Testing for the Interconnection Channel in Global VAR models*

Bertrand Candelon† LIDAM/
LFIN, UCLouvain
and
Angelo Luisi‡§
Department of Economics, Ghent University

Abstract

In a globalized world, modeling macro-financial interconnections is fundamental for meaningful inferences. Global Vector Autoregressive models (*GVAR*s) offer an easy and intuitive framework to deal with foreign information when modeling local markets/economies. Local *VAR*s are augmented by the weighted average of foreign counterparts, employing pre-specified distance matrices (W) justified by economic theory, but not empirically tested. We therefore design a Likelihood Ratio Test for the validity of the proposed distance proxy. In the empirical application regarding euro area sovereign bond yields, we show that existing literature neglected a fundamental feature, the sign of the interconnection. Interestingly, the non-rejected matrix outlines the presence of contagion and flight-to-quality mechanisms in the euro area sovereign bond market well before the euro area debt crisis.

Keywords: Likelihood Ratio Test, GVAR, Interdependence, Spillovers, Fragmentation
JEL Classification Codes: C12, C32, C52, E43, H63

*The authors thank Francisco Blasques, Christian Brownlees, Carlo Favero, Christian Hafner, Heiko Hesse, Fabrizio Iacone, Gert Peersman, Mohammad Hashem Pesaran, Eugen Pircalabelu, Frank Smets, Jonas Striaukas, Andrzej Szczepaniak (discussant), Jean-Charles Wijnandts, and the participants at the 12^{th} International Workshop of Methods in International Finance Network (MIFN), the 19^{th} EconomiX-CNRS and Université de Paris Nanterre Econometrics day, the IESEG seminar series, the 14^{th} Financial Risks International Forum, the 7^{th} RCEA Time Series Workshop, the 8^{th} Annual Conference of the International Association for Applied Econometrics (IAAE), and the UCLouvain and Ghent University internal 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.

†*Email Address*: candelonb@gmail.com
‡*Email Address*: angelo.luisi@ugent.be
§Corresponding author. *Address*: Sint-Pietersplein 5, 9000, Ghent, Belgium.

1 Introduction

In a globalized world, local economies/markets are strongly interconnected. Interconnections can result from various sources, common risk factors, shared resources, cross-border effects, spillovers, contagion, flight-to-quality. Global Vector Autoregressive models (*GVAR*s) offer 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 reduced form *VAR* representation of the world, useful for forecasting, scenario analysis, and describing shock diffusion.

Since the seminal contribution of Pesaran et al. (2004), *GVAR* models have been employed in forecasting macroeconomic and financial variables (Pesaran et al., 2009; Greenwood-Nimmo et al., 2012), in credit risk modeling (Pesaran et al., 2006; Castrén et al., 2010), in identifying the interconnection between uncertainty and economic activity (Cesa-Bianchi et al., 2020), as well as in studying international macroeconomic linkages (Dees et al., 2007; Favero et al., 2011). For an extensive literature review, see Di Mauro and Pesaran (2013), Pesaran (2015), or Chudik and Pesaran (2016). Moreover, theoretical justifications regarding *GVAR* models have been provided, both from an econometric perspective - Dees et al. (2007) show that *GVAR* local models represent a suitable approximation of global factor models, while Chudik and Pesaran (2011) derive the conditions under which high-dimension *VAR* models' unknown parameters would deliver local models in a *GVAR* fashion - and from a macro-financial perspective - Dees et al. (2009) show that in the context of New Keynesian Phillips Curves, the global perspective provides several advantages compared with the standard statistical procedures employed in the literature such as

Hodrick-Prescott filters (Hodrick and Prescott, 1997) for the estimation of steady states.

However, a central discussion on $GVAR$ models regards the correct specification of the links among market-/country-specific blocks. Traditionally, the underlying assumption in $GVAR$ modeling is that the more two markets/countries interact in terms of trade, the more interconnected they are. However, this assumption has been challenged in empirical literature. Lane and Shambaugh (2010) show how, in financial applications, weighting schemes based on financial asset exposures perform better than trade based ones. Favero and Missale (2012) and Favero (2013) show how, in the context of sovereign yields modeling, fiscal fundamentals better reflect investors' expectations compared with traditional channels. Gross et al. (2018) modify $GVAR$ models including several different weighting schemes for macro-financial analyses. Moreover, Gross (2019) shows how different estimated $GVAR$ weights are when compared to trade based ones in macroeconomic applications.

The choice of different weighting schemes 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 ones. To the best of our knowledge, no formal test to resolve this issue exists.

Specification tests in $GVAR$ literature (see, for example Dees et al., 2007) regard parameters' stability and lag length decisions once the weighting scheme has been chosen. Cross-sectional dependence tests (Bailey et al., 2016) only guide the choice between sparse or dense interaction matrices, and outline the correct estimation procedure to follow (see, Elhorst et al., 2021). However, they do not propose a validation test, thus not distinguishing between empirically valid sparse or dense interconnections.

In this paper, we propose a formal test for the empirical validity of the proposed channel of interconnection (summarized in the weights employed in each local $VARX^*$, collected in the so-called W matrix). Exploiting the specific estimation procedure in $GVAR$ modeling, we can test the reduced form $GVAR$ representation of the world as a W-dependent-restricted estimation procedure of reduced form VAR, and build a Likelihood Ratio (LR) test for the restrictions imposed by the $GVAR$ setting. Therefore, we offer a testing strategy for the validity of the proposed set of weights derived from the economic literature. The finite sample properties (size and power) of this new test are assessed via Monte Carlo simulations, proving that the test distribution indeed converges to the typical $\chi^2_{(m)}$ distribution where m represents the number of restrictions imposed by the $GVAR$ setting.

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

As a preview of our results, we first show that the existing specification strategy based on the goodness-of-fit to justify the proposed interconnection matrix does not rule out naïve W matrices. Moreover, we show that existing literature has designed weights focusing on whether the interconnection among nodes is strong or weak, neglecting the importance of the sign of the interconnection. Specifically, in our empirical setting, the sign offers 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 (Beber et al., 2009; 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 $GVAR$ representation and its local-to-general estimation procedure. It also develops the LR test for the validity of the W matrix and evaluates the finite and large samples properties (size and power) of the validity test using Monte Carlo methods. Section 3 describes the empirical illustration, Section 4 presents the results, and Section 5 concludes the paper.

2 Methodology

2.1 Local *VARX** estimation, and the Global *VAR* solution

Following Pesaran et al. (2004), let us consider a system of N nodes, representing countries or regions, indexed by $i = 1, 2, \ldots, N$, observed over a certain period T, indexed by $t = 1, 2, \ldots, T$. Each node features k_i local variables collected in the $k_i \times 1$ vector $Y_{i,t}$. We can collect all the node-specific variables in the $k \times 1$ vector $Y_t = (Y'_{1t}, \ldots, Y'_{Nt})'$ with $k = \sum_1^N k_i$. Local Vector Autoregressions get augmented by the the so-called *star variables* $Y^*_{i,t}$ of dimension $k^*_i \times 1$. *Star variables* are built as weighted averages of foreign counterparts' variables for each node i. The objective is to obtain small scale country specific conditional models to estimate.

Foreign weighted averages are of the form,

$$Y^*_{i,t} = \tilde{W}_i Y_t. \tag{1}$$

The matrix $\tilde{W}_i$ has dimension $k^*_i \times k$ and collects the country specific weights,[1] each row-element-sum is equal to 1, and it measures the *ex ante* defined interaction of each node with the foreign counterparts. Typically, in Global *VAR* literature, trade shares between countries are assumed to be the channel of transmission, assuming hence that the more two countries trade, the more interconnected they are.[2]

The resulting local Vector Autoregressive augmented models are of the following form,[3]

$$Y_{i,t} = \Phi_i Y_{i,t-1} + \Lambda_{i0} Y^*_{i,t} + \Lambda_{i1} Y^*_{i,t-1} + \epsilon_{i,t}, \tag{2}$$

where $\epsilon_{i,t} \overset{i.i.d.}{\sim} (0, \Sigma_i)$ is the $k_i \times 1$ idiosyncratic residual term for node i with variance-

[1] The weights are assumed to be non-negative.

[2] Weights can also be variable specific (for example, trade shares for macro variables and financial flows for financial variables) or time varying (see, for example, Cesa-Bianchi et al., 2012).

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

covariance matrix Σ_i of dimension $k_i \times k_i$, Φ_i is the $k_i \times k_i$ matrix of lagged coefficients, and Λ_{i0} and Λ_{i1} are the $k_i \times k_i^*$ matrices collecting the coefficients associated with the foreign variables.

It is worth noting that the estimation takes place at the local level in (2). Thus, to obtain the *GVAR* representation of the global economy, we need to simultaneously solve for all the domestic variables. Therefore, by stacking all local Vector Autoregressive augmented models in (2) we obtain the equivalent representation,

$$Y_t = \tilde{\Phi} Y_{t-1} + \tilde{\Lambda}_0 Y_t^* + \tilde{\Lambda}_1 Y_{t-1}^* + \epsilon_t, \tag{3}$$

where $Y_t^* = ({Y^*}'_{1,t}, \dots, {Y^*}'_{N,t})'$ is the $k^* \times 1$ vector collecting all the node specific *star variables*, with $k^* = \sum_{i=1}^N k_i^*$, $\tilde{\Phi} = diag(\Phi_1, \dots, \Phi_N)$ is of dimension $k \times k$, $\tilde{\Lambda}_0 = diag(\Lambda_{10}, \dots, \Lambda_{N0})$ and $\tilde{\Lambda}_1 = diag(\Lambda_{11}, \dots, \Lambda_{N1})$ are of dimension $k \times k^*$, and $\epsilon_t = (\epsilon'_{1,t}, \dots, \epsilon'_{N,t})'$ is the $k \times 1$ vector obtained by stacking all the country specific residual terms from (2).

Moreover, from (1), we can re-write the vector Y_t^* in terms of the local variables' vector Y_t as

$$Y_t^* = \tilde{W} Y_t, \tag{4}$$

with $\tilde{W} = (\tilde{W}'_1, \dots, \tilde{W}'_N)'$ being $k^* \times k$-dimensional matrix.

Therefore, we can specify (3) in terms of the local variables as

$$Y_t = \tilde{\Phi} Y_{t-1} + \tilde{\Lambda}_0 \tilde{W} Y_t + \tilde{\Lambda}_1 \tilde{W} Y_{t-1} + \epsilon_t. \tag{5}$$

Re-arranging the terms we obtain

$$[I_k - \tilde{\Lambda}_0 \tilde{W}] Y_t = [\tilde{\Phi} + \tilde{\Lambda}_1 \tilde{W}] Y_{t-1} + \epsilon_t. \tag{6}$$

We can define $G = I_k - \tilde{\Lambda}_0 \tilde{W}$, and $H = \tilde{\Phi} + \tilde{\Lambda}_1 \tilde{W}$, and express (6) as

$$G Y_t = H Y_{t-1} + \epsilon_t. \tag{7}$$

The matrix G is generally full rank and hence non-singular (Pesaran et al., 2004).[4] Therefore, we can express the $GVAR$ representation of the world as the reduced form VAR

$$Y_t = G^{-1}HY_{t-1} + G^{-1}\epsilon_t, \tag{8}$$

or, equivalently,

$$Y_t = \Pi Y_{t-1} + \eta_t, \tag{9}$$

$$\Pi = G^{-1}H, \tag{10}$$

$$\eta_t = G^{-1}\epsilon_t. \tag{11}$$

The $GVAR$ estimation at the local level of the parameters characterizing (2), and the subsequent representation in (6) and (8), offers an interesting avenue for evaluating the validity of the *ex ante* specified interconnection matrix $\tilde{W}$.

Indeed, let's first consider the case of a typical *Structural VAR* representation of the form,

$$B_0Y_t = B_1Y_{t-1} + \omega_t, \tag{12}$$

with B_0 and B_1 being $k \times k$ matrices. B_1 collects the autoregressive slope coefficients, while B_1 the instantaneous relation among target variables. The $k \times 1$ vector ω_t is assumed to be white noise, with its elements mutually uncorrelated with non-singular diagonal covariance matrix Σ_ω. The reduced form representation is of the form,

$$Y_t = \underbrace{B_0^{-1}B_1}_{A_1}Y_{t-1} + \underbrace{B_0^{-1}\omega_t}_{v_t}. \tag{13}$$

Compared with the $GVAR$ procedure outlined above, the parameters A_1 and v_t are typically estimated in the reduced form VAR, and subsequently, the matrix B_0 (or B_0^{-1}) is identified from economic theory, institutional knowledge, or other external constraints.[5]

[4]The case of rank deficient matrix G is discussed in Pesaran (2015).

[5]For an extensive treatment of *Structural VAR* and identification strategies, see Kilian and Lütkepohl (2017).

The Likelihood level (whether considered at the structural or reduced form) reached by estimating A_1 and v_t is not *structure*-specific as, whatever assumption is made with regards to the structure of the system, $B_0^{-1}B_1 = A_1$ and $B_0^{-1}\omega_t = v_t$.

Conversely, in the *GVAR* case, in order to obtain the parameters G, H, and ϵ_t, first local parameters in (2) are estimated, and only subsequently, the reduced form *VAR* representation in (8) is retrieved. Therefore, each specification of the *ex ante* proposed matrix $\tilde{W}$ implies different estimated local parameters and thus final reduced form representation.

It is important to underline that compared with standard *Structural VAR*s, the local *VARX** specifications do not require Σ_i to be diagonal (in (13) Σ_ω is required to be diagonal). Therefore, the residual terms $\epsilon_{i,t}$ cannot be considered structural shocks in the sense that each of the k variables is driven by k distinct unidentified shocks.

2.2 Testing Strategy: the Likelihood Ratio Test

By looking at the *GVAR* reduced form representation in (9), we can exploit the specific local-to-general estimation procedure for testing any $\tilde{W}$ matrix employed empirically. The reduced form *GVAR* representation in (9) assumes that the *ex ante*-imposed W matrix, and the subsequently estimated parameters, once the model is solved, maps the local *VARX** systems into a valid global reduced form *VAR* representation.

Our objective is to design a testing strategy for $H_0 : \tilde{W} = W$ vs $H_1 : \tilde{W} \neq W$, with W indicating the specific matrix employed for the local estimation. We exploit the intuition that each W matrix proposed implies a different level of the reduced form log-Likelihood function once the local parameters are estimated. Such a testing logic is not possible in the *Structural VARs'* framework as the identification strategy does not affect the level of the log-Likelihood.[6] Specifically, the reduced form *VAR* serves as the general representation of

[6] In particular, first the unrestricted parameters are estimated, the identification assumptions are then proposed and the model solved to arrive to the interconnected local specifications.

the world that the $GVAR$ model estimates, once the specific channel of interconnection is specified.

$GVAR$ is therefore a W-specific nested model of the reduced form VAR, that represents the unrestricted *benchmark* for evaluating the $GVAR$ assumptions regarding the channel of interconnections. If indeed the interconnection channel proposed is empirically valid, the log-Likelihood value attained by the $GVAR$ estimation and the value obtained by directly estimating the reduced form VAR will not be statistically different from each other.

Assuming that the probability density function of the process is known, we indicate with $lnL(\theta_u)$ the log-Likelihood function of the reduced form VAR representation of the system characterized in (9) with no restrictions on the parameters' space $\theta_u = \{\Pi, \Sigma_\eta\}$. The log-Likelihood function of the $GVAR$ system as in (6) is indicated as $lnL_T(\theta_r)$ with $\theta_r = \{\tilde{W}, \tilde{\Phi}, \tilde{\Lambda}_0, \tilde{\Lambda}_1, \Sigma_\epsilon\}$. We include $\tilde{W}$ among the parameters to explicitly give evidence that the specific channel of interconnections proposed impacts the log-Likelihood as outlined before.[7] After the estimation, the model is solved to arrive to the reduced form specification in (9) featuring (10) and (11). It is therefore clear that the validity of the interconnection matrix can be assessed through the Likelihood Ratio Test (LRT) statistic computed as,

$$LR = -2T[lnL(\hat{\theta}_r|\tilde{W} = W) - lnL(\hat{\theta}_u)], \tag{14}$$

with $lnL(\hat{\theta}_r|\tilde{W} = W)$ being the log-Likelihood level attained after the estimation of the $GVAR$ parameters once the interconnection channel has been specified. The unrestricted VAR maximum log-Likelihood level is $lnL(\hat{\theta}_u)$. Wilks (1938) result for large-sample distribution of the Likelihood Ratio Test statistic ensures that such a LRT statistic is asymptotically distributed as a $\chi^2_{(m)}$. The degrees of freedom, m, are computed as the difference in the parameters' spaces of the unrestricted VAR, and the $GVAR$ models.

The proposed LR test in (14) is an asymptotic test. The unrestricted VAR to estimate,

[7] In fact, Gross (2019) estimates the interconnection channel jointly with the other parameters.

and the $GVAR$ solution to derive, feature a number of parameters that would result in over-rejecting the true interaction matrix $\tilde{W}$ (if correctly identified) in empirical exercises. As outlined in the following simulation section, the over-rejection increases the larger the number of items/nodes, and the shorter the time series. It thus becomes fundamental to consider LR test corrections in finite samples.

Importantly, existing literature offers solutions to address the short sample distortions through (i) bootstrapping techniques (Kilian, 1998; Kim, 2014), (ii) applying a Bartlett correction, also possible jointly with bootstrapping techniques (see, for example, Lagos and Morettin, 2004; Canepa and Godfrey, 2007), and/or (iii) applying high-dimensional corrections (see, for example, Bai et al., 2009, 2013).

2.3 Size and Power Analyses

To analyze the behavior of the validity test in finite samples, size and power analysis is proposed via Monte Carlo simulations. The aim is to provide evidence that indeed, $GVAR$ models are W-dependent restricted procedures to estimate global reduced form $VARs$. Once the empirical convergence in large sample to the asymptotic significance level of the correspondent theoretical χ^2 is verified, we propose the size-adjusted power. The power will be verified for two sets of errors in the identification of the interconnection channel, when half of the weights are wrongly identified and when all the weights are wrongly identified. The average identification errors proposed are small in size, in order to verify the capability of the test of rejecting small departures from the theoretically correct model.

2.3.1 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, the $GVAR$ model 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.[8] Several *DGPs* are considered to evaluate the behavior of our validity test under the null hypothesis. The specific parameters employed are available in the Supplementary Material.

The first *DGP* (DGP_1) corresponds to a stationary 4-nodes *GVAR* model. 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 $\tilde{\Lambda}_1$ and $\tilde{\Phi}$). 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-nodes *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 unrestricted *VAR* slope parameters to estimate are $10 \times 10 = 100$ while the *GVAR* parameters' space features $10 \times 2 = 20$ elements (i.e., the elements of the diagonal matrices $\tilde{\Lambda}_1$ and $\tilde{\Phi}$). The degrees of freedom of the validity LR test are thus 80.

DGP_3 corresponds to the 10-nodes vector error correction model proposed by Favero (2013). The parameters chosen are the ones reported in Favero (2013), the weights used are fixed and obtained from the average of the time-varying weights employed in the original paper. It is worth noting that two separate $\tilde{W}$s are employed. Local errors are normally distributed with zero mean and 0.5 variance. The unrestricted *VEC* slope parameters

[8] Codes in *R* are available from the authors upon request.

to estimate are $10 \times 10 = 100$ while the $GVEC$ parameters' space features $10 \times 3 = 30$ elements (i.e., the elements of the two diagonal matrices $\tilde{\Lambda}_1$, as two interconnection matrices are employed, and $\tilde{\Phi}$). The degrees of freedom of the validity LR test are thus 70.

DGP_4 corresponds to a stationary 11-nodes $GVAR$ 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 unrestricted VAR slope parameters to estimate are $11 \times 11 = 121$ while the $GVAR$ parameters' space features $11 \times 3 = 33$ elements (i.e., the elements of the diagonal matrices $\tilde{\Lambda}_0$, $\tilde{\Lambda}_1$, and $\tilde{\Phi}$). The degrees of freedom of the validity LR test are 88.

To derive the empirical size of this test, $10,000$ replications per DGP are generated, and we report in Table 1 the rejection frequencies 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.

The rejection frequencies for the 4-node 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.[9]

[9] The bootstrap procedure is presented in the Supplementary Material.

Table 1: **Rejection frequencies under the null hypothesis.** This table shows the rejection frequencies, as a %, of the null hypothesis considering several DGPs. DGP_1 corresponds to the 4-node stationary case, DGP_2 corresponds to the 10-node stationary case, DGP_3 corresponds to the non-stationary case proposed by Favero (2013), and DGP_4 corresponds to the 11-country 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.

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	47.18	32.63	25.98	12.36	5.85

2.3.2 Power analysis

For the power analysis, the four previous $DGPs$ are retained along with their parameters. Instead of simulating under the null hypothesis, alternative Ws are employed. Several misspecifications are considered. For DGP_1, W is simply composed of equal weights (i.e., 0.33). For DGP_2, DGP_3, and DGP_4, we randomly draw the alternative matrices.[10] To highlight the magnitude of the difference between the W matrix considered and that under the null hypothesis ($\tilde{W}$), we compute the average absolute distance ($dist(\mathbf{W}, \tilde{\mathbf{W}}) = \frac{1}{\#of weigths} \sum_{i=1}^{N} \sum_{j=1}^{N} |w_{ij} - \tilde{w}_{ij}|$). The same 5 sample sizes considered in

[10]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 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 the size. Rejection frequencies, which correspond to the size-adjusted power, are reported in Table 2.

Table 2: **Rejection frequencies under alternative hypotheses.** This table shows the rejection frequencies as a % of the null hypothesis when simulating several DGPs with different misspecified transmission matrices $(\tilde{W})$. DGP_1 corresponds to the 4-node stationary case, DGP_2 corresponds to the 10-node stationary case, DGP_3 corresponds to the non-stationary case proposed by Favero (2013), and DGP_4 corresponds to the 10-node stationary case with a contemporaneous term. Half (All) indicates that the misspecification is imposed for half (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 $T = 100, 150, 200, 500, 1,000$. 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.

	$dist(\mathbf{W}, \tilde{\mathbf{W}})$		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	0.032	0.060	6.86	9.90	8.15	13.74	9.49	18.55	20.15	55.33	46.62	94.28

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 0.99 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 often not rejected in small samples. This result is coherent with the literature considering corrections to the short sample problems of asymptotic tests, as outlined in Subsection 2.2.

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_4 contains a contemporaneous term that deteriorates the power, whereas DGP_3 is nearly non-stationary, also impacting the rejection frequency. 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., lower $dist(\mathbf{W}, \tilde{\mathbf{W}})$) always exhibit a lower power than in the case in which all the weights are misspecified (i.e., higher $dist(\mathbf{W}, \tilde{\mathbf{W}})$).

3 Empirical Analysis

To illustrate the importance of the proposed testing strategy 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, for example, Sgherri and Zoli, 2009; Favero and Missale, 2012) 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.[11] 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.

3.1 Data description

The empirical exercise concerns 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.[12] 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.

[11] W indicates the specific matrix employed, not to be confounded with $\tilde{W}$ which is the valid interconnection matrix.

[12] 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).

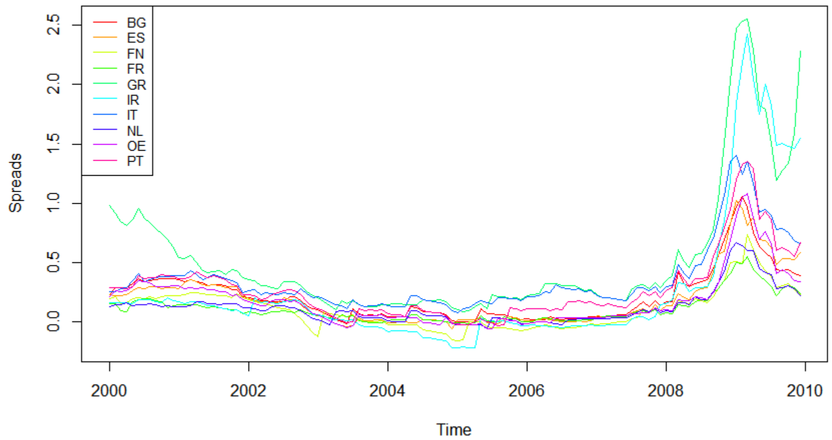

Figure 1: **Long-term sovereign bond spreads over the German *Bund*.** This figure shows the behavior of 10-Year sovereign bond spreads over the German benchmark from January 2000 to December 2009, in % points. The countries under analysis are Austria (OE), Belgium (BG), Finland (FN), France (FR), Greece (GR), Ireland (IR), Italy (IT), the Netherlands (NL), Portugal (PT), and Spain (ES). Frequency: Monthly. Source: *DataStream*.

3.2 Modeling sovereign bond spreads: The choice of *W*

Our aim is to evaluate the validity of the W matrices employed to model sovereign yield spreads, taking into account the importance of non-local, 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 the empirical validity of the matrices of interactions proposed by existing literature, and the empirical capability of rejecting matrices that are not valid.

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

$$\Delta Y_{i,t} = \beta_{i0} + \beta_{i1} Y_{i,t-1} + \beta_{i2}(Baa_{t-1} - Aaa_{t-1}) + \beta_{i3}\Delta(Baa_t - Aaa_t) + e_{i,t}, \tag{15}$$

where $\Delta Y_{i,t} = Y_{i,t} - Y_{i,t-1}$, $Y_{i,t}$ is the 10-year sovereign yield spread between country i and Germany (the usual reference in the euro area), $e_{i,t}$ is the white noise residual term for country i with finite variance, $Baa - Aaa$ represents the time varying global risk aversion as measured by the long term US $Baa - Aaa$ corporate bond spreads. This framework is employed in Favero (2013) as a *benchmark* to assess the relevance of foreign information, as it does not include any weighted average of foreign counterparts. We will refer to it as *basic* model.

As anticipated, Favero (2013) proposes to augment the specification in (15) in a *GVAR* fashion. Specifically, he identifies $w_{ij,t}$ 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 and deficit-to-GDP ratio 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 (i.e., 3% for deficit, and 60% for debt). The deficit- ($def_{i,t}$) and debt-to-GDP ($debt_{i,t}$) 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. It is then possible to build two distance matrices as follows,

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

$$dist^{def}_{ji,t} = |def_{j,t} - def_{i,t}|/3,$$

$$dist^{debt}_{ji,t} = |debt_{j,t} - debt_{i,t}|/60,$$

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

$$\Delta Y_{i,t} = \beta_{i0} + \beta_{i1} Y_{i,t-1} + \beta_{i2}(Baa_{t-1} - Aaa_{t-1}) + \beta_{i3}\Delta(Baa_t - Aaa_t) + \beta_{i4}(debt_{i,t} - debt_{bd,t}) + \beta_{i5}(def_{i,t} - def_{bd,t}) + \beta_{i6} Y^{*,debt}_{i,t-1} + \beta_{i7} Y^{*,def}_{i,t-1} + e_{i,t}, \tag{16}$$

where $Y^{*,k}_t = \sum_{j \neq i} w^k_{ji,t} Y_{j,t}$ with $k = \{def, debt\}$, $\Delta(Baa_t - Aaa_t) = (Baa_t - Aaa_t) - (Baa_{t-1} - Aaa_{t-1})$. The intuition would be that the more two countries are similar 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, expressed as distances from the German values (in (16), *bd* indicates Germany).

In order to show that existing strategies regarding model adequacy in $GVAR$ models are not capable of conclusively rejecting naïve W matrices, we consider a Global model including the country ranking maintained by the International Federation of Association Football in 2020 as interconnection channel.[13] 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.[14] The model thus has the following form,

$$\Delta Y_{i,t} = \beta_{i0} + \beta_{i1} Y_{i,t-1} + \beta_{i2}(Baa_{t-1} - Aaa_{t-1}) + \beta_{i3}\Delta(Baa_t - Aaa_t) + \beta_{i4} Y^{*,FIFA}_{i,t-1} + e_{i,t}, \tag{17}$$

where $Y^{*,FIFA}_{i,t} = \sum_{j \neq i} w^{FIFA}_{ij} Y_{j,t}$. 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.

[13] 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.

[14] W matrices are reported in the Supplementary Material.

3.2.1 Assessing Model Adequacy using the Goodness-of-fit

Typically, the relevance of including foreign information in country specific systems is assessed by comparing the *basic* model in (15) vis-à-vis the specific $GVAR$ model proposed. In Table 3, we report the Adjusted R-squared for the *basic* model, the naïve global model featuring *FIFA*-ranking based weigths along with Favero (2013) debt- and deficit-to-GDP ratio based weights described above. The adjusted R-squared measure of the goodness-of-fit ensures that simply adding a larger number of covariates does not imply automatically a better performance of a model.

Table 3: **Adjusted R-squared for the country systems.** This table reports the Adjusted R^2 values calculated for the different models (Basic, FIFA ranking, Favero, 2013). The largest adjusted R^2 per country are in bold. The basic model is the one with no spillovers among country blocks.

Country	Adj. R^2		
	Basic	FIFA	Favero (2013)
Belgium	0.06	**0.08**	**0.08**
Spain	**0.15**	0.12	0.12
Finland	0.15	0.18	**0.19**
France	0.13	**0.17**	0.14
Greece	0.16	0.19	**0.31**
Ireland	0.28	0.29	**0.36**
Italy	0.23	**0.25**	0.20
Netherlands	0.25	**0.29**	0.27
Austria	0.19	0.21	**0.37**
Portugal	0.16	0.23	**0.30**

First, employing the Adjusted R-squared to compare the *FIFA*-ranking *GVAR* model with the *basic* one, we would directly admit that a W matrix based on (future) football performance is a good indicator of the interconnections observed in the euro area. This reasoning is though biased. Specifically, in a setting such as the one of the European Monetary Union, where countries are structurally interconnected, comparing a framework that admits no interconnections with one that admits interconnections, we risk to obtain misleading results. In fact, such a comparison can only testify the *rejection* of the *basic* model in favour of *any* model admitting interconnections.

Second, when comparing the goodness-of-fit measures of the *FIFA*-ranking-based interconnections with the debt- and deficit-to-GDP ratio as in Favero (2013), we can outline the second drawback of such a model adequacy decision. Specifically, although numerically the Adjusted R-squared is larger for more countries in the case of the framework proposed by Favero (2013), the *FIFA* ranking exhibits better performances for relevant countries such as France, Italy, and the Netherlands, while exhibiting the same performance for Spain and Belgium. This testing strategy for the interconnection channel does not even rule out naïve matrices. Its capability of rejecting one interconnection channel compared to another one in competitive models is very limited. Therefore, it appears necessary the design of a testing procedure for the empirical validity of the interconnection channel.

3.2.2 Estimating interconnections and the sign dilemma

Before introducing the results from our testing strategy in the empirical exercise, it is worth noting that it is possible, in *GVAR* modeling, to estimate the interconnection weights.

Specifically, Gross (2019) proposes to estimate the W matrix jointly with the local *GVAR* coefficients. The procedure starts from local models of the form,

$$\Delta Y_{i,t} = \beta_{i0} + \beta_{i1} Y_{i,t-1} + \beta_{i2} \sum_{i \neq j} w_{ij} Y_{j,t-1} + e_{i,t}, \tag{18}$$

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

$$\min_{\Gamma_i, w_{ij}} \sum_{t=1}^{T} e_{i,t}^2 \tag{19}$$

subject to, $w_{ij} \geq 0$ for $i \neq j$, $w_{ii} = 0$ and $\sum_{j=1}^{N} w_{ij} = 1$. Γ_i is the vector collecting all the local parameters of (18), and w_{ij} are the measures of the tightness of the interconnection between countries in a standard *GVAR* framework. Note that each w_{ij} is constrained to be non-negative. The estimated W 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).

However, central to the discussion regarding sovereign bonds in the euro area is the contemporaneous presence of two different mechanisms, "contagion" and "flight-to-quality". 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 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" (Favero and Missale, 2012).

In order to account for such heterogeneity while keeping the weights non-negative, it would be important to split the interaction matrix 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.[15]

To control for this characteristic, and moreover to ensure we have a *benchmark* to assess the empirical capability of our test regarding not rejecting a valid matrix, we propose an additional weighting scheme. Specifically, the unrestricted *VEC* model (featuring all the countries and augmented with the US long-term corporate bond spread factor as common variable) estimation is performed, and the off-diagonal elements (i.e., cross-country coefficients) 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. Importantly, we need to make sure that our LR test indeed does not reject a matrix that is correctly identified, therefore, using the $t-test$ of the unrestricted representation, we make sure that the weights reflect the desirable characteristics for the cross-country weights. This $GVEC$ model is thus of the form

$$\Delta Y_{i,t} = \beta_{i0} + \beta_{i1} Y_{i,t-1} + \beta_{i2}(Baa_{t-1} - Aaa_{t-1}) + \beta_{i3}\Delta(Baa_t - Aaa_t) + \beta_{i4} \sum_{i \neq j} w_{ij}^{+} Y_{j,t-1} + \beta_{i5} \sum_{i \neq j} w_{ij}^{-} Y_{j,t-1} + e_{i,t}. \quad (20)$$

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.[16]

[15]The results do not change if we decide to normalize the weights to sum to 1 in total or by W matrix.

[16]To the best of our knowledge, only Aquaro et al. (2021) consider, in the context of real estate markets, the possibility of a positive and a negative interaction matrix.

4 Results

4.1 Testing naïve vs valid interconnections

In the previous section we described the empirical setting, and the different interconnection matrices considered for our Likelihood Ratio Test based strategy. In this section, we present the results from our validity test and outline the importance of such a model adequacy check. The new validity test (14) is implemented for the four different W matrices, considering both the asymptotic and the bootstrapped critical values at the standard level of 95%. Table 4 reports the results.

Table 4: W **validity test. Naïve vs valid matrices.** This table reports the Likelihood Ratio Test statistics (Test Stat.) calculated for the different W matrices (FIFA, Favero, 2013; Gross, 2019, T-stat based). The asymptotic critical values (Asymp. CV 95%) correspond to the 95% quantile of a χ^2 distribution with adequate degrees of freedom. The bootstrapped critical value (Boot. CV 95%) corresponds to the 95% quantile of the distribution obtained using the sequence presented in the Supplementary Material. When the null hypothesis is not rejected at a confidence level of 95% at the bootstrapped level, the test statistic appears in bold.

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

First, the test rejects the null hypothesis of the validity of the W matrix based on the $FIFA$ ranking. This result is satisfying, as finding evidence that sovereign yield spreads' interdependence could be based on the future performance of national football teams would be puzzling. The Likelihood Ratio Test therefore rejects the naïve W matrix, a requirement the goodness-of-fit approach outlined in the previous section did not fulfil.

Second, we want to make sure that our test does not reject a matrix that is empirically valid. Indeed, the partitioned version of W outlined in Section 3.2.2 provides an ideal setting for checking the empirical performance of our test.[17] 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 (20).

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 co-movements 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 the Supplementary Material, 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

[17] The W matrices are reported in the Supplementary Material.

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 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 non-rejected W matrix provides important insights regarding sovereign bond spreads spillover effects inside the euro area even before the outburst of the sovereign debt crisis.

An interesting confirmation regarding the importance of the sign when modeling sovereign bond spreads in the euro area in a $GVAR$ framework comes from the employment of Gross (2019) estimation technique. The estimated W should be accurate from an empirical perspective, at the cost of its economic interpretation. However, the test rejects the validity of this estimated transmission matrix. The reason for such rejection comes from the imposition of the non-negativity of each element of the estimated interconnection matrix. As explained in the previous section, such a constraint forces the 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 Gross (2019) approach causing the rejection of the estimated W.

4.2 Testing the validity of the proposed interconnections

In the previous subsection we made sure that our testing strategy indeed rejects naïve interconnection matrices and does not reject valid ones. Moreover, we outlined the importance of sign heterogeneity of the global counterparts when modeling sovereign bond spreads inside the euro area. It is now time to discuss the debt- and deficit-to-GDP $GVAR$ framework proposed by Favero (2013).

Looking at Table 5, we notice that the fiscal fundamental distance framework gets rejected by our test.[18] Economically, it implies that fiscal indicators (debt- and deficit-to-GDP ratios) alone are not adequate factors to explain the interdependence among sovereign yield spreads across euro area countries. However, as proven in the previous subsection, the rejection might be caused by the non-negativity constraint on the interconnection

[18] 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.

Table 5: W **validity test. The sign dilemma.** This table reports the Likelihood Ratio Test statistics (Test Stat.) calculated for the *FIFA*, and Favero (2013) W matrices, adjusted by the sign. The asymptotic critical values (Asymp. CV 95%) correspond to the 95% quantile of a χ^2 distribution with adequate degrees of freedom. The bootstrapped critical value (Boot. CV 95%) corresponds to the 95% quantile of the distribution obtained using the sequence presented in the Supplementary Material. FIFA Adj. corresponds to a model that employs the weights based on the FIFA ranking, but featuring two W matrices, one collecting weights for countries exhibiting a negative sign in the unrestricted VAR, and one collecting countries exhibiting a positive sign in the unrestricted VAR. Favero (2013) Adj. corresponds to a model that employs the weights as in Favero (2013), but featuring two W matrices, one collecting weights for countries exhibiting a negative sign in the unrestricted VAR, and one collecting countries exhibiting a positive sign in the unrestricted VAR. When the null hypothesis is not rejected at a confidence level of 95% at the bootstrapped level, the test statistic appears in bold.

W	Test Stat.	Asymp. CV 95%	Boot. CV 95%
Favero (2013)	303.02	90.53	197.97
FIFA Adj.	281.04	90.53	206.04
Favero (2013) Adj	**127.35**	67.51	138.87

matrices being too tight for the setting considered. Therefore, we assess the validity of the interconnection channels proposed when adjusting by the sign. Specifically, we keep the division of foreign counterparts into positive and negative spillover groups as in (20) while employing the weights proposed by Favero (2013). Very interestingly, we can see from Table 5 that such a framework does not get rejected at the bootstrapped level. Intuitively, this result implies that indeed the channels proposed by Favero (2013) are valid interconnection proxies. The more two countries are similar in terms of debt- and deficit-to-GDP ratios, the more interconnected they are. However, the interconnection with foreign counterparts exhibits an heterogeneous effect. Indeed, the non-rejected framework allows new inferences related to the coexistence of “fight-to-quality” and “contagion” effects among countries.

It is now important to make sure that the sign is not the only reason why the modified Favero (2013) framework does not get rejected by the test. We therefore modify the *FIFA* ranking accordingly. Again, we divide foreign counterparts into the two groups implied by the framework in (20), and then apply the weights from the *FIFA* ranking distance. As we can see from Table 5, even when adjusted by the sign of the interconnections, the *FIFA* ranking gets rejected both at the asymptotic and the bootstrapped level. Therefore, although central for making sure that the interconnection matrix is valid, the sign inclusion does not *automatically* imply the non rejection of the W.

5 Conclusion

$GVAR$ models represent an effective and intuitive econometric framework 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 exploiting the local-to-general logic of $GVAR$ estimation process, it is possible to clearly identify the restrictions imposed to characterize the specific transmission channel. Therefore, we design a novel model adequacy LR test based strategy to empirically validate the proposed W matrix. The asymptotic properties of the test are assessed via Monte Carlo methods.

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 as in Favero (2013).

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 proposed 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.

The new test presented in this paper paves the way for a wide range of economic implementations of *GVAR* models. 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.

SUPPLEMENTARY MATERIAL

Testing for the Interconnection Channel in Global VAR models: Supplementary Material. The supplementary material to the paper *Testing for the Interconnection Channel in Global VAR models* reports the Simulated Data Generating Processes employed for the simulation exercise, the W matrices tested in the empirical exercise, and the Bootstrapped Likelihood Ratio test procedure implemented for the empirical exercise.

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.

Aquaro, M., N. Bailey, and M. H. Pesaran (2021). Estimation and inference for spatial models with heterogeneous coefficients: an application to US house prices. *Journal of Applied Econometrics 36*(1), 18–44.

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.

Canepa, A. and L. Godfrey (2007). Improvement of the quasi-likelihood ratio test in ARMA models: some results for bootstrap methods. *Journal of Time Series Analysis 28*(3), 434–453.

Castrén, O., S. Dées, and F. Zaher (2010). Stress-testing euro area corporate default probabilities using a global macroeconomic model. *Journal of Financial Stability 6*(2), 64–78.

Cesa-Bianchi, A., M. H. Pesaran, and A. Rebucci (2020). Uncertainty and economic activity: A multicountry perspective. *The Review of Financial Studies 33*(8), 3393–3445.

Cesa-Bianchi, A., M. H. Pesaran, A. Rebucci, T. Xu, and R. Chang (2012). China's emergence in the world economy and business cycles in Latin America [with comment]. *Economía 12*(2), 1–75.

Chudik, A. and M. H. Pesaran (2011). Infinite-dimensional VARs and factor models. *Journal of Econometrics 163*(1), 4–22.

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.

Dees, S., F. d. Mauro, M. H. Pesaran, and L. V. Smith (2007). Exploring the international linkages of the euro area: a global VAR analysis. *Journal of Applied Econometrics 22*(1), 1–38.

Dees, S., M. H. Pesaran, L. V. Smith, and R. P. Smith (2009). Identification of new Keynesian Phillips curves from a global perspective. *Journal of Money, Credit and Banking 41*(7), 1481–1502.

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.

Elhorst, J. P., M. Gross, and E. Tereanu (2021). Cross-sectional dependence and spillovers in space and time: Where spatial econometrics and global var models meet. *Journal of Economic Surveys 35*(1), 192–226.

Favero, C., F. Giavazzi, and J. Perego (2011). Country heterogeneity and the international evidence on the effects of fiscal policy. *IMF Economic Review 59*(4), 652–682.

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

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.

Greenwood-Nimmo, M., V. H. Nguyen, and Y. Shin (2012). Probabilistic forecasting of output growth, inflation and the balance of trade in a GVAR framework. *Journal of Applied Econometrics 27*(4), 554–573.

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

Gross, M., J. Henry, and W. Semmler (2018). Destabilizing effects of bank overleveraging on real activity - An analysis based on a threshold MCS-GVAR. *Macroeconomic Dynamics 22*(7), 1750–1768.

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.

Hodrick, R. J. and E. C. Prescott (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, credit, and Banking*, 1–16.

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

Kilian, L. and H. Lütkepohl (2017). *Structural vector autoregressive analysis*. Cambridge University Press.

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

Lagos, B. M. and P. A. Morettin (2004). Improvement of the likelihood ratio test statistic in ARMA models. *Journal of Time Series Analysis 25*(1), 83–101.

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

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

Pesaran, M. H. (2015). *Time Series and Panel Data Econometrics*, Chapter 33: Theory and Practice of GVAR Modelling, pp. 900–935. Oxford University Press.

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, B.-J. Treutler, and S. M. Weiner (2006). Macroeconomic dynamics and credit risk: a global perspective. *Journal of Money, Credit and Banking*, 1211–1261.

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*.

Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The annals of mathematical statistics 9*(1), 60–62.

Testing for the Interconnection Channel in Global VAR models:

Supplementary Material

Bertrand Candelon*
LIDAM/LFIN, UCLouvain
and
Angelo Luisi†‡
Department of Economics, Ghent University

Abstract

The supplementary material to the paper *Testing for the Interconnection Channel in Global VAR models* reports the Simulated Data Generating Processes employed for the simulation exercise, the W matrices tested in the empirical exercise, and the Bootstrapped Likelihood Ratio test procedure implemented for the empirical exercise.

* *Email Address*: candelonb@gmail.com
† *Email Address*: angelo.luisi@ugent.be
‡ Corresponding author. *Address*: Sint-Pietersplein 5, 9000, Ghent, Belgium.

1 Simulated Data Generating Processes

We list the different coefficients and W matrices employed for the simulation exercise, both for the size and power sections.

The first DGP is a 4-node stationary case with no contemporaneous term. The coefficients are in tables from 1 to 2. The second DGP is a 10-node stationary case with no contemporaneous term. The coefficients are in tables from 3 to 6. The third DGP is based on the 10-country non-stationary case of Favero (2013). Therefore, it features two different W matrices (computed here as the averages of the debt- and deficit-to-GDP time varying fiscal distance proposed in the paper). The coefficients are in tables from 7 to 13. The fourth DGP is an 11-node case with contemporaneous term. The coefficients are in tables from 14 to 18.

Table 1: $\tilde{\Phi}$ and $\tilde{\Lambda}_1$ of DGP_1

$\tilde{\Phi}$	1	2	3	4
1	0.60	0.00	0.00	0.00
2	0.00	0.50	0.00	0.00
3	0.00	0.00	0.66	0.00
4	0.00	0.00	0.00	0.56

$\tilde{\Lambda}_1$	1	2	3	4
1	0.75	0.00	0.00	0.00
2	0.00	0.60	0.00	0.00
3	0.00	0.00	0.36	0.00
4	0.00	0.00	0.00	-0.85

Table 2: $\tilde{W}$ matrices, on the left the correct and on the right the misspecified for DGP_1

$\tilde{W}_{correct}$	1	2	3	4
1	0.00	0.31	0.25	0.44
2	0.15	0.00	0.42	0.43
3	0.37	0.25	0.00	0.38
4	0.35	0.45	0.20	0.00

$\tilde{W}_{misspecified}$	1	2	3	4
1	0.00	0.33	0.33	0.33
2	0.33	0.00	0.33	0.33
3	0.33	0.33	0.00	0.33
4	0.33	0.33	0.33	0.00

Table 3: $\tilde{\Phi}$ of DGP_2

$\tilde{\Phi}$	1	2	3	4	5	6	7	8	9	10
1	-0.114	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	-0.538	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	-0.039	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.310	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	-0.067	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	-0.375	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	-0.250	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.571	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.385	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.249

Table 4: $\tilde{\Lambda}_1$ of DGP_2

$\tilde{\Lambda}_1$	1	2	3	4	5	6	7	8	9	10
1	-0.379	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.063	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.323	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	-0.494	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.104	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	0.012	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	-0.393	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.218	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.183	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.172

Table 5: W correct matrix for DGP_2

	1	2	3	4	5	6	7	8	9	10
1	0.000	0.056	0.162	0.029	0.132	0.178	0.080	0.131	0.073	0.160
2	0.219	0.000	0.221	0.014	0.183	0.109	0.002	0.199	0.049	0.004
3	0.228	0.154	0.000	0.153	0.002	0.055	0.138	0.049	0.000	0.221
4	0.006	0.173	0.171	0.000	0.140	0.168	0.111	0.072	0.127	0.031
5	0.062	0.115	0.128	0.016	0.000	0.035	0.109	0.180	0.173	0.182
6	0.133	0.019	0.262	0.064	0.178	0.000	0.108	0.078	0.104	0.054
7	0.153	0.096	0.091	0.205	0.045	0.118	0.000	0.116	0.081	0.095
8	0.104	0.082	0.163	0.127	0.155	0.132	0.160	0.000	0.034	0.042
9	0.079	0.217	0.129	0.048	0.144	0.006	0.150	0.027	0.000	0.200
10	0.115	0.113	0.135	0.155	0.032	0.133	0.064	0.135	0.119	0.000

Table 6: W misspecified matrix for DGP_2

	1	2	3	4	5	6	7	8	9	10
1	0.000	0.128	0.094	0.000	0.129	0.026	0.162	0.171	0.141	0.150
2	0.000	0.000	0.137	0.152	0.125	0.208	0.141	0.140	0.079	0.018
3	0.060	0.098	0.000	0.102	0.104	0.119	0.087	0.166	0.123	0.142
4	0.046	0.090	0.018	0.000	0.232	0.164	0.115	0.237	0.022	0.077
5	0.091	0.007	0.216	0.053	0.000	0.200	0.215	0.096	0.113	0.008
6	0.182	0.131	0.011	0.070	0.158	0.000	0.144	0.066	0.188	0.050
7	0.035	0.110	0.226	0.011	0.229	0.029	0.000	0.128	0.024	0.208
8	0.181	0.026	0.031	0.160	0.064	0.112	0.182	0.000	0.082	0.162
9	0.116	0.156	0.160	0.103	0.088	0.168	0.072	0.100	0.000	0.038
10	0.182	0.082	0.185	0.015	0.152	0.135	0.034	0.203	0.012	0.000

Table 7: $\tilde{\Phi}$ of DGP_3

$\tilde{\Phi}$	1	2	3	4	5	6	7	8	9	10
1	-0.161	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	-0.156	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	-0.295	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	-0.415	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	-0.1	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	-0.239	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	-0.279	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.405	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.38	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.28

Table 8: $\tilde{\Lambda}_{1,debt}$ of DGP_3

$\tilde{\Lambda}_{1,debt}$	1	2	3	4	5	6	7	8	9	10
1	0.02	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.08	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.14	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.068	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.261	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	-0.044	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	-0.059	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.02	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.216	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.235

Table 9: $\tilde{\Lambda}_{1,deficit}$ of DGP_3

$\tilde{\Lambda}_{1,deficit}$	1	2	3	4	5	6	7	8	9	10
1	-0.019	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	-0.028	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.035	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.057	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	-0.246	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	0.486	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	0.185	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.09	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.033	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.011

Table 10: W_{debt} correct matrix for DGP_3

	1	2	3	4	5	6	7	8	9	10
1	0.000	0.185	0.027	0.043	0.092	0.194	0.057	0.134	0.211	0.058
2	0.222	0.000	0.062	0.049	0.058	0.044	0.231	0.139	0.098	0.097
3	0.127	0.165	0.000	0.072	0.085	0.165	0.083	0.123	0.108	0.070
4	0.044	0.046	0.018	0.000	0.213	0.030	0.291	0.052	0.060	0.245
5	0.087	0.054	0.018	0.222	0.000	0.108	0.220	0.086	0.084	0.122
6	0.204	0.213	0.072	0.049	0.119	0.000	0.046	0.123	0.126	0.049
7	0.050	0.040	0.018	0.273	0.204	0.029	0.000	0.074	0.092	0.220
8	0.159	0.146	0.039	0.069	0.105	0.121	0.084	0.000	0.207	0.071
9	0.237	0.070	0.023	0.076	0.081	0.136	0.098	0.200	0.000	0.079
10	0.061	0.088	0.020	0.262	0.157	0.036	0.253	0.053	0.072	0.000

Table 11: $W_{deficit}$ correct matrix for DGP_3

	1	2	3	4	5	6	7	8	9	10
1	0.000	0.044	0.039	0.086	0.276	0.040	0.303	0.052	0.071	0.089
2	0.018	0.000	0.349	0.108	0.016	0.111	0.014	0.191	0.113	0.079
3	0.025	0.388	0.000	0.063	0.023	0.228	0.021	0.120	0.062	0.070
4	0.034	0.096	0.034	0.000	0.024	0.039	0.020	0.100	0.261	0.393
5	0.313	0.045	0.037	0.065	0.000	0.037	0.333	0.048	0.060	0.063
6	0.037	0.203	0.306	0.084	0.033	0.000	0.030	0.112	0.095	0.100
7	0.340	0.039	0.034	0.057	0.343	0.035	0.000	0.042	0.053	0.057
8	0.033	0.231	0.121	0.128	0.027	0.024	0.083	0.000	0.191	0.161
9	0.038	0.125	0.047	0.318	0.029	0.061	0.025	0.164	0.000	0.194
10	0.039	0.080	0.045	0.426	0.027	0.055	0.023	0.133	0.172	0.000

Table 12: W_{debt} misspecified matrix for DGP_3

	1	2	3	4	5	6	7	8	9	10
1	0.000	0.092	0.012	0.106	0.172	0.101	0.043	0.066	0.222	0.186
2	0.190	0.000	0.032	0.151	0.221	0.002	0.162	0.137	0.029	0.076
3	0.140	0.067	0.000	0.058	0.167	0.030	0.136	0.184	0.083	0.135
4	0.071	0.156	0.178	0.000	0.180	0.012	0.150	0.089	0.001	0.163
5	0.066	0.005	0.272	0.184	0.000	0.183	0.002	0.066	0.164	0.058
6	0.000	0.150	0.007	0.175	0.172	0.000	0.141	0.169	0.112	0.073
7	0.166	0.040	0.076	0.142	0.158	0.020	0.000	0.044	0.134	0.221
8	0.190	0.199	0.096	0.014	0.190	0.046	0.129	0.000	0.079	0.057
9	0.083	0.043	0.162	0.102	0.096	0.218	0.047	0.125	0.000	0.123
10	0.059	0.069	0.098	0.078	0.154	0.120	0.146	0.125	0.151	0.000

Table 13: $W_{deficit}$ misspecified matrix for DGP_3

	1	2	3	4	5	6	7	8	9	10
1	0.000	0.061	0.076	0.088	0.242	0.144	0.054	0.160	0.006	0.168
2	0.023	0.000	0.174	0.124	0.121	0.146	0.167	0.034	0.143	0.069
3	0.157	0.138	0.000	0.210	0.036	0.036	0.073	0.117	0.091	0.141
4	0.175	0.064	0.051	0.000	0.147	0.025	0.150	0.181	0.061	0.146
5	0.268	0.180	0.086	0.023	0.000	0.109	0.048	0.077	0.022	0.188
6	0.010	0.166	0.136	0.003	0.209	0.000	0.160	0.061	0.090	0.164
7	0.091	0.069	0.174	0.008	0.176	0.013	0.000	0.162	0.143	0.164
8	0.072	0.242	0.029	0.138	0.004	0.226	0.110	0.000	0.147	0.033
9	0.071	0.079	0.055	0.158	0.031	0.124	0.022	0.236	0.000	0.224
10	0.199	0.007	0.164	0.177	0.123	0.184	0.023	0.073	0.050	0.000

Table 14: $\tilde{\Phi}$ of DGP_4

$\tilde{\Phi}$	1	2	3	4	5	6	7	8	9	10	
1	-0.095	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	-0.448	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	-0.032	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.258	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	-0.056	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	-0.312	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	-0.208	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.476	0.000	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.321	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.208	0.000
11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.379

Table 15: $\tilde{\Lambda}_0$ of DGP_4

$\tilde{\Lambda}_0$	1	2	3	4	5	6	7	8	9	10	
1	0.063	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.323	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	-0.494	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.104	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.012	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	-0.393	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	-0.218	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.183	0.000	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.172	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.219	0.000
11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.314

Table 16: $\tilde{\Lambda}_1$ of DGP_4

$\tilde{\Lambda}_1$	1	2	3	4	5	6	7	8	9	10	
1	-0.355	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.162	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.396	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	-0.097	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.158	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	-0.133	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	0.304	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.422	0.000	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.432	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.439	0.000
11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.274

Table 17: W correct matrix for DGP_4

	1	2	3	4	5	6	7	8	9	10	
1	0.000	0.122	0.002	0.223	0.055	0.004	0.228	0.154	0.153	0.002	0.055
2	0.088	0.000	0.031	0.000	0.142	0.006	0.166	0.164	0.134	0.161	0.107
3	0.086	0.152	0.000	0.037	0.069	0.129	0.144	0.018	0.040	0.122	0.202
4	0.177	0.185	0.089	0.000	0.013	0.177	0.043	0.120	0.073	0.053	0.070
5	0.039	0.147	0.092	0.087	0.000	0.197	0.043	0.113	0.112	0.078	0.091
6	0.099	0.079	0.155	0.121	0.148	0.000	0.126	0.153	0.033	0.040	0.047
7	0.181	0.108	0.040	0.120	0.005	0.125	0.000	0.022	0.166	0.118	0.116
8	0.133	0.152	0.031	0.131	0.063	0.133	0.117	0.000	0.179	0.030	0.031
9	0.059	0.095	0.074	0.114	0.189	0.069	0.055	0.158	0.000	0.027	0.162
10	0.194	0.066	0.156	0.198	0.133	0.063	0.017	0.080	0.035	0.000	0.057
11	0.018	0.152	0.010	0.165	0.135	0.003	0.208	0.159	0.060	0.090	0.000

Table 18: W misspecified matrix for DGP_4

	1	2	3	4	5	6	7	8	9	10	
1	0.000	0.006	0.190	0.016	0.144	0.034	0.165	0.168	0.110	0.109	0.059
2	0.035	0.000	0.054	0.094	0.180	0.190	0.044	0.183	0.150	0.042	0.029
3	0.110	0.150	0.000	0.139	0.122	0.066	0.088	0.147	0.051	0.044	0.083
4	0.173	0.034	0.064	0.000	0.120	0.147	0.069	0.198	0.143	0.043	0.010
5	0.149	0.007	0.001	0.162	0.000	0.174	0.127	0.090	0.143	0.011	0.135
6	0.196	0.078	0.062	0.110	0.057	0.000	0.058	0.201	0.067	0.054	0.116
7	0.128	0.133	0.070	0.052	0.004	0.176	0.000	0.033	0.110	0.081	0.211
8	0.038	0.153	0.119	0.060	0.059	0.105	0.120	0.000	0.150	0.158	0.040
9	0.006	0.124	0.115	0.073	0.119	0.093	0.079	0.083	0.000	0.155	0.153
10	0.035	0.152	0.125	0.195	0.024	0.179	0.104	0.046	0.112	0.000	0.028
11	0.080	0.123	0.029	0.008	0.117	0.149	0.151	0.152	0.067	0.124	0.000

2 The W Matrices

Table 19: **Estimated W matrix using Gross (2019) procedure.** This table reports the estimated W matrix estimated 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.

	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

Table 20: W **matrix using T-value related weights, negative.** 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.

	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

Table 21: W **matrix using T-value related weights, positive.** 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.

	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

Table 22: W **matrix using FIFA related weights.** This table reports the 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.

	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

3 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

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

   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 as follows (abstaining here from the deterministic components)

$$\Delta Y_t = \hat{B} Y_{t-1} + \hat{e}_t, \tag{2}$$

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

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 outlined 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 $\tilde{Y}_{it}$ using the first actual observations as starting values for the different series. The subsequent bootstrapped observations are

computed as:

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

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

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 α% 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 α%.

References

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

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

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *The Annals of Statistics 21*(1), 255–285.