University of Heidelberg

Does Joint Modelling of the World Economy Pay Off? Evaluating Global Forecasts from a Bayesian GVAR

Jonas Dovern, Martin Feldkircher,
and Florian Huber

Does Joint Modelling of the World Economy Pay Off? Evaluating Global Forecasts from a Bayesian GVAR [*]

Jonas Dovern[1], Martin Feldkircher[2], and Florian Huber[2]

[1]Alfred-Weber-Institute for Economics, Heidelberg University
[2]Oesterreichische Nationalbank (OeNB)

Abstract

We analyze how modeling international dependencies improves forecasts for the global economy based on a Bayesian GVAR with SSVS prior and stochastic volatility. To analyze the source of performance gains, we decompose the predictive joint density into its marginals and a copula term capturing the dependence structure across countries. The GVAR outperforms forecasts based on country-specific models. This performance is solely driven by superior predictions for the dependence structure across countries, whereas the GVAR does not yield better predictive marginal densities. The relative performance gains of the GVAR model are particularly pronounced during volatile periods and for emerging economies.

1 Introduction

Forecasters dealing with global economic policy issues or risk management need to have a good understanding of how to optimally model national economies that are strongly integrated with the world economy. Global economic projections must, above all, be coherent across countries.[1]

Against this backdrop, global vector autoregressive (GVAR) models (Pesaran et al., 2004) have been proposed as a forecasting tool (Pesaran et al., 2009) because such models cover all major economies of the world and their interactions. While forecast coherence is potentially the major advantage of GVAR models as a forecasting tool, past assessments of the forecast performance of GVAR models have been conducted variable by variable, i.e., from a univariate perspective. However, for forecasters seeking to forecast a range of variables (e.g., GDP growth in all countries) or to review the joint predictive density of the entire system, such univariate evaluation approaches clearly miss important aspects.[2] The idea is very intuitive: Two vector forecasts with the same accuracy in terms of marginal predictive densities could perform very differently in terms of the copula that links the marginal distributions and captures the dependencies between the vector elements. In fact, even a model that performs significantly worse compared to a benchmark model in terms of predicting the marginal densities could yield superior joint predictive densities.

We add to the literature by assessing the forecast performance of a GVAR model from a multivariate perspective: we analyze whether the rich model structure of GVAR models pays off by yielding superior predictive joint densities relative to

---

[1]Think, for instance, of the business-cycle forecasts published by the International Monetary Fund (IMF) for all member states in its World Economic Outlook (WEO). The IMF has to make sure that the individual country-specific forecasts are coherent in that they reflect the same view about the outlook of the world economy and historical experience with co-movement of business cycles across countries. A recent investigation reports that top-down elements are important in the production process of the WEO (IEO, 2014).

[2]This point was raised already in a comment by Swanson (2009) in response to the seminal paper by Pesaran et al. (2009).

a benchmark that consists of country-specific forecast models. Our paper makes several contributions. First, it is the first analysis of GVAR forecast performance in terms of the overall joint predictive density of all model variables. Second, we suggest to decompose the joint predictive density into the marginal densities and a copula term to analyze the source of performance gains in more detail. Finally, we enhance the Bayesian estimation of GVAR models. We modify the Bayesian GVAR model with stochastic volatility as outlined by Huber (2014) and introduce a variant of the stochastic search variable selection (SSVS) prior. Using a restricted stochastic volatility model in the spirit of Carriero et al. (2012) and a conjugate version of the SSVS prior leads to significant computational gains that are essential when estimating large-scale models such as the GVAR.

Our paper is linked to a number of companion papers that develop a Bayesian framework for GVAR models and analyze the forecast performance of such Bayesian GVAR models. Crespo Cuaresma et al. (2014a) put forward a Bayesian version of the GVAR model and analyze how different prior setups perform in terms of yielding accurate density forecasts. They find that the SSVS prior, on which we concentrade in this paper, exhibit the best forecasting ability among a wide range of priors. Huber (2014) shows how a Bayesian GVAR model can be augmented to account for stochastic volatility. Finally, Dovern and Huber (2015) show, in a complementary paper, that the GVAR model used in the paper at hand yields better turning point prediction than country-specific time-series models.

The main results of this paper can be summarized as follows: First, the GVAR model outperforms, on average, the country-specific Bayesian VAR models with stochastic volatility, which are used as a benchmark, in terms of the quality of the joint predictive density. This holds true also for all types of variables (financial and real) separately. Second, these performance gains are driven by better predictions for the dependence structure across variables and countries rather than by improved predictive marginal densities. Third, the forecast performance of both the GVAR

and the country-specific benchmark models is strongly time-varying, and forecasts deteriorate during the global recession. However, the loss in forecast accuracy during turbulent times is much smaller for the GVAR model than for the benchmark models. That is, incorporating international linkages among countries is essential for forecasting and in particular so during crisis times. Last, and taking a regional angle, the gains of using a world-wide model seem particularly pronounced for emerging economies. More specifically, while taking cross-country linkages between advanced economies into account seems sufficient to generate accurate forecasts for these countries, forecasts for emerging economies tend to improve when based on a truly global sample. Taken at face value this means that modeling spillovers from advanced to emerging economies tends to improve forecasts for these countries, whereas the opposite is not true.

The remainder of this paper is structured as follows. Section 2 summarizes the relevant literature. Section 3 presents the model framework that we use. It explains the structures of the GVAR model as well as our benchmark model and describes our estimation strategy. Section 4 documents the data set that we use and specifies the choices that we make with respect to the priors of the Bayesian GVAR model. Section 5 describes the design of our out-of-sample forecast analysis and documents the evaluation criteria that we use. Section 6 contains the empirical results. It documents the baseline results and shows how the performance of GVAR forecasts varies over time and across different country groups. Finally, Section 7 concludes the paper.

2   Related Literature

Our paper builds on the growing literature on GVAR models. This type of model, which was proposed by Pesaran et al. (2004) and further developed by Dées et al. (2007a) and Dées et al. (2007b), is designed to model dependencies across a wide range of macroeconomic variables from a large number of countries. GVAR models

3

capture the dynamics of a large part of the world economy by linking country-specific vector autoregressive (VAR) models to each other. GVAR models were first developed as a tool for financial stress-testing following the Asian crisis in the late 1990s. Since then, they have been heavily used to analyze the dynamic effects of a wide range of country-specific or global shocks.[3]

More recently, GVAR models have been proposed as a tool for forecasting global economic activity (Pesaran et al., 2009). Most studies attest GVAR models a forecast performance that is about as good or slightly better, on average, as that of simpler benchmark models. Pesaran et al. (2009) find superior performance of GVAR forecasts (relative to univariate benchmark models) for output growth and inflation when they obtain GVAR forecasts by averaging over different model specifications.[4] In contrast, the performance is more mediocre for forecasting financial variables such as interest rates or equity prices. Greenwood-Nimmo et al. (2012) use a GVAR model with country-specific intercept shifts to produce probabilistic forecasts for a number of variables in major economies. They show that forecasts based on the GVAR model outperform univariate benchmark models especially over long forecast horizons. Crespo Cuaresma et al. (2014a) estimate a Bayesian GVAR model and show that forecasts based on this model outperform forecasts from several univariate benchmark models for virtually all variables and different forecast horizons. Two other recent contributions have looked at the forecast performance of GVAR models for small open economies: Assenmacher (2013) shows that a small GVAR model for the Swiss economy delivers forecasts that are comparable in accuracy with widely used benchmark forecasts. de Waal et al. (2013) show that a richer GVAR

---

[3]GVAR models have, for instance, been used to analyze the international transmission of oil price shocks (Cashin et al., 2014), house price shocks (Cesa-Bianchi, 2013), credit supply shocks (Eickmeier and Ng, 2015), cost-push shocks (Galesi and Lombardi, 2013), financial stress shocks (Dovern and van Roye, 2014), monetary policy shocks (Feldkircher and Huber, 2015), liquidity shocks during the Great Recession of 2007-2009 (Chudik and Fratzscher, 2011), and for stress-testing of the financial sector (Castrén et al., 2010). For a more complete overview, see Chudik and Pesaran (2014).

[4]It remains unclear from their paper how much the performance of the univariate models could be improved by averaging over different models/specifications. The study by Pesaran et al. (2009) has been updated with similar findings by Smith (2013).

model delivers better forecasts for the South African economy than a small version including only the three main trading partners. They show, however, that Bayesian country-specific VAR models and univariate autoregressive models tend to beat the GVAR models in terms of forecast accuracy.

Our work relates also to the literature on econometric models that account for stochastic volatility. The recent synchronous increase in the volatility of many macroeconomic variables suggests that traditional linear models fail to properly capture salient features of the data. Clark (2011) uses a standard VAR augmented by stochastic volatility (SV) to produce predictive densities for a set of macroeconomic aggregates. He concludes that allowing for stochastic volatility generally improves the accuracy of density forecasts. Recently, Carriero et al. (2012) proposed a Bayesian VAR with stochastic volatility, which exploits the fact that most macroeconomic variables obey the same pattern of realized volatility. They conclude that imposing a factor structure on the latent log-volatilities helps to improve the accuracy of the density forecasts at little additional costs in terms of computational demands. Huber (2014) proposes a GVAR model with a factor SV structure. He reports that allowing each country's volatility to be driven by a country-specific latent factor improves forecasts of GDP and short-term interest rates, while leading to no improvements for forecasts of inflation, real exchange rates and long-term interest rates.

The third related strand of literature is concerned with multivariate evaluation of vector forecasts. Since we are going to evaluate density forecasts, the methods that we use are most closely linked with the literature on multivariate evaluation of such forecasts which was initiated among others by Diebold et al. (1999) and which is surveyed in Gneiting (2008). Our study is complementary to a number of recent papers that address the issue of evaluating vector-valued *point* forecasts from a multivariate perspective (see, e. g., Eisenbeis et al., 2002; Bauer et al., 2003; Sinclair and Stekler, 2013; Müller-Dröge et al., 2014).

3 Model Framework

GVAR models (Pesaran et al., 2004) are designed to capture the dynamics of a large part of the world economy by linking country-specific vector autoregressive models to each other using trade weights. Though GVAR models are linear, they allow for a range of different interdependencies between variables and countries, such as long-run relationships consistent with theory, short-run spillover effects or cross-sectional dependence in the error structure. Thus, they offer a fair degree of flexibility in modeling the business-cycle dynamics of the world economy in a coherent fashion. In the following sections, we briefly sketch the assumptions that we make about the specification of the GVAR model, and we explain what priors we use for the Bayesian estimation of the model and how we generate density forecasts based on the estimated GVAR model.

3.1 Global Vector Autoregressive Models

Basically, a GVAR model consists of a number of country-specific models that are combined to form the global model. In a first step, the country-specific models can be estimated individually under certain restrictions. In a second step, the GVAR model is "solved" by combining the individual models. Consider a sample of $N+1$ different countries. Let $x_{i,t}$ be the $k_i \times 1$ vector of domestic random variables for country $i = 0, \ldots, N$ and time $t = 1, \ldots, T$. For each country, we consider a VAR model which is augmented with a set of (weakly exogenous) foreign variables (VARX*). To enhance readability of the notation, we restrict all formulas to a VARX*(1,1) specification although we allow for three lags in our empirical application:

$$x_{i,t} = a_{i0} + a_{i1}t + \Phi_{i1}x_{i,t-1} + \Lambda_{i0}x^*_{i,t} + \Lambda_{i1}x^*_{i,t-1} + \varepsilon_{i,t}. \tag{3.1}$$

Here, a_{i0}, a_{i1}, Φ_{i1}, Λ_{i0} and Λ_{i1} are properly sized coefficient matrices measuring the impact of deterministic components, lagged domestic variables and foreign variables.

6

The foreign variables are calculated as a weighted average of the domestic variables of all other countries based on trade weights, $x^*_{i,t} = \sum_{j=0}^{N} w_{ij} x_{j,t}$, under the restriction that $w_{ii} = 0$ and $\sum_{j=0}^{N} w_{ij} = 1$. The assumption that $x^*_{i,t}$ is weakly exogenous at the individual country level reflects the belief that most countries are small relative to the world economy.

Most existing papers on GVAR models make the assumption of a fixed covariance matrix for $\varepsilon_{i,t}$. We are skeptical about this approach, tough. Especially for fast-moving financial variables, this assumption proves inadequate in the presence of sudden shifts in the level of volatility and volatility-clustering. Therefore, we assume that $\varepsilon_{i,t} \sim N(0, \Sigma_{i,t})$. Following Carriero et al. (2012) and Huber (2014), we use the following stochastic volatility specification

$$\Sigma_{i,t} = \exp(h_{i,t}) \times \Sigma_i \tag{3.2}$$

$$h_{i,t} = \kappa_i + \rho_i(h_{i,t-1} - \kappa_i) + \eta_i u_{i,t} \tag{3.3}$$

$$u_{i,t} \sim N(0,1), \tag{3.4}$$

where $h_{i,t}$ denotes a country-specific log-volatility process, with κ_i being its unconditional mean and ρ_i an autoregressive parameter.[5] Finally, η_i denotes the variance of the log-volatility process. To identify the model it is sufficient to set $h_{i,0} = 0$. Note that we assume that the variance of the entire VARX* is driven by a single latent process. This is justified because most macro variables exhibit similar volatility dynamics and especially when it comes to predicting financial crises the correlation between variables tends to increase dramatically.[6] Huber (2014) showed empirically that (for a smaller information set) the differences between a full stochastic volatility

---

[5]We depart from the macroeconomic literature by assuming that the log-volatility follows an autoregressive process of order one rather than a random walk. This choice, motivated in Eisenstat and Strachan (2014) implies that in the limit, the level of the volatility is bounded in probabilistic terms.

[6]Even though the number of additional parameters to be estimated would not increase dramatically by assuming that k_i log-processes govern the behavior of $\Sigma_{i,t}$, we would lose several convenient features of the estimation algorithm, which would induce significant computational costs.

specification and the simplified variant proposed above tend to be rather small for most countries considered.

Pesaran et al. (2004) show how the $N + 1$ country models can be combined to yield a global VAR representation. Defining a $(k_i + k_i^*) \times 1$ vector $z_{i,t} = (x'_{i,t}, x_{i,t}^{*}{}')'$, collecting all contemporaneous terms on the left-hand side and ignoring deterministic terms for notational simplicity, we can rewrite equation (3.1) as

$$A_i z_{i,t} = B_i z_{i,t-1} + \varepsilon_{i,t}, \tag{3.5}$$

with $A_i = (I_{k_i}, -\Lambda_{i,0})$ and $B_i = (\Phi_{i1}, \Lambda_{i1})$ denoting $k_i \times (k_i + k_i^*)$-dimensional matrices, respectively. We collect all endogenous variables in a $k \times 1$ global vector $x_t = (x'_{0,t}, x'_{1,t}, \ldots, x'_{N,t})'$, where $k = \sum_{i=0}^{N} k_i$ is the total number of endogenous variables in the GVAR. By defining a suitable $(k_i + k_i^*) \times k$ linking matrix W_i, it is possible to rewrite $z_{i,t}$ exclusively in terms of x_t and W_i. More specifically, the linking matrix is set such that the following equality holds

$$z_{it} = W_i x_t. \tag{3.6}$$

Inserting equation (3.6) into equation (3.5) and stacking the models for all countries yields

$$G x_t = H x_{t-1} + \varepsilon_t, \tag{3.7}$$

where $G = ((A_0 W_0)', \ldots, (A_N W_N)')'$, $H = ((B_0 W_0)', \ldots, (B_N W_N)')'$, and $\varepsilon_t = (\varepsilon'_{0,t}, \ldots, \varepsilon'_{N,t})'$. Multiplying from the left by G^{-1} yields the reduced-form GVAR representation:

$$x_t = F x_{t-1} + e_t, \tag{3.8}$$

where $F = G^{-1} H$ and $e_t = G^{-1} \varepsilon_t$. Equation (3.8) resembles a standard first-order reduced form VAR. The structure of the VARX* model induces restrictions on the parameter matrix F. In addition to the restrictions imposed on F, the structure of

8

the model has important implications for the specific form of the variance-covariance matrix of e_t, $\Sigma_{e,t}$. In the present application, $\Sigma_{e,t}$ is a positive definite matrix, given by

$$\Sigma_{e,t} = G^{-1}\Sigma_{\varepsilon,t}G^{-1\prime}. \tag{3.9}$$

$\Sigma_{\varepsilon,t}$ is a block-diagonal matrix which consists of the country-specific variance-covariance matrices $\Sigma_{i,t}$. The block-diagonality of $\Sigma_{\varepsilon,t}$ is predicated by the fact that the inclusion of the weakly exogenous variables accounts for cross-country correlation and renders the estimation problem of the $N + 1$ submodels embarrassingly parallel, providing significant computational advantages.

3.2   Selection of Priors and Estimation

We consider a Bayesian version of the GVAR model as introduced by Crespo Cuaresma et al. (2014a) and rely on the stochastic search variable selection (SSVS) prior on the coefficients, which Crespo Cuaresma et al. (2014a) demonstrated to yield excellent forecasts. Since we are going to use a slightly more complex framework allowing for SV, we modify the prior setup from Crespo Cuaresma et al. (2014a) and propose a (conditionally) conjugate version of the SSVS prior. This allows us to reduce the computational burden significantly. Additionally, as noted above, the factor structure of the log-volatilities permits us to exploit a convenient Kronecker structure of the likelihood function.

For prior implementation it proves convenient to rewrite the VARX* in equation (3.1) in the following way

$$x_{i,t} = \Gamma_i Z_{it} + \varepsilon_{it}, \tag{3.10}$$

where $Z_{i,t} = (1, t, x'_{i,t-1}, x^{*}_{i,t}{}', x^{*}_{i,t-1}{}')'$ is a K_i-dimensional data vector, where $K_i = 2 + k_i + 2k_i^*$. The $k_i \times K_i$ matrix of stacked coefficients is denoted by $\Gamma_i = (a_{i0}, a_{i1}, \Phi_{i1}, \Lambda_{i0}, \Lambda_{i1})$. Furthermore, let us denote the vectorized coefficient matrix as $\Psi_i = \text{vec}(\Gamma_i)$. Our stochastic volatility setup can be implemented in a straight-

9

forward fashion by dividing $x_{i,t}$ and $Z_{i,t}$ by $\exp(h_{it}/2)$, where the normalized data matrices are denoted by $\tilde{x}_{i,t}$ and $\tilde{Z}_{i,t}$. Finally, we denote full data matrices by $\tilde{Z}_i = (\tilde{Z}'_{i,0}, \tilde{Z}'_{i,1}, \ldots, \tilde{Z}'_{i,T})'$ and $\tilde{x}_i = (\tilde{x}'_{i,0}, \tilde{x}'_{i,1}, \ldots, \tilde{x}'_{i,T})'$.

As advocated in Carriero et al. (2012), we use the following conjugate prior setup on Ψ_i and Σ_i^{-1}

$$\Psi_i | \Sigma_i^{-1}, \psi_i \sim \mathcal{N}(\underline{\mu}_{\Psi_i}, \Sigma_i \otimes \underline{V}_{\Psi_i}), \tag{3.11}$$

$$\Sigma_i^{-1} \sim \mathcal{W}(\underline{v}_i, \underline{S}_i), \tag{3.12}$$

where $\psi_i = (\psi_{i1}, \ldots, \psi_{iK_i})'$ denotes a vector of binary random variables. This implies that conditional on ψ_i, the prior on the dynamic coefficients is a standard conjugate prior with mean $\underline{\mu}_{\Psi_i}$ and variance $\Sigma_i \otimes \underline{V}_{\Psi_i}$. Knowledge of ψ_i fully specifies $\underline{V}_{\Psi_i}$. More specifically, $\underline{V}_{\Psi_i}$ is a diagonal matrix where the $(j,j)^{th}$ element is given by v_j, with

$$v_j = \begin{cases} \tau_{0j}^2 & \text{if } \psi_{ij} = 0 \\ \tau_{1j}^2 & \text{if } \psi_{ij} = 1. \end{cases} \tag{3.13}$$

τ_{0j} and τ_{1j} are prior-variances set such that $\tau_{1j} \gg \tau_{0j}$. By setting $\underline{\mu}_{\Psi_i}$ equal to zero, equation (3.13) implies that we impose a mixture prior on the different variable types in (3.1) similar to the prior discussed in Koop (2013). To see this, let us define the element corresponding to the lth equation and the jth variable as $[\Gamma_i]_{lj}$. Under the assumption that the prior mean is set equal to zero, we impose the following mixture prior

$$[\Gamma_i]_{lj} | \psi_{ij} \sim (1 - \psi_{ij}) \mathcal{N}(0, \sigma_{ill} \tau_{0,ij}^2) + \psi_{ij} \mathcal{N}(0, \sigma_{ill} \tau_{1,ij}^2), \tag{3.14}$$

for $l = 1, \ldots, k_i, \ j = 1, \ldots, K_i$ and $\sigma_{ill} = [\Sigma_i]_{ll}$. Thus ψ_{ij} controls which of the two Normal distributions applies for the jth regressor. Since the Normal distributions differ only by their variances, and the variances are set such that $\tau_{1j} \gg \tau_{0j}$, we

either impose a dogmatic or a loose prior on $[\Gamma_i]_{lj}$.[7] In contrast to the traditional implementation of the SSVS prior in the spirit of George et al. (2008), the conjugate SSVS specification either includes or excludes a given explanatory variable in all equations of a particular country model. The SSVS prior thus allows us to account for model uncertainty at the individual country level. This is important since our sample spans a broad set of different economies, each being fundamentally different from each other. Thus assuming too much homogeneity in terms of model specification could distort inference and lead to a poor out-of-sample forecasting performance. In addition, uncertainty with respect to the appropriate lag-length is also tackled through the SSVS prior.

For Σ_i^{-1} we impose a standard Wishart prior with scaling matrix $\underline{S}_i$ and degrees of freedom $\underline{v}_i$. Note that we assume prior dependence between Σ_i^{-1} and Ψ_i.

Finally, we also have to impose prior distributions on the parameters of the log-volatility equation. Following Frühwirth-Schnatter and Wagner (2010) and Kastner and Frühwirth-Schnatter (2014), we impose the following set of priors:

$$\kappa_i \sim \mathcal{N}(\underline{\mu}_{\kappa_i}, \underline{V}_{\kappa_i}), \tag{3.15}$$

$$\frac{\rho_i + 1}{2} \sim \mathcal{B}(a_1, b_1), \tag{3.16}$$

$$\eta_i \sim \mathcal{G}(\frac{1}{2}, \frac{1}{2}R_{\eta_i}). \tag{3.17}$$

We use a Normal prior on the level of the log-volatility process κ_i, a Beta prior on the autoregressive parameter ρ_i, and a non-conjugate Gamma prior on the variance of the log-volatility specification with a scalar hyperparameter R_{η_i}.

Due to the conjugacy of the prior setup, "dummy observations" can be used to implement the prior (see Bańbura et al., 2010; Koop, 2013). This implies that the conjugate SSVS prior can be implemented using the following artificial observations

---

[7]Thus, if ψ_{ij} equals zero, the prior variance is set to a very small value, indicating that the posterior distribution of the associated coefficient will shrink towards zero. Otherwise, the coefficient could be anything on the real line.

$$\underline{x}_i = \begin{pmatrix} \underline{V}_{\Psi_i}^{-\frac{1}{2}} \underline{A} \\ \text{diag}(s_1, ..., s_{k_i}) \end{pmatrix}, \quad \underline{Z}_i = \begin{pmatrix} \underline{V}_{\Psi_i}^{-\frac{1}{2}} \\ 0_{k_i \times K_i} \end{pmatrix}, \qquad (3.18)$$

where $\underline{V}_{\Psi_i}^{-\frac{1}{2}}$ is defined by $\underline{V}_{\Psi_i} = \underline{V}_{\Psi_i}^{\frac{1}{2}} \underline{V}_{\Psi_i}^{\frac{1}{2}}$ and $\underline{A}$ equals the zero matrix.[8] Following Doan et al. (1984), $s_1, \ldots, s_{k_i}$ are the residual standard deviations from a set of k_i univariate autoregressions of order four for the elements of $x_{i,t}$. Finally, $0_{k_i \times K_i}$ denotes a $k_i \times K_i$ matrix of zeros.

Combining the "dummy observations" with the real data, we obtain the following quantities:

$$\overline{x}_i = \begin{pmatrix} \tilde{x}_i \\ \underline{x}_i \end{pmatrix}, \quad \overline{Z}_i = \begin{pmatrix} \tilde{Z}_i \\ \underline{Z}_i \end{pmatrix}. \qquad (3.19)$$

It can be shown, that this prior setup gives rise to the following posterior distributions:

$$\Psi_i | \Sigma_i^{-1}, \psi_i, h_i, \mathcal{D}_{iT} \sim \mathcal{N}(\overline{\mu}_{\Psi_i}, \overline{V}_{\Psi_i}), \qquad (3.20)$$

$$\Sigma_i^{-1} | \Psi_i, \psi_i, h_i, \mathcal{D}_{iT} \sim \mathcal{W}(\overline{v}_i, \overline{S}_i), \qquad (3.21)$$

where $h_i = (h_{i,0}, \ldots, h_{i,T})$ denotes the full history of log-volatilities and $\mathcal{D}_{iT}$ denotes the data specific to country i up to time T. The posterior mean and variance for Ψ_i are given by $\overline{V}_{\Psi_i} = (\overline{Z}_i' \overline{Z}_i)^{-1}$ and $\overline{\mu}_{\Psi_i} = \overline{V}_{\Psi_i} \overline{Z}_i' \overline{x}_i$, the posterior scaling matrix for Σ_i^{-1} is given by $\overline{S}_i = (\overline{x}_i - \overline{Z}_i \overline{\mu}_{\Psi_i})'(\overline{x}_i - \overline{Z}_i \overline{\mu}_{\Psi_i})$, and the posterior degrees of freedom are simply $\overline{v}_i = T + \underline{v}_i$.

Due to conjugacy, Ψ_i and Σ_i^{-1} can be simulated using a simple Gibbs sampler by drawing sequentially from the appropriate Normal and Wishart distributions. Unfortunately, the conditional posterior distributions of the remaining parameters are not available in closed form. Hence, for these parameters we have to use other posterior simulation strategies. For ψ_i, we depart from Brown et al. (1998b) and

---

[8]The first k_i rows and columns of $\underline{A}$ correspond to the prior means on the first lag of the endogenous variables, which is set equal to zero in the present application.

Koop (2013) and use a Markov Chain Monte Carlo Model Composition step in the spirit of Brown et al. (1998a). In short, this implies starting with the full model (i. e., all elements of ψ_i are set equal to one) and then using a simple birth-death sampler to explore the posterior distribution (Madigan and York, 1995). The history of log-volatilities and the corresponding parameters are sampled using an ancillarity sufficiency interweaving strategy proposed in Kastner and Frühwirth-Schnatter (2014).

This algorithm allows us to draw from the country-specific joint posterior distributions. However, for the subsequent forecasting exercise, we need estimates for the global posterior distribution, i. e., the posterior quantities for the parameters of equation (3.8), $p(F, \Sigma_{e,t}, h|\mathcal{D}_T)$, with $\mathcal{D}_T$ denoting the full data set available for all countries up to time T and h is the $(N+1) \times T$ matrix collecting all log-volatilities. Drawing from the individual country posterior distributions and using the algebra outlined above allows us to obtain valid draws from the posterior distribution of F and $\Sigma_{e,t}$. Knowing $p(F, \Sigma_{e,t}, h|\mathcal{D}_T)$, we can use Monte Carlo integration to obtain any quantity of interest like predictive densities or impulse response functions.

3.3 Benchmark Model

With a view to assessing the potential superiority of GVAR models over country-specific models, we select a set of country-specific Bayesian vector autoregression (BVAR) models with stochastic volatility as a benchmark.[9] We obtain this benchmark by 'shutting down' the interaction between variables from different countries in our baseline GVAR.

More formally, we restrict the impact of the foreign variables to zero in all equations of the model by setting Λ_{i0} and Λ_{i1} equal to zero matrices in equation (3.1). It follows that G has a block-diagonal structure such that also the covariance matrix

---

[9]We believe that this is a more natural benchmark than traditional univariate models (as, e. g., in Pesaran et al., 2009), which might be overly restrictive for macroeconomic variables such as growth rates or inflation.

$\Sigma_{e,t}$, given in equation (3.9) remains block-diagonal. In other words: the benchmark model neither allows for contemporaneous nor lagged spillover effects between any pair of two countries.

It is worth emphasizing that since the benchmark model is nested within the GVAR, prior choices are only of minor importance for the empirical application that follows.[10] A comparison based on "classical" estimated GVAR models is difficult since our model accounts for heteroscedasticity in a flexible manner.[11]

4 Data & Prior Implementation

4.1 Data Overview

We extend the data set used in previous studies (Dées et al., 2007a;b) with respect to country coverage and time span. More specifically, we use quarterly data for 36 countries (see Table 1) spanning the period from 1979q2 to 2013q4.

[Table 1 about here.]

The domestic variables that are covered in our analysis are real activity (GDP), the change of the consumer price level, real equity prices, the real exchange rate, and short- and long-term interest rates. Thus, our data set features the same variables as Dées et al. (2007a,b) and Pesaran et al. (2009). We follow the bulk of the literature by including oil prices as a global control variable and by using bilateral trade flows as the measure of economic proximity between countries. That is, we construct the weakly exogenous foreign variables based on the average bilateral annual trade flows in the period from 1980 to 2003, which denotes the end of our initial estimation sample.[12] Following Crespo Cuaresma et al. (2014a) and Feldkircher and Huber

---

[10]This holds true as long as we assume the same set of priors for both specifications.

[11]For a comparison between Bayesian GVARs and classical GVARs see Crespo Cuaresma et al. (2014b).

[12]Note that recent contributions (Dovern and van Roye, 2014; Eickmeier and Ng, 2015) suggest using financial data to compute foreign variables related to the financial side of the economy (e. g., interest rates or credit volumes). Since our data sample starts in the early 1980s, reliable data on financial flows—such as portfolio flows or foreign direct investment—are not available. See the

(2015), we calculate foreign counterparts for all variables including the bilateral real exchange rate vis-á-vis the U.S. dollar.[13]

The U.S. model deviates from the other country models in that the oil price is determined within that country-model. The dominant role of the U.S. economy for global financial markets is often accounted for by including only a limited set of weakly exogenous variables. It is not necessary to impose this ad-hoc restriction in our Bayesian approach since it entails variable selection at the country model level.

Finally we correct for outliers in countries that witnessed crisis-induced, extraordinarily strong movements in some of the variables contained in our data. We opted to smooth the relevant time series in these cases.[14] In particular, this applies to the short-term interest rate in Argentina during the periods 1983q4, 1984q3 to 1985q2, 1989q2, and 1990q1. The data are described in more detail in Table 2.[15]

[Table 2 about here.]

4.2 Prior Implementation

To implement the prior setup discussed above, we have to specify suitable hyperparameters for all priors. For the prior on Ψ_i, we scale the prior variances using the so-called "semi-automatic" approach put forward in George et al. (2008). This implies that the prior variances are scaled by the respective least squares variance of the parameter in question. However, due to the fact that we only have to specify K_i variance parameters, we follow Koop (2013) and use the maximum of the variance corresponding to each variable type across all equations.

---

appendix of Feldkircher and Huber (2015) for a sensitivity analysis with respect to the choice of weights.

[13]Carriero et al. (2009) show that controlling for co-movements of currencies is important to improve macroeconomic forecasts.

[14]Alternatively one could introduce a set of dummy variables for the observations in question as in Crespo Cuaresma et al. (2014a) or Feldkircher and Huber (2015). While this might improve the behavior of the country-models' residuals, it still poses the risk that outlier effects could be carried over to other country models via the trade-weighted foreign variables.

[15]With the exception of long-term interest rates, the cross-country coverage of all variables is above 90%. Long-term interest rates are hardly available for emerging markets that do not feature well developed capital markets.

The prior on the precision matrix, Σ_i^{-1}, is rendered fairly uninformative by setting $\underline{S}_i = I_{k_i}$ and $\underline{v}_i = k_i$. Robustness checks lead to the conclusion that this specific prior choice proves to be non-influential for the qualitative ordering of the models discussed.

For the autoregressive parameter in the log-volatility equation ρ_i, we set $a_1 = 25$ and $b_1 = 1.5$, resulting in a prior mean of 0.89 and standard deviation of 0.08. This puts most prior mass on positive values of ρ_i. This prior choice implements the view that the log-volatility follows a fairly persistent autoregressive process rather than a random walk. Due to the fact that our time series are rather short for a model that deals with stochastic volatility, the choice of the hyperparameter is rather crucial in empirical applications. However, robustness checks indicate that the qualitative ordering of the models remains unaffected by changes in the quantitative results. The prior on the level of the log-volatility is set to be uninformative given the scale of data. This implies that we set $\underline{\mu}_{\kappa_i} = 0$ and $\underline{V}_{\kappa_i} = 10$. Finally, for the Gamma prior of the variance of the log-volatility equation, we set $R_{\eta_i} = 1$. This choice is of minor influence in our application.

5 Design of Out-of-Sample Forecast Analysis

Given the posterior distribution of the model parameters, simulating predictive densities for different forecast horizons is straightforward. More formally, the r-steps-ahead predictive density for the model in (3.8) in period τ is given by

$$p(x_{\tau+r}|\mathcal{D}_\tau) = \int_\Xi p(x_{\tau+r}|\Xi, \mathcal{D}_\tau)p(\Xi|\mathcal{D}_\tau)d\Xi, \tag{5.1}$$

where $\Xi = \{F, \Sigma_{e,t}, h\}$ denotes all estimated parameters of the GVAR model. As mentioned above, estimates for (5.1) are readily available using a numerical approximation.

16

We propose the following recursive forecasting exercise. Beginning in $\tau = t_0$, we re-estimate the GVAR and simulate the predictive density $p(x_{\tau+r}|\mathcal{D}_\tau)$ for $r = 1, 4$ for each period. This procedure is repeated until $\tau = T - r$ is reached, producing a sequence of predictive densities for the verification period. Our initial estimation period ranges from 1979q2 to τ_0=2003q4. The hold-out sample consists of 40 observations, covering the time span from 2004q1 to T=2013q4.

To investigate the accuracy of predictive densities, researchers have opted for a plethora of different loss functions. In this paper, we focus on the log predictive score (LPS), which has been motivated, for instance, in Geweke and Amisano (2010). It has the advantage that it is widely used and that it is a so-called proper scoring rule (Gneiting, 2008).

More specifically, the log predictive score at time τ is the r-step-ahead predictive density conditional on the estimated parameters and past data and evaluated at the actual outcome for time $\tau + r$, $x_{\tau+r}^O$:

$$\text{LPS}(x_{\tau+r}^O|\mathcal{D}_\tau) = \log p(x_{\tau+r} = x_{\tau+r}^O|\mathcal{D}_\tau). \tag{5.2}$$

Unfortunately, $p(x_{\tau+r} = x_{\tau+r}^O|\mathcal{D}_\tau)$ does not have a closed form solution for $r > 1$ (Adolfson et al., 2007). We proceed like Adolfson et al. and assume that it can be approximated by a multivariate normal density whose parameters can be estimated from the predictive sample. The resulting approximation is given by

$$\widetilde{LPS}(x_{\tau+r}^O|\mathcal{D}_\tau) \approx - 0.5[k \log(2\pi) + \log|\overline{\Omega}_{\tau+r|\tau}|$$
$$+ (x_{\tau+r}^O - \overline{x}_{\tau+r|\tau})'\overline{\Omega}_{\tau+r|\tau}^{-1}(x_{\tau+r}^O - \overline{x}_{\tau+r|\tau})], \tag{5.3}$$

with $\overline{\Omega}_{\tau+r|\tau}$ denoting the posterior variance-covariance matrix and $\overline{x}_{\tau+r}$ the mean vector of the predictive density.

In our analysis, we also look at the forecast performance in terms of the predictive density for all variables of a specific type. To investigate the corresponding variable-

specific LPS we have to integrate out the effects of all other variables. Under the assumption of a multivariate normal predictive density, the marginal distribution of a subset of variables can be obtained by dropping the (ir-)relevant rows and columns of $\overline{\Omega}_{\tau+r|\tau}$, $\overline{x}_{\tau+r}$, and $x^O_{\tau+r}$. Thus, computation of the LPS corresponding to variables of type m also simply boils down to evaluating the density of a multivariate normal distribution. Formally, this implies replacing $\overline{x}_{\tau+r|\tau}, \overline{\Omega}_{\tau+r|\tau}$ and $x^O_{\tau+r}$ in equation (5.3) with $\overline{x}_{m,\tau+r|\tau}$, $\overline{\Omega}_{m,\tau+r|\tau}$ and $x^O_{m,\tau+r}$ where $x^O_{m,\tau+r}$ is a vector which consists exclusively of variables of type m. Likewise, $\overline{x}_{m,\tau+r|\tau}$ denotes the posterior mean of the predictive densities of variables of type m and $\overline{\Omega}_{m,\tau+r|\tau}$ denotes the posterior variance-covariance matrix corresponding to type-m variables. Note that by the same reasoning the variance and the mean of each of the k marginal predictive densities are given by the proper elements of the main diagonal of $\overline{\Omega}_{\tau+r|\tau}$ and $\overline{x}_{\tau+r}$, say $\left[\overline{\Omega}_{\tau+r|\tau}\right]_{ll}$ and $[\overline{x}_{\tau+r}]_l$ for $l = 1, 2, \ldots, k$.

To analyze the source of the overall forecast performance in more detail, we make use of the following result. It is straightforward to decompose the log score of the joint predictive density into the sum of the log scores of the predictive marginal (normal) densities and the log score of the predictive (Gaussian) copula density. The first term can serve as a measure of how well each of the variables is forecast integrating out all other variables.[16] The second term can serve as a measure of how well the dependence structure among all variables of the model is forecast. Formally, we have

$$\widetilde{LPS}_{\tau+r|\tau}(x^O_{\tau+r}|\mathcal{D}_\tau) = \sum_{l=1}^{k} \log p(x^O_{l,\tau+r}|\mathcal{D}_\tau)$$
$$+ \log c\left(P^{-1}(x^O_{1,\tau+r}|\mathcal{D}_\tau), \ldots, P^{-1}(x^O_{k,\tau+r}|\mathcal{D}_\tau)\right), \qquad (5.4)$$

where $p(x^O_{l,\tau+r}|\mathcal{D}_\tau)$ denotes the predictive marginal density of the lth variable evaluated at the final outcome and conditional on current information, $P(x^O_{l,\tau+r}|\mathcal{D}_\tau)$

---

[16]Using the unweighted sum over all marginal densities implies an equal weighting scheme across variables.

18

denotes the corresponding *cdf*, and $c(\cdot)$ is the probability density function of the Gaussian copula.

6 Empirical Results

6.1 Baseline Comparison: Average Forecast Accuracy

As stated in the previous section, we use the log scores of the joint predictive densities to compare the forecast performance of different models from a multivariate perspective. Table 3 shows these average joint log predictive scores for the GVAR forecasts and our benchmark forecasts based on country-specific stochastic volatility BVAR models. It is evident from the results that the GVAR model outperforms the BVAR models in terms of forecast performance for both forecast horizons. This holds true for the overall joint predictive density as well as for the joint predictive densities of each variable group conditional on the forecasts for the respective other variable groups.

[Table 3 about here.]

The average log score of the overall predictive densities of 1-step-ahead forecasts is 314.6 for the GVAR model compared to 271.6 for the BVAR models. Similarly, for 4-steps-ahead forecasts the corresponding average log score for the GVAR model (169.6) is much higher than that of the BVAR models (123.8).[17]

Where do these gains come from? A priori, there are two possibilities: First, it is possible that the GVAR performs better, on average, in terms of providing good forecasts for the predictive marginal densities for each of the variables. Second, performance gains could stem from more accurate predictions for the dependence structure among the variables in the system. These two effects can be analyzed

---

[17]Note that we cannot formally test the statistical significance of the difference in accuracy by using the test proposed by Amisano and Giacomini (2007), which only applies for a rolling sample scheme during the out-of-sample forecast analysis. In this paper we conduct the forecasting exercise based on an "expanding" window, which is more frequently used among practitioners.

based on the decomposition given in equation (5.4). A gain along the second dimension would not be very surprising because the collection of benchmark BVAR models does not model cross-country dependencies by assumption whereas the GVAR model is designed to capture such interactions. Looking at the second and third row of Table 3 reveals that the second effect does indeed drive the overall forecasting performance of the GVAR: the log score of the predictive copula density for the GVAR model is much higher for both forecast horizons than that of the BVAR model. Interestingly, the gains are much higher for the 4-steps-ahead forecasts (69.6) than for the short-term forecasts (42.5).[18] At the same time, the log scores for the predictive marginal densities indicate that they do not contribute to the superior overall forecast performance of the GVAR model. On the contrary: the GVAR performance for a forecast horizon of 4 quarters is much weaker along this dimension relative to the benchmark. Thus, the bottom line seems to be that the increased number of parameters in the GVAR model leads to a deterioration of forecast performance in terms of marginal predictive distributions which is outweighed by the gains that are realized in terms of better predictions of the cross-country dependencies.

Further results (shown in the lower part of Table 3) indicate that these performance gains are not driven by improved GVAR forecasts of some variable types. By contrast, the GVAR outperforms forecasts from local, country-specific VARs for all macro-variables considered in this paper. Focusing, for instance, on the log score of the joint predictive density of all GDP variables of the sample conditional on the forecasts for the other variables reveals that the GVAR model provides better forecasts for this group of variables for the short forecast horizon (61.9 vs 59.0) and for the longer forecast horizon (27.2 vs. 22.0). The difference seems to be especially pronounced when focusing on the conditional joint predictive densities of the short-term interest rate (i_S) and equity prices (eq) respectively.

---

[18] The difference is even more astonishing when comparing the differences of the predictive log copula scores to the levels of the overall predictive log scores, which are much smaller for the 4-steps-ahead forecasts.

Overall, the results point to strong evidence of the GVAR framework outperforming country-specific benchmark VAR forecasts by a great margin. More specifically, this result is driven by the better predictions of the cross-country dependencies between variables which is a salient feature of the GVAR framework.

6.2 Performance Variation over Time

There is empirical evidence that the predictive accuracy of forecast models fluctuates over time both in terms of absolute accuracy (which obviously varies because shocks of different sizes hit the economy at different times) and relative accuracy (when different models are compared).[19]

This is why we also investigate whether and, if so, how the relative forecast performance of the GVAR model varies over time. On the one hand it could be argued that the complexity of the GVAR model captures the dynamics that once in a while cause volatile business-cycle fluctuations – thus its relative forecast performance could be especially pronounced during turbulent times. On the other hand it is well known that simpler forecast models tend to outperform complexer models especially during volatile periods because they are able to adjust more quickly to structural change.

Figure 1 shows that the log predictive scores of both models indeed vary strongly over time, with large decreases of the log scores for both models standing out during the period of the Great Recession. Three other aspects are noteworthy. First, the joint log predictive score of the GVAR model is consistently higher than that of the BVAR models over all periods of the evaluation sample and for both forecast horizons. Second, the deterioration of the log predictive scores during the Great Recession is—not surprisingly—more moderate for the medium-term forecasts. Finally, the deterioration of the log predictive scores during the Great Recession is much more severe for the BVAR models than for the GVAR model. In other words,

---

[19]See, for instance, Del Negro et al. (2014) or Geweke and Amisano (2010).

the relative forecast performance of the GVAR model increased during the Great Recession. Two features are likely to contribute to this finding. First, the GVAR model reflects information from countries that fall into the recession earlier than others, thus anticipating the economic slowdown in the other countries. Second, the multiple channels of cross-country interaction built into the GVAR model allow anticipating the synchronicity of events during a global downturn.

[Figure 1 about here.]

A decomposition of the log score of the joint predictive density for each forecasting period as described in equation (5.4) can be interpreted as a confirmation of the two hypotheses. Figure 2 shows that two effects are at play which explain the good relative performance during the Great Recession. First, the decline of log scores for the marginal predictive densities is not as sharp in case of the GVAR model as it is for the BVAR benchmark. Second and very interestingly, the GVAR log scores for the predictive copula density actually *increase* during the Great Recession while they slightly decline for the BVAR benchmark. Thus, it seems that non-systematic, idiosyncratic shocks played a smaller role during the global downturn, which enhances the benefit of using a GVAR model that captures the *systematic* dependencies across countries.

[Figure 2 about here.]

An analysis of the different groups of variables reveals some interesting similarities and differences (Figure 3). First, the log predictive scores for the variable-type-specific conditional joint densities are higher for the GVAR than for the BVAR models in all periods for all types of variables.[20] Second, the forecast performance sharply falls during the Great Recession virtually across the board. Third, this fall in forecast performance is equal in magnitude for both types of models with two

---

[20]For the presentation of results we focus on the short-run forecasts. Results for the 4-steps-ahead forecasts are very similar and available upon request.

exceptions, namely the joint predictive distribution of short-term interest rates and of equity prices, which drops more sharply in the BVAR models than in the GVAR model. This indicates that the very large differences of average log predictive densities that were documented for short-term interest rates and equity prices in the previous section mainly reflect the strong forecasting record of the GVAR during times of economic crisis.

[Figure 3 about here.]

6.3 A Closer Look at Regions in the GVAR

Undoubtedly, the countries of our sample are very different with respect to how strongly they are integrated into the world economy and how strongly they drive the global business cycle.[21] Therefore, we want to analyze the gains in terms of forecast performance from using the GVAR model for forecasting different regions. To keep things simple and straightforward, we split the sample into a group of advanced economies (AE) and a group of emerging economies (EE).[22]

Specifically, we set all trade weights, w_{ij}, equal to 0 if country i is in a different country group than country j. All other trade weights are adjusted accordingly to sum to unity. In essence, this yields two separate regional GVAR models for advanced and emerging economies, respectively. These models take information about the interconnectedness of countries within each of the groups into account while neglecting any spillover effects from advanced to emerging economies, and vice versa. The analysis shows how relevant it is to use information from different country groups when the single aim is to forecast variables from a particular group of countries.

To measure this relevance, we report the difference of log predictive scores. Since at this point we focus on forecasting all variables from one particular country group,

---

[21]See, for instance, Kose et al. (2003) for an extensive exploration of the importance of the global business cycle for a wide range of countries.

[22]The group of AE comprises the 15 European countries and the 5 other developed economies. The remaining 16 countries form the group of EE.

we use the joint log predictive score for this subset of variables once the predictions for the variables of the other country group have been integrated out from the overall joint predictive density. We compute this measure for the newly constructed regional GVAR model and subtract the corresponding log score for the conventional GVAR model. Thus, negative values indicate an inferior forecast performance of the regional GVAR models while positive values indicate that moving to the more parsimonious model actually improves the quality of forecasts.

Figure 4 shows the differences of log scores for both country groups and the two forecast horizons. Evidently, neither the regional GVAR models nor the GVAR model dominates the other model over all periods. Both country groups exhibit periods with negative and positive log score differences. It is striking that the gains from using a truly global model increases especially during the time of the Great Recession.[23]

[Figure 4 about here.]

Furthermore, looking at the average over time reveals that taking spillover effects across country groups into account is more important for business-cycle forecasts for emerging economies than for advanced economies. For the 1-step-ahead forecasts, the mean difference is -16.8 for advanced economies but -67.0 for emerging economies, i.e., switching off transmission channels between advanced and emerging economies leads to a much larger deterioration of short-term forecast performance for emerging economies relative to advanced economies. Likewise for 4-steps-ahead forecasts: the average forecast performance for advanced economies increases when moving from the GVAR to the regional GVAR model (mean log score difference of 26.7) while it decreases slightly for emerging economies (-18.9).

---

[23]This results is similar to the finding in section 6.2, which showed that the GVAR model outperformed the country-specific BVAR models especially during the Great Recession.

7 Conclusion

In this paper, we analyze the ability of the GVAR model to generate accurate joint predictive densities for all variables of the system. To this end, we estimate a Bayesian GVAR model with SSVS priors and stochastic volatility and perform an out-of-sample forecast analysis. Our GVAR specification allows for a computationally efficient estimation of the model—which is essential when dealing with such high dimensional multi-country models in a context that requires repeated estimation.

Our results suggest that GVAR forecasts consistently improve the joint log predictive score compared to benchmark forecasts that are based on country-specific stochastic volatility BVAR models. This result holds true for the entire evaluation period of our forecasting sample but especially so for the Great Recession when the GVAR outperforms the benchmark by a great margin. Thus, it seems that modelling cross-country linkages is particularly important during times of global economic turbulences. By using a decomposition of the joint predictive density into the marginal densities and a copula term, we show that the performance gains of the GVAR are exclusively driven by better predictions of the cross-country dependencies. This decomposition approach might be of independent interest also in other forecast evaluation setups.

We proceed by estimating two regional GVAR models for advanced and emerging economies, respectively. A forecast analysis based on these models indicates no deterioration of forecasts for advanced economies when emerging economies are excluded in the regional GVAR. By contrast, shutting down spillover effects from advanced to emerging economies significantly deteriorates the quality of forecasts for emerging economies. This implies that while advanced economies could be modeled more parsimoniously taking into account only within-region spillover effects, forecasts for emerging economies can be significantly improved by considering information from advanced economies. However, for both regions the truly global GVAR framework featuring all economies in the sample excels in terms of forecast accuracy during the

Great Recession 2008/09. Thus, our results only partially confirm the evidence in Kose et al. (2012) who find no strong links between the business cycles of advanced and emerging economies, respectively.

Our findings suggest that there are clear gains from using GVAR models when the aim is to coherently forecast the economies of a range of countries. Modeling cross-country linkages seems to pay off relative to the use of country-specific benchmark models. Our results concerning the minor contribution of modeling emerging economies when forecasting advanced economies suggest that further gains in forecast performance could be achieved by properly trimming the bilateral weights that are used to link different countries. However, our results also suggest that this applies mostly during tranquil times and that a parsimonious GVAR might do worse when the world economy is hit by a truly global shock such as the Great Recession, which is in line with the evidence in Crespo Cuaresma et al. (2014a). We leave these issues for future research.

References

Adolfson, Malin, Jesper Lindé, and Mattias Villani, "Forecasting performance of an open economy DSGE model," *Econometric Reviews*, 2007, *26* (2-4), 289–328.

Amisano, Gianni and Raffaella Giacomini, "Comparing Density Forecasts via Weighted Likelihood Ratio Tests," *Journal of Business & Economic Statistics*, 2007, *25* (2), 177–190.

Assenmacher, Katrin, "Forecasting the Swiss Economy with a Small GVAR Model," in Filippo di Mauro and M. Hashem Pesaran, eds., *The GVAR Handbook*, Oxford University Press, 2013, pp. 244–254.

Bańbura, Marta, Domenico Giannone, and Lucrezia Reichlin, "Large Bayesian vector auto regressions," *Journal of Applied Econometrics*, 2010, *25* (1), 71–92.

Bauer, Andy, Robert A. Eisenbeis, Daniel F. Waggoner, and Tao Zha, "Forecast evaluation with cross-sectional data: The Blue Chip Surveys," *Economic Review*, 2003, (Q2), 17–31.

Brown, Philip J., Marina Vannucci, and Tom Fearn, "Bayesian wavelength selection in multicomponent analysis," *Journal of Chemometrics*, 1998, *12* (3), 173–182.

__ , __ , and __ , "Multivariate Bayesian variable selection and prediction," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 1998, *60* (3), 627–641.

Carriero, Andrea, George Kapetanios, and Massimiliano Marcellino, "Forecasting exchange rates with a large Bayesian VAR," *International Journal of Forecasting*, 2009, *25* (2), 400–417.

__ , **Todd E. Clark, and Massimiliano Marcellino**, "Common drifting volatility in large Bayesian VARs," 2012.

Cashin, Paul, Kamiar Mohaddes, Maziar Raissi, and Mehdi Raissi, "The

differential effects of oil demand and supply shocks on the global economy," *Energy Economics*, 2014, *44* (C), 113–134.

Castrén, Olli, Stéphane Dées, and Fadi Zaher, "Stress-testing euro area corporate default probabilities using a global macroeconomic model," *Journal of Financial Stability*, 2010, *6* (2), 64–78.

Cesa-Bianchi, Ambrogio, "Housing Cycles and Macroeconomic Fluctuations: A Global Perspective," *Journal of International Money and Finance*, October 2013 2013, *37*, 215–238.

Chudik, Alexander and M. Hashem Pesaran, "Theory and Practice of GVAR Modeling," CESifo Working Paper Series 4807, CESifo Group Munich 2014.

__ **and Marcel Fratzscher**, "Identifying the global transmission of the 2007-2009 financial crisis in a GVAR model," *European Economic Review*, April 2011, *55* (3), 325–339.

Clark, Todd E., "Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility," *Journal of Business & Economic Statistics*, 2011, *29* (3).

Crespo Cuaresma, Jesús, Martin Feldkircher, and Florian Huber, "Forecasting with Bayesian Global Vector Autoregressive Models: A Comparison of Priors," Working Papers 189, Oesterreichische Nationalbank (Austrian Central Bank) 2014.

__ , __ , **and** __ , "Forecasting with Bayesian Global Vector Autoregressive Models: A Comparison of Priors," Working Papers 189, Oesterreichische Nationalbank (Austrian Central Bank) March 2014.

de Waal, Annari, Renee van Eyden, and Rangan Gupta, "Do we need a global VAR model to forecast inflation and output in South Africa?," Working Papers 2013-46, University of Pretoria, Department of Economics 2013.

Dées, Stéphane, Filippo di Mauro, M. Hashem Pesaran, and L. Vanessa Smith, "Exploring the international linkages of the euro area: a global VAR

analysis," *Journal of Applied Econometrics*, 2007, *22* (1).

— , **Sean Holly, M. Hashem Pesaran, and L. Vanessa Smith**, "Long Run Macroeconomic Relations in the Global Economy," *Economics - The Open-Access, Open-Assessment E-Journal*, 2007, *1* (3), 1–20.

Del Negro, Marco, Raiden B. Hasegawa, and Frank Schorfheide, "Dynamic Prediction Pools: an Investigation of Financial Frictions and Forecasting Performance," Unpublished manuscript 2014.

Diebold, Francis X., Jinyong Y. Hahn, and Anthony S. Tay, "Multivariate density forecast evaluation and calibration in financial risk management: High frequency returns on foreign exchange," *Review of Economics & Statistics*, 1999, *81* (4), 661–673.

Doan, Thomas, Robert Litterman, and Christopher A. Sims, "Forecasting and conditional projection using realistic prior distributions," *Econometric Reviews*, 1984, *3* (1), 1–100.

Dovern, Jonas and Björn van Roye, "International Transmission and Business-Cycle Effects of Financial Stress," *Journal of Financial Stability*, 2014, *13* (1), 1–17.

— **and Florian Huber**, "Global Prediction of Recessions," unpublished manuscript, Oesterreichische Nationalbank 2015.

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

Eisenbeis, Robert, Daniel Waggoner, and Tao Zha, "Evaluating Wall Street Journal survey forecasters: a multivariate approach," Working Paper 2002-8, Federal Reserve Bank of Atlanta 2002.

Eisenstat, Eric and Rodney W. Strachan, "Modelling Inflation Volatility," CAMA Working Papers 2014-68, Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, The Australian National University 2014.

Feldkircher, Martin, "A global macro model for emerging Europe," *Journal of*

Comparative Economics, 2015, *forthcoming*.

__ **and Florian Huber**, "The international transmission of US shocks – Evidence from Bayesian global vector autoregressions," *European Economic Review*, 2015, *forthcoming*.

Frühwirth-Schnatter, Sylvia and Helga Wagner, "Stochastic model specification search for Gaussian and partial non-Gaussian state space models," *Journal of Econometrics*, 2010, *154* (1), 85–100.

Galesi, Alessandro and Marco J. Lombardi, "External shocks and international inflation linkages," in Filippo di Mauro and M. Hashem Pesaran, eds., *The GVAR Handbook: Structure and Applications of a Macro Model of the Global Economy for Policy Analysis*, Oxford University Press, 2013, pp. 255–270.

George, Edward I., Dongchu Sun, and Shawn Ni, "Bayesian stochastic search for VAR model restrictions," *Journal of Econometrics*, 2008, *142* (1), 553–580.

Geweke, John and Gianni Amisano, "Comparing and evaluating Bayesian predictive distributions of asset returns," *International Journal of Forecasting*, 2010, *26* (2), 216–230.

Gneiting, Tilmann, "Editorial: Probabilistic forecasting," *Journal of the Royal Statistical Society Series A*, 2008, *171* (2), 319–321.

Greenwood-Nimmo, Matthew, Viet Hoang Nguyen, and Yongcheol Shin, "Probabilistic forecasting of output growth, inflation and the balance of trade in a GVAR framework," *Journal of Applied Econometrics*, 2012, *27* (4), 554–573.

Huber, Florian, "Density Forecasting using Bayesian Global Vector Autoregressions with Common Stochastic Volatility," Department of Economics Working Paper Series 179, WU Vienna University of Economics and Business August 2014.

Independent Evaluation Office of the International Monetary Fund (IEO), "IMF Forecasts: Process, Quality, and Country Perspectives," Evaluation Report, Washington, D.C. February 12, 2014.

Kastner, Gregor and Sylvia Frühwirth-Schnatter, "Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models," *Computational Statistics & Data Analysis*, 2014, *76*, 408–423.

Koop, Gary M., "Forecasting with medium and large Bayesian VARs," *Journal of Applied Econometrics*, 2013, *28* (2), 177–203.

Kose, M. Ayhan, Christopher Otrok, and Charles H. Whiteman, "International Business Cycles: World, Region, and Country-Specific Factors," *American Economic Review*, 2003, *93* (4), 1216–1239.

__ , __ , **and Eswar Prasad**, "Global Business Cycles: Convergence Or Decoupling?," *International Economic Review*, 2012, *53* (2), 511–538.

Madigan, David and Jeremy York, "Bayesian graphical models for discrete data," *International Statistical Review*, 1995, *63*, 215–232.

Müller-Dröge, Hans Christian, Tara M. Sinclair, and Herman O. Stekler, "Evaluating Forecasts of a Vector of Variables. A German Forecasting Competition," Working Paper, George Washington University 2014.

Pesaran, M. Hashem, Schuermann Til, and Scott M. Weiner, "Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model," *Journal of Business & Economic Statistics*, 2004, *22* (2), 129–162.

__ , **Til Schuermann, and L. Vanessa Smith**, "Forecasting economic and financial variables with global VARs," *International Journal of Forecasting*, 2009, *25* (4), 642–675.

Sinclair, Tara M. and Herman O. Stekler, "Examining the quality of early GDP component estimates," *International Journal of Forecasting*, 2013, *29* (4), 736–750.

Smith, L. Vanessa, "Short- and Medium-Term Forecasting Using 'Pooling' Techniques," in Filippo di Mauro and M. Hashem Pesaran, eds., *The GVAR Handbook*, Oxford University Press, 2013, pp. 114–130.

Swanson, Norman R., "Comments on "Forecasting economic and financial vari-

ables with global VARs"," *International Journal of Forecasting*, 2009, *25* (4), 697–702.

Table 1: Country coverage of GVAR model

Europe	Other Developed	Emerging Asia	Latin America	Mid-East and Africa
Austria (AT)	Australia (AU)	China (CH)	Argentina (AR)	Turkey (TR)
Belgium (BE)	Canada (CA)	India (IN)	Brazil (BR)	Saudi Arabia (SA)
Germany (DE)	Japan (JP)	Indonesia (ID)	Chile (CL)	South Africa (ZA)
Spain (ES)	New Zealand (NZ)	Malaysia (MY)	Mexico (MX)	
Finland (FI)	United States (US)	Korea (KR)	Peru (PE)	
France (FR)		Philippines (PH)		
Greece (GR)		Singapore (SG)		
Italy (IT)		Thailand (TH)		
Netherlands (NL)				
Portugal (PT)				
Denmark (DK)				
Great Britain (GB)				
Switzerland (CH)				
Norway (NO)				
Sweden (SE)				

Notes: ISO-2 country codes in brackets.

Table 2: Data description

| Variable	Description	Min.	Mean	Max.	Coverage
y	Real GDP, average of 2005=100. Seasonally adjusted, in logarithms.	2.173	4.298	5.400	100%
Δp	Consumer price inflation in q-o-q terms. CPI seasonally adjusted.	-0.157	0.027	1.790	100%
e	Nominal exchange rate vis-à-vis the U.S. dollar, deflated by national price levels (CPI).	-7.591	-2.818	5.459	97.2%
eq	Equity prices deflated by CPI.Seasonally adjusted, in logarithms.	-6.927	3.698	5.857	100%
i_S	Typically 3-months-market rates, rates per annum.	-1.006	0.190	52.360	94.4%
i_L	Typically government bond yields, rates per annum.	0.000	0.077	0.306	61.1%
$poil$	Price of oil, seasonally adjusted, in logarithms.	-	-	-	-
Trade flows	Bilateral data on exports and imports of goods and services, annual data.	-	-	-	-

Notes: Summary statistics pooled over countries and time. The coverage refers to the cross-country availability per country, in %. Data are from the IMF's IFS data base and national sources. Trade flows stem from the IMF's DOTS data base. For more details see the data appendix in Feldkircher (2015).

Table 3: Average joint log predictive scores

	1-step-ahead		4-steps-ahead	
	GVAR	BVAR	GVAR	BVAR
Overall	314.576	271.608	169.622	123.758
Marginals	270.791	270.326	149.596	173.356
Copula	43.785	1.281	20.025	-49.598
y	61.945	58.976	27.176	22.002
Δp	64.888	61.271	35.159	32.319
i_S	50.680	41.762	17.770	10.323
i_L	59.291	56.768	26.967	24.125
e	39.606	36.066	18.491	13.857
eq	27.948	15.465	-2.257	-14.309

Notes: Average log predictive scores over the hold-out sample. Joint and variable-specific log scores have been obtained by using the multivariate normal approximation to the predictive density as in Adolfson et al. (2007). Variable-specific log scores are computed by integrating out the effects of other variables. The scores labeled "*Marginals*" and "*Copula*" refer to the decomposition of the overall log predictive score presented in equation (5.4)
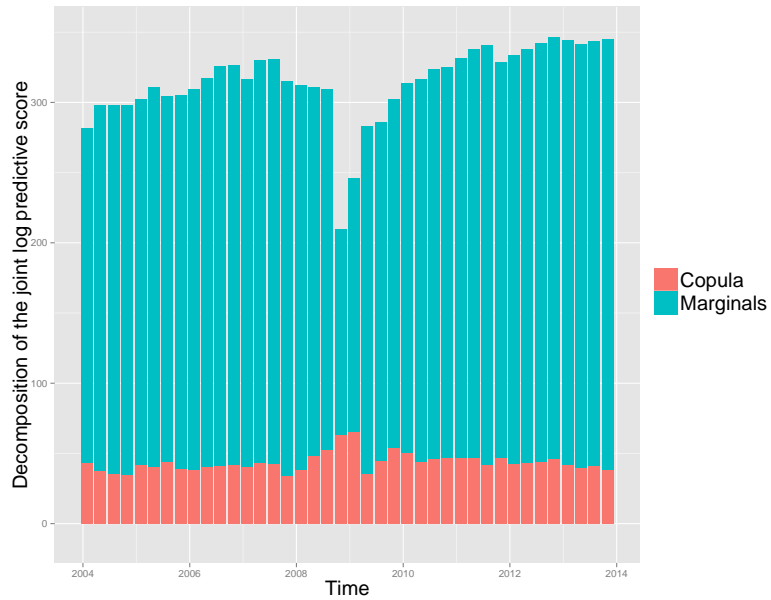
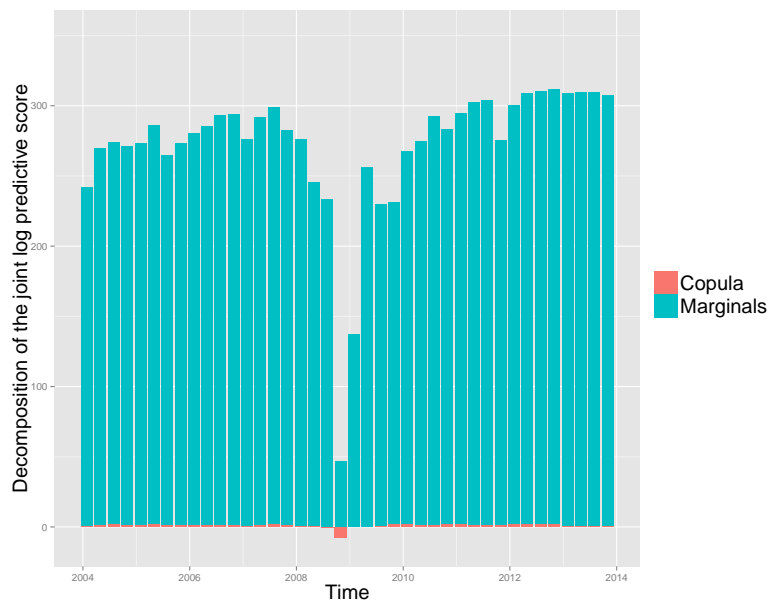**Figure 1:** Evolution of the joint log predictive score over time

(a) 1-step

(b) 4-steps
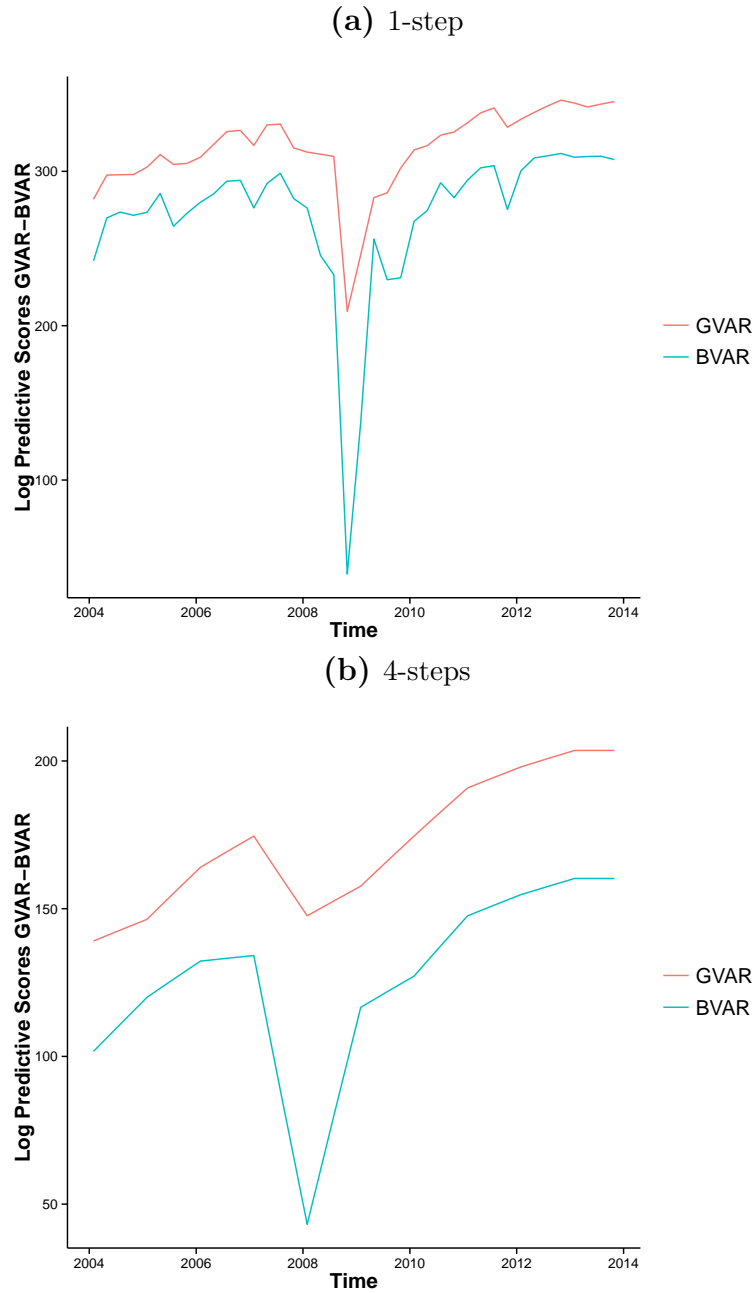
Notes: Evolution of the joint (for all variables and countries in the sytem) log predictive score for the hold-out sample. Results are based on the quadratic approximation to the log predictive score.

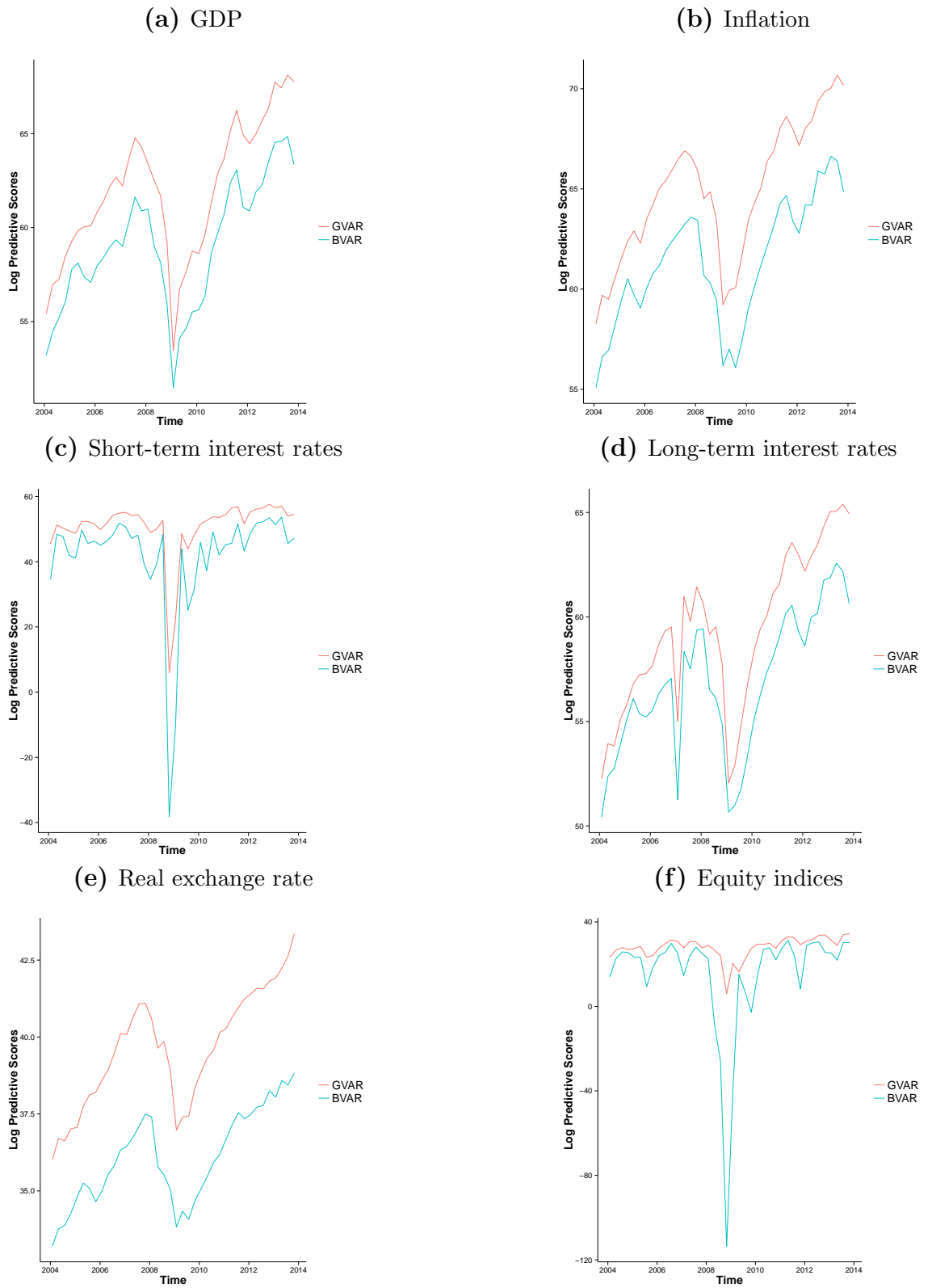**Figure 2:** Components of the joint log predictive score over time

(a) GVAR

(b) BVAR

Notes: Evolution of both components of the (1-step-ahead) joint log predictive score (given in equation (5.4)) over the hold-out sample. Results are based on the quadratic approximation.

Figure 3: Variable-specific log predictive scores over time

(a) GDP

(b) Inflation

(c) Short-term interest rates

(d) Long-term interest rates

(e) Real exchange rate

(f) Equity indices

Notes: Evolution of the variable-specific (1-step-ahead) log predictive score for the hold-out sample. Results are based on the quadratic approximation to the log predictive score after integrating out the effects of other variables.

38

Figure 4: Performance of regional GVAR models

(a) 1-step

(b) 4-steps
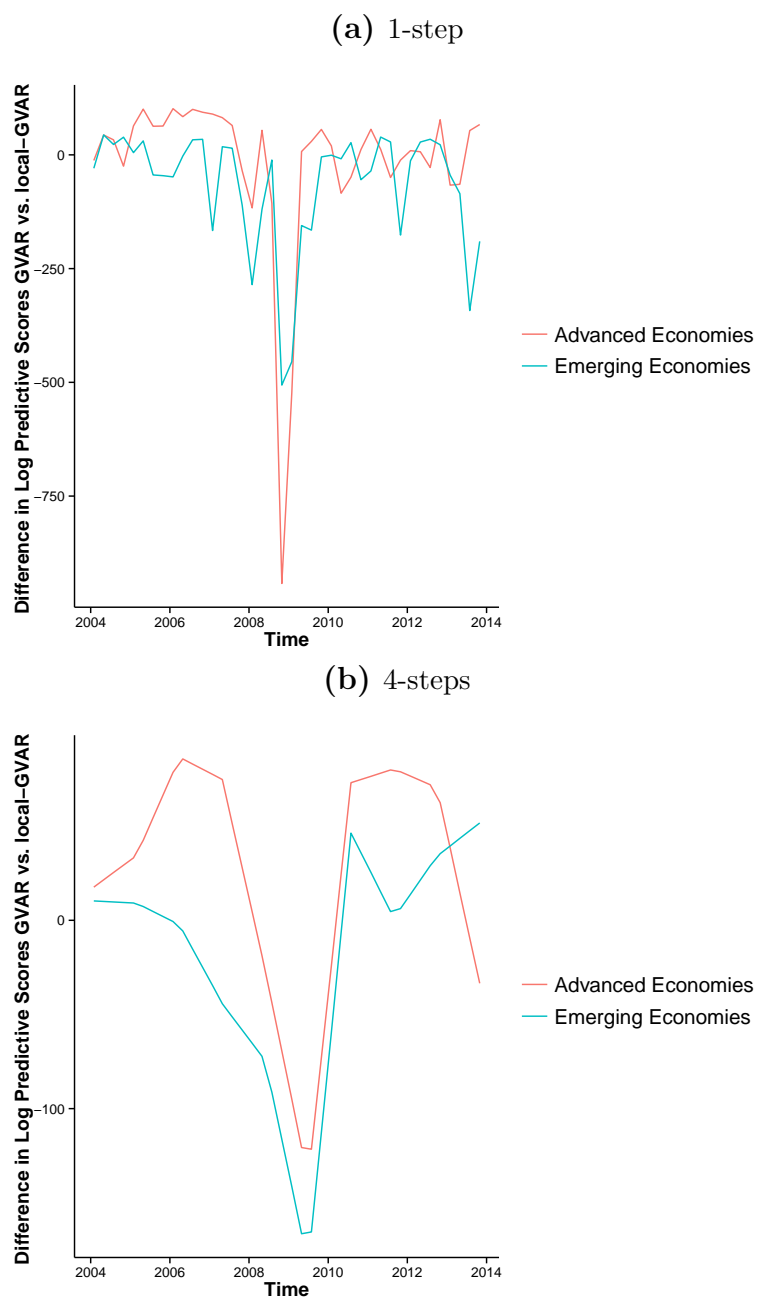
Notes: The figure shows the difference between the region-specific joint log predictive score for the regional GVAR models and the corresponding log score for the benchmark truly global GVAR. Region-specific joint log predictive scores are obtained by integrating out the effects of the other region. Positive values indicate that the regional GVAR models outperform the benchmark GVAR model.