Testing for the Validity of W in $GVAR$ models

Bertrand Candelon[a,1], Angelo Luisi[a,1]

[a] *Université catholique de Louvain, Louvain Finance.*

Abstract

Global *VAR*s are one of the most established econometric frameworks to analyse cross-country/markets interconnections. Nevertheless, they are based on pre-specified channels of interdependence (proxied by the matrix of distance W) that are never empirically tested. This paper develops a simple Likelihood Ratio Test for the validity of the proposed channels, assesses its asymptotic size and power, and proposes a bootstrapped alternative to avoid finite sample distortions. In the empirical application regarding euro area sovereign bond yields modelling, the most popular channels get rejected. The nonrejected W outlines the importance of contagion and flight-to-quality mechanisms even before the sovereign debt crisis.

Keywords: *Global VAR, Structural VAR, Likelihood Ratio Test, Interdependence*

JEL Classification: C12, C32, C52, E43, H63

Email addresses: `candelonb@gmail.com` (Bertrand Candelon), `angelo.luisi@uclouvain.be` (Angelo Luisi)

[1] Voie du Roman Pays, 34, 1348 Louvain-la-Neuve.

1. Introduction

Global vector autoregressive (*GVAR*) models are one of the most widely implemented econometric frameworks for analyzing interrelations among economies or markets. Interconnections can result from various sources, common risk factors, shared resources, cross-border effects, spillovers, contagion, and asymmetric shock diffusion. *GVAR* offers a simple but coherent way to address these features, allowing for economic interpretation of the interconnection channels. *GVAR* builds on local vector autoregressive (*VAR*) models, augmented by the so-called *star* variable, that is, the weakly exogenous weighted average of foreign variables, resulting in local $VARX^*$ models (Harbo et al., 1998; Pesaran et al., 2000). Once estimated, local systems are solved simultaneously to obtain a large *VAR* representation of the world, which is useful for forecasting, scenario analysis, and describing shock diffusion.

The crucial issue of *GVAR* models is therefore the correct specification of the links among country-specific blocks through the transmission matrix W that provides the weights used to build the foreign weakly exogenous variable in each $VARX^*$. The proposed matrices are usually taken *ad hoc* from the economic literature. For example, in the seminal *GVAR* contribution of Pesaran et al. (2004), the authors consider bilateral relative trade shares among 25 countries grouped into 11 regions. The interconnectedness and shock transmission among countries are therefore based on the trade channel.[2] Such an assumption is also made in *GVAR* applications for forecasting (Pesaran et al., 2009; Cuaresma et al., 2016), real estate and housing markets (Vansteenkiste and Hiebert, 2011), or in regional studies (Ong and Sato, 2018).[3] The choice of the transmission matrix is also a key point for *GVAR* applications in finance, where there are multiple channels. Recent studies have considered financial flows and trade (Lane and Shambaugh, 2010; Chen et al., 2016) or different forms of exposure (Eickmeier and Ng, 2015).

[2] The role of trade in the transmission of shocks or crises is extensively supported by the literature; see, for example, Glick and Rose (2002).

[3] See Di Mauro and Pesaran (2013) or Chudik and Pesaran (2016) for an extensive literature review of *GVAR* applications in different fields.

In the case of euro area government bond yields, Favero (2013) recently proposed a *GVAR* representation to capture the time-varying interdependence among spreads in the euro area. He reveals that the nature of comovements among European bond yields is different from that of macroeconomic variables, as the observed instability in their dynamics is much stronger than that of gross domestic products (*GDP*s), for example. Therefore, countries with higher trade shares would not necessarily display tighter linkages in sovereign bond markets. Consequently, Favero (2013) proposes measuring interconnections in terms of fiscal fundamentals. Intuitively, the more similar two countries are in terms of debt- and deficit-to-GDP ratios, the tighter the interconnection between their sovereign yields. Favero (2013) shows that a *GVAR* featuring fiscal distance based W allows disentangling foreign influence from local and common determinants impacting sovereign bond spreads, and also provides better forecasts than those based only on global and local factors (see for example Beber et al., 2009; Favero et al., 2010).

The choice of different interaction matrices W leads to different inferences, structural analyses, and forecasts. Hence, the key question raised by these studies is how to adequately determine *GVAR* weights and/or test for the validity of the proposed W matrix. To the best of our knowledge, no formal test to resolve this issue exists. The existing literature only follows alternative routes to describe some fundamental features that the correct interaction matrix should have. For example, Gross (2019) develops a framework to estimate W together with local coefficients. Although this approach has the advantage of proposing a statistically relevant interaction matrix, the proposed W loses the intuitive economic interpretation of the linkages that makes *GVAR* so popular. Another strand of the literature focuses on cross-sectional dependence tests (Bailey et al., 2016). However, those tests only guide the choice of a sparse (dense) interaction matrix,[4] without validating the proposed W, and thus they do not distinguish between empirically valid sparse (dense) matrices.

In this paper, we develop a formal test for the validity of the proposed channel of trans-

[4]They also outline the correct estimation procedure to follow; see Elhorst et al. (2020).

mission. We express the $GVAR$ specification as a restricted case of the more general class of structural VAR ($SVAR$). It is thus possible to build a likelihood ratio (LR) test for the restrictions imposed by $GVAR$. In doing so, we offer the possibility of testing the validity of the proposed set of W matrices derived from the economic literature. The asymptotic properties (size and power) of this new test are assessed via Monte Carlo simulations. A bootstrapped version is also proposed to address issues that arise when the number of restrictions becomes important relative to the number of observations.

In the empirical part, we consider the euro area sovereign bond yields example as in the pre-sovereign debt crisis period in Favero (2013), which provides an interesting financial framework to reveal the importance of the choice of W.

As a preview of our results, we show that existing tests do not rule out naïve W matrices, for example, we consider one based on the relative distance in terms of points in the ranking of the *Fédération Internationale de Football Association* (FIFA). Our test instead rejects the FIFA ranking based matrix, but also that based on fiscal fundamentals used by Favero (2013) and the estimated W obtained by implementing the numerical procedure of Gross (2019). An interaction matrix W based on the $t-$statistics of the slope coefficients of the unrestricted version of the reduced-form VAR is valid and thus nonrejected by our test. Compared with the matrices proposed by Favero (2013) and Gross (2019), our W is also split into two parts: one weighting the foreign counterparts with a positive effect and another weighting foreign groups of countries associated with negative effect on local sovereign yields. The test therefore provides useful insights regarding the degree of spillover (see for example Afonso et al., 2012) in the euro area and on the potential "flight-to-quality" or "contagion" effects observed during the subsequent sovereign debt crisis (Metiu, 2012). Interestingly, we find signs of both mechanisms in Euro Area sovereign bond market well before the outburst of the sovereign debt crisis. This result allows reexamining existing literature on financial integration in the Euro Area in the period preceding 2010 (Baele et al., 2004).

The paper proceeds as follows: Section 2 addresses the methodology. It describes the restrictions imposed on the $SVAR$ to obtain the $GVAR$ representation. It also develops the LR test for the validity of the W matrix. Section 3 is devoted to Monte Carlo simulations to evaluate the properties (size and power) of the proposed validity test in finite and large samples. Section 4 concerns the empirical illustration, while Section 5 concludes the paper.

2. Methodology

2.1. GVAR specification

Following Pesaran et al. (2004), let us consider a system of N nodes, representing countries or regions, indexed by $i = 1, \dots, N$, observed over a certain period T, indexed by $t = 1, \dots, T$. Each node features k_i local variables x_{it} and k_i^* node-specific weighted averages of foreign variables $x_{it}^* = \sum_{i \neq j} w_{ij} x_{jt}$, which are considered to be weakly exogenous. The weights w_{ij} of each $i \neq j$ represent the relative intensity measure employed in this literature. Specifically, the larger the weight of foreign counterpart j, the tighter the interconnection between i and j.[5] Moreover, weights can also be variable specific (for example, trade shares for macro variables and financial flows for financial variables).

Local models can then be written as follows:[6]

$$x_{i,t} = \Phi_i x_{i,t-1} + \Lambda_{i0} x_{i,t}^* + \Lambda_{i1} x_{i,t-1}^* + \epsilon_{it}, \tag{1}$$

where $\epsilon_{it} \overset{i.i.d.}{\sim} N(0, \Sigma_i)$ is the idiosyncratic residual term for node i, Φ_i is the matrix of lagged coefficients, and Λ_{i0} and Λ_{i1} are the coefficients associated with the foreign variables. Note that the $GVAR$ parameters are estimated at this local level.

[5] Weights are always positive and sum to unity, i.e., $\sum_{i \neq j} w_{ij} = 1$ for each i.

[6] For ease of explanation, we abstract here from deterministic components, time trends, and additional lags.

2.2. GVAR: a restricted VAR

To obtain the *GVAR* representation of the global economy, we need to simultaneously solve for all the domestic variables. Let us define the $(k_i + k_i^*) \times 1$ vector:

$$z_{i,t} = \begin{pmatrix} x_{i,t} \\ x_{i,t}^* \end{pmatrix} \tag{2}$$

and rewrite (1) as:

$$A_i z_{i,t} = B_i z_{i,t-1} + \epsilon_{it} \tag{3}$$

with $A_i = (I_{k_i},\ -\Lambda_{i0})$ and $B_i = (\Phi_i,\ \Lambda_{i1})$. The dimensions of A_i and B_i are $k_i \times (k_i + k_i^*)$, and A_i has full row rank (Pesaran et al., 2004).

Collecting all the node-specific variables in the $(k \times 1)$ vector $x_t = (x'_{0t}, x'_{1t}, \dots, x'_{Nt})'$ for $i = 1, \dots, N$ with $k = \sum_1^N k_i$, we can rewrite:

$$z_{i,t} = W_i x_t. \tag{4}$$

By substituting (4) into (3), we obtain:

$$A_i W_i x_t = B_i W_i x_{t-1} + \epsilon_{it}, \tag{5}$$

with A_i and B_i being of dimension $k_i \times k$. By stacking all the equations together, we obtain the *SVAR* representation:

$$G x_t = H x_{t-1} + \epsilon_t, \tag{6}$$

with the following conditions: $G = \begin{pmatrix} A_0 W_0 \\ A_1 W_1 \\ \vdots \\ A_N W_N \end{pmatrix}$ and $H = \begin{pmatrix} B_0 W_0 \\ B_1 W_1 \\ \vdots \\ B_N W_N \end{pmatrix}$.

G is a $k \times k$-dimensional matrix, generally full rank and hence nonsingular (Pesaran et al., 2004). Let us define the interaction matrix $\tilde{W}$ as:

$$x_t^* = \tilde{W} x_t. \tag{7}$$

$\tilde{W}$ has dimension $k^* \times k$, with $k^* = \sum_1^N k_i^*$, each row-element-sum is equal to 1,[7] and it measures the *ex ante* defined interaction of each node with the foreign counterparts. This matrix is composed of main diagonal blocks of zero elements, given that no self-interaction is admitted. We can therefore equivalently express the matrices G and H as:

$$G = [I_k - \tilde{\Lambda}_0 \tilde{W}] \tag{8}$$

$$H = [\tilde{\Phi} + \tilde{\Lambda}_1 \tilde{W}] \tag{9}$$

where $\tilde{\Lambda}_0 = blkdiag(\Lambda_{10}, \Lambda_{20}, \ldots, \Lambda_{N0})$, $\tilde{\Phi} = blkdiag(\Phi_1, \Phi_2, \ldots, \Phi_N)$, and $\tilde{\Lambda}_1 = blkdiag(\Lambda_{11}, \Lambda_{21}, \ldots, \Lambda_{N1})$.[8] The dimension of $\tilde{\Phi}$ is $k \times k$. $\tilde{\Lambda}_0$ and $\tilde{\Lambda}_1$ are $k \times k^*$-dimensional matrices.

The *GVAR* solution has the following form:

$$x_t = \Pi + v_t, \tag{10}$$

where the parameters are determined in the following way:

$$x_t = G^{-1} H x_{t-1} + G^{-1} \epsilon_t. \tag{11}$$

Therefore, the *GVAR* solution is the reduced-form VAR representation of a specific *SVAR*. To estimate the local models, weighted averages of foreign counterparts have to be weakly

[7]The weights are thus normalized.
[8]Derivations of an illustrative example can be found in Appendix A.

exogenous. As noted in Pesaran et al. (2004), weak exogeneity in $GVAR$ requires that $\sum_{j \neq i} w_{ji}^2 \approx 0$. This relation is violated when the cross-section is not large enough or when some weights are too large (Favero, 2013). In such a case, a $GVAR$ with only lagged foreign variables has to be considered, therefore imposing:

$$G = I_k, \tag{12}$$

which implies $\Lambda_{i0} = 0$ for all $i = 1, \ldots, N$.[9]

2.3. Likelihood ratio test

Assuming that the probability density distribution function of the process is known, we can estimate the unrestricted reduced-form VAR parameters (characterizing (10) with no restrictions imposed) through maximization of the likelihood function $lnL(\hat{\theta})$ with $\hat{\theta} = \{\Pi, \Sigma_v\}$. In the VAR context, it is common to assume that $v_t \overset{i.i.d.}{\sim} N(0, \Sigma_v)$, implying that y_t are jointly normal. Therefore, the unrestricted likelihood function is of the form:

$$lnL_T(\hat{\theta}) = -\frac{N}{2} ln 2\pi - \frac{1}{2} ln|\Sigma_v| - \frac{1}{2(T-p)} \sum_{t=p+1}^{T} v_t' \Sigma_v^{-1} v_t. \tag{13}$$

The estimation of the $GVAR$ parameters composing the parameter space $\tilde{\theta} = \{\tilde{\Phi}, \tilde{\Lambda}_0, \tilde{\Lambda}_1, \Sigma_\epsilon\}$ is performed at the local $VARX^*$ level in (1). Under the weak exogeneity of the cross-sectional weighted averages x_{it}^*, node-specific models can be estimated consistently through ordinary least squares or reduced-rank procedures (Pesaran et al., 2004). Once the local parameters are estimated, and the model solved as outlined in Section 2.2, we obtain the $GVAR$ solution as in (11), featuring (8) and (9). Given the correspondence outlined in the previous section, we can retrieve the likelihood value for the $GVAR$ representation as:

$$lnL_T(\tilde{\theta}) = -\frac{N}{2} ln 2\pi - \frac{1}{2} ln|\Sigma_\epsilon| + ln|G| - \frac{1}{2(T-p)} \sum_{t=p+1}^{T} \epsilon_t' \Sigma_\epsilon^{-1} \epsilon_t. \tag{14}$$

[9] No contemporaneous foreign interaction through the *star* variable is allowed to avoid endogeneity issues.

The specific parameter space reduction and estimation operated in the local *VARX** models, together with the mapped solution, ensures that the reduced-form representation of the *GVAR* model as in (6) with (8) and (9) is a specific restricted case of the more general reduced-form unrestricted *VAR* as of (10).

It is thus possible to establish an LR test statistic:

$$LR = -2T[logL(\tilde{\theta}) - logL(\hat{\theta})] \tag{15}$$

that is asymptotically distributed as χ^2_m. As usual, m represents the number of restrictions imposed on the unrestricted *VAR* parameters to obtain the *GVAR* representation.

Remark 1: The LR test can be extended to nonstationary *GVAR* models.

As Favero (2013) noted, the *GVAR* model accommodates a long-run solution and the existence of cointegration between x_{it} and x^*_{it}, thus allowing flexibility for covariation in both the cross-sectional and time-series dimensions at the local level. The global vector error correction (*GVEC*) model can be written as follows:

$$\Delta x_{i,t} = \Pi_i z_{i,t-1} + \Lambda_{i0} \Delta x^*_{i,t} + \epsilon_{i,t}, \tag{16}$$

where $z_{i,t-1}$ is defined as (2) and $\Pi_i = (I_{k_i} - \Phi_i - \Lambda_{i0} - \Lambda_{i1})$.

The estimation of country-specific systems in (16) through OLS suffers from the fact that one or more of the factors used may have unit roots, which represents a failure to acknowledge the possibility of long-run relationships between the levels of local and foreign variables. Cointegration can be easily overcome by estimating local error-correction models through reduced rank regression (Pesaran et al., 2004).

Nonstationarity relates also to the global solution as of (11). Therefore, the correspondence between the unrestricted *VAR* (or *VEC*) and the global *GVAR* (or *GVEC*) solution must be respected to ensure the properties of the LR test.[10]

As an example, in the simulation exercise and in the empirical application, we consider the framework used by Favero (2013) to model euro area long-term sovereign bond spreads. The model is of the form:

$$\Delta x_{i,t} = \beta_i x_{i,t-1} + \lambda_{i1} x^*_{i,t-1} + e_{i,t}. \tag{17}$$

Therefore, the corresponding global solution after stacking all $\Delta x_{i,t}$ in a vector labeled ΔX_t would be of the form:

$$\Delta X_t = \tilde{B} X_{t-1} + e_t, \tag{18}$$

where $\tilde{B} = [diag(\beta_1, \ldots, \beta_N) + diag(\lambda_{11}, \ldots, \lambda_{N1})\tilde{W}]$, and $\tilde{W}$ is defined as (7). The corresponding likelihood function will be:

$$lnL_T(\hat{\theta}) = -\frac{N}{2} ln 2\pi - \frac{1}{2} ln|\Sigma_e| - \frac{1}{2(T-1)} \sum_{t=2}^{T} (\Delta X_t - \tilde{B} X_{t-1})' \Sigma_e^{-1} (\Delta X_t - \tilde{B} X_{t-1}) \tag{19}$$

and the corresponding unrestricted vector error correction model is:

$$\Delta X_t = B X_{t-1} + \eta_t, \tag{20}$$

featuring the likelihood function:

$$lnL_T(\tilde{\theta}) = -\frac{N}{2} ln 2\pi - \frac{1}{2} ln|\Sigma_\eta| - \frac{1}{2(T-1)} \sum_{t=2}^{T} (\Delta X_t - B X_{t-1})' \Sigma_\eta^{-1} (\Delta X_t - B X_{t-1}). \tag{21}$$

Remark 2: The test can accommodate a finite sample procedure. The LR test in (15) is an asymptotic test. The unrestricted *VAR* to estimate and the *GVAR* solution to derive feature a number of parameters that would result in over-rejecting the true interaction matrix

[10]See Lütkepohl (2005) for an extensive treatment of maximum likelihood estimation of vector error correction models.

$\tilde{W}$ (if correctly identified) in empirical exercises. As outlined in the simulation section, the over-rejection increases the larger the number of items/nodes and shorter the time series is. It therefore becomes fundamental to consider LR test corrections in finite samples.

Therefore, a bootstrapped version of the test can be effectively employed. The description of the bootstrapped LR test procedure can be found in Appendix C. The intuition is that in short samples, asymptotic tests have undesirable properties (see, for example, Kilian, 1998; Kim, 2014, for extensive discussions on the reliability of short sample tests in time-series analysis). Quantitatively, the distortion can be sizable (as shown in the simulation exercise in the next section). Therefore, a bootstrap alternative allows for the correction of the size distortion and satisfactory power properties. Moreover, the wild bootstrap alternative proposed allows accounting also for unknown forms of heteroskedasticity (Hafner and Herwartz, 2000). When the number of unrestricted VAR parameters is of the same magnitude as or larger than the sample size, the LR test might no longer converge to a χ^2 distribution. In this case, corrections have to be applied (see Bai et al., 2013, 2009).

3. Size and Power Analyses

To analyze the behavior of the validity test in finite and large samples, size and power analysis is proposed via Monte Carlo simulations.

3.1. Size analysis

The level of significance of our test is assessed through Monte Carlo methods. Therefore, we simulate several draws of the data generating processes ($DGPs$) under the null hypothesis. For each draw, $GVAR$ is estimated under the assumption that the correct $\tilde{W}$ matrix of interactions is employed. We then estimate the unrestricted VAR model and compute the LR test.[11] Several $DGPs$ are considered to evaluate the behavior of our validity test under the null hypothesis:[12]

[11]Codes in R are available from the authors upon request.
[12]The specific parameters are available in the supplementary material.

- The first DGP (DGP_1) corresponds to a stationary 4-country, single-item $GVAR$ model.[13] Given the low number of countries, no contemporaneous weighted average interconnection is allowed (as it would violate the weak exogeneity condition). The parameters are randomly drawn such that the roots of the lag polynomial strictly lie within the unit circle. Local errors are normally distributed with zero mean and 0.5 variance. In such a setting, the unrestricted VAR admits 16 slope parameters to estimate, while the $GVAR$ specification admits 8 (i.e., the elements of the diagonal matrices Λ_1 and Φ). Therefore, the degrees of freedom of the validity LR test are 8, given that no restrictions are imposed on the estimated covariance matrix of the errors.

- DGP_2 corresponds to a stationary 10-country, single-item $GVAR$ model. We still do not allow for any contemporaneous term (this assumption will be relaxed in DGP_4). The parameters are randomly drawn such that the roots of the lag polynomial strictly lie within the unit circle. Local errors are normally distributed with zero mean and 0.5 variance. The degrees of freedom of the validity LR test are 80.

- DGP_3 corresponds to the vector error correction model proposed by Favero (2013), i.e., with 10 countries. The weights used are fixed and obtained from the average of the time-varying weights employed in the original paper. Note that no contemporaneous term is present and that two separate $\tilde{W}$s are employed. Local errors are normally distributed with zero mean and 0.5 variance. The degrees of freedom of the validity LR test are 70.

- DGP_4 corresponds to a stationary 10-country, 2-item $GVAR$ model. Instantaneous $GVAR$ interconnection is present, and the parameters are randomly drawn such that the roots of the lag polynomial strictly lie within the unit circle. Local errors are normally distributed with zero mean and unit variance. The degrees of freedom of the validity LR test are 320.

[13] The detailed correspondence for a 4-country, single-item case can be found in the online appendix.

To derive the empirical size of this test, 10,000 replications per DGP are generated, and we report in Table 1 the rejection frequency in percentage points of the null hypothesis. Several sample sizes are considered: $T = 100, 150, 200, 500, 1,000$. We discard 100 burn-in observations included to ensure that the rejection frequencies are free of any initial value dependence.

Table 1: Rejection frequency under the null hypothesis

T	**100**	**150**	**200**	**500**	**1,000**
DGP_1	6.89	6.40	6.02	5.35	5.19
DGP_2	19.38	12.53	10.36	6.75	5.89
DGP_3	45.00	27.92	18.92	9.54	6.91
DGP_4	95.26	67.18	43.40	12.46	8.06

Notes: Rejection Frequency as a % of the null hypothesis considering several DGPs. DGP_1 corresponds to the 4-node, 1-item stationary case, DGP_2 corresponds to the 10-node, 1-item stationary case, DGP_3 corresponds to the nonstationary case proposed by Favero (2013), and DGP_4 corresponds to the 10-country, 2-item stationary case with a contemporaneous term. Simulations are performed with sample size $T = 100, 150, 200, 500, 1,000$. For each sample size and DGP, 10,000 simulations are performed for a sample size of $T + 100$. The first 100 observations are discarded to avoid potential bias due to initial value dependence.

The rejection frequencies for the 4-node, single-item case (DGP_1) are very close to the nominal size even when the sample size is very small ($T = 100$). When increasing the number of nodes and allowing for nonstationarity (DGP_2 and DGP_3), the rejection frequency grows in short samples. It also appears that the convergence speed to 5% is slower and the nominal size is reached only when the sample size is very large ($T > 1,000$). Similar findings are observed when including a contemporaneous term. For the latter cases, a bootstrap version of the LR test is suggested, especially since the test statistic is pivotal.[14]

3.2. Power analysis

For the power analysis, the four previous $DGPs$ are retained along with their parameters. Instead of simulating under the null hypothesis, alternative $\tilde{W}$s are employed. Several misspecifications are considered. For DGP_1, $\tilde{W}$ is simply composed of equal weights (i.e., 0.33).

[14]The bootstrap procedure is presented in Appendix C.

For DGP_2, DGP_3, and DGP_4, we randomly draw the alternative matrices.[15] To highlight the magnitude of the difference between the $\tilde{W}$ considered and that under the null hypothesis (W), we compute the average absolute distance ($dist(\mathbf{W}, \tilde{\mathbf{W}}) = \frac{1}{\#ofweigths} \sum_{i=1}^{n} \sum_{j=1}^{n} |w_{ij} - \tilde{w}_{ij}|$). The same 5 sample sizes considered in the previous section are considered to evaluate the impact of the number of observations on the power. Simulations are performed under the same conditions regarding replications ($10,000$) and burn-in dimension (100 observations) as in the case of size. Rejection frequencies, which correspond to the size-adjusted power, are reported in Table 2.

Table 2: Rejection frequencies under alternative hypotheses

	$dist(\mathbf{W}, \tilde{\mathbf{W}})$		T=100		T=150		T=200		T=500		T=1,000	
	Half	All	Half	All	Half	All	Half	All	Half	All	Half	All
DGP_1	0.048	0.084	34.45	52.13	52.96	73.54	68.68	88.08	99.18	99.98	100.00	100.00
DGP_2	0.041	0.079	8.66	11.78	11.45	18.12	14.18	24.23	39.94	69.99	82.43	98.47
DGP_3	0.040	0.080	6.59	14.27	7.45	21.86	9.50	33.56	20.54	88.10	48.00	99.97
DGP_4	0.031	0.065	6.48	7.68	7.27	9.47	8.23	12.12	17.13	34.04	39.85	75.33

Notes: Rejection Frequency as a % of the null hypothesis when simulating several DGPs with different misspecified transmission matrices ($\tilde{W}$). DGP_1 corresponds to the 4-node, 1-item stationary case, DGP_2 corresponds to the 10-node, 1-item stationary case, DGP_3 corresponds to the nonstationary case proposed by Favero (2013), and DGP_4 corresponds to the 10-country, 2-item stationary case with a contemporaneous term. Half (resp. all) indicates that the misspecification is imposed for half (resp. all) of the weights. The distance to the null hypothesis is calculated as $dist(\mathbf{W}, \tilde{\mathbf{W}}) = \frac{1}{\#ofweigths} \sum_{i=1}^{n} \sum_{j=1}^{n} |w_{ij} - \tilde{w}_{ij}|$. Simulations are performed with sample size $T = 1,000$. In each case, $10,000$ simulations are performed, for a sample size $T + 100$. The first 100 observations are discarded to avoid the potential bias due to initial value dependence.

As expected, the size-adjusted power is positively linked to the sample size. For the one-node system without contemporaneous term in DGP_1, the rejection frequency increases quickly and reaches a large number from $T = 500$. When the number of parameters increases, rejection decreases. In the finite sample, the power is fairly low, indicating that the null hypothesis of validity is too often accepted in small samples. This result is coherent with the literature considering corrections to the short sample problems of asymptotic tests, as outlined in the second remark in Subsection 2.3.

[15] The specific matrices employed are available in the supplementary material. We employ a different scheme for DGP_1 to ensure that the distance from the null hypothesis is comparable in magnitude to that of the other DGPs.

The size-adjusted power also depends on the characteristics of the DGP considered. The rejection frequency reaches its maximum for DGP_1, followed by DGP_2 and DGP_3. DGP_3 contains a contemporaneous term that deteriorates the power, whereas DGP_2 is nearly nonstationary, also impacting the rejection frequency. DGP_4 consistently exhibits the lowest power. This ranking is due to the number of parameters to be estimated, the largest for multinode $DGPs$ such as DGP_4. The difference is, however, not large compared with that exhibited by DGP_3, considering that the misspecification (measured through the relative absolute distance) for all countries is of magnitude 0.065 compared with the 0.080 of DGP_3.

Coherently, the power is also linked to the distance to the null hypothesis. The cases in which only half of the W elements deviate from the null hypothesis (i.e., low distance) always exhibit a lower power than in the case in which all the weights are misspecified (i.e., higher distance).

4. Empirical Analysis

To illustrate the importance of our new validity test in $GVAR$ modeling, we consider the case of the sovereign bond spreads of the euro area countries. Modeling government bonds has become very popular since the onset of the European sovereign debt crisis. Analysts have sought to identify the respective shares of local specific factors, based on fiscal fundamentals and growth, and common factors, corresponding to global appetite for risk. Studies based on local VAR representations (see Sgherri and Zoli, 2009; Favero and Missale, 2012, inter alii) have measured these shares. In a seminal analysis, Favero (2013) demonstrated the superiority of $GVAR$ models based on fiscal fundamentals, as they allow us to quantify a third factor, in line with the uncovered interest rate parity condition: the expectations of exchange rate fluctuations associated with the risk of the dissolution of the euro area. This factor has been crucial since the European sovereign debt crisis outburst. Indeed, in a $GVAR$ framework, sovereign bonds' interdependence is well captured by considering each country's spread as a function of the other European government bond spreads via the interconnection matrix W.

In addition, this framework easily accommodates the presence of other fundamental factors, such as time-varying global risk aversion (see, for example, Codogno et al., 2003; Geyer et al., 2004), without overlooking the importance of local factors.

4.1. The different transmission matrices W

Our aim is to evaluate the validity of the W matrices employed to model sovereign yield spreads, taking into account the importance of nonlocal, European risk factors. The purpose is twofold. On the one hand, we show that, based on existing empirical evaluations, we cannot conclusively rule out naïve W matrices. On the other hand, we seek to assess using our new LR test the existing literature regarding the empirical relevance of the proposed matrices of interactions and the capability of rejecting matrices that are not valid.

Our starting point is a simple *VEC(1)* model, which does not allow for spillovers among country blocks. In this case, sovereign bond spreads are modeled as time-varying long-run equilibrium-reverting processes of the form:

$$\Delta x_{i,t} = \beta_{i0} + \beta_{i1} x_{i,t-1} + \beta_{i2}(Baa_{t-1} - Aaa_{t-1}) + \beta_{i3}\Delta(Baa_t - Aaa_t) + e_{i,t} \tag{22}$$

where $x_{i,t}$ is the sovereign yield spread between country i and Germany (the usual reference in the euro area), $Baa - Aaa$ represents the time varying global risk aversion as measured by the long term US $Baa - Aaa$ corporate bond spreads.[16] However, this model does not take into account the importance of foreign factors influencing local economies, which are particularly relevant in the euro area.

We then augment this specification in a *GVEC(1)* model fashion, as proposed by Favero (2013), considering different W matrices.

The second model includes a naïve W matrix. To this end, we consider the country

[16] Baa and Aaa are two of the ratings assigned by the rating agency Moody's to long term corporate bonds reflecting credit worthiness. Aaa is given to an obligor with extremely strong capacity to meet its financial commitments, Baa to an obligor with adequate capacity. The wider the differential between Baa and Aaa long term rates, the more risk averse investors are (Favero and Missale, 2012).

ranking maintained by the International Federation of Association Football in 2020.[17] The elements of W are based on the inverse of the absolute difference in points in the ranking between countries i and j. Weights are then normalized to sum to 1.[18] The model thus has the following form:

$$\Delta x_{i,t} = \beta_{i0}+\beta_{i1}x_{i,t-1}+\beta_{i2}(Baa_{t-1}-Aaa_{t-1})+\beta_{i3}\Delta(Baa_t-Aaa_t)+\beta_{i4}\sum_{i\neq j} w_{ij}x_{j,t-1}+e_{i,t}, \quad (23)$$

The naïve intuition would be that the more two countries will behave similarly (at least 10 years in the future) in terms of men's national football team performance, the more they are interdependent today. This W matrix is assumed to have almost no effect on the transmission effect among government bond spreads.

The third framework corresponds to that proposed by Favero (2013). He imposes $w_{ij,t}$ as time-varying weights corresponding to the distance between countries i and j at time t, in terms of differences in fiscal fundamentals. In other words, he considers the public debt-to-GDP ratio $debt_{it}$ and deficit-to-GDP ratio def_{it} of each country i. For each fiscal indicator, a distance is built as the absolute difference between the values of countries i and j normalized by the value imposed by the Maastricht criteria. The deficit- and debt-to-GDP ratios are taken as a simple average of the actual values and the forecasts (as known at time t) published by the European Commission to reflect also the importance of the future fiscal outlook ($dist^{def}_{jit} = E_t\left(\left|def^j_t - def^i_t\right|\right)/3$ and $dist^{debt}_{jit} = E_t\left(\left|debt^j_t - debt^i_t\right|\right)/60$). It is then possible to build two distance matrices as follows:

$$w^k_{ji,t} = \frac{w^{*,k}_{ji,t}}{\sum_{j\neq i} w^{*,k}_{ji,t}}, \;\; w^{*,k}_{ji,t} = \frac{1}{\text{dist}^k_{ji,t}}, \;\; with \;\; k = \{def, debt\}.$$

Favero (2013)'s model is thus the following:

[17]We consider the men's ranking. Original data are downloaded from https://www.fifa.com/fifa-world-ranking/ranking-table/men/ on June, 1, 2020.

[18]Matrix W is reported in Appendix B.

$$\Delta x_{i,t} = \beta_{i0} + \beta_{i1} x_{i,t-1} + \beta_{i2} \left(Baa_{t-1} - Aaa_{t-1} \right) + \beta_{i3} E_t \left(debt_t^i - debt_t^{bd} \right) + \\ + \beta_{i4} E_t \left(def_t^i - def_t^{bd} \right) + \beta_{i5} \Delta \left(Baa_t - Aaa_t \right) + \beta_{i6} x_{i,t-1}^{*,b} + \beta_{i7} x_{i,t-1}^{*,d} + e_{i,t}, \quad (24)$$

where $x_t^{*,k} = \sum_{j \neq i} w_{ji,t}^k x_t^j$. The intuition would be that the more similar two countries are in terms of fiscal fundamentals, the more interdependent they are. Note that Favero (2013) also includes debt- and deficit-to-GDP ratios as explanatory variables in the model.

The third W corresponds to the estimated interaction matrix as proposed by Gross (2019). Instead of considering an *ad hoc* matrix, Gross (2019) proposes to estimate it jointly with the local $GVAR$ coefficients. The procedure starts from local models of the form:

$$\Delta x_{i,t} = \beta_{i0} + \beta_{i1} x_{i,t-1} + \beta_{i2} \sum_{i \neq j} w_{ij} x_{j,t-1} + e_{i,t}, \quad (25)$$

and estimates the weights together with the local parameters according to the following constrained optimization problem:

$$\min_{\Gamma_i, w_{ij}} \sum_{t=1}^{T} e_{i,t}^2 \quad (26)$$

subject to

$w_{ij} \geq 0$ for $i \neq j$, $w_{ii} = 0$

and $\sum_{j=0}^{N} w_{ij} = 1$.

Γ_i is the vector collecting all the local parameters of (25), and w_{ij} are the measures of the tightness of the interconnection between countries in a standard $GVAR$ framework. The numerical sequential quadratic programming optimization outlined in Gross (2019) is implemented.[19] Note that each w_{ij} is constrained to be nonnegative. The estimated W

[19]We thank Marco Gross for sharing with us the codes for the quadratic programming constrained optimization.

matrix then forces the transmission mechanism from foreign countries to be uniquely in one direction (either positive or negative, based on the sign of the estimated β_{i2}). Therefore, only the relative intensity of foreign counterparts in the W matrix is estimated through this procedure (the larger the weight, the more interdependent the countries are, given the global effect on the local economy).

The last interaction matrix considered relaxes this hypothesis. Specifically, the interaction matrix is split into two different matrices, one collecting the relative intensity of interconnection with foreign countries positively impacting local economies and the other collecting the countries with negative impact.[20] To the best of our knowledge, this approach has never been considered before in the *GVAR* literature. To this end, the unrestricted *VEC* model (as in (18), augmented with the US long-term corporate bond spread factor) estimation is performed, and the off-diagonal elements of the $\tilde{B}$ matrix are tested via simple $t-tests$. The off-diagonal coefficients' $t-test$s provide important information not only regarding the significance of the specific relationships in the multivariate linear model (proxies we will use for the magnitude of each w_{ij}) but also regarding the sign of the specific cross-country coefficients. Based on this information, we can therefore consider two different W matrices, that associated with a negative coefficient (W^{-}) and that associated with a positive coefficient (W^{+}) in the local $VARX^{*}$ models. The $GVEC$ model featuring those characteristics is thus of the form:

$$\Delta x_{i,t} = \beta_{i0} + \beta_{i1} x_{i,t-1} + \beta_{i2}(Baa_{t-1} - Aaa_{t-1}) + \beta_{i3}\Delta(Baa_t - Aaa_t) + \beta_{i4}\sum_{i\neq j} w_{ij}^{+} x_{j,t-1} + \beta_{i5}\sum_{i\neq j} w_{ij}^{-} x_{j,t-1} + e_{i,t}. \quad (27)$$

Intuitively, the diffusion of a shock via the country-specific blocks is asymmetric and could exhibit either positive or negative spillovers. In the case of shock transmission in the context

[20]The results do not change if we decide to normalize the weights to sum to 1 in total or by W matrix. Given the weak exogeneity issue, however, it is advisable, as we do in this paper, to have an overall normalization. Therefore, per country, we ensure that the sum of weights impacting positively and negatively sums to 1. Note that the weights are always nonnegative, and the β coefficients will determine the signs of the effect.

of euro area sovereign bonds, allowing for such a mechanism is crucial. Beber et al. (2009) and Candelon and Tokpavi (2016) found that shock into a particular country's sovereign yield can decrease core European countries' sovereign bond yields due to outflows of money from peripheral countries in the euro zone and inflows to financially sounder countries. This effect is labelled the "flight-to-quality". Moreover, rising yields in one country might lead to higher yield levels for other peripheral countries in the euro zone, usually labeled "contagion". Therefore, the proposed split of the transmission matrices allows for an asymmetric transmission of shocks. However, the existing literature has only focused on one part of the problem, that is, the identification of the magnitude of the interaction, while overlooking the importance of the sign.

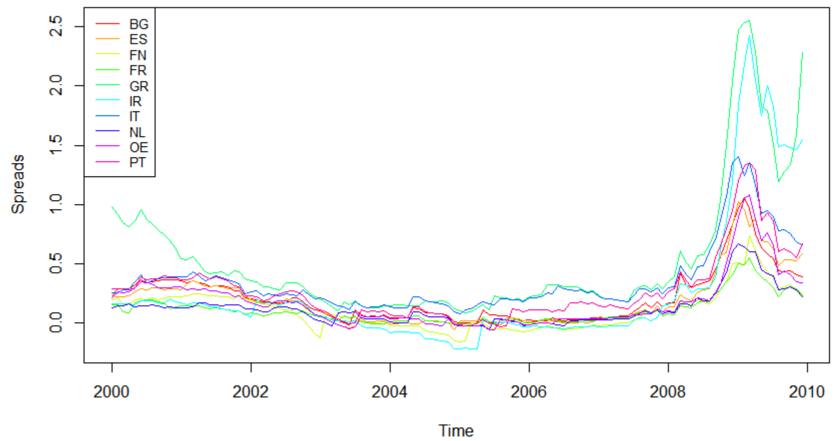
4.2. Data description

Models are estimated considering 10-year sovereign bond interest-rate spreads on German *Bunds* for Austria, Belgium, Finland, France, Greece, Ireland, Italy, the Netherlands, Portugal, and Spain. Data are extracted from Datastream, are of monthly frequency, and cover the period from January 2000 to December 2009 (120 observations). The sample corresponds to the period before the sovereign debt crisis to match the data used in Favero (2013). The euro area sovereign crisis period that began at the end of 2009, and the post-crisis period are excluded. The US corporate long-term $Baa - Aaa$ spreads are extracted from the $FRED$ database of the Federal Reserve Bank of St. Louis. Figure 1 illustrates the evolution of the interest-rate spreads for the countries over time. It is noticeable that that EA sovereign bond interest rates share similar path, supporting hence the idea of a nominal convergence as promoted by the Maastricht treaty. The dispersion tends to increase at the edge of the sovereign debt crisis. The period of investigation is therefore limited to the pre-crisis period.

4.3. Results

The new validity test (15) is implemented for the four different W matrices, considering both the asymptotic and the bootstrapped critical values at the standard level of 95%. Table

Figure 1: 10-year sovereign bond spreads over the German *Bund* in % (2000:1-2009:12)



Note Countries are Austria (OE), Belgium (BG), Finland (FN), France (FR), Greece (GR), Ireland (IR), Italy (IT), the Netherlands (NL), Portugal (PT), and Spain (ES). Source: *DataStream*

Table 3: W validity test

W	Test Stat.	Asymp. CV 95%	Boot. CV 95%
FIFA	334.40	101.88	257.03
Favero (2013)	303.02	90.53	197.97
Gross (2019)	290.52	101.88	246.63
T-Stat based	**83.42**	90.53	175.25

Notes: The likelihood ratio test statistic is calculated for the different W matrices (FIFA, Favero, 2013; Gross, 2019, T-stat based) as in (15). The asymptotic critical value corresponds to the 95% quantile of a χ^2 distribution with adequate degrees of freedom. The bootstrapped critical value corresponds to the 95% quantile of the distribution obtained using the sequence presented in Appendix C. When the null hypothesis is not rejected at a confidence level of 95%, the test statistic appears in bold.

3 reports the results.

To complete the results of the tests, we also report in Table 4 the adjusted R^2 of the country systems for the different W matrices, that provides a measure of the quality of the model fit.

The test rejects the null hypothesis of the validity of the W based on the $FIFA$ ranking. This result is satisfying, as it would have been strange to find evidence that sovereign yield spreads' interdependence could be based on the future performance of national football teams. Nevertheless, Table 4 interestingly reveals that the inclusion of a transmission matrix, even

Table 4: Adjusted R-squared for the country systems

Country	Adj. R^2				
	Basic	FIFA	Favero (2013)	Gross (2019)	T-stat based
Belgium	0.06	0.08	0.08	0.16	**0.36**
Spain	0.15	0.12	0.12	0.25	**0.39**
Finland	0.15	0.18	0.19	0.26	**0.28**
France	0.13	0.17	0.14	0.24	**0.31**
Greece	0.16	0.19	0.31	0.18	**0.51**
Ireland	0.28	0.29	0.36	0.37	**0.58**
Italy	0.23	0.25	0.20	0.26	**0.32**
Netherlands	0.25	0.29	0.27	0.26	**0.41**
Austria	0.19	0.21	0.37	0.43	**0.57**
Portugal	0.16	0.23	0.30	0.38	**0.44**

Note: Adjusted R^2 values are calculated for the different W matrices (Basic, FIFA ranking, Favero, 2013; Gross, 2019, T-stat based). The largest adjusted R^2 per country are in bold. The basic model is the one with no spillovers among country blocks as in (22).

if based on $FIFA$ rankings, improves the adjusted R^2 for all countries (except Spain[21]) compared to the specification without spillovers. This first outcome is interesting, as most of the validation analyses conducted in the $GVAR$ literature are concerned with tests regarding the relevance of including (or excluding) the *star* variable. If even a naïve matrix improves the adjusted R^2, the comparison between a "basic" closed economy framework (only including local variables) and any open economy framework (such as $GVAR$, given the inclusion of a weighted average of foreign counterparts) is biased. A test only based on the inclusion (or exclusion) of the *star* variable is therefore not informative regarding the specific W matrix employed.[22] Instead, our LR test examines the empirical validity of the specific channel of interdependence proposed. This distinction is fundamental. Specifically, even if the in-sample fit performance improves when employing the $GVAR$ framework based on the $FIFA$

[21]When including the debt- and deficit-to-GDP ratios as regressors as in the traditional model in Favero (2013), the only country not improved by the $GVAR$ based on FIFA ranking is Greece instead of Spain. These results are available from the authors upon request.

[22]We could say that this class of tests is only informative regarding the fact that including foreign variables in the analysis is crucial and, therefore, that the closed economy assumption is rejected. However, it does not provide any judgement regarding the validity of the specific open economy framework proposed.

ranking transmission channel, our test rejects the $FIFA$-ranking-based W matrix, revealing its empirical invalidity.

Considering the framework proposed in Favero (2013), the validity of the transmission matrix is also rejected at both the asymptotic and bootstrapped levels. It therefore signals that this $GVAR$ design is not empirically valid. Economically, these findings suggest that fiscal indicators (debt- and deficit-to-GDP ratios) are not adequate factors to explain the interdependence among sovereign yield spreads across euro area countries. Even more interesting, the results in Table 4 indicate that the adjusted R^2 using this W matrix is not always larger than that obtained with the $FIFA$ ranking W matrix. In the cases of France, Italy, and the Netherlands, it is actually worse, while no improvement is shown in the case of Spain and Belgium. Considering fiscal indicators instead of football rankings hence does not improve the sample fit of the model in any conclusive sense.[23]

The third experiment considers the estimated W matrix following the methodology developed in Gross (2019). W should therefore be more accurate from an empirical perspective, at the cost of its economic interpretation. The test also rejects the validity of this estimated transmission matrix. The reason for this rejection lies in the fact that Gross (2019) imposes all w_{ij} to be nonnegative and collected in a single W matrix. As explained in the previous section, this simplification forces the constrained optimization to choose the more relevant set of countries on the basis of a single β coefficient responsible for the local effect of all the foreign counterparts. Intuitively, the coexistence of "fight-to-quality" and "contagion" effects among countries documented in the literature is not possible. This constraint constitutes a limitation of the Gross (2019) approach. In the case of sovereign bond spreads within the euro area, this constraint causes the rejection of the estimated W. Nevertheless, according to the adjusted R^2, the estimated W model outperforms all the previous models (except for

[23]The slow moving debt- and deficit-to-GDP ratios employed by Favero (2013) do not change the characteristics of the LR test. As a matter of fact, if we consider a fixed weights $GVAR$ framework where the weights are the average of the employed time varying debt- and deficit-to-GDP ratios, the LR test statistic value is 303.27 (compared with the 303.02 of the time varying specification). The validity of this framework also gets rejected both considering the asymptotic and the bootstrapped critical values.

Greece and the Netherlands), thus signaling its superiority in terms of in-sample fit.

The final experiment is performed using the partitioned version of W outlined in section 4.1.[24] As expected, the test does not reject at the 95% confidence level the validity of the unrestricted $t - statistics$ based and partitioned W matrix as of (27). Table 4 confirms this finding, as it shows that the adjusted R^2 values for country-specific systems are consistently the highest.

The outcome of the test corroborates the idea that the heterogeneity within the euro zone sovereign bond market was relevant even before the occurrence of the sovereign debt crisis in late 2009. This is very insightful, as most of the existing literature describes the period under analysis as financially integrated (Baele et al., 2004). According to this literature, investors perceived sovereign bonds in the euro area as perfect substitutes, given the observed comovements at low levels of government financing rates (see Figure 1). Our results instead show that symptoms of fragmentation were already relevant in the euro area sovereign bond market well before the outbreak of the sovereign debt crisis.

From the W matrices reported in Appendix B, we can identify the features that will constitute the core of the fragmentation argument following the euro area debt crisis. In our *GVAR* framework, "contagion" would mean the centrality of large weights in the W^+ matrix of financially fragile economies, characterized by high levels of debt- and deficit-to-GDP ratios (namely, Greece, Italy, Ireland, Portugal, and Spain). The "flight-to-quality" mechanism would instead arise if we detected a centrality in the W^- matrix of weights associated with countries belonging to fragile (sound) economies for sound (fragile) economies.

According to the results, Portugal and Ireland are the countries where contagion is the most relevant feature of foreign influence (52% and 49%, respectively). Greece was already showing itself to be the most fragile economy in the euro zone. This country, which subsequently showed the highest sovereign rates, experienced almost equal influence from the "contagion" and "flight-to-quality" effects (accounting for 65% of all foreign influence), with

[24]The W matrices are reported in Appendix B.

Spain (22%) and Italy (12%) being the highest weighted for the "contagion" effect and Austria being the highest weighted (29%) for the "flight-to-quality" effect. Spain and Italy show milder influences of the two effects at this early stage, although 35% of Italian foreign influence is represented by a negative relationship with financially sounder economies. Spain instead shows a close relationship to that of Italy (31% of foreign influence) for the "contagion" effect.

In the case of financially sounder economies, 36% of the foreign influence for the Netherlands comes from the "flight-to-quality" effect, while Austria has the highest weight in the "flight-to-quality" effect coming from Ireland. Finland has the least influence from the two effects, being mostly related to core countries (32% from Belgium). Interestingly, among the financially sounder economies, Belgium and France are related to foreign economies similarly to financially fragile economies, even if they have been less affected by the debt crisis afterwards. Whereas in the case of France, it appears that after the outbreak of the sovereign debt crisis, the specific policies implemented allowed the country to resist the suggested signs of fragility, in the case of Belgium, we note a "flight-to-quality" effect that might have improved its financing rates coming from Ireland (27% of foreign influence).

It is thus important to test for the validity of the transmission matrix also from an economic perspective. Indeed, the nonrejected W matrix provides important insights regarding sovereign bond spreads spillover effects inside the euro area even before the outburst of the sovereign debt crisis.

5. Conclusion

GVAR models constitute one of the most important econometric tools to analyze global interdependence in a multi-country/market environment. *GVAR*s rely on a specific interaction matrix W, which determines the tightness of the interconnection among units. The direct interpretability of this channel for the transmission of shocks represents one of the reasons for its popularity. However, the W matrices employed in this literature are proposed *ad hoc* and justified by the economic literature but never empirically tested.

In this paper, we prove that by expressing *GVAR* as a restricted version of the more general class of *SVAR* models, it is possible to clearly identify the restrictions imposed to characterize the specific transmission channel. Therefore, a simple LR test is designed to empirically validate the proposed W matrix. The asymptotic properties of the test are assessed via Monte Carlo methods. We also propose a bootstrapped version of the test to correct for the bias that asymptotic tests display when the number of observations becomes comparable to the number of parameters to estimate.

To demonstrate the importance of empirically validating the assumed matrix of interactions, we apply our new test to the *GVAR* modeling of sovereign bond yields for the euro area countries before the outbreak of the sovereign bond debt crisis in late 2009.

First, we prove that without testing for the matrix of interaction, we cannot rule out even naïve transmission channels. In fact, we show that when employing existing tests based on in-sample fit performance, we fail to reject a matrix based on the FIFA rankings of national football teams. Moreover, when compared to a more economic-based interaction matrix such as that based on debt- and deficit-to-GDP ratios (as proposed by Favero, 2013), we cannot conclusively decide which performs better. We conclude that existing tests are concerned with the rejection of closed economy models (featuring no interdependence among countries) without checking whether the specific transmission channel is empirically valid.

Second, by employing our LR test, we reject the W matrix based on FIFA rankings, thereby excluding any involvement of future football performance as proxy for cross-country interdependence. Interestingly, our test also rejects the debt- and deficit-to-GDP-based interaction model proposed by Favero (2013) and the estimated W matrix proposed by Gross (2019).

We find that the LR test does not reject a matrix featuring a clear distinction between positive and negative interdependence. As a result, when modeling interconnections among euro area sovereign bond spreads, we must recognize that a country can be the source of increasing yields for some countries and of decreasing yields for others. Even if the sample considered is prior to the sovereign debt crisis, signs of likely “contagion” and “flight-to-quality” effects were thus already evident.

From the nonrejected matrices, we can clearly see how the rates of Portugal and Ireland were closely interconnected to foreign countries with high levels of debt- and deficit-to-GDP ratios. In the case of Greece, the “flight-to-quality” effect (defined as interconnection to financially sounder countries with a negative sign) was also evident. Italy and Spain showed milder “peripheral” characteristics. For core countries such as the Netherlands and Austria, the “flight-to-quality” effect from which they would subsequently benefit (in terms of lowering yields while premia were increasing in peripheral countries) was already present in the pre-crisis data. While Finland seems not to be subject to these effects, Belgium and France showed some “peripheral” behavior.

The new test presented in this paper paves the way for a wide range of economic implementations of *GVAR*. Specifically, testing the interaction matrix would preserve the intuitive interpretation of the proposed channel for the transmission of shocks among local markets while ensuring that the restrictions imposed are supported by empirical data. Inferences based on such a matrix would therefore be both theoretically sound and empirically valid.

Acknowledgments

The authors thank Carlo Favero, Christian Hafner, Eugen Piracabelu, Jean-Charles Wijnandts, and the participants in the 12^{th} International Workshop of Methods in International Finance Network (MIFN), the 19^{th} EconomiX-CNRS and Université de Paris Nanterre Econometrics day, and the Louvain Finance seminars for useful comments and remarks. This research was conducted as part of the research program entitled “Financial and Extrafinancial Risk Modeling” under the aegis of the Europlace Institute of Finance, a joint initiative with Insti7. The usual disclaimer applies.

References

Afonso, A., D. Furceri, and P. Gomes (2012). Sovereign credit ratings and financial markets linkages: application to European data. *Journal of International Money and Finance 31*(3), 606–638.

Baele, L., A. Ferrando, P. Hördahl, E. Krylova, and C. Monnet (2004). Measuring European financial integration. *Oxford Review of Economic Policy 20*(4), 509–530.

Bai, Z., D. Jiang, J.-F. Yao, and S. Zheng (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *The Annals of Statistics 37*(6B), 3822–3840.

Bai, Z., D. Jiang, J.-f. Yao, and S. Zheng (2013). Testing linear hypotheses in high-dimensional regressions. *Statistics 47*(6), 1207–1223.

Bailey, N., S. Holly, and M. H. Pesaran (2016). A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence. *Journal of Applied Econometrics 31*(1), 249–280.

Beber, A., M. W. Brandt, and K. A. Kavajecz (2009). Flight-to-quality or flight-to-liquidity? Evidence from the euro-area bond market. *The Review of Financial Studies 22*(3), 925–957.

Candelon, B. and S. Tokpavi (2016). A non-parametric test for Granger-causality in distribution with application to financial contagion. *Journal of Business & Economic Statistics 34*(2), 240–253.

Chen, Q., A. Filardo, D. He, and F. Zhu (2016). Financial crisis, US unconventional monetary policy and international spillovers. *Journal of International Money and Finance 67*, 62–81.

Chudik, A. and M. H. Pesaran (2016). Theory and practice of GVAR modelling. *Journal of Economic Surveys 30*(1), 165–197.

Codogno, L., C. Favero, and A. Missale (2003). Yield spreads on EMU government bonds. *Economic Policy 18*(37), 503–532.

Cuaresma, J. C., M. Feldkircher, and F. Huber (2016). Forecasting with global vector autoregressive models: A Bayesian approach. *Journal of Applied Econometrics 31*(7), 1371–1391.

Di Mauro, F. and M. H. Pesaran (2013). *The GVAR handbook: Structure and applications of a macro model of the global economy for policy analysis.* OUP Oxford.

Eickmeier, S. and T. Ng (2015). How do US credit supply shocks propagate internationally? A GVAR approach. *European Economic Review 74*, 128–145.

Elhorst, J. P., M. Gross, and E. Tereanu (2020). Cross-sectional dependence and spillovers in space and time: where spatial econometrics and global VAR models meet. *Journal of Economic Surveys.*

Favero, C. and A. Missale (2012). Sovereign spreads in the eurozone: which prospects for a Eurobond? *Economic Policy 27*(70), 231–273.

Favero, C., M. Pagano, and E.-L. Von Thadden (2010). How does liquidity affect government bond yields? *Journal of Financial and Quantitative Analysis 45*(1), 107–134.

Favero, C. A. (2013). Modelling and forecasting government bond spreads in the euro area: a GVAR model. *Journal of Econometrics 177*(2), 343–356.

Geyer, A., S. Kossmeier, and S. Pichler (2004). Measuring systematic risk in EMU government yield spreads. *Review of Finance 8*(2), 171–197.

Glick, R. and A. K. Rose (2002). Does a currency union affect trade? The time-series evidence. *European Economic Review 46*(6), 1125–1151.

Gross, M. (2019). Estimating GVAR weight matrices. *Spatial Economic Analysis 14*(2), 219–240.

Hafner, C. M. and H. Herwartz (2000). Testing for linear autoregressive dynamics under heteroskedasticity. *The Econometrics Journal 3*(2), 177–197.

Harbo, I., S. Johansen, B. Nielsen, and A. Rahbek (1998). Asymptotic inference on cointegrating rank in partial systems. *Journal of Business & Economic Statistics 16*(4), 388–399.

Kilian, L. (1998). Small-sample confidence intervals for impulse response functions. *Review of Economics and Statistics 80*(2), 218–230.

Kim, J. H. (2014). Testing for parameter restrictions in a stationary VAR model: a bootstrap alternative. *Economic Modelling 41*, 267–273.

Lane, P. R. and J. C. Shambaugh (2010). Financial exchange rates and international currency exposures. *American Economic Review 100*(1), 518–40.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis.* Springer Science & Business Media.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *The Annals of Statistics*, 255–285.

Metiu, N. (2012). Sovereign risk contagion in the eurozone. *Economics Letters 117*(1), 35–38.

Ong, S. L. and K. Sato (2018). Regional or global shock? A global VAR analysis of Asian economic and financial integration. *The North American Journal of Economics and Finance 46*, 232–248.

Pesaran, M. H., T. Schuermann, and L. V. Smith (2009). Forecasting economic and financial variables with global VARs. *International Journal of Forecasting 25*(4), 642–675.

Pesaran, M. H., T. Schuermann, and S. M. Weiner (2004). Modeling regional interdependencies using a global error-correcting macroeconometric model. *Journal of Business & Economic Statistics 22*(2), 129–162.

Pesaran, M. H., Y. Shin, and R. J. Smith (2000). Structural analysis of vector error correction models with exogenous I(1) variables. *Journal of Econometrics 97*(2), 293–343.

Sgherri, S. and E. Zoli (2009). Euro area risk during the crisis. *International Monetary Fund Working Paper 09/222*.

Vansteenkiste, I. and P. Hiebert (2011). Do house price developments spillover across euro area countries? Evidence from a global var. *Journal of Housing Economics 20*(4), 299–314.

Appendix A. *GVAR* Correspondence with $\tilde{\boldsymbol{W}}$, $\tilde{\boldsymbol{\Lambda}}_0$, $\tilde{\boldsymbol{\Lambda}}_1$, and $\tilde{\boldsymbol{\Phi}}$.

Following the example in Pesaran et al. (2004), we consider a simple global model with three nodes featuring three variables each, output, prices, and exchange rates:

$$x_t = \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{pmatrix} = \begin{pmatrix} y_{1t} \\ p_{1t} \\ y_{2t} \\ p_{2t} \\ e_{2t} \\ y_{3t} \\ p_{3t} \\ e_{3t} \end{pmatrix}, \quad z_{1t} = \begin{pmatrix} y_{1t} \\ p_{1t} \\ y_{1t}^* \\ p_{1t}^* \\ e_{1t}^* \end{pmatrix}, \quad z_{it} = \begin{pmatrix} y_{it} \\ p_{it} \\ e_{it} \\ y_{it}^* \\ p_{it}^* \end{pmatrix} \quad i = 2, 3.$$

The W_i matrices are:

$$W_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & w_{12} & 0 & 0 & w_{13} & 0 & 0 \\ 0 & 0 & 0 & w_{12} & 0 & 0 & w_{13} & 0 \\ 0 & 0 & 0 & 0 & w_{12} & 0 & 0 & w_{13} \end{pmatrix},$$

$$W_2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ w_{21} & 0 & 0 & 0 & 0 & w_{23} & 0 & 0 \\ 0 & w_{21} & 0 & 0 & 0 & 0 & w_{23} & 0 \end{pmatrix}$$

and

$$W_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ w_{31} & 0 & w_{32} & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{31} & 0 & w_{32} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The matrices $A_i = (I_{k_i}, -\Lambda_{i0})$ and $B_i = (\Phi_i, \Lambda_{i1})$ are:

$$A_1 = \begin{pmatrix} 1 & 0 & -\lambda^{10}_{11} & -\lambda^{10}_{12} & -\lambda^{10}_{13} \\ 0 & 1 & -\lambda^{10}_{21} & -\lambda^{10}_{22} & -\lambda^{10}_{23} \end{pmatrix}, \quad A_i = \begin{pmatrix} 1 & 0 & 0 & -\lambda^{i0}_{11} & -\lambda^{i0}_{12} \\ 0 & 1 & 0 & -\lambda^{i0}_{21} & -\lambda^{i0}_{22} \\ 0 & 0 & 1 & -\lambda^{i0}_{31} & -\lambda^{i0}_{32} \end{pmatrix},$$

$$B_1 = \begin{pmatrix} \phi^1_{11} & \phi^1_{12} & \lambda^{11}_{11} & \lambda^{11}_{12} & \lambda^{11}_{13} \\ \phi^1_{21} & \phi^1_{22} & \lambda^{11}_{21} & \lambda^{11}_{22} & \lambda^{11}_{23} \end{pmatrix}, \quad B_i = \begin{pmatrix} \phi^i_{11} & \phi^i_{12} & \phi^i_{13} & \lambda^{i1}_{11} & \lambda^{i1}_{12} \\ \phi^i_{21} & \phi^i_{22} & \phi^i_{23} & \lambda^{i1}_{21} & \lambda^{i1}_{22} \\ \phi^i_{31} & \phi^i_{32} & \phi^i_{33} & \lambda^{i1}_{31} & \lambda^{i1}_{32} \end{pmatrix}.$$

The resulting matrices G and H, obtained by stacking matrices $G_i = A_i W_i$ and $H_i = B_i W_i$, are of the form:

$$G = \begin{pmatrix} 1 & 0 & -\lambda^{10}_{11}w_{12} & -\lambda^{10}_{12}w_{12} & -\lambda^{10}_{13}w_{12} & -\lambda^{10}_{11}w_{13} & -\lambda^{10}_{12}w_{13} & -\lambda^{10}_{13}w_{13} \\ 0 & 1 & -\lambda^{10}_{21}w_{12} & -\lambda^{10}_{22}w_{12} & -\lambda^{10}_{23}w_{12} & -\lambda^{10}_{21}w_{13} & -\lambda^{10}_{22}w_{13} & -\lambda^{10}_{23}w_{13} \\ -\lambda^{20}_{11}w_{21} & -\lambda^{20}_{12}w_{21} & 1 & 0 & 0 & -\lambda^{20}_{11}w_{23} & -\lambda^{20}_{12}w_{23} & 0 \\ -\lambda^{20}_{21}w_{21} & -\lambda^{20}_{22}w_{21} & 0 & 1 & 0 & -\lambda^{20}_{21}w_{23} & -\lambda^{20}_{22}w_{23} & 0 \\ -\lambda^{20}_{31}w_{21} & -\lambda^{20}_{32}w_{21} & 0 & 0 & 1 & -\lambda^{20}_{31}w_{23} & -\lambda^{20}_{32}w_{23} & 0 \\ -\lambda^{30}_{11}w_{31} & -\lambda^{30}_{12}w_{31} & -\lambda^{30}_{11}w_{32} & -\lambda^{30}_{12}w_{32} & 0 & 1 & 0 & 0 \\ -\lambda^{30}_{21}w_{31} & -\lambda^{30}_{22}w_{31} & -\lambda^{30}_{21}w_{32} & -\lambda^{30}_{22}w_{32} & 0 & 0 & 1 & 0 \\ -\lambda^{30}_{31}w_{31} & -\lambda^{30}_{32}w_{31} & -\lambda^{30}_{31}w_{32} & -\lambda^{30}_{32}w_{32} & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$H = \begin{pmatrix} \phi^1_{11} & \phi^1_{12} & \lambda^{11}_{11}w_{12} & \lambda^{11}_{12}w_{12} & \lambda^{11}_{13}w_{12} & \lambda^{11}_{11}w_{13} & \lambda^{11}_{12}w_{13} & \lambda^{11}_{13}w_{13} \\ \phi^1_{21} & \phi^1_{22} & \lambda^{11}_{21}w_{12} & \lambda^{11}_{22}w_{12} & \lambda^{11}_{23}w_{12} & \lambda^{11}_{21}w_{13} & \lambda^{11}_{22}w_{13} & \lambda^{11}_{23}w_{13} \\ \lambda^{21}_{11}w_{21} & \lambda^{21}_{12}w_{21} & \phi^2_{11} & \phi^2_{12} & \phi^2_{13} & \lambda^{21}_{11}w_{23} & \lambda^{21}_{12}w_{23} & 0 \\ \lambda^{21}_{21}w_{21} & \lambda^{21}_{22}w_{21} & \phi^2_{21} & \phi^2_{22} & \phi^2_{23} & \lambda^{21}_{21}w_{23} & \lambda^{21}_{22}w_{23} & 0 \\ \lambda^{21}_{31}w_{21} & \lambda^{21}_{32}w_{21} & \phi^2_{31} & \phi^2_{32} & \phi^2_{33} & \lambda^{21}_{31}w_{23} & \lambda^{21}_{32}w_{23} & 0 \\ \lambda^{31}_{11}w_{31} & \lambda^{31}_{12}w_{31} & \lambda^{31}_{11}w_{32} & \lambda^{31}_{12}w_{32} & 0 & \phi^3_{11} & \phi^3_{12} & \phi^3_{13} \\ \lambda^{31}_{21}w_{31} & \lambda^{31}_{22}w_{31} & \lambda^{31}_{21}w_{32} & \lambda^{31}_{22}w_{32} & 0 & \phi^3_{21} & \phi^3_{22} & \phi^3_{23} \\ \lambda^{31}_{31}w_{31} & \lambda^{31}_{32}w_{31} & \lambda^{31}_{31}w_{32} & \lambda^{31}_{32}w_{32} & 0 & \phi^3_{31} & \phi^3_{32} & \phi^3_{33} \end{pmatrix}.$$

The off-diagonal elements of the matrices result also from the following operations:

$$\tilde{\Lambda}_0\tilde{W} = \begin{pmatrix} \lambda^{10}_{11} & \lambda^{10}_{12} & \lambda^{10}_{13} & 0 & 0 & 0 & 0 \\ \lambda^{10}_{21} & \lambda^{10}_{22} & \lambda^{10}_{23} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda^{10}_{11} & \lambda^{10}_{12} & 0 & 0 \\ 0 & 0 & 0 & \lambda^{10}_{21} & \lambda^{10}_{22} & 0 & 0 \\ 0 & 0 & 0 & \lambda^{10}_{31} & \lambda^{10}_{32} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda^{20}_{11} & \lambda^{20}_{12} \\ 0 & 0 & 0 & 0 & 0 & \lambda^{20}_{21} & \lambda^{20}_{22} \\ 0 & 0 & 0 & 0 & 0 & \lambda^{20}_{31} & \lambda^{20}_{32} \end{pmatrix} \begin{pmatrix} 0 & 0 & w_{12} & 0 & 0 & w_{13} & 0 & 0 \\ 0 & 0 & 0 & w_{12} & 0 & 0 & w_{13} & 0 \\ 0 & 0 & 0 & 0 & w_{12} & 0 & 0 & w_{13} \\ w_{21} & 0 & 0 & 0 & 0 & w_{23} & 0 & 0 \\ 0 & w_{21} & 0 & 0 & 0 & 0 & w_{23} & 0 \\ w_{31} & 0 & w_{32} & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{31} & 0 & w_{32} & 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$\tilde{\Lambda}_1\tilde{W} = \begin{pmatrix} \lambda_{11}^{11} & \lambda_{12}^{11} & \lambda_{13}^{11} & 0 & 0 & 0 & 0 \\ \lambda_{21}^{11} & \lambda_{22}^{11} & \lambda_{23}^{11} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{11}^{11} & \lambda_{12}^{11} & 0 & 0 \\ 0 & 0 & 0 & \lambda_{21}^{11} & \lambda_{22}^{11} & 0 & 0 \\ 0 & 0 & 0 & \lambda_{31}^{11} & \lambda_{32}^{11} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_{11}^{21} & \lambda_{12}^{21} \\ 0 & 0 & 0 & 0 & 0 & \lambda_{21}^{21} & \lambda_{22}^{21} \\ 0 & 0 & 0 & 0 & 0 & \lambda_{31}^{21} & \lambda_{32}^{21} \end{pmatrix} \begin{pmatrix} 0 & 0 & w_{12} & 0 & 0 & w_{13} & 0 & 0 \\ 0 & 0 & 0 & w_{12} & 0 & 0 & w_{13} & 0 \\ 0 & 0 & 0 & 0 & w_{12} & 0 & 0 & w_{13} \\ w_{21} & 0 & 0 & 0 & 0 & w_{23} & 0 & 0 \\ 0 & w_{21} & 0 & 0 & 0 & 0 & w_{23} & 0 \\ w_{31} & 0 & w_{32} & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{31} & 0 & w_{32} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We can therefore express matrices G and H as:

$$G = [I_8 - \tilde{\Lambda}_0\tilde{W}],$$

$$H = [\tilde{\Phi} + \tilde{\Lambda}_1\tilde{W}],$$

with I_8 being the 8×8 identity matrix and

$$\tilde{\Phi} = \begin{pmatrix} \phi_{11}^1 & \phi_{12}^1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \phi_{21}^1 & \phi_{22}^1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \phi_{11}^2 & \phi_{12}^2 & \phi_{13}^2 & 0 & 0 & 0 \\ 0 & 0 & \phi_{21}^2 & \phi_{22}^2 & \phi_{23}^2 & 0 & 0 & 0 \\ 0 & 0 & \phi_{31}^2 & \phi_{32}^2 & \phi_{33}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \phi_{11}^3 & \phi_{12}^3 & \phi_{13}^3 \\ 0 & 0 & 0 & 0 & 0 & \phi_{21}^3 & \phi_{22}^3 & \phi_{23}^3 \\ 0 & 0 & 0 & 0 & 0 & \phi_{31}^3 & \phi_{32}^3 & \phi_{33}^3 \end{pmatrix}.$$

Appendix B. The *W* Matrices

The different W values for our empirical analysis are reported below.

Table B.5: Estimated W matrix using Gross (2019) procedure

	BG	ES	FN	FR	GR	IR	IT	NL	OE	PT
BG	0.000	0.000	0.000	0.000	0.000	0.062	0.000	0.000	0.938	0.000
ES	0.392	0.000	0.189	0.000	0.000	0.000	0.419	0.000	0.000	0.000
FN	0.967	0.033	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FR	0.503	0.037	0.000	0.000	0.000	0.000	0.379	0.082	0.000	0.000
GR	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
IR	0.000	0.000	0.000	0.000	0.043	0.000	0.790	0.167	0.000	0.000
IT	0.000	0.000	0.000	0.000	0.000	0.431	0.000	0.000	0.239	0.330
NL	0.000	0.000	0.000	0.000	0.000	0.484	0.000	0.000	0.516	0.000
OE	0.613	0.160	0.000	0.000	0.086	0.000	0.000	0.140	0.000	0.000
PT	0.453	0.027	0.000	0.089	0.002	0.000	0.429	0.000	0.000	0.000

Notes: This table reports the estimated W matrix as in Gross (2019). BG stands for Belgium, ES for Spain, FN for Finland, FR for France, GR for Greece, IR for Ireland, IT for Italy, NL for the Netherlands, OE for Austria and PT for Portugal.

Table B.6: W matrix using T-value related weights, negative

	BG	ES	FN	FR	GR	IR	IT	NL	OE	PT
BG	0.000	0.000	0.000	0.000	0.000	0.270	0.000	0.000	0.130	0.066
ES	0.000	0.000	0.000	0.000	0.000	0.163	0.000	0.032	0.020	0.186
FN	0.000	0.000	0.000	0.091	0.000	0.145	0.000	0.070	0.101	0.000
FR	0.000	0.000	0.035	0.000	0.000	0.137	0.000	0.000	0.087	0.113
GR	0.000	0.000	0.031	0.000	0.000	0.043	0.000	0.000	0.285	0.150
IR	0.000	0.000	0.073	0.041	0.000	0.000	0.000	0.000	0.083	0.180
IT	0.000	0.000	0.100	0.000	0.000	0.159	0.000	0.037	0.217	0.086
NL	0.000	0.066	0.000	0.000	0.016	0.161	0.000	0.000	0.053	0.113
OE	0.000	0.000	0.029	0.073	0.000	0.200	0.000	0.000	0.000	0.112
PT	0.000	0.000	0.010	0.000	0.000	0.231	0.000	0.010	0.000	0.000

Notes: This table reports the estimated W matrix for the negative β coefficients tested by $t-stat$. BG stands for Belgium, ES for Spain, FN for Finland, FR for France, GR for Greece, IR for Ireland, IT for Italy, NL for the Netherlands, OE for Austria and PT for Portugal.

Table B.7: W matrix using T-value related weights, positive

	BG	ES	FN	FR	GR	IR	IT	NL	OE	PT
BG	0.000	0.192	0.014	0.022	0.103	0.000	0.127	0.076	0.000	0.000
ES	0.054	0.000	0.142	0.080	0.012	0.000	0.311	0.000	0.000	0.000
FN	0.318	0.181	0.000	0.000	0.044	0.000	0.020	0.000	0.000	0.030
FR	0.121	0.144	0.000	0.000	0.074	0.000	0.191	0.097	0.000	0.000
GR	0.085	0.217	0.000	0.056	0.000	0.000	0.120	0.012	0.000	0.000
IR	0.011	0.190	0.000	0.000	0.109	0.000	0.186	0.127	0.000	0.000
IT	0.237	0.083	0.000	0.067	0.014	0.000	0.000	0.000	0.000	0.000
NL	0.103	0.000	0.138	0.038	0.000	0.000	0.311	0.000	0.000	0.000
OE	0.098	0.112	0.000	0.000	0.157	0.000	0.104	0.116	0.000	0.000
PT	0.131	0.054	0.000	0.086	0.079	0.000	0.391	0.000	0.008	0.000

Notes: This table reports the estimated W matrix for the positive β coefficients tested by $t-stat$. BG stands for Belgium, ES for Spain, FN for Finland, FR for France, GR for Greece, IR for Ireland, IT for Italy, NL for the Netherlands, OE for Austria and PT for Portugal.

Table B.8: W matrix using FIFA related weights

	BG	ES	FN	FR	GR	IR	IT	NL	OE	PT
BG	0.000	0.107	0.036	0.432	0.039	0.050	0.087	0.086	0.054	0.110
ES	0.018	0.000	0.009	0.023	0.010	0.015	0.078	0.071	0.018	0.758
FN	0.031	0.047	0.000	0.034	0.516	0.119	0.054	0.054	0.098	0.047
FR	0.380	0.125	0.035	0.000	0.037	0.049	0.096	0.094	0.054	0.129
GR	0.031	0.048	0.475	0.034	0.000	0.142	0.055	0.056	0.112	0.048
IR	0.033	0.062	0.092	0.037	0.120	0.000	0.076	0.078	0.440	0.060
IT	0.014	0.078	0.010	0.018	0.011	0.019	0.000	0.756	0.023	0.071
NL	0.014	0.072	0.011	0.018	0.012	0.019	0.765	0.000	0.024	0.066
OE	0.035	0.070	0.075	0.040	0.093	0.433	0.091	0.094	0.000	0.069
PT	0.018	0.768	0.009	0.025	0.010	0.015	0.072	0.066	0.017	0.000

Notes: This table reports the estimated W matrix based on the 2020 $FIFA$ ranking. BG stands for Belgium, ES for Spain, FN for Finland, FR for France, GR for Greece, IR for Ireland, IT for Italy, NL for the Netherlands, OE for Austria and PT for Portugal.

Appendix C. Bootstrapped Likelihood Ratio Test

Bootstrapping procedures are widely used when size distortions are encountered, especially in finite samples. We explain the different steps implemented for the empirical exercise.

1. Estimate

$$\Delta x_{i,t} = \hat{\beta}_i x_{i,t-1} + \hat{\lambda}_{i1} x^*_{i,t-1} + \hat{e}_{i,t}. \tag{C.1}$$

   The coefficients $(\hat{\beta}_i, \hat{\lambda}_{i1})$ and the residuals $\hat{e}_{i,t}$ of this *VEC* models are retrieved using the seemingly unrelated regression estimator.

2. Once estimated, we can rewrite the global *VEC* models using (18) as follows (abstaining here from the deterministic components):

$$\Delta X_t = \hat{B} X_{t-1} + \hat{e}_t, \tag{C.2}$$

   where $\hat{B} = [diag(\hat{\beta}_1, \ldots, \hat{\beta}_N) + diag(\hat{\lambda}_{11}, \ldots, \hat{\lambda}_{N1})\tilde{W}]$, and $\tilde{W}$ is defined as (7).

3. Draw with replacement a sequence of residuals $\{\tilde{e}_{1,t}, \ldots, \tilde{e}_{N,t}\}_{t=2}^{T}$. The sequence of resampled errors with replacement is obtained using the wild bootstrap procedure out-

lined in Mammen (1993). This method allows robust statistical inference when unknown forms of heteroskedasticity are present in the data. Specifically, $\{\tilde{e}_{1,t}, \ldots, \tilde{e}_{N,t}\}_{t=2}^{T} = \{k_t\hat{e}_{1,t}, \ldots, k_t\hat{e}_{N,t}\}_{t=2}^{T}$, with k_t being a random sequence with zero mean and unit variance. The distribution proposed by Mammen (1993) is of the following form:

$$k_t = \begin{cases} \frac{1+\sqrt{5}}{2}, & \text{with probability } p = \frac{\sqrt{5}-1}{2\sqrt{5}} \\ \frac{1-\sqrt{5}}{2}, & \text{with probability } 1-p. \end{cases}$$

4. Generate the bootstrapped data sample $\tilde{x}_{it}$ using the first actual observations as starting values for the different series. The subsequent bootstrapped observations are computed as:

$$\Delta\tilde{X}_t = \hat{B}\tilde{X}_{t-1} + \tilde{e}_t, \tag{C.3}$$

with $\tilde{X}_{t-1} = \tilde{X}_{t-2} + \Delta\tilde{X}_{t-1}$

5. Estimate the unrestricted *VEC* model and the *GVEC* model on the bootstrapped sample, and calculate the LR value $\tilde{LR}_1^*$.

6. Repeat $3-5$ a large number of times BOO (in our case, we repeat it $1,000$ times), and build the distribution of the LRs $\{\tilde{LR}_i^*\}$ of dimension BOO.

7. The $\alpha\%$ critical value is the α percentile of $\{\tilde{LR}_i^*\}$. When the test statistics exceed this critical value, the null hypothesis of the validity of the W matrix is rejected at $\alpha\%$.