Crisis-Contingent Dynamics of Connectedness: An SVAR-Spatial-Network "Tripod" Model with Thresholds (Job Market Paper)

By Hang Sun [*]

In recent years a growing number of methodologies have been proposed to empirically measure the connectedness among financial entities. However, few of them attempt to profile the dynamics of the financial connectedness. This paper aims to fill this gap and proposes a novel and systematic model to portray not only the connectedness in a given regime but also its transitions across different regimes. The model is based on an improved version of a "tripod" model, which unifies structural vector auto-regressions (SVARs), spatial models, and network models under one framework. I introduce a transition mechanism into my model, thus making it possible to observe how the interconnections among financial entities vary with certain threshold variables. My model may be applied to various issues regarding the financial and non-financial interconnections among certain entities. As an illustration, I show how my model can shed some new light on the modeling of the recent Eurozone contagions. I reveal a clear causal map of the propagation of shocks among the stock markets in five selected Eurozone countries in both the contagion and non-contagion periods, which are determined automatically by a CDS index. I confirm the existence of contagion among the stock markets in these countries in recent years, and reject the hypothesis that the contagion is solely conducted by the cross-border interbank channel. And the unique roles played by each selected market under both the contagion and non-contagion regimes are efficiently summarized by network analyses.
JEL: C30;C59;G01;G15
Keywords: connectedness, SVAR, spatial weight, network, contagion, financial integration

I. Introduction

Economic entities are often interconnected and these interconnections can have significant economic consequences. Particularly, the interconnections among financial entities, whether they are asset markets, firms, financial institutions, or economies, have drawn much attention since the 2008 financial crisis. A growing

[*] Sun: Maastricht University School of Business and Economics, P.O. Box 616, 6200 MD, Maastricht, The Netherlands, h.sun@maastrichtuniversity.nl. The usual disclaimer applies.

number of methodologies have been proposed to empirically measure the financial connectedness, e.g., Billio et al. (2012), Diebolda and Yılmaz (2014), Giraitis et al. (2016), and Barigozzi and Brownlees (2016). Following this strand of literature, Scidá (2016) finds that financial connectedness can not only be depicted with network models but also with spatial econometric models, and both approaches are unified by an underlying structural vector auto-regression (structural VAR, SVAR) model. In this paper, I call her model a "SVAR-spatial-network" model, or a "tripod" model, as it is an integration of these three econometric models. However, most of these existing literature only tries to depict a static state of the financial connectedness and seldom attempts to model the dynamic evolution of the connectedness of a financial system. As an exception, Giraitis et al. (2016) indeed considers a dynamic model, but unlike most other literature, they assume the connections among the entities are known a priori, which is not realistic for many applications.

This paper contributes to the literature by proposing a novel and systematic way to empirically depict and analyze the connectedness among economic or financial entities in a given regime *and* its transitions across different regimes. Thus, the *dynamics* of connectedness is portrayed in a regime-switching model. With this model, we can figure out how the interconnections in an economic or financial system are decided by some threshold variables that we are interested in, and make comparisons among the interconnections under different regimes.

The model in this paper is rooted in the tripod model of Scidá (2016). Since it contains a regime-switching mechanism that is described by a threshold function, my model may be called a SVAR-spatial-network model with thresholds, or a "*threshold* tripod" model. As a secondary contribution of this paper, the threshold tripod model improves the original model of Scidá (2016) in various ways, on top of the dynamic regime-switching mechanism. First, Scidá (2016) does not give a satisfactory solution to the identification problem of a SVAR model, which is never "an easy problem to address" in any SVAR literature. In this paper, I propose to solve the identification problem in the SVAR part, which is fundamental to the whole tripod model, with a novel strategy that exploits the normal-mixture structure of the residuals. Second, I show that the tripod model actually provides a benchmark to test a broad class of interesting hypotheses on the spatial weight matrices. And third, since the networks implied by the tripod model in this paper are usually complete and often have negative edges, I suggest that such networks can still be analyzed by some novel approaches devised recently.

The threshold tripod model should be particularly useful for the applications where it is as important to profile transitions of connectedness as the connectedness per se. A typical example is the empirical modeling of financial contagions: a large body of literature agrees that a definition of contagion should emphasize the change of entity interrelationship between crisis and non-crisis periods, rather

than the relationship in each period.[1] In fact, as an illustration of my model, I show how it sheds some new light on the topic of financial contagions in the EU, which has become a hotly discussed topic since the start of the Eurozone crisis in late 2009. While many existing studies rely on manual choice of crisis start points and end points, I pick a European CDS index as the threshold variable and let the threshold variable determine two distinctive regimes where the five selected countries, namely Germany, Italy, Spain, Ireland, and Greece, have different interconnections. Under each regime, the mutual connectedness among the five countries is portrayed clearly. Furthermore, I construct statistical tests of two hypotheses. The first test confirms the existence of contagion among the countries; and the second test demonstrates that the contagions are not solely spread through the cross-border interbank channel. The existence of contagion is reconfirmed by global network measures. Moreover, centrality analyses and clustering analyses reveal clearly the distinctive roles played by the selected countries under different regimes.

Besides the contagion among markets in different countries, my model may be applied to profiling the transitions of many other instances of financial interconnections as well. For instance, many studies focus on the interconnections among financial institutions or firms (e.g. Diebolda and Yılmaz, 2014, Barigozzi and Brownlees, 2016). With my model, we may clearly see how these interconnections among institutions or firms change in different situations. And as another interesting application of my model, we may also try to see how the spillover effect across different asset markets shifts. As an example, the volatility transmission among different commodity markets is decided by various macroeconomic and financial factors and is known to shift significantly in different environment (Nazlioglu et al., 2011), it might be interesting to draw a dynamic sketch of it over time with the model in this paper.

Applications of the threshold tripod model are certainly not limited to the finance area. In various fields of economics, we encounter issues concerning interconnections of entities . For example, there are a large number of studies of the connections among firms or regions through which knowledge spillover takes place(e.g. Lee, 2006, Maliranta et al., 2009). It is possible to document the transitions of such connections as well.

A not-so-technical overview of the threshold tripod model is presented in the following section. And the rest of this paper is organized as follows. Section III reviews the basics of an SVAR model and shows how the T-SVAR model used in the following sections is identified. Section IV introduces the restricted SAR model and the way to construct tests based on it. Section V presents the network analyses. Section VI reports the empirical results of the application of the model in the Eurozone markets. Section VII concludes this paper. And in the end, the appendices collect the necessary mathematical proofs.

---

[1]See, e.g., Forbes and Rigobon (2002), Dungey et al. (2005), Chiang et al. (2007), Caporin et al. (2013). For an in-depth discussion, see Section VI.A.

II.   A Not-So-Technical Overview of the Threshold Tripod Model

Before I start to introduce the model formally, I present an overview of it in a descriptive language here.

As its name suggests, the SVAR-spatial-network model is composed of three steps. First, the entities of interest form an SVAR model. As we identify the SVAR, the responses of the entities to the structural shocks with clear economic interpretations are revealed. This provides a baseline for the following steps. Second, by adding some extra restrictions to the SVAR model, we get a spatial auto-regression (SAR) model, which shows clearly how the entity interconnections implied by the SVAR structure. And third, we use certain network analytical techniques to further analyze the connectedness shown by the SAR model. Since the tripod model in this paper has thresholds that model the dynamics of the interconnections, I add threshold functions to both the SVAR and SAR steps, and thereby upgrade them to threshold SVAR (T-SVAR) and threshold SAR (T-SAR) models. In this section, I try to introduce these steps one by one.

A.   The T-SVAR Step

Now it has become a popular practice to start inferring the structure of connectedness of a system from a VAR-type model. Most of the existing literature, such as Billio et al. (2012) and Diebolda and Yılmaz (2014), uses reduced-form VARs. In contrast, the tripod model starts from a *structural* VAR model. Compared to reduced-form VARs, a SVAR allows the variables in the system to react instantaneously to shocks from elsewhere, which is closer to reality if the frequency of observations is not so high that the shocks and responses cannot be separated temporally.

In this paper, I add a threshold function to the SVAR model, thus upgrading it to a T-SVAR model. Regimes are decided by the values of the threshold function that maps one or more exogenous variables to the probabilities of each state. There is no requirement of the form of the threshold function except that its value must be a probability, i.e., a real number ranging from zero to one. Therefore, we can pick all kinds of threshold functions of convenience, which can be transilient (similar to Hansen, 1997), smooth (see van Dijk et al., 2002, for a survey), or even more complex forms. And there is no restriction on the number of threshold variables. This feature of my model can in fact be employed to construct a regime indicator based on the shift of variable connectedness. If we scratch out the structural part, in fact the remaining reduced-form VAR part of the model will similar to the threshold VAR models widely used in empirical studies such as Li and St-Amant (2010) and Afonso et al. (2011).

Compared with reduced-form VARs, there is always an identification problem accompanying SVAR models, as the data do not provide enough information to estimate all the parameters in an SVAR, unless we impose some additional assumptions on the model a priori. With proper prior assumptions, all the pa-

rameters of an SVAR can be estimated, and then we call the SVAR "identified". Econometricians have proposed a number of identification strategies, but common strategies such as identifications based on recursive restrictions or sign restrictions, either require very strong prior assumptions that are rarely justified in reality, or does not result in unique identifications (see Kilian, 2013, for a review). Scidá (2016) suggests to use a machine-learning algorithm to select an optimum set of restrictions, but the restrictions are still in the recursive form, which is often overly strong in reality.

In my model, the SVAR is identified by assuming that the residuals follow normal-mixture distributions.[2] I assume the simplest case, where the normal-mixture distribution of the reduced-form residuals under each regime follow a mixture of two normal distributions with different covariances and zero mean. This is *not* a stronger assumption than the regular assumption of normal-distributed residuals. Rather, it is an even *weaker* assumption than the regular assumption. An ordinary normal distribution is a special case of a mixture-normal distribution and mixture distributions can be used to approach random distributions in the real world (e.g. Kon, 1984, on distributions of asset returns). In an SVAR model with normal-mixture distributed residuals, Lanne and Lütkepohl (2010) show that we can achieve the unique local identification without any prior restriction on the interrelationship among the variables like traditional identification strategies. Indeed, this only leads to local rather than global identifications and therefore we cannot know which variable is responding to which shock merely from identification results.[3] But fortunately this problem can still be handled in actual empirical applications. In this paper, I follow Kohonen (2013) and approach full identifications of the T-SVAR model by referring to some external source of information.

B.   The T-SAR Step

In the tripod model, the connectedness is recorded by the spatial weight matrix in a SAR model. The spatial weight matrix records the inverse distance between each pair of entities in the system, and the inverse distance can either be an inverse of the real geographical distance, or some closeness in the figurative sense, such as the amount of trades. There is in fact no substantial diffeeence between the "inverse distance" and the "connectedness" discussed in this paper. Unlike traditional spatial modeling where the user must designate a spatial weight matrix (e.g. Tonzer, 2015), the spatial weight matrix in the tripod model becomes endogenous and can be estimated from the SVAR identification obtained in the previous SVAR step. This step can be compared with some recent attempts to

---

[2]In order to understand what is a normal-mixture distribution, we can imagine the return of a certain stock follows a different distribution in different situations in the future, and each situation takes place in the future at given probabilities. Since we are not sure which situation will actually take place, the distribution of the return of the stock conditioned on present information will be a mixture of the normal distributions in different situations, or a "normal-mixture" distribution.

[3]This problem is called "label switching" problem by Maciejowska (2010). She offers a detailed proof of it in the appendix.

infer spatial weight matrices from data such as Manresa (2015) and Lam and Souza (2015).

Specifically, the SAR in the tripod model is obtained by imposing some additional linear constraints on the estimation of the SVAR. This step is carried out with a "minimum distance (MD)" approach, which is simple to understand, and easy to implement. Since I consider tripod models with thresholds in this paper, a similar MD approach can also be employed to obtain a threshold SAR (T-SAR) model from the T-SVAR estimation.

Moreover, I show that the thresholds in the threshold tripod model enable us to construct tests of many interesting hypotheses. I focus on two types of tests in this paper. First, I test whether the spatial structure contained in an economic system is identical across different regimes. Second, I test whether the difference between spatial structures under different regimes follows a certain pattern. Since spatial weights are a measure of the connectedness of the system, the two tests are actually equivalent to tests of the existence of the transition of connectedness, and the way of the transition of the connectedness. Tests on other types of restrictions are also possible. For example, it is possible to test whether a given spatial weight matrix is misspecified or not.

C.   The Network Analysis Step

In this step, we can further explore the entity interconnections revealed by the T-SAR model using some network analysis techniques. If we recall that a network is just defined as a group of entities and the connectedness among them, then the system we study can be regarded as a blend of different networks under different regimes and the transitions among the regimes can be regarded as the transitions among different networks.[4]

However, the networks generated by the tripod model in this paper are characterized by several quite unusual traits, including non-unitary edge[5] weights, directed edges, and often negatively weighted edges. Furthermore, the networks are usually complete, which means every pair of nodes in the network is to some extent connected. These traits bring some difficulty to potential applications of classical network analysis techniques as many of them are actually devised to handle common incomplete networks with undirected edges with positive unitary weights . Despite the difficulty, I show that such unusual networks can still be handled by some novel methods that have been devised recently. First, there are the simple strength measure and the PN measure for the centrality analyses (see Everett and Borgatti, 2014), i.e., the analysis of how important is the role of a certain node in the networks; second, McAssey and Bijma (2015) propose a

---

[4]This kind of network blends is also called "overlapping networks" (Battiston et al., 2013) or "overlay networks" (De Domenico et al., 2013). Kivelä et al. (2014) provides a detailed review of the broader class of multilayer networks.

[5]Usually a network is composed by "nodes" and "edges." Each node represents an entity in the network and an edge represents a connection between the two nodes on its two ends.

novel clustering measure for complete and weighted networks, with which we can measure how close are the neighbors of a node connected mutually and to how close the whole network is interconnected; and finally, Scidá (2016) suggests that when we estimate the SAR model, we automatically obtain a global measure for the influence strength of the whole network, which is called the ρ-measure. Scidá (2016) also proposes many other analyses, but most of them can only be applied to networks with all positively weighted edges. Since I try to provide some analyses that still work even when there are negative edges, I switch to a new set of analyses and only keep the ρ-measure in this paper.

III.   SVAR and T-SVAR Models

Before introducing my model formally, I introduce the notations used throughout the remaining parts of this paper. I_a is the identity matrix of order a. O_a is the $a \times a$ matrix where all the elements are zeros. Both AB and $A \times B$ are matrix products of matrices A and B. $\mathbf{1}_a$ is the a-dimensional vector of ones, and $\mathbf{0}_a$ is the a-dimensional vector of zeros. $A \otimes B$ is the Kronecker product of matrices A and B. $A \odot B$ is the Hadamard product[6] of A and B. $\text{diag}(a_1, a_2, \dots)$ is the diagonal matrix where diagonal elements are $a_1, a_2, \dots$, $\text{diag}(X)$ is the diagonal matrix where diagonal elements are respectively the elements of vector X, and $\text{diag}(A_1, A_2, \dots)$ is the block diagonal matrix where $A_1, A_2, \dots$ are on the main diagonal. $\text{vec}(A)$ is the standard vectorization of matrix A.

A.   A Review of SVAR Models

In this section, I present a brief review of the SVAR model, which is the basis of the foundation of the SVAR-spatial-network "tripod" model.

As the main advantage of an SVAR model, it allows us to capture the instantaneous responses of variables to shocks. For instance, if an SVAR system at a daily frequency is composed of two stock market indices, then the SVAR model can reveal how one market responses to a market-specific shock in the other market on the same day. This feature of the SVAR model is especially valuable for financial issues, as many financial variables such as asset prices respond to shocks very promptly. Even for the economic systems where variables respond slowly, this can still be very useful, because many data can only be obtained at an even lower frequency. There responses provide very useful clues for the inference of connectedness of the system.

In the rest part of this paper, I assume $Y_t = (y_{1t}, y_{2t}, \dots, y_{Kt})'$ is a K-dimensional vector representing the endogenous variables observed at time $t = 1, 2, \dots, T$, and $X_t = (x_{1t}, x_{2t}, \dots, x_{Mt})'$ is a vector with the M dimensions representing exogenous variables observed at respective time t. Additionally, let $\{u_t\}_{t=1}^T$ be the K-dimensional error term, then the *reduced-form* model with p lags studied in this paper can be defined as $Y_t =$

---

[6]This is also known as the Schur product, or entrywise product.

$\phi(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}, X_t, u_t)$.[7] Correspondingly, the *structural-form* model can be specified as $\psi(Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}, X_t, u_t) = 0$.

In the baseline SVAR model, let the function ϕ be linear and u_t be a white noise. Then the reduced-form VAR model becomes

$$Y_t = c + \sum_{i=1}^{p} A_i Y_{t-i} + A_X X_t + u_t,$$

where c is a K-dimensional constant term, $A_1, A_2, \ldots, A_p$ are autoregressive coefficient matrices with size $K \times K$, A_X is the exogenous coefficient matrix with size $K \times M$, and $u_t \sim (0, \Sigma_u)$. Obviously, this reduced-form model is equivalent to various structural form models, and in the SVAR literature, we consider the simple case, where the function ψ is linear and additively separable, i.e.,

$$\mathsf{A}Y_t - \mathsf{A}(c + \sum_{i=1}^{p} A_i Y_{t-i} + A_X X_t + u_t) = 0,$$

where A is a invertible $K \times K$ matrix. Hereinafter, I call A the "structural matrix". Defining $\epsilon_t \equiv \mathsf{A}u_t$, then the model can be rearranged as

(1)
$$\mathsf{A}Y_t = \mathsf{A}c + \sum_{i=1}^{p} \mathsf{A}A_i Y_{t-i} + \mathsf{A}A_X X_t + \epsilon_t.$$

It is easy to see that $\epsilon \sim (0, \mathsf{A}\Sigma_u \mathsf{A}')$. If we further require ϵ_t to be orthogonal such that its covariance matrix Σ_ϵ is diagonal, then the model in equation (1) can be called an SVAR model. A diagonal Σ_ϵ indicates that all pairs of components in the error term ϵ_t have zero correlation, therefore the shocks to the economic system represented by the SVAR model are uncorrelated, too. Such shocks are called "structural shocks". Equation (1) is in fact the "A-model" in Lütkepohl (2007) plus some exogenous variables.

In many papers and textbooks about SVAR, the structural matrix A is chosen to let Σ_ϵ be an identity matrix I_K. Under this convention, the structural shocks all have the same size, which provides empirical studies with much convenience in comparing the effects of different shocks. However, since in this paper I do not emphasize on comparing the effects of economic shocks, I pick another way of normalization that is more convenient for further research in the following sections. I require all the diagonal elements of A to be one, but do not restrict the scale of Σ_ϵ. It is easy to see that my normalization is equivalent to the conventional normalization: if A is a structural matrix following my rule of normalization, and $\Sigma_\epsilon = \mathrm{diag}(\sigma_{11}, \sigma_{22}, \ldots, \sigma_{KK})$, then $\mathsf{A} \times \mathrm{diag}(\sigma_{11}^{1/2}, \sigma_{22}^{1/2}, \ldots, \sigma_{KK}^{1/2})$ is obviously a structural matrix following the conventional rule of normalization such that Σ_ϵ is

[7]Surely, it is possible to make Y_t dependent on lags of X_t. In this case, we can stack the lags of X_t in one vector and this definition still applies.

an identity matrix. The most significant advantage of my way of normalization can be shown by rearranging equation (1) as

$$(2) \qquad Y_t = GY_t + \mathsf{A}c + \sum_{i=1}^{p} \mathsf{A}A_i Y_{t-i} + \mathsf{A}A_X X_t + \epsilon_t,$$

where $G \equiv I_K - \mathsf{A}$. Since all diagonal elements of A have been set to one, the diagonal elements of G will be zero. This will be the most convenient form to display an SVAR model in the next section, but within this section, I still focus on the specification in the form of equation (1).

There are two groups of unknown parameters in the SVAR model in equation (1). The first group, $\{c, A_1, A_2, \ldots, A_p, A_X, \Sigma_u\}$, is inherited from the reduced-form VAR; and the second group, $\{\mathsf{A}, \Sigma_\epsilon\}$, contains the additional parameters that are exclusive to the structural form model. Now let us assume that we have anyway obtained a proper estimation of the first group of parameters,[8] then we will only need to find a way to get an estimation of A and Σ_ϵ to fully pin down the structural model. If we can find such an estimation, we say the SVAR model is identified.

However, the identification step is never easy. Straightforwardly, if we assume the distribution of u_t is completely decided by Σ_u and there is no restriction on the free elements in A and Σ_ϵ, then the SVAR model is not identifiable. The reason is intuitive: A and Σ_ϵ are regulated by the equation

$$\Sigma_\epsilon = \mathsf{A}\Sigma_u \mathsf{A}'.$$

Since A is invertible, this equation can be rewritten as

$$(3) \qquad \Sigma_u = \mathsf{A}^{-1}\Sigma_\epsilon \mathsf{A}'^{-1}.$$

Since Σ_u is a covariance matrix and therefore symmetric, it possesses $K(K+1)/2$ free elements. But in the right hand of the equation, A has $K^2 - K$ free elements and Σ_ϵ has K free elements. Putting them together, the right side of equation (3) requires K^2 free elements, whereas the left side can only offer $K(K+1)/2$. Hence, there will be infinite number of solutions to the equation, and the SVAR model fails to be identified. Lütkepohl (2007) also offers a formal rank condition of the identifiability of the SVAR model discussed here.

The traditional wisdom to solve the problem is to cut the number of free parameters in the right side of equation (3) by imposing additional restrictions. As the most straightforward way to do this, we may simply set the lower- or upper-triangle of A to be zero, and thereby A and Σ_ϵ can be immediately obtained by the means of a Cholesky decomposition. However, this is a very strong set of restrictions, as it assumes a recursive causal chain among the variables. In other

---

[8]The estimation methods of reduced-form VARs can be found in Lütkepohl (2007).

words, it assumes we know that among the K variables, Variable A is known to be instantly affected by all K structural shocks, Variable B by certain $K - 1$ structural shock, Variable C by certain $K - 2$ structural shocks, etc. But in practice, it is really difficult to predicate the such relationships among the variables. Scidá (2016) suggests a machine learning algorithm called "PC algorithm" to find the optimal recursive order of the variables[9], however, we still have to make sure that the recursive causal chain is really there among the variables before we apply the algorithm. Besides the recursive restrictions, there are other identification strategies that rely on a priori restrictions, and most of them suffer from similar problems in imposing the restrictions, as is echoed by Kilian (2013). Such problems are quite undesirable, as the very purpose of the tripod model is to reveal the connectedness among variables or entities without knowing much about their interrelationship a priori. In this paper, the identification problem is handled in a novel way, by exploiting the structure of Σ_u and extracting more information from it, which enables us to obtain a unique identification without imposing strong a priori restrictions on the model. This methodology for the SVAR model is described in the following part, and in Section III.C I show that the generalized case of the SVAR model with thresholds can be also identified with it.

B. Identification of an SVAR Model with Normal-Mixture Residuals

Although the main purpose of this paper is to establish a model describing the transitions of connectedness, which is based on a SVAR model with thresholds, in order to illustrate how the identification strategy works, I first introduce how a non-threshold SVAR model with normal-mixture residuals is identified by exploiting the distribution structure of residuals. The idea is pioneered by Lanne and Lütkepohl (2010), who show normal-mixture residuals can be utilized to construct tests of given identification results.

From previous discussion, we have already seen that an SVAR is identified by finding a proper "structure", which combines the orthogonal structural shocks into the overt reduced-formed innovations. A mixture distribution can be regarded as a mixture of several components of simple distributions. The SVAR model is identified, if we assume that the "structure", i.e., the way that the orthogonal shocks are combined into reduced-form innovations is the same for all the components.

Consider the reduced-form VAR model

$$Y_t = c + \sum_{i=1}^{p} A_i Y_{t-i} + A_X X_t + v_t,$$

---

[9]For details of the PC algorithm, see Spirtes et al. (2000, p.117–119).

where

$$v_t = \begin{cases} v_{1t} \sim N(0, \Sigma_1) & \text{with probability } \gamma, \\ v_{2t} \sim N(0, \Sigma_2) & \text{with probability } 1 - \gamma, \end{cases}.$$

Now the reduced-form error term v_t follows a mixture of two normal distributions with zero mean and different covariance matrices. If $\Sigma_1 = \Sigma_2$, or $\gamma = 0, 1$, then the normal-mixture distribution reduces to a regular normal distribution. The covariance matrix of v_t is also the weighted combination of the covariance of its two components, $\gamma\Sigma_1 + (1 - \gamma)\Sigma_2$. Proposition 1 in Lanne and Lütkepohl (2010) demonstrates that given the normal-mixture distributed v_t, there exists a diagonal matrix $\tilde{\Lambda} = \text{diag}(\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_K)$ with $\tilde{\lambda}_i > 0$ for all $i = 1, 2, \ldots, K$ and an invertible matrix $\tilde{\Omega}$ such that $\Sigma_1 = \tilde{\Omega}\tilde{\Omega}'$ and $\Sigma_2 = \tilde{\Omega}\tilde{\Lambda}\tilde{\Omega}'$. Furthermore, if every pair of diagonal elements of $\tilde{\Lambda}$ is different, the decomposition $\tilde{\Omega}$ is unique up to the inversion of the signs of all elements in the same column.[10] Suppose the diagonal elements of $\tilde{\Omega}$ are $(\tilde{\omega}_{11}, \tilde{\omega}_{22}, \ldots, \tilde{\omega}_{KK})$, due to my rule of normalization, I further rewrite the decompositions of covariances as $\Sigma_1 = \Omega\Lambda_1\Omega'$ and $\Sigma_2 = \Omega\Lambda_2\Omega'$, where Ω has all-one diagonal elements, $\Lambda_1 = \text{diag}(\tilde{\omega}_{11}^2, \tilde{\omega}_{22}^2, \ldots, \tilde{\omega}_{KK}^2)$ and $\Lambda_2 = \Lambda_1\tilde{\Lambda}$. Then the covariance matrix of v_t can thus be written as

$$(4) \qquad \Sigma_v = \gamma\Omega\Lambda_1\Omega' + (1 - \gamma)\Omega\Lambda_2\Omega' = \Omega(\gamma\Lambda_1 + (1 - \gamma)\Lambda_2)\Omega'.$$

Suppose that this reduced-form model is equivalent to the SVAR model specified by equation (1) so that

$$(5) \qquad \Sigma_v = \mathsf{A}^{-1}\Sigma_\epsilon\mathsf{A}'^{-1}.$$

Comparing equation (5) with (4), we find

$$\mathsf{A} = \Omega^{-1}.$$

The SVAR model is identified up to this moment. Indeed, we may still invert the column signs of Ω arbitrarily, but this is no problem in practice, as we can always distinguish between a positive shock and a negative shock easily.

The parameters in this SVAR model can be locally estimated by maximum likelihood (ML) estimation. Denote $\epsilon = \{\epsilon_t\}_{t=1}^T$. Collecting all the parameters in the model except the mixture weight γ in the vector θ, then the likelihood function is

$$(6) \qquad L^{\text{SVAR}}(\theta, \gamma; \epsilon) = \prod_{t=1}^T L_t^{\text{SVAR}}(\theta, \gamma; \epsilon_t) = \prod_{t=1}^T f(\epsilon_t; \theta, \gamma),$$

---

[10]That is to say, if two decompositions $\breve{\Omega}$ and $\tilde{\Omega}$ both satisfy $\Sigma_1 = \breve{\Omega}\breve{\Omega}' = \tilde{\Omega}\tilde{\Omega}'$ and $\Sigma_2 = \breve{\Omega}\breve{\Lambda}\breve{\Omega}' = \tilde{\Omega}\tilde{\Lambda}\tilde{\Omega}'$, then there must exists a diagonal matrix J, all diagonal elements of which are either 1 or -1, such that $\breve{\Omega} = \tilde{\Omega}J$.

where the one-time density function is

$$
\begin{aligned}
f(\epsilon_t; \theta, \gamma) = & \gamma(2\pi)^{-K/2} \det(\Omega\Lambda_1\Omega')^{-1/2} \exp\left(-\frac{1}{2}\epsilon_t'(\Omega\Lambda_1\Omega')^{-1}\epsilon_t\right) \\
& + (1-\gamma)(2\pi)^{-K/2} \det(\Omega\Lambda_2\Omega')^{-1/2} \exp\left(-\frac{1}{2}\epsilon_t'(\Omega\Lambda_2\Omega')^{-1}\epsilon_t\right).
\end{aligned}
$$
(7)

However, it should be noticed that we can only obtain *local* estimation through ML. To illustrate the problem, let us assume $\hat{\Omega}$, $\hat{\Lambda}_1$ and $\hat{\Lambda}_2$ are a local ML estimation that maximizes the likelihood function in equation (6). If we permute the i-th and j-th columns of $\hat{\Omega}$ by right-multiply $\hat{\Omega}$ by the elementary column-switching operator E_{ij}, then $\Omega^* = \Omega E_{ij}$, $\Lambda_1^* = E_{ij}\hat{\Lambda}_1 E_{ij}'$ and $\Lambda_2^* = E_{ij}\hat{\Lambda}_1 E_{ij}'$ are also a group of local ML estimation. Since $E_{ij}E_{ij}' = I_K$, we have $\hat{\Omega}\hat{\Lambda}_1\hat{\Omega}' = \Omega^*\Lambda_1^*\Omega^{*'}$ and $\hat{\Omega}\hat{\Lambda}_2\hat{\Omega}' = \Omega^*\Lambda_2^*\Omega^{*'}$. Therefore both groups of estimations maximize the one-time density function in equation (7) for all t and thus the overall likelihood function. This is called the "label switching" problem in Maciejowska (2010). She provides a formal proof of the problem in the appendix. Since $\Omega = \mathsf{A}^{-1}$ represents the simultaneous effects of each structural shocks, the arbitrariness of the column permutation of Ω implies that we can only know how each structural shock affects the economic system, but do not know which structural shock identified by the SVAR model corresponds to which economic shock in reality.

Following Kohonen (2013), I solve this label switching problem with recourse to ordering information contained in some proxy variable in this paper. Kohonen (2013) argues that if we can find a proxy variable that provides information on the ordering of the diagonal elements of Σ_ϵ, then it becomes possible for us to pick the correct *global* estimation of the SVAR model among the local ML estimations. For example, if Y_t represents the returns of the stock markets in several countries, then the size of return shocks indicated by Σ_ϵ can be proxied by the strength of news in each country. Should we find some variable showing the strength of news, the diagonal elements of Σ_ϵ should have the same order as the strength of news. In this way, those local estimations where the diagonal elements of the estimated $\hat{\Sigma}_\epsilon = \hat{\gamma}\hat{\Lambda}_1 + (1-\hat{\gamma})\hat{\Lambda}_2$ is ordered incorrectly will be filtered out.

So far I have shown how to identify an SVAR model with normal-mixture residuals. In following sections, I will show that a SVAR actually contains all the information I need for revealing the connectedness among the entities. However, since this paper aims to describe how the connectedness transitions in different circumstances, we need a model that internalizes the transitions. In this paper, the transitions of connectedness are modeled by a threshold SVAR (T-SVAR) model, which I present in the following section.

In this part, I show how to construct a T-SVAR model. In the T-SVAR model, the data-generating process is partitioned by thresholds into N different regimes. Under each regime, the system is interconnected in a certain way, which is described by an SVAR model. As all the N regimes are estimated simultaneously, we obtain N different ways of interconnections of the entities or variables we are interested in, and the transitions among them are described by the threshold function. The T-SVAR model is still identified through its normal-mixture residuals.

Since in my model the threshold function is restricted to a number between zero and one, the threshold model has a probabilistic interpretation in economics. In the model without thresholds, the connectedness is certain for all the time. After the threshold function is introduced, the connectedness is no long certain, and we can only know that the connectedness is one among several possible states, and the probabilities for each possible state are decided by the threshold function, which is a certain reflection of the exogenous threshold variables. The dynamics of the connectedness is thus portrayed by the time-varying probability distributions decided by the threshold variables.

Assume there is an H-dimensional vector of exogenous threshold variables $Q_t = (q_{1t}, q_{2t}, \ldots, q_{Ht})'$ for time $t = 1, 2, \ldots, T$, and the unnormalized threshold function $\tau_\delta^* : \mathbb{R}^H \mapsto [0, \infty)^N$, where δ represents the unknown parameter(s) of τ_δ^*. I further normalize τ_δ^* by the sum of its value and get the normalized threshold function

$$(8) \qquad \tau(Q_t; \delta) \equiv \tau_\delta(Q_t) = \frac{\tau_\delta^*(Q_t)}{\mathbf{1}_N' \tau_\delta^*(Q_t)}.$$

This means there are N regimes, and τ_δ can map the threshold variables to the probability of the system to be in each regime. For simplicity, in the rest of this paper, when I mention the "threshold function", I always refer to the normalized threshold function. The components of the threshold function are denoted as $\tau(Q_t; \delta) = (\tau_1(Q_t; \delta)', \tau_2(Q_t; \delta)', \ldots, \tau_N(Q_t; \delta)')'$.

In practice, the threshold function can take any form. For example, the transilient form with one threshold variable is

$$(9) \qquad \tau_b^{\text{Tran}}(Q_t) = \begin{cases} (1, 0)', & Q_t \geq b, \\ (0, 1)', & Q_t < b, \end{cases}$$

where b is the breakpoint. This implies that the system has probability one to be under Regime 1 if the threshold variable $Q_t \geq b$, and under Regime 2 otherwise. We can also use the the multivariate logistic form with any number of variable(s)

as the threshold function

$$(10) \qquad \tau_\beta^{\text{Logi}}(Q_t) = \begin{pmatrix} 1/(1 + e^{-\beta'Q_t}) \\ 1 - 1/(1 + e^{-\beta'Q_t}) \end{pmatrix},$$

where β is the vector of scales of the logistic function.[11] This threshold function allows the system to switch between two regimes smoothly, and the location and speed of the switch are decided by parameters β.

Note that according to the definition of the threshold function (8), the threshold variables Q_t is given exogenously, but the parameters δ is unkown a priori. This means before we estimate the T-SVAR model, we only known different values of Q_t correspond to different probability distributions of regimes, but do not know how they correspond. The parameter values are estimated from the data. For instance, in the transilient threshold function (9), we only know that when $Q_t \geq b$, the system is at Regime 1, and let the data decide the value of b.

With these definitions, we consider the reduced-form and corresponding structural-form threshold model with N regimes and p lags:

$$Y_t = \Phi(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}, X_t, r_t)\tau_\delta(Q_t),$$

and

$$(11) \qquad \Psi(Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}, X_t, r_t)\tau_\delta(Q_t) = \mathbf{0}_K,$$

where r_t is still the error term, and the value of $\Phi(\cdot) = (\phi_1(\cdot), \phi_2(\cdot), \ldots, \phi_N(\cdot))$ and $\Psi(\cdot) = (\psi_1(\cdot), \psi_2(\cdot), \ldots, \psi_N(\cdot))$ are matrices of size $K \times N$ denoting the model conditioned on each regime. Notice that I *do not* have restrictions on the columns of $\Phi(\cdot)$ and $\Psi(\cdot)$ here, so they can be either mutually equal, partly equal, or completely unequal. In the application part of this paper, I let all their columns to be equal. Certainly, other kinds of restrictions are all possible.

The T-SVAR model considered in this paper can still be specified by assuming Ψ is linear and additively separable. Its summation form is:

$$(12) \qquad \begin{aligned} &\sum_{n=1}^{N} \mathsf{A}_n Y_t \tau_n(Q_t; \delta) = \\ &\sum_{n=1}^{N} \left(\mathsf{A}_n c^{(n)} + \sum_{i=1}^{p} \mathsf{A}_n A_i^{(n)} Y_{t-i} + \mathsf{A}_n A_X^{(n)} X_t + \varepsilon_{nt} \right) \tau_n(Q_t; \delta), \end{aligned}$$

where A_n denotes the structural matrix conditioned on regime n, $\tau_n(Q_t; \delta)$ is the probability for the system to be under regime n, $c^{(n)}$, $A_i^{(n)}$, and $A_X^{(n)}$ is au-

---

[11]In actual applications of this multivariate ligistic threshold function, the actual threshold variables are usually the desired threshold variables Q_t together with a constant term, say, $Q_t^{\text{Actual}} = (1, Q_t')'$, in order to keep a intercept term.

toregressive coefficients under regime n, and $\varepsilon_{nt} = \mathsf{A}_n r_t$ is the structural shocks conditioned on regime n.

The key to the identification of the T-SVAR model still rests on the distributional assumption of r_t. In the T-SVAR model I assume the conditional distribution of r_t is a time-constant normal-mixture distribution, and hence its unconditional distribution becomes a mixture of N normal-mixture distributions, i.e.

$$(13) \qquad \Sigma_{r,t}|_{Q_t;\delta} = \sum_{n=1}^{N} (\gamma_n \Sigma_{1n} + (1 - \gamma_n)\Sigma_{2n})\tau_n(Q_t;\delta).$$

Recall the results about the covariance decompositions in Section III.B, and then we know there are matrices $\{\Omega_n\}_{n=1}^{N}$, $\{\Lambda_{1n}\}_{n=1}^{N}$, and $\{\Lambda_{2n}\}_{n=1}^{N}$ such that $\Omega_n \Lambda_{1n} \Omega_n' = \Sigma_{1n}$ and $\Omega_n \Lambda_{2n} \Omega_n' = \Sigma_{2n}$ for all $n = 1, 2, \ldots, N$ under regular conditions. Equation (13) can thereby be rewritten as

$$(14) \qquad \Sigma_{r,t}|_{Q_t;\delta} = \sum_{n=1}^{N} \Omega_n(\gamma_n \Lambda_{1n} + (1 - \gamma_n)\Lambda_{2n})\Omega_n' \tau_n(Q_t;\delta).$$

Denote

$$(15) \qquad \Upsilon_n = \Omega_n(\gamma_n \Lambda_{1n} + (1 - \gamma_n)\Lambda_{2n})^{1/2},$$

then equation (14) can also be written as

$$\Sigma_{r,t}|_{Q_t;\delta} = \sum_{n=1}^{N} \Upsilon_n \Upsilon_n' \tau_n(Q_t;\delta).$$

Further, as assumed, for all n, we have $\epsilon_{nt} = \mathsf{A}_n r_t$ conditioned on regime n, therefore the unconditional relationship between the reduced-form and structural-form residuals will be

$$r_t|_{Q_t;\delta} = \sum_{n=1}^{N} (\mathsf{A}_n^{-1}\varepsilon_{nt})\tau_n(Q_t;\delta),$$

and

$$(16) \qquad \Sigma_{r,t}|_{Q_t;\delta} = \sum_{n=1}^{N} (\mathsf{A}_n^{-1}\Sigma_\varepsilon^{(n)}\mathsf{A}_n'^{-1})\tau_n(Q_t;\delta),$$

where $\Sigma_\varepsilon^{(n)}$ denotes the covariance matrix of ε_{nt}. Comparing equations (16) and (14), we get

$$\mathsf{A}_n = \Omega_n^{-1},$$

15

for all $n = 1, 2, \ldots, N$. And meanwhile, we have

$$\mathsf{A}_n^{-1} \Sigma_\varepsilon^{(n)\frac{1}{2}} = \Upsilon_n I_K^{\frac{1}{2}}.$$

This implies that Υ_n is actually the decomposition of the reduced-form covariance matrices if we alternatively require the covariance matrices of the structural shocks ε_{nt} to be unit matrices. In fact, Υ_n is the matrix that people are usually most interested in with traditional SVAR analyses, because it represents the instantaneous response of each variable to unit-sized structural shocks.

The T-SVAR model can still be locally estimated with ML. Denote $\varepsilon_t = \{\varepsilon_{nt}\}_{t=1}^T$ and $\varepsilon = \{\varepsilon_{nt}\}_{t=1,2,\ldots,T;n=1,2,\ldots,N}$. Stacking the mixture weights $(\gamma_1, \gamma_2, \ldots, \gamma_N)$ in vector Γ, and rest unknown parameters in the T-SVAR model except δ in vector Θ, then we get the likelihood function

$$(17) \qquad L^{\text{T-SVAR}}(\Theta, \Gamma, \delta; \varepsilon) = \prod_{t=1}^T L_t^{\text{T-SVAR}}(\Theta, \Gamma, \delta; \varepsilon_t) = \prod_{t=1}^T \varphi(\varepsilon_t; \Theta, \Gamma, \delta),$$

where the one-time density function

$$\varphi(\varepsilon_t; \Theta, \Gamma, \delta) = \sum_{n=1}^N \varphi_n(\varepsilon_{nt}; \Theta, \Gamma) \tau_n(Q_t; \delta),$$

and the one-time density function conditioned on regime n is

$$\varphi_n(\varepsilon_{nt}; \Theta, \Gamma) =$$
$$\gamma_n (2\pi)^{-K/2} \det(\Omega_n \Lambda_{1n} \Omega_n')^{-1/2} \exp\left(-\frac{1}{2} \epsilon_t' (\Omega_n \Lambda_{1n} \Omega_n')^{-1} \varepsilon_{nt} \right)$$
$$+ (1 - \gamma_n)(2\pi)^{-K/2} \det(\Omega_n \Lambda_{2n} \Omega_n')^{-1/2} \exp\left(-\frac{1}{2} \varepsilon_{nt}' (\Omega_n \Lambda_{2n} \Omega_n')^{-1} \varepsilon_{nt} \right).$$

After we obtain the local estimations of the T-SVAR model by maximizing the likelihood function, the global estimation can still be selected with the method proposed by Kohonen (2013) and described in Section III.B in this paper.

Finally, before I finish this section, I have a caveat here concerning the computer implementation of the ML estimation. Although in theory it is nothing more than finding a maximum of the log-likelihood function, in practice, direct numerical optimization algorithms work poorly for the SVAR models with normal-mixture residuals no matter there are thresholds or not, given the complexity of the likelihood functions. In order to achieve a high-quality optimum, I suggest using the expectation-maximization (EM) method. Maciejowska (2010) offers an in-depth discussion about the technical issues in the ML estimation of these mixture models.

The T-SVAR model is the foundation of the tripod model in this paper. In

the following section, I show how to transform a T-SVAR model to a spatial auto-regression model, which portrays the connectedness in a direct-viewing way.

IV.   From Structural Models to Spatial Models

A SVAR model depicts the instantaneous and lagging responses of the variables to structural shocks, and we want to further understand what interconnections the impulse-response relationship implies. Scidá (2016) demonstrates that an identified SVAR model actually contains a spatial auto-regression (SAR) process, and the spatial weight matrix in a SAR model is a good indicator of the connectedness among the entities of interest. Similarly, a T-SVAR model also contains a threshold SAR (T-SAR) process, which implies different spatial weights, and thereby different states of connectedness, under different regimes. The way to transform a T-SVAR model to a T-SAR is shown in Section IV.A. After we obtain the T-SAR model, we may construct statistical tests of various interesting hypotheses, which is discussed in Section IV.B.

A.   The SAR Model

In this section, I start from the SAR model

$$(18) \qquad Y_t = \rho W Y_t + c_W + \sum_{i=1}^{p} B_i Y_{t-i} + B_X X_t + \epsilon_t,$$

where ρ is a scalar scaling coefficient, $W = (w_{ij})_{K \times K}$ is a spatial weight matrix, $c_W, \{B_i\}_{i=1}^{p}$ and B_X are autoregressive coefficients, and ϵ_t is still the orthogonal error term with zero mean and a diagonal covariance matrix Σ_ϵ. Following the convention in spatial models, I assume W is row-standardized to unitary sums and its diagonal elements are all zeros. W represents the inverse distances, or in other words, spatial connectedness among the entities in an economic system, and the "spatial connectedness" can either be a inverse measure of real geographical distances among the entities, or some figurative connection such as the amount of mutual trade among economies. Specifically, the element w_{ij} in W indicates the connection from the j-th entity to the i-th entity.[12] And since W is standardized by ρ, ρ serves as an overall measure of the connectedness strength of the system.

If we compare equation (18) with equation (2), we may find their similarity without much difficulty. Noticing the diagonal elements of $G = I_K - \mathsf{A}$ are zeros, too, equation (18) is nothing more than (2) with row-standardizing constraints on the sans-diagonal structural matrix G. Certainly, in the SVAR model specified by equation (2) is unknown a priori and in spatial models W is usually exogenously

---

[12]If the connections are not directional, then W will be symmetric and thus $w_{ij} = w_{ji}$ for all $i \neq j$. But the connections can be directional, too. In fact, the connections considered in this paper are mostly directional.

given, but still, nothing prevents us to estimated a special SAR model with endogenous a spatial weight, or test an existing spatial weight matrix against the endogenous one implied by the data. In fact, the relationship between the SAR model with endogenous spatial weights and the SVAR model specified by equation (2) has been formally demonstrated by Proposition 3.1 in Scidá (2016). Because in this paper it is very unlikely that the estimation of G would have a row where all the elements are zeros due to the special SVAR identification strategy, I only present a corollary of this proposition below:

Proposition 1. *Assume the SVAR model (2) is identified, and no row in G is fully composed by zeros. Then the SAR model (18) is a constrained SVAR model (2) with $(K-1)$ independent linear restrictions on G given by*

$$\mathcal{R}G\mathbf{1}_K = \mathbf{0}_{K-1},$$

where the $(K-1) \times K$ matrix

$$\mathcal{R} = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

This proposition is an immediate corollary of Proposition 3.1 in Scidá (2016).

Based on the finding above, I proceed to the SAR model with thresholds. Consider the N-regime T-SAR model

(19)
$$\sum_{n=1}^{N} Y_t \tau_n(Q_t; \delta) =$$
$$\sum_{n=1}^{N} \left(\rho_n W_n Y_t + c_W^{(n)} + \sum_{i=1}^{p} B_i^{(n)} Y_{t-i} + B_X^{(n)} X_t + \varepsilon_{nt} \right) \tau_n(Q_t; \delta),$$

where ρ_n, W_n, $c_W^{(n)}$, $B_i^{(n)}$, and $B_X^{(n)}$ are respectively the scaling coefficient, spatial weight matrix, and autoregressive coefficients under regime n. And still, ε_{nt} is the structural error term.

Applying Proposition 1 to each regime in equation (12), we have for each regime n,

$$\mathcal{R}(I_K - \mathsf{A}_n)\mathbf{1}_K = \mathbf{0}_{K-1}.$$

Noticing $\mathcal{R}I_K\mathbf{1}_K \equiv \mathbf{0}_{K-1}$, this can be rewritten as

$$\mathcal{R}\mathsf{A}_n\mathbf{1}_K = \mathbf{0}_{K-1},$$

for all $n = 1, 2, \ldots, N$, or equivalently in the stacking form

$$\mathscr{R}\mathscr{A}\mathbf{1}_{NK} = \mathbf{0}_{N(K-1)},$$

where $\mathscr{R} = I_N \otimes \mathcal{R}$ and $\mathscr{A} = \mathrm{diag}(\mathsf{A}_1, \mathsf{A}_2, \ldots, \mathsf{A}_N)$. This conclusion turns the estimation of the T-SAR model into a restricted ML estimation program:

$$\max \quad L^{\text{T-SVAR}}(\Theta, \Gamma, \delta; \varepsilon),$$
$$\text{s.t.} \quad \mathscr{R}\mathscr{A}\mathbf{1}_{NK} = \mathbf{0}_{N(K-1)}.$$

However, while the constrained ML gives the parameter estimations of the T-SAR model, in practice we have to perform a very complex constrained non-linear optimization procedure; even the unconstrained non-linear optimization in the T-SVAR model is already technically challenging, let alone the restricted non-linear optimization. Therefore, in this paper, I employ the minimum distance (MD) approach suggested by Scidá (2016) to obtain the estimation constrained T-SAR model from the unconstrained T-SVAR model.

Let $\underline{\mathcal{A}} = (\mathrm{vec}(\mathsf{A}_1)', \mathrm{vec}(\mathsf{A}_2)', \ldots, \mathrm{vec}(\mathsf{A}_N)')'$ collect all the elements of $\{\mathsf{A}_n\}_{n=1}^N$ or $\mathscr{A}$, and

$$\underline{\mathcal{G}} \equiv \begin{pmatrix} \mathrm{vec}(G_1)' \\ \mathrm{vec}(G_2)' \\ \vdots \\ \mathrm{vec}(G_N)' \end{pmatrix} = \begin{pmatrix} \mathrm{vec}(I_K - \mathsf{A}_1)' \\ \mathrm{vec}(I_K - \mathsf{A}_2)' \\ \vdots \\ \mathrm{vec}(I_K - \mathsf{A}_N)' \end{pmatrix}.$$

Denote $\mathcal{E} = (e_{ij})_{K \times K}$ be a matrix where $e_{ij} = 0$ for all $i = j$ and $e_{ij} = 1$ for all $i \neq j$. Let S be a $(K^2 - K) \times K^2$ row-selection matrix which deletes all the all-zero rows in $\mathrm{diag}(\mathrm{vec}(\mathcal{E}))$ such that $S \times \mathrm{diag}(\mathrm{vec}(\mathcal{E}))$ does not have all-zero rows[13]. Then we can see the matrix

$$\mathcal{S} = I_N \otimes S$$

can pick out all the non-diagonal elements of $\{\mathsf{A}_n\}_{n=1}^N$ and stack them in the vector $\mathcal{G}$ by letting

$$\mathcal{G} = \mathcal{S}\underline{\mathcal{G}} = \mathcal{S}\underline{\mathcal{A}}.$$

In fact, $\mathcal{G}$ is the parameter vector that collects all the unknown parameters in $\{\mathsf{A}_n\}_{n=1}^N$ or $\mathscr{A}$ in the T-SVAR model as all the fixed parameters in $\underline{\mathcal{G}}$ are removed. Assume we have already obtained an estimation of $\mathcal{G}$ in the unconstrained T-SVAR model denoted by $\mathcal{G}^*$, and the squared error matrix of $\mathcal{G}^*$ is $\Sigma_{\mathcal{G}}^{*}$[14]. Now the MD estimation of $\mathcal{G}$ in the constrained T-SAR model is given by minimizing the weighted distance between the constrained $\mathcal{G}$ and the estimated unconstrained

---

[13]In fact, now we have $S \times \mathrm{diag}(\mathrm{vec}(\mathcal{E})) = S$.

[14]Since $\mathcal{G}^*$ is actually obtained through the ML estimation in this paper, $\Sigma_{\mathcal{G}}^*$ can be easily calculated via estimating Fisher's information of $\mathcal{G}^*$.

$\mathcal{G}^*$:

$$\begin{aligned}(20) \qquad & \min_{\mathcal{G}} \quad (\mathcal{G} - \mathcal{G}^*)'\Sigma_{\mathcal{G}}^{*-1}(\mathcal{G} - \mathcal{G}^*), \\ & \text{s.t.} \quad \mathscr{R}\mathscr{A}\mathbf{1}_{NK} = \mathbf{0}_{N(K-1)}.\end{aligned}$$

Note that although the constraint $\mathscr{R}\mathscr{A}\mathbf{1}_{NK}$ is not explicitly a function of $\mathcal{G}$, it can actually be rewritten as a linear combination of $\mathcal{G}$, which I show later in this section. Now the problem has been turned into the minimization of a quadratic form with linear constraints, which actually has a closed-form solution and is thus far less burdensome in computation than the constrained ML estimator. Moreover, it has been demonstrated by Property 9 in Gourieroux and Monfort (1989) that the MD estimator is asymptotically equivalent to the constrained ML estimator, which guarantees its efficiency.

In order to find the closed-form solution of the program (20), we first need to rewrite the constraints $\mathscr{R}\mathscr{A}\mathbf{1}_{NK} = \mathbf{0}_{N(K-1)}$ as a clear linear combination of the parameter vector $\mathcal{G}$. Concerning this, I show the following proposition:

Proposition 2. *Let* $\text{vec}(G_n^{\wedge 1})$, $n = 1, 2, \ldots, N$ *be a* K^2*-dimensional vector whose* i*-th component is* 1 *if the* i*-th component of* $\text{vec}(G_n)$ *is not zero, and* 0 *otherwise, and*

$$\mathcal{H}^{\wedge 1} = \begin{pmatrix} \text{vec}(G_n^{\wedge 1})\mathbf{1}'_{K-1} & & & \\ & \text{vec}(G_2^{\wedge 1})\mathbf{1}'_{K-1} & & \\ & & \ddots & \\ & & & \text{vec}(G_N^{\wedge 1})\mathbf{1}'_{K-1} \end{pmatrix},$$

The the constraint in the MD program (20) can be equivalently written as:

$$(21) \qquad \mathscr{R}\mathscr{A}\mathbf{1}_{NK} = \mathcal{H}\mathcal{G} = \mathbf{0}_{N(K-1)},$$

where

$$\mathcal{H} = (I_N \otimes (\mathbf{1}'_K \otimes \mathcal{R})') \odot \mathcal{H}^{\wedge 1}.$$

Now that the MD program becomes a classical minimization problem of quadratic forms with linear constraints, the MD estimator is given by its solution, as is described in the following proposition with $\mathscr{U} = \mathbf{0}_{N(K-1)}$:

Proposition 3. *Suppose the estimation* $\hat{\mathcal{G}}$ *of a certain model is given by the following MD program:*

$$\begin{aligned} & \min_{\mathcal{G}} \quad (\mathcal{G} - \mathcal{G}^*)'\Sigma_{\mathcal{G}}^{*-1}(\mathcal{G} - \mathcal{G}^*), \\ & \text{s.t.} \quad \mathscr{H}\mathcal{G} = \mathscr{U}, \end{aligned}$$

where the $d \times N(K^2 - K)$ *matrix* $\mathscr{H}$ *is the* d *distinctive linear constraints imposed on the parameter vector* $\mathcal{G}$, *and* $\mathscr{U}$ *is the* d*-dimensional vector indicating the values*

of the linear constraints, then we have

(22) $\qquad \hat{\mathcal{G}} = \mathcal{G}^* - \Sigma_{\mathcal{G}}^* \mathcal{H}'(\mathcal{H}\Sigma_{\mathcal{G}}^*\mathcal{H}')^{-1}\mathcal{H}\mathcal{G}^* + \Sigma_{\mathcal{G}}^*\mathcal{H}'(\mathcal{H}\Sigma_{\mathcal{G}}^*\mathcal{H}')^{-1}\mathcal{U}.$

Particularly, equation (22) reduces to

$$\hat{\mathcal{G}}^{\mathrm{T-SAR}} = \mathcal{G}^* - \Sigma_{\mathcal{G}}^*\mathcal{H}'(\mathcal{H}\Sigma_{\mathcal{G}}^*\mathcal{H}')^{-1}\mathcal{H}\mathcal{G}^*,$$

when the model is an T-SAR model without any additional restrictions and thus $\mathcal{H} = \mathcal{H}$, $\mathcal{U} = \mathbf{0}_{N(K-1)}$.

Since $\hat{\mathcal{G}}^{\mathrm{T-SAR}}$ contains the estimation of all the unknown parameters of $\{G_n\}_{n=1}^N$ in equation (19), the estimations of $\{\rho_n\}_{n=1}^N$ and $\{W_n\}_{n=1}^N$ can be easily obtained by standardizing the rows of $\{\rho_n\}_{n=1}^N$ by its row sums: for all $n = 1, 2, \ldots, N$,

$$\hat{\rho}_n = \frac{\mathbf{1}_K' \hat{G}_n \mathbf{1}_K}{K},$$

and

$$\hat{W}_n = \hat{\rho}_n^{-1}\hat{G}_n,$$

where $\hat{\rho}_n$, $\hat{W}_n$, and $\hat{G}_n$ are respectively the estimations of ρ_n, W_n, and G_n under regime n.

The estimation $\hat{W}_n$ portrays how the variables or entities are interconnected under Regime n. Based on these results, we may carry out some hypothesis tests to draw some clearer conclusions.

B. Hypothesis Tests

In this section I present how to construct a broad class of hypothesis tests of additional restrictions in the T-SAR model.

Throughout this section I denote $G_n = \rho_n W_n$ for $n = 1, 2, \ldots, N$, the elements in $\{G_n\}_{n=1}^N$ are stacked in $\underline{\mathcal{G}}$, and the free parameters $\mathcal{G} = \mathcal{S}\underline{\mathcal{G}}$ satisfies the restrictions of a T-VAR model so that $\mathcal{H}\mathcal{G} = \mathbf{0}_{N(K-1)}$.

In this paper, I focus on a special class of null hypotheses. Let V be a given $K \times K$ matrix with the same row sums and all its diagonal elements are zeros, and $\Xi_1, \Xi_2, \ldots, \Xi_N$ be N given subsets of $\mathbb{R}^1$. Then a null hypothesis considered in this paper is always such a proposition: there exists at least one group of coefficients $(\xi_1, \xi_2, \ldots, \xi_N)' \in \Xi_1 \times \Xi_2 \times \cdots \times \Xi_N \subseteq \mathbb{R}^K$ such that

(23) $$\sum_{n=1}^N \xi_n G_n = V.$$

Such propositions in fact cover a broad class of interesting hypotheses. For example, assume $N = 2$. Then if we let $V = V^{\mathrm{Exo}}$ be a given spatial matrix, Ξ_1 be

21

$\mathbb{R}^1$, and Ξ_2 be $\{0\}$, then the hypothesis equals

$$\mathsf{H}_0^{\text{One-Regime}}: \quad W_1 = V^{\text{Exo}};$$

if $V = O_K$, $\Xi_1 = \mathbb{R}^1$, and $\Xi_2 = \{-1\}$, then the hypothesis equals

$$(24) \qquad\qquad\qquad \mathsf{H}_0^{\text{Equal}}: \quad W_1 = W_2.$$

And if we let $V = V^{\text{Exo}}$, $\Xi_1 = \Xi_2 = \mathbb{R}^1$, then under the null hypotheses, there exist ξ_1^{Diff} and ξ_2^{Diff} such that

$$(25) \qquad W_2 = -\frac{\xi_1^{\text{Diff}}\rho_1}{\xi_2^{\text{Diff}}}\rho_2 W_1 + \frac{1}{\xi_2^{\text{Diff}}\rho_2}C^{\text{W}} \equiv \zeta_1^{\text{Diff}}W_1 + \zeta_2^{\text{Diff}}V^{\text{Exo}},$$

where $\zeta_1^{\text{Diff}} = -\xi_1^{\text{Diff}}\rho_1/\xi_2^{\text{Diff}}$, and $\zeta_2^{\text{Diff}} = 1/\xi_2^{\text{Diff}}\rho_2$. This implies W_2 is a kind of "aggregation" of W_1 and C^{W}. I denote this null hypothesis as $\mathsf{H}_0^{\text{Diff}}$.

I show this type of null hypotheses can be tested with likelihood ratio (LR) tests against a less restricted model. Since $\{G_n\}_{n=1}^N$ are estimated through the MD approach, we have to first obtain the estimations with additional restrictions under the null hypotheses and then evaluate the likelihood of the estimations under the null hypotheses.

The main problem of the estimation of the restricted model under the null hypothesis is the nonlinearity of the additional restrictions. They appear to be linear but in fact are not, as some ξ_n-s are uncertain. This being said, we may still convert it to a linear problem. First rewrite the restrictions (23) in the vectorized form. Vectorizing both sides of equation (23), we get

$$\xi_1\text{vec}(W_1) + \xi_2\text{vec}(W_2) + \cdots + \xi_N\text{vec}(W_N) = \text{vec}(V).$$

Then let $\xi = (\xi_1, \xi_2, \ldots, \xi_N)'$. Deleting all the fixed diagonal elements and applying multiplication rulues of block matrices, we obtain

$$(26) \qquad\qquad\qquad (\xi' \otimes I_K)\mathcal{G} = \mathcal{V},$$

where $\mathcal{V} = S\text{vec}(V)$ is the remaining components after deleting all the diagonal elements of V from $\text{vec}(V)$.

However, we still cannot simply put the additional restrictions in equation (26) and the basic restrictions for a T-SAR model in equation (21) together and let the restriction matrix under the null hypothesis be the simple combination of the two groups of restrictions. This is because some restrictions in equations (26) and (21) are equivalent - in equation (21) we require the row sums of each G_n, $n = 1, 2, \ldots, N$ to be the same; and when the restrictions in equation (26) are satisfied, since V has the same row sums, as long as $(N-1)$ in the N G_n-s have the same row sums, the remaining one must have the same row rums, too. Formally,

22

under the null hypothesis, if V and $G_1, G_2, \ldots, G_{N-1}$ all have the same row sums so that $\mathcal{R}V\mathbf{1}_K = \mathcal{R}G_1\mathbf{1}_K = \mathcal{R}G_2\mathbf{1}_K = \cdots = \mathcal{R}G_{N-1}\mathbf{1}_K = 0$, we have

$$\mathcal{R}G_N\mathbf{1}_K = \mathcal{R}(V - \xi_1 G_1 - \xi_2 G_2 - \cdots - \xi_{N-1}G_{N-1})\mathbf{1}_K = 0,$$

too. Therefore, we need to remove these redundant restrictions in order to estimate the restricted model under the null hypothesis. Actually, the final restrictions needed for the T-SAR model under the null hypothesis are constructed in the following way:

Without loss of generality, assume that Ξ_N is not restricted to $\{0\}$. Denote

$$\tilde{\mathcal{H}}^{\wedge 1} = \begin{pmatrix} \text{vec}(G_1^{\wedge 1})\mathbf{1}'_{K-1} & & & \\ & \text{vec}(G_2^{\wedge 1})\mathbf{1}'_{K-1} & & \\ & & \ddots & \\ & & & \text{vec}(G_{N-1}^{\wedge 1})\mathbf{1}'_{K-1} \end{pmatrix},$$

and

$$\tilde{\mathcal{H}} = (I_{N-1} \otimes (\mathbf{1}'_K \otimes \mathcal{R})') \odot \tilde{\mathcal{H}}^{\wedge 1},$$

then the final restrictions of model under the null hypothesis (23) are

(27)
$$\mathcal{K}_\xi \mathcal{G} = \mathcal{V},$$

where

$$\mathcal{K}_\xi = \begin{pmatrix} \tilde{\mathcal{H}} \\ \xi' \otimes I_K \end{pmatrix},$$

and

$$\mathcal{V} = \begin{pmatrix} \mathbf{0}_{(N-1)(K-1)} \\ \mathcal{V} \end{pmatrix}.$$

It should be stressed again that the restrictions in equation (27) are not linear, because ξ are unknown parameters and $\mathcal{K}_\xi$ depends on ξ. But we can convert them to linear restrictions. If ξ is known, then we get an estimation of $\mathcal{G}$ by applying Proposition 3:

$$\hat{\mathcal{G}}_\xi^{\text{H0}} = \mathcal{G}^* - \Sigma_\mathcal{G}^* \mathcal{K}_\xi'(\mathcal{K}_\xi \Sigma_\mathcal{G}^* \mathcal{K}_\xi')^{-1}\mathcal{K}_\xi \mathcal{G}^* + \Sigma_\mathcal{G}^* \mathcal{K}_\xi'(\mathcal{K}_\xi \Sigma_\mathcal{G}^* \mathcal{K}_\xi')^{-1}\mathcal{V},$$

and $\hat{\mathcal{G}}_\xi^{\text{H0}}$ is a function of ξ. The final estimation, $\hat{\mathcal{G}}^{\text{H0}}$, can be obtained by searching for a minimum of the weighted distance between $\hat{\mathcal{G}}_\xi^{\text{H0}}$ and $\mathcal{G}^*$, adjusting ξ within the given space $\Xi_1 \times \Xi_2 \times \cdots \times \Xi_N$. Hence, the estimation of the T-VAR model under the null hypothesis (23) is given as:

$$\hat{\mathcal{G}}^{\text{H0}} = \underset{\hat{\mathcal{G}}_\xi^{\text{H0}}, \xi \in \Xi_1 \times \Xi_2 \times \cdots \times \Xi_N}{\arg\min} (\hat{\mathcal{G}}_\xi^{\text{H0}} - \mathcal{G}^*)'\Sigma_\mathcal{G}^{*-1}(\hat{\mathcal{G}}_\xi^{\text{H0}} - \mathcal{G}^*).$$

23

Now that we have obtained $\hat{\mathcal{G}}^{\text{H0}}$, we try to test the restrictions against the a less restricted model under the alternative hypothesis. Suppose the estimation of $\mathcal{G}$ under the alternative hypothesis is $\hat{\mathcal{G}}^{\text{H1}}$. In this paper, I assume the model under the alternative hypothesis is also a T-SAR model with less or no additional restrictions in the form of equation (23). Therefore, $\hat{\mathcal{G}}^{\text{H1}}$ can be obtained and evaluated in the same way as $\hat{\mathcal{G}}^{\text{H0}}$, for which reason I only discuss $\hat{\mathcal{G}}^{\text{H0}}$ in the following paragraphs.

As mentioned above, I test the null hypothesis with an LR test. In order to construct the LR statistic, we must first evaluate $\hat{\mathcal{G}}^{\text{H0}}$ in the likelihood function. Recalling that the T-SAR model is nothing more than a constrained T-SVAR model, we can express the likelihood function with the same form as $L^{\text{T-SVAR}}$ and the same parameter sets Θ, Γ, and Δ. Thus, the likelihood function

$$L^{\text{H0}}(\hat{\Theta}^{\text{H0}}, \hat{\Gamma}^{\text{H0}}, \hat{\delta}^{\text{H0}}; \hat{\varepsilon}^{\text{H0}}) = L^{\text{T-SVAR}}(\hat{\Theta}^{\text{H0}}, \hat{\Gamma}^{\text{H0}}, \hat{\delta}^{\text{H0}}; \hat{\varepsilon}^{\text{H0}}),$$

where $L^{\text{T-SVAR}}$ has been given in equation (17), $\hat{\Theta}^{\text{H0}}$, $\hat{\Gamma}^{\text{H0}}$, and $\hat{\delta}^{\text{H0}}$ are the estimations of respective parameters under the null hypothesis. The problem is, $\hat{\mathcal{G}}^{\text{H0}}$ is merely a part of $\hat{\Theta}^{\text{H0}}$, and the remaining parameters in Θ, together with Γ and δ, are not obtained through the MD estimation. I try to solve this problem by returning to the ML problem

$$(28) \qquad L^{\text{H0}}(\hat{\Theta}^{\text{H0}}, \hat{\Gamma}^{\text{H0}}, \delta^{\text{H0}}; \hat{\varepsilon}^{\text{H0}}) = \max_{\Theta \setminus \mathcal{G}, \Gamma, \delta} L^{\text{H0}},$$

where the set difference $\Theta \setminus \mathcal{G}$ is the parameters in Θ but not in $\mathcal{G}$.

Given the large number of unknown parameters in the maximization problem (28), it is still a rather computationally demanding task.

Fortunately, for three of the most widely used specifications of the reduced-form model, I show that this procedure may be significantly simplified by a generalized lest square (GLS) estimator of the autoregressive coefficients. In the first specification, all the autoregressive coefficients of the reduced-form model under different regimes are free to vary. In the second specification, the autoregressive coefficients of the reduced-form model are assumed to be all the same across different regimes. And the third specification is the mixture of the previous two.

In order to apply the GLS estimator, first let us restore the T-SAR model (19) to the reduced-form

$$(29) \qquad Y_t = \sum_{n=1}^{N} \left(c^{(n)} + \sum_{i=1}^{p} A_i^{(n)} Y_{t-i} + A_X^{(n)} X_t \right) \tau_n(Q_t; \delta) + r_t,$$

and the covariance matrix of r_t has been given by equation (14). For simplicity, in following discussions I denote $\Sigma_{r,t}|_{Q_t;\delta}$ as $\Sigma_{r,t}$. Moreover, let $\mathcal{C}$ collect all the unknown autoregressive coefficients $\{c^{(n)}, A_i^{(n)}, A_X^{(n)} : n = 1, 2, \ldots, N, i = 1, 2, \ldots, p\}$, and ϑ collect all the rest unknown parameters in $\{\Theta \setminus \mathcal{G}, \Gamma, \delta\}$. We

can see that $\vartheta = \{\Lambda_{11}, \Lambda_{12}, \ldots, \Lambda_{1N}, \Lambda_{21}, \Lambda_{22}, \ldots, \Lambda_{2N}, \gamma_1, \gamma_2, \ldots, \gamma_N, \delta\}$.

With these preparations, let us discuss the following three specifications:

Specification A. In this specification, all the autoregressive coefficients of the reduced-form model under different regimes are free to vary. That is to say, for any $n_1, n_2 = 1, 2, \ldots, N$ and $i = 1, 2, \ldots, p$, we *do not* require $c^{(n_1)} = c^{(n_2)}$, $A_i^{(n_1)} = A_i^{(n_2)}$, or $A_X^{(n_1)} = A_X^{(n_2)}$.

The reduced-form (29) can collapse to

$$(30) \qquad Y_t = \sum_{n=1}^{N} C_n \tilde{Z}_{nt} \tau_n(Q_t; \delta) + r_t,$$

where $C_n = (c^{(n)}, A_1^{(n)}, A_2^{(n)}, \ldots, A_p^{(n)}, A_X^{(n)})$ is a stack of autoregressive coefficients, and $\tilde{Z}_t = (1, Y_{t-1}', Y_{t-2}', \ldots, Y_{t-p}', X_t')'$. Further let $C^{\text{SpecA}} = (C_1, C_2, \ldots, C_n)$, and

$$Z_t^{\text{SpecA}} = \tau_\delta(Q_t) \otimes \tilde{Z}_t = \begin{pmatrix} \tau_1(Q_t; \delta) \tilde{Z}_t \\ \tau_2(Q_t; \delta) \tilde{Z}_t \\ \vdots \\ \tau_N(Q_t; \delta) \tilde{Z}_t \end{pmatrix},$$

equation (30) is rewritten as

$$Y_t = C^{\text{SpecA}} Z_t^{\text{SpecA}} + r_t.$$

Vectorizing both side, we get

$$Y_t = (Z_t^{\text{SpecA}'} \otimes I_K) \text{vec}(C^{\text{SpecA}}) + r_t.$$

This is a typical multivariate regression model with time-varying error covariances.

Assuming that ϑ is known, then we can write the GLS estimator as

$$\text{vec}\hat{C}_\vartheta^{\text{SpecA}} = (\hat{\Sigma}_Z^{\text{SpecA}})^{-1} \text{vec}(\hat{\Sigma}_Y^{\text{SpecA}}),$$

where

$$\hat{\Sigma}_Z^{\text{SpecA}} = \frac{1}{T} \sum_{t=1}^{T} (Z_t^{\text{SpecA}} Z_t^{\text{SpecA}'} \otimes \Sigma_{r,t}^{-1}),$$

and

$$\hat{\Sigma}_Y^{\text{SpecA}} = \frac{1}{T} \sum_{t=1}^{T} \Sigma_{r,t}^{-1} Y_t Z_t^{\text{SpecA}'}.$$

Specification B. In this specification, all the autoregressive coefficients of the

25

reduced-form model are the same across different regimes. Therefore, for any $n_1, n_2 = 1, 2, \ldots, N$ and $i = 1, 2, \ldots, p$, we require $c^{(n_1)} = c^{(n_2)} = c$, $A_i^{(n_1)} = A_i^{(n_2)=A_i}$, and $A_X^{(n_1)} = A_X^{(n_2)} = A_X$. Then the threshold function can be removed from equation (29), yielding

$$(31) \qquad Y_t = c + \sum_{i=1}^{p} A_i Y_{t-i} + A_X X_t + r_t.$$

Now, let $C^{\text{SpecB}} = (c, A_1, A_2, \ldots, A_p, A_X)$, and $Z^{\text{SpecB}} = (1, Y'_{t-1}, Y'_{t-2}, \ldots, Y'_{t-p}, X'_t)'$, then the model (31) can also be written as

$$Y_t = (Z_t^{\text{SpecB}'} \otimes I_K)\text{vec}(C^{\text{SpecB}}) + r_t.$$

The GLS estimator follows similarly with Specification A:

$$\text{vec}\hat{C}_\vartheta^{\text{SpecB}} = (\hat{\Sigma}_Z^{\text{SpecB}})^{-1}\text{vec}(\hat{\Sigma}_Y^{\text{SpecB}}),$$

where

$$\hat{\Sigma}_Z^{\text{SpecB}} = \frac{1}{T}\sum_{t=1}^{T}(Z_t^{\text{SpecB}} Z_t^{\text{SpecB}'} \otimes \Sigma_{r,t}^{-1}),$$

and

$$\hat{\Sigma}_Y^{\text{SpecB}} = \frac{1}{T}\sum_{t=1}^{T} \Sigma_{r,t}^{-1} Y_t Z_t^{\text{SpecB}'}.$$

Specification C. It is also easy to specify the reduced-form model (29) as a mixture of Specifications A and B. There are many possible specifications in this category; as an example, consider the following case:

Assume that the intercepts $c_{(1)}, c^{(2)}, \ldots, c^{(n)}$ can freely vary under different regimes, whereas the other autoregressive coefficients are the same across different regimes. This is a typical varying-incept threshold model that is often seen in literature. Still let $A_i^{(n_1)} = A_i^{(n_2)} = A_i$, and $A_X^{(n_1)} = A_X^{(n_2)} = A_X$. Similar with Specifications A and B, let $C^{\text{SpecC}} = (c^{(1)}, c^{(2)}, \ldots, c^{(N)}, A_1, A_2, \ldots, A_p, A_X)$, and $Z^{\text{SpecC}} = (\tau_1(Q_t; \delta), \tau_2(Q_t; \delta), \ldots, \tau_N(Q_t; \delta), Y'_{t-1}, Y'_{t-2}, \ldots, Y'_{t-p}, X'_t)'$, then we can rewrite the model of this specification in the vectorized form again:

$$Y_t = (Z_t^{\text{SpecC}'} \otimes I_K)\text{vec}(C^{\text{SpecC}}) + r_t.$$

The GLS estimator goes as always:

$$\text{vec}\hat{C}_\vartheta^{\text{SpecC}} = (\hat{\Sigma}_Z^{\text{SpecC}})^{-1}\text{vec}(\hat{\Sigma}_Y^{\text{SpecC}}),$$

26

where

$$\hat{\Sigma}_Z^{\text{SpecC}} = \frac{1}{T} \sum_{t=1}^{T} (Z_t^{\text{SpecC}} Z_t^{\text{SpecC}'} \otimes \Sigma_{r,t}^{-1}),$$

and

$$\hat{\Sigma}_Y^{\text{SpecC}} = \frac{1}{T} \sum_{t=1}^{T} \Sigma_{r,t}^{-1} Y_t Z_t^{\text{SpecC}'}.$$

Besides the three specifications discussed above, it is always possible to specify the reduced-form model (29), but then we have to construct the GLS estimators for them ad hoc.

No matter what specification we use, we have now obtained the GLS estimation of $\text{vec}(C)$, which is a function of unknown parameters ϑ. Patilea and Raïssi (2012) has proved that these GLS estimators are asymptotically consistent. Recalling that $\text{vec}(C)$ is in fact a vector composed by the autoregressive coefficients $\mathcal{C}$, I denote it as $\hat{\mathcal{C}}_\vartheta$. It follows that

$$L^{\text{H0}}(\hat{\Theta}^{\text{H0}}, \hat{\Gamma}^{\text{H0}}, \hat{\delta}^{\text{H0}}; \hat{\varepsilon}^{\text{H0}})$$
$$= L^{\text{H0}}(\hat{\mathcal{G}}^{\text{H0}}, \hat{\mathcal{C}}^{\text{H0}}, \hat{\vartheta}^{\text{H0}})$$
$$= \max_{\vartheta}(\hat{\mathcal{G}}^{\text{H0}}, \hat{\mathcal{C}}_\vartheta, \vartheta).$$

As long as we obtain $L^{\text{H0}}(\hat{\Theta}^{\text{H0}}, \hat{\Gamma}^{\text{H0}}, \hat{\delta}^{\text{H0}}; \hat{\varepsilon}^{\text{H0}})$ and $L^{\text{H1}}(\hat{\Theta}^{\text{H1}}, \hat{\Gamma}^{\text{H1}}, \hat{\delta}^{\text{H1}}; \hat{\varepsilon}^{\text{H1}})$, the LR statistic is given by

$$\text{LR} = 2(L^{\text{H1}}(\hat{\Theta}^{\text{H1}}, \hat{\Gamma}^{\text{H1}}, \hat{\delta}^{\text{H1}}; \hat{\varepsilon}^{\text{H1}}) - L^{\text{H0}}(\hat{\Theta}^{\text{H0}}, \hat{\Gamma}^{\text{H0}}, \hat{\delta}^{\text{H0}}; \hat{\varepsilon}^{\text{H0}})).$$

LR can be compared with a χ^2 distribution. The degree of freedom of the χ^2 distribution can be calculated by comparing the numbers of restrictions of the null and alternative hypotheses. Taking the T-SAR model without any additional restrictions as the benchmark, the model under the null or alternative hypothesis with restrictions in equation (27) has at most $K^2 - (K-1)$ more restrictions on $\mathcal{G}$ than the benchmark. But since there are also a certain number of free parameters $\xi_1, \xi_2, \ldots, \xi_N$, and they must be subtracted from the total number of additional restrictions. For example, under $\mathsf{H}_0^{\text{Equal}}$, since ξ is a free parameter, there should be $K^2 - (K-1) - 1$ more restrictions on $\mathcal{G}$ than the benchmark. Following steps of the test do not generate more restrictions; therefore the final number of additional restrictions under a hypothesis in the form of equation (23) is $K^2 - (K-1)$ minus the number of free parameters in $\xi_1, \xi_2, \ldots, \xi_N$. The degree of freedom of the χ^2 distribution follows.

V. Network Analyses

While the connectedness among economic entities can be described by the spatial weight matrices in the T-SAR model, there are ways to analyze the connect-

edness further. The large amount of literature on network has proposed many methods to analyze the connectedness in details. In this paper, I follow Scidá (2016) and build networks from the spatial weight matrices estimated in Section IV.A. This completes the final part of the tripod model.

Similar to Scidá (2016), I take the transposes of the spatial weight matrices $W_1', W_2', \ldots, W_N'$ as the network adjacency matrices $\mathfrak{A}_1, \mathfrak{A}_2, \ldots, \mathfrak{A}_N$. That is to say, for each regime $n = 1, 2, \ldots, N$, I assume there is a network defined by the adjacency matrix $\mathfrak{A}_n = W_n'$. The transpose in fact arises from the different conventions in different strands of literature: in spatial econometrics, the element $w_{n,ij}$ on the i-th row and j-th column of the spatial weight matrix W_n usually denotes the influence of the j-th variable on the i-th variable, whereas in network studies, the element $\mathfrak{a}_{n,ij}$ on the i-th row and j-th column of the adjacency matrix $\mathfrak{A}_n = (\mathfrak{a}_{n,ij})_{K \times K}$ usually denotes the influence of the i-th variable on the j-th variable. The diagonal elements of $\mathfrak{A}_n$ are set to zeros, which means that there are no "loops", which connects a node[15] to itself.

Adjacency matrices defined in this way have several significant features. First, they are usually asymmetric, which results in directed networks. Second, they are weighted networks, where the strength of each edge is usually not one. Third, they are usually complete, which means usually every pair of nodes are somehow connected. Fourth, they often have negatively weighted edges. On the contrary, networks studied in traditional network studies typically lack these features. A large number of methods are devised for analyzing networks that are undirected, non-weighted, incomplete, and do not have negative edges. Certainly, there are also many ways to analyze networks with a part of the four features, but it is indeed not easy to find some method to analyze a network with all the four features. Despite all the difficulties, in this paper I show that it is still possible to apply several analyses to the highly featured networks built from spatial weight matrices.

First, there are centrality measures. Centrality measures are the most popular measure to assess nodes in a network. Generally speaking, they measure how well is a node connected to others. For example, the simplest centrality measure is the degree centrality, which counts the number of neighbors of each node. This measure obviously does not work for our complete network as all nodes in a complete network have the same number of neighbors. Scidá (2016) also propose to use three other types of popular centrality measures, namely closeness centrality, betweenness centrality, and Bonacich power centrality. Among them the closeness centrality and the betweenness centrality do not work for our networks with negative edges, because they evaluate the shortest path between each pair of nodes, and the common definition of "shortest path" fails when there are negative edges, through which the length of the path decreases. And finally, the Bonacich power centrality works only for networks where all edges are simultaneously positive or

---

[15]A network is composed by "nodes" and "edges". Nodes are the entities involved in the network, and edges are the connections between the nodes.

simultaneously negative.

In this paper, I propose to use two centrality measures that can handle the complete networks with both positive and negative edges. The first is the simplest and most intuitive measure by generalizing the degree centrality to weighted networks. Since the networks we consider are also directed networks, we have both the "in-degrees" and the "out-degrees", which respectively measure how much influence a node receives from other nodes, and how much influence a node exerts on other nodes. Formally, for the node k in network n we have the in-degree

$$\mathfrak{c}_{nk}^{\mathrm{d,in}} = \frac{\sum_{i=1}^{K} \mathfrak{a}_{n,ik}}{K - 1},$$

and the out-degree

$$\mathfrak{c}_{nk}^{\mathrm{d,out}} = \frac{\sum_{i=1}^{K} \mathfrak{a}_{n,ki}}{K - 1}.$$

However, while the generalized degree centrality can measure how well a node is connected to its neighbors, it neglects the "quality" of the neighbors. If we believe that a node should also have high centrality when all its neighbors are well connected to other nodes, then we need a centrality measure that evaluates not only how well a node is connected, but also how well its neighbors are connected. This is exactly why the Bonacich centrality measure, Hubbell centrality measure,.etc., are devised.[16] However, since none of these measures can handle networks with both positive and negative edges, following Bonacich and Lloyd (2004), Everett and Borgatti (2014) devises the PN centrality measure. The PN centrality is basically a generalized version of the Hubbell measure that considers the positive and negative edges separately and then sum both sides together. Formally, for the network n, we have the in-PN measure

$$\mathfrak{C}_n^{\mathrm{PN,in}} = \left(I_K - \frac{1}{4(K-1)^2} \mathfrak{A}_n \mathfrak{A}_n' \right)^{-1} \left(I_K + \frac{1}{2(n-1)} \mathfrak{A}_n \right) \mathbf{1}_K,$$

and the out-PN measure

$$\mathfrak{C}_n^{\mathrm{PN,out}} = \left(I_K - \frac{1}{4(K-1)^2} \mathfrak{A}_n' \mathfrak{A}_n \right)^{-1} \left(I_K + \frac{1}{2(n-1)} \mathfrak{A}_n' \right) \mathbf{1}_K,$$

where both $\mathfrak{C}_n^{\mathrm{PN,in}}$ and $\mathfrak{C}_n^{\mathrm{PN,out}}$ are K-dimensional vectors containing the PN centralities for all the nodes $1, 2, \ldots, K$.

Furthermore, because the adjacency matrix $\mathfrak{A}_n$, $n = 1, 2, \ldots, N$ is always column-standardized, we will always get the same in-degree and in-PN centralities for all the nodes. Therefore, in this paper, we only consider the out-degree and out-PN centrality measures.

---

[16]See Everett and Borgatti (2014), for a review of these measures.

Besides centrality measures, the clustering analysis can also reveal some important aspects of a network. Different from the centralities, clustering is both a local analysis of the position of each node in the network and a global analysis of the overall connectedness of the entire network. For each node $i = 1, 2, \ldots, K$, we have a clustering coefficient for the node. While the centrality measures how well the node is connected to other nodes, the clustering coefficient of the node measures how well its neighbors are mutually connected. If the neighbors of a node is closely connected, then the node has a larger clustering coefficient, and vice versa. Moreover, if we take the mean of the clustering coefficients of all the nodes, we get the average clustering coefficient for the whole network, which is a good measure of the overall connectedness of the whole network.

However, traditional definitions clustering coefficients fail again for a complete network. The reason is easy to understand: since all the nodes in a complete network are connected, the clustering coefficients of all the nodes will reach their maximum value, one. Therefore, the clustering coefficients for a complete network must be constructed in a different way.

In this paper, I use the clustering coefficient recently constructed by McAssey and Bijma (2015), which takes the edge weights into consideration and therefore works well for complete networks. Formally, for $k = 1, 2, \ldots, K$ and $n = 1, 2, \ldots, N$, the clustering coefficient of node i in network n is given as

$$\mathscr{C}_{nk} = \frac{(\mathfrak{A}_n + \mathfrak{A}'_n)^3_{kk}}{2((\mathfrak{A}_n + \mathfrak{A}'_n)(\mathbf{1}_K \mathbf{1}'_K - I_K)(\mathfrak{A}_n + \mathfrak{A}'_n))_{kk}},$$

where the subscript kk denotes the k-th diagonal element of the matrix. Further, the average clustering coefficient for the whole network is

$$\bar{\mathscr{C}} = \frac{\sum_{k=1}^{K} \mathscr{C}_k}{K}.$$

And finally, Scidá (2016) suggests that the estimation of the scaling parameter in the T-SAR model, $\hat{\rho}_n$, $n = 1, 2, \ldots, N$ are themselves natural indicators of the strength of the interactions among the variables under regime n. This is another global measure of the overall connectedness among the variables.

VI.   An Application: Financial Contagion in the Eurozone

A.   Introduction

In this section, I present an application of my model, which is on contagion among the stock markets in the Eurozone.[17]

---

[17]The application is mostly coded in R, together with a little C. I thank the authors of R and the R packages listed below: actuar, combinat, doParallel, fields, foreach, ggnetwork, ggplot2, lubridate, Matrix, matrixcalc, mixtools, mnormt, network, nloptr, plyr, profvis, reshape2, turboEM, and vars.

The Eurozone markets have been in turmoil for nearly a decade. At first, they were influenced considerably by the 2008 Financial Crisis; and not after long, the European countries found that they were in trouble again due to debt problems in a handful of countries such as Ireland and Greece. During this period, the contagion among international asset markets received much attention from both the academia and the public.

However, there is yet no consensus on the definition of contagion till now. I identify three categories of definitions. First, recent literature tend to define contagion as the intensification of shock transmissions among markets during turbulent periods. Especially, many studies model the contagion under this definition empirically by profiling the structural breaks of bivariate correlations of assets, such as Forbes and Rigobon (2002), Caporale et al. (2005), Chiang et al. (2007), and Caporin et al. (2013). Second, some authors anchor the contagion on economic fundamentals, and the contagion is defined as shock transmissions in excess of what can be explained by fundamental. Dungey and Martin (2007) is an example of this category of literature. Third, some literature does not strictly differentiate contagion and spillover. They do not assume that the shocks are transmitted in a different way during crisis periods. This is a rather broad definition and seems to be popular only in earlier literature, such as Edwards (1998).

In this paper, I basically follow the first category of definition of contagion. I define contagion as the increase of connectedness among the markets in the turbulent periods. However, my methodology is distinguished from most existing literature in this category, as I do not consider the statistical correlations among the markets; rather, I focus on the connectedness inferred from the contemporaneous impulse-response relationships with the threshold tripod model.[18] Comparing with those contagion measures based on correlations, my measure has several clear features: First, my model is based on contemporaneous impulse-response relationships, which show a clear causal map of the propagation of shocks. It is widely known that correlation is different from causality; so when we observe a correlation between two markets, we cannot know what causes the correlation. But in my model, it is easy to observe how the markets are connected by mutual shocks and responses. Related to this, while correlation analyses suffer from the heteroskedasticity problem and must be corrected (see Forbes and Rigobon, 2001), my model allows the heteroskedasticity explicitly. Second, the network analysis techniques can efficiently summarize the interrelationship among the markets as well as the unique roles played by each market in both tranquil and turbulent periods. Moreover, it is easy to see *how* the market interrelationship changes in different periods from a network angle, which provides us with clues for the channels of contagion. Third, my model allows the data to decide the different states of the market interrelationship. Therefore, we do not need to set the criterion of

---

[18]In fact, the correlation can be a measure of connectedness as well. For example, the weighted gene co-expression network analysis (WGCNA), which has been widely applied in bioinformatics, builds networks on correlations. It might be interesting to investigate market correlations from the network angle, too.

crisis manually. Further, in this application, I use a smooth threshold function, which does not draw a clear-cut difference between crisis and non-crisis periods. This makes my measure of contagion more flexible.

Following Kohonen (2013), I select five representative countries to study the contagion in Eurozone stock markets, namely Greece (GR), Germany (DE), Italy (IT), Spain (ES), and Ireland (IE). And I let the data determine the contagion and non-contagion regimes according to the CDS data. With the threshold tripod model, I confirm the existence of contagion among the stock markets in these countries in recent years, and reject the hypothesis that the contagion is solely conducted by the cross-border interbank channel. Further analyses reveal the distinctive roles played by the five countries in both the contagion and the non-contagion regimes clearly.

B. Data and Model Setup

I study stock market logarithm returns in the five representative countries. The market indices are respectively Athex Composite, DAX 30, FTSE MIB, IBEX 35, and Ireland SE Iverall, all of which are obtained from Datastream. The observation period ranges from January 2nd, 2006 to May 26th, 2016, therefore there are 2,714 observations in total. For computational reasons, I multiply all the log returns by 100.

Figure B1 shows the trends of the data in the observed years. Note that all the data have been rescaled by their initial values. It is easy to notice that these markets are to a large extent interrelated. Especially, all the countries saw a large simultaneous downturn during the 2008 crisis , and a relatively less significant one during the recent European debt crisis. Moreover, the German market and the Greek market show extremely different trends, which hint the different roles played by the two countries.

[Insert Figure B1 here.]

I assume there are two different regimes, and the regime with generally higher market connectedness is called "Regime 2", and the other "Regime 1". Contagion can be thus viewed as the difference of the market connectedness between the two regimes. I then choose the iTraxx Europe index as the threshold variable, which is retrieved from Datastream. It measures the overall default risk in Europe. In fact, default risk is the most common crisis indicator used in contagion literature, e.g., Kalbaskaa and Gątkowski (2012), Manasse and Zavalloni (2013), among others. It is also used by De Bruyckere et al. (2013) in their study of the Eurozone contagion. Their historical trends are plotted in Figure B2. For the threshold function, I use the multivariate logistic function described in equation (10)

$$(32) \qquad \tau_1^{\text{Logi}}(\text{iTraxx}_t; \beta) = \frac{1}{1 + e^{-\beta_0 - \beta_1 \text{iTraxx}_t}},$$

and
$$\tau_2^{\text{Logi}}(\text{iTraxx}_t; \beta) = 1 - \tau_2(\text{iTraxx}_t; \beta),$$
which allows the system to transition between each other smoothly. The location and speed of the transition are decided by the coefficients of the multivariate logistic function β_0 and β_1. Given other parameters constant, the larger the absolute value of β_1 is, the faster the system switches between the regimes.

[Insert Figure B2 here.]

Moreover, I assume all the autoregressive coefficients in the reduced-form VAR are constant. This means the relationship between different markets are the same across different regimes barring the different instantaneous effects in the same day when shocks take place. This assumption can be supported by the so-called "non-crisis-contingent"[19] category of contagion theories in Forbes and Rigobon (2001), which argue "any large cross-market correlations after a shock are a continuation of linkages that existed before the crisis," such as trade. In fact, in this application part I test whether the Eurozone contagion relies on the cross-border interbank market linkage. The lag order is set to one.

It should be noted that I do not include any factor that predicts the returns in this application. The tripod model framework definitely allows me to do so, as I always allow exogenous variables in the model description in previous sections.This makes this application a phenomenological description of past events, and it is probably not very reliable to use these findings to forecast the future. Some authors propose to use factor models to study contagions (e.g. Dungey and Martin, 2007). Indeed, a proper factor model may do a better job in revealing the underlying economic mechanisms of the contagion, but the specification of factors is never easy and can sometimes be controversial. In order to keep the application of my model clear and illustrative in this paper, I leave the potential factor contagion models by the threshold tripod approach for future research.

C.   The Local Identification

I first estimate the two-regime T-SVAR model of market log returns with constant reduced-form autoregressive coefficients, multivariate logistic threshold function, and one lag. Without any additional information, I can achieve the local identifications. But due to the "label-switching" problem, these local identifications of Ω_1 and Ω_2 do not make economic sense at this moment, because we can always arbitrarily permute the columns of the local estimations of them.[20] Nevertheless, since we do not need the global identification to obtain the estimations of the parameters in the threshold functions and the mixture weights, we can still know now that the estimations of the mixture weights under Regimes 1 and 2 are

---

[19]This is the name given by Forbes and Rigobon (2001) to a class of economic *theories*. There is no relationship between this name and the title of this paper at all, as they refer to different things like apples and pears.

[20]For the meaning of the parameters, see Section III.C.

respectively $\hat{\gamma}_1 = 0.310$ and $\hat{\gamma}_1 = 0.118$, and the estimations of the threshold function parameters in equation (32) are respectively $\hat{\beta}_0 = 2.703$, and $\hat{\beta}_1 = -0.037$, where the variable with hats are the estimations of respective parameters without hats. We can see that the coefficient for iTraxx, β_1 is negative. This implies that an increase in iTraxx will lead to a higher probability for the system to be under Regime 2, where the markets have closer interconnections as I show in the following sections. The trend of the threshold function can be viewed more clearly in Figure B3. This figure reveals two major clusters of high-interconnection periods: the 2008 crisis, the following Eurozone crisis. Interestingly, during the period starting from late 2015, there is likely to be a recurrence of the Eurozone crisis, though not as significant as the previous two crises.

[**Insert Figure B3 here.**]

D. From Local Identification to Global Identification

Now that we have obtained local identification, we may proceed to the global identification of the T-SVAR model.

Based on the theoretical model of King and Wadhwani (1990), Kohonen (2013) suggests to use the strength of news in each market, which is extracted from Google Trends Indices (GTIs), as the proxy variable to attain the correct permutation of the columns of covariance matrix decompositions $\hat{\Omega}_1$ and $\hat{\Omega}_2$. The idea is simple: the GTI of a certain term indicates the relative volumes of weekly queries of the term on Google. Therefore, a more volatile GTI of a market-related term will imply stronger news in this period, and stronger news will lead to stronger innovations in asset returns, as are described by the diagonal elements of $\Sigma_\varepsilon^{(1)} = \gamma_1\Lambda_{11} + (1 - \gamma_1)\Lambda_{21}$ and $\Sigma_\varepsilon^{(2)} = \gamma_1\Lambda_{12} + (1 - \gamma_1)\Lambda_{22}$.[21] Once the correct $\hat{\Lambda}_{11}$, $\hat{\Lambda}_{21}$, $\hat{\Lambda}_{12}$ and $\hat{\Lambda}_{22}$ are pinned down, the column permutations of $\hat{\Omega}_1$ and $\hat{\Omega}_2$ must follow, and vice versa. In fact, Google Trends has already been widely employed to proxy news strength and predict asset volatilities (e.g. Vlastakis and Markellos, 2012, Dimpfl and Jank, 2016). In this paper, the global identification is achieved with the help of GTI, too. The search terms are also the same as Kohonen (2013), which can be found in Table B1. The trends of the GTIs for these search terms can be found in Figure B4. Obviously, this plot shows the all the markets in this study receive extreme attention during the Eurozone crisis period.

[**Insert Figure B4 here.**]

In order to apply the Google Trends data to the T-SVAR model, we have to first extract the information on news strength from the raw data series, and then use it to pin down the correct permutations of the columns of $\hat{\Omega}_1$ and $\hat{\Omega}_2$.

Considering that the increases of GTI are likely to be caused by news, whereas the decreases of GTI are usually results of *no news*, I define the news strength for a market as the mean of the positive logarithm changes of respective GTI.

---

[21]See equation (14).

34

Moreover, as I consider two different regimes in my model, the news strength under each regime is actually measured by the weighted mean of the positive logarithm changes of GTI, where the weights are the probability for the system to be in respective regime at each time. Specifically, the news strength for market i under Regime n is given by

$$\text{NewsStrength}_{in} = \frac{\sum_{t=1}^{T} \tau_n^{\text{Logi}}(\text{iTraxx}_t; \hat{\beta}) \max\{\Delta \ln \text{GTI}_{it}, 0\}}{\sum_{t=1}^{T} \tau_n^{\text{Logi}}(\text{iTraxx}_t; \hat{\beta})}.$$

Since the GTI data are at a weekly frequency, concerned about that some news may have long-tail influence in the following week, I also try to replace $\Delta \ln \text{GTI}_{it}$ with the residuals of the AR(1) processes formed by them, and the results turn out to be the same. I report the results in Table B2.

[**Insert Table B2 here.**]

Now that we have the orders of the news strength for each market in under the two regimes, we need to find correct permutations of columns of $\hat{\Omega}_1$ and $\hat{\Omega}_2$ such that the diagonal elements of $\hat{\Sigma}_\varepsilon^{(1)}$ and $\hat{\Sigma}_\varepsilon^{(2)}$ are in the same order as the news strength orders in Table B2. Since there are $5! = 120$ possible column permutations either for $\hat{\Omega}_1$ or $\hat{\Omega}_2$, what we need to do now is to calculate the corresponding $\hat{\Sigma}_\varepsilon^{(1)}$ and $\hat{\Sigma}_\varepsilon^{(2)}$ of all the 120 possible permutations. After that, we can pick out the permutations which correspond to $\hat{\Sigma}_\varepsilon^{(1)}$ and $\hat{\Sigma}_\varepsilon^{(2)}$ with correct ordering of diagonal elements. Luckily, in this example, for both Regimes 1 and 2, there in respectively only one correct permutations of $\hat{\Omega}_1$ and $\hat{\Omega}_2$. The unique global identification of the T-SVAR model is thus achieved.

Once we achieve the global identification, we can calculate the estimations of Υ_1 and Υ_2 in equation (15), which have immediate and important economic interpretations as the instantaneous response of each variable to unit-sized structural shocks. The estimation results are reported in Table B3. The upper half of the table corresponds to $\hat{\Upsilon}_1$ and the lower half to $\hat{\Upsilon}_2$. For each half, the element in i-th row and j-th column indicates the instantaneous response of the i-th market to the unit-sized structural shock originating from Market j. For example, the element in the 1st row and 2nd column in the upper half of the table is -0.319, then it means under Regime 1, when a unit-sized structural shock takes place in the German market, the Greek instantly has a response that has a size of -0.319. The off-diagonal elements actually capture the cross-market spillover effects under each regime. It is interesting to notice that for both regimes the response of the Greek market to positive German news and the response of Irish market to Spanish news are negative. This may be a clue of the roles played by the two countries in the observation periods: even when there is positive news elsewhere, the markets in Greece and Ireland can still begin to go bear themselves. In other word, they are the two likely places where contagion can originate. Nevertheless, it is still not very easy to draw other conclusions only from this table. But this table provides us with a starting point, and in the following sections, I show how

we can unearth more clear and interesting facts from it.

[**Insert Table B3 here.**]

E. *The Spatial Model*

Now that we have obtained the estimations of the T-SVAR model, we can proceed to the T-SAR model implied by the T-SVAR by applying Propositions 2 and 3. The MD estimations of W_1 and W_2, the spatial weight matrix under Regimes 1 and 2 can be found in Table B4.

[**Insert Table B4 here.**]

It is much easier to interpret the spatial weights than Υ_1 and Υ_2. Notice that each row in W_1 and W_2 indicates how "close" the influences from other countries can affect a certain country, and each column indicates how "close" the influences from the corresponding country to other countries. Now compare $\hat{W}_1$ and $\hat{W}_2$ in Table B4 column by column, and it is not difficult to find some interesting facts. First, under Regime 1 the Greek market has quite small influences on every other market, as none of the weights of Greece to other markets exceeds 0.1. However, under Regime 2, the influences of the Greek market surge to more than 0.2, and even reach as high as 0.48 to the Italian market and 0.44 to the Spanish market. This finding confirms the common impression that Greece plays a very important role in exporting contagion during the recent crises. At the same time, many would also have the impression that Germany is the main supporter of the countries in crises, and this is confirmed by the second column in $\hat{W}_2$, too, as the weights of Germany to all other markets except Italy are quite large. Especially, it is Ireland, rather than Greece, receives the largest influence from Germany, and the weight 0.71 is surprisingly high, which hints a very close Germany-Ireland financial bond. However, the significant role played by Germany seems to only exist under Regime 2, and under Regime 1, German influences are much smaller. In other words, Germany is more like the firefighter for the Eurozone who only appears when there is fire. As are shown by the third and fourth column, the influences of the Italian and the Spanish markets generally decay from Regime 1 to Regime 2, which reflects that the crises in both countries are relatively weaker. The weight from Italy to Ireland under Regime 1, 0.86, is quite striking, and this may be a result of the fact that the Irish crisis starts much earlier than the Eurozone crisis, so that many observations when Ireland is in crisis are classified to the periods with higher probability of Regime 1. This guess is also supported by the results in the fifth column, where Ireland shows large influences on other countries under Regime 1 already. And under Regime 2, although the closeness from Ireland to Greece and Italy increases, the closeness from Ireland to both Germany and Spain decreases, and in general, the role of the Irish market does not seem to change too much across different regimes. A panoramic portrayal of the interconnections among the five countries revealed by the spatial weigh matrices is displayed in Figure B5. We can easily perceive from this figure that

36

the lines in the right sub-plot representing Regime 2 are darker than the left sub-plot, which indicates generally closer connections among variables under Regime 2.

[Insert Figure B5 here.]

I then try to test whether the interconnections among the markets are statistically different under different regimes. The null hypothesis is thus set as

$$\mathsf{H}_0^{\text{Equal}}: \quad W_1 = W_2,$$

versus the alternative hypothesis H_1 : the T-SAR model without any additional restriction on W_1 and W_2. This is exactly the example shown in equation (24) in Section IV.B. Since the reduced-form autoregressive coefficients are assumed to be constant across regimes in this application, we can use the results of Specification B in Section IV.B. For this null hypothesis, we have $\text{LR} = 3.133 \times 10^5$. Compared with a χ-square distribution with 20 degrees of freedom, the null hypothesis $\mathsf{H}_0^{\text{Equal}}$ is rejected with the p-value less than 1×10^{-3}.

Next, I present a test for a more interesting null hypothesis: the contagion, defined as the difference between the connectedness of the markets under Regimes 1 and 2, is spread through the cross-border interbank market network. If we regard Regime 1 as the regime without contagion, and Regime 2 as the regime with contagion, then this hypothesis can be interpreted within the T-SAR model. Specifically, under Regime 1, shocks are spread through the network described by W_1, therefore W_1 represents the natural non-contagion-contingent interconnections among the markets even when there is no contagion. Under Regime 2, the non-contagion-contingent interconnections represented by W_1 are still there. However, now there is the contagion, besides the existing interconnections W_1, shocks can now been spread through a new contagion-contingent network, denoting it as $V^{\text{Interbank}}$. As a result of the co-existence of the non-contagion-contingent network W_1 and the contagion-contingent network $V^{\text{Interbank}}$ under Regime 2, the interconnections among the markets now become W_2. Therefore, if we assume the contagion is spread through network V, then we should be able to express W_2 as a linear combination of W_1 and $V^{\text{Interbank}}$. Formally, we can write the null hypothesis as $\mathsf{H}_0^{\text{Diff}}$: There exists at least one $\zeta_1^{\text{Diff}} \in \mathbb{R}^1$ and one $\zeta_2^{\text{Diff}} \in \mathbb{R}^1$ such that

$$W_2 = \zeta_1^{\text{Diff}} W_1 + \zeta_2^{\text{Diff}} V^{\text{Interbank}}.$$

We can see that this null hypothesis is also exactly the example shown in equation (25) in Section IV.B. As long as $V^{\text{Interbank}}$ is given, the null hypothesis can be tested against H_1 : the T-SAR model without any additional restriction on W_1 and W_2.

In this paper, I construct $V^{\text{Interbank}}$ from the cross-border interbank claim data reported by Bank for International Settlements in the same way as Tonzer (2015). Suppose the element in the i-th row and j-th column of $V^{\text{Interbank}}$ is $v_{ij}^{\text{Interbank}}$,

then

$$(33) \qquad v_{ij}^{\text{Interbank}} = \frac{\sum_{t=1}^{T} \text{ForeignClaim}_{ijt}}{\sum_{t=1}^{T} \sum_{j \neq i} \text{ForeignClaim}_{ijt}},$$

where $\text{ForeignClaim}_{ijt}$ is the interbank positions of country i's banking system towards banks in country j at time t. The actual value of $V^{\text{Interbank}}$ is reported in Table B5. Not unexpectedly, all the countries have large claim proportions in Germany and small claim proportions in Greece.

[Insert Table B5 here.]

Now, under the help of the the results of Specification B in Section IV.B again, we obtain the LR statistic for the null hypothesis $\mathsf{H}_0^{\text{Diff}}$: $\text{LR} = 2.979 \times 10^6$. Compared with a χ-square distribution with 21 degrees of freedom, the null hypothesis $\mathsf{H}_0^{\text{Equal}}$ is rejected with the p-value less than 1×10^{-3}. Note that even after the adjustment due to the difference in degrees of freedom between $\mathsf{H}_0^{\text{Equal}}$ and $\mathsf{H}_0^{\text{Diff}}$, the LR statistic for $\mathsf{H}_0^{\text{Diff}}$ is still far larger than $\mathsf{H}_0^{\text{Equal}}$. This implies that it is even worse to assume the contagion is *purely* spread through the cross-border interbank network than to assume there is no contagion at all.

However, one should notice that by this test I *do not* argue that the Eurozone contagion is not spread through the cross-border interbank network. First, the matrix $V^{\text{Interbank}}$ constructed in this paper is only one of the simplest possible constructions of the the cross-border interbank network; and after all, I only argue that $V^{\text{Interbank}}$ cannot *solely* explain the channel of the Eurozone contagion. If we compound $V^{\text{Interbank}}$ with some other networks, we may still get a good portrayal of the contagion channels.

F.   Network Analyses

In this part I present the results of the network analyses. Since the in-measures are not informative in my model, all the measures mentioned in this section are out-measures.

The centrality measures are shown in Table B6. In fact, this table can be regarded as a formal summary of the analysis of the spatial weight matrices in the previous section, as those findings are all reconfirmed here. First, the centrality of Greece grows tremendously from Regime 1 to 2, and the centrality of Germany also sees a large increase. These are quite consistent with the impression about the important roles played by Greece and Germany in the recent Eurozone crisis. The centralities of Italy and Spain drop significantly from Regime 1 to 2, and Ireland maintains highly central in the network under both regimes. All these findings are in full accordance with the analyses of spatial weight matrices in Section VI.E. Additionally, under Regime 2, the most two central countries are Greece and Ireland. Recall that in Section VI.D I show that the markets in only these two countries can still slump even there are positive news everywhere, we

can say Greece and Ireland are the two most important origins of contagion among the selected five countries. These finding hold no matter we use the simple degree centrality measure or the PN centrality measure, which does not only measure how much influence a country has on its neighbors, but also how much influence the country's neighbors have on neighbors' neighbors.

[**Insert Table B6 here.**]

Then I present the results of the clustering analyses in Table B7. First look at the average clustering coefficients of the whole network - the result here is not surprising at all. Since the average clustering coefficients measure the overall connectedness of the entire network, it is no wonder the markets in the five countries become more interconnected under Regime 2, which is associated with crises. However, when we look at the clustering coefficients of each individual country, we will notice some interesting facts that cannot be revealed by centrality analyses. Under Regime 1, the country with the highest clustering coefficient is Greece. Recall that the clustering coefficient indicates how the neighbors are interconnected, the high clustering coefficient of Greece actually suggests that it plays a rather peripheral role under Regime 1 as other countries are closely interconnected even without Greece. In contrast, Italy has the smallest clustering coefficient under Regime 1, which may be a result of its strong centrality under the regime. However, when the system enter Regime 2, the clustering coefficients of these countries change drastically. Italy has become the most peripheral country with the highest clustering coefficient; and now it is Germany that has the smallest clustering coefficient and contributes the most in conglutinating other these countries together. Moreover, although the centrality of Ireland increases from Regime 1 to 2, its clustering coefficient increases, too. This hints that the importance of Ireland may not necessarily increase from Regime 1 to 2.

[**Insert Table B7 here.**]

And finally, the estimations of ρ_1 and ρ_2 can be used as measures of the interconnection strength among the markets, too. For Regime 1, the estimation $\hat{\rho}_1 = 0.876$, and for Regime 2, $\hat{\rho}_1 = 1.044$. For another time, we find the markets under Regime 2 are more closely interconnected than under Regime 1.

VII.   Conclusions

A broad class of economic and financial models can be generalized as the investigation of connectedness of economic variables or entities. Historically, different strands of literature, such as vector auto-regressions, spatial models, and network models all profile the connectedness in different ways. In recent years, there have been a few pioneering attempts to integrate these different strands of literature in one unified framework. And this paper attempts to bridge the structural vector auto-regression, spatial auto-regression, and network models together.

However, sometimes we do not care only about how economic or financial systems are interconnected in isolation, but also how they transition from different

regimes of connectedness in different circumstances. A typical example among others is the empirical modeling of financial contagions, which are usually defined as the difference between the interconnections among entities under in different financial environments, rather than the interconnections as such. Catering to such problems, I try to upgrade the SVAR-spatial-network model with a transition mechanism by introducing the thresholds in the model, which provides considerable convenience for studying the transitions of connectedness of a system.

The SVAR-spatial-network model in this paper can be naturally broken down to three major parts, namely the SVAR, the spatial model, and the network analyses. For the SVAR part, first I show how to obtain an unique identification of an SVAR model without strong prior assumptions on the causal structure by assuming normal mixture distributions of the error terms. This establishes the groundwork for the whole paper. And second, I introduce transitions to the SVAR model by upgrading the SVAR to a threshold SVAR (T-SVAR) model. The threshold function relies on some exogenous variables and can take various forms. For the part of spatial modeling, I demonstrate that a threshold spatial auto-regression (T-SAR) is equivalent to the T-SVAR model with certain linear restrictions. This allows us to estimate the spatial weight matrices, which are good measures of the connectedness among the variables, from the T-SVAR model even without any prior knowledge of the spatial weights. And I then describe how to use the T-SAR model to construct likelihood-ratio tests of a large class of interesting hypotheses. Finally, for the network analyses, based on the intrinsic similarity between spatial models and networks, I show that we can use some newly devised tools for network analyses to reveal additional information of the model.

I then present an application of the model in the paper to the analysis of recent contagions in the Eurozone. I picked the stock markets in five countries, namely Greece, Germany, Italy, Spain, and Ireland as representative countries. In the SVAR part, I establish a crisis indicator based on the European CDS index that can automatically separate the observations into a mixture of two different regimes according to different interconnection patterns of the five markets. Also, with the help of the normal mixture assumption of error terms, I make a clear portrayal of the spillover effects among the markets under each regimes. In the spatial part, I obtain the estimations of the spatial weight matrices in the T-SAR model under different regimes. The spatial weight matrices are good measure of the connectedness among the five markets. And furthermore, concerning the relationship between the system connectedness under different regimes, I construct two formal statistical tests. Defining the contagion as the difference between the connectedness of the five-country system under the two regimes, the first test rejects the hypothesis that contagions do not exist; and the second test rejects the hypothesis that the contagion channel can be solely explained by a cross-border interbank network in Europe. In the end, I carry out the centrality, clustering, and the ρ-analyses of the networks revealed by the T-SAR model. The network average clustering coefficients and ρ-s under different regimes further confirm that

the markets are interconnected differently under different regimes; and the centrality measures and the individual clustering coefficients show the distinctive roles played by the selected five markets under different regimes.

Certainly, the application in studying the contagions in the Eurozone presented here is still far from the depletion of the potential of my model. In fact, in this paper I always try to describe the model in its most general form so as to maker it applicable to as many situations as possible. This being said, their is still considerable room for the improvement of the model in the future. For example, the spatial auto-regression model considered in this paper has a very simple form compared with many other spatial-time series models, and viewing from the angle of the network literature, the network considered in this paper is also quite simple among the immense world of multilayer networks. There is much left for future research.

<div align="center">MATHEMATICAL PROOFS</div>

<div align="center">*A1. Proof of Proposition 2*</div>

Let $\mathcal{H}_{nn} = (\mathbf{1}'_K \otimes \mathcal{R})' \odot \mathrm{vec}(G_n^{\wedge 1})\mathbf{1}'_{K-1}$, then we have

$$\mathcal{H} = \mathrm{diag}(\mathcal{H}_{11}, \mathcal{H}_{22}, \ldots, \mathcal{H}_{NN}).$$

Futher let $\mathcal{G}_n$ be the $((n-1)(K^2 - K) + 1)$-th to the $n(K^2 - K)$-th components of $\mathcal{G}$, then

$$\mathcal{H}\mathcal{G} = \begin{pmatrix} \mathcal{H}_{11}\mathcal{G}_1 \\ \mathcal{H}_{22}\mathcal{G}_2 \\ \vdots \\ \mathcal{H}_{NN}\mathcal{G}_N \end{pmatrix}.$$

By Theorem 3.1 in Scidá (2016), we know that for $n = 1, 2, \ldots, N$, $\mathcal{H}_{nn}\mathcal{G}_n = \mathbf{0}_{K-1}$ is equivalent to $\mathcal{R}\mathsf{A}_n\mathbf{1}_K = \mathbf{0}_{K-1}$, then $\mathcal{H}\mathcal{G} = \mathbf{0}_{N(K-1)}$ is equivalent to

$$\mathscr{R}\mathscr{A}\mathbf{1}_{NK} = \begin{pmatrix} \mathcal{R}\mathsf{A}_1\mathbf{1}_K \\ \mathcal{R}\mathsf{A}_2\mathbf{1}_K \\ \vdots \\ \mathcal{R}\mathsf{A}_N\mathbf{1}_K \end{pmatrix} = \mathbf{0}_{N(K-1)}.$$

$\square$

<div align="center">*A2. Proof of Proposition 3*</div>

The optimization of the MD program can be solved by the Lagrangian method. Suppose that $\mathscr{U}$ has U components. Consider the Lagragian function

$$\mathcal{L} = (\hat{\mathcal{G}} - \mathcal{G}^*)' \Sigma_{\mathcal{G}}^{*-1}(\hat{\mathcal{G}} - \mathcal{G}^*) - \ell'(\mathscr{H}\hat{\mathcal{G}} - \mathscr{U}),$$

wherer ℓ is the Lagragian multiplier. First-order conditions follow:

(A1)
$$\frac{\partial \mathcal{L}}{\partial \hat{\mathcal{G}}'} = 2\Sigma_{\mathcal{G}}^{*-1}(\hat{\mathcal{G}} - \mathcal{G}^*) - \mathscr{H}'\ell = \mathbf{0}_{N(K^2-K)},$$

and

(A2)
$$\frac{\partial \mathcal{L}}{\partial \ell} = \mathscr{H}\hat{\mathcal{G}} - \mathscr{U} = \mathbf{0}_U.$$

Premultiplying equation (A1) by $\mathscr{H}\Sigma_{\mathcal{G}}^*$, we get

$$2\mathscr{H}(\hat{\mathcal{G}} - \mathcal{G}^*) - \mathscr{H}\Sigma_{\mathcal{G}}^*\mathscr{H}'\ell = \mathbf{0}_U.$$

Using equation (A2), it follows that

$$\ell = 2(\mathscr{H}\Sigma_{\mathcal{G}}^*\mathscr{H}')^{-1}\mathscr{H}\hat{\mathcal{G}} - 2(\mathscr{H}\Sigma_{\mathcal{G}}^*\mathscr{H}')^{-1}\mathscr{U}.$$

Substituting it back to equation (A1), we have

(A3) $\quad \hat{\mathcal{G}} = \mathcal{G}^* - \Sigma_{\mathcal{G}}^*\mathscr{H}'(\mathscr{H}\Sigma_{\mathcal{G}}^*\mathscr{H}')^{-1}\mathscr{H}\mathcal{G}^* + \Sigma_{\mathcal{G}}^*\mathscr{H}'(\mathscr{H}\Sigma_{\mathcal{G}}^*\mathscr{H}')^{-1}\mathscr{U}.$

Following equation (A3), it is obvious that

$$\hat{\mathcal{G}}^{\mathrm{T-SAR}} = \mathcal{G}^* - \Sigma_{\mathcal{G}}^*\mathcal{H}'(\mathcal{H}\Sigma_{\mathcal{G}}^*\mathcal{H}')^{-1}\mathcal{H}\mathcal{G}^*,$$

when $\mathscr{H} = \mathcal{H}$, and $\mathscr{U} = \mathbf{0}_{N(K-1)}$. $\qquad\square$

*

REFERENCES

Afonso, A., J. Baxa, and M. Slavik (2011). Fiscal developments and financial stress: A threshold VAR analysis. ECB Working Paper No. 1319.

Barigozzi, M. and C. T. Brownlees (2016). NETS: Network estimation for time series. Available at SSRN: http://ssrn.com/abstract=2249909.

Battiston, F., V. Nicosia, and V. Latora (2013). Metrics for the analysis of multiplex networks. arXiv preprint arXiv:1308.3182.

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

Bonacich, P. and P. Lloyd (2004). Calculating status with negative relations. *Social Networks 26*(4), 331–338.

Caporale, G. M., A. Cipollini, and N. Spagnolo (2005). Testing for contagion: A conditional correlation analysis. *Journal of Empirical Finance 32*(3), 476–489.

Caporin, M., L. Pelizzon, F. Ravazzolo, and R. Rigobon (2013). Measuring sovereign contagion in europe. NBER Working Paper No. 18741.

Chiang, T. C., B. N. Jeon, and H. Li (2007). Dynamic correlation analysis of financial contagion: Evidence from Asian markets. *Journal of International Money and Finance 26*(7), 1206–1228.

De Bruyckere, V., M. Gerhardt, G. Schepens, and R. V. Vennet (2013). Bank/sovereign risk spillovers in the European debt crisis. *Journal of Banking & Finance 37*(12), 4793–4809.

De Domenico, M., A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas (2013). Mathematical formulation of multilayer networks. *Physical Review X 3*(4), 041022.

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

Dimpfl, T. and S. Jank (2016). Can Internet search queries help to predict stock market volatility? *European Financial Management 22*(2), 171–192.

Dungey, M., R. Frya, B. González-Hermosilloc, and V. L. Martind (2005). Empirical modelling of contagion: A review of methodologies. *Quantitative Finance 5*(1), 9–24.

Dungey, M. and V. L. Martin (2007). Unravelling financial market linkages during crises. *Journal of Appllied Econometrics 22*(1), 89–119.

Edwards, S. (1998). Interest rate volatility, capital controls, and contagion. NBER Working Paper No. 6756.

Everett, M. G. and S. P. Borgatti (2014). Networks containing negative ties. *Social Networks 38*, 111–120.

Forbes, K. and R. Rigobon (2001). Measuring contagion: Conceptual and empirical issues. In *International Financial Contagion*, pp. 43–66. Springer, US.

Forbes, K. J. and R. Rigobon (2002). No contagion, only interdependence: Measuring stock market comovements. *The Journal of Finance 57*(5), 2223–2261.

Giraitis, L., G. Kapetanios, A. Wetherilt, and F. Žikeš (2016). Estimating the dynamics and persistence of financial networks, with an application to the sterling money market. *Journal of Applied Econometrics 31*(1), 58–84.

Gourieroux, C. and A. Monfort (1989). A general framework for testing a null hypothesis in a "mixed" form. *Econometric Theory 5*(1), 63–82.

Hansen, B. E. (1997). Inference in TAR models. *Studies in Nonlinear Dynamics & Econometrics 2*(1), 1–16.

Kalbaskaa, A. and M. Gątkowski (2012). Eurozone sovereign contagion: Evidence from the CDS market (2005–2010). *Organization 83*(3), 657–673.

Kilian, L. (2013). Structural vector autoregressions. In *Handbook of Research Methods and Applications in Empirical Macroeconomics*, pp. 515–554. Edward Elgar, UK.

King, M. A. and S. Wadhwani (1990). Transmission of volatility between stock markets. *Review of Financial Studies 1*(3), 5–33.

Kivelä, M., A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter (2014). Multilayer networks. *Journal of Complex Networks 2*(3), 203–271.

Kohonen, A. (2013). On detection of volatility spillovers in overlapping stock markets. *Journal of Empirical Finance 22*, 140–158.

Kon, S. J. (1984). Models of stock returns—a comparison. *The Journal of Finance 39*(1), 147–165.

Lam, C. and P. C. Souza (2015). Estimating the spatial weight matrix using the adaptive LASSO. Working Paper.

Lanne, M. and H. Lütkepohl (2010). Structural vector autoregressions with non-normal residuals. *Journal of Business & Economic Statistics 28*(1), 159–168.

Lee, G. (2006). The effectiveness of international knowledge spillover channels. *European Economic Review 50*(8), 2075–2088.

Li, F. and P. St-Amant (2010). Financial stress, monetary policy, and economic activity. Bank of Canada Working Paper 2010, 12.

Lütkepohl, H. (2007). *New Introduction to Multiple Time Series Analysis*. Berlin Heidelberg: Springer-Verlag.

Maciejowska, K. (2010). Estimation methods comparison of SVAR model with the mixture of two normal distributions—monte carlo analysis. ECO Working Paper 2010/27.

Maliranta, M., P. Mohnen, and P. Rouvinen (2009). Is inter-firm labor mobility a channel of knowledge spillovers? Evidence from a linked employer–employee panel. *Industrial and Corporate Change 18*(6), 1161–1191.

Manasse, P. and L. Zavalloni (2013). Sovereign contagion in Europe: Evidence from the CDS market. Quaderni DSE Working Paper N° 863.

Manresa, E. (2015). Estimating the structure of social interactions using panel data. Working Paper.

McAssey, M. P. and F. Bijma (2015). A clustering coefficient for complete weighted networks. *Network Science 3*(2), 183–195.

Nazlioglu, S., C. Erdem, and U. Soytas (2011). Volatility spillover between oil and agricultural commodity markets. *Energy Economics 33*(3), 497–503.

Patilea, V. and H. Raïssi (2012). Adaptive estimation of vector autoregressive models with time-varying variance: Application to testing linear causality in mean. *Journal of Statistical Planning and Inference 142*(11), 2891–2912.

Scidá, D. (2016). Structural VAR and financial networks: A minimum distance approach to spatial modeling. Working Paper.

Spirtes, P., C. Glymour, and R. Scheines (2000). *Causation, Prediction, and Search* (2nd ed.). Cambridge, Massachusetts: The MIT Press.

Tonzer, L. (2015). Crossborder interbank networks, banking risk and contagion. *Journal of Financial Stability 18*, 19–32.

van Dijk, D., T. Teräsvirtab, and P. H. Fransesa (2002). Smooth transition autoregressive models—a survey of recent developments. *Econometric Reviews 21*(1), 1–47.

Vlastakis, N. and R. N. Markellos (2012). Information demand and stock market volatility. *Journal of Banking & Finance 36*(6), 1808–1821.

Figure B1. : Stock Market Indices

This figure reports daily closing values from January 2nd, 2006 to May 26th, 2015 of the representative stock market indices in five Eurozone countries. All the data are rescaled by their initial values.

Figure B2. : iTraxx Values
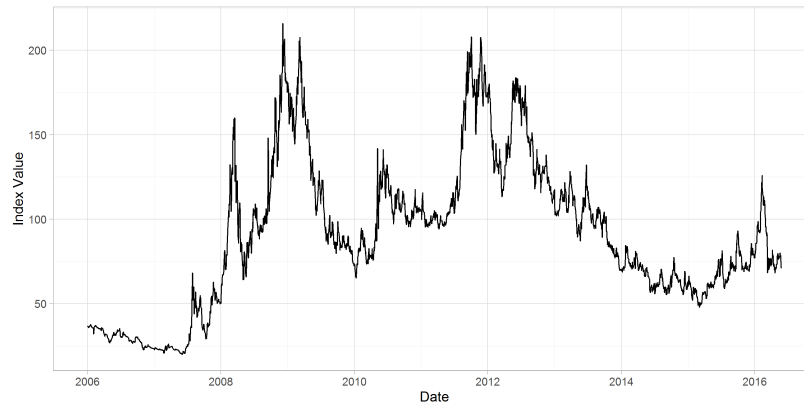
This figure reports daily closing values of iTraxx Europe from January 2nd, 2006 to May 26th, 2015.

Figure B3. : Values of the Threshold Function
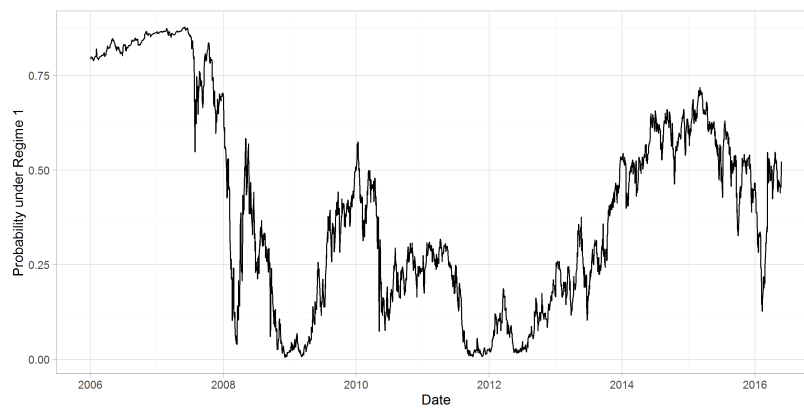
This figure shows the probability for the system to be under Regime 1, which has generally closer market interconnections from January 2nd, 2006 to May 26th, 2015.

47

Figure B4. : Goolge Trends Indices
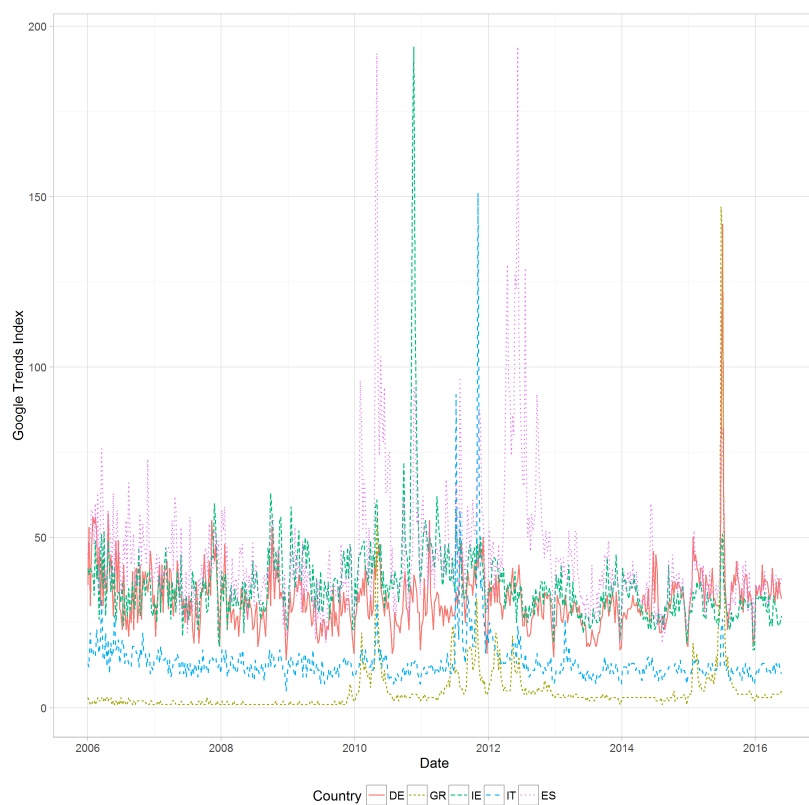
This figure shows the Google Trends Indices for the search terms listed in Table B1 from January 2nd, 2006 to May 26th, 2015. All series keep their original scales.

Figure B5. : An Overview of the Connectedness among the Five-Country System
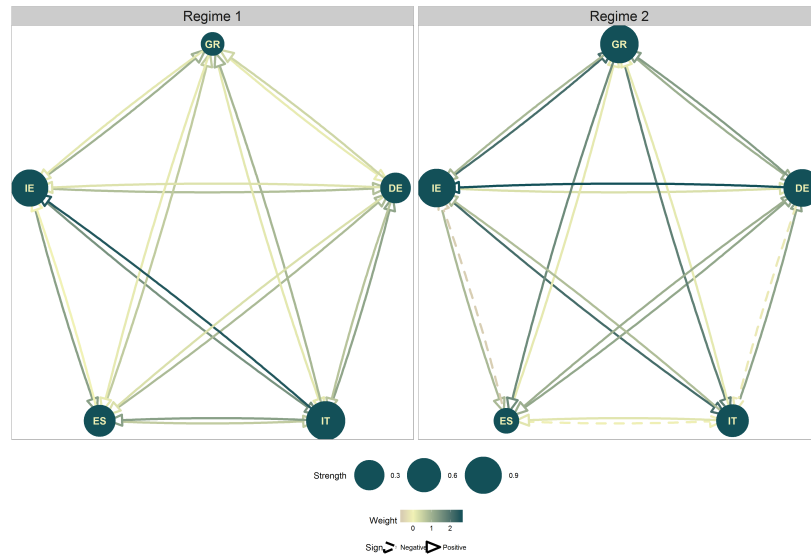
This figure is an overview of the five-country system studied in this paper. The lines represent the closeness among the countries revealed by the T-SAR model, among which dashed lines are negative values. A darker line indicates a closer connection between the countries connected by the line. And the sizes of the balls where country names are printed show the PN-centralities of each country.

Table B1—: Search Terms for the Google Trends Indices

Country	Term
Greece	Greece economy OR Greece debt OR Greece stock market
Germany	Germany economy OR Germany debt OR Germany stock market
Italy	Italy economy OR Italy debt OR Italy stock market
Spain	Spain economy OR Spain debt OR Spain stock market
Ireland	Ireland economy OR Ireland debt OR Ireland stock market

This table shows the search terms for Google Trends Indices for each country .

Table B2—: News Strength

Regime	Country	Log Change Measure		AR(1) Residual Measure	
		News Strength	Rank	News Strength	Rank
	Greece	0.1361	5	0.1344	5
	Germany	0.0973	3	0.0903	3
1	Italy	0.0941	4	0.0861	4
	Spain	0.0971	2	0.0895	2
	Ireland	0.0826	1	0.0772	1
	Greece	0.1123	5	0.1170	5
	Germany	0.0808	2	0.0776	2
2	Italy	0.1013	4	0.0967	4
	Spain	0.0984	3	0.0938	3
	Ireland	0.0744	1	0.0716	1

This table shows the news strength for each country under Regimes 1 and 2 estimated from logarithm changes of Google Trends Indices and the AR(1) residuals thereof. Then then are ranked from the smallest to the biggest.

Table B3—: Instantaneous Responses to Unit-Sized Structural Shocks

Matrix	Responding Variable	Structural Shock Label				
		GR	DE	IT	ES	IE
$\hat{\Upsilon}_1$	GR	3.355	−0.319	0.803	0.141	0.602
	DE	1.268	1.543	1.186	0.977	0.809
	IT	1.270	1.000	1.018	0.531	2.088
	ES	1.141	0.350	1.392	1.344	1.819
	IE	0.978	0.384	2.774	−0.382	0.915
$\hat{\Upsilon}_2$	GR	3.642	−0.259	1.117	0.133	0.340
	DE	1.376	1.251	1.648	0.926	0.458
	IT	1.378	0.810	1.415	0.503	1.181
	ES	1.239	0.284	1.934	1.274	1.029
	IE	1.061	0.312	3.855	−0.362	0.518

This table shows estimations of Υ_1 and Υ_2, which stand for the instantaneous responses of each market return to unit-sized structural shocks. The upper half of the table corresponds to $\hat{\Upsilon}_1$ and the lower half to $\hat{\Upsilon}_2$. For each half, the element in i-th row and j-th column indicates the instantaneous response of the i-th market to the unit-sized structural shock originating from Market j. For example, the element in the 1st row and 2nd column in the upper half of the table is -0.319, then it means under Regime 1, when a unit-sized structural shock takes place in the German market, the Greek instantly has a response that has a size of -0.319.

Table B4—: Estimations of Spatial Weight Matrices

Matrix	Spatial Weight, To	Spatial Weight, From				
		GR	DE	IT	ES	IE
$\hat{W}_1$	GR	0	0.144	0.309	0.204	0.343
	DE	0.037	0	0.407	0.290	0.266
	IT	0.054	0.238	0	0.200	0.509
	ES	0.040	0.092	0.435	0	0.432
	IE	0.059	0.081	0.859	0.002	0
$\hat{W}_2$	GR	0	0.360	0.038	0.034	0.568
	DE	0.278	0	0.330	0.312	0.079
	IT	0.482	−0.041	0	−0.014	0.573
	ES	0.435	0.281	0.050	0	0.233
	IE	0.241	0.707	0.216	−0.164	0

This table shows estimations of the spatial matrices W_1 and W_2. The upper half of the table corresponds to $\hat{W}_1$ and the lower half to $\hat{W}_2$. Since these matrices are asymmetric, I differentiate the weights from a country to the weights to a country.

Table B5—: Value of $V^{\text{Interbank}}$

	GR	DE	IT	ES	IE
GR	0	0.690	0.136	0.060	0.114
DE	0.028	0	0.350	0.377	0.245
IT	0.052	0.841	0	0.119	0.033
ES	0.008	0.506	0.372	0	0.113
IE	0.070	0.431	0.287	0.211	0

This table reports the value of $V^{\text{Interbank}}$ used in this paper, For the meaning of the elements, see equation (33).

Table B6—: Centrality Measures

Country	Degree Measure		PN Measure	
	Regime 1	Regime 2	Regime 1	Regime 2
GR	0.111	1.071	1.119	2.365
DE	0.326	0.975	1.341	2.264
IT	1.182	0.473	2.287	1.573
ES	0.409	0.126	1.426	1.122
IE	0.911	1.084	1.954	2.405

This table reports the degree and PN centralities for each market under different regimes.

Table B7—: Clustering Coefficients

Country	Regime 1	Regime 2
GR	1.092	0.891
DE	1.043	0.867
IT	0.535	1.373
ES	1.000	1.203
IE	0.711	0.885
Network Average	0.876	1.043

This table displays the clustering coefficients for each country as well as the network average clustering coefficients under different regimes.