

Huber, Florian; Crespo-Cuaresma, Jesus; Feldkircher, Martin

Conference Paper

Forecasting with Bayesian global vector autoregressive models: A comparison of priors

54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Huber, Florian; Crespo-Cuaresma, Jesus; Feldkircher, Martin (2014) : Forecasting with Bayesian global vector autoregressive models: A comparison of priors, 54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124192>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Forecasting with Bayesian Global Vector Autoregressive Models: A Comparison of Priors*

Jesús Crespo Cuaresma^{a, b, c, d}, Martin Feldkircher^e, and Florian Huber^a

^aVienna University of Economics and Business (WU)

^bWittgenstein Centre for Demography and Human Capital (WIC)

^cInternational Institute for Applied Systems Analysis (IIASA)

^dAustrian Institute of Economic Research (WIFO)

^eOesterreichische Nationalbank (OeNB)

This version: November 2013

Abstract

This paper puts forward a Bayesian version of the global vector autoregressive model (B-GVAR) that accommodates international linkages across countries in a system of vector autoregressions. We compare the predictive performance of B-GVAR models for the one- and four-quarter ahead forecast horizon for standard macroeconomic variables (real GDP, inflation, the real exchange rate and interest rates). Our results show that taking international linkages into account improves forecasts of inflation, real GDP and the real exchange rate, while for interest rates forecasts of univariate benchmark models remain difficult to beat. Our Bayesian version of the GVAR model outperforms forecasts of the standard cointegrated VAR for practically all variables and at both forecast horizons. The comparison of prior elicitation strategies indicates that the use of the stochastic search variable selection (SSVS) prior tends to improve out-of-sample predictions systematically. This finding is confirmed by density forecast measures, for which the predictive ability of the SSVS prior is the best among all priors entertained for all variables at all forecasting horizons.

Keywords: Global vector autoregressions, forecasting, prior sensitivity analysis.

JEL Codes: C32, F44, E32, O54.

*The views expressed in this paper are not necessarily those of the Oesterreichische Nationalbank. We would like to thank Sylvia Frühwirth-Schnatter and the participants in an internal research seminar at the Vienna University of Economics and Business for helpful comments. Email: jcrespo@wu.ac.at, martin.feldkircher@oebn.at and fhuber@wu.ac.at.

1 Introduction

The rise in international trade and cross-border financial flows over the last decades implies that countries are more than ever exposed to economic shocks from abroad, a fact that was recently demonstrated by the global financial crisis. Macroeconomic tools that treat countries as isolated from the rest of the world may hence miss important information for forecasting and counterfactual analysis which can lead to important biases. Global vector autoregressive (GVAR) models constitute a useful econometric framework that accommodates spillovers from the global economy in a tractable manner. GVAR specifications consist of single country models that are stacked to yield a comprehensive representation of the world economy. The structure of GVAR models offers a framework to analyze the spatial propagation of foreign shocks, while the fact that the country models are estimated separately ensures that country specificities are also incorporated in the analysis.

The empirical literature on GVAR models has been largely influenced by the work of M. Hashem Pesaran and co-authors (Pesaran et al., 2004; Garrat et al., 2006). In a series of papers, these authors examine the effect of U.S. macroeconomic impulses on selected foreign economies employing agnostic, structural and long-run macroeconomic relations to identify the shocks. (Pesaran et al., 2004; Dees et al., 2007a,b). Recent papers have advanced the literature on GVAR modelling in terms of country coverage (Feldkircher, 2013), identification of shocks (Eickmeier and Ng, 2011) and the specification of international linkages (Eickmeier and Ng, 2011; Chudik and Fratzscher, 2011; Galesi and Sgherri, 2009).

Most of the existing applications of GVAR models concentrate on the quantitative assessment of the propagation of macroeconomic shocks using historical data, while the forecasting performance of such specifications has been the object of a very limited number of contributions hitherto. Pesaran et al. (2009) evaluate GVAR forecasts in an out-of-sample exercise and propose to pool GVAR forecasts over different estimation windows and model specifications in order to account for potential structural breaks and misspecifications. Pesaran et al. (2009) conclude that taking global links across economies into account using GVAR models leads to more accurate out-of-sample predictions than those from univariate specifications for output and inflation. The results, however, are less spectacular for interest rates, the exchange rate and financial variables and the authors find strong cross-country heterogeneity in the performance of GVAR forecasts. Han and Ng (2011) employ a GVAR model to forecast macroeconomic variables in five Asian economies. They find that (one-step ahead) forecasts from GVAR models outperform those of stand-alone VAR specifications for short-term interest rates and real equity prices. Greenwood-Nimmo et al. (2012) concentrate on predicted directional changes to evaluate the forecasting performance of GVAR specifications and confirm their superiority as compared to univariate benchmark models in long-run forecast horizons.

A second strand of the literature advocates the estimation of large VARs using Bayesian techniques. More specifically, Bańbura et al. (2010) assess the forecasting performance of a typical monetary VAR based on more than 100 macroeconomic variables and sectoral information. They show that forecasts of these large-scale models can outperform small benchmark VARs when the degree of shrinkage is set accordingly. More related to our work, Giannone and Reichlin (2009); Alessi and Bańbura (2009) propose to exploit these shrinkage properties and estimate Bayesian VARs with a large cross-section. Alessi and Bańbura (2009) find that the Bayesian VAR in an international framework as well as dynamic factor models are able to yield accurate one to four step ahead forecasts for quarterly data.

In this contribution we combine the virtues of the two strands mentioned above and propose a Bayesian GVAR (B-GVAR) model. Akin to the GVAR framework we assume that links among economies are determined exogenously, while we borrow strength from the Bayesian literature in estimating the individual country models. This allows us to keep the virtues of the GVAR framework offering a coherent way for policy and counterfactual analysis. Our model includes standard variables that are typically employed in small-country VARs such as output, inflation, short- and long-term interest rates, the real exchange rate and the oil price as a global control variable (see e.g., Dees et al., 2007a,b; Pesaran et al., 2004, 2009, 2007, among others). We compare the predictive ability of B-GVAR models with prior specifications resembling those which resurface frequently in Bayesian VAR empirical studies: the normal-conjugate prior, a non-informative prior on the coefficients and the variance, the inverse Wishart prior that assumes prior independence, the Minnesota prior (Litterman, 1986), the single-unit prior, which accommodates potential cointegration relationships (Doan et al., 1984; Sims and Zha, 1998) and the stochastic search variable selection (SSVS) prior George et al. (2008). By inheriting the properties of their single-country VAR counterparts, B-GVAR models are expected to be less prone to overfitting (Giannone and Reichlin, 2009) and allow the researcher to include prior beliefs in the model, while still taking the long-run comovement of variables into account. We compare our battery of priors using rolling window forecasts for one quarter and four quarters ahead. We benchmark the predictive ability of B-GVAR models to that of univariate autoregressive models (standard first-order autoregressions and random walk specifications), as well as to the standard cointegrated GVAR approach.

Our results indicate that taking cross-country interlinkages into account improves one-quarter and four-quarter ahead forecasts for most variables in the specification. The predictive ability gains of B-GVAR models compared to the univariate benchmark model are particularly pronounced for output and exchange rate forecast, while improvements for inflation and (short-term) interest rates are moderate. In addition, Bayesian variants of GVAR models tend to outperform the cointegrated GVAR at both forecast horizons. Comparing across priors, our results support the use of the Minnesota prior on the one hand, and the prior that implements stochastic search variable selection, on the other hand. While both priors excel in one quarter-ahead forecasts of output, the exchange rate and long-term interest rates, the SSVS prior dominates all other priors also at the one-year ahead horizon.

The paper is structured as follows. Section 2 provides a brief description of the global VAR model, while Section 3 introduces the prior specifications we evaluate. In Section 4 we illustrate the data and perform a prior comparison based on forecasts for output, inflation, the real exchange rate and short- and long-term interest rates. Finally, Section 5 concludes.

2 The GVAR Model

GVAR specifications constitute a compact representation of the world economy designed to model multilateral dependencies among economies across the globe. In principle, a GVAR model comprises *two layers* via which enable the specification to capture cross-country spillovers. In the first layer, separate multivariate time series models – one per country – are estimated. In the second layer, the country models are stacked to yield a global model that is able to trace the spatial propagation of a shock as well as its temporal dynamics.

The first layer is composed by country-specific local VAR models, enlarged by a set of weakly exogenous and global variables (VARX* model). Assuming that our global economy consists

of $N + 1$ countries, we estimate a VARX* of the following form for every country $i = 0, \dots, N$:

$$x_{it} = a_{i0} + a_{i1}t + \Phi_i x_{i,t-1} + \Lambda_{i0} x_{it}^* + \Lambda_{i1} x_{i,t-1}^* + \pi_{i0} d_t + \pi_{i1} d_{t-1} + \varepsilon_{it} \quad (1)$$

Here, x_{it} is a $k_i \times 1$ vector of endogenous variables in country i at time $t \in 1, \dots, T$, Φ_i denotes the $k_i \times k_i$ matrix of parameters associated with the lagged endogenous variables and Λ_{ik} are the coefficient matrices of the k_i^* weakly exogenous variables, of dimension $k_i \times k_i^*$. Furthermore, $\varepsilon_t \sim N(0, \Sigma_i)$ is the standard vector error term, d_t denotes the vector of *strictly exogenous* variables, which are linked to the vector of exogenous variables through the matrices π_{i0} and π_{i1} and t is a deterministic trend component. In order to keep the exposition of the model as simple as possible, we set the lag length of the country-specific VAR models to one lag. If $\Lambda_{i0}, \Lambda_{i1}, \pi_0$ and π_1 are composed exclusively by zero elements, the specification boils down to that of a standard VAR model (with a deterministic linear trend if $a_{i1} \neq 0$).

The weakly exogenous or *foreign* variables, x_{it}^* , are constructed as a weighted average of their cross-country counterparts,

$$x_{it}^* := \sum_{j \neq i} \omega_{ij} x_{jt}^*, \quad (2)$$

where ω_{ij} denotes the weight corresponding to the pair of country i and country j . The weights ω_{ij} reflect economic and financial ties among economies, which are usually proxied using data on (standardized) bilateral trade (see e.g., Eickmeier and Ng, 2011, for an application using a broad set of different weights). The assumption that the x_{it}^* variables are weakly exogenous at the individual level reflects the belief that most countries are small relative to the world economy.

Following Pesaran et al. (2004), the country-specific models can be rewritten as

$$A_i z_{it} = a_{i0} + a_{i1}t + B_i z_{i,t-1} + \pi_0 d_t + \pi_1 d_{t-1} + \varepsilon_{it}, \quad (3)$$

where $A_i := (I_{k_i} - \Lambda_{i0})$, $B_i := (\Phi_i - \Lambda_{i1})$ and $z_{it} = (x_{it} \ x_{it}^*)'$. By defining a suitable link matrix W_i of dimension $(k_i + k_i^*) \times k$, where $k = \sum_{i=0}^N k_i$, we can rewrite z_{it} as $z_{it} = W_i x_t$, with x_t (the so-called global vector) being a vector where all the endogenous variables of the countries in our sample are stacked. Substituting (3) in (1) and stacking the different local models leads to the global equation, which is given by

$$\begin{aligned} x_t &= G^{-1} a_0 + G^{-1} a_1 t + G^{-1} H x_{t-1} + G^{-1} \pi_0 d_t + G^{-1} \pi_1 d_{t-1} + G^{-1} \varepsilon_t \\ &= b_0 + b_1 t + F x_{t-1} + \Gamma_0 d_t + \Gamma_1 d_{t-1} + e_t, \end{aligned} \quad (4)$$

where $G = (A_0 W_0 \cdots A_N W_N)'$, $H = (B_0 W_0 \cdots B_N W_N)'$ and a_0, a_1, π_0 and π_1 contain the corresponding stacked vectors containing the parameter vectors of the country-specific specifications. The eigenvalues of the matrix F , which is of prime interest for forecasting and impulse response analysis, are assumed to lie within the unit circle in order to ensure stability of (4). The framework outlined above deviates from the work pioneered by Pesaran et al. (2004) in that we do not impose cointegration relationships in the individual country-specific models. Consequently, we impose that all the eigenvalues of Φ_i in (3) lie within the unit circle for all $i = 0, \dots, N$, a somewhat stricter condition than the one imposed by Pesaran et al. (2004).

Solving equation (4) forward and taking conditional expectations, we arrive at the expression needed for calculating n -steps ahead predictions,

$$\hat{x}_{T+n} = \mathbf{E} \left(x_{T+n} | x_T, \bigcup_{\tau=1}^n d_{t+\tau} \right) = F^n x_T + \sum_{\tau=0}^{n-1} F^\tau [b_0 + b_1(T + n - \tau)] + \sum_{\tau=0}^{n-1} F^\tau [\Gamma_0 d_{T+n-\tau} + \Gamma_1 d_{T+n-\tau-1}]. \quad (5)$$

The companion matrix F^n converges to a zero matrix rapidly as $n \rightarrow \infty$. Consequently, information about the initial state of the system becomes less important as n increases.

3 The B-GVAR: Priors over Parameters

Bayesian analysis of the GVAR model requires the elicitation of prior distributions for all parameters of the model. We use several prior structures that have been developed for VAR specifications for the individual country-specific models together with standard prior settings for the parameters corresponding to (weakly) exogenous variables and combine the posterior results in the B-GVAR model.¹ For prior implementation, it proves to be convenient to work with the parameter vector $\Psi_i = (a'_{i0} \ a'_{i1} \ \text{vec}(\Phi_i)' \ \text{vec}(\Lambda_{i0})' \ \text{vec}(\Lambda_{i1})' \ \text{vec}(\pi_{i0})' \ \text{vec}(\pi_{i1})')'$.

The Natural Conjugate Prior

We start with the simplest prior for the coefficients of the country-specific VARX* models, which is the natural conjugate prior. In the VARX* framework, the inverse of the variance-covariance matrix of the error term in the country-specific model is assumed to follow a Wishart distribution, so that

$$\Sigma_i^{-1} \sim W(\underline{S}^{-1}, \underline{v}) \quad (6)$$

where \underline{S}^{-1} is the prior scale matrix and \underline{v} are the prior degrees of freedom. The prior on the coefficients given Σ_i is multivariate Gaussian,

$$\Psi_i | \Sigma_i \sim N(\underline{\Psi}, \Sigma_i \otimes \underline{V}). \quad (7)$$

where \underline{V} is the prior variance on the coefficients. The posterior distribution of $\Psi_i | \Sigma$ is normal and the posterior distribution of the precision matrix is of Wishart form. The marginal posterior of Ψ_i , which is obtained after integrating out Σ of the conditional posterior in (7), is distributed as a multivariate t-distribution.² The diffuse prior is obtained by letting the prior variance on the coefficients approach infinity, which implies that the prior precision on the coefficients, \underline{V}^{-1} , approaches the zero matrix. As a result, our posterior effectively becomes the likelihood, implying that we arrive at the standard OLS results.

Such a prior specification is used in a BVAR setting by, e.g., Kadiyala and Karlsson (1993), where a forecasting comparison between different prior structures is performed. They conclude that the performance of BVAR models using the natural conjugate prior for Canadian GDP is good, despite its simplicity.

¹Karlsson (2012) provides an excellent overview about Bayesian VAR models.

²For a complete derivation of this result, see Zellner (1973).

The Minnesota Prior

Following Koop and Korobilis (2010), the Minnesota prior for Ψ_i is given by

$$\Psi_i \sim N(\underline{\Psi}_{Mn}, \underline{V}_{Mn}). \quad (8)$$

We summarize our beliefs about the dynamics of the variables in the system in $\underline{\Psi}_{Mn}$. The parameters of each equation in the country-specific VARX* models are centered a priori on the values corresponding to naive random walk specifications for the endogenous variables. The prior mean corresponding to the lag of the endogenous variables is thus a priori set to one and the rest of the parameters to zero. Favouring unit root behaviour for macroeconomic variables complies with standard practices of applied econometricians, given the typically high persistence of such time series (Sims and Zha, 1998).

The prior variance-covariance matrix associated with the coefficients, \underline{V}_{Mn} , is assumed to be diagonal. If \underline{V}_i denotes the block of \underline{V}_{Mn} corresponding to equation i , then its diagonal elements are specified as

$$\underline{V}_{i,jj} = \begin{cases} \frac{\alpha_1}{r^2} & \text{for the parameter of the } r\text{-th lag of the endogenous variable} \\ \frac{\alpha_2 \sigma_{ii}}{r^2 \sigma_{jj}} & \text{for coefficients on lag } r \text{ of variable } j \neq i \text{ for } r = 1, \dots, p \\ \alpha_3 \sigma_{ii} & \text{for coefficients on exogenous, weakly exogenous and deterministic variables} \end{cases}$$

where σ_{ii} refers to the standard error of a univariate autoregression for the corresponding variable and r^2 is a deterministic function of the lag length. Consequently, the strength of the prior belief in the random walk is governed by α_1 . The hyperparameter r^2 increasingly tightens the variance on the prior for distant lags, reflecting the belief that longer lag lengths have a detrimental effect on the forecasting performance of the variables under study. The term $\frac{\sigma_{ii}}{\sigma_{jj}}$ serves as a scaling factor to adjust for differences in the units in which the variables are measured. Note that the Minnesota prior in the form described above allows for an asymmetric treatment of the parameters in the different equations by varying the hyperparameters α_1, α_2 and α_3 . In particular, it is a common practice to elicit these hyperparameters in such a way that the variance of the coefficients for other explanatory variables is looser compared to those for lags of the dependent variable.

Owing to its flexibility, the Minnesota prior is widely applied in empirical studies. In the framework outlined above, prior elicitation boils down to the specification of the hyperparameters $\alpha_1, \alpha_2, \alpha_3$ and replacing Σ_i , the variance-covariance matrix of the error term, with an estimate $\hat{\Sigma}_i$. By treating Σ_i as known, the posterior of the Minnesota prior is given analytically, although this implies that uncertainty surrounding the estimate of Σ_i is ignored. Note also that $\hat{\Sigma}_i$ is typically non-diagonal, which implies that we allow for correlation across the equations of each country-specific model. Elicitation of the hyperparameters can be done by simple trial and error, using recommendations found in the literature or using hyperpriors (see Giannone et al., 2012).

Recent contributions by Bańbura et al. (2010) and Carriero et al. (2009) use large VARs with variants of the Minnesota prior to forecast macroeconomic quantities. Both of these studies highlight the strong forecasting performance of Bayesian VAR models which rely on this prior specification. Álvarez and Ballabriga (1994) show that the Minnesota prior is also capable of capturing long-run patterns and that adding long-run restrictions to the prior does not

necessarily improve forecasts.

The Single Unit Root (SUR) Prior

The Minnesota prior specification can be implemented using dummy observations. This approach, used in the early Bayesian VAR literature (Litterman, 1986), mimics the behaviour of the natural conjugate prior. We start by rendering equation (1) in the standard regression form,

$$Y_i = X_i\Psi_i + \varepsilon_i \quad (9)$$

where Y_i contains the stacked endogenous variables for country i and X_i is the corresponding design matrix. Following Sims and Zha (1998), the single-unit-root prior is implemented by concatenating artificial observations (*dummy observations*) at the beginning of the available sample. We denote these dummy observations by $y^* = \bar{Y} \underline{\psi}$ for the endogenous variables and $X^* = \bar{X} \underline{\vartheta}$ for the explanatory variables, where \bar{Y} and \bar{X} are the corresponding sample means over some predetermined time period at the beginning of the sample, while $\underline{\psi}$ and $\underline{\vartheta}$ are hyperparameters controlling the tightness of the prior. Stacking the artificial and real data yields

$$r_i = Z_i\rho + \epsilon_i^* \quad (10)$$

where $r_i = (Y_i \ y^*)'$ and $Z_i = (X_i \ X^*)'$. Estimation of equation (10) is carried out using Theil-Goldberger mixed estimation³ (Theil and Goldberger, 1961) and the likelihood of the dummy data can be interpreted as its corresponding prior distribution. This prior pushes the variables in the country-specific VAR towards their unconditional (stationary) mean, or toward a situation where there is at least one single unit root present. That is, either the process has a unit root or it is stationary and starts near its mean - implying a penalty for models with inherent initial 'transient' dynamics (Sims, 1992).

This prior structure has the advantage of being simple, allowing for cointegrating relationships in the data and allowing for analytical posterior solutions. In contrast to the traditional implementation of the Minnesota prior, however, it is not possible to treat different variables in an asymmetric fashion, implying that the shrinkage capabilities of the single unit root prior are limited as compared to those of the standard Minnesota prior implementation.⁴

The independent Normal-Wishart Prior

As noted by Koop and Korobilis (2010), one drawback of the natural conjugate prior is that it relies on the Kronecker structure of the prior variance covariance matrix, which hinders the use of structural VAR or restricted VAR specifications. By assuming prior independence between the precision and the coefficients we can gain flexibility in the structure of the prior, albeit losing analytical tractability and thus having to rely on Gibbs sampling to carry out posterior inference.

³This boils down to applying OLS to the augmented regression equation in (10), which implies $\hat{\rho} = (Z_i'Z_i)^{-1}Z_i'r_i$.

⁴The so-called sum-of-coefficients or no-cointegration prior (Doan et al., 1984) can also be implemented using dummy observations. This prior centers the sum of coefficients on the lagged endogenous variables around unity and the rest of the parameters in the model around zero. Since in our application the maximum lag length is one, the sum-of-coefficients prior would mimic the setting of the Minnesota prior. Consequently, we excluded the prior from our forecasting exercise.

The general structure of this type of prior is given by

$$p(\Psi_i, \Sigma_i^{-1}) = p(\Psi_i)p(\Sigma_i^{-1})$$

where $\Psi_i \sim N(\underline{\Psi}, \underline{V}_{NW})$ and $\Sigma_i^{-1} \sim W(\underline{S}^{-1}, \underline{v})$. Assuming that the moments in the prior for Ψ_i are given by those of the Minnesota prior, this specification leads to the so-called Minnesota-Wishart prior. While the joint posterior density is not of well-known form, the conditional posteriors are Gaussian for Ψ_i and inverse Wishart for Σ_i , which allows for Gibbs sampling in a straightforward manner.⁵

Stochastic search variable selection (SSVS) Prior

The SSVS prior uses a mixture of normals on each coefficient of the VARX*:

$$\Psi_{i,j}|\delta_{i,j} \sim (1 - \delta_{i,j})N(0, \iota_{0j}^2) + \delta_{i,j}N(0, \iota_{1j}^2) \quad (11)$$

Here, $\delta_{i,j}$ denotes a binary random variable corresponding to coefficient j in country model i . It equals one if the coefficient is included in the corresponding model and zero if it is a priori excluded from the respective country model. The normal distribution corresponding to $\delta_{i,j} = 0$ is typically specified with ι_{0j}^2 close to zero, thus aiming at shrinking the coefficient towards zero. On the other hand, the prior variance of the normal distribution for $\delta_{i,j} = 1$, ι_{1j}^2 , is set to a comparatively large value. A relatively uninformative prior is thus used on coefficient j conditional on inclusion. The intuition behind the SSVS prior is the following: Given the value of $\delta_{i,j}$, a variable is excluded or included in the model. This is done by imposing a dogmatic prior on $\Psi_{i,j}$ if $\delta_{i,j} = 0$ and a diffuse prior if $\delta_{i,j} = 1$. The first case implies that the corresponding coefficient could be safely regarded as zero, implying that the variable is not included in the model.

To simplify the prior implementation it is possible to rewrite equation (11) using a multivariate normal distribution

$$\Psi_i|\delta \sim N(D\underline{\Psi}_i|HH) \quad (12)$$

where $D = \text{diag}(\delta_{i,1}, \delta_{i,2}, \dots, \delta_{i,n})$ and $I_\iota = \text{diag}(\iota_{i,1}, \iota_{i,2}, \dots, \iota_{i,n})$. $\iota_{i,j}$ is the prior variance on the coefficients depending on $\delta_{i,j}$. The specification of the hyperparameters is done in a *default semi-automatic* approach, proposed by George et al. (2008), where $\iota_{i,0j}^2$ and $\iota_{i,1j}^2$ are scaled using the respective standard deviations of the VARX* estimated by OLS.

Estimation of this model requires Markov Chain Monte Carlo (MCMC) methods, although the conditional posteriors of $\delta_{i,j}$, Ψ_i and Σ_i is known. This implies that we can employ a simplified version of the Gibbs sampler outlined in George et al. (2008), where we start drawing Ψ_i from its full conditional posterior, which follows a Normal distribution. In the next step, we draw the latent variable $\delta_{i,j}$ from a Bernoulli distribution and in the last step we draw Σ_i from an inverse Wishart distribution.⁶ This algorithm is repeated n times and the first n_{burn}

⁵See the Appendix for a description of the design of the Markov Chain Monte Carlo algorithm to draw from the posterior distribution in this case.

⁶In contrast to the implementation in George et al. (2008), we impose a inverse Wishart prior on Σ_i and depart from using a restriction search over Σ_i .

draws are discarded as burn-ins.⁷ Averaging the draws of the $\delta_{i,j}$ leads to posterior inclusion probabilities for each variable j . Note that, in contrast to the Minnesota prior, where the prior variance on the coefficients differs between different types of variables, the SSVS prior effectively applies an individual degree of shrinkage to every coefficient on the country level. That is, cross-country heterogeneity is taken into account within the GVAR by allowing for potentially different country VARX* models.

4 Prior Settings and the Predictive Ability of B-GVAR Models: Results

4.1 Data and Model Specification

We compare the different prior structures described in Section 3 in terms of out-of-sample predictions using B-GVAR models estimated with a sample which covers a large part of the global economic activity in the world. Our data set contains quarterly observations for 45 countries and 1 regional aggregate, the euro area (EA)⁸. Table 1 presents the country coverage of our sample, which includes emerging economies, advanced economies and the most important oil producers and consumers across the globe.

Advanced Economies (11):	US, EA, UK, CA, AU, NZ, CH, NO, SE, DK, IS
Emerging Europe (14):	CZ, HU, PL, SK, SI, BG, RO, HR, AL, RS, TR, LT, LV, EE
CIS and Mongolia (6):	RU, UA, BY, KG, MN, GE
Asia (9):	CN, KR, JP, PH, SG, TH, ID, IN, MY
Latin America (5):	AR, BR, CL, MX, PE
Middle East and Africa (1):	EG

Abbreviations refer to the two-digit ISO country code.

Table 1: Country coverage

The 46 economies in our sample represent more than 90% of the global economy in terms of GDP in 2010, which improves significantly upon the country coverage in Dees et al. (2007a), whose sample amounted to 78% of aggregate production in the same year.⁹ As compared to other contributions in the literature, our time span covers the period of the global financial crisis and its aftermath. We have 72 quarterly observations by country spanning the period 1995Q1 to 2012Q4. The *domestic* variables that are used in our analysis comprise data on real activity, change in prices, the real exchange rate, and short- and long-term interest rates (Dees et al., 2007a,b; Pesaran et al., 2004, 2009, 2007). We follow the bulk of the literature in including oil prices as a *global* control variable.

⁷A more detailed description of this Gibbs sampler can be found in the Appendix.

⁸The country composition on which the data on the euro area is based changes with time. While historical time series are based on data of the ten original euro area countries, the most recent data are based on 17 countries. The results of our analysis remain qualitatively unchanged if we use a consistent set of 14 euro area member states throughout the sample period instead of the rolling country composition for the data on the euro area, as the relative economic size of these three countries is quite small.

⁹These figures are based on nominal GDP and are taken from the IMF's World Economic Outlook database, April 2012.

The variables used in the model are briefly described in Table 2. Most of the data are available with wide country coverage, with the exception of government yields. Since local capital markets in emerging economies (in particular in Eastern Europe) are still developing, data on interest rates are hardly available for these countries. For those countries for which data at the beginning of the sample period were missing we use an expectation maximization algorithm to impute the values.¹⁰

Table 2: Data description

Variable	Description	Min.	Mean	Max.	Coverage
y	Real GDP, average of 2005=100. Seasonally adjusted, in logarithms.	3.465	4.516	5.194	100%
Δp	Consumer price inflation. CPI seasonally adjusted, in logarithms.	-0.258	0.020	1.194	100%
e	Nominal exchange rate vis-à-vis the US dollar, deflated by national price levels (CPI).	5.699	-2.220	5.459	97.8%
i_S	Typically 3-months-market rates, rates per annum.	0.000	0.100	4.332	93.5%
i_L	Typically government bond yields, rates per annum.	0.006	0.060	0.777	39.1%
$poil$	Price of oil, seasonally adjusted, in logarithms.	-	-	-	-
Trade flows	Bilateral data on exports and imports of goods and services, annual data.	-	-	-	-

Summary statistics pooled over countries and time.

The coverage refers to the cross-country availability per country, in %.

We follow Pesaran et al. (2004, 2009); Dees et al. (2007a) and choose weights based on bilateral trade to calculate the weakly exogenous variables given in equation (4). Since world trade collapsed with the onset of the global recession, trade weights are computed using an average of bilateral exports and imports over the period from 2006 to 2008. With the exception of the real exchange rate, we take the trade weighted counterparts as weakly exogenous variables into the country models for all variables. In line with the literature, the oil price is determined within the U.S. country model. Since our data span is rather short, untreated outliers can have a serious impact on the overall stability and the results of the model. We therefore introduce a set of dummy variables in the country-specific specifications to control for outliers. These account for the fact that some countries witnessed extraordinarily high interest rates at the beginning of the sample period (which returned steadily to 'normal' levels) and that some economies (Russia or Argentina, for instance) were exposed to one-off crisis events. We identify the largest deviations from 'normal' times per country and use interaction terms to take care of unusually large historical observations. The exact specifications of the country models are provided in the Appendix.

¹⁰A more detailed account of the imputation method and the data sources is provided in Feldkircher (2013).

Our benchmark models comprise the random walk, the AR(1) model, the traditional cointegrated GVAR (*Pesaran*) and the B-GVAR with a diffuse prior (*Diffuse*). The benchmark models thus span prominent univariate models, the standard model employed in the literature and the B-GVAR without fusing relevant prior knowledge into the analysis (i.e., estimated using standard maximum likelihood methods).

4.2 Priors and Predictive Ability: Results

The exact values of the hyperparameters for the different GVAR prior specifications are the following:

- The prior mean for the natural conjugate prior $\underline{\Psi}$ is set to zero for all coefficients. The prior variance on the coefficients equals $\underline{V} = \text{diag}(5^2)$, implying a loose prior on the coefficients. For the inverse variance-covariance matrix, the prior scale matrix equals $\underline{S}^{-1} = \text{diag}(1)$ and the prior degrees of freedom $\underline{\nu}$ are set equal to the number of endogenous variables in country i .
- For the Minnesota prior, the hyperparameters $\underline{\alpha}_1, \underline{\alpha}_2$ and $\underline{\alpha}_3$ are set equal to 0.5, 0.5 and 10^2 respectively. This implies that we have a relatively loose prior on the dynamic coefficients of the VAR and a uninformative prior on the exogenous, weakly exogenous and deterministic variables.¹¹ The maximum-likelihood estimate $\hat{\Sigma}$ is used for the variance-covariance matrix.
- In case of the single unit root prior, we set $\underline{\psi} = 5$ and $\underline{\vartheta} = 2$. Those values imply that there is a strong tendency to return to the initial state of the process, which is consistent with cointegration. By setting $\underline{\psi} \neq 0$ we also allow for co-persistence between the variables. This implies that when all variables are fixed at the initial level, they tend to stay there.
- The hyperparameters for the traditional implementation of the independent Normal-Wishart prior are the same as the ones used for the natural conjugate prior. For the Minnesota-Wishart version we rely on the same set of parameters used in the standard Minnesota implementation, where we also rely on the same prior for the variance-covariance matrix as in the natural conjugate/independent Normal-Wishart case.
- For the SSVS prior, we set the prior inclusion probability for each variable equal to 0.5, which implies that a priori, every variable is deemed equally likely to enter the model. We set $\iota_{i,0j} = 0.1\sigma_j$ and $\iota_{i,1j} = 10\sigma_j$ and rely on the semi-automatic approach described in George et al. (2008) to scale the hyperparameters, where σ_j is the standard error attached to coefficient j based on a VARX* estimated by OLS.

The initial estimation period ranges from 1995Q1 to 2008Q4 and we use the period 2009Q1-2012Q4 as out-of-sample hold-out observations for the comparison of predictive ability across specifications. We base our comparison on recursive one-quarter-ahead and one-year-ahead (four-quarters-ahead) predictions obtained by reestimating the models over the rolling window defined by the beginning of the available sample and the corresponding period in the hold-out sample. In addition to comparing point forecasts, we also obtain predictive densities (see the

¹¹Treating the weakly exogenous variables in a different way does not have a big impact on the predictive abilities of the final model.

Appendix for the computational details of the estimation of predictive densities for B-GVAR models) and the continuous rank probability score (CRPS) is used to evaluate the goodness of the density forecasts. The CRPS is given by

$$\text{CRPS}(F, \tilde{y}) := \int_{-\infty}^{\infty} [F(\gamma) - H(\gamma - \tilde{y})]^2 d\gamma, \quad (13)$$

where $H(\cdot)$ denotes the Heavyside function, which equals one if the argument is positive and zero otherwise, $F(\cdot)$ denotes the predictive cumulative distribution function and \tilde{y} are the actual observations. The CRPS is the integral of the Brier score (Hersbach, 2000) at all possible thresholds γ (Gneiting et al., 2005). Assuming that $F(\gamma)$ is the CDF of a normal distribution allows closed form solutions of the integral in (13). Loosely speaking, the CRPS measures the difference between the predicted and the sample cumulative distributions (Carney and Cunningham, 2006) and thus penalises predictions that are far away from the actual observations.¹²

Table 3: Relative Forecasting Performance, One-Quarter-Ahead: Root Mean Square Error and Continuous Rank Probability Score

	NC	M	SUR	IW	M-IW	SSVS	Diffuse	Pesaran	AR
y	1.3580 (0.0390)	0.6075 (0.0105)	0.8529 (0.0182)	1.1874 (0.0413)	0.6157 (0.0345)	0.5996 (0.0075)	0.7023 (0.0148)	0.9059 -	0.7856 -
Δp	1.0236 (0.0358)	0.8825 (0.0116)	0.7156 (0.0187)	1.0100 (0.0395)	0.9172 (0.0349)	0.8126 (0.0079)	1.1100 (0.0128)	1.2493 -	1.0139 -
e	0.6072 (0.0502)	0.4870 (0.0474)	0.5784 (0.0772)	0.5869 (0.0520)	0.7126 (0.0542)	0.4802 (0.0301)	0.8803 (0.0544)	0.8519 -	0.7980 -
i_S	1.1084 (0.0385)	0.7299 (0.0115)	0.8296 (0.0424)	1.0303 (0.0417)	0.7659 (0.0395)	0.6851 (0.0099)	1.1312 (0.0177)	0.8516 -	0.5744 -
i_L	1.1558 (0.0355)	0.4635 (0.0102)	0.4984 (0.0112)	1.1127 (0.0391)	0.7903 (0.0336)	0.5696 (0.0060)	0.7798 (0.0105)	0.8450 -	0.6162 -

Notes: The figures refer to the ratio of the RMSE corresponding to the model to the RMSE of random walk predictions. Average CRPS in parentheses. Results based on rolling forecasts over the time period 2009Q1-2012Q4. NC stands for the normal conjugate prior, M stands for the Minnesota prior, SUR stands for the single unit root prior, IW stands for the independent normal Wishart prior, M-IW stands for the Minnesota-Wishart version of the IW prior, SSVS stands for the SSVS prior, Diffuse stands for the model estimated using maximum likelihood, Pesaran stands for the cointegrated GVAR specification and AR stands for the first-order autoregressive model. Bold figures refer to the lowest value across models for a given endogenous variable.

Table 3 and Table 4 show the relative forecasting performance of models based on the different prior specifications with respect to random walk predictions (with drift for real GDP and without drift for the rest of the variables). In addition to the standard average root mean squared error based on point forecasts, we calculate the average CRPS for each endogenous variable and report simple (unweighted) cross-country averages. The largest relative improvements in terms of predictive ability with respect to the naive forecasts embodied in the random walk benchmark take place for real GDP. B-GVAR models tend to systematically improve predictions for most of the variables in the model with respect to both the naive autoregressive and the random walk forecasts. The interest rate variables are the only exception, with the univariate autoregressive model performing particularly well both in one-quarter-ahead and one-year-ahead predictions.

The relatively weak performance of the GVAR model based on maximum likelihood (the column ‘Diffuse’ in Tables 3 and 4) suggests that the specification of priors plays a central role

¹²Note that for the evaluation of point forecasts, the CRPS reduces to the well-known Mean Absolute Deviation (MAD) criterion (Hersbach, 2000).

Table 4: Relative Forecasting Performance, Four-Quarters-Ahead: Root Mean Square Error and Continuous Rank Probability Score

	NC	M	SUR	IW	M-IW	SSVS	Diffuse	Pesaran	AR
y	1.0090 (0.0399)	0.4857 (0.0187)	0.9999 (0.0229)	0.8847 (0.0423)	0.5799 (0.0360)	0.3651 (0.0148)	0.7192 (0.0224)	0.9332 -	0.5198 -
Δp	0.9204 (0.0356)	0.8538 (0.0084)	1.2064 (0.0178)	0.9126 (0.0393)	0.9172 (0.0346)	0.7442 (0.0081)	1.2205 (0.0118)	1.1408 -	1.0895 -
e	0.7658 (0.0571)	0.6174 (0.0566)	0.9803 (0.0782)	0.7607 (0.0585)	0.7126 (0.0598)	0.5328 (0.0419)	1.4179 (0.0609)	0.5898 -	0.6679 -
i_S	0.8301 (0.0383)	0.5984 (0.0115)	1.0405 (0.0398)	0.7876 (0.0412)	0.7659 (0.0392)	0.4466 (0.0103)	0.9439 (0.0169)	0.7686 -	0.5791 -
i_L	0.7541 (0.0354)	0.7174 (0.0056)	0.6917 (0.0107)	0.7527 (0.0390)	0.7903 (0.0333)	0.4277 (0.0064)	0.7448 (0.0070)	0.6464 -	0.5012 -

Notes: The figures refer to the ratio of the RMSE corresponding to the model to the RMSE of random walk predictions. Average CRPS in parentheses. Results based on rolling forecasts over the time period 2009Q1-2012Q4. NC stands for the normal conjugate prior, M stands for the Minnesota prior, SUR stands for the single unit root prior, IW stands for the independent normal Wishart prior, M-IW stands for the Minnesota-Wishart version of the IW prior, SSVS stands for the SSVS prior, Diffuse stands for the model estimated using maximum likelihood, Pesaran stands for the cointegrated GVAR specification and AR stands for the first-order autoregressive model. Bold figures refer to the lowest value across models for a given endogenous variable.

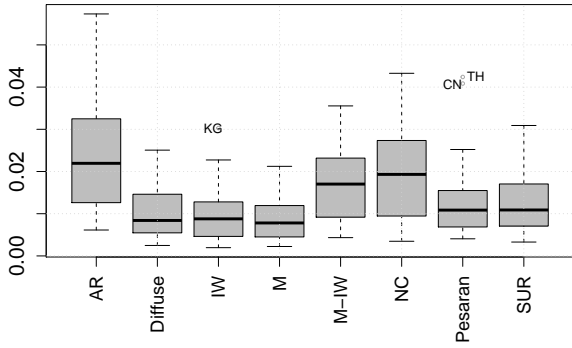
as a determinant of the predictive ability of B-GVAR structures. The two Minnesota prior structures (the standard Minnesota prior and the Minnesota-inverse-Wishart prior) and the SSVS prior present the best predictive performance across variables and forecasting horizons. Especially on the one-year-ahead horizon, the SSVS prior specification performs extraordinarily well, outperforming all other specifications to a large extent. These are the only priors which systematically improve over the the random walk predictions over all variables and at both forecasting horizons. In addition, the B-GVAR with SSVS priors beats all other models for all variables at all horizons in terms of forecast ability based on the predictive density.

The flexibility of the SSVS prior appears to pay off in terms of predictive ability. Since each country is allowed to have its own equation-wise structure, country specifics can play out strongly, which is one of the features of GVARs that make them attractive in the first place. Our results suggest that accommodating country-specifics while controlling for international linkages in the data can help boost the forecasting performance of global macroeconomic models. Another way of endogenizing the elicitation for the prior hyperparameters in SSVS priors was recently put forward in Giannone et al. (2012). They advocate imposing a prior on the hyperparameters of the Minnesota prior. Posterior inference for the hyperprior is done using a simple MCMC scheme, where the marginal likelihood is maximized with respect to the hyperparameters. This implies that the appropriate degree of shrinkage is selected in a data-driven fashion. Our findings, however, reveal the important role of country-specifics in the GVAR framework, which is corroborated by the performance of the SSVS prior in terms of both forecast accuracy and the density forecast performance measures.

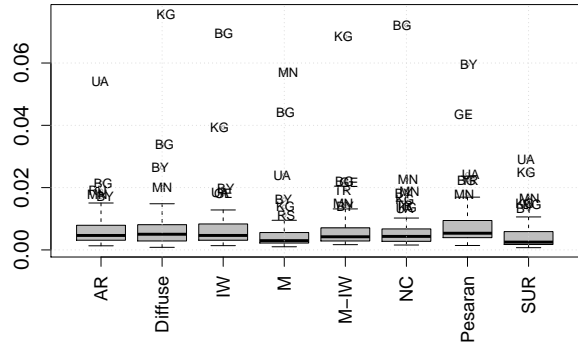
Figures 1 and 2 show boxplots depicting the distribution of average root mean squared errors across countries in our sample. The superiority of forecasts based on B-GVAR models with Minnesota and SSVS priors is also visible in terms of the dispersion of prediction errors across economies, both at the one-quarter-ahead and four-quarters-ahead horizons. The predictive ability of the best performing specification appears thus relatively homogeneous across variables and economies. Although no discernible robust pattern is present linking the level of

Figure 1: Cross-country distribution of RMSEs - 1-step ahead forecasts.

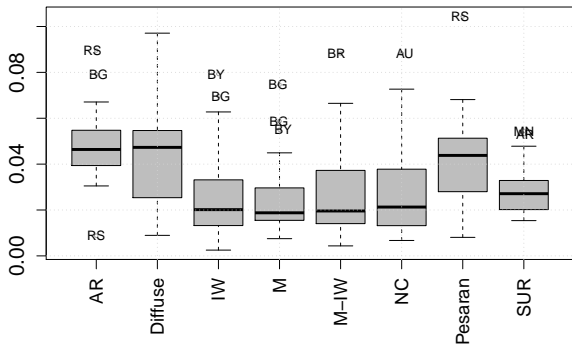
(a) Real GDP



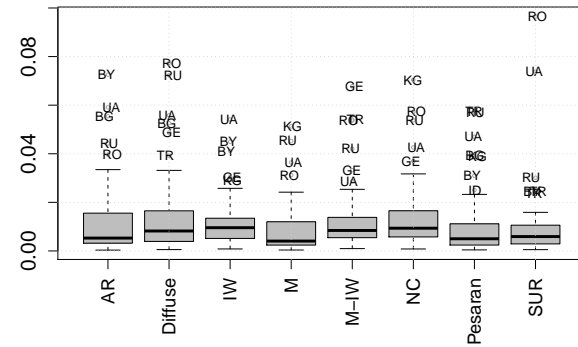
(b) Inflation



(c) Real Exchange Rate



(d) Short-term interest rates



(e) Long-term interest rates

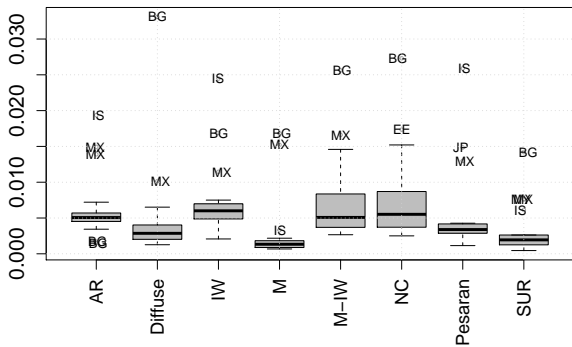
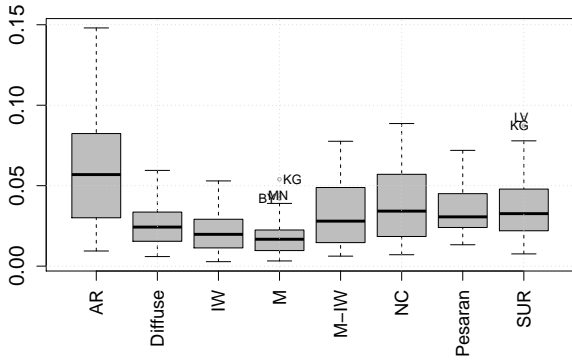
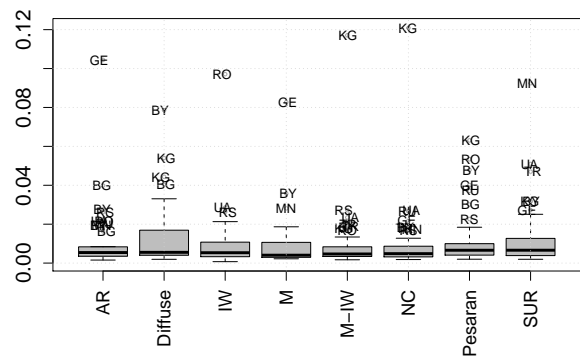


Figure 2: Cross-country distribution of RMSEs - 4-step ahead forecasts.

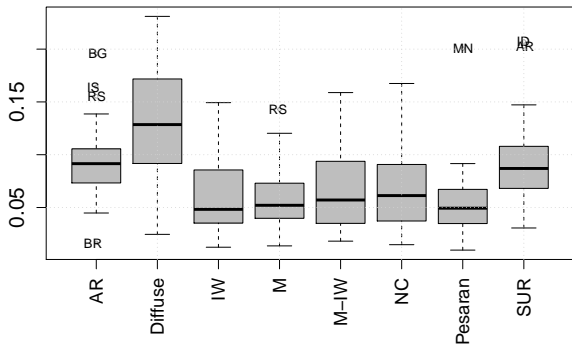
(a) Real GDP



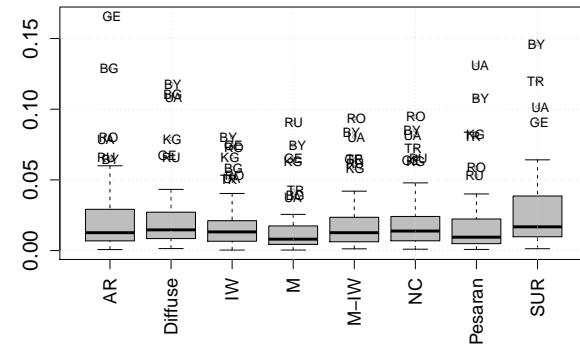
(b) Inflation



(c) Real Exchange Rate



(d) Short-term interest rates



(e) Long-term interest rates

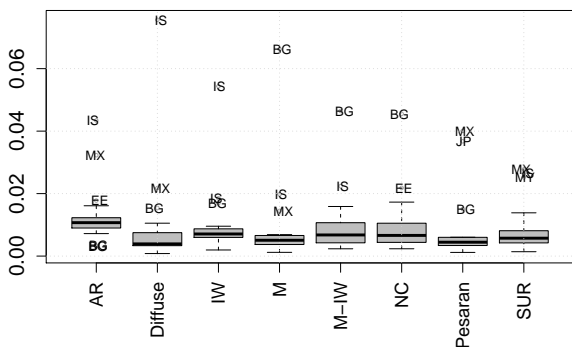
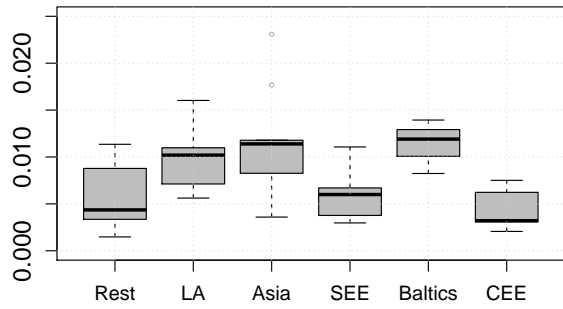
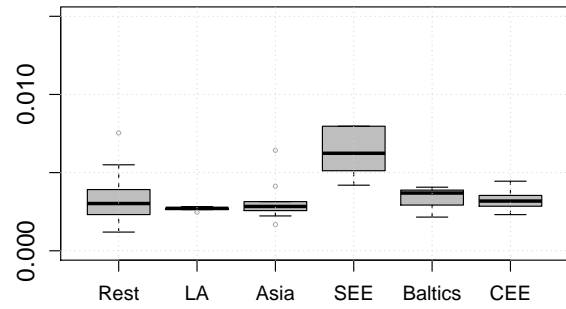


Figure 3: Aggregate RMSE Distribution across country groups - 1-step ahead forecasts.

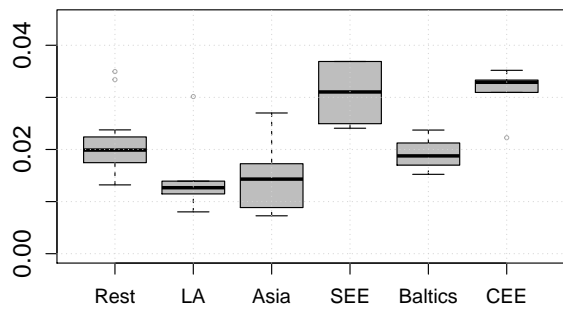
(a) Real GDP



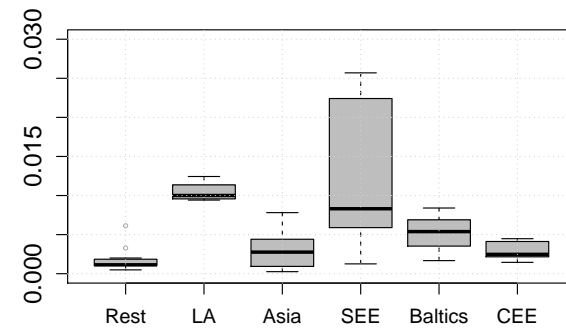
(b) Inflation



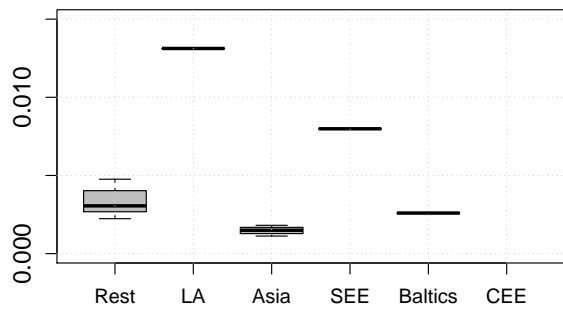
(c) Real Exchange Rate



(d) Short-term interest rates



(e) Long-term interest rates



development of the country (as measured by income per capita) and the average prediction error, the outlying prediction errors tend to correspond to middle-income emerging markets. This is the case for practically all of the specifications and prior structures used and for all the variables being predicted. Given the relatively higher volatility of macroeconomic fundamentals in emerging markets for the sample at hand, the generality of such a result across specifications and priors does not appear surprising for the setting used in the forecasting comparison.

Figure 3 sheds some light on the aggregate forecasting performance of the B-GVAR with the SSVS prior across country groups. Three regularities are worth emphasizing. First, for real GDP, short-term interest rates and inflation, the B-GVAR shows the strongest performance in RMSE terms for developed countries, whereas forecasts for emerging market economies tend to be noisier in general. Second, the intra-group dispersion of RMSEs is remarkably low for most groups and variables. This is especially true for inflation forecasts in Latin America, Asia and CEE, where the RMSE distribution appears to be extremely homogeneous. Finally, figure 3d helps to explain the weak performance of the B-GVAR when it comes to forecasting short-term interest rates. Especially for the Latin American- and southeastern European country blocks, short-term interest rate predictions tend to show a high intra-group RMSE dispersion and a higher overall average RMSE as compared to the other blocks.

4.3 Robustness Checks and Additional Results

We check the robustness of the results discussed above considering different time frames for the estimation and forecasting subsamples. Increasing the length of the forecasting sample tends to have a negative effect on the predictive accuracy of the models under consideration. In particular, when the sample is expanded to the extent that the beginning of the financial crisis is incorporated in the forecasting window, forecasts tend to become more volatile for most prior specifications. For both the SSVS prior, Minnesota prior and the single unit root prior, the results stay however extremely robust. Including the beginning of the crisis in the hold-out sample translates into a 4 to 5 percent loss of predictive accuracy in terms of RMSE for the Minnesota, SSVS and single unit root priors. This compares to a 20 to 40 percent loss in accuracy for the traditional cointegrated GVAR and the B-GVAR with a diffuse prior. Increasing the length of the forecasting window thus does not change our conclusions concerning the relative performance of the prior specifications and that the induced shrinkage tends to make predictions more robust.

Another issue that deserves further scrutiny has to do with the estimation of the models in (log) levels or in growth rates. For standard BVARs, Carriero et al. (2013) conduct a forecasting exercise and compare the predictive accuracy of specifications in levels versus specifications in growth rates, reaching the conclusion that the model in growth rates generally yields more accurate forecasts. However, our results suggest that a (log) level specification delivers better results in terms of forecasting accuracy (see Table A1). The SSVS and Minnesota priors, which are also the specifications performing best for most variables when the system is estimated in first differences, lead to prediction errors which are about 10 to 30 percent higher in terms of RMSE as compared to the specifications in levels. For GDP forecasts, the level specification leads to RMSEs which are around 30 percent lower, while for inflation and exchange rates these differences are rather small (around 5 percent). The RMSEs for short and long term interest rate forecasts are around 10 and 8 percent higher for the growth rate specification. For the 4-steps ahead forecasts, the differences in RMSE terms become even larger, from around 45 percent for GDP to around 15 percent for the real exchange rates.

We also changed the prior setting by eliciting the hyperparameters of the selected priors using pre-sample information. In particular, for countries where data prior to 1995 are available, we use the pre-1995 observations to elicit the hyperparameters of the Minnesota and single-unit-root prior. In the case of the SUR prior, we use the average of the pre-1995 observations to construct the dummy observations. For the Minnesota prior, we scale the prior-variance matrix on the coefficients using the OLS residuals retrieved from univariate regressions on the pre-1995 part of the data.¹³ The results indicate that using pre-sample information to construct the prior improves the forecasts only marginally on average. However, for some countries like the United States and Great Britain, forecasts based on the setting employing pre-sample information tend to be more accurate. For other countries – especially emerging market economies – the opposite is the case. Interestingly, for the particular countries where the predictive ability is relatively worse (Latvia, Slovakia, Hungary) no pre-sample information was used since data prior 1995 are not available. The forecasting performance of these countries is therefore negatively influenced by including pre-sample information for their trading partners. On average, the differences are rather small, indicating no systematic difference between our baseline results and the pre-sample priors.¹⁴

5 Conclusions

In this paper we develop a Bayesian GVAR model and assess its out-of-sample predictive ability for macroeconomic variables employing seven different prior settings. Our results indicate that taking international linkages among the economies into account significantly improves forecasts for real output, the real exchange rate and inflation. For short-term and long-term interest rates, forecast gains relative to the univariate model are modest. With the exception of the predictive results related to the real exchange rate, these results are well in line with findings of Pesaran et al. (2009). On the other hand, Bayesian GVARs tend to outperform the standard cointegrated GVAR model put forward in Pesaran et al. (2004). This result is mainly driven by the shrinkage capabilities inherent in Bayesian priors, while in comparison the standard cointegrated GVAR is prone to overfitting. Forecasts for the cointegrated GVAR tend to be better at the four-quarter-ahead relative to the one-quarter-ahead horizon and might even further improve at longer horizons (Greenwood-Nimmo et al., 2012). Within the class of Bayesian GVARs the choice of prior on the parameters plays an important role in predictive ability. In particular, the SSVS prior (and to a lesser extent, the Minnesota prior) tends to present the best forecasting ability among the specifications entertained. This prior systematically outperforms forecasts from the naive benchmark model, the cointegrated GVAR model and other B-GVARs. This holds true for all variables – although to a different quantitative degree – and for both forecasting horizons considered.

Our findings thus recommend the use of Bayesian GVAR coupled with the stochastic search variable selection prior for global macroeconomic specifications aimed at forecasting. An additional virtue of our proposed model is the explicit treatment of uncertainty regarding variable choice for the individual country models. The strong forecasting performance of the SSVS prior

¹³This implies using the time span from 1980Q1 to 1994Q4 (if available) to retrieve the σ_i s. If the number of pre-sample observations for a country is not sufficient to ensure proper estimation, we switch back to the standard setting for that country.

¹⁴ Figure A1 in the Appendix illustrates the differences by showing the relative RMSE of the SUR prior when the period from 1980 to 1995 is used to inform the prior with respect to the RMSE without pre-sample information.

implies that global analysis requires a proper treatment of the individual country models in the first place. These results are robust with respect to different hold-out sample periods, including the outbreak of the global financial crisis, specification of the variables in first differences rather than in levels and using pre-sample observations as training samples for the priors examined in the paper. Our results do not reveal any discernible systematic patterns in terms of the relationship between predictive errors and country characteristics. This suggests that the model is equally suitable for emerging markets and advanced economies.

References

- Alessi L, Bańbura M. 2009. Comparing global macroeconomic forecasts. Mimeo.
- Álvarez LJ, Ballabriga FC. 1994. BVAR Models in the Context of Cointegration. A Monte Carlo Experiment. *Bank of Spain, Research Department WP Nr. 9405*.
- Bañbura M, Giannone D, Reichlin L. 2010. Large bayesian vector auto regressions. *Journal of Applied Econometrics* **25**: 71–92.
- Carney M, Cunningham P. 2006. Evaluating Density Forecasting Models. *Dublin, Trinity College Dublin, Department of Computer Science 2006-21*.
- Carriero A, Clark TE, Marcellino M. 2013. Bayesian VARs: specification choices and forecast accuracy. *Journal of Applied Econometrics* .
- Carriero A, Kapetanios G, Marcellino M. 2009. Forecasting exchange rates with a large Bayesian VAR. *International Journal of Forecasting* **25**: 400–417.
- Chudik A, Fratzscher M. 2011. Identifying the global transmission of the 2007-2009 financial crisis in a GVAR model. *European Economic Review* **55**: 325–339.
URL <http://ideas.repec.org/a/eee/eecrev/v55y2011i3p325-339.html>
- Dees S, di Mauro F, Pesaran HM, Smith LV. 2007a. Exploring the international linkages of the euro area: a global VAR analysis. *Journal of Applied Econometrics* **22**.
- Dees S, Holly S, Pesaran HM, Smith VL. 2007b. Long Run Macroeconomic Relations in the Global Economy. *Economics - The Open-Access, Open-Assessment E-Journal* **1**: 1–20.
URL <http://ideas.repec.org/a/zbw/ifweej/5582.html>
- Doan TR, Litterman BR, Sims CA. 1984. Forecasting and Conditional Projection Using Realistic Prior Distributions. *Econometric Reviews* **3**: 1–100.
- Eickmeier S, Ng T. 2011. How do credit supply shocks propagate internationally? A GVAR approach. Discussion Paper Series 1: Economic Studies 2011,27, Deutsche Bundesbank, Research Centre.
URL <http://ideas.repec.org/p/zbw/bubdp1/201127.html>
- Feldkircher M. 2013. A Global Macro Model for Emerging Europe. *Oesterreichische Nationalbank Working Paper Series 185/2013*.
URL <http://ideas.repec.org/p/onb/oenbwp/185.html>
- Galesi A, Sgherri S. 2009. Regional Financial Spillovers Across Europe: A Global VAR Analysis. IMF Working Papers 09/23, International Monetary Fund.
URL <http://ideas.repec.org/p/imf/imfwpa/09-23.html>
- Garrat A, Lee K, Pesaran MH, Shin Y. 2006. *Global and National Macroeconometric Modelling: A Long-Run Structural Approach*. Oxford University Press.
- George EI, Sun D, Ni S. 2008. Bayesian stochastic search for var model restrictions. *Journal of Econometrics* **142**: 553–580.
- Giannone D, Lenza M, Primiceri GE. 2012. Prior selection for vector autoregressions. Technical report, National Bureau of Economic Research.
- Giannone D, Reichlin L. 2009. Comments on "Forecasting economic and financial variables with global VARs". *International Journal of Forecasting* **25**: 684–686.
URL <http://ideas.repec.org/a/eee/intfor/v25y2009i4p684-686.html>
- Gneiting T, Raftery AE, Westveld III AH, Goldman T. 2005. Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. *Monthly Weather Review* **133**: 1098–1118.
- Greenwood-Nimmo M, Nguyen VH, Shin Y. 2012. Probabilistic Forecasting of Output Growth, Inflation and the Balance of Trade in a GVAR Framework. *Journal of Applied Econometrics* **27**: 554–573.
- Han F, Ng TH. 2011. ASEAN-5 Macroeconomic Forecasting Using a GVAR Model. *Asian Development Bank, Working Series on Regional Economic Integration* **76**.
- Hersbach H. 2000. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting* **15**: 559–570.
- Kadiyala K, Karlsson S. 1993. Forecasting with generalized bayesian vector auto regressions. *Journal of Forecasting* **12**: 365–378.

- Karlsson S. 2012. Forecasting with Bayesian Vector Autoregressions. *Örebro University Working Paper* **12/2012**.
- Koop G, Korobilis D. 2010. *Bayesian multivariate time series methods for empirical macroeconomics*, volume 3. Now Publishers Inc.
- Litterman R. 1986. Forecasting with Bayesian vector autoregressions - Five years of experience. *Journal of Business and Economic Statistics* **5**: 25–38.
- Pesaran MH, Schuermann T, Smith LV. 2007. What if the UK or Sweden had joined the euro in 1999? An empirical evaluation using a Global VAR. *International Journal of Finance and Economics* **12**: 55–87.
- Pesaran MH, Schuermann T, Smith LV. 2009. Forecasting economic and financial variables with global VARs. *International Journal of Forecasting* **25**: 642–675.
- Pesaran MH, Schuermann T, Weiner SM. 2004. Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model. *Journal of Business and Economic Statistics, American Statistical Association* **22**: 129–162.
- Sims CA. 1992. Bayesian Inference for Multivariate Time Series with Trend. *presented at the American statistical Association meetings* .
- Sims CA, Zha T. 1998. Bayesian Methods for Dynamic Multivariate Models. *International Economic Review* **39**: 949–968.
- Theil H, Goldberger AS. 1961. On pure and mixed statistical estimation in economics. *International Economic Review* **2**: 65–78.
- Zellner A. 1973. *An introduction to Bayesian inference in econometrics*. Wiley.

A Appendix

A.1 MCMC Algorithms

Inverse Wishart Prior Posterior MCMC

In order to analyze the joint posterior, we rely on a Gibbs sampler with the following design:

Gibbs sampler for the independent-normal Wishart prior

1. Draw $\Psi_i^{(k)}$ from the Normal $p(\Psi_i|y, \Sigma)$
2. Draw $\Sigma_i^{-1, (k)}$ from the Wishart $p(\Sigma_i^{-1}|y, \Psi_i)$
3. Repeat steps (1)-(2) N times and discard the first N_{burn} as burn-ins

SSVS Prior Posterior MCMC

For independent Bernoulli priors on $\delta_{i,j}$, the Gibbs sampling scheme is given by the following algorithm:

Gibbs sampler for the independent-normal Wishart prior

1. Draw $\Psi_i^{(k)}$ from $N(\bar{\Psi}_i, \bar{V}_i)$, where $\bar{V}_i = [(DD)^{-1} + \Sigma_i^{-1} \otimes (X_i'X_i)^{-1}]$
2. Given $\Psi_i^{(k)}$, sample $\delta_{i,j}$ from $\delta_{i,j}|\Psi_i^{(j)}, \delta_{-(i,j)} \sim Ber(u_{1,ij}/(u_{1,ij} + u_{0,ij}))$, where $\delta_{-(i,j)}$ denotes all draws for δ except for the j th coefficient. Furthermore, $u_{0,ij} = p(\Psi_i|\delta_{-(i,j)}, \delta_{i,j} = 0)p_{ij}$ and $u_{1,ij} = p(\Psi_i|\delta_{-(i,j)}, \delta_{i,j} = 1)(1 - p_{ij})$, where p_{ij} is the prior inclusion probability of coefficient j in country i .
3. Draw $\Sigma_i^{(k)}$ from the inverse Wishart $p(\Sigma_i|y, \Psi_i, \delta)$
4. Repeat steps (1)-(2) N times and discard the first N_{burn} as burn-ins

Simulation of the Predictive Density

Obtaining an estimate for the global variance-covariance matrix is difficult, especially in a Bayesian setting. In order to avoid this problem, we simulate $u_{i,t}$ at the local level and transform them by using the posterior draws for G . The algorithm for this procedure is described below.

Sampling from the Predictive Density

1. Use $\Sigma_i^{(j)}$ to draw $\tilde{u}_{T+1,i}^{(j)}, \dots, \tilde{u}_{T+h,i}^{(j)}$ from $u_{i,t} \sim N(0, \Sigma_i^{(j)})$
2. Draw $\Psi_i^{(j)}$ from the posterior of country i and solve the global model as described in Section 2 to obtain $F = F^{(j)}$
3. Stack $u_{T+1,i}^{(j)}, \dots, u_{T+h,i}^{(j)}$ for $i = 1, \dots, N$ and transform using $e_t = G^{-1}\tilde{u}_t^{(j)}$
4. Calculate $\tilde{x}_{T+n} = F^n x_T + \sum_{\tau=0}^{n-1} F^\tau [b_0^{(j)} + b_1^{(j)}(T+n-\tau)] + \tilde{u}_{T+h,i}^{(j)}$
5. Repeat steps (1)-(3) N_{pred} times

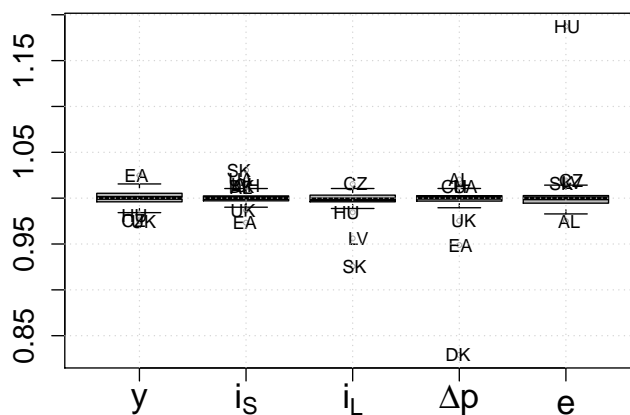
A.2 Robustness

Table A1: Relative Forecasting Performance, 1 Quarter Ahead, Difference Specification: Root Mean Square Error

	NC	M	SUR	IW	M-IW	SSVS	Diffuse
y	1.2450	0.6567	1.5753	0.8433	0.65735	0.8351	0.8031
Δp	1.0748	0.8905	0.8949	1.0382	0.9860	0.8315	1.2431
e	0.6285	0.5279	0.6101	0.6071	0.7373	0.4333	1.0211
i_S	1.2414	0.8028	1.1116	1.4733	0.8808	0.7034	1.5044
i_L	1.2713	0.5191	0.6727	1.6691	1.0637	0.7430	1.0449

Notes: The figures refer to the ratio of the RMSE corresponding to the model to the RMSE of random walk predictions. Results based on rolling forecasts over the time period 2009Q1-2012Q4. NC stands for the normal conjugate prior, M stands for the Minnesota prior, SUR stands for the single unit root prior, IW stands for the independent normal Wishart prior, M-IW stands for the Minnesota-Wishart version of the IW prior, Diffuse stands for the model estimated using maximum likelihood, Pesaran stands for the cointegrated GVAR specification and AR stands for the first-order autoregressive model. Bold figures refer to the lowest value across models for a given endogenous variable.

Figure A1: Cross-country distribution of relative RMSEs for the SUR with pre-sample information over the SUR without - 1-step ahead forecasts.



A.3 Specification of the country models

Countries	Domestic Variables	Foreign Variables	Dummy Variables
EA	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$Dp \times eaD(07Q4-08Q4), eaD(07Q4-08Q4)$
US	$y, \Delta p, i_s, i_l, poil$	$y^*, \Delta p^*, i_s^*, i_l^*, poil^{**}$	$Dp \times usD(07Q4-08Q4), usD(07Q4-08Q4)$
UK	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
JP	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
CN	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
CZ	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times czD(97Q1-97Q2), czD(97Q1-97Q2)$
HU	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
PL	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
SI	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times siD(95Q1-96Q4), siD(95Q1-96Q4)$
SK	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
BG	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times bgD(95Q1-97Q2), Dp \times bgD(95Q1-97Q2)$ $bgD(95Q1-97Q2)$
RO	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times roD(96Q4-97Q3, 98Q1, 98Q4-99Q2),$ $Dp \times roD(96Q4-97Q3, 98Q1, 98Q4-99Q2),$ $e \times roD(96Q4-97Q3, 98Q1, 98Q4-99Q2),$ $roD(96Q4-97Q3, 98Q1, 98Q4-99Q2)$
EE	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times eeD(97Q4-99Q1), eeD(97Q4-99Q1)$
LT	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times ltD(95Q1-96Q4), ltD(95Q1-96Q4)$
LV	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times lvD(95Q1-96Q4), lvD(95Q1-96Q4)$
HR	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times hrD(95Q1-96Q2), hrD(95Q1-96Q2)$
AL	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times alD(97Q1-98Q3), alD(97Q1-98Q3)$
RS	$y, \Delta p, e$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times rsD(98Q4-00Q4), rsD(98Q4-00Q4)$
RU	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times ruD(95Q1-95Q3, 98Q3), ruD(95Q1-95Q3, 98Q3)$
UA	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times uaD(98Q3-99Q4), uaD(98Q3-99Q4)$
BY	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
GE	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times geD(95Q1-96Q4), geD(95Q1-96Q4)$
MN	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times mnD(95Q1-98Q1), mnD(95Q1-98Q1)$
KG	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
AR	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times arD(01Q4-02Q3), e \times arD(01Q4-02Q3),$ $arD(01Q4-02Q3)$
BR	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times brD(95Q1-95Q4), Dp \times brD(95Q1-95Q4),$ $brD(95Q1-95Q4)$
CL	$y, \Delta p, e$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
MX	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$Dp \times mxD(95Q1-98Q4), mxD(95Q1-98Q4)$
PE	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$peD(98Q3)$
KR	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times krD(97Q4-98Q2), Dp \times krD(97Q4-98Q2),$ $krD(97Q4-98Q2)$
PH	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
SG	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
TH	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times thD(97Q3-98Q2), thD(97Q3-98Q2),$ $e \times thD2(95Q1-98Q3), thD2(95Q1-98Q3)$
IN	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$Dp \times inD(98Q4-99Q4), inD(98Q4-99Q4)$
ID	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times idD(97Q3-98Q2), e \times idD(97Q3-98Q2),$ $Dp \times idD(97Q3-98Q2), idD(97Q3-99Q2)$
MY	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$e \times myD(95Q1-97Q4), eD(95Q1-97Q4)$
AU	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
NZ	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
TR	$y, \Delta p, e, i_s$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	$i_s \times trD(00Q4-01Q1), trD(00Q4-01Q1)$
EG	$y, \Delta p, e$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
CA	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
CH	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
NO	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
SE	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
DK	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-
IS	$y, \Delta p, e, i_s, i_l$	$y^*, \Delta p^*, e^*, i_s^*, i_l^*, poil^{**}$	-

Notes: The table represents the general specification and variable cross-country variable coverage of our GVAR model. Throughout the paper we have used 1 lag for endogenous, weakly exogenous and strictly exogenous variables only.