



A New Look at Cross-Country Aggregation in the Global VAR Approach: Theory and Monte Carlo Simulation

Halil Ibrahim Gunduz^{1,2} · Furkan Emirmahmutoglu³ · M. Eray Yucel⁴

Accepted: 11 February 2024
© The Author(s) 2024

Abstract

Requirements to understand and forecast the behavior of complex macroeconomic interactions mandate the use of high-dimensional macroeconomic models. The Global Vector Autoregressive (GVAR) modeling technique is very popular among them and it allows researchers and policymakers to take into account both the complex interdependencies that exist between various economic entities and the global economy through the world's trade and financial channels. However, determining the cross-section unit size while using this approach is not a trivial task. In order to address this issue, we suggest an objective procedure for the detection of the size of the cross-country aggregation in GVAR models. While doing so, we depart from the Akaike Information Criterion (AIC) and propose an analytical modification to it, mainly employing an ad hoc approach without violating Akaike's main principles. To supplement the theoretical results, small sample performances of those procedures are studied in Monte Carlo experiments as well as implementing our approach on real data. The numerical results suggest that our ad hoc modification of AIC can be used to determine the structure of the cross-section unit dimension in GVAR models, allowing the researchers and policymakers to build parsimonious models.

Keywords Global VAR · Cross-Country aggregation · Model selection · Akaike information criterion · Ad-hoc approach

1 Introduction

International high-dimensional datasets with relatively large cross-sectional (countries/regions; N) and time (T) dimensions are frequently used in economics. These datasets provide a wealth of information to both academics and practitioners in the field, helping them to understand economic developments. Although such large datasets provide more information, they also reveal interdependencies, such as how individual economies depend on the global system through various channels. Factors like resource sharing, political developments, trade, labour, and capital movements affect

Extended author information available on the last page of the article

relations between countries. Global shocks and technological advances can also cause economies to become interdependent. It is therefore important to take account of these interdependencies in models with a large number of variables.

Due to increased economic and financial integration among countries, macroeconomic policy spillovers have garnered significant attention in extensive research in recent decades. Various approaches have been employed to address these channels: Decompositions such as Kalman filtering (e.g., Canova & Marrinan, 1998; Gregory et al., 1997; Kose et al., 2003; Lumsdaine & Prasad, 2003), unobserved factor models (e.g., Bai & Ng, 2007; Giannone et al., 2005; Stock & Watson, 1999, 2002), Factor Augmented Vector Auto Regressions (FAVAR) (e.g., Bernanke et al., 2005; Stock & Watson, 2005), large-scale Bayesian VARs (e.g., Banbura et al., 2010; Carriero et al., 2009; De Mol et al., 2008; Giacomini & White, 2006), Panel VARs (PVAR) (e.g., Abrigo & Love, 2016; Canova & Ciccarelli, 2013; Love & Zicchino, 2006) have all been used.

Dees et al. (2007) (DdPS) highlight the usefulness of unobserved factor models but note the difficulties in identifying these factors and assigning economic interpretations to them. DdPS also emphasize that even when all the common factors are considered, significant residual interdependencies remain unexplained due to policy and trade spillovers. In an analysis of the UK economy using the FAVAR approach, Lagana and Mountford (2005) found that while the inclusion of additional variables through factor augmentation helps to solve the price puzzle, it creates new puzzles related to the counterintuitive effects of interest rate changes on house prices, equity prices, and the exchange rate. Furthermore, Kapetanios and Pesaran (2007) conducted Monte Carlo experiments and found that estimators using cross-sectional averages perform better in capturing common correlated effects than those based on principal components.

Another widely used econometric method for analyzing high-dimensional time series data is the Global Vector Autoregressive (GVAR) modelling approach. First proposed by Pesaran et al. (2004) (PSW), it was presented as a workable solution to create a coherent global model of the world economy. DdPS developed the theoretical justification for this approach to approximate a global factor mode.¹ The GVAR approach addresses the challenge of dimensionality by decomposing large-dimensional VARs into a smaller set of conditional models, which are interconnected through cross-sectional averages. Compared to Bayesian and Panel VARs, the GVAR approach provides an intuitive framework for understanding cross-country linkages and does not impose any restrictions on the dynamics of individual country sub-models. However, modelling a complex system like the global economy using the GVAR approach invariably involves many difficulties.

The most fundamental source of difficulty is the dimension and structure of the N . After PSW and DdPS, many studies have aimed to explain the basic variables of a *target* country. This involves prioritizing the model results of a single country in the dataset over which the GVAR modelling approach is used. When the data contain a large N in a GVAR model, data management and computation are very

¹ See Chapter 33 of Pesaran (2015) for a survey.

time-consuming. Hence, the complexity of a GVAR model stems from the composition of N , i.e., the aggregation of the cross-sections.

Since PSW introduced the model, there have been extensive applications of cross-country aggregation in the GVAR approach. Cesa-Bianchi et al. (2012) used the interdependence between China, Latin America, and the rest of emerging Asia, excluding China and India. They found that the long-run impact of a shock to China's GDP on Latin American economies has tripled since the mid-1990s, while the long-run impact of a shock to U.S. GDP has decreased by about 50%. Assenmacher (2013) re-examined the DdPS dataset to model the Swiss economy with only its three highest trading partners (accounting for about 80% of trade), excluding other countries. Cashin et al. (2014) utilized the GVAR model to examine the economic effects of oil-price shocks, distinguishing between supply- and demand-driven shocks. They examined 38 countries/regions, from 1979Q2 to 2011Q2, and created two regions: the six countries of the Gulf Cooperation Council (GCC) and the Euro Area block, same as in the DdPS. They found that the macroeconomic consequences of oil price shocks varied significantly, depending on the source of the shock and whether the country was an oil importer or exporter.

De Waal et al. (2015) estimated two models for South Africa, one restricted to South Africa and its three main trading partners (accounting for about 55% of trade) and the other considering all countries in the DdPS data. This approach contrasts with the lower bound on the share of trading partners share reported in Assenmacher (2013). Georgiadis (2016) employed a GVAR modelling approach that includes 61 economies from 1999Q1 to 2009Q4. He aggregates the economies of Estonia, Latvia, and Lithuania into a Baltic (BAL) region, while the economies of Venezuela, Ecuador, and Saudi Arabia are classified as an Oil Exporting Countries (OPC) block. His study focused on the important role of the US in the global economy, in particular the cross-border spillovers of its economic activities. The findings suggest that the economic influence of the US is as strong abroad as it is domestically, with the exception of a few economies such as some Australasian, African, and Latin American emerging markets and China. Spillovers were particularly significant for Russia, the Baltics States, Greece, Ireland, and Luxembourg. In general, spillovers to non-advanced economies were smaller than those to advanced economies, but still significant. The study raises the question of whether global welfare could be improved if US policymakers took these spillovers into account, given the role of the US dollar as the global reserve currency.

A recent study by Zahedi et al. (2022) focused on the growing influence of China on the global economy. They constructed a Global Vector Autoregressive (GVAR) dataset with data from 42 countries covering the period 2000Q1–2019Q3. The study treated China and the US as individual countries, while European countries were grouped together, and the rest were considered as the rest of the world (ROW). The findings show that China's increasing economic prosperity and global trade share have led to statistically significant spillover effects on the global economy. This effect is tied to China's substantial growth in investment and activity, which has boosted its global trade.

Modelling the *target* country and other important countries, such as *reference* country, individually, and aggregating of the remaining countries into 5–10 blocks or

regions makes the analysis more manageable. We consider the challenge of determining the appropriate dimension for cross-country aggregation in GVAR models and selecting the optimal model from the available options when using the same dataset. In our world of advanced computational capabilities, models of any size can be estimated for policy analysis. Policy-making institutions do not face practical difficulties in developing high-dimensional models and using them to generate forecasts. However, the availability of high computing power should not diminish the importance of the principle of parsimony. Therefore, in this paper, we propose an Akaike-type information criterion (AIC) and an ad hoc modification of this criterion to identify the optimal model from a set of competing models. The best approximating model is obtained by minimizing this criterion.

From a purely economic point of view, it is essential to reduce a GVAR model to its most parsimonious form. As will be discussed in detail in the subsequent sections, the GVAR approach involves a network comprising a *target* country, its influential economic and trade partners, and an aggregated entity referred to as ROW. For any *target* country, it is intuitive that the inclusion of more partners would improve the statistical quality of a GVAR. However, improvements in the statistical and analytical quality of scenario analyses do not necessarily occur simultaneously. In order to ensure a robust scenario analysis with concrete assumptions, it is important to distinguish carefully between the most impactful partners and the rest of the world. This raises the question of whether to include a partner country as a separate entity or as part of the rest of the world in the GVAR model. This paper aims to provide a reliable starting point for answering this question by proposing a useful modification of the AIC.

In this paper, we also individually examine the selection of the dimension of the GVAR model for the developing countries. Our analysis is based on the updated DdPS dataset, which includes data from 33 countries from 1979Q2 to 2019Q4. The empirical applications yield findings that underscore the importance of the US, China, the Euro, and Japan for developing economies. As major participants in the global economy, these countries exert significant influence on developing economies worldwide through their economic relationships and interactions. They play a crucial role as major trading partners, sources of investment, and providers of technology and expertise for many developing countries.

The rest of the paper is structured as follows: Sect. 2 outlines the well-known GVAR model settings and some preliminary remarks on information criteria. Section 3 presents an AIC-type information criterion and its ad hoc version. Section 4 conducts Monte Carlo simulation studies to examine the performance of the proposed criterion. Section 5 contains a real data analysis. Section 6 provides a conclusion. The proofs are provided in the Appendix, and further details of the simulation experiments and empirical results are given in the Supplementary Materials.

2 An Overview of the GVAR Modelling Approach

In this section we provide an overview of the GVAR modelling approach due to Pesaran et al. (2004) to lay down the basis of our work in the subsequent sections. We consider that there are N countries (or regions) in our settings, $i = 1, 2, \dots, N$. The

aim here is to explain the sources of the changes in the $(k_i \times 1)$ vector of the country-specific domestic (endogenous) variables vector, \mathbf{x}_{it} . If a total of k , $k = \sum_{i=1}^N k_i$, endogenous variables are to be estimated using the standard VAR model, the problem known generally as the curse of dimensionality arises.² The GVAR modelling approach solves this problem in two steps. In the first step, each country-specific conditional VARX* model—domestic variables are modelled as a VAR, augmented by the vector of foreign-specific variables—is estimated individually. In the second step, the estimated conditional country models are combined into a large global model.

2.1 Country-Specific VARX* Models

We consider the following VARX* (p_i, q_i) model as follows:

$$\mathbf{x}_{it} = \mathbf{a}_{i0} + \mathbf{a}_{i1}t + \sum_{\ell=1}^{p_i} \Phi_{i\ell} \mathbf{x}_{i,t-\ell} + \Lambda_{i0} \mathbf{x}_{it}^* + \sum_{\ell=1}^{q_i} \Lambda_{i\ell} \mathbf{x}_{i,t-\ell}^* + \varepsilon_{it} \tag{2.1}$$

where \mathbf{x}_{it} and \mathbf{x}_{it}^* are $(k_i \times 1)$ vector of domestic variables specific to countries and $(k_i^* \times 1)$ vector of foreign variables, respectively.³ $\Phi_{i\ell}$ are $(k_i \times k_i)$ matrices of lagged coefficients, and $\Lambda_{i\ell}$ are $(k_i \times k_i^*)$ matrices of coefficients associated with the foreign specific variables.

The foreign variables, denoted by \mathbf{x}_{it}^* , is the weighted average of domestic variables

$$\mathbf{x}_{it}^* = \sum_{n=1}^N w_{in} \mathbf{x}_{nt}$$

where w_{in} be a series of weights with $w_{ii} = 0$ and $\sum_{n=1}^N w_{in} = 1$. ε_{it} is a $(k_i \times 1)$ vector of country-specific errors with $E[\varepsilon_{it}] = 0$ and Σ_{ii} , $(k_i \times k_i)$ a nonsingular covariance matrix. The relationship between the shock of h th and h' th variables of the i th country is $\sigma_{ii,hh'} = cov(\varepsilon_{iht}, \varepsilon_{ih't})$. More specifically,

$$\varepsilon_{it} \sim IID(0, \Sigma_{ii}) \tag{2.2}$$

where $\Sigma_{ii} = (\sigma_{ii,hh'})$ and the cross-sectional dependence of shocks between i th and n th country is expressed as

$$\Sigma_{in} = cov(\varepsilon_{it}, \varepsilon_{nt}) = E(\varepsilon_{it} \varepsilon_{nt}') \tag{2.3}$$

where $i \neq n$ and $\Sigma_{in} = (\sigma_{in,hh'})$.

The parameters of the VARX* (p_i, q_i) model in Eq. (2.1) are estimated for all countries separately. To do this, $(k_i + k_i^*) \times 1$ dimensional vector of domestic and

² See Chudik and Pesaran (2016) for further details.

³ When observable variables such as oil prices are added to the model, their contemporaneous and lagged values is on the right-hand side of equation.

country-specific foreign variables, $\mathbf{z}_{it} = (\mathbf{x}'_{it}, \mathbf{x}^{*t})'$, is defined and Eq. (2.1) is rewritten as

$$\mathbf{A}_{i0}\mathbf{z}_{it} = \mathbf{a}_{i0} + \mathbf{a}_{i1}t + \sum_{\ell=1}^p \mathbf{A}_{i\ell}\mathbf{z}_{it-\ell} + \varepsilon_{it} \tag{2.4}$$

where $\mathbf{A}_{i0} = (\mathbf{I}_{k_i}, -\mathbf{\Lambda}_{i0})$, $\mathbf{A}_{i\ell} = (\mathbf{\Phi}_{i\ell}, \mathbf{\Lambda}_{i\ell})$, $p = \max_i(p_i, q_i)$, and, for $\ell > p_i$, $\mathbf{\Phi}_{i\ell} = \mathbf{0}$, and, for $\ell > q_i$, $\mathbf{\Lambda}_{i\ell} = \mathbf{0}$.

The error-correction representation, VECMX^* , of Eq. (2.1) can be equivalently written as

$$\Delta\mathbf{x}_{it} = \mathbf{a}_{i0} + \mathbf{a}_{i1}t + \mathbf{\Lambda}_{i0}\Delta\mathbf{x}_{it}^* - \mathbf{\Pi}_i\mathbf{z}_{i,t-1} + \sum_{\ell=1}^p \mathbf{H}_{i\ell}\Delta\mathbf{z}_{i,t-\ell} + \varepsilon_{it} \tag{2.5}$$

where

$$\mathbf{\Pi}_i = \mathbf{A}_{i0} - \sum_{\ell=1}^p \mathbf{A}_{i\ell}, \text{ and } \mathbf{H}_{i\ell} = -(\mathbf{A}_{i,\ell+1} + \mathbf{A}_{i,\ell+2} + \dots + \mathbf{A}_{i,\ell+p}).$$

The estimation theory of VAR models with weakly exogenous $I(1)$ variables was developed by Harbo et al. (1998) and PSW. PSW indicates that the weak exogeneity assumption can be tested. This assumption is generally not rejected when the weights used to create the (*) variables are granular,⁴ and when the target country's economy is small compared to the rest of the world.

For unrestricted constant and restricted trend coefficient, the error correction model in Eq. (2.5) can be written as

$$\Delta\mathbf{x}_{it} = \mathbf{c}_{i0} - \boldsymbol{\alpha}_i\boldsymbol{\beta}'_i[\mathbf{z}_{i,t-1} - \gamma_i(t-1)] + \mathbf{\Lambda}_{i0}\Delta\mathbf{x}_{it}^* + \sum_{\ell=1}^p \mathbf{\Gamma}_{i\ell}\Delta\mathbf{z}_{i,t-\ell} + \varepsilon_{it} \tag{2.6}$$

where $\forall i \in N$, the dimensions of $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ are $k_i \times r_i$ and $(k_i + k_i^*) \times r_i$, respectively. And also, the ranks of these matrices are $\text{rank}(\boldsymbol{\alpha}_i) = \text{rank}(\boldsymbol{\beta}_i) = r_i$.

By adjusting $\boldsymbol{\beta}_i$ as $\boldsymbol{\beta}_i = (\boldsymbol{\beta}'_{ix}, \boldsymbol{\beta}'_{ix^*})'$, the possibility of cointegration both within domestic variables and between domestic and foreign variables, and consequently across domestic variables of different countries is allowed as follows:

$$\boldsymbol{\beta}'_i(\mathbf{z}_{it} - \gamma_it) = \boldsymbol{\beta}'_{ix}\mathbf{x}_{it} + \boldsymbol{\beta}'_{ix^*}\mathbf{x}^*_{it} - (\boldsymbol{\beta}'_i\boldsymbol{\gamma}_i)t$$

The error correction model in Eq. (2.6) is estimated by reduced rank regression.⁵

⁴ Pesaran (2006) stated the conditions that the weights should be such that (i) : $w_i = O\left(\frac{1}{N}\right)$, (ii) : $\sum_{i=1}^N w_i = 1$, (iii) : $\sum_{i=1}^N |w_i| < K$,

⁵ See Pesaran (2015) for further details.

2.2 Solution Strategy

In the second step of the GVAR approach, individually estimated country models are combined to create a large Global VAR model. For this purpose, “link” matrices \mathbf{W}_i are used, which contains country-specific weights. For all i , this large GVAR model’s dimension is $(k_i + k_i^*) \times k$, where $k = \sum_{i=1}^N k_i$ and $k^* = \sum_{i=1}^N k_i^*$ are the total number of endogenous and exogenous variables in the entire system, respectively. This arrangement enables the rewriting of the global model in terms of \mathbf{x}_t , which is a $(k \times 1)$ vector of all variables:

$$\mathbf{z}_{it} = \mathbf{W}_i \mathbf{x}_t \tag{2.7}$$

Substituting the right-hand side of this expression into Eq. (2.4) produces

$$\mathbf{A}_{i0} \mathbf{W}_i \mathbf{x}_t = \mathbf{a}_{i0} + \mathbf{a}_{i1} t + \sum_{\ell=1}^p \mathbf{A}_{i\ell} \mathbf{W}_i \mathbf{x}_{t-\ell} + \epsilon_{it} \tag{2.8}$$

Its closed form is

$$\mathbf{G}_0 \mathbf{x}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \sum_{\ell=1}^p \mathbf{G}_\ell \mathbf{x}_{t-\ell} + \epsilon_t \tag{2.9}$$

where, for $\ell = 0, 1, 2, \dots, p$,

$$\mathbf{G}_\ell = \begin{pmatrix} \mathbf{A}_{1,\ell} \mathbf{W}_1 \\ \mathbf{A}_{2,\ell} \mathbf{W}_2 \\ \vdots \\ \mathbf{A}_{N,\ell} \mathbf{W}_N \end{pmatrix}, \mathbf{a}_0 = \begin{pmatrix} \mathbf{a}_{10} \\ \mathbf{a}_{20} \\ \vdots \\ \mathbf{a}_{N0} \end{pmatrix}, \mathbf{a}_1 = \begin{pmatrix} \mathbf{a}_{11} \\ \mathbf{a}_{21} \\ \vdots \\ \mathbf{a}_{N1} \end{pmatrix} \text{ and } \epsilon_t = \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \vdots \\ \epsilon_{Nt} \end{pmatrix}$$

If \mathbf{G}_0 is non-singular matrix, which depends on weight matrices and estimation of parameters, the GVAR(p) is obtained by premultiplying both sides of the equation given in (2.9) by \mathbf{G}_0^{-1} :

$$\mathbf{x}_t = \mathbf{b}_0 + \mathbf{b}_1 t + \sum_{\ell=1}^p \mathbf{H}_\ell \mathbf{x}_{t-\ell} + \mathbf{u}_t \tag{2.10}$$

where

$$\mathbf{b}_0 = \mathbf{G}_0^{-1} \mathbf{a}_0, \mathbf{b}_1 = \mathbf{G}_0^{-1} \mathbf{a}_1 \\ \mathbf{H}_\ell = \mathbf{G}_0^{-1} \mathbf{G}_\ell, \mathbf{u}_t = \mathbf{G}_0^{-1} \epsilon_t$$

PSW states that the number of cointegrations in the global model cannot exceed the sum of the number of long-term cointegrations in the conditional country models.

2.3 Cross-Country Aggregation in GVAR Modelling

The GVAR modelling approach is a flexible method for taking into account various global links. However, modelling a large number of countries can pose difficulties in terms of data management, computation and data analysis. To make the analysis more manageable, one possible approach is to model the countries of high importance individually, and aggregate the remaining countries into one or more different regions. This section discusses how to perform such regional aggregation.

Suppose the dataset contains R countries/regions, $i = 1, 2, \dots, R$, and i th region is composed of N_i countries. Let \mathbf{x}_{ijt} and \mathbf{x}_{ijt}^* , for $j = 1, 2, \dots, N_i$, denote the vectors of country-specific domestic and foreign variables in the j th country of the i th region, respectively. In this case, Eq. (2.1) can be rewritten as:

$$\mathbf{x}_{ijt} = \mathbf{a}_{ij0} + \mathbf{a}_{ij1}t + \sum_{\ell=1}^{p_i} \Phi_{ij\ell} \mathbf{x}_{ij,t-\ell} + \Lambda_{ij0} \mathbf{x}_{ijt}^* + \sum_{\ell=1}^{q_i} \Lambda_{ij\ell} \mathbf{x}_{ij,t-\ell}^* + \epsilon_{ijt} \quad (2.11)$$

Here, the source of heterogeneity in the parameter vectors $\Phi_{ij\ell}$, Λ_{ij0} and $\Lambda_{ij\ell}$ comes from distinct country-specific model coefficients. Regional weights are used to aggregate the model.

The weight of the j th country in the i th region is w_{ij}^0 , and the sum of the weights of the countries in a region is $\sum_{j=1}^{N_i} w_{ij}^0 = 1$. Using the weighted regional variables,

$$\mathbf{x}_{it} = \sum_{j=1}^{N_i} w_{ij}^0 \mathbf{x}_{ijt}$$

we can derive the regional model similar to the VARX* model provided in Eq. (2.1):

$$\mathbf{x}_{it} = \mathbf{a}_{i0} + \mathbf{a}_{i1}t + \sum_{\ell=1}^{p_i} \Phi_{i\ell} \mathbf{x}_{i,t-\ell} + \Lambda_{i0} \mathbf{x}_{it}^* + \sum_{\ell=1}^{q_i} \Lambda_{i\ell} \mathbf{x}_{i,t-\ell}^* + \epsilon_{it} \quad (2.12)$$

where

$$\mathbf{a}_{i0} = \sum_{j=1}^{N_i} w_{ij}^0 \mathbf{a}_{ij0} \mathbf{a}_{i1} = \sum_{j=1}^{N_i} w_{ij}^0 \mathbf{a}_{ij1}$$

and

$$\epsilon_{it} = \sum_{j=1}^{N_i} w_{ij}^0 \epsilon_{ijt}$$

Region-specific foreign variables, \mathbf{x}_{it}^* , can be formed using regional trade weights or country-specific trade weights. After PSW and DdPS, most studies in the literature focused on explaining the basic variables of a *targeted* country. This raises the question of how to organize countries other than the *targeted* one in the cross-section dimension. Assenmacher (2013) modelled only three trading

partners along with Switzerland itself, and utilizing the DdPS dataset (GVAR-DdPS). The model included the three countries with the highest trade rates (approximately 80%) excluding the others. This more restricted version of the GVAR-DdPS dataset was justified by its ease of analysis, with fewer parameters estimated for a smaller number of countries. However, the study didn't elaborate on different model alternatives or the reasons for choosing one.

Similarly, De Waal et al. (2015) aimed to explain the South African economy using two forms of GVAR-DdPS, restricted and unrestricted. The restricted model included data from GVAR-DdPS data only for South Africa and its three main trading partners (approximately 55%), while the unrestricted model used all GVAR-DdPS data. The reason for setting up the models in this way contradicts the lower limit of the trade partnership share in Assenmacher (2013). The study stated that the trade partnership for South Africa is closer to 80%, and thus nearly the entire GVAR-DdPS dataset should be used, eliminating the distinction between the restricted and unrestricted model. However, no conclusion was reached as to which of the two models should be selected.

A common feature of these studies is the exclusion of what is referred to in the models as the 'rest of the world'. This refers to the aggregation of countries into a region, achieved by restricting the country dimension in the dataset, which may lead to model specification errors. Competing models with different country components are generally not considered when explaining the dynamics of the *target* country's economy. To the best of our knowledge, the literature on the GVAR modelling approach has not considered exploring the strengths and weaknesses of multiple potential models derived from geographical, cultural or political regions. Given the multiple model options, the choice of the most suitable model is covered by the topic of model selection. The selection of the most appropriate model among the GVAR models containing different countries and/or regions, all derived from the same dataset, is identified as a significant gap in the literature. Therefore, model selection criteria can be used to eliminate this uncertainty. Specifically, an Akaike-type information criterion and an ad hoc modification of this criterion are proposed to allow the selection among all possible GVAR model specifications that can be generated within the same dataset.

3 Information Criteria in Model Selection of GVAR

Based on the framework developed in Sect. 2 by Pesaran et al. (2004), we introduce our proposed information criteria in this section. At the beginning, we adopt the general principle of Gourieroux and Monfort (1995) to derive the Akaike-type information criterion. To simplify to computations, we assume that $p_i = q_i = p$ and cross-sectional dimension contains N countries (or regions), with $i \in \mathcal{S}_{(N)} \equiv \{1, \dots, N\} \subseteq \mathbb{N}$.⁶ As previously defined in Eq. (2.4), the vector \mathbf{z}_{it} vector contains all variables in the whole system. It can be re-expressed as

⁶ In this case, there is no loss of generality.

$\mathbf{z}_t = (\mathbf{x}'_t, \mathbf{x}^{*t})'$, where \mathbf{x}_t and \mathbf{x}^{*t} are $(k \times 1)$ and $(k^* \times 1)$ vectors respectively, and are compatible with the definitions given in Sect. 2.2. The variable \mathbf{z}_t , namely

$$\mathbf{z}_t = \sum_{\ell=1}^p \Phi_{\ell} \mathbf{z}_{t-\ell} + \mathbf{a}_0 + \mathbf{a}_1 t + \epsilon_t \tag{3.1}$$

is subject to the curse of dimensionality introduced in Sect. 2.1. The unit root and cointegration properties of \mathbf{z}_t are expressed with the vector error correction (VEC) model structure of the VAR(p) model:

$$\Delta \mathbf{z}_t = -\Pi \mathbf{z}_{t-1} + \sum_{\ell=1}^{p-1} \Gamma_{z\ell} \Delta \mathbf{z}_{t-\ell} + \mathbf{a}_0 + \mathbf{a}_1 t + \epsilon_t \tag{3.2}$$

where

$$\Pi = \mathbf{I}_{(k+k^*)} - \sum_{\ell=1}^p \Phi_{\ell}, \quad \Gamma_{z\ell} = - \sum_{\tau=\ell+1}^p \Phi_{\tau} \tag{3.3}$$

for $\ell = 1, \dots, p - 1$, $\{\Gamma_{z\ell}\}_{\ell=1}^{p-1}$ and Π are $(k + k^*) \times (k + k^*)$ matrices of short and long run responses.

3.1 Information Criterion

$(k + k^*) \times 1$ Dimensional vector ϵ_t in Eq. (3.1) is closed form of the VEC model of ϵ_{it} , given in Eq. (2.2). We make the following assumption.

Assumption 1 (*Weak Cross-Sectional Dependence in Errors*). ϵ_t is independent and identically multivariate normally distributed, $\epsilon_t \sim IIDN(0, \Sigma)$, and Σ , is positively defined, $\|\Sigma\|_2 > 0$.

Remark 1 *High-dimensional time series models, such as GVAR, widely subject to error cross-sectional dependence. In addition, Assumption 1 allows the likelihood analysis of the VEC model given in (3.2) and conditional expression of the model.*⁷

When ϵ_t is rearranged to be $\mathbf{z}_t = (\mathbf{x}'_t, \mathbf{x}^{*t})'$, $\epsilon_t = (\epsilon'_{xt}, \epsilon'_{x^*t})'$, variance–covariance matrix of the model in accordance with Assumption 1 is

$$E(\epsilon_t \epsilon'_t) = \Sigma = \begin{pmatrix} \sum_{xx} & \sum_{xx^*} \\ \sum_{x^*x} & \sum_{x^*x^*} \end{pmatrix}$$

where the dimensions of \sum_{xx} , \sum_{xx^*} , \sum_{x^*x} and $\sum_{x^*x^*}$ matrices are $(k \times k)$, $(k \times k^*)$, $(k^* \times k)$ and $(k^* \times k^*)$, respectively. And then, ϵ_{xt} is expressed by ϵ_{x^*t} ⁸.

⁷ See [Gourieroux and Monfort \(1995\)](#) and [Pesaran et al. \(2000\)](#) for more information.

⁸ See [Bierens \(2004\)](#) for the transformation.

$$\epsilon_{xt} = \sum_{xx^*} \sum \epsilon_{x^*x^*x^*t}^{-1} + v_t \tag{3.4}$$

where $v_t \sim IIDN(0, \Sigma_{vv})$, $\Sigma_{vv} = \Sigma_{xx} - \Sigma_{xx^*} \Sigma_{x^*x^*}^{-1} \Sigma_{x^*x}$, v_t and ϵ_{x^*t} are unrelated.

For Δx_t , we have conditional model:

$$\Delta x_t = -\Pi_{xx^*} z_{t-1} + \Lambda_0 \Delta x_t^* + \sum_{\ell=1}^{p-1} \Lambda_\ell \Delta z_{t-\ell} + c_0 + c_1 t + v_t \tag{3.5}$$

where $\Pi_{xx^*} = \Pi_x - \Sigma_{xx^*} \Sigma_{x^*x^*}^{-1} \Pi_{x^*}$, $\Lambda_0 = \Sigma_{xx^*} \Sigma_{x^*x^*}^{-1}$, for $\ell = 1, \dots, p-1$, $\Lambda_\ell = \Gamma_\ell - \Sigma_{xx^*} \Sigma_{x^*x^*}^{-1} \Gamma_\ell^*$, $c_0 = a_{x0} - \Sigma_{xx^*} \Sigma_{x^*x^*}^{-1} a_{x^*0}$ and $c_1 = a_{x1} - \Sigma_{xx^*} \Sigma_{x^*x^*}^{-1} a_{x^*1}$.

Following Pesaran et al. (2000), it assumes that $\{x_t^*\}_{t=1}^\infty$ is weakly exogenous, namely

$$\Pi_{x^*} = 0$$

so $\Pi_{xx^*} = \Pi_x$.

More specifically, we have

$$\Delta x_t = -\Pi_x z_{t-1} + \Lambda_0 \Delta x_t^* + \sum_{\ell=1}^{p-1} \Lambda_\ell \Delta z_{t-\ell} + c_0 + c_1 t + v_t \tag{3.6}$$

and

$$\Delta x_t^* = \sum_{\ell=1}^{p-1} \Lambda_\ell \Delta z_{t-\ell} + a_{x^*0} + \epsilon_{x^*t} \tag{3.7}$$

The $\{x_t^*\}_{t=1}^\infty$ and $\{x_t\}_{t=1}^\infty$ processes are Garratt et al., (2006) state that trend coefficients must be restricted in order to eliminate the quadratic trend and ensure that the trend in the deterministic part of x_t is linear. In this case, $a_{x^*1} = 0$ and so $c_1 = a_{x1}$.

The equation system given in (3.2) has different dynamics, for endogenous variables, Δx_t , and for exogenous variables, Δx_t^* , and thus consists of two parts. Granger and Lin (1995) and Pesaran et al. (2000) describe variables with $I(1)$ characteristics associated with the concept of weak exogeneity as “long-run forcing”.

The $\{x_t^*\}_{t=1}^\infty$ and $\{x_t\}_{t=1}^\infty$ processes are consistent with the conditional and the marginal model definitions, respectively.⁹ Thus, the information criteria proposed for the GVAR modelling approach can be considered within the framework of the likelihood approach described by Pesaran et al. (2000).

To simplify the derivation, unrestricted intercept and restricted trend model option given in Eq. (2.6), $c_0 \neq 0$ and $c_1 = \Pi_x \gamma$, be used in Eq. (3.6) and we have

$$\Delta x_t = c_0 + \Pi_x \tilde{z}_{t-1} + \Lambda_0 \Delta x_t^* + \sum_{\ell=1}^{p-1} \Lambda_\ell \Delta z_{t-\ell} + v_t \tag{3.8}$$

⁹ See Pesaran (2015) for the conditional and the marginal model concepts.

where $\tilde{\mathbf{z}}_{t-1} = (t, \mathbf{z}'_{t-1})'$, and $\mathbf{\Pi}_{\tilde{\mathbf{x}}} = \mathbf{\Pi}_x(-\gamma, I_{(k+k^*)})$. In addition, $\mathbf{\Pi}_x = \boldsymbol{\alpha}_x \boldsymbol{\beta}'$, $\mathbf{\Pi}_x$, $\boldsymbol{\alpha}_x$ and $\boldsymbol{\beta}$ are $k \times (k + k^*)$, $(k \times r)$ and $(k + k^*) \times r$ dimensional matrices, respectively. $r = \sum_{i=1}^N r_i$ is the total number of cointegrations in the GVAR system.

When the VEC model given in Eq. (3.8) is stacked for T observation,

$$\Delta \mathbf{X} = \mathbf{c}'_{0T} + \Lambda \Delta \mathbf{Z}_- + \mathbf{\Pi}_{\tilde{\mathbf{x}}} \tilde{\mathbf{Z}}_{-1} + \mathbf{V} \tag{3.9}$$

where $\Delta \mathbf{X} \equiv (\Delta \mathbf{x}_1, \dots, \Delta \mathbf{x}_T)$, l_T is $T \times 1$ dimensional ones vector, $\Delta \mathbf{X}^* \equiv (\Delta \mathbf{x}_1^*, \dots, \Delta \mathbf{x}_T^*)$, $\Delta \mathbf{Z}_{-\ell} \equiv (\Delta \mathbf{z}_{1-\ell}, \dots, \Delta \mathbf{z}_{T-\ell})$, $\ell = 1, \dots, p - 1$, $\Lambda \equiv (\Lambda_0, \Lambda_1, \dots, \Lambda_{p-1})$, $\Delta \mathbf{Z}_- \equiv (\Delta \mathbf{X}^{*'} , \Delta \mathbf{Z}_{-1}' , \dots, \Delta \mathbf{Z}_{1-p}')'$, $\tilde{\mathbf{Z}}_{-1} \equiv (T, \mathbf{Z}_{-1})'$ $\mathbf{Z}_{-1} \equiv (\mathbf{z}_0, \dots, \mathbf{z}_{T-1})$, and $\mathbf{V} \equiv (\mathbf{v}_1, \dots, \mathbf{v}_T)$.

In order to utilize the model presented in Eq. (3.9) instead of the unconditional VEC($p - 1$) model given in Eq. (3.2), which presents a dimension problem when estimated directly, the log-likelihood function is split into two parts: the conditional and the marginal model.

We can further explain the above expression with the following lemma:

Lemma 1 *If the assumption of weak exogeneity is valid, $\mathbf{\Pi}_{x^*} = 0$, the log-likelihood function of the unconditional model in Eq. (3.2) can be derived from the conditional and the marginal models given in Eq. (3.6) and Eq. (3.7):*

$$\ell(\boldsymbol{\theta}_{\Delta z}, \boldsymbol{\Sigma}) = \ell_v(\boldsymbol{\theta}, \boldsymbol{\Sigma}_{vv}) + \ell_*(\boldsymbol{\theta}^* \boldsymbol{\Sigma}_{x^*x^*})$$

where $\boldsymbol{\theta}_{\Delta z}$, $\boldsymbol{\theta}$ and $\boldsymbol{\theta}^*$ are vectors of parameters in unconditional, conditional and marginal models, respectively.

The proof of Lemma 1 is given in Appendix.

The log-likelihood function of the conditional model in Eq. (3.9) is

$$l(\boldsymbol{\theta}) = \frac{-kT}{2} \log(2\pi) - \frac{T}{2} \log |\boldsymbol{\Sigma}_{vv}| - \frac{1}{2} tr(\boldsymbol{\Sigma}_{vv}^{-1} \mathbf{V} \mathbf{V}') \tag{3.10}$$

where $\boldsymbol{\theta}$ is a $K \times 1$ dimensional vector,

$$\boldsymbol{\theta} = \left(\text{vec}(\Lambda)', \text{vec}(\boldsymbol{\alpha}_x)', \text{vec}(\boldsymbol{\beta}_x)', \text{vech}(\boldsymbol{\Sigma}_{vv})' \right)'$$

and

$$K = \left[k \times [k^* + (k + k^*)(p - 1)] + kr + (k + k^*) \times r + \left(\frac{k(k + 1)}{2} \right) \right] \tag{3.11}$$

In Eq. (3.2), we define $\Delta \mathbf{z}_t$ as a continuous random variable, $\Delta \mathbf{Z}$, where $\Delta \mathbf{z}_t = (\Delta \mathbf{x}'_t, \Delta \mathbf{x}^{*'}_t)'$, characterized by the probability density function $f(\Delta \mathbf{z}, \boldsymbol{\theta}_{\Delta z})$. The parameter vector $\boldsymbol{\theta}_{\Delta z}$ is $(K_{\Delta z} \times 1)$ dimensional, with $\boldsymbol{\theta}_{K_{\Delta z}} \in \Theta \subset \mathbb{R}^{K_{\Delta z}}$. Let's assume the true parameter vector for $\boldsymbol{\theta}_{\Delta z0}$ is $\boldsymbol{\theta}_{\Delta z}$, which is unknown. This vector originates

from the probability density function $f(\Delta \mathbf{z}, \boldsymbol{\theta}_{\Delta z;0})$, which correctly reflects observed sample data. We select a $\boldsymbol{\theta}_{\Delta z}$ such that the actual parameter vector closely approximates $\boldsymbol{\theta}_{\Delta z;0}$. Thus, the probability density functions $f(\Delta \mathbf{z}, \boldsymbol{\theta}_{\Delta z})$ and $f(\Delta \mathbf{z}, \boldsymbol{\theta}_{\Delta z;0})$ are expressed as the rival model and the actual model, respectively. Using Lemma 1, these models are described as $f(\Delta \mathbf{x}, \boldsymbol{\theta})$ and $f(\Delta \mathbf{x}, \boldsymbol{\theta}_0)$.

The Kullback–Leibler (K-L) information distance is generally used to derive an Akaike-type information criterion.¹⁰ In order to use the K-L information to measure the distance between the two models, the following assumptions should be made. These assumptions are also known as regularity conditions (RC).¹¹ We consider that i' , be the i' th element of $\boldsymbol{\theta}$.

Assumption 2 (RC1) $\forall \boldsymbol{\theta} \in \Theta$,

$$\frac{\partial \log f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \theta_{i'}}, \frac{\partial^2 \log f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \theta_{i'} \partial \theta_{j'}}, \frac{\partial^3 \log f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \theta_{i'} \partial \theta_{j'} \partial \theta_{k'}}$$

the derivatives exist for all $\Delta \mathbf{x}$ and $i', j', k' = 1 \dots K$.

Assumption 3 (RC2) For $\boldsymbol{\theta}_0 \in \Theta$, functions of $G(\Delta \mathbf{x})$, $H(\Delta \mathbf{x})$ and $K(\Delta \mathbf{x})$ exist,

$$\left| \frac{\partial f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \theta_{i'}} \right| \leq G_{i'}(\Delta \mathbf{x}), \left| \frac{\partial^2 f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \theta_{i'} \partial \theta_{j'}} \right| \leq H_{i'j'}(\Delta \mathbf{x}), \left| \frac{\partial^3 f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \theta_{i'} \partial \theta_{j'} \partial \theta_{k'}} \right| \leq K_{i'j'k'}(\Delta \mathbf{x})$$

inequalities exist for $\Delta \mathbf{x}$ and $i', j', k' = 1 \dots K$, and

$$\int G_{i'}(\Delta \mathbf{x}) d\Delta \mathbf{x} < \infty, \int H_{i'j'}(\Delta \mathbf{x}) d\Delta \mathbf{x} < \infty, \int K_{i'j'k'}(\Delta \mathbf{x}) d\Delta \mathbf{x} < \infty$$

Assumption 4 (RC3) For $\boldsymbol{\theta} \in \Theta$,

$$\mathbf{B}(\boldsymbol{\theta}) = E \left[\left(\frac{\partial \log f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \log f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right) \right] \tag{3.12}$$

exist and $\| \mathbf{B}(\boldsymbol{\theta}) \|_2 > 0$.

Remark 2 According to RC1, $\frac{\partial \log f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ has Taylor series expansion as a function of $\boldsymbol{\theta}$. $\frac{\partial \log f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ is considered score vector. RC2 allows derivatives of $\int f(\Delta \mathbf{x}, \boldsymbol{\theta}) d\Delta \mathbf{x}$ and $\int \left[\frac{\partial \log f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] d\Delta \mathbf{x}$ to be obtained according to $\boldsymbol{\theta}$. Finally, RC3 requires the score vector to have a finite variance.

Lemma 2 If the regularity conditions are provided in RC1-RC3, the score vector, $\mathbf{S}(\boldsymbol{\theta}) = \frac{\partial \log f(\Delta \mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$, has zero mean and constant variance.

¹⁰ For further details of K-L information, see Burnham and Anderson (2002).

¹¹ See Hogg et al. (2005) for detailed information.

The proof of Lemma 2 is given in Appendix.

For VEC model given in (3.9), the K-L information is

$$\begin{aligned} K - L(\theta_0; \theta) &= E[\log f(\Delta \mathbf{X}, \theta_0) - \log f(\Delta \mathbf{X}, \theta)] \\ &= \int f(\Delta \mathbf{x}, \theta_0) \log f(\Delta \mathbf{x}, \theta_0) d\Delta \mathbf{x} - \int f(\Delta \mathbf{x}, \theta) \log f(\Delta \mathbf{x}, \theta) d\Delta \mathbf{x} \\ &= H(\theta_0; \theta_0) - H(\theta_0; \theta) \end{aligned} \quad (3.13)$$

where E indicates the expectation of the $\Delta \mathbf{X}$ according to $f(\Delta \mathbf{X}, \theta_0)$.

When the expression is arranged, the K-L is

$$K - L(\theta_0; \theta) = \int f(\Delta \mathbf{x}, \theta_0) \log \left(\frac{f(\Delta \mathbf{x}, \theta_0)}{f(\Delta \mathbf{x}, \theta)} \right) d\Delta \mathbf{x} \quad (3.14)$$

Proposition 1 *If the regularity conditions are provided in RC1-RC3, for the VEC model given in Eq. (3.9), $K - L(\theta_0; \theta)$ is a non-negative measure, the amount of information of the distance of $f(\Delta \mathbf{x}, \theta)$ to $f(\Delta \mathbf{x}, \theta_0)$ probability density function.*

The proof of Proposition 1 is given in Appendix.

Corollary 1 *Proposition 1 provides that the analytical characteristics of the amount of $K - L(\theta_0; \theta)$ information for the VEC model given in Eq. (3.9) are.*

- (i) *If $f(\Delta \mathbf{x}, \theta_0) \neq f(\Delta \mathbf{x}, \theta)$, $K - L(\theta_0; \theta) > 0$.*
- (ii) *If, and only if $f(\Delta \mathbf{x}, \theta_0) = f(\Delta \mathbf{x}, \theta)$, if the model is specified correctly, $K - L(\theta_0; \theta) = 0$.*
- (iii) *Because $\Delta \mathbf{X}$ is independent and identically distributed, $K - L(\theta_0; \theta)$ is additive.*

f_0 and f will be used instead of $f(\Delta \mathbf{x}, \theta_0)$ and $f(\Delta \mathbf{x}, \theta)$ for ease of spelling. If g is an another competing model other than f , Eq. (3.14) is rewritten as

$$K - L(f_0; f) = H(f_0) - H(f_0; f)$$

$$K - L(f_0; g) = H(f_0) - H(f_0; g)$$

$$K - L(f_0; f) - K - L(f_0; g) = H(f_0; g) - H(f_0; f).$$

When comparing competing models, the actual correct model f_0 does not need to be known. Furthermore, the f_0 model is never truly known. Therefore, the priority

in the amount of K-L information given in Eq. (3.14) will be the estimation of the competing model f . The K-L information amount is expressed as:

$$K - L(f_0; \hat{f}) = H(f_0) - H(f_0; \hat{f}).$$

The estimate of the amount of K-L information is random because it depends on \hat{f} . We focus on the expected amount of K-L information¹²

$$E[K - L(f_0; \hat{f})] = \int f_0(\Delta \mathbf{x}) \log f_0(\Delta \mathbf{x}) d\Delta \mathbf{x} - E \left[\int f_0(\Delta \mathbf{x}) \log \hat{f}(\Delta \mathbf{x}) d\Delta \mathbf{x} \right] \tag{3.15}$$

where the expectation is on \hat{f} .

As mentioned earlier, the first term is not necessary when choosing between competing models. Thus, minimizing the expected amount of K-L information equates to maximizing the second term. Multiplying the second term by -2 , the purpose function¹³ becomes:

$$A = -2E \left[\int f_0(\Delta \mathbf{x}) \log \hat{f}(\Delta \mathbf{x}) d\Delta \mathbf{x} \right] \tag{3.16}$$

where the integral is based on the f_0 , which is the true the population random variable $\Delta \mathbf{X}$. As mentioned earlier, the true probability density function will never be known. However, estimates are made based on the sample drawn from $\Delta \mathbf{X}$, $\Delta \tilde{\mathbf{x}}$. Hence, Eq. (3.16) can be rewritten as:

$$A = -2E [\log \hat{f}(\Delta \tilde{\mathbf{x}})] \tag{3.17}$$

Here, $\Delta \tilde{\mathbf{x}}$ is an independent sample drawn from the random variable $\Delta \mathbf{X}$ with a probability density function of f_0 . Thus, A represents the expected estimate of log-likelihood. Burnham and Anderson (2002) argue that models with low A values have good out-of-sample log-likelihood.

$$\mathbf{v}_t = \Delta \mathbf{x}_t - \mathbf{c}_0 - \Pi_{\tilde{\mathbf{x}}} \tilde{\mathbf{z}}_{t-1} - \Lambda_0 \Delta \mathbf{x}_t^* - \sum_{\ell=1}^{p-1} \Lambda_{\ell} \Delta \mathbf{z}_{t-\ell}$$

and when the log-likelihood function of the conditional model given in Eq. (3.10) is rewritten as:

$$l(\theta) = \frac{-kT}{2} \log(2\pi) - \frac{T}{2} \log |\Sigma_{vv}| - \frac{1}{2} \sum_{t=1}^T \mathbf{v}'_t \Sigma_{vv}^{-1} \mathbf{v}_t \tag{3.18}$$

The maximum value of the log-likelihood function of the system is obtained with the parameters that are actually correct:

¹² See Gourieroux and Monfort (1995) for further details.

¹³ The multiplication of log-likelihood by -2 is similar to parameter sizing of χ^2 .

$$l(\theta_0) = \frac{-kT}{2} \log(2\pi) - \frac{T}{2} \log |\Sigma_{vv,0}|$$

Therefore, the desired value of the A function is:

$$A_0 = kT \log(2\pi) + T \log |\Sigma_{vv,0}| \tag{3.19}$$

This will be the value obtained if there is no estimation error.¹⁴

Lemma 3 *If the conditional VEC model given in Eq. (3.9) with a parameter count of K and variance–covariance matrix $\Sigma_{vv,0}$ is actually estimated with $\hat{f}(\Delta x)$ while the correct model is f_0 ,*

$$A = kT(\log(2\pi) + 1) + T \log |\Sigma_{vv,0}| + K \tag{3.20}$$

and

$$E[-2\ell_T(\hat{\theta})] = kT(\log(2\pi) + 1) + T \log |\Sigma_{vv,0}| - K \tag{3.21}$$

The proof of Lemma 3 is provided in Appendix.

The target function A in Eq. (3.20) appears to exceed the desired value of A_0 by K parameters. The second term represents the cost of an additional parameter estimate, measured by the expected amount of K-L information. When parameters are estimated instead of using actual parameter values, the K-L amount increases linearly with the number of parameters, K .

This result is represented differently in Eq. (3.21). The sample log-likelihood function is less than the desired value of A_0 by K parameters. This represents the cost of over-fitting within the sample. The sample log-likelihood is a measure of in-sample compatibility and is therefore less than the population’s log-likelihood. Taken together, these two statements indicate that the target value of the sample logarithmics is $2K$ less than A . This reflects the combined cost of over-fitting and parameter estimation.

By combining expression Eqs. (3.20) and (3.21), we get:

$$AIC = -2 \times (\log - likelihood) + 2 \times (number\ of\ parameter)$$

which forms the general structure of the AIC. The *likelihood* is obtained by estimating parameters, where $-2 \times (\log - likelihood)$ signifies *the goodness of fit* and $2 \times (number\ of\ parameter)$ is considered the penalty term.¹⁵

Theorem 1 *GVAR-Akaike information criterion (GAIC): In the VEC model given in Eq. (3.9), under the validity of Lemma 1–3, $\{\mathfrak{M}_{\tilde{k}_{(N)}}: \tilde{k}_{(N)} = \tilde{k}_{(N),1}, \tilde{k}_{(N),2}, \dots, \tilde{k}_{(N),K}$*

¹⁴ The population variance-covariance matrix is unknown, but calculations can be easily shown if $\Sigma_{vv,0}$ is known.

¹⁵ See Cavanaugh and Neath (2019) for further details.

$\in \mathbb{Z}^+, N \in \mathbb{N}$ is a set of all competing models generated according to the country/regional structure and

$$\begin{aligned} \text{GAIC}(\tilde{k}_{(N)}) &= -2 \log \ell \left(\hat{\theta}_{\tilde{k}_{(N)}} \right) + 2\tilde{k}_{(N)} \\ &= k(\log(2\pi) + 1) + \log \left| \hat{\Sigma}_{vv,0} \right| + \frac{2\tilde{k}_{(N)}}{T} \end{aligned} \tag{3.22}$$

is an estimator of twice the target function given in Eq. (3.21) minimized to select an $\mathfrak{M}_{\tilde{k}_{(N)}}$ model from among the competing models.

k is the number of endogenous variables and $\tilde{k}_{(N)}$ is the number of parameters that vary depending on the country/region dimension N .

The proof of Theorem 1 is given in the Appendix. All models are estimated, and the one with the lowest GAIC value is considered the best. The advantages of AIC include its ease of calculation, implementation, and interpretation. Essentially, it provides an intuitive approach to model selection, accounting for the cost (or penalty) of each additional parameter estimate.

Despite these advantages, the criterion has well-documented disadvantages in the literature. Cavanaugh and Neath (2019) argue that when the number of parameters is larger than the number of observations, the penalty term on the right-hand side of the criterion is significantly affected, leading to biased results. This situation is particularly evident when the sample size is increased. Shibata (1989) emphasizes that the AIC criterion is sensitive to the number of parameters, and Linhart and Zucchini (1986) claim that the criterion can be heavily biased towards incorrect model selection as the number of parameters increases.

The GAIC criterion, developed for the GVAR modelling approach, is also likely to face similar challenges. This is particularly true as the number of individual country/region models increases, along with the number of endogenous and, therefore, exogenous variables, causing the size of the parameter vector to exceed the number of observations. This feature is particularly prominent in macroeconomic analyses and in datasets used in the GVAR modelling approach. The finite sample characteristics observed for the Akaike-type criteria were also observed for the GAIC criterion. Section 3.2 proposes an *ad hoc* solution for the GAIC criterion based on these observations.

3.2 An Ad hoc Modification to Information Criterion

For the GVAR model with a number of parameters, K , GAIC in Eq. (3.22) is:

$$\text{GAIC}(K) = k(\log(2\pi) + 1) + \log \left| \hat{\Sigma}_{vv,0} \right| + \frac{2K}{T} \tag{3.23}$$

and the row-dimension of the vector of parameters is defined as:

$$K = \left[k \times [k^* + (k + k^*)(p - 1)] + kr + (k + k^*) \times r + \left(\frac{k(k + 1)}{2} \right) \right] \tag{3.24}$$

The total number of endogenous and exogenous variables in the GVAR system is defined by average values as follows:

$$k = \sum_{i=1}^N k_i = N\bar{k}$$

and

$$k^* = \sum_{i=1}^N k_i^* = N\bar{k}^*$$

So we have

$$\text{GAIC}(K) = N\bar{k}(\log(2\pi) + 1) + \log \left| \hat{\Sigma}_{vv,0} \right| + \frac{2K}{T} \tag{3.25}$$

From this point of view, applying same approach to K in Eq. (3.24) obtained the following equation:

$$K = (p - 1)N^2\bar{k}^2 + pN^2\bar{k}\bar{k}^* + 2rN\bar{k} + rN\bar{k}^* + \frac{1}{2}N^2\bar{k}^2 + \frac{1}{2}N\bar{k} \tag{3.26}$$

Thus, the $\text{GAIC}(K)$ criterion given in Eq. (3.25) is defined in terms of three different functions as follows:

$$\text{GAIC}(K) = f_1(N, \bar{k}) + f_2(\hat{\Sigma}_{vv,0}) + f_3(N^2, N, p, r, \bar{k}, \bar{k}^*, 2T^{-1})$$

Firstly, we consider the function $f_2(\hat{\Sigma}_{vv,0})$ that takes into account the stochastic structure. This function is defined as follows:

$$f_2(\hat{\Sigma}_{vv,0}) = \log \left| \hat{\Sigma}_{vv,0} \right|$$

where $\hat{\Sigma}_{vv,0}$ be a matrix of $k \times k$ dimension. The estimate of the variance–covariance matrix,

$$\hat{\Sigma}_{vv,0} = \sum_{t=1}^T \hat{v}_t \hat{v}_t' / T \tag{3.27}$$

is calculated from the estimated residuals of country/region specific models.

If the total number of endogenous variables in the GVAR model, $k = N\bar{k}$, is greater than time dimension T , $\hat{\Sigma}_{vv,0}$ may not be positively defined.¹⁶ This will be a serious problem in the calculation of the criteria.

The matrix is given in Eq. (3.27) is the estimation of the large-dimension variance-covariance matrices mentioned by Bailey et al., (2019), is in keeping with the problems and solutions that are being encountered. It can be defined as follows:

$$\hat{\Sigma}_{vv,0} = \begin{pmatrix} \widehat{cov}(\hat{v}_{1t}, \hat{v}_{1t}) & \widehat{cov}(\hat{v}_{1t}, \hat{v}_{2t}) & \dots & \widehat{cov}(\hat{v}_{1t}, \hat{v}_{Nt}) \\ \widehat{cov}(\hat{v}_{2t}, \hat{v}_{1t}) & \widehat{cov}(\hat{v}_{2t}, \hat{v}_{2t}) & \dots & \widehat{cov}(\hat{v}_{2t}, \hat{v}_{Nt}) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{cov}(\hat{v}_{Nt}, \hat{v}_{1t}) & \widehat{cov}(\hat{v}_{Nt}, \hat{v}_{2t}) & \dots & \widehat{cov}(\hat{v}_{Nt}, \hat{v}_{Nt}) \end{pmatrix} \tag{3.28}$$

One solution recommendation is to be restricted to the diagonal elements of this matrix. In addition, PSW states that the GVAR modelling approach enables weak cross-sectional dependency, $N \rightarrow \infty$,

$$\frac{\sum_{i,n=1}^N \sigma_{in,hh'}}{N} \rightarrow 0$$

In the framework of these approaches, to derive an *ad-hoc* modification of GAIC, a transformation of the function $f_2(\hat{\Sigma}_{vv,0})$ is applied as follows.¹⁷

$$f_2^{ad-hoc}(Bdiag\hat{\Sigma}_{vv,0}) = \frac{\sum_{i=1}^N \log|\hat{\Sigma}_{v_i v_i, 0}|}{N} \tag{3.29}$$

Secondly, we examine $f_1(N, \bar{k})$ and $f_3(N^2, N, p, r, \bar{k}, \bar{k}^*, 2T^{-1})$ functions, which are deterministic structures depending on N . To obtain *ad-hoc* versions of these functions, we rearrange them to remove the cross-sectional dimension effect so that they do not depend on N . Thus, these functions can be obtained as follows:

$$f_1^{ad-hoc}(\bar{k}) = \frac{N\bar{k}(\log(2\pi) + 1)}{N} = \bar{k}(\log(2\pi) + 1) \tag{3.30}$$

and

$$f_3^{ad-hoc}(\bar{K}) = \frac{\varrho \times \dim(K)}{N} \tag{3.31}$$

where *dim* denote dimension expression and $\varrho = 2/T$.

Finally, $GAIC(K)^{ad-hoc}$ is obtained as the sum of these three functions as follows:

¹⁶ See Smith and Galesi (2014) for further details.

¹⁷ *Bdiag* means that non-diagonal elements are restricted and the main diagonal matrices are taken as they were given in (3.32).

$$\text{GAIC}(K)^{ad-hoc} = f_1^{ad-hoc}(\bar{k}) + f_2^{ad-hoc}(B\text{diag}\hat{\Sigma}_{vv,0}) + f_3^{ad-hoc}(\bar{K}) \quad (3.32)$$

4 Small Sample Evidence

This section examines the finite sample performances of the model selection criteria previously defined as GAIC and $\text{GAIC}(K)^{ad-hoc}$, respectively. We adopt an approach similar to that of Fujikoshi and Satoh (1997). In this approach, we consider the performance of the information criteria when the country dimension of the true model is larger (smaller) than the country dimension of the candidate models. The objective of this approach is to generate competing models against the true model in all circumstances, whether the dimension of the model parameters is large or small. Our main aim is to measure the sensitivity of these criteria across all possible model alternatives, as might be encountered in empirical studies. We will first introduce the setting, then present and discuss the findings.

4.1 Design of Simulation

The Data Generation Process (DGP) is based on four main scenarios. Our aim is to simulate events commonly found in macroeconometric applications when designing these scenarios. The primary reason for this approach is the variability in the structure and content of the cross-sectional country dimension within the same dataset in the GVAR model, depending on the *target* country whose economy is being analyzed. For instance, when considering the economy of *target* country *A*, some countries need to be modeled individually, while others within the same dataset must be aggregated. Conversely, the content and composition of the cross-sectional dimension should be changed when the aim is to analyze the *target* country *B* in the same dataset.¹⁸ As a result, our primary focus is on *target* and *reference* countries, as stated in Sect. 1. The following scenarios will outline the main considerations for the formulation of the DGPs:

Scenarios

1. The size of *target* country's economy is smaller than that of other countries. Furthermore, the relationship between *target* and *reference* country is stronger than the relationship between other countries and *reference* country
 2. The size of *target* country's economy is smaller than that of other countries. Furthermore, the relationship between *target* and *reference* country is weaker than the relationship between other countries and *reference* country
 3. The size of *target* country's economy is smaller than that of *reference* country, but it is bigger than other countries. Furthermore, the relationship between *target* and *reference* country is stronger than the relationship between other countries and *reference* country
-

¹⁸ *Country* is used to denote the *cross-section unit* size of simulated GVAR model.

Scenarios

4. The size of *target* country's economy is smaller than that of *reference* country, but it is bigger than other countries. Furthermore, the relationship between *target* and *reference* country is weaker than the relationship between other countries and *reference* country

The DGPs are created in response to the aforementioned scenarios.. The cross-sectional unit dimension, which represent the countries, is numbered and labeled as follows: $i = 1, 2, 3, 4, 5, 6, 7$. Four different DGPs are simulated in the following direction:

DGP1: The economy of the *target* country, labeled as 6, is smaller than all the other country's economies ($6 < 1 < 2 < 3 < 4 < 5 < 7$). The relationship between the *target* country (label 6) and countries labeled as 1, 2, 3, 4, and 5 is weak, while the relationship between *reference* country (label 7) and *target* country is strong.

DGP2: The economy of the *target* country, labeled as 6, is smaller than all the other countries' economies ($6 < 1 < 2 < 3 < 4 < 5 < 7$). The relationship between the *target* country and the other countries is strong, but the relationship between the *target* country and the *reference* country (label 7) is weak.

DGP3: The economy of the *target* country, labeled as 5, is the second largest economy after the *reference* country's economy ($6 < 1 < 2 < 3 < 4 < 5 < 7$). The relationship between the *target* country (label5) and countries labeled as 1, 2, 3, 4 and 6 is weak, while its relationship with the *reference* country (label 7) is strong .

DGP4: The economy of the *target* country, labeled as 5, is the second largest economy after the *reference* country's economy ($6 < 1 < 2 < 3 < 4 < 5 < 7$). The relationship between the *target* country (label 5) and countries labeled as 1, 2, 3, 4, and 6 is strong, but its relationship with the reference country (label 7) is weak.

A common feature across all DGPs is that the economic size ranking of the countries holds for all DGPs. Table 1 illustrates this situation by using predetermined values assigned to the variable D . The values assigned to this variable only serve to provide an order of magnitude¹⁹: $6 < 1 < 2 < 3 < 4 < 5 < 7$.

The predetermined variable "trade" has been generated to represent the link matrix \mathbf{W}_i , as described in Sect. 2.2. In **DGP1**, the relationship between the *target* country (label 6) and the countries labeled 1, 2, 3, 4 and 5 is weak. However, its relationship with the *reference* country's economy (label 7) is stronger, as shown in Table 2 below. For example, the amount of artificial trade carried out by the *target* country (label 6) has conducted with the *reference* country (label 7) is 800.

After that, trade weights are calculated using these values, which are presented in Table 3.²⁰ As can be seen from Table 3, the importance value of the *reference* country (label 7) for the *target* country (label 6) is 53%. Thus, the *reference* country is more important for the *target* country than other countries are.

As can be seen from Table 4, the relationship between the *target* country (label 6) and the countries labeled 1, 2, 3, 4, and 5 is strong, while the relationship with the

¹⁹ Arbitrary assignment of values, provided that the ranking remains the same, will not cause any changes in calculations.

²⁰ Due to rounding errors, the sum of some cells may not be exactly one.

Table 1 Predetermined values of D

i	1	2	3	4	5	6	7
D	10,000	20,000	30,000	40,000	50,000	5000	70,000

reference country's economy (label 7) is weak in **DGP2**. The amount of artificial trade carried out by the *target* country (label 6) has conducted with the *reference* country (label 7) is 50.

The trade weights are calculated using these values, which are shown in Table 5. In **DGP 2**, the importance of the *reference* country (label 7) for the *target* country (label 6) is 7%. The importance of the *reference* country for the *target* country is less than the importance of the other countries.

In **DGP3**, we use the trade weights which obtained in **DGP2**. The relationship between the *target* country (label 5) in **DGP3** and the countries labeled 1, 2, 3, 4 and 6 is weak, but the relationship is strong with the *reference* country's economy (label 7) is 27%, as shown in Table 5.

In Table 6, the amount of artificial trade carried out by the *target* country (label 5) with the *reference* country (label 7) is 500. In **DGP4**, the relationship between the *target* country (label 5) and the countries labeled 1, 2, 3, 4 and 6 is strong, but the relationship with the *reference* country's economy (label 7) is weak, 13%, as shown in Table 7.

In all DGPs, we examine two endogenous (domestic) variables per countries i (where $i = 1, \dots, N$ and $k_i = 2$), denoted as $\mathbf{x}_{it} = (x_{i1}, x_{i2})'$. Therefore, the dimension of the vector of all endogenous variables, \mathbf{x}_t , is $(2N \times 1)$, and is denoted as $\mathbf{x}_t = (\mathbf{x}'_{1t}, \mathbf{x}'_{2t}, \dots, \mathbf{x}'_{Nt})'$. The initial value of all variables for all countries is set to zero, $\mathbf{x}_{-101} = 0$. In the simulations, $T \in \{100, 200, 500\}$ and 2000 replications are run for each experiment.

Because DdPS derives an approximation of GVAR to a global factor model, all DGPs were generated from the canonical global factor model, for $t = -101, -100, \dots, 0, 1, 2, \dots, T$,

$$\mathbf{x}_{it} = \delta_{i0} + \Gamma_{if} f_t + \xi_{it} \quad (4.1)$$

and

$$\delta_{i0} \sim IIDN(1, 1),$$

$$\Gamma_{if} = \begin{pmatrix} \gamma_{i11} \\ \gamma_{i21} \end{pmatrix} \sim IID \begin{pmatrix} N(0, 0.50) \\ N(0.5, 0.50) \end{pmatrix}$$

where δ_{i0} and Γ_{if} are vectors of (2×1) constant coefficients and loadings of factor, respectively.

Unobservable global factors vector f_t is generated as

$$f_t = f_{t-1} + \eta_{ft}$$

Table 2 Values for country-level trade relations for **DGP1**

	1	2	3	4	5	6	7
1	0	250	350	450	650	100	800
2	160	0	400	500	700	120	1000
3	200	300	0	550	750	140	1200
4	250	350	450	0	800	150	1400
5	310	400	500	600	0	200	1600
6	350	100	150	250	350	0	600
7	400	500	600	700	1200	800	0

Table 3 Values for country level trade rates for **DGP1**

	1	2	3	4	5	6	7
1	0	0.13	0.14	0.15	0.15	0.07	0.12
2	0.10	0	0.16	0.16	0.16	0.08	0.15
3	0.12	0.16	0	0.18	0.17	0.09	0.18
4	0.15	0.18	0.18	0	0.18	0.10	0.21
5	0.19	0.21	0.20	0.20	0	0.13	0.24
6	0.21	0.05	0.06	0.08	0.08	0	0.09
7	0.24	0.26	0.24	0.23	0.27	0.53	0

Table 4 Values for country-level trade relations for **DGP2 and DGP3**

	1	2	3	4	5	6	7
1	0	250	350	450	650	100	800
2	160	0	400	500	700	120	1000
3	200	300	0	550	750	140	1200
4	250	350	450	0	800	150	1400
5	310	400	500	600	0	200	1600
6	350	100	150	250	350	0	600
7	400	500	600	700	1200	50	0

Table 5 Values for country level trade rates for **DGP2 and DGP3**

	1	2	3	4	5	6	7
1	0	0.13	0.14	0.15	0.15	0.13	0.12
2	0.10	0	0.16	0.16	0.16	0.16	0.15
3	0.12	0.16	0	0.18	0.17	0.18	0.18
4	0.15	0.18	0.18	0	0.18	0.20	0.21
5	0.19	0.21	0.20	0.20	0	0.26	0.24
6	0.21	0.05	0.06	0.08	0.08	0	0.09
7	0.24	0.26	0.24	0.23	0.27	0.07	0

Table 6 Values for country-level trade relations for **DGP4**

	1	2	3	4	5	6	7
1	0	250	350	450	650	100	800
2	160	0	400	500	700	120	1000
3	200	300	0	550	750	140	1200
4	250	350	450	0	800	150	1400
5	310	400	500	600	0	200	1600
6	350	100	150	250	350	0	600
7	400	500	600	700	500	800	0

Table 7 Values for country-level trade rates for **DGP4**

	1	2	3	4	5	6	7
1	0	0.13	0.14	0.15	0.15	0.07	0.12
2	0.10	0	0.16	0.16	0.16	0.08	0.15
3	0.12	0.16	0	0.18	0.17	0.09	0.18
4	0.15	0.18	0.18	0	0.18	0.10	0.21
5	0.19	0.21	0.20	0.20	0	0.13	0.24
6	0.21	0.05	0.06	0.08	0.08	0	0.09
7	0.24	0.26	0.24	0.23	0.27	0.53	0

where $\eta_{f,t} \sim IIDN(0, 1)$, (2×1) vector of country/region specific effects ξ_{it} , $\xi_t = (\xi'_{1,t}, \xi'_{2,t}, \dots, \xi'_{N,t})'$, is generated as

$$\xi_t = (\mathbf{I}_{2N} - \mathbf{\Pi}_\xi) \xi_{t-1} + \mathbf{u}_t \quad (4.2)$$

where $\mathbf{u}_t = (\mathbf{u}'_{1t}, \mathbf{u}'_{2t}, \dots, \mathbf{u}'_{Nt})'$, and $\mathbf{u}_t \sim IIDN_{2N}(0, \mathbf{I})$.²¹ For $i, n = 1, 2, \dots, N$, $(2N \times 2N)$ matrix of long-run relationship, $\mathbf{\Pi}_\xi = \{\Pi_{\xi,in}\}$, is generated as

$$\mathbf{\Pi}_\xi = \alpha_\xi \beta'_\xi$$

where α_ξ means speed of adjustment, $\alpha_\xi = \{\alpha_{\xi,in}\}$, and long-run coefficients are β_ξ , $\beta_\xi = \{\beta_{\xi,in}\}$. Their dimensions are $(2N \times r)$. $\forall i$, we set $r = \sum_{i=1}^N r_i$ and $r_i = 1$.

For validity of following condition,

$$Rank(\mathbf{\Pi}_\xi) = Rank(\alpha_\xi) = Rank(\beta_\xi) = r < 2N$$

$\forall i \in N$, we set

²¹ The features that this model should provide are included in the Supplementary Material.

$$\alpha_{\xi,i} = \begin{pmatrix} -0.5 \\ 0 \end{pmatrix}, \beta_{\xi,i} = \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

$$\alpha_{\xi} = \begin{pmatrix} \alpha_{\xi 1} & 0 & \dots & 0 \\ 0 & \alpha_{\xi 2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_{\xi N} \end{pmatrix}$$

and

$$\beta_{\xi} = \begin{pmatrix} \beta_{\xi 1} & 0 & \dots & 0 \\ 0 & \beta_{\xi 2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \beta_{\xi N} \end{pmatrix}$$

The following description gives the details of the simulation. For $y = 1, 2, 3, 4$, $\mathcal{S}_N^{DGP_y} = \{i : i \in N\}$ shows countries/regions in the y -th DGP . The true model is denoted as TM . For $m = 1, 2, 3, 4, 5$, EM_m refers to the m -th Estimated Model (which means Rival or Candidate Model), and ROW_b refers to the b -th Rest of the World.²² To understand the approach in the simulation and to facilitate the interpretation of the results, we discuss a sample DGP ²³:

- For y -th DGP , TM contains 4 countries/regions. These are ROW , country 5, *target* country, number 6, and *reference* country, number 7, respectively:

$$\mathcal{S}_{N, TM=4}^{DGP_{y=1}} = \{ \mathcal{S}_{N, ROW_1}^{DGP_{y=1}}, 5, 6, 7 \}.$$

- For y -th DGP , the region ROW consists of aggregation of countries 1, 2, 3, 4:

$$\mathcal{S}_{N, ROW_1}^{DGP_{y=1}} = \{1, 2, 3, 4\}.$$

- For y -th DGP , countries, estimated in first rival model:

$$\mathcal{S}_{N, EM_1}^{DGP_{y=1}} = \{1, 2, 3, 4, 5, 6, 7\},$$

$$\mathcal{S}_{N, ROW_0}^{DGP_{y=1}} = \{ \}.$$

- For y -th DGP , countries, estimated in second rival model:

$$\mathcal{S}_{N, EM_2}^{DGP_{y=1}} = \{ \mathcal{S}_{N, ROW_1}^{DGP_{y=1}}, 3, 4, 5, 6, 7 \},$$

²² The same definitions are made for all DGPs.

²³ All other cases are included in Supplementary Material.

$$\mathcal{S}_{N,ROW_1}^{DGP_{y=1}} = \{1, 2\}.$$

- For y -th DGP , countries, estimated in third rival model:

$$\mathcal{S}_{N,EM_3}^{DGP_{y=1}} = \left\{ \mathcal{S}_{N,ROW_2}^{DGP_{y=1}}, 4, 5, 6, 7 \right\},$$

$$\mathcal{S}_{N,ROW_2}^{DGP_{y=1}} = \{1, 2, 3\}.$$

- For y -th DGP , countries, estimated in fourth rival model:

$$\mathcal{S}_{N,EM_4}^{DGP_{y=1}} = \left\{ \mathcal{S}_{N,ROW_3}^{DGP_{y=1}}, 5, 6, 7 \right\},$$

$$\mathcal{S}_{N,ROW_3}^{DGP_{y=1}} = \{1, 2, 3, 4\}.$$

- For y -th DGP , countries, estimated in fifth rival model:

$$\mathcal{S}_{N,EM_5}^{DGP_{y=1}} = \left\{ \mathcal{S}_{N,ROW_4}^{DGP_{y=1}}, 6, 7 \right\},$$

$$\mathcal{S}_{N,ROW_4}^{DGP_{y=1}} = \{1, 2, 3, 4, 5\}.$$

4.2 The Results of Monte Carlo Experiments

Tables 8, 9, 10, 11, 12, 13 and 14 report the results of the GAIC and $GAIC^{ad-hoc}$ criterion in Monte Carlo experiments with various DGP specifications as mentioned above.²⁴ These tables display the frequencies at which the criteria select correct and incorrect models. The rows indicate the TM , while the columns show the EM . Cells marked with * denote the frequency of correctly determining the model based on the criterion. Other cells display the selection frequency results of the candidate model.

Table 8 reports the results for $DGP_{y=1}$. When the dimension of the TM is 3 and $T = \{100, 200, 500\}$, the GAIC criterion selects this model is approximately 100%. However, for other results marked with *, increasing the dimension of the TM worsens the performance of the GAIC criterion. It also exhibits bias towards selecting the alternative candidate model with 3 countries. This reflects the over-parameterization of an Akaike-based criterion, as described in Sect. 3.2. For example, when the dimension of the TM is 4 and $T = \{100, 200, 500\}$, the value of the criterion increases, favoring the selection of the nearest alternative, which is the candidate

²⁴ In simulation study, we use an open-source Matlab toolbox developed by Smith and Galesi (2014), can be obtained from the web at: <https://sites.google.com/site/gvarmodellng/>. All documents including datasets and Matlab codes on the simulation designs are given in supplementary materials.

Table 8 Average selected model frequencies of GAIC and GAIC^{ad-hoc} in **DGPI**

T = 100											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	90.55*	2.75	0.65	0.55	5.5
4	99.6	0.4*	0	0	0	4	30.55	50.20*	2.40	1.65	15.2
5	100	0	0*	0	0	5	36.65	2.90	42.05*	1.1	14.30
6	100	0	0	0*	0	6	21.80	1.65	1.35	65.4*	9.8
7	100	0	0	0	0*	7	10.65	7.85	35.2	7.05	39.25*
T = 200											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	91*	2.15	0.35	0.1	6.4
4	100	0*	0	0	0	4	54.9	22.30*	2.35	0.90	19.55
5	100	0	0*	0	0	5	58.65	1	20.65*	0.65	18.65
6	100	0	0	0*	0	6	39.75	1.6	0.85	46.2*	11.6
7	100	0	0	0	0*	7	5.15	7.15	40.75	4.90	42.05*
T = 500											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	87*	2.3	0.55	0.15	9.75
4	99.6	0.4*	0	0	0	4	78.9	1.10*	1.85	0.65	17.5
5	100	0	0*	0	0	5	77.55	1	6.85*	0.15	14.4
6	100	0	0	0*	0	6	69.20	1.7	0.85	14.7*	13.55
7	100	0	0	0	0*	7	5.55	12.35	28.40	3.6	50.10*

TM and EM mean true and estimated model, respectively

* highlights that TM coincides with the EM

Table 9 The relationship with countries for $S_{N, TM}^{DGP_{y=1}}$

	ROW	5	6	7
ROW	0	0.65	0.34	0.67
5	0.37	0	0.13	0.24
6	0.17	0.08	0	0.09
7	0.45	0.27	0.53	0

model with 3 countries. More specifically, the GAIC criterion is inherently sensitive to the number of parameters

Table 10 The relationship with countries for $\mathcal{S}_{N,EM_3}^{DGP_{y=1}}$

	ROW	6	7
ROW	0	0.47	0.91
6	0.26	0	0.09
7	0.74	0.53	0

Table 11 The relationship with countries for $\mathcal{S}_{N,DM=4}^{DGP_{y=1}}$ and $\mathcal{S}_{N,DM=4}^{DGP_{y=2}}$

	ROW	5	6	7	ROW	5	6	7	
ROW	0	0.65	0.34	0.67	ROW	0	0.65	0.67	0.67
5	0.37	0	0.13	0.24	5	0.37	0	0.26	0.24
6	0.17	0.08	0	0.09	6	0.17	0.08	0	0.09
7	0.45	0.27	0.53	0	7	0.45	0.27	0.07	0

When the dimension of TM is 3 and $T = \{100, 200, 500\}$, the $GAIC^{ad-hoc}$ results in Table 8 reveal a selection frequency of approximately 91%, similar to the GAIC. If the dimension of TM is 4 and $T = 100$, the selection frequency of TM for GAIC and $GAIC^{ad-hoc}$ is % 4 and % 50, respectively. Increasing T diminishes the selection frequency of TM for $GAIC^{ad-hoc}$ to 51%, 22% and 1.10%. Although similar to GAIC, the rate of decrease is slower. When evaluated in $DGP_{y=2}$, the frequencies are approximately 56%, 37% and 15%, respectively.

Tables 9 and 10 show that, in the case of $DGP_{y=1}$ with TM dimension is 4, the relationship between *target* country (label 6) and the country (label 5) is weak. In contrast, the *target* country (label 6) has a strong relationship with *ROW*. Table 11 reveals that the relationship between *target* country (label 6) and *ROW* increases from 34 to 67%. These observations demonstrate that the strength of relationships between countries significantly influences model selection.

For $TM = 4, 5$ in $DGP_{y=1,2}$, with $T = 100, 200$ and 500, $GAIC^{ad-hoc}$ is sensitive to candidate models with $EM = 3$. In contrast, for $TM = 6, 7$ in $DGP_{y=1,2}$, when $T = 100, 200$ and 500, $GAIC^{ad-hoc}$ is more robust and shows less sensitivity to candidate models with $EM = 3$. Comparing the performance of $GAIC^{ad-hoc}$ for various sets²⁵ reveals some important findings. The most notable difference between $TM = 4, 5$ and $TM = 6, 7$ in $DGP_{y=1,2}$ is the need to consider almost all countries individually. Under these sets, $GAIC^{ad-hoc}$ gives successful results in small sample performances. Therefore, the ad hoc modification for GAIC produces more successful outcomes in situations where the country dimension is large and countries need to be treated individually in the GVAR modelling approach. Indeed, this is also observed for $TM = 5, 6$ in $DGP_{y=3,4}$.

²⁵ $\mathcal{S}_{N, TM=6}^{DGP_{y=1,2}} = \{\mathcal{S}_{N, ROW_1}^{DGP_{y=1,2}}, 3, 4, 5, 6, 7\}$ with $\mathcal{S}_{N, ROW_1}^{DGP_{y=1,2}} = \{1, 2\}$ and $\mathcal{S}_{N, EM=3}^{DGP_{y=1,2}} = \{\mathcal{S}_{N, ROW_3}^{DGP_{y=1}}, 6, 7\}$ with $\mathcal{S}_{N, ROW_3}^{DGP_{y=1,2}} = \{1, 2, 3, 4, 5\}$

Table 12 Average selected model frequencies of GAIC and GAIC^{ad-hoc} in DGP2

T = 100											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	91.2*	3	0.65	0.65	4.5
4	99.75	0.25*	0	0	0	4	26.8	55.8*	2.6	1.5	13.3
5	100	0	0*	0	0	5	36.45	4	44.35*	1.35	14.3
6	100	0	0	0*	0	6	26.20	2.95	1.5	57.45*	11.9
7	100	0	0	0	0*	7	11.45	7.55	30.1	7.15	43.75*
T = 200											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	91*	2.15	0.35	0.1	6.4
4	100	0*	0	0	0	4	54.9	22.30*	2.35	0.90	19.55
5	100	0	0*	0	0	5	58.65	1	20.65*	0.65	18.65
6	100	0	0	0*	0	6	39.75	1.6	0.85	46.2*	11.6
7	100	0	0	0	0*	7	5.15	7.15	40.75	4.90	42.05*
T = 500											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	88*	2.45	0.5	0.2	9.3
4	100	0*	0	0	0	4	67.1	15.1*	2.05	0.90	14.85
5	100	0	0*	0	0	5	77.45	1	6.5*	0.25	14.55
6	100	0	0	0*	0	6	58.75	1.55	1	27.2*	11.5
7	100	0	0	0	0*	7	3.85	13.20	25.65	3.85	53*

TM and EM mean true and estimated model, respectively

* highlights that TM coincides with the EM

A similar result to that of the GAIC criterion presented in Table 8 can be found in Tables 11, 12 and 13. As mentioned earlier, this scenario is consistent with the Akaike literature.

In $DGP_{y=3}$, $T = 100, 200$ and 500 , while for $TM = 5$ the results given in Table 13 show that the selection frequency of TM for GAIC^{ad-hoc} are 87%, 80% and 78%. For $TM = 6$, the results are 90%, 87% and 78%, respectively. The same results are given in Table 14 for $DGP_{y=4}$. For $T = 100, 200$ and 500 , and $TM = 5$, the results indicate that the selection frequency of TM for GAIC^{ad-hoc} are 87%, 76% and 77%. If $TM = 6$, these proportions are 92%, 88% and 77%, respectively. Comparing these findings with the results for $TM = 5, 6$ in $DGP_{y=1,2}$, $T = 100, 200$ and 500 , we conclude that:

Table 13 Average selected model frequencies of GAIC and GAIC^{ad-hoc} in DGP3

T = 100											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	61.7*	1.05	3.3	33.75	0.2
4	99.55	0.45*	0	0	0	4	1.8	27.75*	9	61.25	0.2
5	100	0	0*	0	0	5	5.6	0.9	87.15*	0.4	5.95
6	100	0	0	0*	0	6	0.85	1.5	6.25	90.95*	0.45
7	99.95	0.05	0	0	0*	7	7.95	17.1	14.30	60	0.65*
T = 200											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	51.35*	0.45	2.25	45.8	0.15
4	100	0*	0	0	0	4	0.5	5.8*	8.35	85.25	0.1
5	100	0	0*	0	0	5	9.85	0	80.05*	0.05	9.65
6	100	0	0	0*	0	6	0.45	1.50	9.85	87.85*	0.35
7	100	0	0	0	0*	7	0.6	17	19.1	62.4	0.9*
T = 500											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	26.05*	0.55	4.55	68.85	0
4	100	0*	0	0	0	4	0.6	0.35*	8	91	0.05
5	100	0	0*	0	0	5	7.80	0	78.25*	0.25	13.3
6	100	0	0	0*	0	6	1	2.55	17.75	78.45*	0.25
7	100	0	0	0	0*	7	0.85	14.1	23.7	61	0.35*

TM and EM mean True and Estimated model, respectively

* highlights that TM coincides with the EM

- An increase in the economic size of the *target* country,
- A decrease in the economic size of the *ROW*,

can enhance the successful of the GAIC^{ad-hoc}.

We also examine that the structure used in these DGPs,

$$S_{N, TM=5}^{DGP_{y=1,2}} = \left\{ S_{N, ROW_{TM=5}}^{DGP_{y=1,2}}, 4, 5, 6, 7 \right\},$$

$$S_{N, TM=5}^{DGP_{y=3,4}} = \left\{ S_{N, ROW_{TM=5}}^{DGP_{y=3,4}}, 3, 4, 5, 7 \right\},$$

Table 14 Average selected model frequencies of GAIC and GAIC^{ad-hoc} in DGP4

T = 100											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	61.5*	0.95	3.35	33.95	0.25
4	99.5	0.5*	0	0	0	4	1.7	28.05*	8.7	61.25	0.3
5	100	0	0*	0	0	5	5.80	0.85	87.05*	0.4	5.9
6	100	0	0	0*	0	6	0.80	1.45	5.90	91.55*	0.3
7	99.95	0.05	0	0	0*	7	6.10	16.25	14.45	62.2	1*
T = 200											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	51.3*	0.5	2.2	45.95	0.05
4	100	0*	0	0	0	4	0.65	5.95*	8.5	84.80	0.1
5	100	0	0*	0	0	5	9.5	0	76.25*	0.2	13.6
6	100	0	0	0*	0	6	0.5	1.55	9.25	88.4*	0.3
7	100	0	0	0	0*	7	0.75	15	22.1	61.55	0.6*
T = 500											
GAIC						GAIC ^{ad-hoc}					
TM/EM	3	4	5	6	7	TM/EM	3	4	5	6	7
3	100*	0	0	0	0	3	25.7*	0.6	4.7	69	0
4	100	0*	0	0	0	4	0.55	0.4*	8	90.9	0.15
5	100	0	0*	0	0	5	8.05	1	77.1*	0.25	14
6	100	0	0	0*	0	6	1.20	2.15	19.6	76.9*	0.15
7	100	0	0	0	0*	7	0.85	15.90	22.95	59.85	0.45*

TM and EM mean true and estimated model, respectively

* highlights that TM coincides with the EM

and

$$S_{N,ROW_{TM=5}}^{DGP_{y=1,2}} = \{1, 2, 3\},$$

$$S_{N,ROW_{TM=5}}^{DGP_{y=3,4}} = \{1, 2, 6\},$$

and the following transformation,

$$f_2^{ad-hoc} \left(Bdiag \hat{\Sigma}_{vv,0} \right) = \frac{\sum_{i=1}^N \log \left| \hat{\Sigma}_{v_i v_i, 0} \right|}{N},$$

We present some concluding remarks. For a country with a large economy, shocks occurring in countries with smaller economies are often too inconsequential to warrant individual treatment. Thus, these countries can be aggregated. The economic size of the *target* country serves as a crucial indicator in the aggregation process of smaller economies. This approach allows us to obtain the appropriate model selection without having to estimate numerous models. A similar scenario has been observed under conditions where the 6-country model is correct. Furthermore, our findings are consistent with those obtained in $DGP_{y=3,4}$ and $DGP_{y=1,2}$ where $TM = 3$. It's also worth noting that the size of the *ROW* region is also an important indicator for the *target* country.

Considering the case where $TM = 3$,

$$S_{N, TM=3}^{DGP_{y=1,2}} = \left\{ S_{N, ROW_{TM=3}}^{DGP_{y=1,2}}, 6, 7 \right\},$$

$$S_{N, TM=3}^{DGP_{y=3,4}} = \left\{ S_{N, ROW_{TM=3}}^{DGP_{y=3,4}}, 5, 7 \right\},$$

and

$$S_{N, ROW_{DM=3}}^{DGP_{y=1,2}} = \{1, 2, 3, 4, 5\},$$

$$S_{N, ROW_{DM=3}}^{DGP_{y=3,4}} = \{1, 2, 3, 4, 6\}.$$

In $DGP_{y=1,2}$ while the country (label 5) with the *second largest economy* is part of the *ROW* region, the economy (label 6) with the weakest economy is now less important. As mentioned above, the modification made with $\int_2^{ad-hoc} (Bdiag \hat{\Sigma}_{vv,0})$ inherently depends on the importance of the shocks from the countries. The results observed in the case of $TM = 4$ are similar. When constructing the model for the *target* country, the difference in economic size between the *ROW* region and the *target* country should not be too large.

In $DGP_{y=3,4}$ for $TM = 7$, the selection of the candidate model $EM = 6$ stands out. For this case, we have

$$S_{N, TM=7}^{DGP_{y=3,4}} = \{1, 2, 3, 4, 5, 6, 7\},$$

$$S_{N, EM=5}^{DGP_{y=3,4}} = \left\{ S_{N, ROW_{EM=5}}^{DGP_{y=3,4}}, 2, 3, 4, 5, 7 \right\},$$

and

$$S_{N, ROW_{TM=7}}^{DGP_{y=3,4}} = \{\},$$

$$S_{N, ROW_{EM=5}}^{DGP_{y=3,4}} = \{1, 6\}.$$

In the candidate model alternative, the *ROW* region comprises countries with the smallest economic size, labelled as ‘1’ and ‘6’. Looking at Table 7, we can see that the relationship between the *target* country (label 5) and the country (label 1), is stronger than the relationship between the *target* country, labelled with 5, and the *reference* country (label 7). Moreover, country (label 6) is the second most associated country with country (label 1). In cases where individual country models need to be considered, *GAIC^{ad-hoc}* shows sensitivity to the complexity of the relationships between countries. Even though it may not be quantifiable, the involvement of countries outside the target country, due to factors such as geographical reasons, cannot be filtered out in the GVAR modelling approach. This affects the criteria.

5 Empirical Application: How Many Countries Should be Individually Included in GVAR Models for Developing Countries?

In what follows, we conduct an analysis to demonstrate the usefulness and effectiveness of model selection criteria by investigating the structure of modelling in developing countries within a GVAR framework. Our objective is to highlight the empirical performance of our approach. In doing so, we analyse each developing country individually. We propose a modelling strategy for constructing dynamic multi-country frameworks that consider the structure of trade relations and draws upon the existing GVAR literature. Before presenting our proposed modelling framework, we provide a brief overview of the methodology and data employed.

5.1 Data and Variables

In our econometric analysis, we utilize the updated GVAR-DdPS dataset, which includes data from 33 countries and covers the period from 1979Q2 to 2019Q4. This dataset is drawn from the dataset of Mohaddes and Raissi (2020).²⁶ Table 15 presents the list of countries in the updated GVAR-DdPS database.

The primary focus of this paper lies on developing countries, specifically: Argentina, Brazil, Chile, China, India, Indonesia, Korea, Malaysia, Mexico, Peru, Philippines, Saudi Arabia, South Africa, Thailand, and Turkey. Our modelling strategy is to conduct an analysis and derive results for each country individually.

For country i during the period t , the updated GVAR dataset comprises of quarterly macroeconomic and financial variables. These variables include log real GDP (y_{it}), the inflation rate (p_{it}), short-term interest rate (ρ_{it}^S), long-term interest rate (ρ_{it}^L), log deflated exchange rate (e_{it}), log real equity prices (q_{it}), and quarterly data on oil prices ($poil_t$). Consequently, the vector of endogenous variables, \mathbf{x}_{it} , contains the following variables:

$$\begin{aligned} y_{it} &= \ln(GDP_{it}/CPI_{it}), p_{it} = \ln(CPI_{it}), q_{it} = \ln(EQ_{it}/CPI_{it}), \\ e_{it} &= \ln(E_{it}), \rho_{it}^S = 0.25 \times \ln(1 + R_{it}^S/100), \rho_{it}^L = 0.25 * \ln(1 + R_{it}^L/100) \end{aligned}$$

²⁶ For further details, see <https://www.mohaddes.org/gvar>.

Table 15 Countries in the GVAR Model

<i>Asia and Pacific</i>	<i>North America</i>	<i>Europe</i>
Australia	Canada	Austria
China	Mexico	Belgium
India	United States (US)	Finland
Indonesia		France
Japan	<i>South America</i>	Germany
Korea	Argentina	Italy
Malaysia	Brazil	Netherlands
New Zealand	Chile	Norway
Philippines	Peru	Spain
Singapore		Sweden
Thailand	<i>Middle East and Africa</i>	Switzerland
	Saudi Arabia	Turkey
	South Africa	United Kingdom

where GDP_{it} is the nominal gross domestic product, CPI_{it} is the consumer price index, EQ_{it} is the nominal equity price index, E_{it} is US dollars exchange rate, R_{it}^S is the short rate, and R_{it}^L is the long rate.

We use country-specific trade-weighted averages to calculate country-specific foreign variables:

$$\mathbf{x}_{it}^* = \sum_{j=1}^N w_{ij} \mathbf{x}_{jt}, w_{ii} = 0$$

where, as mentioned above, w_{ij} is a set of weights such that $\sum_{j=1}^N w_{ij} = 1$. For the empirical applications, the trade weights are calculated as the average of the years 2014–2016:

$$w_{ij} = \frac{T_{ij,2014} + T_{ij,2015} + T_{ij,2016}}{T_{i,2014} + T_{i,2015} + T_{i,2016}}$$

where $T_{ij,t}$ represents the bilateral trade between country i and country j in a specific year t ; it is calculated as the average of country i 's exports and imports with country j , $T_{i,t} = \sum_{j=1}^N T_{ij,t}$ for $t = 2014, 2015, 2016$. In the supplementary materials, we provide concise information and specifications of the individual models for developing countries.

5.2 Modelling Strategy and Results

We have formulated a strategy for constructing competing models to analyze individual countries. The key components of this strategy are as follows: (i) a developing country's trade with countries in the 33-country GVAR should not be less than 50% of its total trade with all trading partners, (ii) when a developing

Table 16 Argentina's cumulative trade weights with trading partners

Country*	$\sum_{j=1}^N w_{ij}$	Country*	$\sum_{j=1}^N w_{ij}$
Brazil	0.266	United Kingdom	0.927
China	0.433	Malaysia	0.940
Euro**	0.588	Peru	0.951
US [!]	0.716	South Africa	0.960
Chile	0.751	Australia	0.969
India	0.780	Saudi Arabia	0.978
Mexico***	0.807	Turkey	0.984
Japan	0.827	Philippines	0.990
Canada	0.847	Sweden	0.993
Korea	0.864	Singapore	0.996
Switzerland	0.880	Norway	0.999
Indonesia	0.896	New Zealand	1.000
Thailand	0.912		

(*) denotes that the trade weights are arranged in descending order, from largest to smallest. (**) indicates the threshold where a minimum value of 50% is attained. (!) shows that the requirement of individual US participation is satisfied. (***) signify the fulfillment of the maximum country size condition, where Argentina's trade with trading partners exceeds 80%

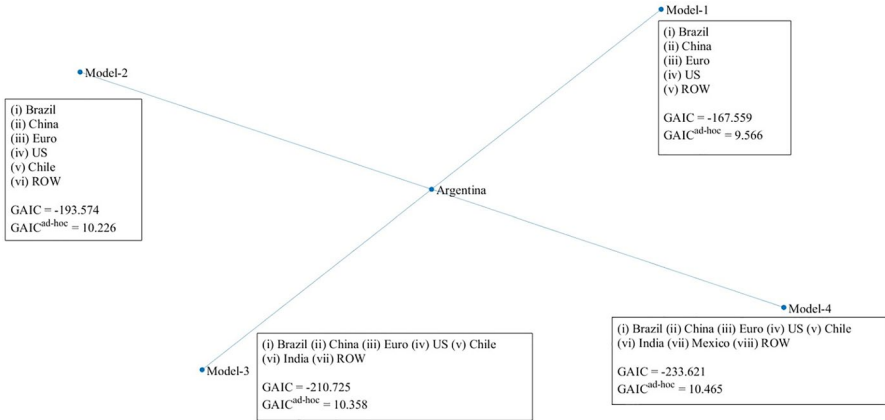


Fig. 1 Competing models for Argentina based on trade relations

country's trade with its trading partners exceeds 80%, we build a model with the largest country size, and (iii) the US is always included as a separate entity in these models. For demonstration purposes, Table 16 shows the cumulative trade weights with other countries for Argentina, as an example from the GVAR DdPS dataset.

<i>GAIC</i>	ARGENTINA	BRAZIL	CHINA	CHILE	INDIA	INDONESIA	KOREA	MALAYSIA	MEXICO	PERU	PHILIPPINES	SOUTH AFRICA	SAUDI ARABIA	THAILAND	TURKEY	AUSTRALIA	CANADA	JAPAN	NORWAY	NEW ZEALAND	SINGAPORE	SWEDEN	SWITZERLAND	UK	US	EURO	
ARGENTINA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
BRAZIL	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
CHINA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
CHILE	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
INDIA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
INDONESIA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
KOREA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
MALAYSIA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
MEXICO	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
PERU	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
PHILIPPINES	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
S. AFRICA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
S. ARABIA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
THAILAND	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
TURKEY	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange

Note: The yellow and orange boxes represent the countries chosen and not chosen to be included in the model, respectively.

Fig. 2 The results of GAIC-model selection criterion for developing countries

<i>GAIC^{ad-hoc}</i>	ARGENTINA	BRAZIL	CHINA	CHILE	INDIA	INDONESIA	KOREA	MALAYSIA	MEXICO	PERU	PHILIPPINES	SOUTH AFRICA	SAUDI ARABIA	THAILAND	TURKEY	AUSTRALIA	CANADA	JAPAN	NORWAY	NEW ZEALAND	SINGAPORE	SWEDEN	SWITZERLAND	UK	US	EURO	
ARGENTINA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
BRAZIL	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
CHINA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
CHILE	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
INDIA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
INDONESIA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
KOREA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
MALAYSIA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
MEXICO	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
PERU	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
PHILIPPINES	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
S. AFRICA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
S. ARABIA	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
THAILAND	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
TURKEY	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange

Note: The yellow and orange boxes represent the countries chosen and not chosen to be included in the model, respectively.

Fig. 3 The results of GAIC^{ad-hoc}-model selection criterion for developing countries

Using the information given in Table 16, Fig. 1 illustrates the identification of models for Argentina utilizing this strategy.

According to the modelling strategy based on trade relations, four different competing models are developed. These models consist of *ss*countries represented by square boxes. Model 1 comprises Brazil, China, the Euro area, the US, and the ROW because of the minimum two conditions are satisfied. The Euro area, as in the DdPS, comprises

eight out of the eleven countries that initially adopted the Euro on 1 January 1999. These are Germany, France, Italy, Spain, the Netherlands, Belgium, Austria, and Finland. The ROW region is formed by combining all the remaining countries listed in Table 15, except Brazil, China, the Euro area, and the US. Similarly, in Model 2, the ROW region is formed by combining all the other countries in Table 15, except Brazil, China, the Euro area, the US, and Chile. Model 3 shows that India, like Chile, is added to the model individually. Consequently, using a comparable methodology, the fourth model, is constructed, and Mexico is included as the final individually modelled country, meeting the maximum country size condition (exceeding 80%). Figures 2 and 3 display the results obtained using the above modelling strategy for 15 developing countries, using GAIC and GAIC^{ad-hoc} respectively. The calculated results according to the criteria for these models are shown in the supplementary materials.

On the one hand, the GAIC criterion tends to favor the selection of models corresponding to the largest countries. This tendency aligns with the AIC criterion's behavior in other econometric model structures, such as the lag length selection in VAR models. It is commonly observed that the AIC criterion tends to favor models with a larger number of parameters. This observation is also consistent with the findings presented in the theoretical and simulation experimental sections of this paper.

On the other hand, the GAIC^{ad-hoc} criterion tends to prioritize the selection of models with a fewer number of parameters. This implies that when the trade partnership of developing countries exceeds 50%, sufficient information on the economic structure of these countries is collected. This observation is also consistent with the theoretical framework outlined in the study and the results of the simulation experiments.

Figures 2 and 3 also show findings that reinforce the significance of the US, China, the Euro area, and Japan for developing economies. As major players in the global economy, these developed countries have a significant impact on developing countries through their economic relations and interactions. They play a crucial role as major trading partners, sources of investment, and providers of technology and expertise for many developing countries.

Developing economies have significant links with major economic powers, including the US, China, the euro area and Japan. The US holds a central position in the global economy and is an important trading partner for many developing economies. China, characterized by rapid economic growth, has emerged as a crucial player in global trade. Developing economies have expanded their trade relations with China and integrated themselves into its production chain. The Euro area is also a trading partner for many developing economies. In addition, Japan is a notable trading partner for developing countries, with a robust economy in sectors such as high technology, automobiles and finance. The economic policies of these developed countries, such as interest rates, exchange rates, and trade policies, can have a direct and significant effects on the growth and competitiveness of developing countries.

These relationships with major economic powers present both opportunities and challenges for developing countries. Depending on global economic developments, these developing economies may need to adapt their economic policies and refine their foreign trade strategies. At the same time, these relationships can provide

incentives for developing countries to strengthen their domestic economic foundations and enhance their competitiveness.

6 Conclusion

The main purpose of this study is to propose an approach to the problem of cross-country aggregation within GVAR models. GVAR models are superior to FAVAR models and PVAR models both in terms of modelling flexibility and the ability to aggregate cross-country aggregated estimates. As a result, academic researchers and practitioners tend to favour GVAR models and enjoy the ability to include virtually all countries' data in their models. The availability of vast computational resources, including both the software and hardware, turn the task of multi-country modelling into a trivial one.

Nevertheless, while practicing GVAR modelling, the modeler has the liberty to consider the other economies in the world either individually or within a bundle of countries marked as the 'Rest of the World' (ROW). The common practice is to take into account the close trade partners individually and to bundle the others as ROW. While such an approach seems both realistic and practical, it undermines a deep-rooted principle of econometrics, which is 'parsimony'. In relation to the principle of parsimony, avoidance from model specification errors can be facilitated via proper model selection criteria. Owing to the serious transaction costs in GVAR models, determining the appropriate model at the least cost must be considered a priority for effective day-to-day use of these models. Having noticed the literature gap in GVAR model selection, this study develops, examines and suggests an information criterion for parsimonious model selection in GVARs.

Our proposed information criteria have been developed in two stages. We considered a relatively straightforward Akaike Information Criterion-type criterion in the first stage, which is a criterion that does not distinguish between the cross-section and time dimensions. After observing the dominance of the cross-section dimension on the values of our first criterion, we adjusted it to come up with our second, i.e., proposed, information criterion. The proposed criterion, by construction, is capable of handling the cross-section and time dimensions in a balanced fashion. In that, while determining the appropriate model dimensions, expansion of a GVAR model over either of the dimensions would yield equal degrees of punishment.

Owing to the ad hoc character of the proposed information criterion, it is not feasible to test its qualities by means of formal statistical tests of hypotheses. So, we resort to two separate exercises in which we examine if the criterion works sufficiently well. In the first exercise, the small sample characteristics of the developed criterion are examined through a series of simulation experiments. The basic principle in designing experiments is that the model dimension, which is actually considered correct, both covers a set of competing models and has a subset. Thus, the effects of all possible situations in high-dimensional time series data are reflected, eventually underlining the viability of the proposed

criterion. Namely, the proposed criterion has provided results that coincide with the simulated data generating processes at hand.

The findings out of our first exercise can be summarized as follows: First, the economic size of the target country that is being explained affects the country/region structure of the GVAR model that should be generated. When compared to a country with a large economy and a small country, it is less fragile to be affected by cross-country relations. Secondly, the results of the GVAR modelling approach show sensitivity to the relationship of countries with the country designated as the reference country, regardless of their economic size. The weakening of the relationship with the reference country for the target country allows other countries to participate more individually in the model. Finally, the success of the GVAR modelling approach and in connection with the model selection criteria is higher if the country/region models are individually included in the system rather than the creation of more compact regions that are aggregated.

In the second exercise, based on an updated DdPS dataset of 33 countries from 1979Q2 to 2019Q4, individual GVAR model selection problem is considered. The proposed criterion has successfully yielded the GVAR models that reflect the real-life trade relations of developing countries with their counterparts. In this exercise, the results obtained exhibit a common characteristic: developing countries are more prone to the developments in the US, China, European countries and Japan than the rest of the world. In all these exercises, parsimonious as well as economically intuitive GVAR specifications have been obtained.

All in all, the proposed information criterion of this study seems to be a promising first step in sustaining parsimony in GVARs. Future work may shed more light on a standard set of principles to be applied within the GVAR modelling approach in a way to bring together the forecast performance of models and their degree of parsimony.

Appendix

Proof of Lemma 1

: The log-likelihood function of the unconditional model given in Eq. (3.2) is

$$\ell(\theta_{\Delta z}, \Sigma_{zz}) = -\frac{T(k + k^*)}{2}(\log 2\pi) - \frac{T}{2}\log|\Sigma_{zz}| - \frac{1}{2}\sum_{t=1}^T e_t' \Sigma_{zz}^{-1} e_t \quad (A.1)$$

where $e_t = (e_{xt}', e_{x^*t}')'$, we have

$$e_t' \Sigma^{-1} e_t = e_{xt}' \Sigma^{xx} e_{xt} + 2e_{xt}' \Sigma^{xx*} e_t^* + e_t^{*'} \Sigma^{x^*x^*} e_t^* \quad (A.2)$$

where

$$\Sigma^{-1} = \begin{pmatrix} \Sigma^{xx} & \Sigma^{xx*} \\ \Sigma^{x*x} & \Sigma^{x*x*} \end{pmatrix} \tag{A.3}$$

and

$$|\Sigma| = |\Sigma_{x*x*}| \left| \Sigma_{xx} - \Sigma_{xx*} \Sigma_{x*x*}^{-1} \Sigma_{x*x} \right|$$

From Eq. (3.5), using $v_t = \epsilon_{xt} - \Sigma_{xx*} \Sigma_{x*x}^{-1} \epsilon_{x*t}$, we have

$$\epsilon_t' \Sigma^{-1} \epsilon_t = v_t' \Sigma_{vv}^{-1} v_t + \epsilon_{x*t}' \Sigma_{x*x*}^{-1} \epsilon_{x*t} \tag{A.4}$$

where

$$\Sigma_{vv} = \Sigma_{xx} - \Sigma_{xx*} \Sigma_{x*x*}^{-1} \Sigma_{x*x} \tag{A.5}$$

Furthermore, similar to the process performed in Eq. (3.6):

$$v_t = \epsilon_{xt} - \Lambda_0 \epsilon_{x*t} = \Delta x_t - \Pi_{xx,x*} z_{t-1} - \Lambda_0 \Delta x_t^* - \sum_{\ell=1}^{p-1} \Lambda_\ell \Delta z_{t-\ell} - c_0 - c_1 t \tag{A.6}$$

using the assumption of weak exogeneity, $\Pi_{xx,x*} = \Pi_x$, we also have

$$\ell(\theta_{\Delta z}, \Sigma) = \ell_v(\theta, \Sigma_{vv}) + \ell_*(\theta^* \Sigma_{x*x*}) \tag{A.7}$$

where

$$\begin{aligned} \ell_v(\theta, \Sigma_{vv}) = & -\frac{Tk}{2}(\log 2\pi) - \frac{T}{2} \log |\Sigma_{vv}| - \frac{1}{2} \sum_{t=1}^T \\ & \left(\Delta x_t - \Pi_x z_{t-1} - \Lambda_0 \Delta x_t^* - \sum_{\ell=1}^{p-1} \Lambda_\ell \Delta z_{t-\ell} - c_0 - c_1 t \right)' \Sigma_{zz}^{-1} \\ & \left(\Delta x_t - \Pi_x z_{t-1} - \Lambda_0 \Delta x_t^* - \sum_{\ell=1}^{p-1} \Lambda_\ell \Delta z_{t-\ell} - c_0 - c_1 t \right), \end{aligned}$$

and from Eq. (3.8), we have

$$\begin{aligned} \ell_*(\theta^* \Sigma_{x*x*}) = & -\frac{Tk^*}{2}(\log 2\pi) - \frac{T}{2} \log |\Sigma_{x*x*}| - \frac{1}{2} \sum_{t=1}^T \\ & \left(\Delta x_t^* - \sum_{\ell=1}^{p-1} \Lambda_\ell \Delta z_{t-\ell} - a_{x*0} \right)' \Sigma_{zz}^{-1} \end{aligned}$$

$$\left(\Delta \mathbf{x}_t^* - \sum_{\ell=1}^{p-1} \Lambda_{\ell} \Delta \mathbf{z}_{t-\ell} - \mathbf{a}_{\mathbf{x}*0} \right)$$

Hence, under the validity of the weak exogeneity assumption, the log-likelihood of the VEC ($p - 1$) model in Eq. (3.2) is the sum of the log-likelihood of the models in Eq. (3.8) and Eq. (3.10). Consequently, θ , the conditional model parameter vector, is independent of the marginal model, $\ell_*(\theta^* \Sigma_{y**})$. Finally, MLE estimation of θ in the conditional model is identical to the indirect estimation of parameters in the unconditional model.

Proof of Lemma 2

: Because $f(\Delta \mathbf{x}, \theta)$ is a probability density function,

$$\int f(\Delta \mathbf{x}, \theta) d\Delta \mathbf{x} = 1$$

When the partial derivative is received on both sides of the equality,²⁷ according to θ ,

$$\begin{aligned} \int \frac{\partial f(\Delta \mathbf{x}, \theta)}{\partial \theta} d\Delta \mathbf{x} &= 0 \\ \int \left\{ \frac{1}{f(\Delta \mathbf{x}, \theta)} \right\} \frac{\partial f(\Delta \mathbf{x}, \theta)}{\partial \theta} f(\Delta \mathbf{x}, \theta) d\Delta \mathbf{x} &= 0 \tag{A.8} \\ \int \frac{\partial \log f(\Delta \mathbf{x}, \theta)}{\partial \theta} f(\Delta \mathbf{x}, \theta) d\Delta \mathbf{x} &= 0 \end{aligned}$$

Therefore we have $\int S(\theta) \partial f(\Delta \mathbf{x}, \theta) d(\theta) = 0$ or $E[S(\theta)] = 0$.²⁸

By assumption RC3, the variance of $S(\theta)$ is finite,

$$\begin{aligned} \text{Var}[S(\theta)] &= E[S(\theta)S(\theta)'] \\ &= E \left[\frac{\partial \log f(\Delta \mathbf{x}, \theta)}{\partial \theta} \cdot \frac{\partial \log f(\Delta \mathbf{x}, \theta)}{\partial \theta'} \right] \tag{A.9} \end{aligned}$$

Proof of Proposition 1

: K-L information given in Eq. (3.15) is

$$K - L(\theta_0; \theta) = - \int f(\Delta \mathbf{x}, \theta_0) \log \left(\frac{f(\Delta \mathbf{x}, \theta)}{f(\Delta \mathbf{x}, \theta_0)} \right) d\Delta \mathbf{x} \tag{A.10}$$

where from Eq. (3.2)

$$\theta = \left(\text{vec}(\Lambda)', \text{vec}(\alpha_x)', \text{vec}(\beta_x)', \text{vech}(\Sigma_{vv})' \right)'$$

²⁷ Derivatives and integrals can be replaced by the assumption RC2.

²⁸ The expectation is based on $f(\Delta \mathbf{x})$ s

and

$$\theta_0 = \left(\text{vec}(\Lambda_{00})', \text{vec}(\alpha_{x0})', \text{vec}(\beta_{x0})', \text{vech}(\Sigma_{vv0})' \right)'$$

And as mentioned in Eq. (3.5), since we have $v_{0t} \sim IIDN(0, \Sigma_{vv0})$, and $\frac{f(\Delta\mathbf{x}, \theta)}{f(\Delta\mathbf{x}, \theta_0)}$ is a convex function, using Jensen's inequality,²⁹ for $\log \left[\frac{f(\Delta\mathbf{x}, \theta)}{f(\Delta\mathbf{x}, \theta_0)} \right]$ we get:

$$E \left[\log \left(\frac{f(\Delta\mathbf{x}, \theta)}{f(\Delta\mathbf{x}, \theta_0)} \right) \right] \leq \log E \left[\left(\frac{f(\Delta\mathbf{x}, \theta)}{f(\Delta\mathbf{x}, \theta_0)} \right) \right] \tag{A.11}$$

where

$$\begin{aligned} E \left(\frac{f(\Delta\mathbf{x}, \theta)}{f(\Delta\mathbf{x}, \theta_0)} \right) &= \int \frac{f(\Delta\mathbf{x}, \theta)}{f(\Delta\mathbf{x}, \theta_0)} f(\Delta\mathbf{x}, \theta_0) d\Delta\mathbf{x} \\ &= \int f(\Delta\mathbf{x}, \theta) d\Delta\mathbf{x} = 1 \end{aligned}$$

It is used in (A.11), we finally have

$$\begin{aligned} E \left[\log \left(\frac{f(\Delta\mathbf{x}, \theta)}{f(\Delta\mathbf{x}, \theta_0)} \right) \right] &\leq \log(1) = 0 \\ E[\log f(\Delta\mathbf{x}, \theta)] &\leq E[\log f(\Delta\mathbf{x}, \theta_0)] \end{aligned} \tag{A.12}$$

Proof of Lemma 3

: The unbiasedness of the MLE estimator was obtained by using Eq. (3.18) and from Pesaran et al. (2000), we have

$$\begin{aligned} \int f_0(\Delta\mathbf{x}) \log f(\Delta\mathbf{x}, \hat{\theta}) d\Delta\mathbf{x} &= -\frac{kT}{2} \log 2\pi \\ &\quad - \frac{T}{2} \log \left| T^{-1} \int \Delta\hat{\mathbf{x}} \left(\mathbf{I}_T - \hat{\mathbf{Z}}_{-1}^{*'} \beta_* \right) \right. \\ &\quad \left. \times (\beta_*' \hat{\mathbf{Z}}_{-1}^* \hat{\mathbf{Z}}_{-1}^{*'} \beta_*)^{-1} \beta_*' \hat{\mathbf{Z}}_{-1}^* \Delta\hat{\mathbf{x}}' d\Delta\mathbf{x} \right|, \end{aligned}$$

where $\Delta\hat{\mathbf{x}}$ and $\hat{\mathbf{Z}}_{-1}^{*'}$ are residuals of ordinary least square (OLS) estimation in regression of $\Delta\mathbf{x}$ and \mathbf{Z}_{-1}^* on $(\mathbf{1}_T, \Delta\mathbf{Z}'_*)$, respectively.

We have that

²⁹ See Hogg et al. (2005) for further details.

$$\begin{aligned}
 \int f_0(\Delta \mathbf{x}) \log f(\Delta \mathbf{x}, \hat{\theta}) d\Delta \mathbf{x} &= -\frac{kT}{2} \log 2\pi - \frac{k}{2} \\
 &\quad - \frac{T}{2} \log |\Sigma_{v,v,0}| \\
 &\quad - \text{tr}[I(\theta_0) [[\hat{\theta} - \theta_0 [\hat{\theta} - \theta'_0]]]]
 \end{aligned}
 \tag{A.13}$$

where $I(\theta_0)$ is Fisher information matrix given in RC3 assumption, and also

$$\begin{aligned}
 A &= -2E \left[\int f_0(\Delta \mathbf{x}) \log f(\Delta \mathbf{x}, \hat{\theta}) d\Delta \mathbf{x} \right] \\
 &= kT \log 2\pi + k + T \log |\Sigma_{v,v,0}| \\
 &\quad + \text{tr}E [I(\theta_0) [[\hat{\theta} - \theta_0 [\hat{\theta} - \theta'_0]]]], \\
 &= kT(\log 2\pi + 1) + T \log |\Sigma_{v,v,0}| + K
 \end{aligned}$$

This expression is Eq. (3.21).

In a similar way, substituting $\hat{\beta}$ in Eq. (3.19)

$$-2\mathcal{L}_{\hat{\beta}} = k(\log 2\pi) + T^{-1} \hat{\mathbf{V}} \hat{\mathbf{V}}'$$

and taking expectation we finally have

$$-E[2\mathcal{L}_{\hat{\beta}}] = kT(\log 2\pi + 1) + T \log |\Sigma_{v,v,0}| - K$$

Proof of Theorem 1: Using the Taylor expansion, similar to the asymptotic distribution theory of MLE estimation, we first note that the model $f(\Delta \mathbf{x}, \theta_0)$ is estimated with $\hat{f}(\Delta \mathbf{x}) = f(\Delta \mathbf{x}, \hat{\theta})$.

The target function given in Eq. (3.17) can be rewritten as

$$A = -2E[\log f(\Delta \tilde{\mathbf{x}}, \hat{\theta})]$$

where $\Delta \tilde{\mathbf{x}}$ is an independent sample.

Let $\tilde{\theta}$ be the MLE calculated on this independent sample. This is an independent copy of $\hat{\theta}$.

We can rewrite A as

$$A = -2E[\log f(\Delta \mathbf{x}, \tilde{\theta})] \tag{A.14}$$

From RC2 assumption, consider Hessian

$$\mathbb{H} = -\frac{\partial}{\partial \theta \partial \theta'} E[\log f(\Delta \mathbf{x}, \theta)] > 0$$

and using a second-order Taylor series expansion of the log-likelihood $f(\Delta \mathbf{x}, \tilde{\theta})$ about $\hat{\theta}$, this is³⁰

$$\begin{aligned} \log f(\Delta \mathbf{x}, \tilde{\theta}) &= \log f(\Delta \mathbf{x}, \hat{\theta}) + \frac{\partial}{\partial \theta'} \log f(\Delta \mathbf{x}, \hat{\theta}) (\tilde{\theta} - \hat{\theta}) \\ &\quad - \frac{1}{2} (\tilde{\theta} - \hat{\theta})' \mathbb{H}(\tilde{\theta} - \hat{\theta}) + O_p(T^{-1/2}) \end{aligned} \quad (\text{A.15})$$

because of the first-order condition for the MLE, the second equality holds and we have

$$\log f(\Delta \mathbf{x}, \tilde{\theta}) = \log f(\Delta \mathbf{x}, \hat{\theta}) - \frac{T}{2} (\tilde{\theta} - \hat{\theta})' \mathbb{H}(\tilde{\theta} - \hat{\theta}) + O_p(T^{-1/2}) \quad (\text{A.16})$$

Using *RCI-3*, we get

$$A = E[2 \log f(\Delta \mathbf{x}, \hat{\theta})] + E[T(\tilde{\theta} - \hat{\theta})' \mathbb{H}(\tilde{\theta} - \hat{\theta})] + O(T^{-1/2}),$$

and

$$\begin{aligned} &= -E[2 \log L(\hat{\theta})] + E[T(\tilde{\theta} - \theta)' \mathbb{H}(\tilde{\theta} - \theta)] \\ &+ E[T(\hat{\theta} - \theta)' \mathbb{H}(\hat{\theta} - \theta)] + 2E[T(\tilde{\theta} - \theta)' \mathbb{H}(\hat{\theta} - \theta)] + O(T^{-1/2}) \end{aligned}$$

Under the *RCI-3*, MLE has convergence properties which is proved by Pesaran et al. (2000).

Therefore, we finally have

$$\begin{aligned} A &= -E[2\ell_T(\hat{\theta})] + E[\chi_K^2] + E[\tilde{\chi}_K^2] + O(T^{-1/2}), \\ &= -E[2\ell_T(\hat{\theta})] + 2K + O(T^{-1/2}). \end{aligned}$$

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10614-024-10569-6>.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). We declare that we did not receive any financial support and/or special funding related to the research, writing or publication of this paper.

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

³⁰ O_p denotes boundedness in probability.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abrigo, M. R., & Love, I. (2016). Estimation of panel vector autoregression in Stata. *The Stata Journal*, 16(3), 778–804.
- Assenmacher, K. (2013). Forecasting the Swiss economy with a small GVAR model. In F. Mauro & M. H. Pesaran (Eds.), *The GVAR handbook: Structure and applications of a macro model of the global economy for policy analysis* (pp. 244–254). Oxford University Press.
- Bai, J., & Ng, S. (2007). Determining the number of primitive shocks in factor models. *Journal of Business & Economic Statistics*, 25(1), 52–60.
- Bailey, N., Pesaran, M. H., & Smith, L. V. (2019). A multiple testing approach to the regularisation of large sample correlation matrices. *Journal of Econometrics*, 208(2), 507–534.
- Banbura, M., Giannone, D., & Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1), 71–92.
- Bernanke, B. S., Boivin, J., & Eliasziw, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly Journal of Economics*, 120(1), 387–422.
- Bierens, H. J. (2004). *Introduction to the mathematical and statistical foundations of econometrics*. Cambridge University Press.
- Burnham, K., & Anderson, D. (2002). *Model selection and multi-model inference: A practical information-theoretic approach*. Springer.
- Canova, F., & Ciccarelli, M. (2013). Panel vector autoregressive Models: A survey☆ The views expressed in this article are those of the authors and do not necessarily reflect those of the ECB or the euro-system. *VAR models in macroeconomics—new developments and applications: Essays in honor of Christopher A. Sims* (Vol. 32, pp. 205–246). Emerald Group Publishing Limited.
- Canova, F., & Marrinan, J. (1998). Sources and propagation of international output cycles: Common shocks or transmission? *Journal of International Economics*, 46(1), 133–166.
- Carriero, A., Kapetanios, G., & Marcellino, M. (2009). Forecasting exchange rates with a large Bayesian VAR. *International Journal of Forecasting*, 25(2), 400–417.
- Cashin, P., Mohaddes, K., Raissi, M., & Raissi, M. (2014). The differential effects of oil demand and supply shocks on the global economy. *Energy Economics*, 44, 113–134.
- Cavanaugh, J. E., & Neath, A. A. (2019). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdisciplinary Reviews: Computational Statistics*, 11(3), e1460.
- Cesa-Bianchi, A., Pesaran, M. H., Rebucci, A., Xu, T., & Chang, R. (2012). China's emergence in the world economy and business cycles in latin America. *Economia*, 12(2), 1–75.
- Chudik, A., & Pesaran, M. H. (2016). Theory and practice of gvar modelling. *Journal of Economic Surveys*, 30(1), 165–197.
- De Mol, C., Giannone, D., & Reichlin, L. (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, 146(2), 318–328.
- De Waal, A., Van Eyden, R., & Gupta, R. (2015). Do we need a global VAR model to forecast inflation and output in south africa? *Applied Economics*, 47(25), 2649–2670.
- Dees, S., Mauro, F. D., Pesaran, M. H., & Smith, L. V. (2007). Exploring the international linkages of the euro area: A global var analysis. *Journal of Applied Econometrics*, 22, 1–38.
- Fujikoshi, Y., & Satoh, K. (1997). Modified aic and cp in multivariate linear regression. *Biometrika*, 84(3), 707–716.

- Garratt, A., Lee, K., Pesaran, M. H., & Shin, Y. (2006). *Global and national macroeconomic modeling: A long-run structural approach*. Oxford University Press.
- Georgiadis, G. (2016). Determinants of global spillovers from US monetary policy. *Journal of International Money and Finance*, 67, 41–61.
- Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6), 1545–1578.
- Giannone, D., Reichlin, L., & Sala, L. (2005). Monetary policy in real time. In M. Gertler & K. Rogoff (Eds.), *NBER macroeconomics annual 2004* (Vol. 19, pp. 161–200). MIT Press.
- Gourieroux, C., & Monfort, A. (1995). *Statistics and econometric models*, (Vol. 1). Cambridge University Press.
- Granger, C. W., & Lin, J. L. (1995). Causality in the long run. *Econometric Theory*, 11(3), 530–536.
- Gregory, A. W., Head, A. C., & Raynauld, J. (1997). Measuring world business cycles. *International Economic Review*, 38, 677–701.
- Harbo, I., Johansen, S., Nielsen, B., & Rahbek, A. (1998). Asymptotic inference on cointegrating rank in partial systems. *Journal of Business & Economic Statistics*, 16(4), 388–399.
- Hogg, R. V., McKean, J., & Craig, A. T. (2005). *Introduction to mathematical statistics*. Pearson Education.
- Kapetanios, G., & Pesaran, M. H. (2007). Alternative approaches to estimation and inference in large multifactor panels: small sample results with an application to modelling of asset returns. In G. Phillips & E. Tzavalis (Eds.), *The refinement of econometric estimation and test procedures: Finite sample and asymptotic analysis*. Cambridge, UK: Cambridge University Press.
- Kose, M. A., Otrok, C., & Whiteman, C. H. (2003). International business cycles: World, region, and country-specific factors. *American Economic Review*, 93(4), 1216–1239.
- Lagana, G., & Mountford, A. (2005). Measuring monetary policy in the UK: A factor-augmented vector autoregression model approach. *The Manchester School*, 73, 77–98.
- Linhart, H., & Zucchini, W. (1986). *Model selection*. Wiley.
- Love, I., & Zicchino, L. (2006). Financial development and dynamic investment behavior: Evidence from panel VAR. *The Quarterly Review of Economics and Finance*, 46(2), 190–210.
- Lumsdaine, R. L., & Prasad, E. S. (2003). Identifying the common component of international economic fluctuations: A new approach. *The Economic Journal*, 113(484), 101–127.
- Mohaddes, K., & Raissi, M. (2020). Compilation, revision and updating of the Global VAR (GVAR) database, 1979Q2–2019Q4.
- Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), 967–1012.
- Pesaran, M. H. (2015). *Time series and panel data econometrics*. Oxford University Press.
- Pesaran, M. H., Schuermann, T., & Weiner, S. M. (2004). Modeling regional interdependencies using a global error-correcting macroeconomic model. *Journal of Business & Economic Statistics*, 22(2), 129–162.
- Pesaran, M. H., Shin, Y., & Smith, R. J. (2000). Structural analysis of vector error correction models with exogenous I(1) variables. *Journal of Econometrics*, 97(2), 293–343.
- Shibata, R. (1989). Statistical aspects of model selection. *From data to model* (pp. 215–240). Springer.
- Smith, L., & Galesi, A. (2014). *GVAR Toolbox 2.0*. University of Cambridge, Judge Business School.
- Stock, J. H., & Watson, M. W. (2005). Implications of dynamic factor models for VAR analysis. NBER Working Paper No. 11467.
- Stock, J. H., & Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44(2), 293–335.
- Stock, J. H., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2), 147–162.
- Zahedi, R., Shahmoradi, A., & Taiebnia, A. (2022). The ever-evolving trade pattern: a global VAR approach. *Empirical Economics*, 63, 1–26.

Authors and Affiliations

Halil Ibrahim Gunduz^{1,2}  · Furkan Emirmahmutoglu³  · M. Eray Yucel⁴ 

✉ Halil Ibrahim Gunduz
halil.gunduz@istanbul.edu.tr; halil.gunduz@maastrichtuniversity.nl
Furkan Emirmahmutoglu
f.emirmahmutoglu@hbv.edu.tr
M. Eray Yucel
eray@bilkent.edu.tr

- ¹ Department of Econometrics, Faculty of Economics, Istanbul University, Beyazit, Istanbul, Turkey
- ² School of Business and Economics, Department of Quantitative Economics, Maastricht University, Maastricht, The Netherlands
- ³ Department of Econometrics, Ankara Hacı Bayram Veli University, Besevler, 06500 Ankara, Turkey
- ⁴ Department of Economics, Faculty of Economics, Administrative and Social Sciences, Ihsan Dogramaci Bilkent University, Ankara, Turkey