BBVA Research

Creating Opportunities

N.º 21/08

Working Paper

A Global Vector Autoregressive Model for Banking Stress Testing

Rodolfo Méndez-Marcano

July 2021

A Global Vector Autoregressive Model for Banking Stress Testing

Rodolfo Méndez-Marcano*†

Abstract

This paper describes the Risk-GVAR 1.0, a Global Vector Autoregressive (GVAR) macroeconometric model designed, and successfully employed, to lend support to the internal stress testing exercises that banks, complying with prudential regulations, perform periodically to assess the adequacy of their current levels of capital. Additionally, it provides arguments justifying both the convenience of relying on macroeconometric models in this context and the specific choice of a GVAR.

Key words: macroeconometrics, prudential regulation, risk management, scenario analysis, stress test.
JEL classification: C32, E37, G17, G32

1 Introduction

Many of the biggest financial institutions of the global financial system in 2007 proved massively undercapitalized to cope with the adverse shocks that triggered the global financial crisis of 2007-2008, increasing its costs to society (tax payers bailout, credit crunch, etc). Among the instruments introduced and employed by the banking authorities for preventing something similar from happening again the **Internal Capital Adequacy Assessment Process** (ICAAP) stands out.

The general concept of the ICAAP (see EBA, 2016) was introduced in 2004 as part of the second pillar of the Basel II Accord[1](see BIS, 2005). However, it had been dead letter until the global financial crisis, since an increasing number of national supervisory authorities (starting by the European Union's) have developed detailed practical guidelines and started to compel **systemic banks**[2] to put them into practice.

The objective of the ICAAP is to strengthen the internal risk management capabilities of systemic banks by obliging its top management to, firstly, undertake and get deeply involved in regular exercises of stress testing based on macroeconomic **scenario analysis**, that is, on the evaluation of the effects on the level of capital of the most adverse future macroeconomic scenarios discernible from the relevant information currently

*BBVA Research, rodolfo.mendez@bbva.com. All the data and documentation required for estimating and employing the Risk-GVAR 1.0 (using the GVAR Toolbox 2.0 for MATLAB) are available on `https://doi.org/10.13140/RG.2.2.26252.74882`

†The author would like to thank Jorge Redondo and Nicolas Vazquez for their excellent research assistance, and Sara Baliña, Joxe Mari Barrutiabengoa, Julian Cubero and Rafael Domenech for their revision, suggestions and encouragement.

[1]The Basel Accords are the successive sets of banking regulatory recommendations issued by the Basel Committee on Banking Supervision, formed by the central banks and other banking regulatory authorities of more than 40 countries, including the biggest world economies.

[2]Roughly speaking, a systemic or "too big to fail" bank is one that is both big and interconnected enough to be a potential source of contagion (externalities) of liquidity and solvency problems to the rest of the national or even global financial system.

available (see BIS, 2000); and, secondly, given the results of these exercises to undertake the management actions required to guarantee the bank's solvency in case any of these scenarios (or similar ones) materialize in the future.

The central role of scenario analysis in the ICAAP is what brings **quantitative macroeconomic models** (hereinafter, **macroeconometric models**) to the stage given the computational challenge posed by the need to periodically producing and updating the required menu of adverse scenarios, each one comprehending projections for a large set of macroeconomic variables, in an efficient and transparent way. A macroeconometric model provides an explicit mathematical representation (in the form of a system of equations) of the potential pattern of interrelationships between the set of relevant macroeconomic variables based on equally explicit assumptions and/or empirical evidence. As such, it produces the required macroeconomic projections just by extrapolation in a completely automatic and transparent way through low-cost software and hardware.

This is exemplified by the macroeconometric model presented in this paper, the Risk-GVAR-1.0, a Global Vector Autoregressive Model (GVAR) that has successfully played this role in the context of **one of the largest European commercial banks**[3] since 2016 (when the ICAAP became mandatory for European Union's systemic banks). A bank whose **global scope** (business presence in several countries and continents) makes it necessary a **global or multi-country** macroeconometric model able to produce projections for the wide range of international macroeconomic variables that can potentially affect the value of its well-geographically-diversified global portfolio of assets and liabilities.

The next section presents a schematic description of the typical design and generation process of an adverse macroeconomic scenario of global scope for stress testing purposes, highlighting the stage where the support of a global macroeconometric model is most needed. The third section justifies the choice of a GVAR as the class of macroeconometric model to be used in this context. The fourth section explains in detail the nature of a GVAR model by clarifying its relationship with the highly standard and popular Vector Autoregressive Model (VAR). The fifth section describes and justifies the particular characteristics of the Risk-GVAR-1.0 by comparing it with the seminal and most standard GVAR introduced by Dees, Di Mauro, Pesaran and Smith, 2007 (the DDPS-GVAR). The sixth section illustrates the forecasting performance of the model presented against that of the DDPS-GVAR and slight variations of both. Finally, the seventh section provides a summary and the conclusions.

[3] BBVA

2 Scenario analysis and macroeconometric models

The elaboration of an adverse macroeconomic scenario of global scope for stress-testing the capital position of a given bank follows these steps:

1. Identify a specific shock (risk event) that could potentially cause sizable **adverse macroeconomic effects** (GDP growth reduction, inflation increase, domestic currency depreciation, sovereign risk increase, etc) in some of the countries and/or global markets most relevant for the bank.

   The most clear-cut case would be a **global shock** that could potentially affect (in a direct or indirect way) most countries and markets simultaneously. This is the case with a domestic shocks at some of the largest world economies (the USA, China and/or the European Union) or a shock on some global commodity market, like one that significantly affects the supply of oil (mainly geopolitical conflicts at the Middle East).

2. Under the assumption that the identified risk event will occur in the near future (a US recession in our example), project (based on the historical evidence on such event) the future path that would be most likely followed by the **first layer of relevant macroeconomic variables**, namely, the small group of macroeconomic variables both most directly affected by the shock and that act as the main transmission channel of its effects to the rest of domestic and international macroeconomic variables. Using the example of a US recession, this first layer could be reduced just to the US GDP.

3. Third, given the latter path or projection for the first layer of relevant macroeconomic variables (the US GDP in the selected example), project the future path that will most likely be followed by the **second layer of macroeconomic variables**, namely, the much wider group of macroeconomic variables required for computing the path of the value of the bank's assets, liabilities, and/or related risk indicators.

   This second layer of macroeconomic variables can in turn be divided, in function of standard macroeconomic concepts, into two subgroups: a **core subgroup**, whose evolution is directly affected by the evolution of the first-layer of variables, and a **non-core subgroup** whose evolution is essentially determined by the evolution of the core subgroup, and is only affected by the first layer variables through the core subgroup.

4. Fourth. given the projections for the first- and second-layers of macroeconomic variables, project the future path that will most likely be followed by the value the bank's asset, liabilities and/or related risk indicators.

For the outputs of the first and second steps, macroeconometric models would be of very limited, if any, use. On the one hand, they can usually be obtained from reliable external sources (world or country risk reports by private or multilateral financial institutions like IMF, BIS and credit rating agencies) because their highly generic nature (simultaneously and equally useful for the stress testing exercises of any bank or many of them) and the small set of macroeconomic variables projected. On the other hand, their in-house elaboration involve very different and much more event-specific and costly tools, including the comparative analysis of the historical evidence available for similar international events using, in the simplest case, a descriptive statistical approach, or, alternatively, panel-data logic or probit regressions or similar techniques.

In the third step, the projection of the second layer of macroeconomic variables conditional on the first, macroeconometric models play its most essential role given the large set of macroeconomic variables that need to be frequently projected and the high specificity of the variables that make up this set, which can vary substantially depending on the characteristics of the portfolios composition of the bank being stress-tested. These characteristics make unlikely to find a free external source for the scenarios required, and makes the automation of the process (using standard hardware and software) convenient, if not indispensable, something only possible on the basis of the mathematical structure of a macroeconometric model.

Given the output from the previous steps, the fourth step only requires the application of well established financial engineering concepts, assumptions and tools.

3 Why to choose a VAR model

As will be shown in the next section, a GVAR is no more than a restricted type of VAR, specifically, a feasible multi-country (and potentially global[4]) generalization of the standard single-country VAR. Thus, the reasons for choosing a GVAR against other multi-country macroeconometric models (in order to satisfy the stress testing needs of a global systemic bank) arise from the distinctive characteristics that it shares with a standard VAR model.

Accordingly, this section[5] describes the place of the standard VAR (and GVAR) in the general taxonomy of macroeconometric models as the most straightforward way of describing their distinctive characteristics and highlighting the advantages they offer when used to design the risk macrofinancial scenarios that feed the stress testing exercises of a global systemic bank.

[4]In this context, the adjective "global" refers to a set of countries that represents most of the World GDP.

[5]This section relies heavily on the methodological view of Nobel prize-winning economist, Christopher Sims (see Sims, 1980 and 2002).

Table 1 shows a classification of macroeconometric models along the most relevant dimensions to the present discussion. On the one hand, according to what is represented by their parameters, models can be divided into **reduced form** or **structural form** models. On the other hand, according to the information and method used to assign values to these parameters, models can be divided into **empirical**, **theoretical** or **hybrid**.

TABLE 1: Macroeconometric models
(mi subtitulo)

			Empirical	Theoretical	Hybrid
Reduced form (correlation pattern)			X[1]	X[2] (Derived from structural form)	
Structural form (causality pattern)	Overidentified	Microfounded (Derived from a general equilibrium setup)		X[3]	
		Ad-Hoc/Inconsistent		X[4]	
	Identified (Point/Set)				X[5] (Derived from empirical reduced Form)

[1] Standard VAR & GVAR ; [2] DSGEVAR(∞) ; [3] DSGE, CGE ; [4] SEM ; [5] Standard SVAR

. The table clearly shows that the standard VAR model, and its GVAR extension, are classified as empirical reduced form models, whereas the rest of models fall in the category of either theoretical or hybrid structural form models. In the following paragraphs, both of these categories are dissected along with their sub-categories.

As explained in the introduction, a macroeconometric model is essentially a mathematical representation of the average pattern of interrelationships between a set of macroeconomic variables along its join time evolution. These interrelationships can refer to two different but related elements:

- The set of potentially observable average pairwise correlations[6] between the values of the variables involved at different dates or, equivalently, the average association or co-movement between each pair of values for different dates of the same or different variables that can be inferred just by inspecting the historical join evolution of the variables.

- The set of unobservable average pairwise causal effects between variables that in theory should underly or explain the set of observable correlations referred above.

[6] More precisely: partial correlations, that is, the correlation between the fluctuations of two variables holding fix the rest of relevant variables.

An example would clarify the previous definitions. From the historical quarterly time series of observations of the variation rates (q-o-q) of the oil price (Brent) and the Euro Area's GDP, it is easy to compute the average correlation between the contemporary values of the two variables, which is 28% when data for the period 1994Q1-2019Q2 is used. But this correlation can in theory be broken down into two elements: the average response of the GDP to the causal effect of autonomous fluctuations of the oil price as those originated in oil supply shocks (such as when a military conflict in the Middle East temporarily halts the oil production of a big oil exporter) and the average response of the oil price to the causal effect of autonomous fluctuations of the GDP via its direct effect on the global demand for oil (as when the global financial crisis caused the contraction of the GDP of the Euro Area and other advanced economies and a consequent reduction of global oil demand).

As shown in **Table 1**, the latter distinction is the basis of the differentiation between **reduced form** models versus **structural form** models: A reduced form model represents the pattern of correlations between a set of macroeconomic variables while a structural form model represents the hypothetical underlying pattern of causality between the same set of variables (from which the pattern of correlations is derived).

The following relevant dimension of differentiation between classes of models related to the information and method used for assigning values to the *parameters* of the models required for making them **quantitative** (and only then useful for generating projections). In this dimension, models can be **empirical**, **theoretical** or **hybrid**, whether the information employed corresponds, respectively, to empirical information (the historical time series of observations for the variables of the model), theoretical assumptions (including purely ad-hoc ones), or a combination of both.

If sufficient and relevant empirical information is available, an empirical model is always the most transparent and reliable (in regard to forecasting accuracy) type of model. However, empirical models can only be reduced form models, i.e, models that provide an approximation to the average pattern of correlation between the variables, but not to the underlying pattern of causal interactions[7]. This fact limits the type of projections and scenarios that can be directly generated with a reduced form model to factual or most likely conditional and unconditional projections given the observed historical past.

In turn, counterfactual projections, which assume that some shocks and causal effects present in the past will not be present in the future, require a structural form model that incorporate the required theoretical and/or ad-hoc assumptions for disentangling shocks and causality from correlation and forecasting errors

[7] Going back to the example of the oil price and the EA's GDP: based only on the historical time series observations for both variables only the correlation between the value of one variable and the (contemporary, lagged or leaded) value of the other can be estimated, whereas its break down into the, always hypothetical, causal effects from one variable to the other require making theoretical assumptions as structural models do.

respectively. Now, structural models can be of two classes:

- **Identified structural form**: refers just to complementing an empirical reduced form (i.e, an standard VAR or GVAR model) with the minimum set of consistent theoretical assumptions required to disentangle the specific hypothetical causal interactions of interest from the correlations directly provided by the reduced form (and the associated structural shocks from the corresponding forecasting errors).

- **Overidentified structural form**: ignores the empirical reduced form (i.e, the pattern of correlations discernible from the historical data), and derive a whole hypothetical pattern of causal interactions between the variables from the specific theory (plus ad-hoc assumptions) of preference, finally using empirical information only for filling the *holes* (i.e, the quantitative aspects of the set of parameters that your preferred theory plus ad-hoc assumptions left undetermined).

Given the previous definitions and concepts, we are now prepared to answer the central question of this section: why is an empirical reduced form model (whose par excellence representative is a standard VAR or GVAR) a good, or even the best, class of macroeconometric model for supporting the scenarios-analysis-based stress testing exercises of a systemic bank of global scope, like the one considered here?.

The main argument is that scenario analysis requires credible scenarios, in the sense of being scenarios whose probability of occurrence can be trusted to be significantly higher than zero, and the most solid and objective support for this kind of confidence is the historical evidence, as summarized by an empirical reduced form model[8].

Notwithstanding, there is an immediate counterargument: the previous assertion only makes sense if there is enough historical macroeconomic time series data available to estimate a reduced form model to be feasible and its result reliable.

But the reply is equally straightforward: in the last few decades, hand in hand with the information and digital revolution, enough international macroeconomic data have accumulated and become readily available, making empirical reduced form macroeconometric models feasible and reliable forecasting tools.

[8] A metaphor can help clarify the following arguments: empirical reduced form and theoretical structural models play, respectively, a similar role to observational studies on human populations and experiments with (non-human) animals in the medical research of the effects on human health from the intake of specific substances (like a vaccine). In medical research these tools are seen as complementary more than substitutes or competitors: the results from observational studies on human populations are, in general, more reliable than results from experiments with animals for predicting the effect of a substance on human health on the condition that it is based on enough data (the availability of which requires that the substance be consumed by a large number of people for a long period of time and that enough related data is recorded). However, experiments with animals are an obligatory preliminary approach when the substance has not been consumed by (enough) people in the past and it is potentially dangerous and, the only possible approach, when the researcher need to explore the living body of the experimental subject in dangerous ways to understand the exact channels and mechanisms through which the substance produces its effects.

A second argument and more practical argument is the much higher cost and time involved in the development of theoretical structural models, as well as the higher qualitative and quantitative human capital requirements for their operation.

But there is a second and more complex counterargument: although empirical reduced form models are by construction the best tool for forecasting the most likely scenario and extrapolating less likely events from the past; in contrast with theoretical structural form scenarios, reduced form models cannot simulate many highly specific counterfactual scenarios (i.e, scenarios caused by very specific shocks), less so if they are completely unprecedented, nor they help to select a unique narrative of the underlying causal interactions.

The reply in this case is as follows:

a. Firstly, the statement is not strictly true. It neglects to mention that by constraining the estimation sample, an empirical reduced form model can roughly isolate the causal effects of many specific shocks and then simulate their future effects by extrapolation. The only conditions for this is the existence of historical episodes when these were the only shocks in action and/or the effects of the rest of shocks can be assumed negligible.

b. Secondly, as Antolín-Díaz, Petrella and Rubio-Ramírez (2021) masterfully explain and illustrate, for more specific counterfactual scenarios, you just need to complement your VAR (or GVAR) model with the required identification restrictions, and form the appropriate identified hybrid structural model (i.e, the appropriate Structural Vector Autoregressive model or SVAR).

   Developing a SVAR once you count with a VAR model, takes just a tiny fraction of the cost and time it takes to develop a theoretical structural model, and its only disadvantages are irrelevant for stress testing: A SVAR can only simulate counterfactual scenarios related to the causal effects of shocks experienced in the past or, highly similar to some of them, but these are precisely the only credible counterfactual scenarios given the historical evidence.

c. As for the detailed unique narrative provided by (and inscribed in) each specific theoretical structural model, though valuable, it can be replaced by the more flexible and varied narratives that can be provided based on expert criterion and analysis to justify the conditional projections from a VAR or SVAR model.

4 What is a GVAR?

VAR models (along with the Structural VARs derived from them) has been for decades the *par excellence* empirical approach to macroeconomic analysis and forecasting at the national level, and the GVAR (in itself no more than a restricted VAR) is the result of the effort to extend the use of this approach from a single-country setting to a multi-country, or even global one.

As explained below, it is not feasible to extend a VAR model to a large multi-country setting, while retaining some forecasting reliability, unless enough a priori restrictions are imposed on its parameters. But at the same time, introducing these restrictions can severely compromise the highly valuable empiricalness of the model (and with it, its forecasting reliability, transparency and objectivity), unless they are very few and generally admissible. As will be explained next, the standard GVAR claims to be the optimal solution to this dilemma or trade-off.

4.1 Unrestricted Single-country VAR

For simplicity, let's start with a simple example of a standard VAR, namely a bivariate VAR for the logarithm of the real GDP (y^{arg}) and the logarithm of the CPI (p^{arg}) for Argentina:

$$\underbrace{\begin{pmatrix} y_t^{arg} \\ p_t^{arg} \end{pmatrix}}_{y_t} = \underbrace{\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}}_{C} + \sum_{k=1}^{p} \underbrace{\begin{pmatrix} a_{11k} & a_{12k} \\ a_{21k} & a_{22k} \end{pmatrix}}_{A_k} \underbrace{\begin{pmatrix} y_{t-k}^{arg} \\ p_{t-k}^{arg} \end{pmatrix}}_{y_{t-k}} + \underbrace{\begin{pmatrix} \varepsilon_{y,t}^{arg} \\ \varepsilon_{p,t}^{arg} \end{pmatrix}}_{\varepsilon_t} \tag{1}$$

for $t = 1, 2, 3, ..., T$ (where t represents time, usually in quarters)

and where $\varepsilon_t \sim N[0, \Omega]$ with $\Omega = \begin{pmatrix} \omega_{11k} & . \\ \omega_{21k} & \omega_{22k} \end{pmatrix}$

The model is a system of two linear equations, each expressing the value at date t of one of the endogenous variables (only y^{arg} and p^{arg} in this example) as a linear function of the lagged values (until a maximum of p lags) of both of them plus an error term ($\varepsilon_{y,t}^{arg}$ and $\varepsilon_{p,t}^{arg}$), such that the vector of error terms at date t (ε_t) is assumed to be distributed according to a multivariate normal probability distribution with mean $\mathbf{0}$ (zero vector), variance-covariance matrix Ω and non-serial correlation, $Cov(\varepsilon_t, \varepsilon_{t-j}) = 0$ for any j in the integer line.

Therefore, the parameters to be estimated are given by the elements of the matrices C (dimension $\eta \times 1$), A_k (dimension $\eta \times \eta$) for $k = 1, 2, ...p$ and Ω (also dimension $\eta \times \eta$ but symmetric and positive semi-definite, as corresponding to a variance-covariance matrix) where $\eta = m \times ndvc$, $ndvc$ is the number of domestic variables per country, m the number of countries and p the maximum number of lags.

It should be noted that, by construction, once estimated, the set of parameters of the model (excluding the intercepts, the vector C) will contain the information of all the pairwise partial correlations between the values of the endogenous variables for the same or different dates. More specifically it will contain the partial correlation between the value of any of the endogenous variables at any given date and the contemporary and lagged value (for any lag between 1 and p) of itself or any other variable in the model given the value of the rest of variables[9].

Usually, the parameters, for each admissible value of p[10], are estimated equation by equation[11] using the **Ordinary Least Squares (OLS)** method (which provides consistent estimators in this context) using a sample of as long as possible (usually quarterly) time series of observations for the variables in the model. Finally, the estimated model selected is the one with the value of p that minimizes some likelihood-related information criterion (usually, the Akaike Information Criterion).

4.2 Unrestricted Multi-country VAR

Single-country VARs, like VAR (1), have long been considered a good enough approximation for the study of large and relatively closed economies like the US economy (focus of the seminal applications of the VAR approach in macroeconomics by Sims, 1980, and the bulk of the myriad of subsequent applications), but it is clearly inappropriate (as argued, for example, by Cushman and Zha, 1997) for studying small open economies, significantly dependent as they are on the rest of the world (for example, Argentina, Spain or Canada), or, obviously, when the main interest is analyzing the interaction between the macroeconomic variables of different countries.

However, as illustrated below, the multi-country expansion of a VAR poses a formidable challenge for a satisfactory estimation because with each additional country included the number of **regression parameters** (including intercept) per equation increases (by more than one) whereas the number of observations per variable remains fixed, limited by the, typically short, length of the historical macroeconomic time series available for

[9] The partial correlation between contemporary values of any two variables is not directly given, but are implicit and can be computed in a straightforward manner from the elements of Ω.

[10] A maximum number of lags is fixed a priori (usually 4 for quarterly data, to account for potential seasonality), which is the only a priori restriction on the parameters of a standard VAR, apart from the selection of the variables to include.

[11] Because every equation has the same regressors, there is no efficiency gain in making a system-wide estimation (using for example the SUR method).

estimation[12]. In this way, the addition of countries increasingly worsen the problem of overparameterization[13], already present in a single-country VAR (because of the high number of interrelated domestic macroeconomic variables of interest).

To illustrate this point, we reformulate the previous bivariate VAR for Argentina by adding to it the same two domestic macroeconomic variables (logarithm of GDP and CPI) for three of its MERCOSUR trade partners (Brazil, Uruguay and Paraguay), with the following result:

$$\underbrace{\begin{pmatrix} y_{t-k}^{arg} \\ p_{t-k}^{arg} \\ y_{t-k}^{bra} \\ p_{t-k}^{bra} \\ y_{t-k}^{uru} \\ p_{t-k}^{uru} \\ y_{t-k}^{par} \\ p_{t-k}^{par} \end{pmatrix}}_{y_t} = \underbrace{\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \\ c_8 \end{pmatrix}}_{C} + \sum_{k=1}^{p} \underbrace{\begin{pmatrix} a_{11k} & a_{12k} & a_{13k} & a_{14k} & a_{15k} & a_{16k} & a_{15k} & a_{16k} \\ a_{21k} & a_{22k} & a_{23k} & a_{24k} & a_{25k} & a_{26k} & a_{15k} & a_{16k} \\ a_{31k} & a_{32k} & a_{33k} & a_{34k} & a_{35k} & a_{36k} & a_{15k} & a_{16k} \\ a_{41k} & a_{42k} & a_{43k} & a_{44k} & a_{45k} & a_{46k} & a_{15k} & a_{16k} \\ a_{51k} & a_{52k} & a_{53k} & a_{54k} & a_{55k} & a_{56k} & a_{15k} & a_{16k} \\ a_{61k} & a_{62k} & a_{63k} & a_{64k} & a_{65k} & a_{66k} & a_{15k} & a_{16k} \\ a_{51k} & a_{52k} & a_{53k} & a_{54k} & a_{55k} & a_{56k} & a_{15k} & a_{16k} \\ a_{61k} & a_{62k} & a_{63k} & a_{64k} & a_{65k} & a_{66k} & a_{15k} & a_{16k} \end{pmatrix}}_{A_k} \underbrace{\begin{pmatrix} y_{t-k}^{arg} \\ p_{t-k}^{arg} \\ y_{t-k}^{bra} \\ p_{t-k}^{bra} \\ y_{t-k}^{uru} \\ p_{t-k}^{uru} \\ y_{t-k}^{par} \\ p_{t-k}^{par} \end{pmatrix}}_{y_{t-k}} + \underbrace{\begin{pmatrix} \varepsilon_{y,t}^{arg} \\ \varepsilon_{p,t}^{arg} \\ \varepsilon_{y,t}^{bra} \\ \varepsilon_{p,t}^{bra} \\ \varepsilon_{y,t}^{uru} \\ \varepsilon_{p,t}^{uru} \\ \varepsilon_{y,t}^{par} \\ \varepsilon_{p,t}^{par} \end{pmatrix}}_{\varepsilon_t} \tag{2}$$

$$\text{with } \varepsilon_t \sim N[0,\Omega] \text{ where } \Omega = \begin{pmatrix} \omega_{11} & . & . & . & . & . & . & . \\ \omega_{21} & \omega_{22} & . & . & . & . & . & . \\ \omega_{31} & \omega_{32} & \omega_{33} & . & . & . & . & . \\ \omega_{41} & \omega_{42} & \omega_{43} & \omega_{44} & . & . & . & . \\ \omega_{51} & \omega_{52} & \omega_{53} & \omega_{54} & \omega_{55} & . & . & . \\ \omega_{61} & \omega_{62} & \omega_{63} & \omega_{64} & \omega_{65} & \omega_{66} & . & . \\ \omega_{71} & \omega_{72} & \omega_{73} & \omega_{74} & \omega_{75} & \omega_{76} & \omega_{77} & . \\ \omega_{81} & \omega_{82} & \omega_{83} & \omega_{84} & \omega_{85} & \omega_{86} & \omega_{87} & \omega_{88} \end{pmatrix}$$

Let's assume, for illustrative purposes, a maximum of 4 lags (p=4), and a number of observations per variable of 80 (think of quarterly time series from 2000Q1 to 2019Q4), both common values in practice. Then the inclusion of the additional 3 countries, implies an increase in the number of regression parameters per equation[14] from 9 in VAR (1) to 33 in VAR (2) which, though still outnumbering the available observations per variable, substantially reduce the degrees of freedom and consequently the reliability and forecasting performance of the estimated model.

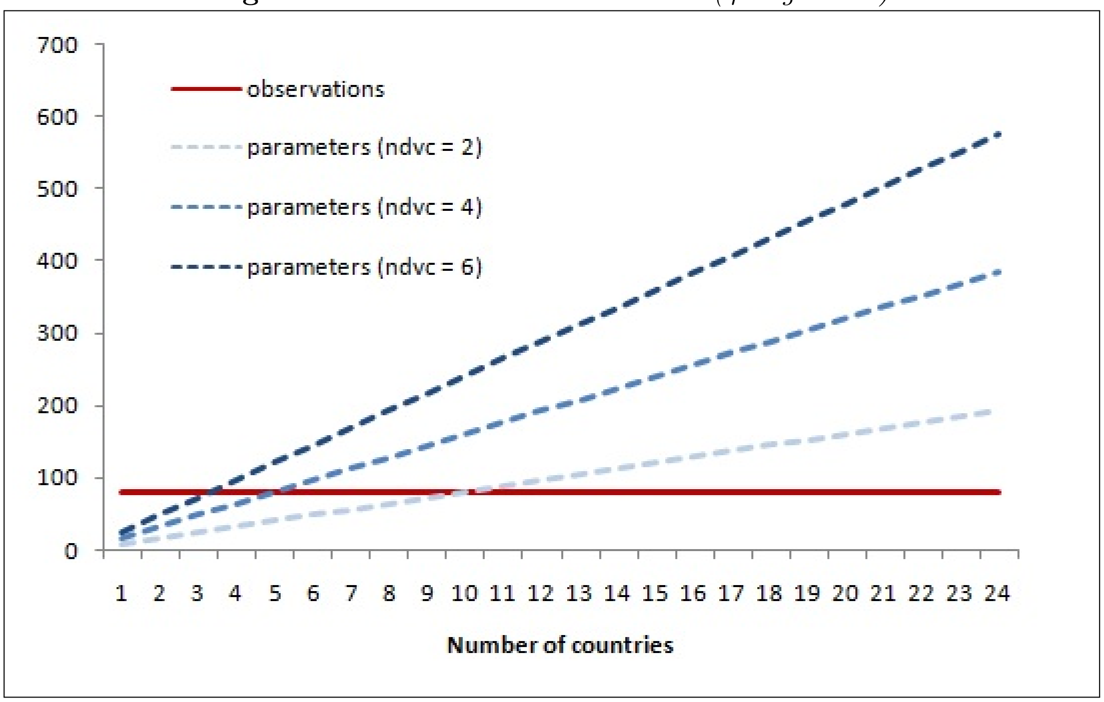
[12] At the level of the whole system of equations, this implies that with each additional country the total number of parameters increases exponentially, whereas the total number of observations increases linearly.

[13] Too many regression parameters per equation to be estimated given the number of observations available per variable as for the OLS estimators to be reliable or even feasible. For one thing, computing OLS estimators becomes mathematically impossible when the number of parameters per equation exceeds the number of observations per variable and, even if that is not the case, the computed estimates and derived forecasts will be unreliable if observations do not outnumber parameters by a large enough amount.

[14] Computed as $\eta \times p + 1$ where, as before, $\eta = m \times ndvc$.

But, as shown in Figure 1, things worsen quickly when additional countries are included to the point of making it unfeasible the estimate an unrestricted VAR with more than a very small number of countries. In the figure, the maximum number of countries for which a VAR can be estimated is equal to or less than the value at the intersection of the red line (fixed number of observations) and the blue broken line corresponding to the number of domestic variable per country (ndvc) considered: with just 2 domestic variables per country (ndvc=2), as in the previous examples, the maximum number of countries is 9, but with 4 domestic variables (ndvc=4) it decreases to 4 countries whereas with 6 (ndvc=6), it decreases to just 3 countries.

Figure 1:*Parameters vs. observations (4-Lags-VAR)*



ndvc = number of domestic variables per country

4.3 GVAR as a feasible Multi-country VAR

The challenge of making it possible to estimate a Global VAR (or an extensive multi-country VAR) requires introducing a set of generally acceptable a priori assumptions (and implied restrictions on parameters) that can produce a substantial reduction both in the number of parameters to be estimated and in its growth with each additional country.

Pesaran et.al. (2001) propose one such set, making it feasible for the first time to reliably estimate a Global (though restricted) VAR, and making it the standard basis of all posterior GVARs, including the seminal and most popular of them: the DDPS-GVAR of Di Mauro, Dees, Pesaran and Smith (2007). The set involves the following assumptions (and implied restrictions), all of which commonly used in international macroeconomic theory and modelling (hence, their relatively easy and general edibility):

- When a domestic macroeconomic variable is significantly and directly affected by a foreign macroeconomic variable, this effect is not directly caused by its value for some individual foreign country but instead by some weighted aggregate (sum or average) of its values for all (or the most relevant) foreign countries, where the appropriate weights will be given by some measure of bilateral macroeconomic interdependence.

  Let's use the Spanish GDP as an illustration: this assumption asserts that it does not depend as much on the GDP of a specific European country (Portugal, France, etc.) as it does on the aggregate of the GDPs of the rest of the countries or the relevant subset of them (mainly the European Union).

- For practical purposes, the appropriate weights can be considered known a priori (so they do not need to be estimated) because they can be properly approximated by a publicly available measure of average "bilateral trade weights", i.e, the average weight of the bilateral trade (imports plus exports of goods and services) between the domestic country and each foreign country divided by the total trade of the domestic country.

- With some exceptions (mainly, the largest world economies: USA, China and/or the EU), for most countries and for practical purposes the causal effects of any domestic macroeconomic variable on (individual or aggregate) foreign macroeconomic variables can be assumed negligible.

The powerful implications of these three assumptions can be illustrated by incorporating them in our previous effort for modelling the macroeconomic interrelationship between Argentina, Brazil, Uruguay and Paraguay. The result is the following interconnected suit of Small Open Economy VARXs[15], which together

[15] A VARX is a VAR that includes some exogenous variables, i.e, variables that appear on the right-hand-side of all the equations of the VAR but not on the-left-hand side of any of them.)

form the following 4-country GVAR:

$$
\begin{aligned}
\begin{pmatrix} y_t^{arg} \\ p_t^{arg} \end{pmatrix} &= \begin{pmatrix} c_1^1 \\ c_2^1 \end{pmatrix} + \sum_{k=1}^{p_1} \begin{pmatrix} a_{11k}^1 & a_{12k}^1 \\ a_{21k}^1 & a_{22k}^1 \end{pmatrix} \begin{pmatrix} y_{t-k}^{arg} \\ p_{t-k}^{arg} \end{pmatrix} + \sum_{k=\mathbf{0}}^{q_1} \begin{pmatrix} b_{11k}^1 & b_{12k}^1 \\ b_{21k}^1 & b_{22k}^1 \end{pmatrix} \begin{pmatrix} y_{t-k}^{arg^*} \\ p_{t-k}^{arg^*} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t}^1 \\ \varepsilon_{p,t}^1 \end{pmatrix} \\
\begin{pmatrix} y_t^{bra} \\ p_t^{bra} \end{pmatrix} &= \begin{pmatrix} c_1^2 \\ c_2^2 \end{pmatrix} + \sum_{k=1}^{p_2} \begin{pmatrix} a_{11k}^2 & a_{12k}^2 \\ a_{21k}^2 & a_{22k}^2 \end{pmatrix} \begin{pmatrix} y_{t-k}^{bra} \\ p_{t-k}^{bra} \end{pmatrix} + \sum_{k=0}^{q_2} \begin{pmatrix} b_{11k}^2 & b_{12k}^2 \\ b_{21k}^2 & b_{22k}^2 \end{pmatrix} \begin{pmatrix} y_{t-k}^{bra^*} \\ p_{t-k}^{bra^*} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t}^2 \\ \varepsilon_{p,t}^2 \end{pmatrix} \\
\begin{pmatrix} y_t^{uru} \\ p_t^{uru} \end{pmatrix} &= \begin{pmatrix} c_1^3 \\ c_2^3 \end{pmatrix} + \sum_{k=1}^{p_3} \begin{pmatrix} a_{11k}^3 & a_{12k}^3 \\ a_{21k}^3 & a_{22k}^3 \end{pmatrix} \begin{pmatrix} y_{t-k}^{uru} \\ p_{t-k}^{uru} \end{pmatrix} + \sum_{k=0}^{q_3} \begin{pmatrix} b_{11k}^3 & b_{12k}^3 \\ b_{21k}^3 & b_{22k}^3 \end{pmatrix} \begin{pmatrix} y_{t-k}^{uru^*} \\ p_{t-k}^{uru^*} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t}^3 \\ \varepsilon_{p,t}^3 \end{pmatrix} \\
\begin{pmatrix} y_t^{par} \\ p_t^{par} \end{pmatrix} &= \begin{pmatrix} c_1^4 \\ c_2^4 \end{pmatrix} + \sum_{k=1}^{p_n} \begin{pmatrix} a_{11k}^4 & a_{12k}^4 \\ a_{21k}^4 & a_{22k}^4 \end{pmatrix} \begin{pmatrix} y_{t-1}^{par} \\ p_{t-1}^{par} \end{pmatrix} + \sum_{k=0}^{q_n} \begin{pmatrix} b_{11k}^4 & b_{12k}^4 \\ b_{21k}^4 & b_{22k}^4 \end{pmatrix} \begin{pmatrix} y_{t-k}^{par^*} \\ p_{t-k}^{par^*} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t}^4 \\ \varepsilon_{p,t}^4 \end{pmatrix}
\end{aligned} \tag{3}
$$

where

$$
\begin{aligned}
y_t^{arg^*} &= \theta_{11} \times y_t^{arg} + \theta_{12} \times y_t^{bra} + \theta_{13} \times y_t^{uru} + \theta_{14} \times y_t^{par} \\
p_t^{arg^*} &= \theta_{11} \times p_t^{arg} + \theta_{12} \times p_t^{bra} + \theta_{13} \times p_t^{uru} + \theta_{14} \times p_t^{par} \\
y_t^{bra^*} &= \theta_{21} \times y_t^{arg} + \theta_{22} \times y_t^{bra} + \theta_{23} \times y_t^{uru} + \theta_{24} \times y_t^{par} \\
p_t^{bra^*} &= \theta_{21} \times p_t^{arg} + \theta_{22} \times p_t^{bra} + \theta_{23} \times p_t^{uru} + \theta_{24} \times p_t^{par} \\
y_t^{uru^*} &= \theta_{31} \times y_t^{arg} + \theta_{32} \times y_t^{bra} + \theta_{33} \times y_t^{uru} + \theta_{34} \times y_t^{par} \\
p_t^{uru^*} &= \theta_{31} \times p_t^{arg} + \theta_{32} \times p_t^{bra} + \theta_{33} \times p_t^{uru} + \theta_{34} \times p_t^{par} \\
y_t^{par^*} &= \theta_{41} \times y_t^{arg} + \theta_{42} \times y_t^{bra} + \theta_{43} \times y_t^{uru} + \theta_{44} \times y_t^{par} \\
p_t^{par^*} &= \theta_{41} \times p_t^{arg} + \theta_{42} \times p_t^{bra} + \theta_{43} \times p_t^{uru} + \theta_{44} \times p_t^{par}
\end{aligned}
$$

with $\theta_{ij} = 0$ for $i = j$ and $\sum_{j=1}^{4} \theta_{ij} = 1$ for $i = 1, 2, 3, 4$, and, as usual, $(\varepsilon_{y,t}^h, \varepsilon_{p,t}^h)' \sim N[0, \Omega_h]$ for $h = 1, 2, 3, 4$

What is achieved in this way?:

- Though both, the VAR (2) model and the GVAR (3) model, summarize the association or correlation between the selected domestic variables for Argentina, Brazil, Uruguay and Paraguay, the GVAR (3) model affords it with much fewer parameters per equation: for p=4, VAR (2) has, as observed before, 33 regression parameters per equation, whereas GVAR (3) has only 9 parameters per equation, the same that the single-country VAR (1) model.

- Much more importantly, **the introduction of additional countries in model (3) does not increase the number of parameters per equation** and, consequently, does not not consume degrees of freedom. This property alone, guarantee the feasibility of making of model (3) a true global model.

But can model (3) really be considered a VAR?: Pesaran et.al. (2001) showed that, in fact, model (3), as a whole, can be straightforwardly re-expressed as a single restricted VAR model. The details are clearly explained by Smith and Galesi (2014).

4.4 Variations

In practice, GVAR specifications (including the seminal DDPS-GVAR and the Risk-GVAR, introduced in this paper) are used to differ in several dimensions from the simple small symmetric GVAR of the previous section, and between each other. These are some of most common variants:

- Even when GVARs are commonly expressed for their use on the levels or log-levels of the macrofinancial variables, most of the time their estimation actually use a version formulated in error-correction-form or first-differences (though usually re-expressed in levels or log-levels form after estimation).

- Not all country satellites should have a block of foreign variables, the most typical example being treating the USA economy as a large closed economy not affected by foreign variables (even in aggregate way) or, as is the case with the DDPS-GVAR, including only one variable in its foreign block (the weighted average of the bilateral real exchange rates of its trade partners).

- Not all country satellites need to have the same variables in their domestic and foreign blocks, the most common example being when some domestic variable is only available for some countries as is the case with the EMBI spread, only available for emerging economies, or that of a representative long-run interest rate, which is not available for most emerging economies.

- A given national satellite does not need to have the same variables in its domestic (endogenous) and foreign (exogenous) blocks.
  - On the one hand, the foreign block frequently include, in addition to aggregated country-variables, also a set of "global variables", i.e, variables whose value is determined by a sum of international transactions and can be treated as exogenous for most countries, the main example being the price of oil and other commodities.
  - On the other, sometimes the foreign block only includes a subset of the domestic variables (always aggregated for the whole set of foreign countries in the model): those considered direct transmission channels from foreign countries. For example, it could be assumed that a country is **directly** affected by the aggregated GDP of its trade partners but not for their individual expenditure components (consumption, investment, etc),whose effect would only be indirect (just through GDP).

- In aggregating national variables for forming the foreign aggregates variables, trade weights (sum of bilateral exports plus imports between total exports and imports) are not the only or exclusive option but can be substituted or combined (using different type of weights for different variables) by export or import weights alone, distance weights or financial trade weights (representing the bilateral trade in financial assets between countries, as in Eickmeier and Ng, 2015). For a survey and evaluation of the main alternatives see Martin and Crespo (2016).

- In certain cases, some of the domestic variables from a particular country are allowed to enter directly (and not as part of a weighted aggregation or average) into the foreign block of the satellites of some or all the rest of countries. The typical example is that of the US domestic variables as is the case with Chudik and Smith, 2013.

5 Examples: Comparing DDPS-GVAR and Risk-GVAR-1.0

The already mentioned DDPS-GVAR of Dees, Di Mauro, Pesaran and Smith (2007) is the departure point for every new GVAR model, and the Risk-GVAR is no exception. This is a natural choice for several reasons: it is the standard and most popular reference, it is the most extensively documented, and it is also the example whose database and interface is included and documented with open source GVAR-Toolbox for MATLAB, the seminal and *par excellence* software for the estimation and use of GVAR models.

However, a careful evaluation reveals that the DDPS-GVAR specification is inadequate for the specific purpose (banking stress testing) and user (BBVA) that are the focus of this paper. For several reasons:

- It shows a very poor performance at forecasting inflation and nominal variables in general, as shown by the backtesting evaluation below.

- It is excessively complex and overparameterized (giving room to many unlikely direct interactions between domestic variables for different countries) as well as large (too many countries, including many of low importance for the world economy), which implies a heavy computational burden and high maintenance costs.

- It excludes some relevant countries (i.e, countries where the user has valuable business operations).

- It excludes banking variables, that are key for performing banking stress testing exercises.

The differences between Risk-GVAR and DDPS-GVAR arise from the efforts to correct this inadequacy, that is, to find a GVAR model better suited to our specific purposes and user. In the following paragraphs,based on the information provided in **Table 2**[16], we compare both models highlighting the aforementioned differences and its motivation[17].

Table 2: Variables notation and description

LABEL	DESCRIPTION	TRANSFORMATION
y	Real GDP	logarithm
p	CPI	logarithm
fx1	exchange rate (LCU/USD)	logarithm
fx2	exchange rate (LCU/EURO)	logarithm
ep1	fx1 - p	none
ep2	fx2 - p	none
epr1	fx1 - p + p_usa	none
epr2	fx2 - p + p_emu	none
eq	equity index	logarithm
R	short-term interest rate (annual,%)	logarithm
LR	long-term interest rate (annual,%)	logarithm
r	$0.25 \times log(1 + R/100)$	none
lr	$0.25 \times log(1 + LR/100)$	none
embi	EMBI spread	logarithm
Cren	Credit stock (LCU)	logarithm
Depon	Deposit stock (LCU)	logarithm
Cre	Cren - p	none
Depo	Depon - p	none
poil	oil price (WTI)	logarithm
pmetal	metals price index	logarithm
praw	raw materials price index	logarithm

The choice of countries and macrofinancial variables to be included in each GVAR model, summarized in the next two tables, was decided based on the purpose of each of them and cost-effectiveness considerations (like keeping management and maintenance costs). In the case of the DDSP-GVAR the original purpose was an agnostic exploration of the international macroeconomic links between the Euro Area and the rest of the world, while for the Risk-GVAR, as already explained, it was producing risk scenarios for the required domestic macrofinancial variables of those countries considered **directly business-relevant** for our end-user bank (namely, those where the bank has a significant part of its business) but also those considered **indirectly business-relevant** (namely those whose macrofinancial evolution substantially affects the evolution of the directly business-relevant countries, as is the case with the largest world economies: USA, EU and China).

Table 3 shows the **countries included in each model**, the following facts standing out:

[16] We follow as close as possible the notation and basic transformations of Dees, Di Mauro, Pesaran and Smith (2007), including their transformation of the interest rates into their instantaneous equivalent in a continuous compounding regime: $r = 0.25 \times log(1 + R)$ where R is the original annual nominal interest rate.

[17] To get a more complete description of the two GVARs refer to the excel interface and User's Guide for their estimation and use with the GVAR-Toolbox for MATLAB of Smith and Galesi. In the case of DDPS-GVAR they are included in the installation files of the toolbox (downloadable here), while in the case of the Risk-GVAR they are available here

- The number of individual countries included in the DDPS-GVAR (33 with 8 grouped as EA) is almost double those in the Risk-GVAR (18 with 7 grouped as EA-ex-Spain) because of the exclusion of the **non** business-relevant-countries from the latter and the inclusion of the business-relevant countries absent in the former, namely, Colombia and Venezuela.

- The bulk of countries in each GVAR has its own individual satellite in it, but there is an exception: the macrofinancial variables of the members of the Euro Area (with the exception of Spain in the case of the Risk-GVAR) are averaged (using PPP-GDP as weights) and treated as the domestic variables for a single "country" (EA and EA-ex-Spain in the cases of the DDPS-GVAR and Risk-GVAR respectively) and henceforth included in a single satellite. The rationale for this aggregation is:

  – In the case of the DDPS-GVAR, to accomplish its original purpose: studying the international linkages of the Euro Area as a whole.

  – In the case of the Risk-GVAR, the fact that Spain is the only country in the Euro Area that is directly-business-relevant. The rest of the Euro Area, seen as a single "country", is only indirectly-business-relevant, i.e, it is only relevant as the main trade partner of Spain and one of the main trade partners for tje rest of the business-relevant countries.

Table 4 shows the variables included in the domestic and foreign block of each national satellite for each GVAR. In the following paragraphs, We highlight the more remarkable contrasts between both GVARs:

- The first column shows all the categories of macroeconomic variables (14) included in at least one of the national satellites of one or both GVARs.

- Only five of these categories appear in both GVARs: y, p, r, lr and poil. Whereas only the DDPS-GVAR include ep, eq, $pmetal$ and $praw$, and only the Risk-GVAR include $epr1$, $epr2$, $embi$, Cre and $Depo$. These differences are as follows:

  – The Risk-GVAR opts for using a standard indicator of the relative value of national currencies, the bilateral real exchange rates (epr1, epr2), instead of the unusual indicator present in DDPS-GVAR (ep). This makes the direct outputs of the model more useful and easier to interpret.

Table 3: Countries and satellites

LABEL	DSDP-GVAR	RISK-GVAR
Argentina	Yes	Yes
Australia	Yes	No
Austria	EA	EA-ex-Spain
Belgium	EA	EA-ex-Spain
Brazil	Yes	Yes
Canada	Yes	No
China	Yes	Yes
Chile	Yes	Yes
Colombia	Not	Yes
Finland	EA	EA-ex-Spain
France	EA	EA-ex-Spain
Germany	EA	EA-ex-Spain
India	Yes	No
Indonesia	Yes	No
Italy	EA	EA-ex-Spain
Japan	Yes	No
Korea	Yes	No
Malaysia	Yes	No
Mexico	Yes	Yes
Netherlands	EA	EA-ex-Spain
Norway	Yes	No
New Zealand	Yes	No
Peru	Yes	Yes
Philippines	Yes	No
South Africa	Yes	No
Saudi Arabia	Yes	No
Singapore	Yes	No
Spain	EA	Yes
Sweden	Yes	No
Switzerland	Yes	No
Thailand	Yes	No
Turkey	Yes	Yes
UK	Yes	No
USA	Yes	Yes
Venezuela	Not	Yes
COUNTRIES	33	18
SATELLITES	25+EA	11+EA-ex-Spain

- The Risk-GVAR excludes the average international price of metals and raw material, because their highly correlation with the (included) oil price make them redundant (if projections for them are required, they can be easily produced by auxiliary regressions linking them to the oil price projection).
  - The Risk-GVAR includes banking variables, specifically the two most fundamental from a commercial banking perspective: total deposits and total loans.
  - The Risk-GVAR excludes the equity price index but includes the EMBI spread of the emerging countries. It is considered that for the purpose of designing risk macrofinancial scenarios the latter variable is more useful. At the macroeconomic level, the equity price index basically moves with the GDP, whereas the EMBI reflects many different risks so imposing restrictions on its future trajectory allows many types of risk scenarios.

- While the DDPS-GVAR allows every domestic variables for each country to have direct effects (proportional to the weight of their bilateral trade) on all the domestic variables of the rest of countries, in order to reduce complexity and increase economic interpretability the Risk-GVAR assumed that the macroeconomic interaction between countries occurs only through GDPs, with the only exception being the US' long-term interest rate and inflation which would directly affect all countries.

- The Risk-GVAR is block-recursive: it is assumed that the block of the USA, China and Europe is not affected by the domestic variables of the rest of countries, which is done by collapsing to zero the weight of the latter in the total trade of the former (and re-scaling the bilateral trade weights of each of these three large economies with the other two so that they add to 1).

Finally, **Table 5** shows the way each GVAR is estimated, that is, the form of its representation, the transformation of the variables used and the sample period. The main differences are:

- The DDPS-GVAR is expressed in error correction form, with all variables treated as I(1) except for p, r and lr which are treated as I(2), and is estimated with quarterly data since 1979Q2.

- The Risk-GVAR is expressed in first difference form (i.e, no cointegration allowed), with r,lr and $embi$ treated as I(0), and the rest of variables treated as I(1), including p, and use a much shorter sample, starting in 1999Q1.

- In this way:

Table 4: Variables included in the national satellites

	Domestic block		Foreign block (which/how)	
LABEL	**DDSP-GVAR**	**RISK-GVAR**	**DDSP-GVAR**	**RISK-GVAR**
y	all	all	all / avg	all / avg
p	all	all	all / avg	all - usa / usa
ep1	all - usa	NONE	usa / avg	NONE
epr1	NONE	all-usa-spain-turkey	NONE	usa/avg
epr2	NONE	turkey	NONE	NONE
eq	all	NONE	all / avg	NONE
r	all	all	all / avg	NONE
lr	developed	developed	all / avg	all - usa / usa
embi	NONE	emerging	NONE	emerging/average
Cre	NONE	all	NONE	NONE
Depo	NONE	all	NONE	NONE
poil	usa	usa	all - usa	all - usa
pmetal	usa	NONE	all - usa	NONE
praw	usa	NONE	all - usa	NONE
MAXIMUM	9	10	9	6

- The Risk-GVAR consider inflation and interest rates as mean-reverting variables, which is more consistent with the standard macroeconomic view and, in turn, justified by the predominance of monetary policy frameworks with focus on inflation targeting. This also explains why the data sample starts in 1999Q1: in previous decades monetary regimes experienced substantial structural changes around the world (the most extreme example being the several hyperinflation episodes).
- The Risk-GVAR gives up to estimate cointegration relationships because the estimation sample is too short for these estimates to be reliable, and because it is reasonable to distrust the stability of such relationships for forecasting purposes (as illustrated by Clements and Hendry, 1995).

Table 5: Differences required for stationarity, sample and cointegration

LABEL	**DDSP-GVAR**	**RISK-GVAR**
y	1 diff	1 diff
p	2 diff	1 diff
ep1	1 diff	No included
epr1	No included	1 diff
epr2	No included	1 diff
eq	1 diff	No included
r	1 diff	0 diff
lr	1 diff	0 diff
embi	No included	0 diff
Cre	No included	1 diff
Depo	No included	1 diff
poil	1 diff	1 diff
pmetal	1 diff	No included
praw	1 diff	No included
COINTEGRATION	YES	NO
SAMPLE	1979Q2-2013Q1	2002Q1-2013Q1

6 Backtesting

This section shows the results of evaluating the forecasting performance[18] of the Risk-GVAR against that of the DDPS-GVAR (and some variants) through a back-testing exercise for a sample of variables (interannual growth rates of quarterly GDP and CPI) and countries (USA, Euro Area, Turkey, Mexico and Argentina).

Specifically we show the forecasts for the period 2011Q2-2019Q2 from models estimated with data through 2011Q1 (both, data used in estimation and forecast as a dotted lines in the figures that follow) along with the historical data finally observed for the whole period (continuous line). The small differences between the data used in estimation and the observed historical data is due to historical data revisions by official sources.

The models to be compared, are not only the Risk-GVAR and the DDPS-GVAR but also the following three variants of the DDPS-GVAR, which together help to shed light on the influence of alternative changes in the specification or estimation sample of both models:

- **First-difference DDPS-GVAR**: excluding cointegration terms from the DDPS-GVAR.
- **Short-sample DDPS-GVAR**: same specification as the DDPS-GVAR but estimation sample starting at 2002Q1 instead of 1979Q2.
- **Quasi-Risk-GVAR**: The result of incorporating in the DDPS-GVAR some distinctive features of the Risk-GVAR. Namely, mean reversion of inflation and interest rates, no cointegration, block-recursivity (USA, China and Euro Area, as a block, is not affected by the rest of the world) and much shorter estimation sample (start at 1999Q1).

  Notwithstanding, this variant still differs from the Risk-GVAR in the set of countries, variables and the historical data used for estimation (the DDPS-GVAR is estimated with the data included in the GVAR-Toolbox-2.0, whereas the Risk-GVAR used data that incorporate all the ex-post revisions by the official sources).

We start by back-testing the original DDPS-GVAR, our benchmark. **Figure 2** shows the forecasts of the DDPS-GVAR for GDP growth (y-o-y,%) and inflation (y-o-y,%) against the observed time series. What stands out here, is the very poor performance of DDPS-GVAR at forecasting inflation: in general, the DDPS-GVAR tends to forecast absurdly explosive paths for inflation.

It could be thought that these explosive inflation forecasts are an artifact of using data from the 1980s because there were substantial structural changes in the inflation regimes in the 1980s and 1990s (for one

[18]The section focuses in **unconditional** forecasts, that is, join forecasts for the whole set of variables of the model (as opposed to conditional forecasts: forecasting a subset of variables given the path for the rest of variables of the model).

thing, in the 1980s several emerging economies experienced hyperinflations and the in the 1990s many central banks adopted inflation targeting systems), but as shown below the estimation with data beginning in 2002Q1 (Short-sample DDPS-GVAR) also produces this explosive behavior.

Figure 3 shows a comparison of the backtesting exercise for alternative variants of the DDPS-GVAR. Again, what stands out is the absurdly explosive nature of most inflation forecasts. This means that neither shortening the estimation sample nor excluding cointegration terms makes the forecasting performance of the DDPS-GVAR acceptable.

Now we introduce the Risk-GVAR in the backtesting exercise: firstly, by comparing in **Figure 4** its forecasting performance against that of the original DDSP-GVAR, and secondly, in **Figure 5**, comparing it with the Quasi-Risk-GVAR. The second comparison, in particular, allows us to explore the source of the superior forecasting performance of the Risk-GVAR against the DDPS-GVAR.

Figure 4 shows the general superior performance of the Risk-GVAR against the DDPS-GVAR, especially for inflation, where explosiveness disappears and forecasts look reasonable given the context (inflation is in most cases systematically overestimated, but that is consistent with the abnormally low inflation of the period).

Figure 5 shows that the Risk-GVAR and the Quasi-Risk-GVAR produce very similar forecasts. This proves that the aforementioned superiority is not an artifact of using revised data for estimation in the former case or having fewer countries, but it does seems related to other distinctive characteristics of Risk-GVAR as the mean-reversion assumption about inflation and interest rates and block-recursivity.

Figure 2: *Unconditional forecasts from DDPS-GVAR*

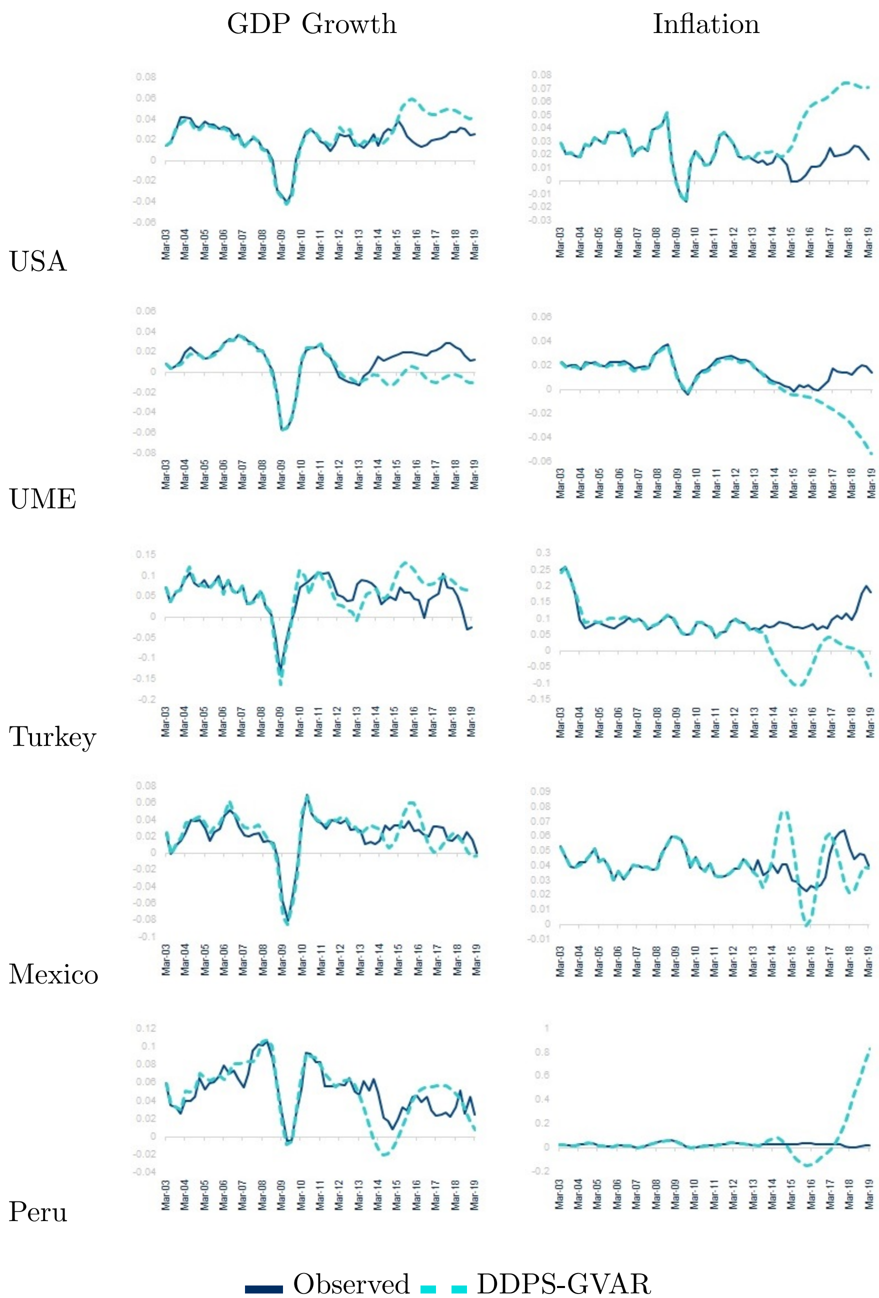


Figure 3: *Unconditional forecasts: DDPS-GVAR versus variants*

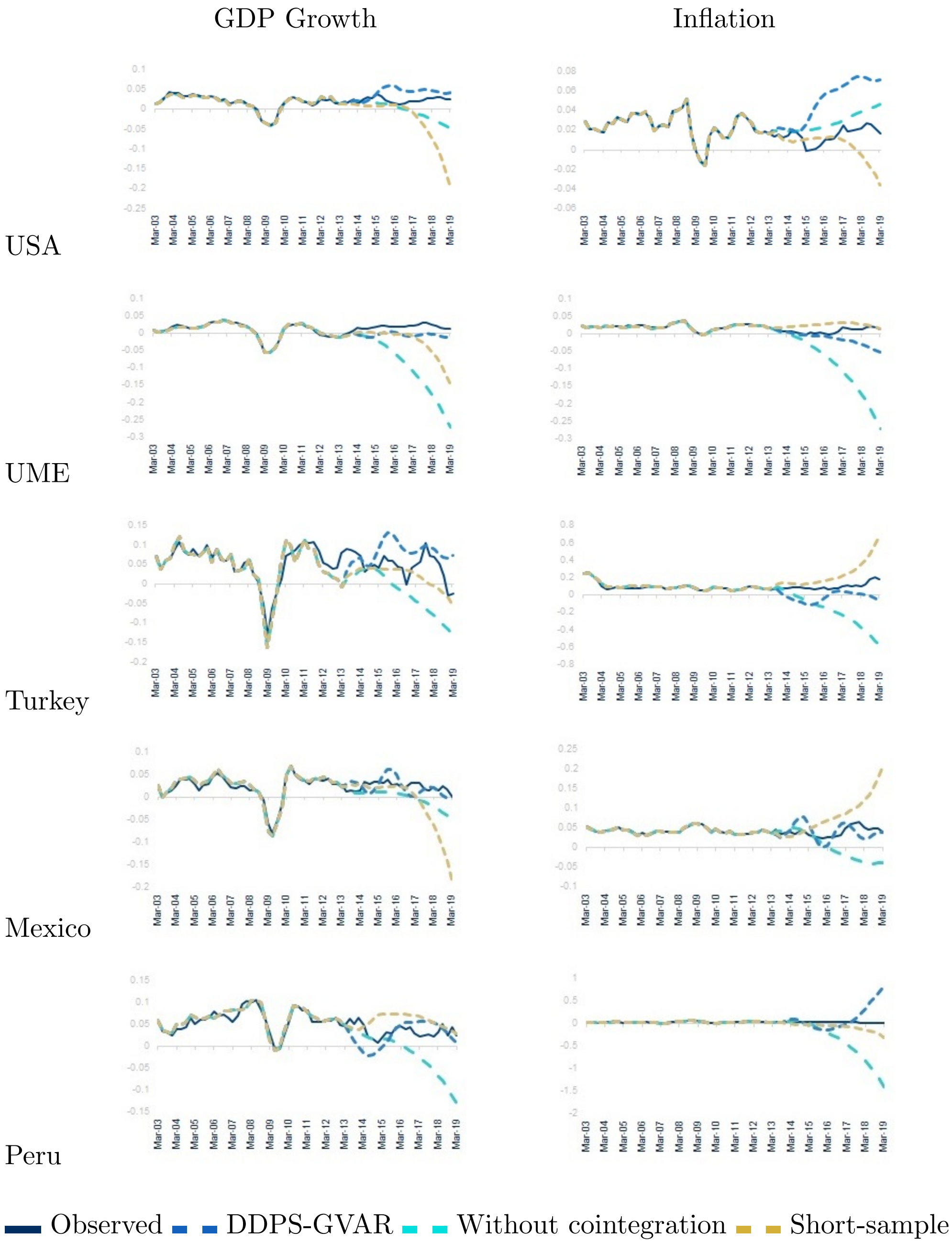


Figure 4: *Unconditional forecasts: DDPS-GVAR versus Risk-GVAR*

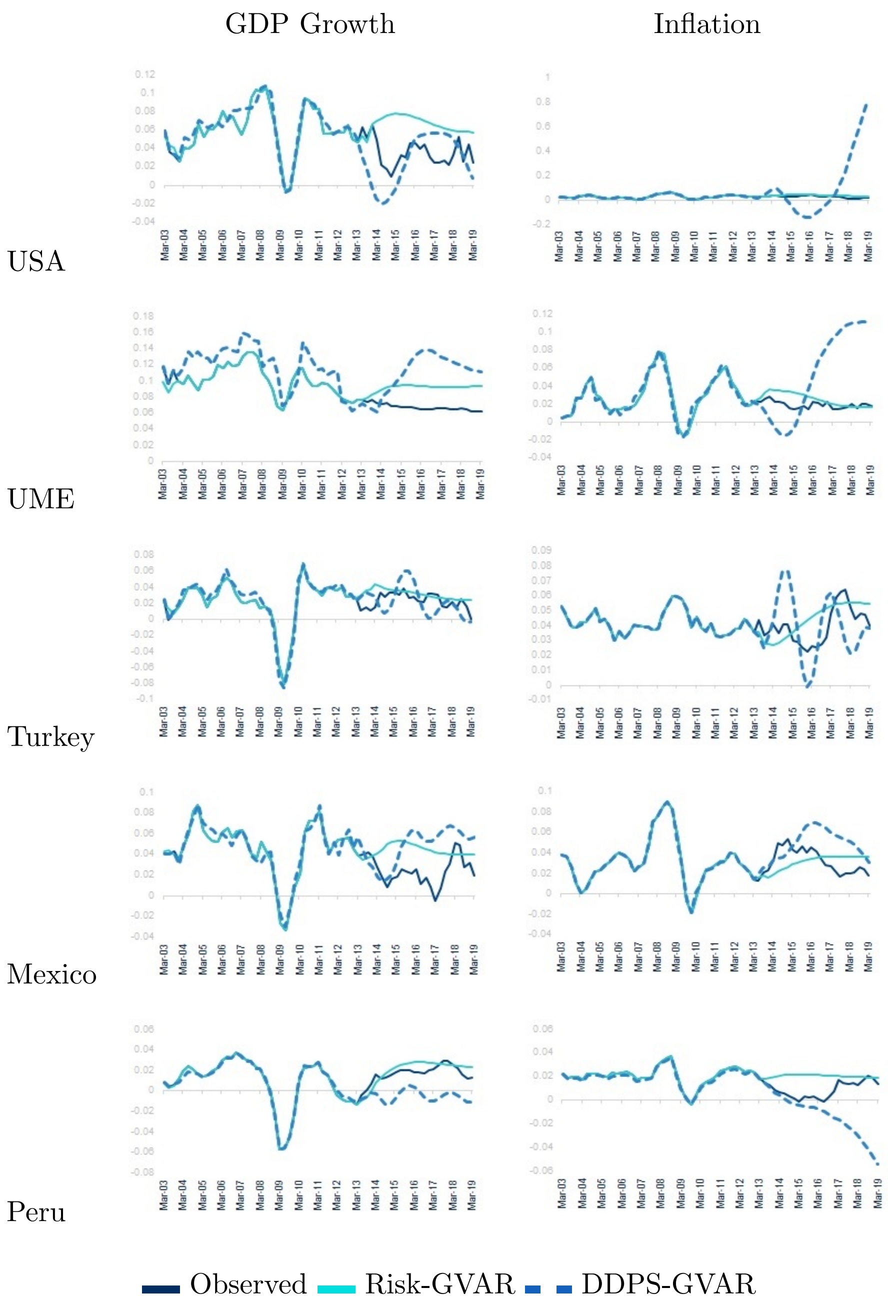
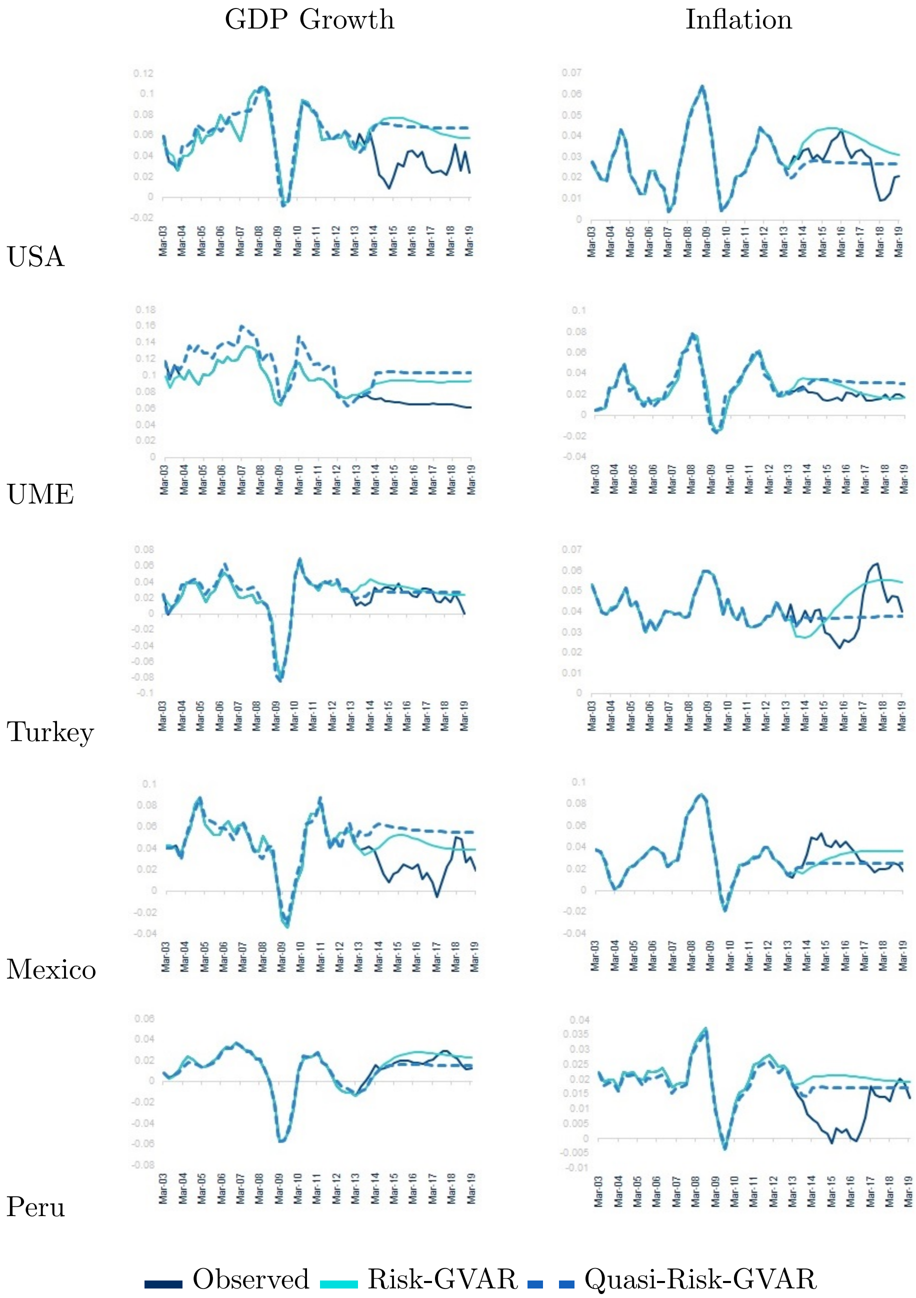


Figure 5: *Unconditional forecasts: Risk-GVAR Versus Quasi-Risk-GVAR*



Finally, the following tables summarize the results of the whole set of backtesting exercises in terms of the **U-Theil** (Theil's U) indicator, defined as the ratio between the root mean squared of the forecast errors (**RMSE**) for the following k periods made by a given model and the **RMSE** of a benchmark model, in our case the DDPS-GVAR. In formal terms:

$$\text{U-Theil(modelname,k)} = \frac{\sqrt{\frac{1}{k}\left[\sum_{i=1}^{k}(g_{t+i} - f_{t+i}^{modelname})^2\right]}}{\sqrt{\frac{1}{k}\left[\sum_{i=1}^{h}(g_{t+i} - f_{t+i}^{DDPS})^2\right]}}$$

Where k is the maximum forecast horizon considered (such that $k = 4$ represents the evaluation of forecasts for the next four periods, i.e, $t+1$, $t+2$, $t+3$, $t+4$, where t is the current date), g_{t+i} represents the observed value to be forecasted at period $t+i$, and $f_{t+i}^{modelname}$ and f_{t+i}^{DDPS} the forecasts of g_{t+i} produced at t (i.e, using only data until date t) by models *modelname* and *DDPS-GVAR*. respectively.

Tables 6 and **7** show the U-Theils resulting from evaluating the forecasts for the period 2013Q2-2016Q2 (until 13 quarters ahead) of the inter-annual growth rates of the quarterly GDP and CPI from the different models.

TABLE 6: U-Theil for GDP growth rate forecasts for the 2013Q2-2016Q2

	Average		Countries							
	All	Excluded-Arg	USA	EMU	China	Turkey	Mexico	Peru	Chile	Argentina
DDPS-GVAR	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Short-sample DDPS-GVAR	2,06	1,27	1,74	1,96	0,57	1,76	2,19	0,32	1,99	4,22
Quasi-Risk-GVAR	0,98	0,37	0,38	0,18	0,92	0,24	0,16	0,13	0,04	2,68
Risk-GVAR	1,01	0,85	0,24	0,11	0,68	2,09	1,38	0,23	1,47	1,46

TABLE 7: U-Theil for CPI growth rate forecasts for the 2013Q2–2016Q2

	Average		Countries							
	All	Excluded-Arg	USA	EMU	China	Turkey	Mexico	Peru	Chile	Argentina
DSPS-GVAR	1,00	1.00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Short-sample DSPS-GVAR	2,56	2.55	0,68	5,04	2,88	2,66	6,66	1,01	3,91	2,72
Quasi-Risk-GVAR	0,59	0.49	0,71	2,45	1,49	0,21	0,88	0,12	1,62	3,15
Risk-GVAR	0,58	0.53	0,78	3,24	1,57	0,22	0,66	0,36	1,57	1,69

On average, the Risk-GVAR and Cuasi-Risk-GVAR show similar performance at forecasting GDP growth as the DDPS-GVAR, but a much better performance at forecasting inflation. The average forecasting performance of the short-sample DDPS-GVAR is much worse than that of the rest of models.

Table 8 and **Table 9** extend the forecasted period to 2019:Q2 (until 25 quarters ahead), i.e, it evaluates the forecasting performance for longer time horizons. Again, on average, the Risk-GVAR and the Quasi-Risk-GVAR show a comparable performance at forecasting GDP-growth but a much better performance at forecasting inflation than the DDPS-GVAR (which is in turn much better than its short-sample counterpart).

TABLE 8: U-Theil for GDP growth rate forecasts for the 2013Q2–2019Q2

	Average		Countries							
	All	Excluded-Arg	USA	EMU	China	Turkey	Mexico	Peru	Chile	Argentina
DSPS-GVAR	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Short-sample-DSPS-GVAR	2,29	1,38	3,18	3,29	0,12	1,33	38,74	0,08	40,64	7,04
Cuasi-Risk-GVAR	1,01	0,72	0,29	0,19	0,77	0,54	3,44	0,74	4,32	2,49
Risk-GVAR	0,91	0,84	0,20	0,14	0,56	1,27	7,90	0,75	23,78	1,26

TABLE 9: U-Theil for CPI growth rate forecasts for the 2013Q2–2019Q2

	Average		Countries							
	All	Excluded-Arg	USA	EMU	China	Turkey	Mexico	Peru	Chile	Argentina
DSPS-GVAR	1,00	1.00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Short-sample-DSPS-GVAR	2,75	3.03	2,84	3,69	1,05	22,88	42,31	0,37	51,06	1,24
Cuasi-Risk-GVAR	0,74	0.73	3,05	1,20	0,58	1,48	2,56	0,05	13,10	0,83
Risk-GVAR	0,69	0.71	3,66	1,69	0,31	0,18	4,19	0,09	12,06	0,57

7 Summary and conclusions

The paper has presented and described the Risk-GVAR-1.0, a macroeconometric model designed for lending support to a systemic global European bank in meeting the requirements of the **Internal Capital Adequacy Assessment Process** (**ICAAP**) which they are being compelled to undertake on a regular basis by the European banking supervisory authorities as part of their regulatory response to the lessons of the **Global Financial Crisis** of 2007-2008.

A key element of the ICAAP is the performance of regular stress test exercises based on scenario analysis, which involves the regular creation of updated risk macroeconomic scenarios for a wide range of international macroeconomic variables, something that can only be done in a cost-efficient way with the support of a global macroeconometric model, which by its mathematical nature makes the process of generating these scenarios transparent, auditable and automatic.

Additionally, arguments have been presented to justify the selection of a Global Vector Autoregressive model (GVAR) as the type of model to develop against the main alternatives. In particular, it has been explained that the empirical nature of a GVAR makes it more appropriate for forecasting purposes than the alternative more theoretical macroeconometric models.

Finally, the superior forecasting performance of the Risk-GVAR-1.0 was shown in comparisons with the most standard and popular GVAR developed by Di Mauro, Dees, Pesaran and Smith (2007).

8 References

Antolín-Díaz, Juan.; Petrella, Ivan. and Rubio-Ramírez, Juan. (2020): "Structural scenario analysis with SVARs", *Journal of Monetary Economics*, in press.

BIS (2000): "Stress Testing by large financial institutions: current practice and aggregation issues", *Bank for International Settlements* (Committee of the Global Financial System). Available online here.

Chudik, Alexander and Smith, Vanessa. (2013), "The GVAR approach and the dominance of US Economy", *Globalization and Monetary Policy Institute Working Paper 136*, Federal Reserve Bank of Dallas.

Clements, Michael. and Hendry, David. (1995), "Forecasting in cointegrated systems", *Journal of Applied Econometrics*, 10(2).

Cushman, David and Zha, Tao. (1997), "Identifying monetary policy in a small open economy under flexible exchange rates", *Journal of Monetary Economics*, 39(3). Link.

Dees, Stephane., Di Mauro, Filippo., Pesaran, Hashem and Smith, Vanessa. (2007), "Exploring the international linkages of the euro area: a global VAR analysis", *Journal of Applied Econometrics, 22(1).*

EBA (2016), "Final Report on Guidelines on ICAAP and ILAAP information collected for SREP purposes",*European Banking Authority*. Link.

EBA (2018), "Guidelines on Institutions' Stress Testing",*European Banking Authority*. Link.

Eickmeier, Sandra and Ng, Tim. (2015), "How do US credit supply shocks propagate internationally? A GVAR approach", *European Economic Review, 74 (February 2015).*

Garrat, Anthony., Lee, Kevin., Pesaran, Hashem and Shin, Yongcheol. (2006), "GLOBAL MACROECONOMETRIC MODELLING: A Long Run Structural Approach", *Oxford University Press.*

Martin, Florian and Crespo Cuaresma, Jesús.(2016), "Weighting schemes in global VAR modelling: a forecasting exercise", *Letters in Spatial and Resource Sciences*, 10, 45-56.

Pesaran, Hashem., Schuermann, Til and Weiner, Scott. (2001), "Modelling regional interdependencies using a Global Error-Correcting Macroeconometric Model", *Journal of Business and Economic Statistics*, 22.

Sims, Christopher. (1980), "Macroeconomics and reality", *Econometrica*, 48(1).

Sims, Christopher. (2002),"The role of models and probabilities in the monetary policy process", BPEA,2(2002). Link.

Smith, Vanessa. and Galesi, Alessandro. (2014), "GVAR Toolbox 2.0 User Guide", included in the installation folder of the GVAR-Toolbox-2.0. Link.

Working paper

2021

21/08 **Rodolfo Méndez-Marcano:** A global vector autoregressive model for banking stress testing.

21/07 **Ali B. Barlas, Seda Guler Mert, Berk Orkun Isa, Alvaro Ortiz, Tomasa Rodrigo, Baris Soybilgen and Ege Yazgan:** Big Data Information and Nowcasting: Consumption and Investment from Bank Transactions in Turkey.

21/06 **Ángel de la Fuente y Rafael Doménech:** El nivel educativo de la población en España y sus regiones: actualización hasta 2019.

21/05 **Saidé Salazar, Jaime Oliver, Álvaro Ortiz, Tomasa Rodrigo and Ignacio Tamarit:**
ESP/ Patrones de Consumo de Efectivo vs Tarjeta en México: una aproximación Big Data.
ING / Cash Vs Card Consumption Patterns in Mexico: A Machine Learning Approach.

21/04 **Ángel de la Fuente:** La financiación autonómica en 2020: una primera aproximación y una propuesta de cara a 2021.

21/03 **Ángel de la Fuente:** Las finanzas autonómicas en 2020 y entre 2003 y 2020.

21/02 **Joxe Mari Barrutiabengoa, J. Julián Cubero and Rodolfo Méndez-Marcano:** Output-side GHG Emissions Intensity: A consistent international indicator.

21/01 **Ángel de la Fuente y Pep Ruiz:** Series largas de VAB y empleo regional por sectores, 1955-2019 Actualización de *RegData-Sect* hasta 2019.

2020

20/17 **Amparo Castelló-Climent and R. Doménech:** Human Capital and Income Inequality Revisited.

20/16 **J.E. Boscá, R. Doménech, J. Ferri, J.R. García and C. Ulloa:** The Stabilizing Effects of Economic Policies in Spain in Times of COVID-19.

20/15 **Ángel de la Fuente:** La evolución de la financiación de las comunidades autónomas de régimen común, 2002-2018.

20/14 **Ángel de la Fuente:** El impacto de la crisis del Covid sobre el PIB de las CCAA en 2020: una primera aproximación.

20/13 **Ali B. Barlas, Seda Guler Mert, Álvaro Ortiz and Tomasa Rodrigo:** Investment in Real Time and High Definition: A Big Data Approach.

20/12 **Félix Lores, Pep Ruiz, Angie Suárez y Alfonso Ugarte:** Modelo de precios de la vivienda en España. Una perspectiva regional.

20/11 **Ángel de la Fuente:** Series largas de algunos agregados económicos y demográficos regionales: Actualización de RegData hasta 2019.

20/10 **Ángel de la Fuente:** La liquidación de 2018 del sistema de financiación de las comunidades autónomas de régimen común.

20/09 **Lucía Pacheco Rodríguez and Pablo Urbiola Ortún:** From FinTech to BigTech: an evolving regulatory response.

20/08 **Federico D. Forte:** Network Topology of the Argentine Interbank Money Market.

20/07 **Ángel de la Fuente:** Las finanzas autonómicas en 2019 y entre 2003 y 2019.

20/06 **Vasco M. Carvalho, Juan R. Garcia, Stephen Hansen, Álvaro Ortiz, Tomasa Rodrigo, José V. Rodríguez Mora and Pep Ruiz:** Tracking the COVID-19 Crisis with High-Resolution Transaction Data.

20/05 **Jinyue Dong and Le Xia:** Forecasting modeling for China's inflation.

20/04 **Le Xia:** Lessons from China's past banking bailouts.

20/03 **Ángel de la Fuente y Pep Ruiz:** Series largas de VAB y empleo regional por sectores, 1955-2018. RegData_Sect FEDEA-BBVA (v5.0_1955-2018).

20/02 **Luis Antonio Espinosa y Juan José Li Ng:**
ESP/ El riesgo del sargazo para la economía y turismo de Quintana Roo y México.
ING / The risk of sargassum to the economy and tourism of Quintana Roo and Mexico.

20/01 **Ángel de la Fuente:** La dinámica territorial de la renta en España, 1955-2018. Los determinantes directos de la renta relativa: productividad, ocupación y demografía.

CLICK HERE TO ACCESS THE WORKING DOCUMENTS PUBLISHED IN
Spanish and English

Creating Opportunities

DISCLAIMER

The present document does not constitute an “Investment Recommendation”, as defined in Regulation (EU) No 596/2014 of the European Parliament and of the Council of 16 April 2014 on market abuse (“MAR”). In particular, this document does not constitute “Investment Research” nor “Marketing Material”, for the purposes of article 36 of the Regulation (EU) 2017/565 of 25 April 2016 supplementing Directive 2014/65/EU of the European Parliament and of the Council as regards organisational requirements and operating conditions for investment firms and defined terms for the purposes of that Directive (MIFID II).

Readers should be aware that under no circumstances should they base their investment decisions on the information contained in this document. Those persons or entities offering investment products to these potential investors are legally required to provide the information needed for them to take an appropriate investment decision.

This document has been prepared by BBVA Research Department. It is provided for information purposes only and expresses data or opinions regarding the date of issue of the report, prepared by BBVA or obtained from or based on sources we consider to be reliable, and have not been independently verified by BBVA. Therefore, BBVA offers no warranty, either express or implicit, regarding its accuracy, integrity or correctness.

This document and its contents are subject to changes without prior notice depending on variables such as the economic context or market fluctuations. BBVA is not responsible for updating these contents or for giving notice of such changes.

BBVA accepts no liability for any loss, direct or indirect, that may result from the use of this document or its contents.

This document and its contents do not constitute an offer, invitation or solicitation to purchase, divest or enter into any interest in financial assets or instruments. Neither shall this document nor its contents form the basis of any contract, commitment or decision of any kind.

The content of this document is protected by intellectual property laws. Reproduction, transformation, distribution, public communication, making available, extraction, reuse, forwarding or use of any nature by any means or process is prohibited, except in cases where it is legally permitted or expressly authorised by BBVA.

ENQUIRIES TO:
BBVA Research: Azul Street. 4. La Vela Building – 4th and 5th floor. 28050 Madrid (Spain).
Tel.:+34 91 374 60 00 y +34 91 537 70 00 / Fax:+34 91 374 30 25
bbvaresearch@bbva.com www.bbvaresearch.com