Central European Journal of Economic Modelling and Econometrics

GVAR: A Case of Spurious Cross-Sectional Cointegration

Piotr Kębłowski*

Submitted: 21.08.2020, Accepted: 30.01.2021

Abstract

Global Vector Autoregressive models came to be used quite widely in empirical studies using macroeconomic non-stationary panel data for the global economy. In this paper, it is shown that when the loading matrix of the cointegrating vectors is not block-diagonal and the cross-sectional spillovers of disequilibrium exist, the use of the GVAR model leads to spurious cross-sectional long-run relationships. Moreover, the results of Monte Carlo simulation show that the GVAR model is outperformed by other valid econometric approaches in terms of the maximum likelihood estimator of long-run coefficients, when the cointegrating vectors matrix is block-diagonal.

Keywords: global VAR, GVAR, panel VAR, PVAR, spurious cross-sectional cointegration

JEL Classification: C10, C18, C33

*University of Łódź, Poland; e-mail: piotr.keblowski@uni.lodz.pl; ORCID: 0000-0002-6671-0756

1 Introduction

The expanding globalization of the world economy strengthens global value chains, international trade, as well as financial market integration. This process inevitably provides a strong rationale for the use of cross-sectional statistical frameworks in analyses of modern macroeconomic phenomena. The increased globalization coincides with dissemination of high-quality consistent panel data that are provided at quarterly or even monthly basis and for long time spans, and are widely used in international comparisons of different countries or regions. This invariably means that the panel counterparts of techniques used for non-stationary time-series data can be applied for empirical datasets, in order to disentangle the spatiotemporal behaviour of economic phenomena. For example, Bussière et al. (2009) model global trade flows in panel of emerging markets and advanced economies, Favero (2013) and Temizsoy and Rojas (2019) investigate determinants of credit rating spreads, whereas Bi and Anwar (2017) examine the impact of US monetary policy shocks on China's economy. The common feature of above mentioned articles is the use of the global VAR model as a statistical framework of analyses. A comprehensive survey of applications associated with global VAR provide Chudik and Pesaran (2016).

The growing interest over the last two decades for analysing multi-country or multi-region data led to development of models that (i) do not impose cross-sectional short- and/or long-run homogeneity, (ii) allow for various types of spatiotemporal interdependencies, (iii) and are able to mimic country- or region-specific responses to shocks that affect the world economy. There are two main strands how to deal with macroeconomic panels, which can be discriminated on the basis of contemporary state-of-the-art macroeconomic-oriented panel models, and which form alternatives to spatial panel data models that simply employ the spatial weights matrix.

On the one hand, the error factor structures and the common correlated effects estimator proposed by Pesaran (2006) or the interactive fixed effects and the principal components estimator proposed by Bai (2009) can be used, primarily when the variables under study are generated by stationary processes. On the other hand, if the processes generating macroeconomic categories are non-stationary and more flexibility is necessary regarding spatiotemporal interdependencies, then the natural candidate for the model embodying the admissible statistical frameworks would be the large-scale VAR model. Since the large-scale VAR model is infeasible in typical empirical cases, it is necessary to impose some panel structure on the large-scale VAR in order to gain feasibility.

In view of the need for the feasible large-scale VAR model with panel structure, Pesaran et al. (2004) propose to utilize the international trade patterns as well as to use a set of conditional country-specific models, by imposing weak-exogeneity of cross-sectional sub-processes. This essentially means that the cross-sectional loading matrix of the cointegrating vectors is block-diagonal and the cross-unit spillovers of disequilibrium are not allowed, as opposed to the cointegrating vectors matrix, where the cross-unit cointegration vectors are accepted, even though they are restricted by the international trade shares. The other solution has been proposed by Larsson and Lyhagen (2007), who suggest to assume that the cointegrating vectors matrix is block-diagonal rather than to restrict the loading matrix.

Both approaches, the global VAR model proposed by Pesaran et al. (2004) as well as the panel VAR model introduced by Larsson and Lyhagen (2007) allow for the instantaneous and the short-run cross-unit interdependencies, but they clearly differ in terms of sources of the long-run cross-sectional dependencies. Definitely, both the cross-sectional cointegrating vectors and the cross-unit spillovers of disequilibrium can exist in general. However, of the two possibilities only one can be excluded a priori for some empirical phenomena on the basis of some established economic theory, i.e. the cross-sectional cointegrating vectors. Moreover, even if the cross-unit cointegrating vectors are predicted by the economic theory due to some international parities, there is usually possibility to transform the problem (real exchange rate instead of nominal, interest rates differentials, e.g.). In case of the cross-sectional spillovers of disequilibrium the economic theory remains blind and any a priori restrictions may be incorrect.

The aforementioned problem calls for an investigation into effects of conditioning model by erroneously assuming weak-exogeneity of cross-sectional sub-processes. To this aim two issues are explored. Firstly, consequences of the use of the GVAR model are examined when the cross-sectional spillovers of disequilibrium exist. Secondly, performance of the maximum likelihood estimator (MLE) of long-run coefficients is studied using the GVAR model, the PVAR model, and the individual VAR models. Therefore, three statistical frameworks that are embodied within the large-scale VAR model are considered, whereas all models that assume cross-sectional short- and/or long-run homogeneity or common factors are excluded.

The remainder of this paper is organised as follows. First, it introduces three statistical settings nested by the large-scale VAR model: the GVAR model, the PVAR model, as well as the individual VAR models applied for each country or region independently. Secondly, it describes a Monte Carlo simulations and the results. The paper closes with concluding remarks.

2 Statistical framework

Firstly, consider the finite-order VAR model:

$$\mathbf{y}_{it} = \sum_{k=1}^{K_i} \mathbf{\Pi}_{ki}\mathbf{y}_{i,t-k} + \mathbf{\Phi}_i\mathbf{d}_t + \boldsymbol{\nu}_{it}, \quad i = 1, 2, \ldots, I, \quad t = 1, 2, \ldots, T, \tag{1}$$

where $\mathbf{y}_{it} = \begin{bmatrix} y_{1it} & y_{2it} & \ldots & y_{Pit} \end{bmatrix}'$ is a $P \times 1$ vector of variables for cross-section i and period t, K_i denotes a cross-section-specific lag order, $\mathbf{\Pi}_{ki}$ is a $P \times P$ matrix of lagged coefficients, $\mathbf{d}_t$ and $\mathbf{\Phi}_i$ stand for a $N \times 1$ vector of deterministic components and $P \times N$ matrix of their coefficients, and $\boldsymbol{\nu}_{it}$ is an $P \times 1$ idiosyncratic error term with zero mean and covariance matrix $\mathbf{\Omega}_i$. Obviously, model (1) represents a pure time-series approach, which severely limits its use for empirical modelling of cross-sectional data.

Next, consider the large-scale VAR model:

$$\mathbf{y}_t = \sum_{k=1}^{K} \mathbf{\Pi}_k \mathbf{y}_{t-k} + \mathbf{\Phi}\mathbf{d}_t + \boldsymbol{\varepsilon}_t, \tag{2}$$

where $\mathbf{y}_t = \begin{bmatrix} \mathbf{y}'_{1t} & \mathbf{y}'_{2t} & \ldots & \mathbf{y}'_{It} \end{bmatrix}'$ is a $IP \times 1$ vector of stacked panel data for period t, $K = \max(K_i)$, $\mathbf{\Pi}_k$ is a $IP \times IP$ matrix of lagged coefficients, $\mathbf{\Phi}$ denotes an $IP \times N$ matrix of deterministic term coefficients and $\boldsymbol{\varepsilon}_t$ is an $IP \times 1$ independently and identically distributed error term with zero mean and covariance matrix $\mathbf{\Omega}$. Clearly, model (2) still constitutes *de facto* a time-series approach, however a wide bunch of short- and long-run cross-sectional dependencies is here allowed. In order to make it easily visible let the error-correction form of model (2) to be explicitly written, in the compact form:

$$\Delta\mathbf{y}_t = \mathbf{\Pi}\mathbf{y}_{t-1} + \sum_{k=1}^{K-1} \mathbf{\Gamma}_k \Delta\mathbf{y}_{t-k} + \mathbf{\Phi}\mathbf{d}_t + \boldsymbol{\varepsilon}_t, \tag{3a}$$

where $\mathbf{\Pi}$ and $\mathbf{\Gamma}_k$ are $IP \times IP$ matrices of long- and short-run multipliers respectively, or in the expanded form:

$$\Delta\mathbf{y}_t = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1I} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{I1} & \mathbf{A}_{I2} & \cdots & \mathbf{A}_{II} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \cdots & \mathbf{B}_{1I} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \cdots & \mathbf{B}_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_{I1} & \mathbf{B}_{I2} & \cdots & \mathbf{B}_{II} \end{bmatrix}' \mathbf{y}_{t-1} +$$

$$+ \sum_{k=1}^{K-1} \begin{bmatrix} \mathbf{\Gamma}_{11,k} & \mathbf{\Gamma}_{12,k} & \cdots & \mathbf{\Gamma}_{1I,k} \\ \mathbf{\Gamma}_{21,k} & \mathbf{\Gamma}_{22,k} & \cdots & \mathbf{\Gamma}_{2I,k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Gamma}_{I1,k} & \mathbf{\Gamma}_{I2,k} & \cdots & \mathbf{\Gamma}_{II,k} \end{bmatrix} \Delta\mathbf{y}_{t-k} + \mathbf{\Phi}\mathbf{d}_t + \boldsymbol{\varepsilon}_t, \tag{3b}$$

where $\mathrm{var}(\boldsymbol{\varepsilon}_t) = \mathbf{\Omega} = \begin{bmatrix} \mathbf{\Omega}_{11} & \mathbf{\Omega}_{12} & \cdots & \mathbf{\Omega}_{1I} \\ \mathbf{\Omega}_{21} & \mathbf{\Omega}_{22} & \cdots & \mathbf{\Omega}_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Omega}_{I1} & \mathbf{\Omega}_{I2} & \cdots & \mathbf{\Omega}_{II} \end{bmatrix}$, $\mathbf{A}_{ij}$ and $\mathbf{B}_{ij}$ are $P \times R_i$ matrices of full rank and R_i denotes cointegration rank for cross-section i. It is straightforward to note that the model (3a)/(3b) allow for cross-sectional dependencies: (i) in the error term, since $\mathrm{cov}(\varepsilon_{it}, \varepsilon_{jt}) = \mathbf{\Omega}_{ij} \neq \mathbf{0}$, (ii) in the short-run co-movements, as quantified by the $\mathbf{\Gamma}_{ij,k}$ matrices, (iii) in the long-run relationships, whose coefficients are expressed by the cointegrating vectors matrix $\mathbf{B}$, (iv) and in the spillovers of disequilibrium, which are measured by the loading matrix $\mathbf{A}$.

Although the model (3a)/(3b) is a flexible form for panels generated by nonstationary (integrated of order one) processes, it becomes infeasible in practice. Since the short-run effects can be concentrated out according to the Frisch-Waugh theorem, the main problem to gain feasibility is how to restrict the long-run coefficients when the long-run cross-sectional co-movements are observed.

One solution has been proposed by Larsson and Lyhagen (2007), who suggest to assume that the $\mathbf{B}$ matrix in the model (3a)/(3b) is block-diagonal, which precludes any cross-sectional long-run relationships, whereas the loading matrix $\mathbf{A}$ is left unrestricted, which means that the cross-sectional spillovers of disequilibrium are allowed. Therefore the model proposed by Larsson and Lyhagen (2007), panel vector autoregression (PVAR), is given by:

$$\Delta\mathbf{y}_t = \mathbf{A}\begin{bmatrix}\mathbf{B}_{11} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{B}_{22} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_{II}\end{bmatrix}'\mathbf{y}_{t-1} + \sum_{k=1}^{K-1}\mathbf{\Gamma}_k\Delta\mathbf{y}_{t-k} + \mathbf{\Phi}\mathbf{d}_t + \boldsymbol{\varepsilon}_t. \tag{4}$$

The model (4) can be estimated by means of the MLE and the switching algorithm proposed by Johansen (1991). This approach can be effectively applied for moderate and large samples ($T > 100$), when the dimension of $\mathbf{y}_t$ (IP) is below 40, see Kębłowski (2016). Moreover, further efficiency gains are possible in small samples when the homogenous cointegrating vectors can be assumed, i.e. that the cointegrating vectors in each cross-section span the same space.

The other possibility is to restrict the loading matrix to be block-diagonal. The most notable model meeting this requirement is the global VAR (GVAR) model proposed by Pesaran et al. (2004) and further developed by Dees et al. (2007). The main proposition is not only to assume that the processes generating individual cross-sections (regions, countries) are weakly-exogenous, which enables in itself the use of conditional model, but also to take advantage of geographical patterns of international trade (trade shares) in order to create a parsimonious model. The latter is simply a direct borrowing from spatial panel data models that use the spatial weights matrix $\mathbf{W}$.

Therefore, the GVAR model is essentially a set of individual conditional models:

$$\mathbf{y}_{it} = \sum_{k=1}^{K_i}\mathbf{\Pi}_{ki}\mathbf{y}_{i,t-k} + \sum_{m=0}^{M_i}\mathbf{\Lambda}_{mi}\mathbf{x}_{i,t-m} + \mathbf{\Phi}_i\mathbf{d}_t + \boldsymbol{\upsilon}_{it}, \tag{5}$$

where $\begin{bmatrix}\mathbf{y}'_{it} & \mathbf{x}'_{it}\end{bmatrix}' = \mathbf{W}_i\mathbf{y}_t$, $\mathbf{x}_{it} = \begin{bmatrix}x_{1it} & x_{2it} & \dots & x_{Qit}\end{bmatrix}'$ is a $Q \times 1$ vector of country-specific foreign variables that are created by country-specific weights w_{ij}, i.e. $\mathbf{x}_{it} = \sum_{j=0}^{I} w_{ij}\mathbf{y}_{jt}$, weights can be based on trade shares or capital flows and meet the requirements: $\underset{i}{\forall} \sum_{j=0}^{I} w_{ij} = 1$, $\underset{i}{\forall} w_{ii} = 0$, and $\mathbf{\Lambda}_{mi}$ is a $P \times Q$ matrix of coefficients associated with foreign variables. Denoting $\mathbf{W}_i \mathbf{y}_t = \mathbf{z}_{it}$ and $K = \max(K_i, M_i)$ the model (5) can be rewritten in the VAR form:

$$\mathbf{G}_{0i}\mathbf{z}_{it} = \sum_{k=1}^{K} \mathbf{G}_{ki}\mathbf{z}_{i,t-k} + \mathbf{\Phi}_i \mathbf{d}_t + \boldsymbol{\upsilon}_{it}, \tag{6}$$

where $\mathbf{G}_{0i} = \begin{bmatrix} \mathbf{I}_{K_i} & -\mathbf{\Lambda}_{0i} \end{bmatrix}$ and $\underset{0<k\leq K}{\forall} \mathbf{G}_{ki} = \begin{bmatrix} \mathbf{\Pi}_{ki} & \mathbf{\Lambda}_{ki} \end{bmatrix}$.

Let us now consider the cointegration restriction within the model proposed by Pesaran et al. (2004). The error-correction form of model (5) is:

$$\Delta\mathbf{y}_{it} = \mathbf{\Pi}_i \mathbf{y}_{i,t-1} + \sum_{k=1}^{K_i-1} \mathbf{\Gamma}_{ki}\Delta\mathbf{y}_{i,t-k} + \mathbf{\Lambda}_i \mathbf{x}_{i,t-1} + \sum_{m=0}^{M_i-1} \mathbf{T}_{mi}\Delta\mathbf{x}_{i,t-m} + \mathbf{\Phi}_i \mathbf{d}_t + \boldsymbol{\upsilon}_{it}, \tag{7a}$$

which can be shortly written as

$$\Delta\mathbf{y}_{it} = \widehat{\mathbf{\Pi}}_i \mathbf{z}_{i,t-1} + \sum_{k=1}^{K-1} \widehat{\mathbf{\Gamma}}_{ki}\Delta\mathbf{z}_{i,t-k} + \mathbf{T}_{0i}\Delta\mathbf{x}_{i,t} + \mathbf{\Phi}_i \mathbf{d}_t + \boldsymbol{\upsilon}_{it}, \tag{7b}$$

where $\widehat{\mathbf{\Pi}}_i = \begin{bmatrix} \mathbf{\Pi}_i & \mathbf{\Lambda}_i \end{bmatrix}$, $\widehat{\mathbf{\Gamma}}_{ki} = \begin{bmatrix} \mathbf{\Gamma}_{ki} & \mathbf{T}_{ki} \end{bmatrix}$, and $\widehat{\mathbf{\Pi}}_i$ and $\widehat{\mathbf{\Gamma}}_{ki}$ are $P \times (P+Q)$ matrices of long- and short-run multipliers respectively. If the reduced rank restriction holds, then, obviously, the long-run multipliers matrix can be decomposed as $\widehat{\mathbf{\Pi}}_i = \mathbf{A}_i {\mathbf{B}_i}'$, where the $P_i \times R_i$ loading matrix $\mathbf{A}_i$ and the $(P_i + Q_i) \times R_i$ cointegrating matrix $\mathbf{B}_i$ are of full rank.

In order to see how the large-scale model (2) encompasses the GVAR model proposed by Pesaran et al. (2004), let us consider the VAR form (6), which can be written compactly as:

$$\mathbf{G}_0 \mathbf{y}_t = \sum_{k=1}^{K} \mathbf{G}_k \mathbf{y}_{t-k} + \mathbf{\Phi}\mathbf{d}_t + \boldsymbol{\upsilon}_t, \tag{8}$$

where $\underset{0\leq k\leq K}{\forall} \mathbf{G}_k = \begin{bmatrix} \mathbf{W}_1'\mathbf{G}_{k1}' & \mathbf{W}_2'\mathbf{G}_{k2}' & \ldots & \mathbf{W}_I'\mathbf{G}_{kI}' \end{bmatrix}'$, $\mathbf{\Phi} = \begin{bmatrix} \mathbf{\Phi}_1' & \mathbf{\Phi}_2' & \ldots & \mathbf{\Phi}_I' \end{bmatrix}'$, $\boldsymbol{\upsilon}_t = \begin{bmatrix} \boldsymbol{\upsilon}_{1t}' & \boldsymbol{\upsilon}_{2t}' & \ldots & \boldsymbol{\upsilon}_{It}' \end{bmatrix}'$, $\mathrm{var}(\boldsymbol{\upsilon}_t) = \mathbf{\Sigma} = \mathrm{diag}(\mathbf{\Sigma}_{11}, \ \ldots, \ \mathbf{\Sigma}_{II})$, so $\mathrm{cov}(\boldsymbol{\upsilon}_{it}, \boldsymbol{\upsilon}_{jt}) = \mathbf{\Sigma}_{ij} = \mathbf{0}$. The VAR solution of (8) is:

$$\mathbf{y}_t = \sum_{k=1}^{K} \mathbf{G}_0^{-1}\mathbf{G}_k \mathbf{y}_{t-k} + \mathbf{G}_0^{-1}\mathbf{\Phi}\mathbf{d}_t + \mathbf{G}_0^{-1}\boldsymbol{\upsilon}_t = \sum_{k=1}^{K} \mathbf{F}_k \mathbf{y}_{t-k} + \widetilde{\mathbf{\Phi}}\mathbf{d}_t + \zeta_t, \tag{9}$$

and $\mathrm{var}(\zeta_t) = \mathbf{G}_0^{-1}\boldsymbol{\upsilon}_t\boldsymbol{\upsilon}_t'{\mathbf{G}_0^{-1}}' = \mathbf{G}_0^{-1}\mathbf{\Sigma}{\mathbf{G}_0^{-1}}'$. Therefore, the GVAR model is a heavily restricted form of the general large-scale VAR model (2).

Finally, let us consider the error-correction form of model (8):

$$\mathbf{G}_0\Delta\mathbf{y}_t = \mathbf{H}\mathbf{y}_{t-1} + \sum_{k=1}^{K-1}\widetilde{\mathbf{G}}_k\Delta\mathbf{y}_{t-k} + \boldsymbol{\Phi}\mathbf{d}_t + \boldsymbol{\upsilon}_t, \tag{10}$$

where $\mathbf{H} = -\mathbf{G}_0 + \sum_{j=1}^{K}\mathbf{G}_k$ and $\underset{0<k<K-1}{\forall}\widetilde{\mathbf{G}}_k = -\sum_{j=k+1}^{K}\mathbf{G}_k$. The long-run multipliers matrix can be decomposed as

$$\mathbf{H} = \begin{bmatrix} \left(-\mathbf{G}_{01} + \mathbf{G}_{11} + \ldots \mathbf{G}_{K1}\right)\mathbf{W}_1 \\ \left(-\mathbf{G}_{02} + \mathbf{G}_{12} + \ldots \mathbf{G}_{K2}\right)\mathbf{W}_2 \\ \vdots \\ \left(-\mathbf{G}_{0I} + \mathbf{G}_{1I} + \ldots \mathbf{G}_{KI}\right)\mathbf{W}_I \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1\mathbf{B}_1{}'\mathbf{W}_1 \\ \mathbf{A}_2\mathbf{B}_2{}'\mathbf{W}_2 \\ \vdots \\ \mathbf{A}_I\mathbf{B}_I{}'\mathbf{W}_I \end{bmatrix} = \mathbf{A}\widetilde{\mathbf{B}}',$$

where $\mathbf{A} = \operatorname{diag}\left(\mathbf{A}_1,\ \mathbf{A}_2,\ \ldots,\ \mathbf{A}_I\right)$ and $\widetilde{\mathbf{B}} = \begin{bmatrix}\mathbf{W}_1{}'\mathbf{B}_1 & \mathbf{W}_2{}'\mathbf{B}_2 & \ldots & \mathbf{W}_I{}'\mathbf{B}_I\end{bmatrix}$.

3 Monte Carlo simulation

In this section, we investigate two issues. Firstly, we focus on main consequences of wrongly assuming that non-stationary sub-processes generating individual cross-sections (regions, countries) are weakly-exogenous for the GVAR framework. Secondly, we examine performance of the maximum likelihood estimator of long-run parameters in GVAR, in PVAR, and in individual VAR models, when the cross-sectional spillovers of disequilibrium exist, but the long-run relationships are restricted to individual cross-sections.

In order to address the aforementioned questions, we use the Monte Carlo simulation technique and the data generating process (DGP) is defined as follows:

$$\Delta\mathbf{y}_t = \mathbf{A}\left(\mathbf{I}_I \otimes \mathbf{B}_{ii}'\right)\begin{bmatrix}\mathbf{y}_{t-1}{}' & \mathbf{j}'\end{bmatrix}' + \sum_{k=1}^{K-1}\boldsymbol{\Gamma}_k\Delta\mathbf{y}_{t-k} + \boldsymbol{\varepsilon}_t, \tag{11}$$

where $\mathbf{j}$ stands for $IP \times 1$ all-ones vector, the number of variables for each cross-section is $P = 5$, the number of long-run relationships for each cross-section is $R = 2$ and we use a second-order VAR – $K = 2$. The error-term is generated by the multivariate normal distribution, $\boldsymbol{\varepsilon}_t \sim N_{IP}\left(\mathbf{0};\boldsymbol{\Omega}\right)$, the covariance matrix is distributed by the inverse Wishart distribution, $\boldsymbol{\Omega} \sim W_{IP}^{-1}\left(\mathbf{I}\cdot\left(df - I\cdot P - 1\right); df\right)$, with number of degrees of freedom $df = I \cdot P + 40$, which ensures that the expected value of diagonal elements of the covariance matrix is unity, $\underset{i=j}{\forall} E\left(\omega_{ij}\right) = 1$, whereas the expected absolute value of non-diagonal elements takes a constant value irrespective of I and P, $\underset{i\neq j}{\forall} E\left(|\omega_{ij}|\right) = 0.125$, which was investigated by means of Monte Carlo simulation. The short-run effects are defined as $\underset{i=j}{\forall}\gamma_{ij,1} = 0.5$ and $\underset{i\neq j}{\forall}\gamma_{ij,1} \sim U(-0.1;0.1)$, the coefficients of the cointegrating vectors matrix are $\underset{i=1,\dots,I}{\forall}\mathbf{B}_{ii} = \begin{bmatrix}1 & -1 & 0 & 0 & 0 & 1\\ 0 & 0 & 1 & -1 & 0 & 1\end{bmatrix}'$ and $\underset{i\neq j}{\forall}\mathbf{B}_{ij} = \mathbf{0}$, and the loadings are $\underset{i=j}{\forall}\mathbf{A}_{ij} = \begin{bmatrix}-0.5 & 0 & 0 & 0 & 0\\ 0 & 0 & -0.5 & 0 & 0\end{bmatrix}'$ and $\underset{i\neq j}{\forall}\mathbf{A}_{ij} = \begin{bmatrix}-0.1 & 0 & 0 & 0 & 0\\ 0 & 0 & -0.1 & 0 & 0\end{bmatrix}'$. The only deterministic term is a constant restricted to the cointegration space. The roots of the autoregressive polynomial are examined in each case in order to exclude the explosive roots. The number of replications equals 100000.

The mentioned above settings of the DGP are not intended to closely mimic any specific empirical example, however, the GVAR model presented in Dees et al. (2007) has 4 to 6 domestic variables and 1 to 4 cointegrating vectors for each cross-section. With respect to our DGP, we have the same number of variables for each cross-section and we have set equal country-specific weights for the link matrices $\mathbf{W}_i$:

$$\underset{i}{\forall}\quad \mathbf{W}_i = \begin{bmatrix} 0 & 1/(I-1) & \cdots & 1/(I-1)\\ 1/(I-1) & 0 & \cdots & 1/(I-1)\\ \vdots & \vdots & \ddots & \vdots\\ 1/(I-1) & 1/(I-1) & \cdots & 0 \end{bmatrix}.$$

Table 1: Frequency of spurious cross-sectional cointegration in GVAR model (significance of a "foreign" variable in a cointegrating vector), restricted MLE

$T\backslash I$	2	3	4	5	6	7	8	9	10
100	0.459	0.530	0.568	0.591	0.605	0.618	0.628	0.640	0.654
200	0.561	0.646	0.687	0.705	0.716	0.722	0.725	0.728	0.733
400	0.686	0.765	0.797	0.810	0.814	0.814	0.812	0.809	0.808
800	0.786	0.849	0.870	0.877	0.878	0.876	0.871	0.866	0.864
2000	0.870	0.913	0.925	0.928	0.929	0.926	0.921	0.917	0.914

Since the cross-sectional cointegrating vectors matrix $\mathbf{B}$ in the DGP is block-diagonal, the cross-unit cointegrating vectors are excluded. However, as reported in Table 1, probability of a given foreign variable to be found as significant (using t-ratio) for a given cointegrating vector exceeds by far the size of the test (5%). In fact, the larger the sample size T and the number of cross-sections I are, the higher the probability of spurious cross-sectional cointegration is. Considering a panel covering $I = 8$ cross-sections and $T = 100$ the probability of a given foreign variable to be spuriously found as significant in a given cointegrating vector is as high as almost 63%. Obviously, the probability that any foreign variable will be found as significant in a given cointegrating vector is even higher.

Table 2: Frequency of spurious cross-sectional cointegration in GVAR model (significance of any “foreign” variable in any cointegrating vector), restricted MLE

$T\backslash I$	2	3	4	≥ 5
100	0.884	0.987	0.999	1
200	0.919	0.996	1	1
400	0.959	1	1	1
800	0.985	1	1	1
2000	0.995	1	1	1

It is worth mentioning that although the t-ratios are widely used for significance testing in the context of coefficients of long-run relationships, the alternative approach would be to employ distance measures between the space spanned by cointegration vectors in the DGP and the estimated one, see e.g. Larsson and Villani (2001).

Table 2 gives frequencies of spurious cross-sectional cointegration for the GVAR, which is defined as the probability of spurious significance of any foreign variable in any cointegrating vector, i.e. that the cointegration space related to given cross-section is spanned by coefficient(s) related to distinct cross-sections. It can be easily noticed that this frequency is almost equal to unity, when the number of cross-sections exceeds 2. Moreover, the results presented in Tables 1 and 2 do not alter significantly, if the unrestricted constant is used in the DGP, i.e. the linear trend is allowed in the data, as well as, when the country-specific weights for the link matrices $\mathbf{W}_i$ are uniformly distributed, which relaxes the assumption of equally-sized economies of regions/countries (additional results available upon request).

The phenomenon of spurious cross-sectional cointegration that occurs within the GVAR framework under cross-sectional spillovers of disequilibrium necessitates a very cautious use of the GVAR model in modelling of regional interdependencies in the global economy when data under study are generated by nonstationary processes. Moreover, it seems also very likely that this phenomenon may affect various tests of weak-exogeneity, which are applied after identification of (spuriously cross-sectional) cointegrating vectors within the GVAR framework.

The next issue to be investigated is performance of the maximum likelihood estimator of the cointegrating vectors matrix $\mathbf{B}$ in GVAR, in PVAR, and in individual VAR models, when the cross-sectional spillovers of disequilibrium exist, i.e. the loading matrix $\mathbf{A}$ is unrestricted, but the $\mathbf{B}$ matrix is block-diagonal. Therefore, we have the misspecified GVAR, the PVAR with the MLE that is efficient (in the limit), as well as a set of individual VAR models with the MLE that is inefficient. The restricted MLE for the VAR is calculated as proposed by Johansen (1991), the restricted MLE for the GVAR is derived as in Harbo et al. (1998), and the restricted MLE for the PVAR is computed as in Larsson and Lyhagen (2007).

Table 3 presents standard deviation of the restricted MLE of $\mathbf{B}$ using the aforementioned models for $I \in \{1, 2, \ldots, 10\}$ and $T \in \{100,\ 150,\ 200,\ 400,\ 800\}$.

Table 3: Standard deviation of the restricted MLE of cointegrating vectors

I	1	2	3	4	5	6	7	8	9	10
$T = 100$										
VAR	0.056	0.084	0.106	0.124	0.143	0.168	0.207	0.307	0.449	0.661
GVAR	0.056	0.116	0.150	0.174	0.210	0.258	0.314	0.402	0.536	0.718
PVAR	0.056	0.087	0.152	0.289	0.591	1.412	-	-	-	-
$T = 150$										
VAR	0.034	0.051	0.063	0.072	0.079	0.085	0.096	0.143	0.227	0.304
GVAR	0.034	0.059	0.072	0.082	0.093	0.103	0.117	0.147	0.202	0.364
PVAR	0.034	0.047	0.057	0.067	0.077	0.091	0.204	0.365	-	-
$T = 200$										
VAR	0.024	0.036	0.045	0.051	0.056	0.060	0.067	0.084	0.154	0.259
GVAR	0.024	0.041	0.049	0.056	0.062	0.068	0.077	0.093	0.128	0.189
PVAR	0.024	0.033	0.038	0.042	0.043	0.045	0.046	0.049	0.063	0.115
$T = 400$										
VAR	0.012	0.017	0.021	0.024	0.026	0.027	0.030	0.038	0.064	0.137
GVAR	0.012	0.019	0.022	0.025	0.027	0.029	0.032	0.037	0.055	0.083
PVAR	0.012	0.015	0.016	0.017	0.016	0.015	0.014	0.012	0.011	0.010
$T = 800$										
VAR	0.006	0.008	0.010	0.011	0.012	0.013	0.014	0.018	0.030	0.069
GVAR	0.006	0.009	0.010	0.012	0.013	0.014	0.015	0.017	0.022	0.049
PVAR	0.006	0.007	0.008	0.008	0.007	0.006	0.006	0.005	0.004	0.004

It can be easily noticed that the PVAR model outperforms the other frameworks in terms of the standard deviation of the restricted MLE of the **B** matrix, except for $T = 100$ as well as $T = 150$ and $I \geq 6$, where the dimensionality effect prevails over efficiency gains of the PVAR model. Considering a panel covering 8 cross-sections and 200 observations for each cross-section, the standard deviation of the restricted MLE of **B** for the PVAR model equals 0.049, whereas for the VAR and the GVAR model it is almost twice as large and equals 0.084 and 0.093 respectively. For large samples counting 400 observations and more, which are, however, quite unusual in typical macro-panels used in practice, and 8 cross-sections standard deviation of the restricted MLE for the VAR and the GVAR model increases even threefold as compared to the respective value for the PVAR model.

Since Pesaran et al. (2004) and Dees et al. (2007) use the unrestricted MLE for estimating their GVAR models, we should consider performance of the unrestricted MLE as well. In order to enable performance comparison, the cointegration vectors are transformed into linear combinations that are as close as possible to subspaces spanned by respective design matrices, as proposed by Johansen and Juselius (1994), see also Kębłowski (2016).

The results in Table 4 reveal that the PVAR model is no longer the best in terms of standard deviation of the unrestricted MLE of **B**, as the PVAR model becomes overparametrized with respect to the number of long-run parameters to be estimated.

Table 4: Standard deviation of the unrestricted MLE of cointegrating vectors

I	1	2	3	4	5	6	7	8	9	10
$T = 100$										
VAR	0.091	0.148	0.199	0.251	0.336	0.430	0.599	0.845	1.148	1.560
GVAR	0.091	0.305	0.402	0.506	0.642	0.801	0.997	1.226	1.515	1.811
PVAR	0.091	0.370	0.808	1.461	2.600	4.352	-	-	-	-
$T = 150$										
VAR	0.053	0.081	0.104	0.126	0.148	0.175	0.260	0.393	0.602	0.923
GVAR	0.053	0.122	0.154	0.189	0.231	0.274	0.370	0.470	0.610	0.802
PVAR	0.053	0.139	0.319	0.543	0.807	1.239	1.965	3.043	-	-
$T = 200$										
VAR	0.037	0.057	0.073	0.087	0.101	0.122	0.152	0.227	0.383	0.673
GVAR	0.037	0.080	0.101	0.122	0.145	0.173	0.211	0.270	0.345	0.485
PVAR	0.037	0.089	0.171	0.303	0.450	0.616	0.835	1.165	1.664	2.408
$T = 400$										
VAR	0.017	0.026	0.033	0.039	0.046	0.054	0.068	0.099	0.183	0.351
GVAR	0.017	0.034	0.042	0.051	0.060	0.072	0.088	0.109	0.141	0.193
PVAR	0.017	0.037	0.062	0.094	0.133	0.183	0.244	0.314	0.384	0.455
$T = 800$										
VAR	0.008	0.012	0.016	0.019	0.022	0.026	0.034	0.050	0.095	0.208
GVAR	0.008	0.016	0.020	0.023	0.028	0.033	0.041	0.053	0.070	0.100
PVAR	0.008	0.017	0.027	0.039	0.053	0.069	0.087	0.106	0.128	0.152

Considering a panel covering 8 cross-sections, the smallest standard deviation of the unrestricted MLE is associated with the individual VAR models. Therefore, the GVAR model is still outperformed by other models, except for the case of $I \geq 9$ and $T \geq 200$.

4 Conclusions

Modelling macroeconomic non-stationary panel data and regional interdependencies in the global economy constitute one of a main strands in the contemporaneous macroeconomic research. In this paper we show that the use of the global VAR model leads to spurious cross-sectional long-run relationships when the cross-sectional spillovers of disequilibrium exist. This phenomenon cannot be considered as a striking feature of the GVAR model, since it is just a straightforward consequence of restrictions imposed for rank factorization of the long-run multipliers matrix. If the long-run multipliers matrix is not block-diagonal and an erroneous assumption of weak-exogeneity of cross-sectional sub-processes is made then the only way how the decomposition of the long-run multipliers matrix can be (approximately) done is by a non-block-diagonal cointegrating vectors matrix.

The results of the Monte Carlo study indicate as well that the GVAR model is outperformed by other valid econometric approaches in terms of the maximum

likelihood estimator of the long-run coefficients, when the cointegrating vectors matrix is block-diagonal. In terms of the standard deviation of the restricted MLE of the $\mathbf{B}$ matrix, the PVAR model performs well for $T = 150$ and $I \leq 6$, and the longer the time span is, the better the small sample properties are. For $T = 200$ the PVAR model clearly outperforms the GVAR and the individual VAR's for all considered values of I. However, it should be distinctly stressed that the PVAR framework requires small I and large T, therefore, it cannot be employed in all $\frac{I}{P}$ cases.
Finally, even though simultaneous investigation on both the cross-sectional cointegrating vectors and the cross-unit spillovers of disequilibrium in the VAR framework without a priori restrictions would require a large T in general, it is feasible for some specific cases in small or moderate samples. For example, Jacobson et al. (2008) consider modification of the PVAR model (4), by allowing for one common variable that spans the sub-spaces for all cross-section, as it is in case of the purchasing power parity.

Acknowledgements

The author wishes to thank two anonymous referees for their valuable comments.

References

[1] Bai J., (2009), Panel data models with interactive fixed effects, *Econometrica* 77, 1229–1279.

[2] Bi Y., Anwar S., (2017), US monetary policy shocks and the Chinese economy: a GVAR approach, *Applied Economics Letters* 24, 553–558.

[3] Bussière M., Chudik A., Sestieri G., (2009), Modelling global trade flows. Results from a GVAR model, ECB Working Paper Series 1087.

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

[5] Dees S., di Mauro F., Pesaran M. H., Smith V., (2007), Exploring the international linkages of the euro area: A global VAR analysis, *Journal of Applied Econometrics* 22, 1–38.

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

[7] Harbo L., Johansen S., Nielsen B., Rahbek A., (1998), Asymptotic inference on cointegrating rank in partial system, *Journal of Business and Economic Statistics* 16, 388–399.

[8] Jacobson T., Lyhagen J., Larsson R., Nessén M., (2008), Inflation, exchange rates and PPP in a multivariate panel cointegration model, *Econometrics Journal* 11, 58–79.

[9] Johansen S., (1991), Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica* 59, 1551–1580.

[10] Johansen S., Juselius K., (1994), Identification of the long-run and the short-run structure. An application to the ISLM model, *Journal of Econometrics* 63, 7–36.

[11] Kębłowski P., (2016), Canonical correlation analysis in panel vector error correction model. Performance Comparison, *Central European Journal of Economic Modelling and Econometrics* 8(4), 203–217.

[12] Larsson R., Lyhagen J., (2007), Inference in panel cointegration models with long panels, *Journal of Business & Economic Statistics* 25, 473–483.

[13] Larsson R., Villani M., (2001), A distance measure between cointegration spaces, *Economics Letters* 70, 21–27.

[14] Pesaran M. H., (2006), Estimation and inference in large heterogeneous panels with multifactor error structure, *Econometrica* 74, 967–1012.

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

[16] Temizsoy A., Montes-Rojas G., (2019), Measuring the effect of monetary shocks on European sovereign country risk: An application of GVAR models, *Journal of Applied Econometrics* 22, 484–503.