Discussion Papers

Collana di E-papers del Dipartimento di Economia e Management – University of Pisa

Davide Fiaschi Angela Parenti

An Estimate of the Degree of Interconnectedness between European Regions: A Bayesian Model Averaging Approach

Discussion Paper n. 171

2013

Discussion Paper n. 171, presentato: **October 2013**

Indirizzo degli Autori:

Davide Fiaschi
Dipartimento di Economia e Management, via Ridolfi 10, 56100 PISA – Italy
Email: dfiaschi@ec.unipi.it

Angela Parenti
Dipartimento di Economia e Management, via Ridolfi 10, 56100 PISA – Italy
Email: ireneb@ec.unipi.it@ec.unipi.it

Please cite as it follows
Davide Fiaschi, and Angela Parenti (2013), "An Estimate of the Degree of Interconnectedness between European Regions: A Bayesian Model Averaging Approach", Discussion Papers del Dipartimento di Economia e Management – University of Pisa, n. 171 (http://www-dse.ec.unipi.it/ricerca/discussion-papers.htm).

Discussion Paper
n. 171

Davide Fiaschi - Angela

An Estimate of the Degree of Interconnectedness between European Regions: A Bayesian Model Averaging Approach

Abstract

This paper provides a methodology based on General Variance Decomposition and Bayesian Model Averaging to estimate the degree of economic interconnectedness across different regions, and applies such methodology to a sample of 199 European NUTS2 regions in the period 1980-2008. The estimated connectedness appears very heterogeneous and not symmetric. The idiosyncratic component is not very significant, as well as the common component. A clear pattern of core-periphery exists but not defined in geographical terms. The country component is not very significant, very heterogeneous across countries, and proportional to countries' size. The degree of interconnectedness positively depends on the time horizon of the analysis. Finally, the comparison of the estimated connectedness matrix with two spatial matrices generally used in spatial econometrics (a first-order contiguity and a distance-based matrix) reveals that both are far from representing the actual interconnectedness between European regions.

Classificazione JEL: C11; R11; R12; O50

Keywords: General Variance Decomposition, Bayesian Model Averaging, spatial matrix.

Contents

I. Introduction

The literature on European regions mainly focuses on growth and convergence in income and productivity (see, e.g., Barro and Sala-i Martin (1991), Barro and Sala-i Martin (2004)), and on the geographic concentration of per capita GDP (see, e.g., Le Gallo and Ertur (2003)). Another strand of literature analyzes the degree of interconnectedness and the diffusion of shocks across European regions. In particular, two distinct approaches have been proposed. The first analyzes the correlation among growth cycles of the European regions (see, e.g, De Nardis et al. (1996)), and try to summarize their interactions and co-movements.

The second, originally developed in the regional science and geography literatures, is based on a spatial weight matrix, where the elements of this matrix represent the direction and strength of spillovers between each pair of units. For example, Le Gallo et al. (2003) study the spatial diffusion process in income growth implied by a spatial error model, assuming that the spatial interaction between regions can be represented by an *exogenous* spatial weight matrix based on the inverse of squared distance between regional centroid.

In spatial econometric, "the spatial weight matrix is the fundamental tool used to model the spatial interdependence between regions. More precisely, each region is connected to a set of neighbouring regions by means of a *spatial pattern* introduced *exogenously* as a spatial weight matrix W" (Le Gallo et al. (2003, p.110)). Therefore, the W matrix represents the *a priori* assumption about interaction strength between regions. However, in many cases considerable attention should be given to specifying the matrix W to represent as far as possible economic process (see Corrado and Fingleton (2012)).

The aim of the paper is twofold: i) to find a methodology which allows to estimate a connectedness matrix without take it as exogenous; ii) to apply the methodology to the per capita GDP growth of 199 European NUTS2 regions (EU15) over the period 1980-2008.

The paper it is organized as it follows: Section explains the methodology to calculated a panel of growth rate volatility of per capita GDP and to estimate the associated connectedness matrix. Section III. discusses an application to this methodology to a large sample of European regions. Section IV. concludes.

II. The Methodology

This section firstly discusses how to estimate a panel of growth rate volatilities (GRV) for a sample of regions (Section II.A.); this panel is then used to perform a general variance decomposition analysis (GVD) on the residuals from a VAR in order to estimated the so-called *connectedness matrix* (Section II.B.). The procedure is largely inspired by Diebold and Yilmaz (2011), with the additional difficulty arising in the estimate of VAR, where the number of observations for each region is generally lower than the number of regions, i.e. we face a typically high-dimensional problem (see Hastie et al. (2008)). We will discuss how Bayesian Model Averaging is particularly adapted to deal with this type of problem.[1] Section II.C. discusses how the perspective of network theory provides crucial insights on the mechanics of interconnection behind the connectedness matrix; finally, Section II.E. compares the proposed approach with the two of the most used matrices in spatial econometrics, i.e. first-order contiguity matrix and matrix based on the inverse-distance.

II.A. Estimation of a Panel of GRV

In this section we discuss how to estimate a panel of GRV for a sample of regions/countries following the approach in McConnell and Perez-Quiros (2000) and Fiaschi and Lavezzi (2011). The basic idea is that the dynamics of growth rate of per capita GDP can be well-approximated by an autoregressive process of order p (denoted by AR(p)):

$$\gamma_{jt} = \mu_j + \phi_1 \gamma_{j,t-1} + ... + \phi_p \gamma_{j,t-p} + \epsilon_{jt}, \tag{1}$$

where ϵ_{jt} is assumed to be normally distributed. Given that ϵ_{jt} follows a normal distribution, an unbiased estimator of the standard deviation of ϵ_{jt}, σ^{ϵ}_{jt}, is given by:

$$\hat{\sigma}^{\epsilon}_{jt} = \sqrt{\frac{\pi}{2}} |\hat{\epsilon}_{jt}|. \tag{2}$$

As pointed out in literature (see McConnell and Perez-Quiros (2000)), the absolute value specification make the estimator more robust to departures from the hypothesized conditional normality of errors than the estimator of the variance $\hat{\epsilon}^2_{jt}$. From Eq. (2) we derive the unbiased estimator of the standard deviation of the growth rate of per capita GDP, σ^{γ}_{jt}. For example, if the growth rate follows an AR(1) process (see Hamilton (1994), p.53), the standard deviation of the growth rate is given by:

$$\hat{\sigma}^{\gamma}_{jt} = \frac{\hat{\sigma}^{\epsilon}_{jt}}{\sqrt{1-\phi_1^2}} = \frac{\sqrt{\frac{\pi}{2}} |\hat{\epsilon}_{jt}|}{\sqrt{1-\phi_1^2}}. \tag{3}$$

This method is easily extended to higher-order AR models (see Hamilton (1994), pp. 58-59.). This methodology allows us to build a panel of GRV by exploiting both the cross-sectional and the time-series dimensions of growth rates.

In the empirical analysis we select for each region in the sample the best order of AR using the EIC criteria, that is a bootstrap version of the Akaike Information Criteria for small samples (see Ishiguro et al. (1997)).

[1] In a companion paper the same problem has been dealt with a LASSO approach (see Fiaschi and Parenti (2013)).

II.B. The Connectedness Matrix

As in Diebold and Yilmaz (2011) we use a vector-autoregressive (VAR) model to represent the process governing the GRV of regions, and estimate the generalized variance decomposition (GVD hereafter) which allows to measure the population *connectedness*, i.e. assessing the share of forecast error variance in a region due to shocks arising *elsewhere*.

The use of VAR implicitly implies that relationships across units of observations are essentially linear, and that the contemporaneous relationships are well represented by pairwise correlations (i.e., the variance-covariance matrix).

Moreover, the use of GVD is subject to some restrictive assumptions, the most notable is the Gaussian distribution of shocks.[2]

Assume that a VAR of order p is a good approximating model of the process governing the GRV of regions:[3]

$$\mathbf{x}_t = \mathbf{c} + \sum_{i=1}^{p} \mathbf{\Phi}_i \mathbf{x}_{t-i} + \epsilon_t, \quad t = 1, ..., T, \tag{4}$$

where $\mathbf{c}$ is a $N \times 1$ vector of constants, $\mathbf{x}_t = (x_{1t}, ..., x_{Nt})'$ is a $N \times 1$ vector of jointly determined dependent variables, and $\mathbf{\Phi}_i,\ i = 1, ..., N$ is the $N \times N$ coefficients matrix.

As in Pesaran and Shin (1998), we assume that:

1. $E(\epsilon_t) = \mathbf{0}$, $E(\epsilon_t \epsilon_t') = \Sigma$ for all t, where $\mathbf{\Sigma} = \{\sigma_{ij},\ j = 1, ..., N\}$ is an $N \times N$ positive definite matrix, and $E(\epsilon_t \epsilon_{t'}') = \mathbf{0}$ for all $t \neq t'$.
2. All roots of $|\mathbf{I}_N - \sum_{i=1}^{p} \mathbf{\Phi}_i z^i| = 0$ fall outside the unit circle, that is $\mathbf{x}_t$ is covariance-stationary.
3. $\mathbf{x}_{t-1}, ..., \mathbf{x}_{t-p},\ t = 1, ..., T$ are not perfectly collinear.

Under Assumption 2 Eq. (4) can be rewritten as the infinite moving average representation:

$$\mathbf{x}_t = \mu + \sum_{i=0}^{\infty} \mathbf{\Theta}_i \epsilon_{t-i}, \quad t = 1, ..., T, \tag{5}$$

where $\mu = (\mathbf{I}_N - \mathbf{\Phi}_1 - \cdots - \mathbf{\Phi}_p)^{-1}\mathbf{c}$ is the mean of the process, and the $N \times N$ coefficient matrices $\mathbf{\Theta}_i$ can be obtained as:

$$\mathbf{\Theta}_i = \mathbf{\Phi}_1 \mathbf{\Theta}_{i-1} + ... + \mathbf{\Phi}_p \mathbf{\Theta}_{i-p}, \quad i = 1, 2, ... \tag{6}$$

with $\mathbf{\Theta}_0 = \mathbf{I}_N$ and $\mathbf{\Theta}_i = 0$ for $i < 0$.

[2]Alternatively, the use of Cholesky-factor identification is sensitive to ordering of the units of observations.

[3]Notation refers to Pesaran and Shin (1998).

To measure the effect of shocks at a given point in time on the expected future values of variables in a dynamical system, Koop et al. (1996) advance the generalized impulse response function. In particular, denoting the history of the economy up to time t as $\mathbf{\Omega}_{t-1}$, the generalized impulse response function of $\mathbf{x}_t$ at horizon H is defined as:

$$\mathbf{GI}_x(H, \delta, \mathbf{\Omega}_{t-1}) = E(\mathbf{x}_{t+H}|\epsilon_t = \delta, \mathbf{\Omega}_{t-1}) - E(\mathbf{x}_{t+H}|\mathbf{\Omega}_{t-1}) = \mathbf{\Theta}_H \delta. \tag{7}$$

If we consider the generalized impulse response function to be conditional not on all the shocks at time t, but on just one of them, we get:

$$\mathbf{GI}_x(H, \delta_j, \mathbf{\Omega}_{t-1}) = E(\mathbf{x}_{t+H}|\epsilon_{jt} = \delta_j, \mathbf{\Omega}_{t-1}) - E(\mathbf{x}_{t+H}|\mathbf{\Omega}_{t-1}). \tag{8}$$

If the vector of random shocks, ϵ_t, has a multivariate normal distribution the conditional expectation of the shocks is a liner function of the single shock δ_j, that is:

$$E(\epsilon_t|\epsilon_{jt} = \delta_j) = (\sigma_{1,j}, ..., \sigma_{Nj})' \sigma_{jj}^{-1} \delta_j = \mathbf{\Sigma} \mathbf{e}_j \sigma_{jj}^{-1} \delta_j; \tag{9}$$

where $\mathbf{e}_j$ is an $N \times 1$ selection vector with unity as its j-th element and zero elsewhere.

Therefore, the vector of *unscaled* generalized impulse response of the effect of a shock in the j-th equation at time t on $\mathbf{x}_t$ is given by:

$$\mathbf{GI}_x(H, \delta_j, \mathbf{\Omega}_{t-1}) = \left(\frac{\mathbf{\Theta}_H \mathbf{\Sigma} \mathbf{e}_j}{\sqrt{\sigma_{jj}}} \right) \left(\frac{\delta_j}{\sqrt{\sigma_{jj}}} \right). \tag{10}$$

In particular, by setting $\delta_j = \sqrt{\sigma_{jj}}$ we get the *scaled* generalized impulse response function:

$$\psi_j^g(H) = \sigma_{jj}^{-\frac{1}{2}} \mathbf{\Theta}_H \mathbf{\Sigma} \mathbf{e}_j; \tag{11}$$

which measure the effect of one standard error shock to the j-th equation at time t on expected values of $\mathbf{x}$ at time $t + H$.

From the above generalized impulses, Pesaran and Shin (1998) derive the generalized (i.e., order-invariant) forecast error variance decomposition, defined as the proportion of the H-step ahead forecast error variance of variable i which is accounted for by innovations in variable j in the VAR. Then, for $H = 1, 2, ...$[4], H-step GVD matrix $D^{gH} = [d_{ij}^{gH}]$ has entries:

$$d_{ij}^{gH} = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' \mathbf{\Theta}_h \mathbf{\Sigma} e_j)^2}{\sum_{h=0}^{H-1} (e_i' \mathbf{\Theta}_h \mathbf{\Sigma} \mathbf{\Theta}_h' e_i)} \tag{12}$$

where e_j is the selection vector, Θ_h is the coefficient matrix of the h-lagged shock vector in the MA representation of the *non-orthogonalized* VAR, $\mathbf{\Sigma}$ is the covariance matrix of the shock in the *non-orthogonalized* VAR, and σ_{ii} its diagonal.

[4]Notice that $H = 1$ actually corresponds to the contemporaneous connectedness.

As in Diebold and Yilmaz (2011) we normalize the GVD matrix by row in order to have unity sums of forecast error variance contribution (remember that the shocks are not necessarily orthogonal). Therefore, the *connectedness matrix* has entries as $\tilde{d}_{ij}^{gH} = \frac{d_{ij}^{gH}}{\sum_{j=1}^{N} d_{ij}^{gH}}$.

	x_1	...	x_N	From Others
x_1	$\tilde{d}_{11}^{gH}$	...	$\tilde{d}_{1N}^{gH}$	$\sum_{j=1}^{N} \tilde{d}_{1j}^{gH}$, $j \neq 1$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_N	$\tilde{d}_{N1}^{gH}$	...	$\tilde{d}_{NN}^{gH}$	$\sum_{j=1}^{N} \tilde{d}_{Nj}^{gH}$, $j \neq N$
To Others	$\sum_{i=1}^{N} d_{i1}^{H}$, $i \neq 1$	...	$\sum_{i=1}^{N} \tilde{d}_{iN}^{gH}$, $i \neq 1$	$\frac{1}{N} \sum i, j = 1^N \tilde{d}_{ij}^{gH}$, $i \neq j$

Table 1: Connectedness Matrix derived from the GVD Matrix.

In particular, $\tilde{d}_{ij}^{gH}$ is the ij-th H-step ahead variance decomposition component, that is the fraction of region's i H-step forecast error variance due to shocks in region j. The *cross*-variance decomposition, that is the off-diagonal elements (i.e., $i \neq j$), measure the *pairwise directional connectedness* and, in general, can be different (i.e., $\tilde{d}_{ij}^{gH} \neq \tilde{d}_{ji}^{gH}$) as the bilateral imports/exports. On the other hand, the diagonal elements (*own connectedness*) measure the fraction of region's i H-step forecast error variance due to shocks arising in the same region.

The connectedness matrix is based on the predictive horizon H that is related to the concept of dynamic connectedness. In particular, the generalized variance decomposition 1-step ahead represents the contemporaneous connectedness. As the predictive horizon increases there is more possibility for connectedness to appear. In this sense, we can distinguish between short-run and long-run connectedness.

II.C. Network Interpretation of Connectedness Matrix

The proposed methodology has a straightforward interpretation in terms of network and of percolation of shocks through it. For the sake of simplicity consider the case with three regions and VAR(1), whose variance-covariance matrix Σ is given by:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & 0 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ 0 & \sigma_{32} & \sigma_{33} \end{bmatrix} \tag{13}$$

from which the GVD matrix at $H = 1$:

$$D^{g1} = \begin{bmatrix} 1 & \frac{\sigma_{12}}{\sigma_{22}} & 0 \\ \frac{\sigma_{21}}{\sigma_{11}} & 1 & \frac{\sigma_{23}}{\sigma_{33}} \\ 0 & \frac{\sigma_{32}}{\sigma_{22}} & 1 \end{bmatrix} \tag{14}$$

In case of VAR(1) $\Theta_0 = I_N$, $\Theta_1 = \Phi$, $\Theta_2 = \Phi^2$, ..., where Φ is the coefficient matrix of VAR.

The network representation related to D^{g1} in Eq. (14) is reported in Fig. (1). The structure of *contemporaneous* network fully reflects the shape of Σ both in terms of existence of links and in terms of their strength. However, differently from Σ, D^{g1} is not symmetric, i.e. the contemporaneous network is both *weighted* and *directional*. In the analysis of connectedness matrix D we generally consider row-standardize elements in the spirit of the methodology of GVD. Adopting a network approach this normalization could have not effect on the analysis only if we focus on the existence of links (as it is the case for the analysis in Section III.D.); a study of the percolation of shocks through network, on the contrary, would request to use the original values; but this is out of the scope of paper.

Assuming that the coefficient matrix Θ_1 of the *1*-lagged shock vector in the MA representation of the *non-orthogonalized* VAR(1) is given by:

$$\Theta_1 = \Phi = \begin{bmatrix} \phi_{11} & 0 & \phi_{13} \\ 0 & 0 & 0 \\ \phi_{31} & 0 & \phi_{33} \end{bmatrix} \tag{15}$$

Therefore, the GVD matrix at $H = 2$ is given by:

$$D^{g2} = \begin{bmatrix} \dfrac{\sigma_{11} + \phi_{11}^2\sigma_{11}}{\sigma_{11} + \phi_{11}^2\sigma_{11} + \phi_{13}^2\sigma_{33}} & \dfrac{\sigma_{12}}{\sigma_{22}}\left[1 + \left(\phi_{11} + \phi_{13}\frac{\sigma_{32}}{\sigma_{12}}\right)^2\right] & \dfrac{\phi_{13}^2\sigma_{33}}{\phi_{33}\phi_{13}\sigma_{33} + \phi_{31}\phi_{11}\sigma_{11}} \\ \\ \dfrac{\sigma_{21}}{\sigma_{11}} & 1 & \dfrac{\sigma_{23}}{\sigma_{33}} \\ \\ \dfrac{\phi_{31}^2\sigma_{11}}{\phi_{11}\phi_{31}\sigma_{11} + \phi_{13}\phi_{33}\sigma_{33}} & \dfrac{\sigma_{32}}{\sigma_{22}}\left[1 + \left(\phi_{33} + \phi_{31}\frac{\sigma_{12}}{\sigma_{32}}\right)^2\right] & \dfrac{\sigma_{33} + \phi_{33}^2\sigma_{33}}{\sigma_{33} + \phi_{33}^2\sigma_{33} + \phi_{31}^2\sigma_{11}} \end{bmatrix} \tag{16}$$

The network representation related to D^{g2} in Eq. (16) is reported in Fig. (2). The structure of network appears crucially affected by Φ both in terms of the emergence of new links and in terms of their strength.

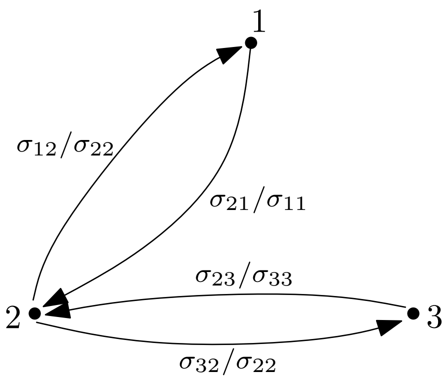

Figure 1: Network representation of GVD matrix at horizon $H = 1$

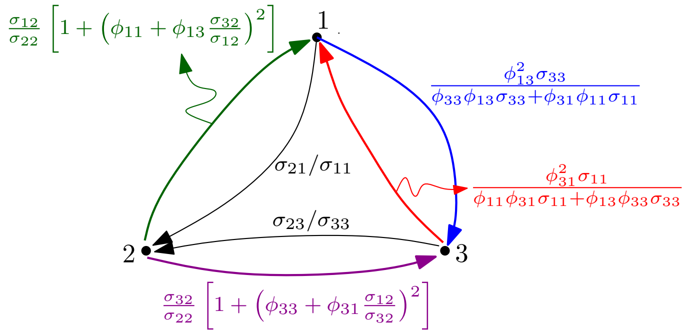

Figure 2: Network representation of GVD matrix at horizon $H = 2$

In particular, new links appear connecting Regions 1 and 3 through the VAR coefficients ϕ_{13} and ϕ_{31}.

VAR coefficients also drive the extent of persistence of shocks over time; for example, d_{11}^{g2} depends on ϕ_{11} (the effect of autoregressive component of Region 1) and ϕ_{13} (the shocks received from Region 3); the coefficients appear to have a power proportional to time horizon (i.e. 2), that is shocks have an exponential decay. It is straightforward to show that a longer time horizon increases the strength of links with an exponential decay (for example d_{11}^{g3} would present terms like ϕ_{11}^4 and ϕ_{13}^2).

Region 2, missing any lag with itself and with the other regions in the VAR, displays a network partially independent of the time horizon considered. In particular, the connectedness from Region 2 to other regions are affected through the contemporaneous covariances σ_{12}, σ_{22}, and σ_{32}, while the connectedness from other regions to Region 2 are not affected.

Relaxing the assumption of VAR(1), for example in favor of VAR(2), increases both the percolation of shocks through network and their persistence, but the qualitative results remain the same.

II.D. The Estimation of Connectedness Matrix via Bayesian Model Averaging

The typical dimensions of datasets used in cross-country and cross-region analysis are such that the number of countries/regions N is much higher than the length of time series T, i.e. we are facing a high-dimensional problem with $N \gg T$. This suggests to maintain the order of VAR at the minimum level equal to 1, i.e. GRV of each region at time t will depend on a constant, on its lagged GRV at time $t-1$, and on the GRV of all other regions at time $t-1$.

We depart from Diebold and Yilmaz (2011) considering also a common component affecting the GRV of all regions in the sample.[5]

[5]In the empirical analysis as common component we will use the GRV estimated for the

Since $N \gg T$, the VAR(1) cannot be estimated by standard OLS: the total number of parameters to be estimated equal to $K = N + 2$, i.e. all lagged GRV of regions plus common component plus constant, is higher than the number of observations T. This problem is here overcome by using a Bayesian Model Averaging (BMA) approach.

Many empirical application involve linear regression model with a large number of potential explanatory variables. Including all the potential variables in the regression can decrease the accuracy of the estimation as long as some of the variables are irrelevant. The traditional approach is to do a sequence of tests with the aim of selecting a single best model which omits all the irrelevant variables. However, "each time a test is carried out, there is a possibility that a mistake will be made (i.e. the researcher will reject the 'better' model for a 'not so good one'). The probability of making a mistake quickly increases as sequences of tests are carried out. Secondly, even if a sequential testing procedure does lead to a selection of the 'best' model, standard decision theory implies that it is rarely desirable to simply present results for this model and ignore all evidence from the 'not quite so good' models. By doing so *model uncertainty* is ignored." (Koop (2003, p.267)).

The logic of BMA is that one should obtain results for every model under consideration and average them, using the posterior model probabilities as weights.[6] In particular, given a dependent variable Y, a number of observations T, and a set of candidate regressors $X_1, ..., X_K$, the variable selection problem is to find the most effective subset of regressors. Consider R different models, M_r for $r = 1, ..., R$, where each one represents a subset of the candidate regressors. Model M_r has the form $Y = \alpha + \sum_{j=1}^{K_r} \beta_j^{(r)} X_j^{(r)} + \epsilon$, where $X_1^{(r)}, ..., X_{K_r}^{(r)}$ is a subset of $X_1, ..., X_K$, $\beta^{(r)} = (\beta_1^{(r)}, ..., \beta_{K_r}^{(r)})$ is a vector of regression coefficients to be estimated, and $\epsilon \sim N(0, \sigma^2)$ is the error term. Finally, $\theta_r = (\alpha, \beta^{(r)}, \sigma)$ is the vector of parameters in M_r.

The likelihood function of model M_r, $p(D|\theta_r, M_r)$, summarizes all the information about θ_r that is provided by the data, D. Using the law of total probability, it is possible to calculate from the likelihood function the *integrated likelihood* $p(D|M_r)$. Denoting $p(M_r)$ the prior probability that M_r is the correct model given that of the models is considered, the *posterior model probability* of M_r, $p(M_r|D)$ is given by the model's share of the total posterior mass:

$$p(M_r|D) = \frac{p(D|M_r)p(M_r)}{\sum_{r=1}^{R} p(D|M_r)p(M_r)}. \tag{17}$$

Therefore, the *posterior inclusion probability* (PIP) of a candidate regressor, $p(\beta_j \neq 0|D)$, can be obtained by summing the posterior model probabilities

growth of the aggregated EU15 per capita GDP.

[6]See for example (Eicher et al., 2011).

across those models that include the regressor. The PIP can be interpreted as the probability that each regressor should be included.

The implementation of the BMA has two main problems. First, if K is the number of potential explanatory variables, the number of possible models is 2^K. Therefore, if K is large, the number of possible models is astronomical. In this respect we follow the literature and we use the so-called MC^3 algorithms for the estimation.

Second, the implementation of BMA in linear regression is subject to the specification of the prior distributions. For BMA, the prior has two parts: a prior for the parameters of each model, and the prior probability of each model (see (Fernandez et al., 2001)). If substantial prior information is available and can be expressed as a probability distribution, this should be used. Often, however, the prior information is small relative to the information in the data, and then it makes sense to use a default prior. Priors on parameters may affect results since they may influence the integrated likelihood $p(D|M_r)$, which is a key component of the posterior model weights. We follow Fernandez et al. (2001) in setting the prior for parameters and, in particular, the Zellner's g-prior is set to $g_r = 1/K^2$ given that in our case $T << K^2$.[7] As regards the model priors, we use the *random theta* prior by Ley and Steel (2009), who suggest a binomial-beta hyperprior on the a priori inclusion probability $p(M_r)$.[8]

The goal of model selection is to choose a model for future prediction. Under the Bayesian approach, the model with highest PIP is not necessarily the optimal predictive model. Barbieri and Berger (2004) show that, for selection among normal linear models, the optimal predictive model is often the *median probability model*, which is defined as the model consisting of those variables which have overall PIP greater than or equal to 1/2 of being in a model.[9] As estimation of VAR(1) we will therefore use the *median probability model*.

II.E. Connectedness Matrix versus Spatial Matrix

As discussed in the introduction one of the main goals of the paper is to get some insights on the shape of spatial matrix $\mathbf{W}$, which in spatial literature measures the spatial dependence across different regions (see (Anselin, 2001) for a general introduction).

$\mathbf{W}$ is generally taken as exogenous in spatial literature and it is specified or in term of geographic contiguity or in terms of geographical distance (see (Anselin, 2001)). Corrado and Fingleton (2012) formulate three main critiques to current literature: i) the values in the cells of $\mathbf{W}$ comprise an explicit hypothesis about the

[7] Values of g_r that are closer to zero imply priors that are less informative.

[8] The estimation is carried out using the *bms* function of the *BMS* R-package.

[9] A different approach not pursued here is proposed by Brown et al. (2002): BMA for variable selection can be well approximated by the *median probability model* when orthogonality of both the prior and x-variables can be assumed; the latter can be however suboptimal when these orthogonality conditions fails.

strength of interlocation connection", in particular, "*a priori* assumption about interaction strength"; ii) "Typically, isotropy is assumed, so that only distance between j and h is relevant, not the direction j to h"; iii) "The potential for dynamic $\mathbf{W}$ matrices poses some problems for estimation, given the assertion that $\mathbf{W}$ is necessarily a fixed entity. While this my not be such an issue for cross-section approaches, [...], with the extension of spatial econometrics to include panel data modelling it may be the case that $\mathbf{W}$ is evolving."

To discuss how our connectedness matrix D^{gH} is strictly related to $\mathbf{W}$ assume that the data generating process of GRV of regions y follows:

$$\mathbf{y}_t = (\mathbf{I}_N - \rho\mathbf{W})^{-1}[\mu_N + \mathbf{X}_t\beta + \mathbf{u}_t] = (\mathbf{I}_N - \rho\mathbf{W})^{-1}[\mu_N + \mathbf{X}_t\beta] + \mathbf{v}_t \tag{18}$$

where μ_N is the vector of fixed effects of length N, $\mathbf{y}_t$ is a vector of length N, $\mathbf{X}_t$ is a matrix of dimensions (N, k), β is a vector of length k of coefficients, and $\mathbf{u}_t$ is the error component of dimension N and $\mathbf{v}_t$ is the spatially filtered error component of dimension N. The error component $\mathbf{u}_t$ is specified as:

$$\mathbf{u}_t = \lambda\mathbf{W}\mathbf{u}_t + \psi\mathbf{W}\mathbf{u}_{t-1} + \phi\mathbf{u}_{t-1} + \epsilon_t, \tag{19}$$

where ϵ_t is the vector of innovations, with $E[\epsilon_t,] = 0$, $E[\epsilon_t\epsilon'_t,] = \sigma^2_\epsilon I_N$, and $E[\epsilon_t\epsilon'_{t'},] = 0$ for each $t' \neq t$. Eq. (19) reflect the possibility that $\mathbf{u}_t$ can display mixed dynamics in both space and time.[10] We follow the literature assuming that $\mathbf{W}$ is the same for the spatially lagged dependent variable and the errors.

From Eq. (19) we derive:[11]

$$\mathbf{u}_t = (\mathbf{I}_N - \lambda\mathbf{W})^{-1}\left\{\sum_{i=0}^{\infty}\phi^i\left[(\mathbf{I}_N + \psi/\phi\mathbf{W})(\mathbf{I}_N - \lambda\mathbf{W})^{-1}\right]^i\epsilon_{t-i}\right\}$$

from which we get the variance-covariance matrix of $\mathbf{u}_t$, $\mathbf{U}$, for all t, i.e.:

$$\begin{aligned}\mathbf{U} = E[\mathbf{u}_t\mathbf{u}'_t] &= \sigma^2_\epsilon(\mathbf{I}_N - \lambda\mathbf{W})^{-1}\times\\ &\times\left[\mathbf{I}_N - \phi^2(\mathbf{I}_N - \lambda\mathbf{W}')^{-1}(\mathbf{I}_N + \psi/\phi\mathbf{W}')(\mathbf{I}_N + \psi/\phi\mathbf{W})(\mathbf{I}_N - \lambda\mathbf{W})^{-1}\right]^{-1}\times\\ &\times(\mathbf{I}_N - \lambda\mathbf{W}')^{-1}\end{aligned} \tag{20}$$

Therefore, the variance-covariance matrix of $\mathbf{v}_t$, $\mathbf{V}$, for all t is given by:

$$\mathbf{V} = E[\mathbf{v}_t\mathbf{v}'_t] = (\mathbf{I}_N - \rho\mathbf{W})^{-1}\mathbf{U}(\mathbf{I}_N - \rho\mathbf{W}')^{-1}. \tag{21}$$

Assuming that the VAR representation is an approximating model of $\mathbf{y}_t$ (see Diebold and Yilmaz (2011, p. 11)), a possible estimation of $\mathbf{V}$ is given by the variance-covariance matrix of the VAR(1) model, i.e. $\hat{\Sigma}$.

---

[10]See Elhorst (2013) for a general introduction to spatial panel models.

[11]We are assuming that the first-order spatial autoregressive process is ergodic.

The approximation of stochastic process through a VAR representation allows to overcome the *incidental parameters* problem discussed, e.g., in Anselin (2002). In our model the total number of parameters to be estimate is equal to $N^2 - N + 5$ (all no-zero elements of W plus λ, ϕ, ψ, ρ, and σ_ϵ^2) and the number of observations are equal to N^2 (the elements of V); under the assumption $N > 5$ it is therefore possible to estimate the elements of $\mathbf{W}$ as well as the other parameters of Eq. (21). But the estimate of $\mathbf{W}$ from $\hat{\mathbf{\Sigma}}$ becomes very unreliable already for small N: for our sample of $N = 199$ observations, the total number of parameters o estimate is equal to 39407 against a number of observations equal to 39601.

However, a comparison between Eqq. (12) and (21) makes clear that D^{gH} and $\mathbf{W}$ are calculated on the same information set, i.e. $\hat{\mathbf{\Sigma}}$.

Moreover, the comparison highlights how spatial panels whose observations refer to variable with different timing (e.g. panel of annual observations versus panel with five-year average observations) should include different spatial matrix reflecting the different degrees of interconnectedness (five-year average observations are likely to have a higher level of interconnectedness). A similar argument is made in network analysis (see Newman (2009)).

III. Empirical Results

In this section we first discuss the estimated GRV of European regions; then we proceed to show the results of the estimate of VAR(1); finally we investigate the properties of the estimated connectedness matrix in light of network approach and in comparison to two popular spatial matrices used in spatial literature.

III.A. Estimated GRV for 199 European Regions

Our sample consists of the growth rates of per capita GDP of 199 European NUTS2 regions belong to EU15[12] over the period 1980-2008.[13]

[12] Appendix A contains the list of regions and the best AR order selected by EIC.

[13] Data on per capita GDP come from Econometrics (2010).

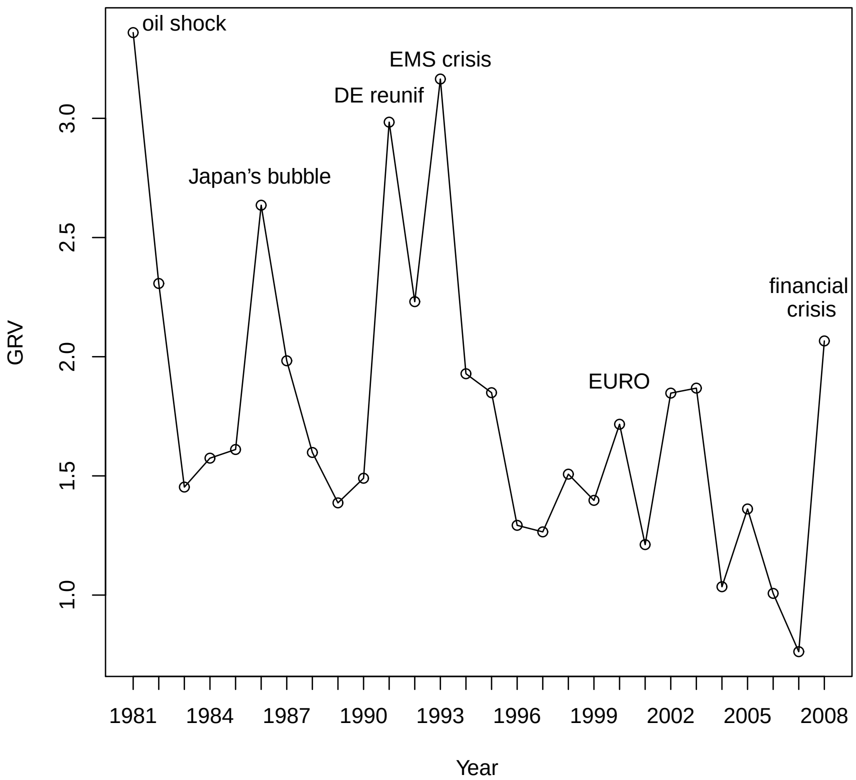

Figure 3: Estimated cross-section averages of GRV for the period 1980-2008 for the sample of 199 European regions

Figure 4 reports the sample *average* volatility for the period 1980-2008. The peaks in volatility can be associated with different shocks occurred during this period, as the 1979 oil crisis, the Japan's bubble economy of the 1980s, the German reunification and the 1992 EMS crisis. The 1999-2002 period correspond to the introduction of euro,[14] while the last peak in volatility displayed in the Fig. 4 reflected the financial crisis of 2007-2008.

Figure 4 report the time averages for each regions in the sample. The overall impression is the presence of a spatial pattern of core-periphery (with regions in the center of Europe displaying less GRV), but it is not clear if there exists also a country component (except for Greece, Norway and Finland).

[14]The euro came into existence on 1 January 1999 even though the currency was only virtually, and in 2002 notes and coins began to circulate.

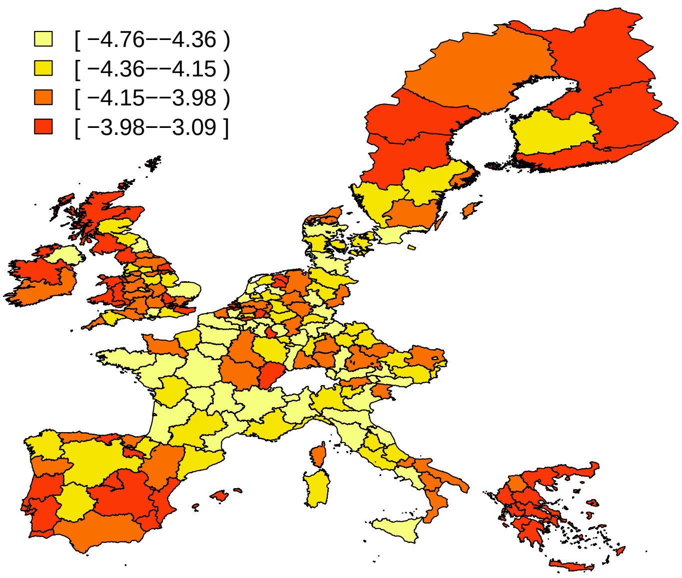

Figure 4: Estimated (log of) time averages of GRV for the sample of 199 European regions

III.B. The Estimate of VAR(1) Coefficients $\hat{\Phi}$

As showed in Section II.C. the connectedness structure appears crucially affected by the estimated VAR coefficients $\hat{\Phi}$ both in terms of the emergence of new links and in terms of their strength.

Figure (5) shows that at most each region has 21 sources of shocks (including is own lagged value and the common factor) in the VAR, and that the volatility of some regions only depends on the constant (i.e, none of the lagged values of volatilities are significantly different from zero).[15] Differently, the distribution of the number of region to which shocks are transmitted is more symmetrically distributed around 17, as showed in Figure (6), and that at most a region is a source of shocks for other 29 regions in the VAR.

[15]As showed in Section II.C. this does not implies that these regions are disconnected given that the $\hat{\Sigma}$ matrix is full.

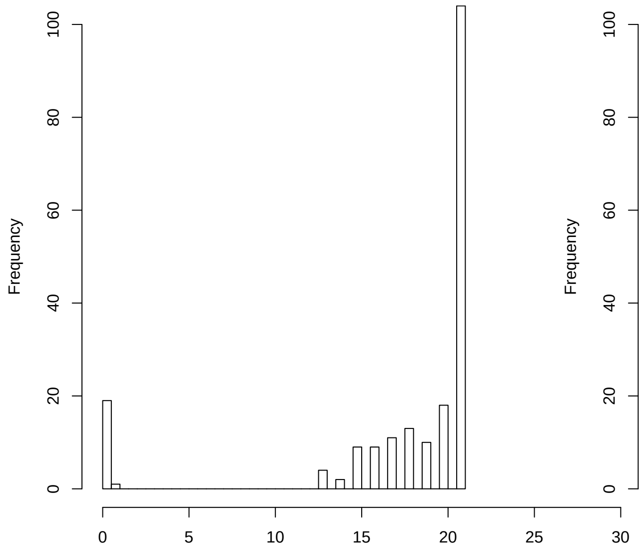

Figure 5: Significant VAR(1) coefficients: From Regions

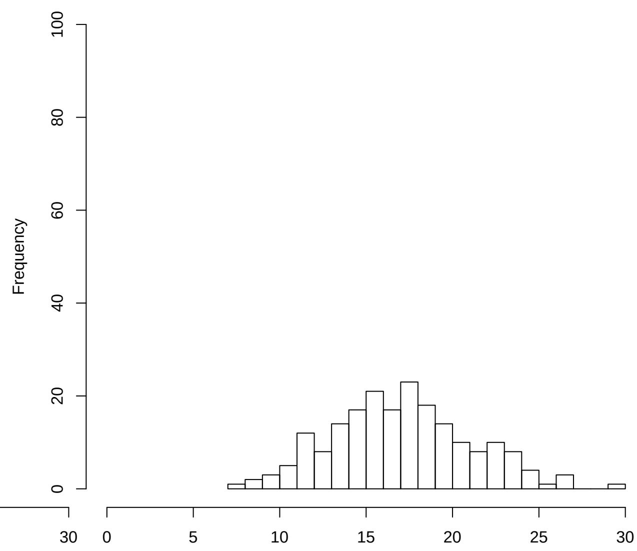

Figure 6: Significant VAR(1) coefficients: To Regions

III.C. Summary of Connectedness Matrix

The full-sample connectedness matrix is 199×199; therefore, it is not possible to look at the entire matrix. In order to get information on the degree of connectedness between European regions we give some summary of the matrix and a network visualization.

First of all we consider the contemporaneous connectedness, that is the variance decomposition 1-year ahead, $\tilde{D}^{g1}$.

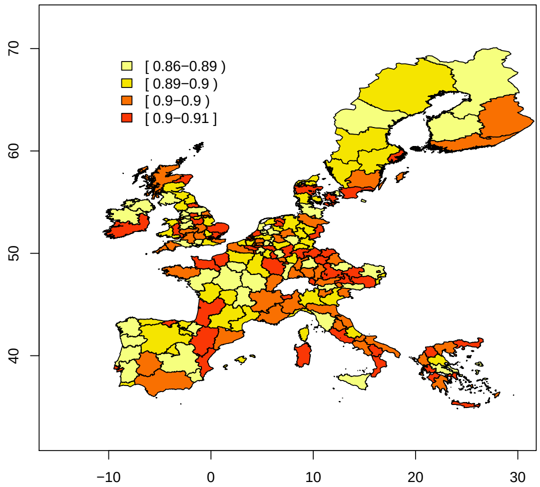

Figure 7: From Regions

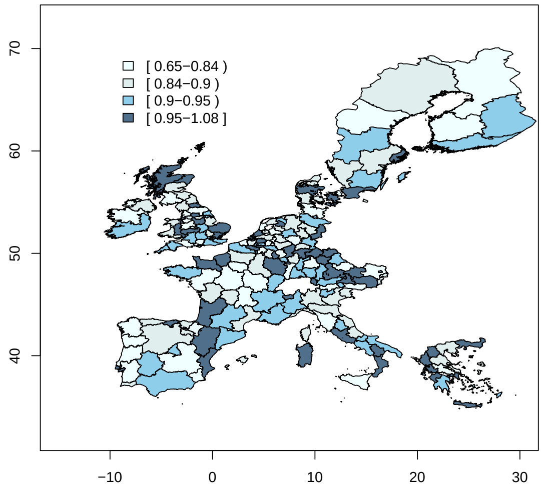

Figure 8: To Regions

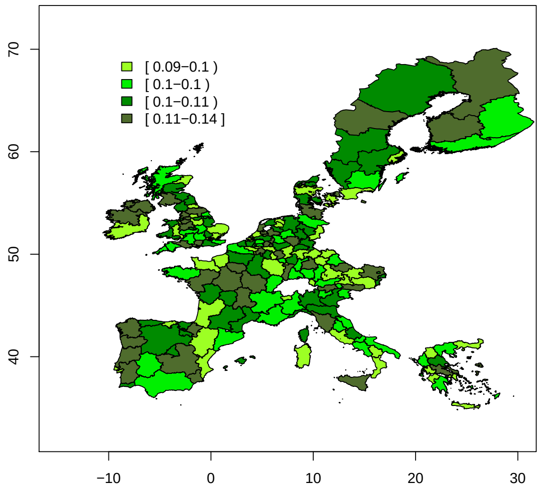

Figure 9: Idiosyncratic Factor

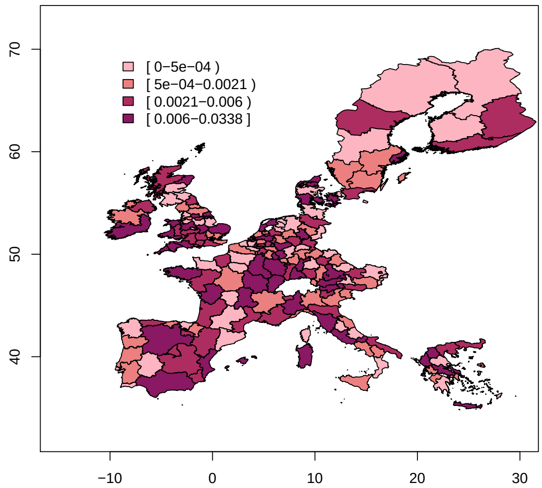

Figure 10: Common Factor

Figure 7 shows the *total connectedness* ***from*** *others*, that is the 1-year forecast-error variance of each region coming from shocks arising in all other regions, while Figure 8 shows the *total connectedness* ***to*** *others*, that is the 1-year forecast-error variance coming from shocks arising in each region and transmitted to all other regions. We notice that regions with high connectedness form others also have high connectedness to other and, therefore, they are highly interconnected. These is the case of Northern regions of the United Kingdom, most of the Southern regions of Italy and Greece, a cluster of neighboring regions across Germany and Denmark, and a cluster of neighboring regions across France and Spain.

Figure 9 shows the *idiosyncratic factor*, that is the fraction of 1-year forecast error variance due to a shock arising in the region itself. The idiosyncratic factor range from about 9% in Dytiki Ellada (GR23) to about 14% in Mellersta Norrland (SE32) and it has a significant impact for regions of Portugal, Norther Ireland and some regions in the center of France. Finally, Figure 10 shows the *common factor*, that is the fraction of 1-year forecast error variance due to a shock that is common for all European regions. The impact of common factor is low, ranging from 0% to 3%. However, some regions as Tuscany (ITE1), Mellersta Norrland (SE32), Dytiki Ellada (GR23), Auvergne (FR72) and Bourgogne (FR26) have both high impact of the idiosyncratic and common factors and, therefore, seem to be poorly interconnected.

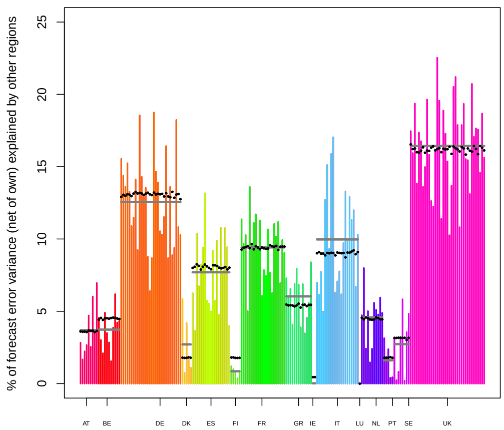

Figure 11: Country Component for each region with the mean value (grey lines) and the value under the null hypothesis of equal impact of all regions (black points).

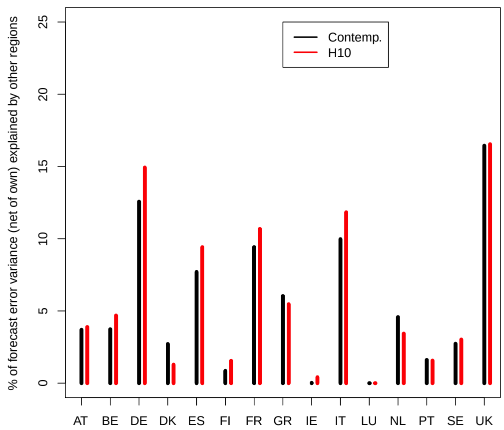

Figure 12: Country Component in the short-run and long-run: aggregate level

Figure 11 shows, for each region, how much of the contemporaneous connectedness is due to shocks arising in regions belonging to the same country (but excluding the region itself). For each region we also report the value of the percentage of forecast error variance explained by the other regions belonging to the same country under the null hypothesis of equal impact of each region in the sample (black points) and the mean value across each country's regions (grey lines).[16] For most of the regions belonging to Belgium, Finland and Sweden the contemporaneous connectedness due to shocks arising in regions belonging to the

[16]In particular, the null hypothesis of absence of country effect for region i is that the share of its total forecast error variance due to the regions belonging to the same country of i should

same country is lower than the value under the null hypothesis of equal impact, implying that these regions tend to be more connected to the rest of Europe while for most of the regions of Denmark, Greece and Italy the country component is more important.

Figure 12 shows, at aggregate level, how much of the forecast error variance can be attributed to the country component both in the short-run (i.e. 1-year forecast error variance) and in the long-run (i.e., 10-years forecast error variance).[17] The long-run connectedness of Germany, Spain, France and Italy is higher than the contemporaneous connectedness. This implies that in the long-run the shocks arising in these countries are more important and, therefore, they show a lower degree of interconnectedness. On the contrary, for Denmark and Netherlands the country component is less important in the long-run.

III.D. Network Approach to the Analysis of Interconnectedness

As stated by Diebold and Yilmaz (2011, p.8) "[...] variance decompositions are networks. More precisely, the variance decomposition matrix D, which defines our connectedness table, and all associated connectedness measure, is a network adjacency matrix A. Hence network connectedness measures can be used in conjunction with variance decompositions to understand connectedness among components". Specifically, variance decompositions define weighted, directed networks.

To analyze connectedness matrices as a networks only links, i.e. pairwise directional connectedness, with a strength greater than 2.12% will be considered; in other words, we consider that region i and region j are linked only if the fraction of region's i H-step forecast error variance due to shocks in region j, d_{ij}^{H}, is higher than 2.12% corresponding to a significance level equal to 0.025.[18] Moreover, we consider only the existence of the link without its weight (that is, we construct from the connectedness matrix a 0-1 adjacent matrix corresponding to a directed but unweighted network).

be equal to:

$$\sum_{j \in N_i^C} \tilde{d}_{i,j}^{g1} = \frac{N_i^C}{N-1}\left(1 - \tilde{d}_{ii}^{g1} - \tilde{d}_{i(N+1)}\right), \tag{22}$$

where N_i^C is the set of regions in the same country of region i, and $\tilde{d}_{i(N+1)}$ is the forecast error variance due to common component.

[17] We still exclude the region itself.

[18] In particular, the 2.12% threshold is derived by simulating the variance-covariance matrix of the VAR(1) under the null hypothesis: $E(\epsilon_t) = \mathbf{0}$, $E(\epsilon_t \epsilon_t') = \hat{\sigma}_i^2 \mathbf{I_N}$ for all t, where $\hat{\sigma}_i^2$ is the estimated variance of the VAR(1) (i.e. diag($\hat{\Sigma}$)), and $E(\epsilon_t \epsilon_{t'}') = \mathbf{0}$ for all $t \neq t'$. Using this simulated variance-covariance matrix we therefore estimate the 1-year ahead GVD matrix under the same null hypothesis and we get the 2.12% threshold which corresponds to the fraction of 1-step ahead forecast error variance, $d_{ij}^{H=1}$, derived from a covariance σ_{ij} that is significantly different from zero at 2.5% significance level. Since under the null-hypothesis also all coefficients of VAR are zero, this threshold holds also for each H-step ahead forecast error variance.

Table 2 reports some basic statistics of the networks for different H (1 (contemporaneous), 5, 10, and 20). We follow the notation in Newman (2009) labelling by n the number of vertices (for our sample the number of regions plus common component), m the number of edges (the number of no-zero links), c the mean degree (i.e. m/c, S the fraction of vertices in the largest (weakly connected) component), l the mean geodesic distance (any two no-connected links are excluded by calculation), d the diameter of network (the length of the longest finite geodesic path), C the average clustering coefficient (based on transitivity in weak form), and r the assortative coefficient.

	N	m	c	S	l	d	C	r
H1	200	1195	5.97	1	3.31	6	0.15	0.83
H5	200	1185	5.92	1	2.97	8	0.14	0.12
H10	200	1232	6.16	1	3.07	7	0.1	0.28
H20	200	1291	6.46	1	3.06	6	0.09	0.41

Table 2: Characteristics of the networks derived from the GVD matrix with H1, H5, H10, H20. N is the number of vertices (for our sample the number of regions plus common component), m the number of edges (the number of no-zero links), c the mean degree (i.e. m/c, S the fraction of vertices in the largest (weakly connected) component), l the mean geodesic distance (any two no-connected links are excluded by calculation), d the diameter of network (the length of the longest finite geodesic path), C the average clustering coefficient (based on transitivity in weak form), and r the assortative coefficient (see Newman (2009)).

Connectedness measured by c (and by the inverse of l) increases with time horizon. Clustering is instead decreasing with time horizon, but remains at high level suggesting a core-periphery structure.

Fundamental characteristic of a network is its degree distribution, that is a description of relative frequencies of nodes that have a given degree d under a certain degree distribution, where the degree of a node is the number of links to other nodes. In particular, in a *regular network* all nodes have the same degree, while a *Poisson random network* has a Poisson degree distribution. Another prominent distribution is a *scale-free distribution* where the relative probabilities of degree of fixed ratio are the same independent of the scale of those degrees. Scale-free distribution has "fat tails", that is there have many more nodes with very small and very large degrees with respect to a Poisson distribution.

Figures 13-15 show the degree distributions (in log-log) of the connectedness matrices with H=1, 5 and 10 years respectively. In all cases, the degree distributions seem to follow a Poisson distribution. This implies that in all variance decomposition matrices (with H=1, 5 and 10 years) there is a great deal of heterogeneity in the degrees of nodes.

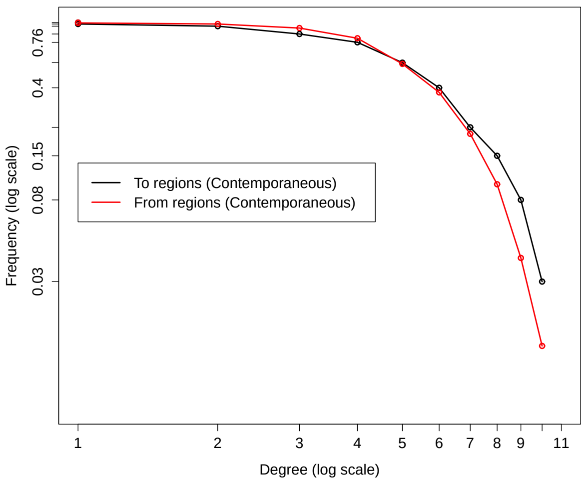

Figure 13: Log Degree H1. Black line is the degree distribution *to* other nodes, while red line is the degree distribution *from* other nodes.

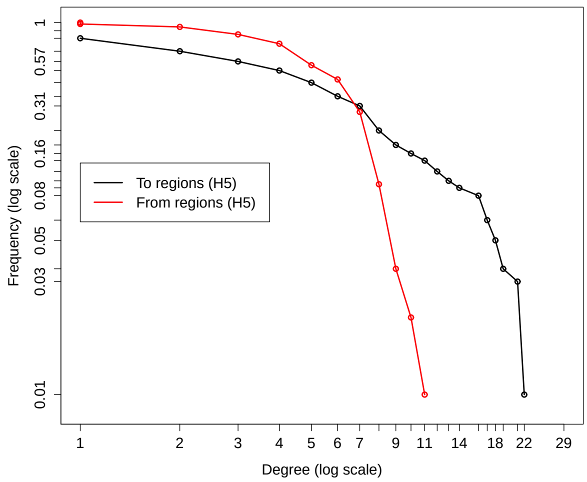

Figure 14: Log Degree H5. Black line is the degree distribution *to* other nodes, while red line is the degree distribution *from* other nodes.

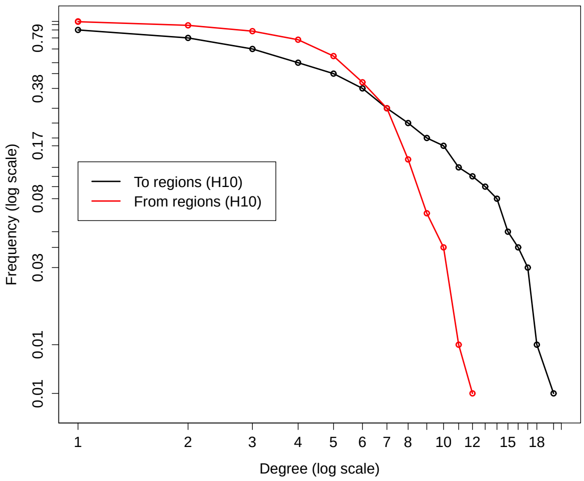

Figure 15: Log Degree H10. Black line is the degree distribution *to* other nodes, while red line is the degree distribution *from* other nodes.

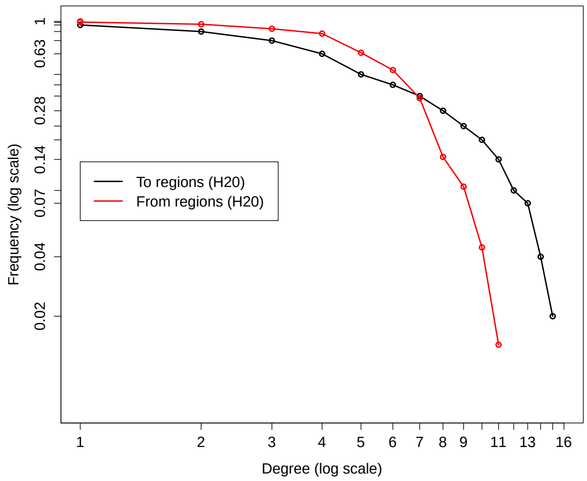

Figure 16: Log Degree H20. Black line is the degree distribution *to* other nodes, while red line is the degree distribution *from* other nodes.

Finally, using a Kamada-Kawai algorithm Figures 17 and 18 a provides a graphical visualization of the estimated network across European regions.[19] The

[19]The aim of the Kamada-Kawai algorithm is to find a set of coordinates in which, for each

colors of the vertices are the same for regions belonging to the same country and correspond to those in Figure (11).

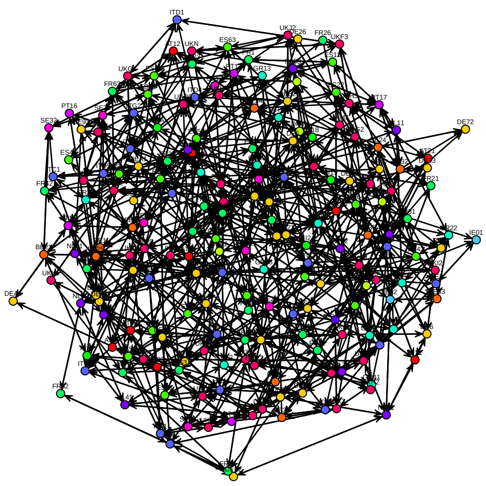

Figure 17: Kamadakawai network with threshold=2.12% for adjacent matrix derived form DgH.Norm.H1

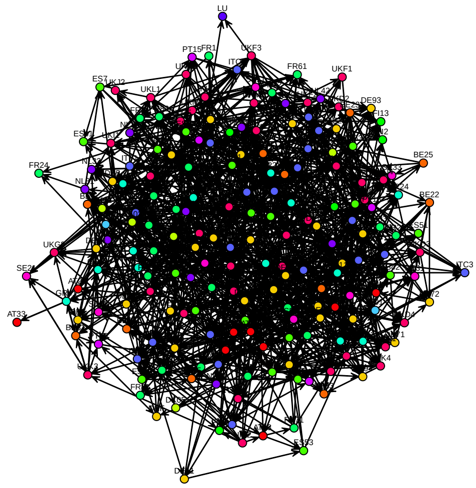

Figure 18: Kamadakawai network with threshold=2.12% for adjacent matrix derived form DgH.Norm.H10

The higher degree of connectedness of network with time horizon of 10 years is evident, as well as the core-periphery structure of network.

III.E. Connectedness Matrix versus Spatial Matrix

In order to compare our connectedness matrix with some spatial weight matrix used in the spatial econometric literature, we construct two networks derived from the GVD of the variance-covariance matrix of a spatial model as the one in Eq. (21) of Section II.E.. In particular we assume two different spatial weight matrices, i.e. a first-order contiguity matrix, W_{cont}, and a distance based matrix with cut-off, $W_{invDistQ1}$, (both row-standardized) whose weights are given by:

$$w_{cont}(i,j) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ share a border} \\ 0 & \text{otherwise} \end{cases}$$

$$w_{invDistQ1}(i,j) = \begin{cases} d_{ij}^{-2} & \text{if } d_{ij} < 370 \text{ miles} \\ 0 & \text{otherwise} \end{cases}$$

pair of nodes, the Euclidean distance is approximately proportional to the geodesic distance between two nodes.

and we calibrate the parameters of the model (ρ=0.32, λ=0.42, ϕ=0.32, ψ=0 and σ_ϵ^2=1) to get networks which are similar to our contemporaneous network in terms of mean degrees (see Table **??**).[20]

	n	m	c	S	l	d	C	r
H1	199	1188	**5.97**	1	3.3	6	0.15	0.83
W_{cont}	199	1082	**5.44**	18	2.47	16	0.56	0.94
$W_{invDistQ1}$	199	1237	**6.22**	5	4.05	21	0.63	0.97

Table 3: Characteristics of the networks derived from the GVD matrix with H1, W_{cont} and $W_{invDistQ1}$.

The network graph visualization is given in Figures (19)-(20). As before, the colors of the vertices are the same for regions belonging to the same country and correspond to those in Figure (11).

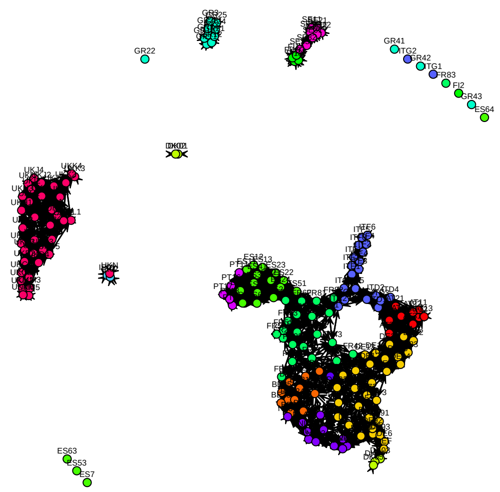

Figure 19: Kamadakawai network with threshold=2.12% for adjacent matrix derived from the spatial model with W_{cont}

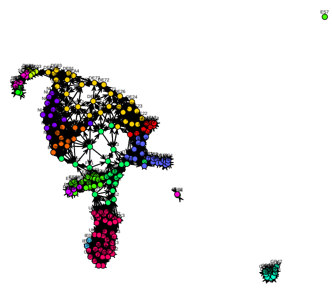

Figure 20: Kamadakawai network with threshold=2.12% for adjacent matrix derived from the spatial model with $W_{invDistQ1}$

[20]Obviously we use the same level of significance on the pairwise directional connectedness equal to 2.12% and we build the 0-1 adjacent matrices corresponding to directed but unweighted networks.

IV. Concluding Remarks

In this paper we have discussed how it is possible to estimate the interconnections between European regions by a connectedness matrix. The latter has the advantage to be immediately interpretable as a network between regions; finally we have investigate how the most popular spatial matrices used in the spatial literature have features completely different with respect to our estimated connectedness matrix.

The use of Bayesian Model Averaging has been motivated by the high-dimensional problem we face; however, this is not the only approach that it is possible to follow; in a companion paper we illustrate how a LASSO approach can produce a better result (see Fiaschi and Parenti (2013)).

As regards the empirical applications a next step is to apply the proposed methodology to the analysis of the interconnection between the dynamics of per capita GDP of countries.

Acknowledgements We are very grateful to PRIN meeting participants in Pisa and Alessandria for their comments. The usual disclaimers apply. D.F. and A. P. have been supported by the Italian Ministry of University and Research as a part of the PRIN 2009 program (grant protocol number 2009H8WPX5).

Affiliation and Address of the Authors

Davide Fiaschi: Dipartimento di Economia e Management, University of Pisa, Pisa, Italy.

Angela Parenti: Dipartimento di Economia e Management, University of Pisa, Pisa, Italy.

References

Anselin, L. (2001). Spatial econometrics. In B. Baltagi (Ed.), *A Companion to Theoretical Econometrics*, pp. 310–330. Oxford: Blackwell.

Anselin, L. (2002). Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural economics 27*(3), 247–267.

Barbieri, M. M. and J. O. Berger (2004). Optimal predictive model selection. *The Annals of Statistics 32*(3), 870–897.

Barro, R. and X. Sala-i Martin (1991). Convergence across states and regions. *Brooking Papers on Economic Activity 1*, 107–182.

Barro, R. J. and X. Sala-i Martin (2004). *Economic Growth.* Cambridge: MIT press.

Brown, P. J., M. Vannucci, and T. Fearn (2002). Bayes model averaging with selection of regressors. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64*(3), 519–536.

Corrado, L. and B. Fingleton (2012). Where is economics in spatial econometrics? *Journal of Regional Science 52*(2), 210–239.

De Nardis, S., A. Goglio, and M. Malgarini (1996). Regional specialization and shocks in europe: some evidence from regional data. *Weltwirtschaftliches Archiv 132*(2), 197–214.

Diebold, F. X. and K. Yilmaz (2011). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *NBER Working Paper 17490*.

Econometrics, C. (2010). *European Regional Database*. Cambridge: Cambridge Econometrics Limited.

Eicher, T. S., C. Papageorgiou, and A. E. Raftery (2011). Default priors and predictive performance in bayesian model averaging, with application to growth determinants. *Journal of Applied Econometrics 26*(1), 30–55.

Elhorst, J. (2013). Spatial panel models. *Handbook of Regional Science*.

Fernandez, C., E. Ley, and M. F. Steel (2001). Benchmark priors for bayesian model averaging. *Journal of Econometrics 100*(2), 381–427.

Fiaschi, D. and A. M. Lavezzi (2011). *Economic Growth and Development (Frontiers of Economics and Globalization, Volume 11)*, Chapter Growth Volatility and the Structure of the Economy, pp. 203–245. Emerald Group Publishing Limited.

Fiaschi, D. and A. Parenti (2013). An estimate of the degree of interconnectedness between european regions: A lasso approach. *Mimeo*.

Hamilton, J. (1994). *Time Series Analysis*. Princeton: Princeton University Press.

Hastie, T., R. Tibshirani, and J. J. H. Friedman (2008). *The Elements of Statistical Learning*. New York: Springer.

Ishiguro, M., Y. Sakamoto, and G. Kitagawa (1997). Bootstrapping log likelihood and eic, an extension of aic. *Annals of the Institute of Statistical Mathematics 49*(3), 411–434.

Koop, G. (2003). *Bayesian econometrics*. J. Wiley.

Koop, G., M. H. Pesaran, and S. M. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics 74*(1), 119–147.

Le Gallo, J. and C. Ertur (2003). Exploratory spatial data analysis of the distribution of regional per capita gdp in europe, 1980-1995. *Journal of Economics 82*, 175–201.

Le Gallo, J., C. Ertur, and C. Baumont (2003). *European Regional Growth*, Chapter A Spatial Econometric Analysis of Geographic Spillovers and Growth for European Regions, 1980-1995, pp. 99–129. Springer.

Ley, E. and M. F. Steel (2009). On the effect of prior assumptions in bayesian model averaging with applications to growth regression. *Journal of Applied Econometrics 24*(4), 651–674.

McConnell, M. and G. Perez-Quiros (2000). Output fluctuations in the united states: What has changed since the early 1980's? *The American Economic Review 90*, 1464–1476.

Newman, M. (2009). *Networks: an Introduction*. Oxford University Press.

Pesaran, H. H. and Y. Shin (1998). Generalized impulse response analysis in linear multivariate models. *Economics letters 58*(1), 17–29.

A List of NUTS2 Regions and selected AR order

Region Code	Region Name	AR order selected by EIC	Region Code	Region Name	AR order selected by EIC
AT11	Burgenland (A)	0	DK01	Hovedstaden	0
AT12	Nieder?sterreich	0	DK02	Sj?lland	0
AT13	Wien	1	DK03	Syddanmark	0
AT21	K?rnten	0	DK04	Midtjylland	0
AT22	Steiermark	0	DK05	Nordjylland	0
AT31	Ober?sterreich	0	ES11	Galicia	0
AT32	Salzburg	0	ES12	Principado de Asturias	0
AT33	Tirol	0	ES13	Cantabria	0
AT34	Vorarlberg	0	ES21	Pais Vasco	0
BE1	R?gion de Bruxelles-Capitale	1	ES22	Comunidad Foral de Navarra	0
BE21	Prov. Antwerpen	0	ES23	La Rioja	0
BE22	Prov. Limburg (B)	0	ES24	Arag?n	1
BE23	Prov. Oost-Vlaanderen	0	ES3	Comunidad de Madrid	1
BE24	Prov. Vlaams Brabant	1	ES41	Castilla y Le?n	0
BE25	Prov. West-Vlaanderen	1	ES42	Castilla-la Mancha	1
BE31	Prov. Brabant Wallon	0	ES43	Extremadura	0
BE32	Prov. Hainaut	0	ES51	Catalu?a	1
BE33	Prov. Li?ge	1	ES52	Comunidad Valenciana	1
BE34	Prov. Luxembourg (B)	1	ES53	Illes Balears	0
BE35	Prov. Namur	0	ES61	Andalucia	1
DE11	Stuttgart	0	ES62	Regi?n de Murcia	0
DE12	Karlsruhe	0	ES63	Ciudad Aut?noma de Ceuta (ES)	0
DE13	Freiburg	0	ES64	Ciudad Aut?noma de Melilla (ES)	0
DE14	T?bingen	0	ES7	Canarias (ES)	1
DE21	Oberbayern	0	FI13	It?-Suomi	0
DE22	Niederbayern	0	FI18	Etel?-Suomi	0
DE23	Oberpfalz	0	FI19	L?nsi-Suomi	0
DE24	Oberfranken	0	FI1A	Pohjois-Suomi	0
DE25	Mittelfranken	0	FI2	?land	0
DE26	Unterfranken	0	FR1	Ille de France	0
DE27	Schwaben	0	FR21	Champagne-Ardenne	0
DE5	Bremen	0	FR22	Picardie	0
DE6	Hamburg	0	FR23	Haute-Normandie	2
DE71	Darmstadt	0	FR24	Centre	0
DE72	Gie?en	0	FR25	Basse-Normandie	0
DE73	Kassel	0	FR26	Bourgogne	0
DE91	Braunschweig	0	FR3	Nord - Pas-de-Calais	0
DE92	Hannover	0	FR41	Lorraine	0
DE93	L?neburg	0	FR42	Alsace	0
DE94	Weser-Ems	0	FR43	Franche-Comt?	0
DEA1	D?sseldorf	0	FR51	Pays de la Loire	0
DEA2	K?ln	0	FR52	Bretagne	0
DEA3	M?nster	0	FR53	Poitou-Charentes	0
DEA4	Detmold	0	FR61	Aquitaine	0
DEA5	Arnsberg	0	FR62	Midi-Pyr?n?es	0
DEB1	Koblenz	0	FR63	Limousin	0
DEB2	Trier	0	FR71	Rh?ne-Alpes	0
DEB3	Rheinhessen-Pfalz	0	FR72	Auvergne	0
DEC	Saarland	0	FR81	Languedoc-Roussillon	0
DEF	Schleswig-Holstein	0	FR82	Provence-Alpes-C?te d'Azur	0
			FR83	Corse	0

Region Code	Region Name	AR order selected by EIC	Region Code	Region Name	AR order selected by EIC
GR11	Anatoliki Makedonia, Thraki	0	PT15	Algarve	0
GR12	Kentriki Makedonia	0	PT16	Centro (PT)	0
GR13	Dytiki Makedonia	0	PT17	Lisboa	0
GR14	Thessalia	0	PT18	Alentejo	0
GR21	Ipeiros	0	SE11	Stockholm	0
GR22	Ionia Nisia	0	SE12	?stra Mellansverige	0
GR23	Dytiki Ellada	0	SE21	Sm?land med ?arna	0
GR24	Sterea Ellada	0	SE22	Sydsverige	0
GR25	Peloponnisos	0	SE23	V?stsverige	0
GR3	Attiki	0	SE31	Norra Mellansverige	2
GR41	Voreio Aigaio	1	SE32	Mellersta Norrland	0
GR42	Notio Aigaio	0	SE33	?vre Norrland	0
GR43	Kriti	0	UKC1	Tees Valley and Durham	0
IE01	Border, Midlands and Western	0	UKC2	Northumberland, Tyne and Wear	0
IE02	Southern and Eastern	0	UKD1	Cumbria	1
ITC1	Piemonte	0	UKD2	Cheshire	0
ITC2	Valle d'Aosta	0	UKD3	Greater Manchester	1
ITC3	Liguria	0	UKD4	Lancashire	1
ITC4	Lombardia	1	UKD5	Merseyside	0
ITD1	Provincia Autonoma Bolzano	0	UKE1	East Yorkshire, Northern Lincolnshire	0
ITD2	Provincia Autonoma Trento	0	UKE2	North Yorkshire	0
ITD3	Veneto	0	UKE3	South Yorkshire	0
ITD4	Friuli-Venezia Giulia	0	UKE4	West Yorkshire	1
ITD5	Emilia-Romagna	0	UKF1	Derbyshire, Nottinghamshire	0
ITE1	Toscana	3	UKF2	Leicestershire, Rutland, Northants	0
ITE2	Umbria	0	UKF3	Lincolnshire	0
ITE3	Marche	0	UKG1	Herefordshire, Worcestershire, Warks	0
ITE4	Lazio	0	UKG2	Shropshire, Staffordshire	0
ITF1	Abruzzo	0	UKG3	West Midlands	1
ITF2	Molise	0	UKH1	East Anglia	1
ITF3	Campania	1	UKH2	Bedfordshire, Hertfordshire	1
ITF4	Puglia	0	UKH3	Essex	1
ITF5	Basilicata	0	UKI1	Inner London	1
ITF6	Calabria	1	UKI2	Outer London	1
ITG1	Sicilia	0	UKJ1	Berkshire, Bucks and Oxfordshire	1
ITG2	Sardegna	0	UKJ2	Surrey, East and West Sussex	1
LU	LUXEMBOURG	1	UKJ3	Hampshire and Isle of Wight	0
NL11	Groningen	0	UKJ4	Kent	0
NL12	Friesland (NL)	0	UKK1	Gloucestershire, Wiltshire, Bristol	1
NL13	Drenthe	2	UKK2	Dorset and Somerset	1
NL21	Overijssel	1	UKK3	Cornwall and Isles of Scilly	1
NL22	Gelderland	0	UKK4	Devon	0
NL31	Utrecht	0	UKL1	West Wales and The Valleys	1
NL32	Noord-Holland	0	UKL2	East Wales	1
NL33	Zuid-Holland	0	UKM2	Eastern Scotland	0
NL34	Zeeland	0	UKM3	South Western Scotland	0
NL41	Noord-Brabant	1	UKM5	North Eastern Scotland	0
NL42	Limburg (NL)	1	UKM6	Highlands and Islands	0
PT11	Norte	0	UKN	Northern Ireland	1