

Weighting schemes in global VAR modelling: a forecasting exercise

Florian Martin¹ · Jesús Crespo Cuaresma^{1,2,3,4}

Received: 26 November 2015 / Accepted: 13 March 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract We provide a comprehensive analysis of the out-of-sample predictive accuracy of different global vector autoregressive (GVAR) specifications based on alternative weighting schemes to address global spillovers across countries. In addition to weights based on bilateral trade, we entertain schemes based on different financial variables and geodesic distance. Our results indicate that models based on trade weights, which are standard in the literature, are systematically outperformed in terms of predictive accuracy by other specifications. We find that, while information on financial linkages helps improve the forecasting accuracy of GVAR models, averaging predictions by means of simple predictive likelihood weighting does not appear to systematically lead to lower forecast errors.

Keywords Global VAR modelling · Forecasting · Global spillovers

JEL Classification C31 · E44

The authors would like to thank two anonymous referees for helpful comments, Sandra Eickmeier and Tim Ng for providing some of the data used in earlier versions of this paper, as well as the participants in the special session in honour of Manfred M. Fischer at the 2015 Congress of the European Regional Science Association (ERSA) in Lisbon for useful discussion of an earlier draft of this piece.

✉ Jesús Crespo Cuaresma
jcrespo@wu.ac.at

¹ Vienna University of Economics and Business (WU), Vienna, Austria

² Wittgenstein Center for Demography and Global Human Capital (IIASA, VID/OEAW, WU), Vienna, Austria

³ International Institute of Applied Systems Analysis (IIASA), Vienna, Austria

⁴ Austrian Institute of Economic Research (WIFO), Vienna, Austria

1 Introduction

In recent years, exploring interdependencies across countries and markets has become one of the major fields of research in economics. In an increasingly globalized world, interactions between economies are becoming more important as business cycles harmonise among major economies and shocks spread quickly across the globe. In the attempt to analyse and understand the dynamics that drive an interconnected global economy, researchers have developed a range of econometric approaches. Modelling the complex linkages of a multitude of countries quickly results in econometric specifications with many more variables and parameters than observations, thus making models which are not sparsely parametrized impracticable.

One way to limit the parameter space in econometric modelling exercises aimed at assessing cross-country linkages is to use combinations of variables in surrounding economies as regressors in order to model the interdependencies among markets. An obvious choice is to use weighted averages of variables as this allows capturing information of other regions in each single country model, much like in a spatial econometric model. The most common approach found in the literature using such a technique is given by global vector autoregressive (or GVAR) specifications (see for example [Pesaran et al. 2004a](#); [Garrat et al. 2006](#); [Dees et al. 2007a](#)). The standard method of averaging foreign variables in GVAR specifications is based on trade weights and in most applications on the share of bilateral trade between two countries in relation to the total trade volume of each one of those countries. Given the importance of trade as a major driver of business-cycle co-movement ([Baxter and Kouparitsas 2005](#)), basing an empirical model on trade weights is certainly a reasonable choice. However, the recent financial crisis has demonstrated that financial linkages appear to be an important determinant of macroeconomic spillovers.

As such, financial linkages have also been put forward as an important channel of international shock transmission. [Van Rijckeghem and Weder \(2001\)](#) argue that multiple financial mechanisms can cause cross-border spillovers. For example, banks that incur losses in one country could see themselves forced to sell assets in other countries in order to fulfil their capital requirements. This logic also applies to other investors receiving margin calls ([Calvo 1999](#)). [Forbes and Chinn \(2004\)](#) come to a mixed conclusion concerning the explanatory power of financial cross-country linkages. They find that direct trade still appears to be the most important driver of regional interdependencies and that financial linkages have a comparably minor effect. Their results suggest that bank lending can have a considerable effect but that its magnitude is very dependent on model specification.

Since the selection of the appropriate weighting scheme in large GVAR models tends to be done in an *ad hoc* fashion, in this contribution we investigate how weighting schemes affect the out-of-sample forecasting performance of standard GVAR models. Given that the choice of the optimal weighting scheme is basically a problem of model uncertainty,¹ we explore whether a simple model averaging approach based on

¹ In the framework of spatial autoregressive models, [LeSage and Pace \(2009\)](#) and [Crespo Cuaresma and Feldkircher \(2013\)](#) provide discussions to model averaging in the context of uncertainty about spatial linkages.

the predictive likelihood of GVAR specifications can outperform individual weighting schemes in terms of out-of-sample predictive ability. We rely on the setting put forward by Eickmeier and Ng (2015), so our results can also be seen as an extension and robustness check of this piece. Unlike most other empirical studies using GVAR models, Eickmeier and Ng (2015) consider linkages in the individual country models that go beyond standard trade weights. Instead, the authors assess whether combinations of trade weights and financial weights can outperform the traditional specification. The results in Eickmeier and Ng (2015) support the superiority of these alternative linkage specifications over the traditional trade weights.

In this study we entertain a greater variety of model specifications as in Eickmeier and Ng (2015), as not only combinations of trade and financial weights are analysed but also models based on pure financial linkages. In addition, simple model averaging techniques based on predictive likelihood weighting are applied in order to deal with the issue of model uncertainty. Our results indicate that GVAR models based on standard trade weights achieve inferior predictive accuracy as compared to simple weighting schemes that rely on information about geodesic distance or bilateral financial linkages. The results of our out-of-sample forecasting exercise do not support the use of averaging forecasts using predictive likelihood as an instrument to achieve improvements in predictive accuracy.

The structure of the paper is as follows. Section 2 gives a brief introduction to GVAR modelling and the estimation framework. Section 3 provides information regarding the data used and the specification of the individual country models. The results of the forecasting exercise are presented in Sects. 4 and 5 concludes the paper.

2 GVAR modelling: the framework

The GVAR approach was originally proposed by Pesaran et al. (2004b) and constitutes a large-dimensional but simple model for modelling complex interrelated systems such as the global economy. One key feature of a GVAR model is that it allows for interdependence at multiple levels, thus allowing national and international dynamics to be empirically evaluated in a consistent and transparent manner. Additionally, GVAR specifications enable to reflect dynamics that are consistent with theory (long-run equilibria), while matching short-run adjustment dynamics.

A GVAR model consists of a set of individual country VARX* models, which are linked in order to yield a global model. For each country a VARX*(p_i, q_i) specification is constructed as follows

$$\mathbf{x}_{it} = \mathbf{a}_{i0} + \mathbf{a}_{i1}t + \Phi_{i1}\mathbf{x}_{i,t-1} + \dots + \Phi_{ip_i}\mathbf{x}_{i,t-p_i} + \Lambda_{i0}\mathbf{x}_{it}^* + \Lambda_{i1}\mathbf{x}_{i,t-1}^* + \dots + \Lambda_{iq_i}\mathbf{x}_{i,t-q_i}^* + \mathbf{u}_{it}, \tag{1}$$

where \mathbf{x}_{it} is a k -dimensional column vector of domestic variables for country i in period t , \mathbf{a}_{i0} is a vector of constants, $\mathbf{a}_{i1}t$ is a linear trend, \mathbf{x}_{it}^* are k^* -dimensional column vectors of weighted foreign variables (assumed weakly exogenous) and \mathbf{u}_{it} is a k -dimensional column vector of serially uncorrelated error terms. Φ_{it} and Λ_{it} are the corresponding coefficient matrices. The foreign variables \mathbf{x}_{it}^* in a GVAR model are constructed as weighted averages of other countries' domestic variables

$$\mathbf{x}_{it}^* = \sum_{j=0}^N w_{ij} \mathbf{x}_{jt}, \quad w_{ii} = 0 \tag{2}$$

with w_{ij} , $j = 0, 1, \dots, N$ being a set of weights such that $\sum_{j=0}^N w_{ij} = 1$.

Defining a vector that stacks domestic and foreign variables, $\mathbf{z}_{it} = (\mathbf{x}_{it} \ \mathbf{x}_{it}^*)'$, we can write Eq. (1) as

$$\mathbf{A}_{i0} \mathbf{z}_{it} = \mathbf{a}_{i0} + \mathbf{a}_{i1} t + \mathbf{A}_{i1} \mathbf{z}_{it-1} + \dots + \mathbf{A}_{ip_i} \mathbf{z}_{it-p_i} + \mathbf{u}_{it}, \tag{3}$$

where

$$\mathbf{A}_{i0} = (\mathbf{I}_{k_i}, -\Lambda_{i0}), \quad \mathbf{A}_{ij} = (\mathbf{\Phi}_{ij}, \Lambda_{ij}) \quad \text{for } j = 1, \dots, p_i.$$

Using the link matrix \mathbf{W}_i , \mathbf{z}_{it} can be written as $\mathbf{z}_{it} = \mathbf{W}_i \mathbf{x}_t$, where \mathbf{x}_t is a $K \times 1$ vector including all endogenous variables of the system and \mathbf{W}_i is a $(k_i + k_i^*) \times k$ matrix which contains the weights capturing bilateral exposures between the countries under investigation.

Using this transformation, Eq. (1) can be written as

$$\mathbf{A}_{i0} \mathbf{W}_i \mathbf{x}_t = \mathbf{a}_{i0} + \mathbf{a}_{i1} t + \mathbf{A}_{i1} \mathbf{W}_i \mathbf{x}_{t-1} + \dots + \mathbf{A}_{ip_i} \mathbf{W}_i \mathbf{x}_{t-p_i} + \mathbf{u}_{it}, \quad \text{for } i = 0, 1, 2, \dots, N, \tag{4}$$

which yields the global model when stacking all the individual country VARX* models. This global specification is specified as

$$\mathbf{G}_0 \mathbf{x}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{G}_1 \mathbf{x}_{t-1} + \dots + \mathbf{G}_p \mathbf{x}_{t-p} + \mathbf{u}_t, \tag{5}$$

where $\mathbf{G}_0 = (\mathbf{A}_{00} \mathbf{W}_0 \ \mathbf{A}_{10} \mathbf{W}_1 \ \dots \ \mathbf{A}_{N0} \mathbf{W}_N)'$, $\mathbf{G}_j = (\mathbf{A}_{0j} \mathbf{W}_0 \ \mathbf{A}_{1j} \mathbf{W}_1 \ \dots \ \mathbf{A}_{Nj} \mathbf{W}_N)'$ for $j = 1, 2, \dots, p$; $\mathbf{a}_0 = (\mathbf{a}_{00} \ \mathbf{a}_{10} \ \dots \ \mathbf{a}_{N0})'$, $\mathbf{a}_1 = (\mathbf{a}_{01} \ \mathbf{a}_{11} \ \dots \ \mathbf{a}_{N1})'$, $\mathbf{u}_t = (\mathbf{u}_{0t} \ \mathbf{u}_{1t} \ \dots \ \mathbf{u}_{Nt})'$ and $p = \max p_i$ across all i . In general $p = \max(\max p_i, \max q_i)$. Premultiplying with \mathbf{G}_0^{-1} yields the autoregressive representation of the GVAR(p) model

$$\mathbf{x}_t = \mathbf{b}_0 + \mathbf{b}_1 t + \mathbf{F}_1 \mathbf{x}_{t-1} + \dots + \mathbf{F}_p \mathbf{x}_{t-p} + \boldsymbol{\epsilon}_t, \tag{6}$$

where $\mathbf{b}_0 = \mathbf{G}_0^{-1} \mathbf{a}_0$, $\mathbf{b}_1 = \mathbf{G}_0^{-1} \mathbf{a}_1$, $\mathbf{F}_j = \mathbf{G}_0^{-1} \mathbf{G}_j$ for $j = 1, 2, \dots, p$ and $\boldsymbol{\epsilon}_t = \mathbf{G}_0^{-1} \mathbf{u}_t$. Once estimates of the parameters are available, Eq. (6) can be solved recursively and used for producing forecasts and constructing impulse response functions in the framework of GVAR models.

The standard procedure to specify and estimate a GVAR model proposed by Pesaran et al. (2004b) starts by estimating each country-specific VARX* model separately in its error correction form,

$$\Delta \mathbf{x}_{it} = \mathbf{c}_{i0} - \boldsymbol{\alpha}_i \boldsymbol{\beta}_i' [\mathbf{z}_{i,t-1} - \boldsymbol{\gamma}_i(t-1)] + \mathbf{A}_{i0} \Delta \mathbf{x}_{it}^* + \boldsymbol{\Gamma}_i \Delta \mathbf{z}_{i,t-1} + \mathbf{u}_{it}, \tag{7}$$

where $\mathbf{z}_{it} = (\mathbf{x}'_{it} \ \mathbf{x}'_{it}^*)'$, $\boldsymbol{\alpha}_i$ is a $k_i \times r_i$ matrix of rank r_i and $\boldsymbol{\beta}_i$ is a $(k_i + k_i^*) \times r_i$ matrix of rank r_i . Conditional on \mathbf{x}_{it}^* the individual VARX* models are estimated

using reduced rank regression, treating the foreign variables as weakly exogenous. Johansen's trace statistic is used to determine the rank order of each country VARX* model. The lag orders of the domestic and foreign variables, p_i and q_i respectively, are determined using Akaike's information criterion, in our application with an assumed maximum lag order of $p = 2$ and $q = 1$.

3 Data and empirical strategy

The dataset used for the forecasting exercise is based on an updated version of that used by Dees et al. (2007b).² The data spans information on 33 countries, representing about 90% of world output.³ The GVAR models entertained are thus composed by 33 individual VARX* models, since we do not aggregate countries to world regions. The following variables enter each country's VARX* model: (log) real GDP, inflation, (log) real equity prices, (log) real exchange rate, short term interest rate, long term interest rate and the (log) price of oil. Inflation is calculated as the first difference in log CPI and the short and long term interest rates are quarterly (not annualized) rates.⁴

Most existing applications of GVAR models make use of trade weights as their default weighting scheme. Eickmeier and Ng (2015) have been among the first to test how the choice of weighting schemes in GVAR models affects the inference carried out for a given set of data. Besides the standard trade weights, Eickmeier and Ng (2015) also consider specifications based on bilateral portfolio investment, foreign direct investment and banking claims. We apply weighting schemes based on the same type of bilateral financial variables and additionally consider weights based on geodesic distance and the cost of trade.⁵ Consequently we employ nine different weighting concepts, namely

- Bilateral trade flows (*Trade*),
- Inward portfolio investment (*PIin*),
- Outward portfolio investment (*PIout*),
- Inward foreign direct investment (*FDIin*),
- Outward foreign direct investment (*FDIout*),
- Inward banking claims (*BCin*),
- Outward banking claims (*BCout*),
- Trade costs (*TC*),
- Geodesic distance (*GD*).

Bilateral portfolio investment data are obtained from Table 8 of the IMF Coordinated Portfolio Investment Survey (CPIS). Bilateral foreign direct investment data

² The dataset is available from the GVAR Toolbox 2.0 (Smith and Galesi 2014). The GVAR Toolbox was employed in the estimation of the models in the forecasting exercise.

³ The economies included in the sample are Argentina, Australia, India, Austria, Belgium, Brazil, Philippines, Canada, China, Chile, Finland, France, Germany, Indonesia, Italy, Japan, Korea, Malaysia, Mexico, Netherlands, Norway, New Zealand, Peru, South Africa, Saudi Arabia, Singapore, Spain, Sweden, Switzerland, Thailand, Turkey, United Kingdom and the USA.

⁴ For sources and detailed information concerning the dataset, see Dees et al. (2007b).

⁵ The actual weights are calculated from inverse geodesic distance and inverse cost of trade.

Table 1 Weighting schemes in the analysis

Concept	Trade weights		Other weighting concept	
	All var.	GDP and Infl.	All var.	Financial var.
Trade	x			
TC			x	
GD			x	
Pfin			x	
PIout			x	
FDIin			x	
FDIout			x	
BCin			x	
BCout			x	
Pfin and trade		x		x
PIout and trade		x		x
FDIin and trade		x		x
FDIout and trade		x		x
BCin and trade		x		x
BCout and trade		x		x

is from Table 6.1-o of the IMF Foreign Direct Investment Survey (CDIS) and data on international claims of banking groups with headquarters in a particular country (banking claims) are from BIS international banking statistics (IBS).⁶ Information on bilateral cost of trade was obtained from the World Bank UNESCAP trade costs database. Geodesic distances have been derived from CEPPI's GeoDist dataset. The trade weight data are sourced from [Dees et al. \(2007b\)](#).

When specifying the GVAR model, financial weights can either be applied to all variables or only to the financial covariates (interest rates, equity prices and exchange rates), while using trade weights for GDP and inflation. [Eickmeier and Ng \(2015\)](#) only consider the latter case, while we allow for both specifications in our forecasting exercise. Therefore, together with the benchmark trade weights model and the models based on trade costs and geodesic distance, 15 different alternative specifications of the weighting scheme are entertained in our analysis (see [Table 1](#)).

Since the financial weights are not available for very long horizons, we use fixed weights throughout the analysis. In order to avoid any potential time-related bias in our comparative analysis we average all weights over the same time span 2010 to 2011. [Table 2](#) contains the correlations for the different weighting matrices. As expected, the correlation coefficients are all positive. With a correlation coefficient of 0.61, inward foreign direct investment has the highest correlation with trade weights. Overall the financial weights are less correlated with each other than with the trade weights. The lowest correlation can be observed between the financial linkages and the ones based

⁶ The CPIS and CDIS datasets occasionally contain negative gross exposures. We set these observations to zero.

Table 2 Cross-correlations of different weighting matrices

Model	Trade	TC	GD	PIin	PIout	FDIin	FDIout	BCin	BCout
Trade	1.00	0.45	0.51	0.56	0.57	0.61	0.52	0.46	0.56
TC	0.45	1.00	0.44	0.25	0.27	0.24	0.22	0.29	0.26
GD	0.51	0.44	1.00	0.21	0.26	0.41	0.41	0.23	0.27
PIin	0.56	0.25	0.21	1.00	0.76	0.63	0.47	0.55	0.59
PIout	0.57	0.27	0.26	0.76	1.00	0.64	0.50	0.55	0.61
FDIin	0.61	0.24	0.41	0.63	0.64	1.00	0.66	0.73	0.50
FDIout	0.52	0.22	0.41	0.47	0.50	0.60	1.00	0.46	0.54
BCin	0.46	0.29	0.23	0.55	0.55	0.73	0.46	1.00	0.36
BCout	0.56	0.26	0.27	0.59	0.61	0.50	0.54	0.36	1.00

on trade costs and geodesic distance. Given the high mobility of capital, this stylized fact can be considered to be in line with standard economic theory.

All country-specific VARX* models, with the exception of that corresponding to the US, contain GDP, inflation, short-term interest rates, long-term government bond yields, equity prices and the bilateral real exchange rate as endogenous variables. Additionally, they include foreign aggregates of these variables and the oil price as weakly exogenous variables. In the forecasting application, the low interest environment together with a widespread economic downturn after the financial crisis often cause negative forecasts for interest rates and inflation. Zero lower bounds for the short-term and long-term interest rates and the rate of inflation have been imposed in the prediction step.

In addition to obtaining forecasts based on GVAR models for a given weighting scheme, we also attempt at integrating the uncertainty about global spillovers making use of predictive likelihood weighting in the spirit of [Kapetanios et al. \(2006\)](#). For this purpose, we construct a training sample within our in-sample period where we obtain measures of out-of-sample predictive accuracy for each variable and each GVAR model. Using these, we construct weights for each model based on the out-of-sample predictive likelihood, so that those models with better predictive ability in the training sample receive higher weights (see for details [Kapetanios et al. 2006](#)). We construct two types of model-averaged predictions, one based on the overall predictive likelihood across all variables and another one where each variable receives a different weight based on the variable-specific forecast errors in the training sample. Furthermore, we also compare the forecasting accuracy of the models entertained with that of a simple unweighted average of predictions.

4 Out-of-sample forecasting results

In order to assess the forecasting ability of GVAR models with different weighting schemes, we use the period starting in the first quarter of 1980 to the first quarter of 2011, leaving eight quarters (2011Q2–2013Q1) available for the out-of-sample

forecasting evaluation. The GVAR model is estimated separately for each one of the weighting schemes described above. We assess forecasting accuracy using the root mean squared forecasting errors (RMSE) relative to the RMSE that would be obtained using a random walk prediction. The Diebold-Mariano test statistic (Diebold and Mariano 2002) is used to address whether differences in forecasting ability are statistically significant. For the forecast averaging techniques used in the study (predictive likelihood averaging based on all variables and predictive likelihood averaging based on variable-specific weights as well as simple unweighted model averaging) we use the period spanned from the second quarter of 2010 to the first quarter of 2011 as a training sample.

The results for the relative RMSEs of the different models for forecasting horizons of one, two, four and eight quarters ahead are shown in Table 3. Values below one indicate that the corresponding model outperforms the benchmark random walk forecast. We present the results based on the forecasting error of all variables and the individual results for GDP, which is often the variable of interest in GVAR modelling exercises. The results for all variables are computed as the unweighted RMSE across all variables.⁷

The first relevant result of our forecasting exercise concerns the predictive ability of GVAR models which employ trade weights to construct global spillovers. Assessing their forecasting performance across all variables, the predictive power of such models is systematically inferior to that of models based on other weighting schemes. In particular, basing the quantitative assessment of global spillovers using weights based on geodesic distance (arguably the simplest of all weighting strategies used) presents the best forecasting performance when concentrating on short-term predictions (one quarter ahead) across all variables, although the difference in forecasting power with respect to the random walk prediction is not statistically significant. An important result of the analysis concerns the lack of forecasting superiority of a single specification across all variables and all prediction horizons. While GVAR models with weights based on geodesic distance outperform all other specifications in short-term forecasting (albeit not significantly so), the model with trade and inward portfolio investments has the best performance when the predictive horizon is two quarters. The trade costs weighting scheme is the best performer for the one year ahead horizon and the model where spillovers are based on weights from inward banking claims outperforms all others when forecasting two years ahead. However, the improvements in predictive ability with respect to the random walk benchmark at the two year ahead horizon do not appear statistically significant.

Weighting schemes based on predictive likelihood are not able to outperform the best single specifications, although averaging across specifications using weights which are based on the overall performance across all variables tends to provide reasonably good forecasts in the short and medium run.

The results for GDP provide evidence concerning the added value of combining trade and financial information to approximate macroeconomic spillovers in GVAR

⁷ For particular applications, a weighting scheme over variables could be desirable. Given the general setting of our exercise, we decided to use an unweighted average in the forecast comparison.

Table 3 Predictive accuracy results

Model	All variables				GDP			
	h = 1	h = 2	h = 4	h = 8	h = 1	h = 2	h = 4	h = 8
Trade	1.03	0.92*	1.04	1.35***	0.97	1.00	1.12*	1.31***
TC	1.16	0.99	0.80**	1.03	1.05	1.15	1.17**	1.03
GD	0.96	0.82***	1.00	2.49***	1.05	1.05	1.22*	1.83***
PIin	0.99	0.77***	0.83	1.05	1.20	1.17	1.06	0.93
PIout	1.04	0.82**	1.01	1.19	1.28**	1.37***	1.39***	1.37**
FDIin	1.05	0.82***	0.92	1.04	1.26**	1.15	1.05	1.06
FDIout	1.07	0.96	1.14	1.39**	1.17	1.26**	1.34***	1.41***
BCin	1.08	0.86**	0.82	0.94	1.31**	1.16	1.01	1.03
BCout	0.97	0.82***	0.90	1.07	1.04	1.07	1.07	1.03
PIin and trade	1.05	0.67***	0.80*	0.99	0.92	0.86	0.89	0.87*
PIout and trade	0.97	0.71***	0.83	1.05	0.99	1.01	1.10	1.15*
FDIin and trade	1.05	0.77***	0.96	1.21**	0.91	0.89	1.05	1.12
FDIout and trade	0.98	0.83***	1.01	1.23**	0.97	0.96	1.09	1.19*
BCin and trade	0.98	0.74***	0.84	1.05	0.98	0.97	0.98	0.99
BCout and trade	1.04	0.89**	1.00	1.23**	1.12	1.16*	1.27**	1.29**
PLMA (all variables)	0.97	0.82***	0.90	1.07*	1.05	1.07	1.07	1.03
PLMA (variable specific)	0.98	0.84***	0.94	1.11	1.16	1.20*	1.24**	1.25**
Unweighted average	0.98	0.88***	1.03	1.38***	1.07	1.08	1.17**	1.32***

Bold figures indicate minima

*** / ** / * : 99 % / 95 % / 90 % significance of Diebold-Mariano test that the forecast differs from the benchmark

models. The best performers in terms of predictive accuracy are GVAR models that combine trade weights with weights based on financial flows. For one quarter ahead predictions the financial variable achieving best forecasts is inward FDI linkages, while for all other horizons it is inward portfolio investments. The forecast improvements, however, are only statistically significant in the long-run prediction horizons. The performance of forecast-averaging techniques is particularly disappointing in the case of GDP predictions, as they cannot outperform the benchmark random walk specification at any of the forecasting horizons assessed in the exercise.

Since the results of the forecasting exercise may be strongly affected by the global financial crisis that started in 2008, we carried out a robustness check using a reduced sample. The in-sample period in this alternative exercise is defined to cover the period 1980Q1–2005Q4, with the last four quarters used to obtain predictive likelihood weights for forecasting averaging. The period 2006Q1–2007Q4, which precedes the financial crisis, is used for the assessment of out-of-sample predictive accuracy. The relative RMSE of the different weighting schemes based on this reduced sample are presented in Table 4 for the full set of variables and for GDP. Although the cumulative results over all variables appear much more stable than with the sample that includes the financial crisis, the heterogeneity by forecasting horizon across variables is large, as exemplified in the results for GDP. The best overall forecasting performance in this

Table 4 Predictive accuracy results: pre-crisis period

Model	All variables				GDP			
	h = 1	h = 2	h = 4	h = 8	h = 1	h = 2	h = 4	h = 8
Trade	1.10	1.32**	1.32*	1.22**	0.89	1.06	1.13	1.19
TC	0.96	1.00	1.26**	2.53***	0.89	0.97	1.04	1.38***
GD	1.16*	1.26**	1.68***	1.13**	0.96	0.93	1.02	1.03
PIin	1.16*	1.42***	1.78***	1.60***	0.87	1.04	0.96	0.97
PIout	0.99	1.19**	1.35***	1.32***	0.98	1.08	1.05	1.07
FDIin	1.24***	1.32**	1.54***	1.42***	0.99	0.97	1.10	1.15**
FDIout	0.84**	0.92	0.92	0.96	0.96	1.05	0.96	0.89
BCin	1.18	1.48*	1.67***	1.54***	0.94	1.11	1.12	1.12
BCout	0.95	0.98	1.19**	1.18***	0.89	0.98	0.97	1.00
PIin and trade	0.89	1.01	1.11	1.10	0.93	1.07	0.98	1.03
PIout and trade	0.95	1.21**	1.31**	1.18*	0.97	1.11	1.12	1.18
FDIin and trade	1.11	1.34**	1.39***	1.23***	0.97	1.21	1.20	1.22*
FDIout and trade	1.06	1.31	1.32*	1.25**	1.02	1.21	1.13	1.17
BCin and trade	0.81**	1.10	1.05	1.05	1.00	1.19	1.15*	1.18*
BCout and trade	1.10	1.35*	1.27*	1.22*	0.95	1.18	1.11	1.14
PLMA (all variables)	0.89	1.01	1.11	1.10	0.93	1.07	0.98	1.03
PLMA (variable specific)	0.88*	0.98	1.08	1.08	0.94	1.02	1.01	1.04
Unweighted average	1.03	1.16**	1.38***	1.30***	0.91	1.02	1.04	1.10

Bold figures indicate minima

*** / ** / *: 99 % / 95 % / 90 % significance of Diebold-Mariano test that the forecast differs from the benchmark

setting is achieved by the GVAR model with weights based exclusively on outward FDI for all forecasting horizons but one quarter ahead. For the one-quarter ahead horizon, it is the hybrid weighting scheme based on trade and inward banking claims that provides the most accurate predictions, although the difference to the model based on outward FDI weights is very small. For the data prior to the crisis, only the short-term predictions (one quarter ahead) appear to improve over the random walk benchmark significantly.

In the case of GDP forecasts, the best predictive accuracy tends to be obtained in GVAR models with financial variables as determinants of the spillover weights, although the particular financial variable differs across forecasting horizons. Inward portfolio investments achieve the most accurate GDP forecasts one quarter and one year ahead, while weights based on outward FDI flows perform best eight quarters ahead. For the two quarters ahead horizon, it is the weighting scheme based on geodesic distances that outperforms the rest. As in the case of the full sample, forecast averaging does not provide significant improvements in out-of-sample predictions for GDP and statistically significant improvements over the random walk are not present.

To sum up, our results indicate that the use of weighting schemes based on trade flows in GVAR models leads to inferior predictive accuracy as compared to other specifications. In particular, simple weighting schemes based on geodesic distance and

schemes that exploit the information of bilateral financial linkages appear particularly promising in terms of providing improvements in predictive accuracy for macroeconomic variables. However, no statistically significant differences in predictive ability with respect to the random walk model are found in short-run for all variables in the specification. We also find that the gains of assessing model uncertainty through predictive likelihood weighting are extremely limited in the context of the prediction exercise carried out. However, it should be highlighted that single models do not tend to systematically improve predictions at all horizons and for all variables and that the forecasting ability of individual specifications can fluctuate strongly across predictive horizons. Forecast averaging presents in general stable predictive ability across predictive horizons, in spite of the fact that it does not rank among the top specifications in terms of horizon-specific predictive ability.

The gains in predictive ability obtained by using spatial spillovers (see for example the results for two quarters ahead for all variables in Table 3) based on geodesic distance can be explained by the exogeneity of such a measure, which makes estimates of the GVAR parameters less affected by potential simultaneity biases. On the other hand, our results provide further evidence of the importance of financial linkages as a determinant of shock transmission across developed economies (Eickmeier and Ng 2015).

5 Conclusions and furthers path of research

Addressing global linkages in macroeconometric models is expected to improve their usefulness for policy analysis and forecasting. In this context, global VAR models are a standard instrument. In this study we present a comprehensive analysis of the role played by different weighting schemes as determinants of the out-of-sample predictive accuracy of GVAR models.

Our results indicate that the performance of standard GVAR models, that employ weights on foreign variables based on bilateral trade, tends to be significantly worse than that of specifications using geodesic distance or financial flows as a basis of the weights. Although there is no single weighting scheme among those entertained in the study that performs best across all variables and forecasting horizons, our results indicate that the information contained in bilateral financial flow variables can be exploited to improve the predictive performance of GVAR specifications. Further research to assess optimal approaches to combining data on trade and financial linkages for GVAR model specifications appears particularly promising in this context.

Our results concerning the possibility of improving predictive performance through forecast averaging are relatively disappointing. Embedding the problem of weight specification in a Bayesian framework (see for a recent contribution on the estimation of Bayesian GVAR models, Crespo Cuaresma et al. 2016) may provide a suitable context to improve the performance of model averaged predictions in future research.

Acknowledgments Open access funding provided by Vienna University of Economics and Business (WU).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution,

and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Baxter, M., Kouparitsas, M.A.: Determinants of business cycle comovement: a robust analysis. *J. Monet. Econom.* **52**(1), 113–157 (2005)
- Calvo, G.A.: Contagion in emerging markets: when wall street is a carrier, pp. 1–22 (1999)
- Crespo Cuaresma, J., Feldkircher, M., Huber, F.: Forecasting with Bayesian Global Vector Autoregressive Models: a comparison of priors. *J. Appl. Econom.* (2016) (forthcoming)
- Crespo Cuaresma, J., Feldkircher, M.: Spatial filtering, model uncertainty and the speed of income convergence in europe. *J. Appl. Econ.* **28**(4), 720–741 (2013)
- Dees, S., Holly, S., Pesaran, H.M., Smith, V.L.: Long run macroeconomic relations in the global economy. *Economics* **1**(3), 1–20 (2007a)
- Dees, S., Mauro, F.D., Pesaran, M.H., Smith, L.V.: Exploring the international linkages of the euro area: a global var analysis. *J. Appl. Econom.* **22**(1), 1–38 (2007b)
- Diebold, F.X., Mariano, R.S.: Comparing predictive accuracy. *J. Bus. Econ. Stat.* **20**(1), 134–144 (2002)
- Eickmeier, S., Ng, T.: How do US credit supply shocks propagate internationally? A GVAR approach. *Eur. Econ. Rev.* **74**, 128–145 (2015)
- Forbes, K.J., Chinn, M.D.: A decomposition of global linkages in financial markets over time. *Rev. Econ. Stat.* **86**(3), 705–722 (2004)
- Garrat, A., Lee, K., Pesaran, M.H., Shin, Y.: Global and national macroeconomic modelling: a long-run structural approach. Oxford University Press, Oxford (2006)
- Kapetanios, G., Labhard, V., Price, S.: Forecasting using predictive likelihood model averaging. *Econ. Lett.* **91**(3), 373–379 (2006)
- LeSage, J.P., Pace, R.K.: Introduction to spatial econometrics. CRC Press Taylor & Francis Group, Boca Raton, London, New York (2009)
- Pesaran, M.H., Schuermann, T., Weiner, S.M.: Modeling Regional Interdependencies using a Global Error-Correcting Macroeconometric Model. *J. Bus. Econ. Stat. Am. Stat. Assoc.* **22**, 129–162 (2004a)
- Pesaran, M.H., Schuermann, T., Weiner, S.M.: Modeling regional interdependencies using a global error-correcting macroeconomic model. *J. Bus. Econ. Stat.* **22**(2), 129–162 (2004b)
- Smith, L., Galesi, A.: Gvar toolbox 2.0. GVAR Toolbox 2.0. <https://sites.google.com/site/gvarmodelling/gvar-toolbox> (2014)
- Van Rijkeghem, C., Weder, B.: Sources of contagion: is it finance or trade? *J. Int. Econ.* **54**(2), 293–308 (2001)