Networks in risk spillovers: a multivariate GARCH perspective*

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

January 2016

Abstract

We propose a spatial approach for modeling risk spillovers using financial time-varying proximity matrices based on observable networks. We show how these methods could be useful in (i) isolating risk channels, risk spreaders and risk receivers, (ii) investigating the role of portfolio composition in risk transfer, and (iii) computing target exposure structures able to reduce the forecasted system variance and thus the risk of the system. Our empirical analysis builds on banks' foreign exposures provided by the Bank of International Settlements (BIS) as a proxy for Euro area cross-country holdings. We find, in the European sovereign bond markets, that Germany, Italy and, to a lesser extent, Greece are playing a central role in spreading risk, and Ireland and Spain are the most susceptible receivers of spillover effects that can be traced back to a physical claim channel: banks' foreign exposures. We additionally show that acting on these physical channels before the sovereign crisis, it would have been possible to have a clear risk mitigation outcome.

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

JEL Classification: C58, G10

*We thank the participants at the International Symposium of Forecasting 2015, Riverside, USA, and the Computational and Financial Econometrics Conference 2015 in Pisa, Italy, for the comments received from them. The usual disclaimers apply. The authors acknowledge financial support from 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 and the SAFE Research Center, funded by the State of Hessen initiative for research, LOEWE.

†University Ca' Foscari Venezia (Italy)

‡University of Padova (Italy)

§SAFE-Goethe University Frankfurt (Germany)

¶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 fields related to contagion and systemic risk. Even though those concepts are almost 20 years old, originating mostly in the description of the currency crisis at the end of the last century (see Forbes and Rigobon (2001); Allen and Gale (2000); Bae et al. (2003); Pericoli and Sbracia (2003) Eisenberg and Noe (2001); De Bandt and Hartmann (2000); Rochet and Tirole (1996); Freixas et al. (2000)), there is still not complete consensus around their definitions, as shown in De Bandt et al. (2009), Bisias et al. (2012), and Forbes (2012). For this reason, we prefer not to use those terms and refer to the phenomena we are investigating as variance and covariance spillovers. Our choice relates with our focus on methodological contributions. According, we show how networks can be integrated within multivariate volatility models for the purpose of analyzing risk spillovers. Furthermore, building on a classical model, we provide several new tools specifically designed to take advantage of the knowledge of the spillover effects, due to the existing connections (network links) across analyzed variables. These tools will allow us to filter networks accounting for the model dynamic, and to disentangle the network spillover contribution from the standard autoregressive-like risk dynamic. Moreover, from a forecasting perspective, we provide the methodology for recovering the optimal network design, with relevant implications for model users.

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 of, or shocks impacting on, a source entity before time t. This definition is signed, directional, 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 the returns (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). A covariance spillover is the contribution to the time-t covariance (thus dependence) of two entities due to the variance (risk) of an entity, the covariance between any two entities, or shocks affecting the covariance evolution, all before time t. Those definitions are clearly related to contagion and systemic risk, but at the same time they are rather restrictive since there could be other symptoms of the latter two broad concepts going beyond what spillover evidences might suggest. Moreover, despite being per se relevant, in the literature the detection of spillovers generally lacks an economic intuition associated with the fundamental transmission channels, motivating the presence of spillovers. For this reason, in this paper we give emphasis to both the detection of spillovers and the identification of their potential sources.

In more detail, we aim at introducing economically grounded drivers of variance and covariance spillovers. For this purpose, we take into account network structures among entities. In fact, networks represent the set of connections existing across entities, and thus the structure from which spillovers might originate.

In developing our methodology, we introduce and exploit a parallel between the network approach and the tools commonly used in spatial econometrics. In particular, the adjacency matrix, i.e. the companion representation of any network, is set equal to the distance matrix in spatial econometrics. Clearly, since the network describes an economic structure (and can be estimated with different and competing approaches), the distance measured by the network is of a purely economic nature. Nevertheless, the parallel allows us to take advantage of all the tools developed in the past decades in the spatial econometrics literature. From an empirical point of view, there has been a surge of financial economics contributions using economically based distances for the interpretation of contagion and systemic risk (see Billio et al. (2015); Bianchi et al. (2015); Keiler and Eder (2013); Schaumburg et al. (2014)). Our work belongs to this strand of literature and aims at building a bridge between the contagion, volatility spillover and network, and systemic risk fields.

The use of distances in spatial statistics and econometrics normally refers to physical elements, with geographical neighboring relationships representing central elements in several areas, for instance for real estate studies. Distances across financial entities are clearly more difficult to measure but, if available, they can be useful for taking advantage of the tools of network science, which studies the existence of relationships (edges) among many entities (nodes). Notably, in network science, entities are neighbors if they are directly or indirectly connected, and the distance is associated with the network structure.

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, most advanced strands of the spatial econometrics literature (see Keiler and Eder (2013); Schaumburg et al. (2014); Caporin and Paruolo (2015)) focus on statistical, economic and financial relationships. The starting point of our paper is the contribution of Caporin and Paruolo (2015) that introduces spatial econometrics tools to the analysis of conditional volatility models, and thus to the estimation and measurement of risks. We take a step further in two different aspects: First, we consider the proximity matrix (a generalization of the spatial distance matrix) used by Caporin and Paruolo (2015) as a network structure. The focus on the risk dimension, therefore, distinguishes us from recent related literature that mostly aims to explain expected returns, conditional to a network structure, with only an indirect description of the drivers of covariances. To our best knowledge, the only papers that introduce a functional dependence of the covariance from a network are those of Bianchi et al. (2015) and (Billio et al., 2015). The first considers inferred network relationship that are only of a statistical nature, while the second is closer to our economic foundation of network links. However, both papers do not focus on risk dynamics but rather on returns dynamics. On the contrary, our model belongs to the GARCH literature and is thus a pure risk model. Moreover, our network relationships are intended to be derived from financially relevant quantities (in our application, cross-border exposures of national banking systems), which are commonly perceived as potential transmission channels of shocks or associated with transmission channels. This choice is made with the aim of investigating and measuring the amount of spillover that could be explained by the transmission channels, as summarized by the network, and of prescribing an intervention in order to reduce spillover and mitigate risk. Our methodology thus relates to previous analyses pointing at what is called direct contagion as opposed to informational contagion (see Allen et al. (2009); Hasman (2013)).

The most important contributions of our work are on the methodological side. We start by generalizing the model of (Caporin and Paruolo, 2015) with the introduction of time-variation in the proximity matrices. This is a consequence of the use of financial quantities in the estimation of networks. By construction, financial variables are time-varying and thus networks are time-varying, with consequences on model estimation. By conditioning on the networks, we are able to overcome the computational burden and the curse of dimensionality. Notably, the model parameters becomes time-varying but still preserving the model feasibility in moderate cross-sectional dimensions, an uncommon feature in the multivariate GARCH literature. With our techniques, we are able to ascertain which part of the spillover is driven by network connections. We analyze this aspect by resorting to a decomposition of the risk in the system into different components. For each conditional variance (covariance), we thus separate the impact of the own past shocks and variances (covariances) from the contribution of shocks and variances (covariances) associated with linked (through the network) assets. This is an important issue from the monitoring point of view as it allows disentangling the role of the network from that of the autoregressive-like variance (covariance) dynamic. The third methodological contribution combines the model estimation outcomes with the financially-driven network. We show how to filter out the network using the significance of the model parameters. In this way, we recover information on the statistically relevant channels of spillover. Finally, by adopting proper forecasting techniques, our modeling framework is capable of proposing policy intervention strategies that aim to mitigate spillover and, in general, risk in the system, by acting on direct transmission channels, represented by the network connections. We show how to estimate the optimal networks by focusing on a specific criterion function.

We also present an empirical analysis that shows the potential benefits deriving from our approach. The empirical analysis concerns the relevance of the network of cross-country banking system exposures in explaining the European sovereign bond spillovers.

Our results underline the role of Ireland and Spain as risk receivers and the importance of Germany, Italy and, to a less extent, Greece as risk spreaders. In addition, we provide an ex ante way to mitigate the risk in the system by representing the system by an equally weighted index, and its risk with a forecasted variance proxy. We use this proxy as the objective function to be minimized, to determine target exposures that, if hypothetically enforced by the regulator before the sovereign bond crisis of the second quarter of 2010, could have limited its impact.

The paper is organized as follows. In Section 2, we introduce the econometric model for the spatial interpretation of risk. In Section 3 we discuss several model developments such as inference-based networks, system variance decomposition, and forecast based otimal target networks. In Section 4, we apply the methodology to bond yields for the major countries of the Euro area during the subprime and sovereign debt crises. Finally, in Section 5 we summarize our findings, outline their usefulness and trace a path for future extensions.

2 Spatial Econometrics of Risk

In this section, we propose a method for introducing financial proximity into the treatment of risk and dependence across financial entities or assets.[1] After describing the model, in the following section, we introduce several tools useful for inference and forecasting.

The first novelty of our methodology is provided by the fact we allow for time-varying nature of the measure of financial proximity. This is in sharp contrast to the usual spatial econometrics definition in which proximity relations are fixed, be they geographical, based for instance on physical measures of distance, as in Anselin (2001) and Elhorst (2003), or fixed economic quantities, such as the industry sectors used in Caporin and Paruolo (2015). However, when phisical distances are replaced by economically based distances, we can easily loose time invariance. Therefore, the introduction of time varying proximity

[1] In this section, we use the words entities, assets, subjects, and nodes as synonyms. The appropriate choice depends on the data analyzed.

relation is a novely but also a need. This aspect makes our estimation procedure more difficult, but at the same time this additional complexity allows us to better explore how the change in these time-varying relationships would impact on the risk and spillover effects across assets. Moreover, this allows us to suggest potential interventions that could be made by policy authorities and/or regulators.

2.1 Proximity and Networks

In the spatial statistics and econometric literature (see Anselin (2001),LeSage and Pace (2009),Elhorst (2003) a proximity matrix is a matrix whose entries are related to some notion of distance between entities. The prototypical example is real geographical distances. These are generally summarized into a weight matrix W, whose entries $[W]_{i,j}$ correspond to the geographical distance involved in moving from i to j. The matrix W is obviously static and symmetric. Making a parallel with network studies, W corresponds to the adjacency matrix of a weighted undirected network. Usually, as discussed in Elhorst (2003), the proximity matrix W is row normalized in order to maintain reasonable magnitudes for the parameters. In addition, in the classical spatial statistics and econometrics literature, the spatial impact is measured by means of a single coefficient ρ that pre-multiplies the weight matrix W. This implies only a common and unique impact across the entities involved in the analyses.[2]

Following Caporin and Paruolo (2015), we consider a more general viewpoint, and introduce the proximity matrix P as a linear combination of a weight matrix W and an identity matrix I, $P = \rho_0 I + \rho_1 W$, where, I is an identity matrix, ρ_1 and ρ_0 are scalars, ρ_1 representing the global impact of a network on the variables of interest, and ρ_0 being a constant common to all the variables. The advantage of this formulation for proximity matrices relates to the possibility of distinguishing between a common constant impact

[2]The two most common specifications are the spatial auto regressive 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).

and the additive one associated with the neighbors. A first and immediate generalization for the use of spatial econometrics in finance, as opposed to economics, is that the parameters are not constrained to be scalars. We generalize the scalar coefficients into diagonal parameter matrices, thereby introducing variable-specific coefficients. Thus, we have $P = diag\,(\rho_0)\,I + diag\,(\rho_1)\,W$. Caporin and Paruolo (2015) refer to such a specification as the *heterogeneous* case, given that each variable/asset has its own response to W, and we can thus introduce heterogeneity into the relationship with neighbors. We stress that, similarly to Caporin and Paruolo (2015), we want variable-specific parameters and this requires the use of diagonal matrix coefficients in the definition of proximity matrices, which thus become affine functions of the network.

We further elaborate on these matrices, going beyond what is discussed in Caporin and Paruolo (2015) in two directions:

- First, we note that these proximity matrices are not flexible enough to deal with financial relationships. In particular, we need to account for time-dependence in the weight matrix W.
- Second, symmetry in W is not necessary, consistent with the adjacency matrix that characterizes a weighted directed network.

The introduction of the time-varying dimension is relevant since financial markets move quickly and thus it is too restrictive to impose stable and time-invariant relationships among financial entities. In the previous literature, many applications of spatial methods to financial markets average these time-varying relationships to obtain a static framework (see, as an example, Schaumburg et al. (2014)). This clearly leads to a relevant information loss.

It is also important to consider asymmetry. In spatial econometrics, the matrix W is, in general, symmetric as, if A is a neighbor of B with a given distance between them, the reverse is also true. However, in a financial framework, symmetry is not usual. As an example, we can consider financial claims to define whether two financial institutions are neighbors. It is highly improbable that the amount of claim that A has on B will be the same amount that B has on A; thus, the relationship is likely to be asymmetric; moreover, considering net claims would mean losing relevant information as investors could perceive the connection differently depending on its direction.

These two features are not new and have already been taken into account in the recent financial network literature, as shown by Billio et al. (2015) and Schaumburg et al. (2014). However, these recent papers miss the opportunity to exploit the combination of asymmetry and vector coefficients. In fact, these two elements lead to a non-commutativity of the resulting model; that is, pre- and post-multiplication of the weight matrix by the coefficient matrix lead to two different models. This allows us to exploit and underline different features of the analyzed series. Accordingly, we consider two alternative representations for the proximity matrices and thus two alternative models:

$$P_L(W_t) = \text{diag}(\rho_{0,L}) I_n + \text{diag}(\rho_{1,L}) W_t \quad (1)$$

$$P_R(W_t) = \text{diag}(\rho_{0,R}) I_n + W_t \text{diag}(\rho_{1,R}) \quad (2)$$

where n is the number of series and I_n is the $n \times n$ identity matrix.

To better highlight our contributions, in Table 1 we briefly compare our approach to the classical spatial econometrics one. To summarize, the differences we introduce are (i) the use of time variation in spatial proximity as opposed to the use of constant and physical proximity relationships, (ii) asymmetry in defining neighboring relationships, making the approach coherent with the use of weighted directed networks in defining proximity, (iii) generalization of the proximity matrices' construction through the exploitation of heterogeneity in defining the impact from neighbors, and (iv) the introduction of left and right multiplication proximity matrices and thus models to highlight different features of the data. We also introduce an additional distinguishing element that refers to the use of a normalization step in the construction of the spatial matrix W. In fact, the presence of time variation requires the use of specific normalization rules, which necessarily go beyond the traditional row normalization. Given its relevance in our framework, we discuss aspects of normalization in a later subsection. We stress here that our approach also differs from the classical one in the choice of normalization.

Table 1: Proximity Models

Classical Spatial Statistics	Our Proximity Model
Static	Time varying
$W_t = W$	W_t
Weighted undirected network	Weighted directed network
$W' = W$	$W_t' \neq W_t$
Proximity is a linear combination	Proximity is a left or right affine function
$\rho_0 + \rho_1 W$	$\text{diag}(\rho_{0,L})\, I + \text{diag}(\rho_{1,L})\, W_t$
	$I\text{diag}(\rho_{0,R}) + W_t\text{diag}(\rho_{1,R})$
ρ_0, ρ_1 scalars	$\rho_{0,L}, \rho_{1,L}, \rho_{0,R}, \rho_{1,R}$ vectors
Row normalization	Economic magnitude $M_{t,j}$ normalization
$\sum_j [W]_{i,j} = 1$	$[W_t]_{i,j} \to \dfrac{[W_t]_{i,j}}{M_{t,j}}$

2.2 Model and Parameter Estimation Procedure

The introduction of proximity matrices allows us to recover the role of networks, once they have been introduced in a dynamic model. Since we are interested in risk analysis, we need to introduce a dynamic variance model. A popular specification adopted for the estimation of conditional variance matrices is the BEKK model of (Engle and Kroner, 1995). 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. We do not further specify the mean model, since we are interested in the risk dynamics. The simplest BEKK model is given by

$$u_t = \Sigma_t^{1/2}\epsilon_t \qquad \epsilon_t \sim \mathcal{N}(0, I_n), \qquad t = 1, \ldots, T \tag{3}$$

$$\Sigma_t = CC' + Au_{t-1}u_{t-1}'A' + B\Sigma_{t-1}B' \qquad M = L, R \tag{4}$$

where C is a lower triangular matrix, $\Sigma_t^{1/2}$ is the Cholesky decomposition of Σ_t,[3] and A and B are $n \times n$ parameter matrices.

The full BEKK model described in equation (4) is computationally unfeasible 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. Unfortunately, despite being feasible, these restricted specifications impose strong limitations on the interpretability of the model outcomes as 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. As discussed in Caporin and Paruolo (2015), the spatial version of the BEKK model has the main advantage of being more parsimonious than the full BEKK case, but at the same time it is much more flexible than the diagonal specification. Moreover, the inclusion of proximity matrices allows us to model, with limited additional parameters, spillovers and feedback effects.

As already anticipated, we extend the spatial-BEKK GARCH model by introducing time variation in the proximity matrices to take into account time-varying weight matrices W_t, and by allowing for two different forms of proximity. The model we consider thus has the following structure:

$$u_t = \Sigma_t^{1/2}\epsilon_t \quad \epsilon_t \sim \mathcal{N}\left(0, I_n\right), \quad t = 1, \ldots, T \tag{5}$$

$$\Sigma_t = CC' + A_M\left(W_t\right) u_{t-1}u_{t-1}' A_M\left(W_t\right)' + B_M\left(W_t\right) \Sigma_{t-1} B_M\left(W_t\right)' \qquad M = L, R \tag{6}$$

where the parameter matrices are proximity matrices as in equations (1) and (2) and thus include either left multiplication or right multiplication. Consequently, the

[3] 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.

parameters matrices might have the following structure

$$
\begin{aligned}
A_L\left(W_t\right) &= A_{0,L} + A_{1,L}W_t = \operatorname{diag}\left(a_{0,L}\right) I + \operatorname{diag}\left(a_{1,L}\right) W_t && (7)\\
B_L\left(W_t\right) &= B_{0,L} + B_{1,L} = W_t = \operatorname{diag}\left(b_{0,L}\right) I + \operatorname{diag}\left(b_{1,L}\right) W_t && (8)\\
A_R\left(W_t\right) &= A_{0,R} + W_t A_{1,R} = I\operatorname{diag}\left(a_{0,R}\right) + W_t\operatorname{diag}\left(a_{1,R}\right) && (9)\\
B_R\left(W_t\right) &= B_{0,R} + W_t B_{1,R} = I\operatorname{diag}\left(b_{0,R}\right) + W_t\operatorname{diag}\left(b_{1,R}\right) && (10)
\end{aligned}
$$

and where $a_{0,M}$, $a_{1,M}$, $b_{0,M}$ and $b_{1,M}$, with $M = L, R$, are $n \times 1$ vectors.Note that the L and R matrices are substitute and do not co-exist in a single model.

We note that the two specifications, with left and right multiplication, only provide different results if the spatial matrices W_t are not symmetric. Under symmetry of W_t, the two specifications lead to the same result.[4]

Within the Spatial-BEKK framework, left and right multiplication parametrizations allow researchers to focus on different aspects of risk propagation. To better understand this aspect, 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 Spatial-BEKK model with right or left multiplication. 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

$$
v_t = \left(I_n + \theta W\right) u_t. \qquad (11)
$$

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:

[4]This is a consequence of symmetry. Suppose we focus on the shock component and assume a constant W. We have $A_L u_{t-1}u_{t-1}' A_L' = \left(A_{0,L} + A_{1,L}W\right) u_{t-1}u_{t-1}' \left(A_{0,L} + W' A_{1,L}\right)$ thanks to the diagonal form of the parameter matrices. Moreover, by symmetry, $\left(A_{0,L} + A_{1,L}W\right) u_{t-1}u_{t-1}' \left(A_{0,L} + W' A_{1,L}\right) = \left(A_{0,L} + W' A_{1,L}\right) u_{t-1}u_{t-1}' \left(A_{0,L} + A_{1,L}W\right)$. The latter is equal to the right multiplication case if $W = W'$.

$$\begin{aligned} v_t &= v_t^0 + v_t^1 && (12)\\ v_{i,t}^0 &= [I_n u_t]_i = u_{i,t} && (13)\\ v_{i,t}^1 &= [\theta W u_t]_i = [W \theta u_t]_i = \theta \sum_{j=1}^{n} \omega_{i,j} u_{j,t} && (14) \end{aligned}$$

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 Spatial-BEKK model and provide a novel decomposition.

We start from the left multiplication case and we focus on the ARCH part of the model as we point at highlighting the role of innovations. We note that

$$\begin{aligned} v_{L,t} = A_L(W) u_t &= (A_{0,L} + A_{1,L} W) u_t = v_{L,t}^0 + v_{L,t}^1 && (15)\\ v_{L,i,t}^0 &= [A_{0,L} u_t]_i = a_{0,L,i} u_{i,t} && (16)\\ v_{L,i,t}^1 &= [A_{1,L} W u_t]_i = a_{1,L,i} \sum_{j=1}^{n} \omega_{i,j} u_{j,t}. && (17) \end{aligned}$$

We have that the $i-$th element of $v_{L,t}$ depends on its own past shock, weighted by the coefficient $a_{0,L,j}$ (direct effect), and on the past shocks of its neighbors weighted by the distance, all loaded with the same coefficient, $a_{1,L,j}$ (indirect effect).

Consequently, bearing in mind we are discussing properties of a conditional covariance model, the left multiplication specification allows us to investigate which are the risk receivers or the systemically fragile entities, since in this case the model parameters emphasize the role of the risk recipients. The model parameters included into the vector $a_{1,L}$ monitors the reaction of risk receivers to shocks originated from the neighbors or, in a network framework, from the connected nodes.

If we consider the right multiplication case and still refer to the ARCH part of the model, we have

$$v_{r,t} = A_R(W)\, u_t = (A_{0,R} + WA_{1,R})\, u_t = v^0_{R,t} + v^1_{R,t} \tag{18}$$

$$v^0_{R,i,t} = [A_{0,R}u_t]_i = a_{0,R,i}u_{i,t} \tag{19}$$

$$v^1_{R,i,t} = [WA_{1,R}u_t]_i = \sum_{j=1}^{n} \omega_{i,j}a_{1,R,j}u_{j,t}. \tag{20}$$

Differently 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). Consequently, with the right multiplication version of the model, the parameters magnify the effect of the source of risk, allowing us to focus on risk spreaders or on systemically important entities. In addition, indirect effects $v^1_{R,i,t}$ now depend on more parameters compared to the left multiplication case. In fact, we can rewrite the indirect effect as follows

$$v^1_{R,i,t} = [W]_{i,.}\,(a_{1,R} \odot u_{t-1}) \tag{21}$$

where $[W]_{i,.}$ the $i-$th row of the W matrix and $\odot$ is the element-by-element matrix product (or Hadamard product). This stresses that the indirect effect depends on the entire vector of parameters $a_{1,R}$.

Since the two competing models, with left and right multiplication, do provide insights into two very different aspects of risk propagation, we suggest estimating both of them on the same dataset. Moreover, using results from two estimated models we can easily recover two different rankings for the countries, one based on their risk spreading effectiveness (systemic importance) and one based on their risk receiving propensity (fragility). This clearly opens the door for further economically relevant insights into risk propagation mechanisms.

Differently from Caporin and Paruolo (2015), our specifications can include time-varying proximity matrices. In this paper, we assume these matrices are known before the estimation of the BEKK model. Consequently, the model evaluation is conditional on the availability of the full sequence W_t for $t = 1, 2, \ldots, T$. The estimated model parameters are time-invariant, and correspond to the diagonal vectors in the proximity matrices. However, the presence of W_t makes the traditional full BEKK parameter matrices time-varying and this is certainly an important innovation compared to the current literature, since we are not aware of a closed-form methodology for estimating BEKK-type models with time-varying parameters (even if the time variation is driven by exogenous terms).[5]

If the W_t are not known, we consider a two-step estimation procedure. The first step focuses on the estimation of the spatial matrices W_t, while the second one is devoted to the estimation of the spatial-BEKK parameters, and is conditional to the first step.

We also highlight that the spatial matrices W_t could have a smoothly evolving pattern; that is, they are time varying but on a lower time scale than that adopted for the evolution of the entities in the system. A similar assumption has already been used in Billio et al. (2015).

The parameter estimation of our spatial-BEKK with time-varying parameters uses quasi maximum likelihood estimation (QMLE) methods with robust standard errors. If we denote by $\theta \equiv (\text{vec}\,(C)\,, a_{0,M}, a_{1,M}, b_{0,M}, b_{1,M})$ the vector of parameters, the log- likelihood is

[5] For DCC-type and BEKK models, time variation could be modeled through a Markov switching mechanism (see Billio and Caporin (2005); Pelletier (2006); Lee and Yoder (2007)), but the estimation would require some approximation due to the path-dependence structure of the models.

$$\ell(\theta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\det(\Sigma_t) - \frac{1}{2}u_t(\Sigma_t)^{-1}u_t'. \tag{22}$$

Following Caporin and Paruolo (2015), a simple identification restriction consists of constraining the first element of the vectors $\mathrm{vec}(C), a_{0,M}, a_{1,M}, b_{0,M}$ and $b_{1,M}$, to be positive. We stress that this identification restriction allows for the presence of coefficients of both signs in $a_{0,M}, b_{0,M}, a_{1,M}$ and $b_{1,M}$. Such a flexibility in the sign allows for the possibility of negative spillovers, that is, terms that reduce the variance. Such a possibility could be of the upmost relevance in portfolio construction as it could boost diversification benefits.

With respect to the asymptotic properties of the estimators, following Boussama et al. (2011) and under the Gaussian assumption for innovations, we can ensure the ergodicity and stationarity of the process by constraining the maximum spectral radius:

$$\max_{t\in[1,T]} \rho\left(A_M(W_t)\otimes A_M(W_t) + B_M(W_t)\otimes B_M(W_t)\right) < 1. \tag{23}$$

Moreover, combining the stationarity constraint with an assumption on the finiteness of the sixth-order moments of innovations u_t, we can obtain the consistency and asymptotic normality of the QMLE procedure (with robust standard errors and conditional on the availability of the W_t matrices); see Hafner and Preminger (2009).[6] The latter result allows us to perform classical inferential and model specification procedures.

[6]The conditions for ergodicity given in Boussama et al. (2011) are less restrictive, allowing for integrated processes as well, but are very difficult to impose in the optimization procedure.

3 Model Developments and Improvements for Risk Analyses

In the following subsections, we present several advantages and potential improvements provided by the spatial-BEKK model with time-varying parameters, focusing on (i) the flexibility allowed by different normalization rules, (ii) the insight provided by inference-based networks, (iii) the decomposition of system variance, and (iv) the determination of potential regulatory intervention thanks to the identification of optimal network relationships through forecasting techniques. Their practical impact is highlighted in the empirical analysis presented in Section 4.

3.1 Normalization Rules

Taking into account 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 be able 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 consider different types of normalizations. We thus suggest employing normalizations that are either time-invariant or divide each row of W_t by an (economic) measure of the magnitude of the entities. The first case refers, for instance, to the max row normalization adopted by Billio et al. (2015), in which the row normalization of spatial matrices W_t considers the maximum row sum over time and not the time-specific row sum. The time-invariant normalization M_i corresponding to the i-th row thus corresponds to

$$M_i = \max_t \left\{ \sum_{j=1}^{N} [W_t]_{i,j} \right\} \tag{24}$$

In the second case, the normalization starts from the availability of a quantity monitoring the size or the (economic) relevance 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 second approach, we stress that the spatial matrices W_t will not have maximum eigenvalue equal to 1 as is standard in spatial econometrics. However, we stress that, within the spatial-BEKK framework, the matrices W_t are just a tool for solving the curse of dimensionality, and therefore the presence of a maximum eigenvalue differing from 1 is not a concern.

3.2 Inference-Based Networks

Within our spatial-BEKK model, time-varying parameter matrices are composed by two elements: the series W_t that we assume to be observed without errors, and the parameters estimates, which we assume to be characterized by an asymptotic normal distribution. By combining these two components, we can revise our knowledge of the underlying network by building several "inference-based" networks, whose characterization derives from the dependence on specific inferential aspects of the BEKK model parameters. We first note that the spatial-BEKK model depends on the W_t matrices but, ex post, the network information might be revised and filtered from two different parameter sets: those associated with the A matrices, the so-called ARCH parameters, and those coming from the B matrices, the GARCH parameters. Moreover, as already discussed, left and right multiplication proximity matrices correspond to risk spreaders (systemic important) and risk receivers (fragile) entities, leading to two different interpretations of the results and also two different inferred structures. Therefore, when considering inference-based networks, we can distinguish between four possible cases, depending on the information used to filter out the network:

- the use of either ARCH or GARCH parameter matrices;

- the use of either left or right multiplication in the proximity matrices.

We focus here on the A matrices with left multiplication and note that the same line of reasoning can be applied to B matrices as well as to right multiplication.

We propose building an inferred network by taking into consideration both the size and the uncertainty associated with the estimated parameters. We define the new inferred weight matrix $\mathcal{W}_t^{A,L}$ and obtain it as follows. First of all, let us consider the estimated parameter matrix

$$\hat{A}_{L,t} = diag\left(\hat{a}_0\right) + diag\left(\hat{a}_{1,L}\right) W_t. \tag{25}$$

Then, focus on the off-diagonal element at position i, j, with $i \neq j$, that is $\hat{a}_{1,L,i}\omega_{t,i,j}$. The test statistic for the null hypothesis of $a_{1,L,i}\omega_{t,i,j} = 0$ is

$$t\text{-}stat\left(a_{1,L,i}\omega_{t,i,j} = 0\right) = \frac{\hat{a}_{1,L,i}\omega_{t,i,j}}{\sqrt{\mathbb{V}\text{ar}\left(\hat{a}_{1,L,i}\omega_{t,i,j}\right)}} \tag{26}$$

$$= \frac{\hat{a}_{1,L,i}\omega_{t,i,j}}{\omega_{t,i,j}\sqrt{\mathbb{V}\text{ar}\left(\hat{a}_{1,L,i}\right)}} = t\text{-}stat\left(a_{1,L,i} = 0\right) \text{ for each } j$$

if we consider the network to be observed without error. This also means that the p-values for the null hypothesis $a_{1,L,i}\omega_{t,i,j} = 0$ and $a_{1,L,i} = 0$ are equal, with the equivalent result for right multiplication.

We thus define as a filtered network, or inference-based network, the network whose adjacency matrix has been filtered with the p-values of the model parameters. If we focus on the ARCH parameters and left multiplication, the filtered network equals

$$\left[\mathcal{W}_t^{A,L}\right]_{i,j} = a_{1,L,i}\omega_{t,i,j} \times \left(1 - p\text{-}value\left(a_{1,L,i}\omega_{t,i,j} = 0\right)\right) = a_{1,L,i}\omega_{t,i,j} \times \left(1 - p\text{-}value\left(a_i^L = 0\right)\right) \tag{27}$$

Using a similar approach we can derive filtered networks from the GARCH parameters and/or from right multiplication. 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. Moreover, the left and right multiplication cases are associated with a focus on systemically important and fragile entities respectively. The possibility of obtaining an inference-based network through the filtration of model parameter estimates can offer further insights into the network's relevance.

3.3 Decomposition of System Variance

The introduction of proximity matrices in the dynamic of BEKK models allows the estimation $\frac{(n+1)n}{2}$ 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 Spatial-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 the Spatial-BEKK a spatio-temporal model, we have a decomposition conditional to the past.

We propose a four term decomposition of the system conditional covariance:

1. *Costant Contribution*: it represents the part of the covariance which is unrelated to the model dynamic and is thus independent from the network;

2. *Direct Contribution*: it 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*: it represents the covariance contribution due to indirect effects, that is due to the network exposures of the assets;

4. *Mixed contribution*: it represents the covariance contribution originating from the quadratic form of the model and 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[7] and we focus on the left multiplication model.

From equation (6), the conditional covariance at time t is given by the sum of three elements: the constant, a quadratic term associated with the shocks; a quadratic term associated with the past conditional covariance.

In our decomposition, the *constant term* is simply given by constant of the conditional covariance, thus CC'.

We now focus on the shock response term, the ARCH part of the model. We remind that we introduce in equations (16) and (17) a definition of direct $v^0_{L,i,t}$ and indirect effects $v^1_{L,i,t}$ within the ARCH part. We now decompose the entire shock response term as follows:

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

We take a closer look at the decomposition focusing on the element i,j of the matrix $A_L\left(W\right)u_{t-1}u'_{t-1}A_L\left(W\right)'$. Note that if $i=j$ we deal with variances, while for $i \neq j$ we consider covariances.

[7] The diagonal elements remain null. Note that if off-diagonal elements are zero, some simplification might be present in the equations we report.

The first element in the ARCH term decomposition refers to the variance (or covariance) own shock:

$$\left[A_{0,L}u_{t-1}u_{t-1}'A_{0,L}'\right]_{i,j} = v_{L,i,t-1}^{0}v_{L,j,t-1}^{0} = a_{0,L,i}a_{0,L,j}u_{it-1}u_{j,t-1}.$$

This is comparable to a direct shock contribution. The fourth term, represents the contribution faced by element i, j of the covariance due to network exposures:

$$\left[A_{1,L}Wu_{t-1}u_{t-1}'W'A_{1,L}'\right]_{i,j} = v_{L,i,t-1}^{1}v_{L,j,t-1}^{1} = a_{1,L,i}a_{1,L,j}\sum_{k=1}^{n}\omega_{i,k}u_{k,t}\sum_{l=1}^{n}\omega_{j,l}u_{l,t},$$

This corresponds to an indirect effect, that is the shocks impact due to the network,

The second and third terms can be interpreted as mixed effects as they combine both direct and indirect elements:

$$\left[A_{1,L}Wu_{t-1}u_{t-1}'A_{0,L}'\right]_{i,j} = v_{L,i,t-1}^{1}v_{L,j,t-1}^{0} = a_{1,L,i}a_{0,L,j}\sum_{k=1}^{n}\omega_{i,k}u_{k,t}u_{j,t}$$

and

$$\left[A_{0,L}u_{t-1}u_{t-1}'W'A_{1,L}'\right]_{i,j} = v_{L,i,t-1}^{0}v_{L,j,t-1}^{1} = a_{0,L,i}a_{1,L,j}u_{j,t}\sum_{k=1}^{n}\omega_{j,k}u_{j,t}.$$

Moving to the GARCH part of the model, similarly to the ARCH part, we first introduce two additional terms, which are associated with the direct and indirect persistence effects. These two terms equal

$$m_{L,i,t}^{0} = \left[B_{0,L}u_t\right]_i = b_{0,L,i}u_{i,t} \tag{28}$$

$$m_{L,i,t}^{1} = \left[B_{1,L}Wu_t\right]_i = b_{1,L,i}\sum_{j=1}^{n}\omega_{i,j}u_{j,t}. \tag{29}$$

They differ from the terms $v_{L,i,t}^{0}$ and $v_{L,i,t}^{1}$ in their dependence on the GARCH parameters. The direct, indirect and mixed contributions we can recover from the GARCH part of the model correspond to covariances between the terms in equations (28) and (29).

In fact, for the indirect contribution originating from the GARCH part of the mode, using the bilinearity of the conditional covariance operator and conditionally on the network W we have:

$$\begin{aligned}
\left[\Omega^{1,1}_{L,t-1}\right]_{i,j} &= \mathbb{C}\text{ov}\left(m^1_{L,i,t-1}, m^1_{L,j,t-1}\right| I_{t-2}, W) \\
&= \mathbb{C}\text{ov}\left(b_{1,L,i}\sum_{j=1}^{n}\omega_{i,k}u_{k,t-1}, b_{1,L,j}\sum_{l=1}^{n}\omega_{j,l}u_{l,t-1}\right| I_{t-2}, W\Bigg) \\
&= b_{1,L,i}\sum_{j=1}^{n}\omega_{i,k}b_{1,L,j}\sum_{l=1}^{n}\omega_{j,l}\mathbb{C}\text{ov}\left(u_{k,t-1}, u_{l,t-1}\right| I_{t-2}, W) \\
&= b_{1,L,i}\sum_{j=1}^{n}\omega_{i,k}b_{1,L,j}\sum_{l=1}^{n}\omega_{j,l}\left[\Sigma_{t-1}\right]_{k,l} \\
&= \left[B_{1,L}W\Sigma_{t-1}W'B'_{1,L}\right]_{i,j}.
\end{aligned}$$

We can recover similar equalities for the direct and mixed contributions. In Table (2) we summarize the elements appearing in the conditional covariance decomposition.

Table 2: Decomposition of $[\Sigma_t]_{i,j}$ in the left multiplication model

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

Further, we highlight that the decomposition is time-varying by construction and it might be also affected by the dynamic in the network structure.

Using the definitions in equations (19) and (20) for right direct and indirect effect, and following the derivation detailed above for the left multiplication case, it is possible to derive a similar decomposition for the right multiplication case.

The variance decompositions outlined above are specific to a single element of the covariance matrix. However, we might be interested in recovering a synthetic measure of the decomposition at the entire covariance level. We propose defining this synthetic (and time-varying) measure starting from a composite representation of the system. Among the many possible choices, we choose the simplest one and thus consider a portfolio characterized by equal weights for each financial institution or entity whose risk is being analyzed using the spatial-BEKK model. The use of different weighting schemes, with potentially better economic explanations, is left for further empirical research.

The conditional variance of the equally weighted portfolio is obtained by averaging the conditional covariance matrix of the system. Therefore, the equally weighted portfolio variance decomposition is equal to[8]

$$\mathbb{V}\text{ar}\left(\frac{1}{n}\mathbf{1}'y_t\middle|\mathcal{I}_{t-1}\right) = \left(\sigma_t^{Constant}\right)^2 + \left(\sigma_t^{Direct}\right)^2 + \left(\sigma_t^{Indirect}\right)^2 + \left(\sigma_t^{Mixed}\right)^2 \tag{30}$$

$$\left(\sigma_t^{Constant}\right)^2 = \frac{1}{n^2}\sum_{i,j=1}^{n}[CC']_{i,j} \tag{31}$$

$$\left(\sigma_t^{Direct}\right)^2 = \frac{1}{n^2}\sum_{i,j=1}^{n}\left(v_{L,i,t-1}^0 v_{L,j,t-1}^0 + \left[\Omega_{L,t-1}^{0,0}\right]_{ij}\right) \tag{32}$$

$$\left(\sigma_t^{Indirect}\right)^2 = \frac{1}{n^2}\sum_{i,j=1}^{n}\left(v_{L,i,t-1}^1 v_{L,j,t-1}^1 + \left[\Omega_{L,t-1}^{1,1}\right]_{ij}\right) \tag{33}$$

$$\left(\sigma_t^{Mixed}\right)^2 = \frac{1}{n^2}\sum_{i,j=1}^{n}\Big(v_{L,i,t-1}^0 v_{L,j,t-1}^1 + \left[\Omega_{L,t-1}^{0,1}\right]_{ij} \tag{34}$$

$$+ \; v_{L,i,t-1}^1 v_{L,j,t-1}^0 + \left[\Omega_{L,t-1}^{1,0}\right]_{ij}\Big) \tag{35}$$

Since the model specifications allow for the possibility of having positive and negative signs on both ARCH and GARCH coefficients, in principle, we expect that diversification benefits could arise from all four contributions. An equivalent decomposition can be derived for the right multiplication case.

[8]Again we report here only the left multiplication case, because the right case is completely analogous

3.4 Optimal Network and Target Exposures

Another interesting aspect of our model relates to the possibility of obtaining, given past information, the optimal network, that is the weight matrix $W^\star$ that minimize the future evolution of the entire system variance. Accordingly, it is possible to define target exposures that minimize the (future) risk in the system.

3.4.1 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 $t \in [1, \dots, T]$ from which estimates for $\hat{C}$, $\hat{A}$, $\hat{B}$ 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 $T+1$ to $T+h$. The first step is to compute the $n \times 1$ vector of filtered innovations (standardized residuals) $\hat{\epsilon}_t$ for each time in the estimation period $t \in [1, \dots, T]$, by multiplying the vector u_t by the inverse of the Cholesky decomposition of the estimated conditional covariance $\hat{\Sigma}_t$:

$$\hat{\epsilon}_t = \hat{\Sigma}_t^{-\frac{1}{2}} u_t, \qquad t \in [1, \dots, T]. \tag{36}$$

The second step is to bootstrap N_B samples of length h from the $n \times T$ matrix of filtered innovations $[\hat{\epsilon}_1, \dots, \hat{\epsilon}_T]$, using 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)[9]. In this way, we obtain the bootstrapped innovations $\tilde{\epsilon}_{T+l}^{[b]}$ with $b \in [1, \dots, N_B]$ and $l \in [1, \dots, h]$.In turn, these allow computing the bootstrapped mean innovations u_{T+l} and the bootstrapped covariances for each l and b:

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

$$\tilde{u}_{T+l}^{[b]} = \hat{\Sigma}_{T+l}^{\frac{1}{2}} \tilde{\epsilon}_{T+l}^{[b]} \tag{37}$$

$$\tilde{\Sigma}_{T+l}^{[b]} = \hat{C}\hat{C}' + \hat{A}\tilde{u}_{T+l}^{[b]}\tilde{u}_{T+l}^{[b\prime]}\hat{A} + B\tilde{\Sigma}_{T+l-1}^{[b]}B' \tag{38}$$

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

$$\hat{\Sigma}_{T+l}^{F} = \frac{1}{N_B}\sum_{b=1}^{N_B} \tilde{\Sigma}_{T+l}^{[b]}. \tag{39}$$

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 exist 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.

3.4.2 Optimal Network

Conditional on the bootstrapped innovations $\tilde{\epsilon}_{T+l}^{[b]}$ with $b \in [1, \ldots, N_B]$ and $l \in [1, \ldots, h]$, and assuming that the network is constant over the forecast horizon, the forecasted covariance path is a function of the network at time T, $W_T \mapsto \hat{\Sigma}_{T+l}^{F}(W_T)\ l \in [1, \ldots, h]$. This raises the interesting possibility of obtaining the target network that can reduce the risk of the system. To define optimal target exposures, we require that they, at least locally, minimize the variance of the system, which we approximate by the equally weighted index of all the series. Such an approach is of particular interest when there exists a 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 certainly available at a daily or even higher frequency.

We thus assume that the network data are available at a lower frequency than the entities data. 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 networks 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 thus require that the average forecasted volatility 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:

$$\min_{\mathrm{vec}W^\star} \frac{1}{h}\sum_{l=1}^{h}\frac{1}{n^2}\mathbf{1}'\hat{\Sigma}^F_{T+l}\left(W^\star\right)\mathbf{1} \tag{40}$$

$$s.t.\, 0 \leq \left[W^\star\right]_{i,j} \leq 1\, for\, i,j = 1\ldots n \tag{41}$$

$$\mathrm{Tr}\left(W^\star\right) = 0 \tag{42}$$

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

$$\sum_{j=1}^{n}\left[W^\star\right]_{i,j} = \sum_{j=1}^{n}\left[W_Q\right]_{i,j}, \tag{43}$$

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. The use of out strength or in strength imposes a redistribution among the receivers or the donors in the network. The choice of preferred constraint depends on the application and on the purposes of the analysis.

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 one to compute the filtered innovations for the forecasted periods, conditional on the true, observed $Q+1$ network:

$$\hat{\epsilon}_{T+l} = \hat{\Sigma}_{T+l}^{-\frac{1}{2}} \left(W_{Q+1}\right) u_{T+l} \qquad l \in [1, \ldots, h] \tag{44}$$

Then, we can reconstruct the us and the proxy for the equally weighted index's conditional variance as if the realized network for the period of interest is the optimal one $W^{\star}$:

$$\tilde{u}^{\star}_{T+l} = \hat{\Sigma}_{T+l}^{\frac{1}{2}} \left(W^{\star}\right) \hat{\epsilon}_{T+l} \tag{45}$$

$$\mathbb{V}\text{ar}\left(\frac{1}{n}\mathbf{1}' y^{\star}_{T+l}\middle| \mathcal{I}_{T+l-1}\right) = \mathbb{V}\text{ar}\left(\frac{1}{n}\mathbf{1}' u^{\star}_{T+l}\middle| \mathcal{I}_{T+l-1}\right) \simeq \left(\frac{1}{n}\mathbf{1}' u^{\star}_{T+l}\right)^2 \tag{46}$$

In this way, we can compare the obtained optimal volatility proxy with the realized volatility proxy, the latter being robust against model misspecification.

Finally, we highlight that the output of the previous optimization also includes the target exposures that can be helpful to policymakers in order to enforce claims redistribution in the financial system. Clearly, the quality and reliability of recommendations depend on the quality of data on which the model has been estimated.

3.5 Implications for Forecasting and Network Evaluation

Our results in terms of inferred networks, system variance decomposition and an optimal network of exposures are relevant and of interest even from a forecasting perspective.

First, the inferred networks can be determined over the parameters estimated on a rolling basis. Consequently, the monitoring of the evolution of the inferred networks could provide relevant insights on the future dynamics of the system. In fact, information on the system's evolution could be obtained by analyzing the sequence $\mathcal{W}_{t,n}^{m,p}$ with $m = A, B$, $p = L, R$, where t refers to the observed weight matrix available at time t and n identifies the size of the estimation window used to fit the spatial-BEKK model. If the spatial matrices evolve at a rate lower than that of the entities, the inferred networks enable the evaluation of the impact of parameter changes on the network structure.

Second, the decomposition of the risk in the system can be achieved in-sample, as well as in an out-of-sample analysis. In this last case, the decomposition of the system variance becomes central in the construction of optimal portfolios of the analyzed assets. The spatial-BEKK model can be used to provide forecasts of the conditional covariance matrix. From these, we can obtain forecasts of the system variance decomposition, thus obtaining

$$\mathbb{V}\mathrm{ar}\left(\frac{1}{n}\mathbf{1}'y_{t+1}\middle|\mathcal{I}_t\right) = \left(\sigma_{t+1}^{Constant}\right)^2 + \left(\sigma_{t+1}^{Direct}\right)^2 + \left(\sigma_{t+1}^{Indirect}\right)^2 + \left(\sigma_{t+1}^{Mixed}\right)^2 \tag{47}$$

In turn, these forecasts might form the basis for the construction of *optimal* portfolios in a forecasting perspective. Moreover, as the forecasts are functions of the network available at time t, impacts on network changes can also easily be obtained. In particular, if we highlight the dependence on the network, we can compute the following differences caused by the change from W_t to $W^\star$:

$$\mathbb{V}\mathrm{ar}\left(\frac{1}{n}\mathbf{1}'y_{t+1}\middle|\mathcal{I}_t, W_t\right) - \mathbb{V}\mathrm{ar}\left(\frac{1}{n}\mathbf{1}'y_{t+1}\middle|\mathcal{I}_t, W^\star\right). \tag{48}$$

These differences can be further decomposed into the differences among the four components of the system variance. This also helps in the evaluation of the optimal target exposures. As an example, if we consider the network representation of the banking system, by comparing the actual network W_t with the target network $W^\star$ and looking at the variance decomposition, regulators could evaluate the total maximum impact they could achieve by moving from the actual design of the network to the optimal one. Moreover, they could decompose the (relative) advantage in the *diagonal*, *holding* and *overlapping* components. Finally, they could determine how minor or partial changes to the network could affect the system variance and what fraction of the maximum gain, that is the maximum variance reduction associated with the optimal network, they would achieve.

4 Empirical Analysis: The Example of European Sovereign Bond Risk Spillovers

To better clarify the advantages and potential of our methodology, we consider an application to publicly available data, and in particular we consider the European sovereign bond yields. We use two different data sources: (i) the changes in the ten-year sovereign bond yields 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, by taking the claims reported by the banking sector as a proxy for the claims of the entire country, we believe we achieve a good compromise, allowing us to evaluate the risk of the sovereign market while also accounting for the presence of interdependence among countries due to foreign claims. The aim of the 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 Greece, Italy, Ireland, Spain, France and Germany from 2006 to 2013, 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. Implicit in this interpolating choice is the assumption that foreign claims variation is much slower than the changes in bond yields. BIS consolidated banking statistics provide internationally comparable measures of national banking systems' exposures to country risk (McGuire and Wooldridge (2005)). Country risk refers to country-wide 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 residency 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 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 counterparty, 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. Due to the mixed nature of the data and the importance of the local banking system in international financial intermediation (see McCauley et al. (2010)), we consider these statistics as a good proxy for cross-country holdings. We expect that, if A reports a claim with B as a counterparty, investors will perceive the sovereign bonds of A to be dependent on the sovereign bonds of B in terms of the claim amount, and the same could be true for our matrix W_t. We report some summary statistics of BIS claims in billions of US Dollars in Table 3.

Table 3: BIS Ultimate Risk Basis Consolidated Banking Statistics: Foreign Claims Summary Statistics (Millions of US Dollars)

	mean						
	FR	DE	GR	IE	IT	PT	ES
FR	136550.2	47211.43	436.2857	7465.476	7329.333	1325.857	11209.24
DE	2890.476	10024.86	7991.143	0	1494.857	126.0476	54.57143
GR	111.0952	640.1905	0.666667	0	275.9048	0	11.66667
IE	8346.81	16033.9	90.7619	0	1799.429	321.1905	2823.143
IT	0	210087.9	1478.286	24707.67	42704.9	8145.429	50101.76
PT	404612.4	204933.5	584.6667	44803.76	0	5213.667	39768.9
ES	125015.8	159029.1	1839.857	15542.52	24742.67	7230.81	28595.71
	s.d.						
	FR	DE	GR	IE	IT	PT	ES
FR	78357.55	8143.641	400.2512	1943.93	2523.868	626.4275	4448.87
DE	768.4744	2428.127	1897.988	0	706.0261	62.18639	26.23656
GR	79.03791	190.3546	0.966092	0	190.2845	0	2.516611
IE	1734.736	2265.344	137.4681	0	754.3077	359.7015	739.2039
IT	0	25274.26	509.562	4681.422	9344.833	1442.989	15052.39
PT	107050.7	32002.6	255.3107	11389.74	0	1557.107	7562.398
ES	22050.09	15443.13	1549.368	4140.974	5260.584	3570.141	14442.01
	min						
	FR	DE	GR	IE	IT	PT	ES
FR	0	35141	61	5385	4211	411	4261
DE	1181	5691	4712	0	198	44	10
GR	8	412	0	0	6	0	8
IE	5554	13150	1	0	782	32	1458
IT	0	172867	673	16264	25578	5813	26523
PT	161227	153721	179	13054	0	3021	29986
ES:Spain	75710	131263	524	6284	14380	2228	17512
	max						
	FR	DE	GR	IE	IT	PT	ES
FR	310131	59015	1468	12680	11687	2494	19925
DE	4257	13973	10963	0	2493	253	87
GR	305	1354	3	0	589	0	17
IE	12007	20245	407	0	3416	1298	4116
IT	0	266302	2579	32451	56983	11778	77135
PT	531133	269532	1034	57326	0	8216	51376
ES	166332	188566	4825	22172	31599	12696	63684
	median						
	FR	DE	GR	IE	IT	PT	ES
FR	117696	46701	341	6738	6289	1185	11510
DE	2920	10771	8403	0	1788	103	54
GR	134	628	0	0	338	0	11
IE	8414	15754	9	0	1665	122	3057
IT	0	201625	1417	24883	44534	7849	49713
PT	446638	201532	606	46669	0	5232	39250
ES	126819	159520	1016	16366	25687	5183	23282

4.1.2 Normalization and Robustness

As discussed in Section 3.1, we consider several alternatives for the economic magnitude to be used for the normalization. Our final choice for 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. The other choices investigated for normalization were no 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[10]. We report some summary statistics of BIS claims normalized by the total claims in Table 4, and network representations for selected periods in the first column of Figure 1.

4.1.3 Sovereign Bond Yields

We use the daily changes in the ten-year yields of sovereign bonds, from 1/3/2006 to 12/30/2013, for France, Germany, Greece, Ireland, Italy, Portugal and Spain, as downloaded from Datastream. The choice of the ten-year maturity is due to data availability, in particular for the Greek bond. As can be seen from Table 5, the asymmetry of some series, in particular that of Greece, but also those of Portugal and Ireland, is striking. Correlation is high between specific pairs, namely, France and Germany, Spain and Italy, and Ireland and Portugal, highlighting the closeness between those economies. Despite all being positive, several correlations display relatively small values. Most interestingly, the smallest correlations are those between Germany and the other European countries (France excluded). Although a multivariate GARCH methodology is not sufficient for handling the big movements in the yield series, we consider it a good approximation for monitoring the risk evolution by accounting for network dependence, leaving for future research the explicit inclusion of jumps in the model.

[10]Estimation results for these alternatives are available upon request.

Table 4: BIS Consolidated Banking Statistics: Ultimate Risk Basis Foreign Claims Normalized by Total Claims by Reporting Country. Summary Statistics

	mean						
	FR	DE	GR	IE	IT	PT	ES
FR	0.000	0.072	0.018	0.020	0.123	0.009	0.050
DE	0.058	0.000	0.011	0.047	0.059	0.012	0.067
GR	0.018	0.023	0.000	0.007	0.009	0.001	0.003
IE	0.041	0.105	0.015	0.000	0.077	0.011	0.051
IT	0.054	0.269	0.007	0.028	0.000	0.010	0.034
PT	0.061	0.072	0.046	0.028	0.042	0.000	0.180
ES	0.049	0.046	0.001	0.016	0.036	0.064	0.000
	s.d.						
	FR	DE	GR	IE	IT	PT	ES
FR	0.000	0.010	0.005	0.004	0.017	0.002	0.004
DE	0.003	0.000	0.001	0.009	0.002	0.002	0.004
GR	0.009	0.010	0.000	0.005	0.007	0.001	0.002
IE	0.006	0.045	0.004	0.000	0.011	0.002	0.001
IT	0.017	0.102	0.004	0.012	0.000	0.006	0.010
PT	0.017	0.022	0.013	0.005	0.014	0.000	0.021
ES	0.011	0.009	0.000	0.004	0.003	0.002	0.000
	min						
	FR	DE	GR	IE	IT	PT	ES
FR	0.000	0.057	0.009	0.014	0.084	0.007	0.045
DE	0.054	0.000	0.009	0.034	0.055	0.010	0.057
GR	0.007	0.010	0.000	0.003	0.002	0.000	0.001
IE	0.034	0.054	0.012	0.000	0.066	0.007	0.048
IT	0.041	0.086	0.000	0.018	0.000	0.006	0.024
PT	0.041	0.033	0.026	0.020	0.024	0.000	0.135
ES	0.027	0.036	0.001	0.010	0.030	0.060	0.000
	max						
	FR	DE	GR	IE	IT	PT	ES
FR	0.000	0.086	0.024	0.027	0.144	0.012	0.057
DE	0.062	0.000	0.014	0.057	0.065	0.015	0.073
GR	0.036	0.050	0.000	0.019	0.024	0.003	0.009
IE	0.052	0.179	0.025	0.000	0.098	0.015	0.054
IT	0.086	0.353	0.012	0.052	0.000	0.022	0.054
PT	0.092	0.104	0.074	0.039	0.062	0.000	0.208
ES	0.063	0.069	0.001	0.022	0.042	0.068	0.000
	median						
	FR	DE	GR	IE	IT	PT	ES
FR	0.000	0.074	0.018	0.019	0.128	0.009	0.049
DE	0.059	0.000	0.011	0.050	0.059	0.011	0.068
GR	0.016	0.020	0.000	0.006	0.005	0.001	0.002
IE	0.039	0.078	0.014	0.000	0.073	0.009	0.051
IT	0.046	0.311	0.009	0.023	0.000	0.007	0.032
PT	0.053	0.075	0.044	0.028	0.037	0.000	0.185
ES	0.052	0.044	0.001	0.017	0.036	0.063	0.000

Table 5: Daily Changes in Ten-Year Sovereign Bond Yields from 1/3/2006 to 12/30/2013. Summary Statistics

	FR	DE	GR	IE	IT	PT	ES
mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000
s.d.	0.014	0.018	0.028	0.014	0.014	0.017	0.014
min	0.097	0.061	-17.405	-0.912	-0.562	-1.332	-0.984
Skewness	6.935	6.602	543.084	26.474	14.103	41.015	14.152
Kurtosis	-0.080	-0.132	-0.907	-0.194	-0.137	-0.265	-0.156
max	0.093	0.094	0.169	0.093	0.091	0.144	0.065
median	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlations							
		DE	GR	IE	IT	PT	ES
	FR	0.69	0.11	0.30	0.39	0.22	0.40
	DE		0.01	0.14	0.03	0.07	0.09
	GR			0.35	0.24	0.39	0.28
	IE				0.48	0.61	0.53
	IT					0.44	0.80
	ES						0.49

4.2 Parameter Estimation

In Table 6, we report QMLE results for the relevant parameters of the model. The estimation is obtained using a numerical constrained optimization in order to satisfy the ergodicity condition (23).[11] We have estimated three models: (i) a restricted diagonal model in which there is no network dependence, (ii) a left multiplication model that allows us to focus on fragile nodes of the network, and (iii) a right multiplication model that underlines the important nodes. As Table 6 shows, both the risk receivers (left multiplication) and risk spreaders (right multiplication) spatial models outperform the diagonal model and this is also formally demonstrated by the likelihood ratio test statistics reported in the table. Notably, the test strongly rejects the null, thus supporting the relevance of networks in variance spillover analysis. Our first important result confirms the relevance of foreign claims in explaining conditional covariances for bond yields. One

[11]Regarding the restriction on the sixth-order moment, we checked, after conducting the estimation, for the finiteness of the fourth moments, following Hafner (2003), but we are not aware of any closed form restriction for sixth-order moments and their derivation is outside the scope of the present paper. The Matlab-based estimation software we used is available upon request.

important aspect to consider is that the period under consideration is a turbulent one, and we cannot exclude that foreign claims are less important in normal times.

Table 6 shows the risk receiving propensity of Spain and Ireland in both the short and long run. The two countries have statistically significant coefficients in the network-related contributions, both when the network is included in the ARCH part of the model and when it affects the GARCH part. Moreover, even though the coefficients are not statistically different from zero, the different sign for Germany [12] evidences a diversifying role of this country in the covariance contribution with respect to the other countries, in both the short and long run. The relevance of Germany is also emphasized in the risk receivers (right multiplication) spatial model, for which it is the most important country in both the short and long run, with coefficients significant at the 1% level. This comes as no surprise, as the German Bund is the European benchmark against which spreads are computed. Italy comes second in this ranking, being important in the long and in the short run at 5%. Italy's relevance can be justified by the extent of its public debt, together with its economic relevance. Then, in the short run and with a lower significance (10%), we also see a role for Greece, despite the fact that it is usually recognized as the source of troubles. In our opinion, these results are mainly driven by the different economic magnitudes of these countries, and can be explained by the argument that the majority of the big swings in the Greek bond can be reabsorbed by the other countries, while small moves in the German and Italian bond markets greatly affect the behavior of other countries' bonds. There is no clear diversification pattern, aside from the fact that the covariance contributions among the important countries are clearly positive.

4.3 Inferred Networks

The estimation of the coefficients of the model provides a good picture of the relevant nodes in the claims network, which are important to monitor. However, it is possible, using the methodology outlined in Section 3.2, to build a graphical representation that

[12] We recall, here, that it is the relative sign that matters, as discussed in subection 2.1.

Table 6: Estimated Relevant Parameters of the Diagonal BEKK (Top Panel), Spatial risk receivers (left multiplication) BEKK (Central Panel) and Spatial risk spreaders(right multiplication) BEKK (Bottom Panel) Models on Daily Changes in the Ten-Year European Sovereign Bond Yields from 1/3/2006 to 12/30/2013

	a_0 DBEKK	b_0 DBEKK		
FR	0.209***	0.978***		
DE	0.222***	0.975***		
GR	0.228***	0.974***		
IE	0.202***	0.979***		
IT	0.201***	0.978***		
PT	0.214***	0.977***		
ES	0.204***	0.979***		
log-likelihood	51406.87			
	a_0 SBEKK L	a_1 SBEKK L	b_0 SBEKK L	b_1 SBEKK L
FR	0.260***	0.000	0.965***	0.000
DE	0.372***	-0.425	0.930***	0.149
GR	0.267***	0.163	0.963***	-0.044
IE	0.165***	0.216*	0.986***	-0.047**
IT	0.212***	0.086	0.977***	-0.027
PT	0.214***	0.096	0.976***	-0.026
ES	0.171***	0.388**	0.984***	-0.083**
log-likelihood	51646.81			
likelihood ratio	479.89***			
	a_0 SBEKK R	a_1 SBEKK R	b_0 SBEKK R	b_1 SBEKK R
FR	0.204***	0.000	0.979***	0.000
DE	0.232***	0.133***	0.973***	-0.034***
GR	0.286***	0.361*	0.958***	-0.100
IE	0.177***	-0.116	0.984***	0.019
IT	0.224***	0.398**	0.972***	-0.101**
PT	0.174***	-0.147	0.984***	0.020
ES	0.253***	0.294	0.968***	-0.067
log-likelihood	51615.64			
likelihood ratio	417.54***			

allows us to monitor directly the edges of the network, that is, the level of exposure between two specific countries. Some of the results for selected periods are shown in Figure 1. The complete representation is available upon request.

Given that in foreign claims networks the strongest linkages are among the major economies (Germany, France and Italy) and between Spain and Portugal, in the filtered network we can appreciate how the map changes after the inference step. The second column in Figure 1 represents the short-term response to shocks in the changes of the bond yields and shows that inference mainly magnifies the role of Spain in the short-run response. A counterintuitive effect is instead the long-run response that may appear to be a second-order effect. In this case, the figure shows that the most fragile country in the long run is Germany and not Ireland or Spain as we would have expected from the coefficient significance. This apparent contradiction comes from the fact that Germany has (i) the lowest p-value among the non-significant ones (0.3), (ii) a coefficient that is almost twice that of Spain, and (iii) strong claim relationships with the other countries. Only the combination of these three effects in the filtered network can hint at the possibility that Germany is fragile when it comes to long-term shocks, thus revealing the usefulness of this kind of representation. The fourth column emphasizes the short-term shock role of risk sources the. It appears to be the most crowded and this can be explained by the fact that, in this case, the magnifying glasses of inference work for the two countries that appear to have the strongest links in terms of foreign claims. In this case, Italy appears to be a bigger source of risk than Germany because of the larger size of its coefficient. The fifth column investigates the risk receivers effectiveness of persistence terms. Connections are not so strong in this case, indicating the presence of negligible long-term sources of risk. Finally, it is interesting to note that the claim network and consequently all the others remain, virtually, the same during the subprime crisis (2008-Q3) and the sovereign bond crisis (2010-Q2). This clearly justifies our decision, in the following, to consider for our target network exercise an estimation sample that includes the subprime crisis and to compute the forecast, on which we optimize, over the sovereign crisis quarter.

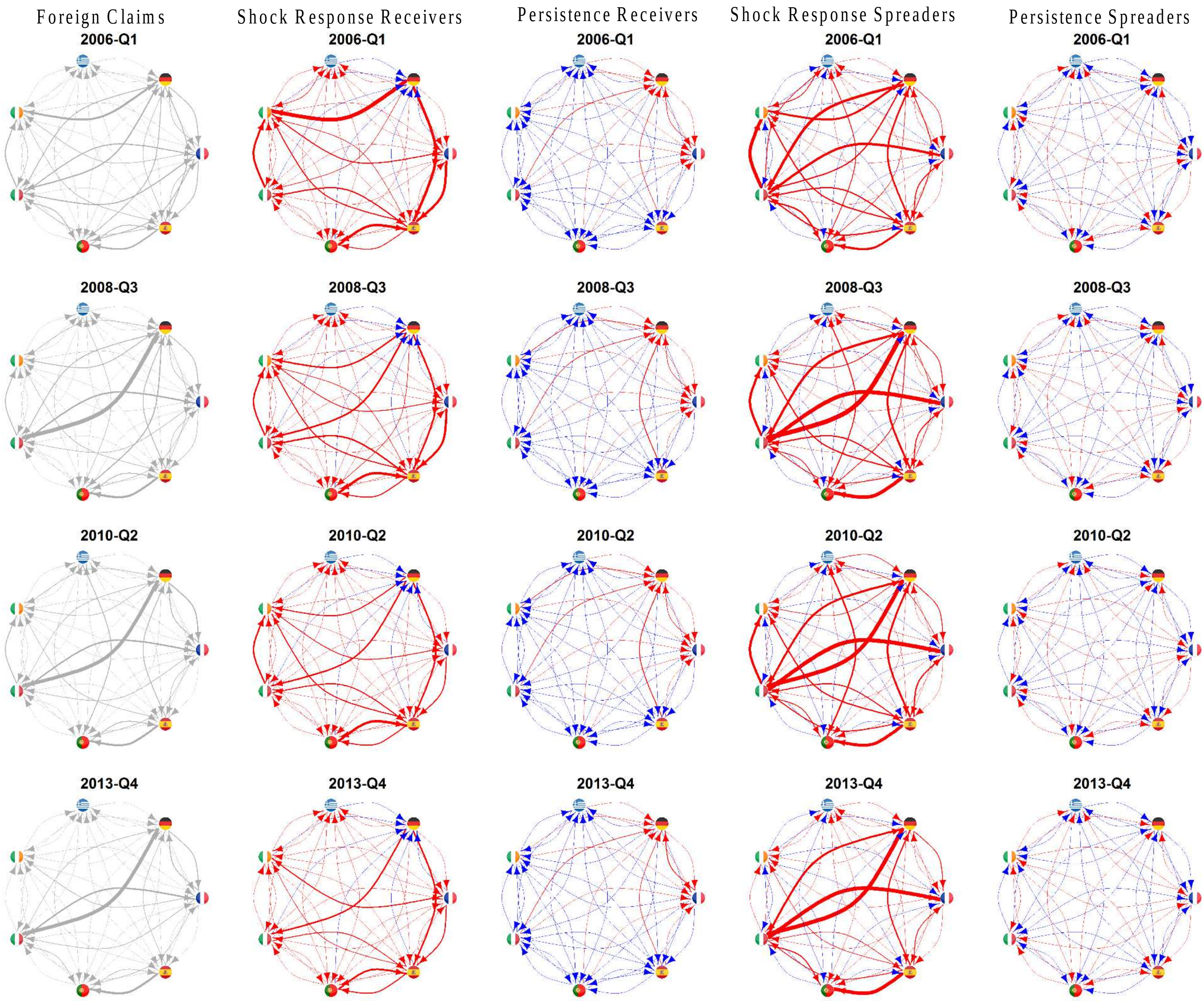


Figure 1: Selected graphical representation of the normalized foreign claims network w_t in the first column, the shock response risk receivers (left multiplication) networks $\mathcal{W}_t^{A,L}$ in the second column, the persistence risk receivers (left multiplication) networks $\mathcal{W}_t^{B,L}$ in the third column, the shock response risk spreaders (right multiplication) networks $\mathcal{W}_t^{A,R}$ in the fourth column and the persistence risk spreaders (right multiplication) networks $\mathcal{W}_t^{B,R}$ in the fifth column, all obtained from data on daily changes in the ten-year sovereign bond yield from 1/3/2006 to 12/30/2013.

4.4 Decomposition of Variance of Equally Weighted Index

To have yet another, point of view, and to show the flexibility of our analysis as an instrument of inquiry, we use the methodology of Subsection 3.3 on the variance decomposition of the risk receivers (left multiplication) and risk spreaders (right multiplication) model. Figure 2 and 3 reports the percentage of the system variance, constant, mixed and indirect contributions. Note the we do not include the direct one since, although being the biggest, it does not depend on the weights. In particular, we want to stress the presence, in tranquil periods, of negative (diversifying) contributions coming from the mixed part. The same contribution has instead, in turbulent periods, positive peaks leading to an increase of risk in the financial system. This is particularly evident for the risk spreaders case (fig.3) for which, during the second Greek bailout, the mixed contribution accounts for more than one fourth of the system variance. The indirect part is, in contrast, negligible, without noticeable diversification benefits. We also note that the relevance of the constant term is high and this could suggest, as we have already argued, that the model is only able to explain the dependence and variability in the data partially.

4.4.1 Estimated Target Exposures

Our methodology allows a proper ex-post analysis of spillover occurrences, in particular after relevant events, such as the default of a financial instituition. In this subsection, we show how our methodology could be of interest for regulatory interventions. In fact, if the inferred networks have an economic and financial motivation as spillover channels, as we showed previously for our bond yield example, the model allow to estimate the impact of the networks on the system variance. Therefore, it is possible to draw policy recommmendations from the estimated model, by focusing on the identification of the optimal (in terms of the risk of the system) network design. In particular, we propose to minimize the forecast path of the conditional system variance, looking for the optimal network structure according to the methodology outlined in Subsection 3.4. In principle, the regulator could then incentivize the achievement of such target exposures, obtaining

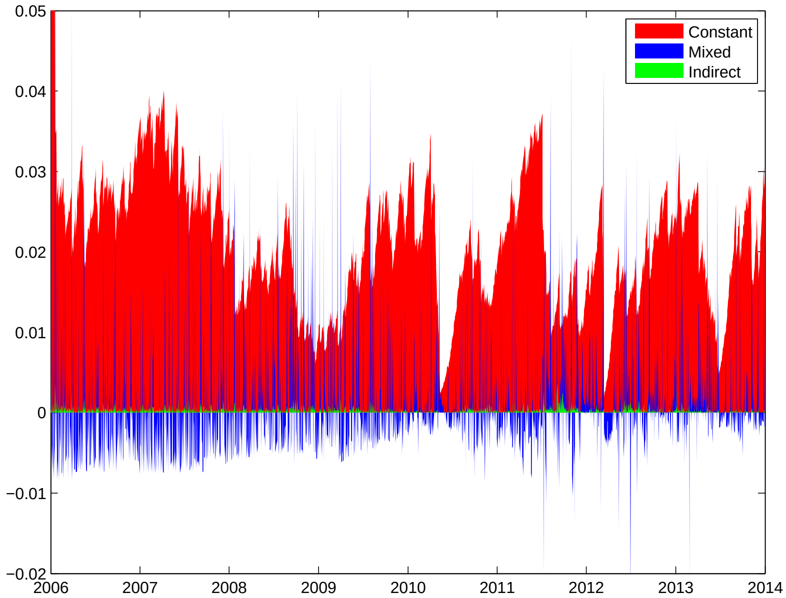


Figure 2: Relative variance decomposition, risk receivers (right multiplication) model, of the equally weighted index, obtained from data on daily changes in the ten-year sovereign bond yields from 1/3/2006 to 12/30/2013, with the direct contribution omitted

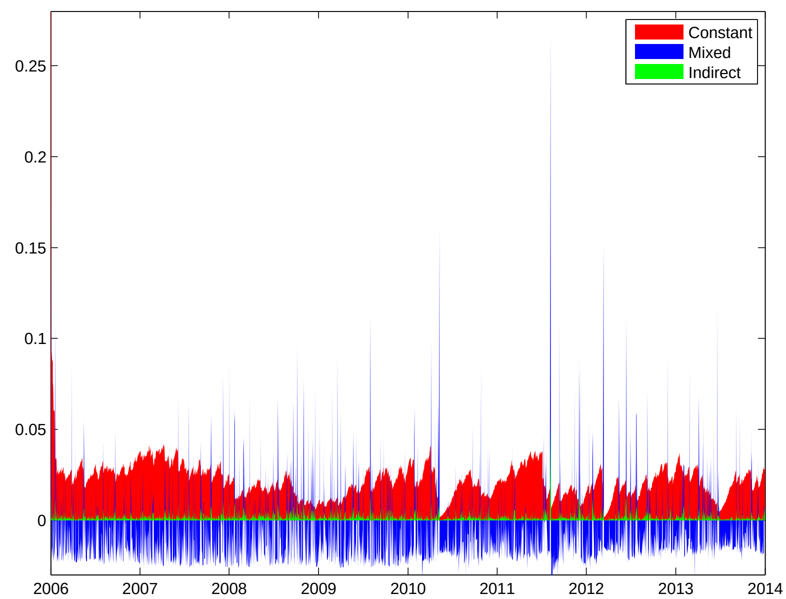
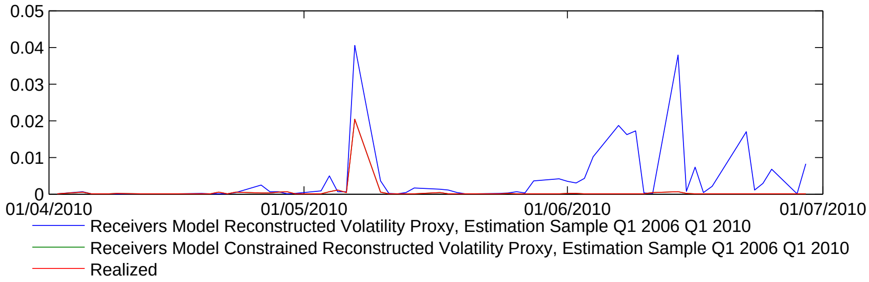


Figure 3: Relative variance decomposition, risk spreaders (right multiplication) model, of the equally weighted index, obtained from data on daily changes in the ten-year sovereign bond yields from 1/3/2006 to 12/30/2013, with the direct contribution omitted

a reduction of market volatility and a mitigation of risk.

As already discussed, we choose to optimize the forecast path in the quarter of the sovereign bond crisis (Q2 2010), while also including the subprime crisis in the estimation sample (from Q1 2006 to Q1 2010). Our results convince us that the spillovers channeled through foreign claims are the same on both occasions. We start by analyzing the reconstructed effect on the variance proxy of the equally weighted index when we change the network of the quarter Q2 2010, with one of the optimal networks coming from a different risk receivers(left multiplication), risk spreaders (right multiplication), constrained and unconstrained model. For the constrained model, equation (43) implies that there is only a redistribution of the claims among the considered countries; for the unconstrained model, the total amount of claims changes for each country.

Figure 4 shows the realized and reconstructed variance proxy of the equally weighted index during the sovereign debt crisis according to equations (45) and (46). Looking at the figure, we can conclude that the entities that were fragile are still playing the same role, but there is a significant reduction of the risk spread by systemically important entities, which generates a sensible reduction of the realized variance proxy in both the constrained and unconstrained cases, with differences among them that seem negligible. Table 7 shows that the optimal network in the fragile (left multiplication) constrained case is the same as the realized one, while the prescription coming from systemically important entities indicates that Portugal should have had larger cross-border exposures to Italy and Germany. In particular, considering the more realistic constrained case, Italy should have invested more across borders and Portugal should have received more investments from the other countries. To evaluate the feasibility of these redistributions, in Table 8 we report, for the constrained risk spreaders (right multiplication) case only, the differences in millions of US Dollars in the amounts needed to achieve the optimal network. In general, the sensible variance reduction we obtain from our calculation is implied by redistributions that are extreme and 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



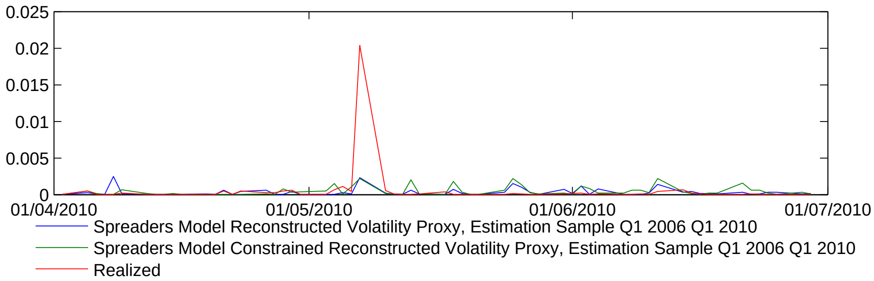


Figure 4: Variance proxy of the equally weighted index during the debt crisis of Q2 2010, obtained from data on daily changes in the ten-year sovereign bond yield.

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 equation (43).

5 Conclusions

This paper illustrates how financial networks can be efficiently integrated within a multivariate GARCH framework for risk analyses both in and out of sample. Our framework, which we refer to as spatial econometrics of risk, for its relation with both spatial econometrics and risk analyses, 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 of assets. Our work builds on the introduction of spatial methods into volatility models, as introduced by Caporin and Paruolo (2015). The model depends on proximity matrices that represent the economic distances among assets, and thanks to

Table 7: Target Exposures, Obtained from Data on Daily Changes in the Ten-Year Sovereign Bond Yield

True Exposures

Location Reporting	FR	DE	GR	IE	IT	PT	ES
FR	0.0000	0.0833	0.0177	0.0135	0.1314	0.0104	0.0510
DE	0.0659	0.0000	0.0105	0.0464	0.0515	0.0125	0.0608
GR	0.0145	0.0424	0.0000	0.0034	0.0041	0.0009	0.0050
IE	0.0346	0.0556	0.0144	0.0000	0.0735	0.0096	0.0484
IT	0.0387	0.2998	0.0060	0.0168	0.0000	0.0053	0.0297
PT	0.0557	0.0290	0.0738	0.0257	0.0258	0.0000	0.1720
ES	0.0211	0.0315	0.0007	0.0111	0.0262	0.0619	0.0000

SBEKK L (Delta wrt true)

Location Reporting	FR	DE	GR	IE	IT	PT	ES
FR	0.0000	0.0610	0.1262	0.1304	0.0167	0.1336	0.0938
DE	-0.0333	0.0000	0.0242	-0.0108	-0.0163	0.0232	-0.0275
GR	0.0199	-0.0073	0.0000	0.0307	0.0300	0.0332	0.0292
IE	0.0768	0.0504	0.0967	0.0000	0.0391	0.1013	0.0637
IT	0.1562	-0.0983	0.1887	0.1774	0.0000	0.1889	0.1652
PT	0.1364	0.1594	0.1187	0.1650	0.1650	0.0000	0.0250
ES	0.0503	0.0354	0.0706	0.0599	0.0455	0.0112	0.0000

SBEKK L Constrained (Delta wrt true)

Location Reporting	FR	DE	GR	IE	IT	PT	ES
FR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
DE	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
GR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
IE	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
IT	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
PT	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ES	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

SBEKK R (Delta wrt true)

Location Reporting	FR	DE	GR	IE	IT	PT	ES
FR	0.0000	0.0943	0.0604	0.0722	0.1250	0.0361	0.0626
DE	0.0548	0.0000	0.0619	0.0168	0.1419	0.0528	0.0382
GR	0.0202	-0.0060	0.0000	0.0316	0.0369	0.0357	0.0289
IE	0.0811	0.0728	0.0827	0.0000	0.0943	0.0827	0.0457
IT	0.1577	-0.0426	0.0374	0.0335	0.0000	0.0473	0.1180
PT	0.1293	0.2014	0.0309	0.0923	0.3019	0.0000	-0.0443
ES	0.0542	0.0515	0.0638	0.0399	0.0819	-0.0167	0.0000

SBEKK R Constrained (Delta wrt true)

Location Reporting	FR	DE	GR	IE	IT	PT	ES
FR	0.0000	-0.0172	0.0184	0.0231	-0.0400	0.0215	-0.0060
DE	-0.0229	0.0000	0.0249	-0.0112	0.0109	0.0217	-0.0234
GR	-0.0039	-0.0308	0.0000	0.0083	0.0093	0.0107	0.0063
IE	0.0041	-0.0131	0.0192	0.0000	-0.0187	0.0233	-0.0148
IT	0.0341	-0.2088	0.0512	0.0408	0.0000	0.0488	0.0339
PT	0.0067	0.0448	-0.0314	0.0201	0.0857	0.0000	-0.1259
ES	0.0040	-0.0046	0.0242	0.0110	0.0068	-0.0414	0.0000

Table 8: Investment Needed to Reach Target Exposures (Millions of USD), Obtained from Data on Daily Changes in the Ten-Year Sovereign Bond Yield

SBEKK R Constrained (Delta wrt true) Millions of USD							
Location Reporting	FR	DE	GR	IE	IT	PT	ES
FR	0	-55512	59560	74745	-129067	69570	-19296
DE	-68505	0	74248	-33318	32680	64723	-69828
GR	-519	-4107	0	1102	1245	1432	847
IE	2259	-7205	10523	0	-10233	12799	-8142
IT	29720	-181735	44571	35510	0	42457	29477
PT	919	6126	-4296	2751	11720	0	-17220
ES	5040	-5725	30357	13855	8499	-52026	0

their presence the model is able to describe and investigate spillover effects. In this work, we focus on proximity matrices that depend on financial/economic networks, and that allow us to capture the interdependence across the modeled variables. In an empirical example, we show that our methodology is suitable for dealing with a network of financial institutions, using both structural and descriptive analyses, as well as being suitable for policy purposes.

We make a number of contributions that go beyond the original contribution of Caporin and Paruolo (2015). We show how we can take advantage of the non-commutativity of matrices in modeling, and focus on both the risk receiving propensity (fragility) and risk spreading effectiveness (systemic importance) of spillovers. We show in the empirical application that our model is indeed able to give a reasonable description of European spillovers during the sovereign crisis, both in terms of country roles, through the significance of coefficients, and in terms of the network description of the events. We evidence the fundamental role of Ireland and Spain as risk receivers and the risk spreading effectiveness of Germany, Italy and, to a lesser extent, Greece as risk spreaders. In this respect, a natural evolution of the model would be to consider a bilateral multiplication model estimating left and right matrices jointly. A richer specification of this sort would have identification issues that would need to be dealt with, and we leave it for further research.

We also propose an interpretation of the right multiplication model, focusing on risk absorbers, in terms of portfolio composition, from which we derive a covariance decomposition that allows us to relate the holding and the overlapping of different portfolios to the conditional variance of the system. Finally, we propose a forecast-based methodology for computing target exposures that could be enforced by the regulator with the aim of reducing the volatility in the system. We are aware of the limits of this optimization and forecasting exercise but it could be considered an important tool for regulators to use to monitor financial stability. According to the recent review by Toniolo and White (2015) of the financial stability mandate across countries and across history, the principal interventions central banks took to maintain financial stability were liquidity provision and monitoring of the systemically important financial institutions. In our paper, we propose a new econometric tool with the ability to help the regulator fulfill the monitoring requirement once the bilateral exposure data of financial institutions have been collected and are available.

References

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

Franklin Allen, Ana Babus, and Elena Carletti. Financial crises: Theory and evidence. *Annu. Rev. Financ. Econ.*, 1(1):97–116, 2009.

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

Luc Anselin. Spatial econometrics. *A companion to theoretical econometrics*, 310330, 2001.

Kee-Hong Bae, G Andrew Karolyi, and Rene M Stulz. A new approach to measuring financial contagion. *Review of Financial studies*, 16(3):717–763, 2003.

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

Daniele Bianchi, Monica Billio, Roberto Casarin, and Massimo Guidolin. Modeling contagion and systemic risk. *Available at SSRN 2537986*, 2015.

Monica Billio and Massimiliano Caporin. Multivariate markov switching dynamic conditional correlation garch representations for contagion analysis. *Statistical methods and applications*, 14(2):145–161, 2005.

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 Panzica, and Loriana Pelizzon. The impact of network connectivity on factor exposures, asset pricing and portfolio diversification. *Mimeo*, 2015.

Dimitrios Bisias, Mark D Flood, Andrew W Lo, and Stavros Valavanis. A survey of systemic risk analytics. *US Department of Treasury, Office of Financial Research*, (0001), 2012.

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.

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

Olivier De Bandt and Philipp Hartmann. Systemic risk: a survey. *Working Paper Series 0035, European Central Bank*, 2000.

Olivier De Bandt, Philipp Hartmann, and José-Luis Peydró. Systemic risk in banking: an update. In A. Berger, P. Molyneux, and Wilson J., editors, *Oxford Handbook of Banking*. Oxford University Press, 2009.

Larry Eisenberg and Thomas H Noe. Systemic risk in financial systems. *Management Science*, 47(2):236–249, 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 Kenneth F Kroner. Multivariate simultaneous generalized arch. *Econometric theory*, 11(01):122–150, 1995.

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

Kristin Forbes and Roberto Rigobon. Measuring contagion: conceptual and empirical issues. In *International financial contagion*, pages 43–66. Springer, 2001.

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.

Christian M Hafner. Fourth moment structure of multivariate garch models. *Journal of Financial Econometrics*, 1(1):26–54, 2003.

Christian M Hafner and Arie Preminger. On asymptotic theory for multivariate garch models. *Journal of Multivariate Analysis*, 100(9):2044–2054, 2009.

Augusto Hasman. A critical review of contagion risk in banking. *Journal of Economic Surveys*, 27(5):978–995, 2013.

Sebastian Keiler and Armin Eder. Cds spreads and systemic risk: A spatial econometric approach. *Discussion Paper N.01/2013 Deutsche Bundesbank*, 2013.

Hsiang-Tai Lee and Jonathan K Yoder. A bivariate markov regime switching garch approach to estimate time varying minimum variance hedge ratios. *Applied Economics*, 39(10):1253–1265, 2007.

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 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.

Denis Pelletier. Regime switching for dynamic correlations. *Journal of econometrics*, 131 (1):445–473, 2006.

Marcello Pericoli and Massimo Sbracia. A primer on financial contagion. *Journal of Economic Surveys*, 17(4):571–608, 2003.

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.

Jean-Charles Rochet and Jean Tirole. Interbank lending and systemic risk. *Journal of Money, Credit and Banking*, 28(4):pp. 733–762, 1996.

Julia Schaumburg, Francisco Blasques, Siem Jan Koopman, and Andre Lucas. Spatial gas models for systemic risk measurement. *Mimeo*, 2014.

Gianni Toniolo and Eugene N White. The evolution of the financial stability mandate: From its origins to the present day. Technical report, National Bureau of Economic Research, 2015.