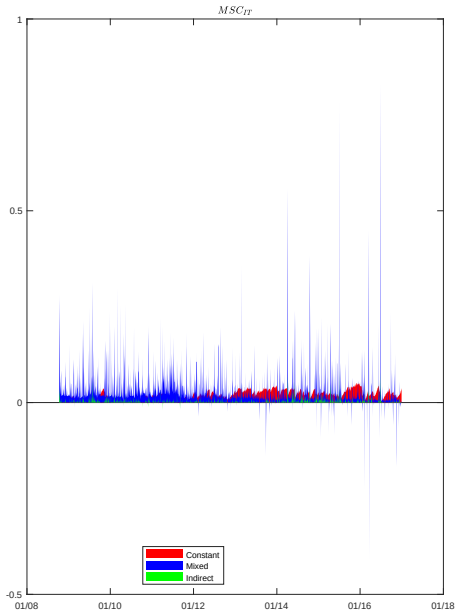

Figure S.18: Relative marginal spillover contribution, of Italy, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

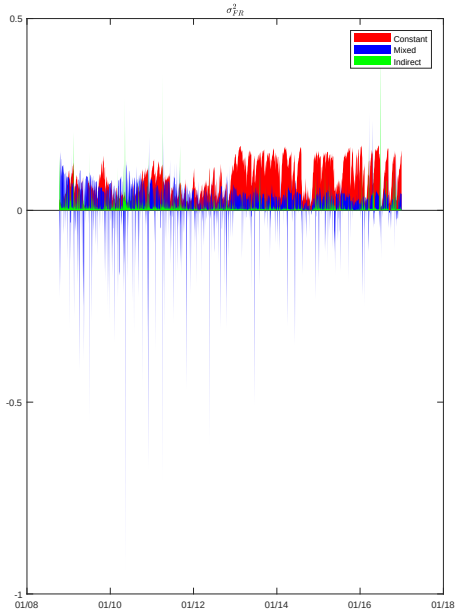

Figure S.19: Relative variance decomposition, of France, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

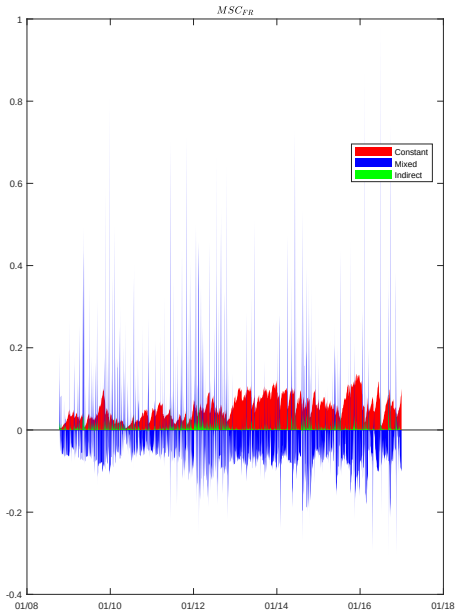

Figure S.20: Relative marginal spillover contribution, of France, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

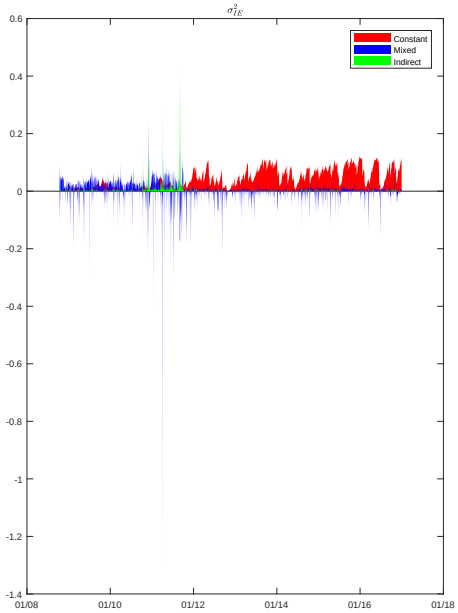

Figure S.21: Relative variance decomposition, of Ireland, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

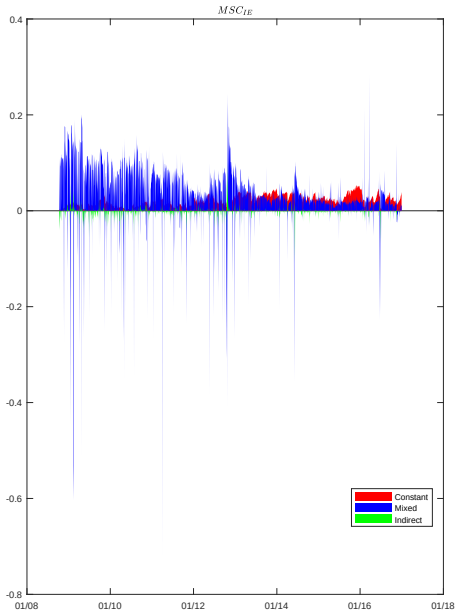

Figure S.22: Relative marginal spillover contribution, of Ireland, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

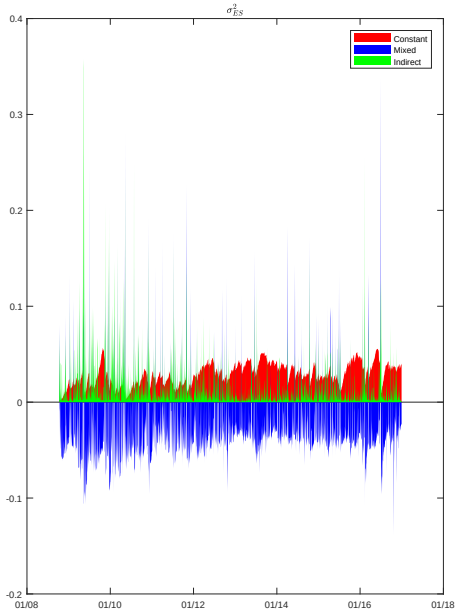

Figure S.23: Relative variance decomposition, of Spain, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

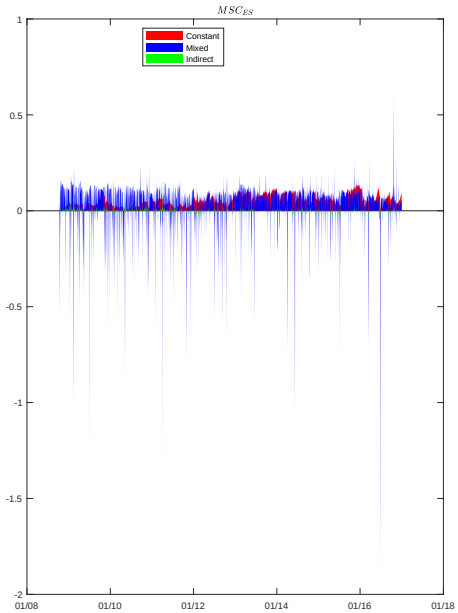

Figure S.24: Relative marginal spillover contribution, of Spain, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

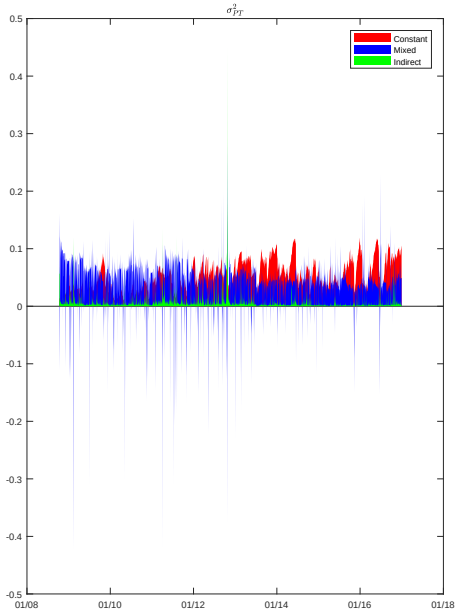

Figure S.25: Relative variance decomposition, of Ireland, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

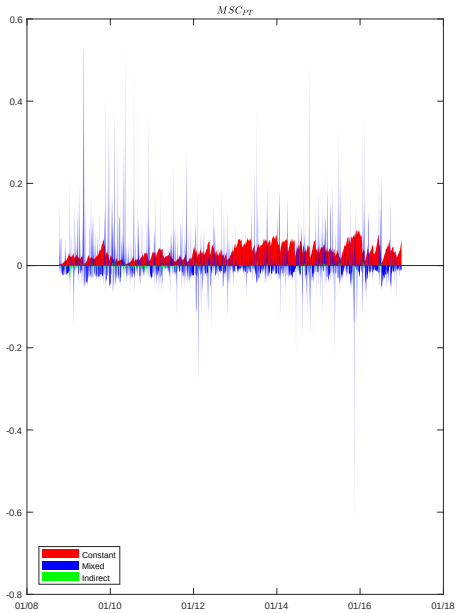

Figure S.26: Relative marginal spillover contribution, of Ireland, obtained from data on daily Changes in the Five-Year EMU Sovereign CDS spreads from 9/10/2008 to 30/12/2016, with the direct contribution omitted

Recent Issues

No. 224	Giulio Girardi, Kathleen W. Hanley, Stanislava Nikolova, Loriana Pelizzon, Mila Getmansky Sherman	Portfolio Similarity and Asset Liquidation in the Insurance Industry
No. 223	Florian Deuflhard	Quantifying Inertia in Retail Deposit Markets
No. 222	Vanessa Endrejat, Matthias Thiemann	Reviving the Shadow Banking Chain in Europe: Regulatory Agency, Technical Complexity and the Dynamics of Co-Habitation
No. 221	Axel H. Börsch-Supan, Klaus Härtl, Duarte N. Leite, Alexander Ludwig	Endogenous Retirement Behavior of Heterogeneous Households Under Pension Reforms
No. 220	Yangming Bao, Martin R. Goetz	Local Peer Effects and Corporate Investment
No. 219	Andreas Hackethal – Christine Laudenbach – Steffen Meyer – Annika Weber	Client Involvement in Expert Advice – Antibiotics in Finance?
No. 218	Florian Hoffmann, Roman Inderst, Marcus Opp	Only Time will Tell: A Theory of Deferred Compensation
No. 217	Maddalena Davoli, Jia Hou	Financial Literacy and Socialist Education: Lessons from the German Reunification
No. 216	Stefano Colonnello, Giuliano Curatola, Alessandro Gioffré	Pricing Sin Stocks: Ethical Preference vs. Risk Aversion
No. 215	Zsuzsa R. Huszár, Zorka Simon	The Pricing Implications of the Oligopolistic Securities Lending Market: A Beneficial Owner Perspective
No. 214	Edin Ibrocevic, Matthias Thiemann	All Economic Ideas are Equal, but Some are more Equal than Others: A Differentiated Perspective on Macroprudential Ideas and their Implementation
No. 213	Klaus Gugler, Michael Weichselbaumer, Christine Zulehner	Effects of Government Spending on Employment: Evidence from Winners and Runners-up in Procurement Auctions