

Forecasting with Bayesian Global Vector Autoregressive Models

A Comparison of Priors

Jesús Crespo Cuaresma

WU Vienna

Martin Feldkircher

OeNB

Florian Huber

WU Vienna

8th ECB Workshop on Forecasting Techniques

ECB

13-14 June

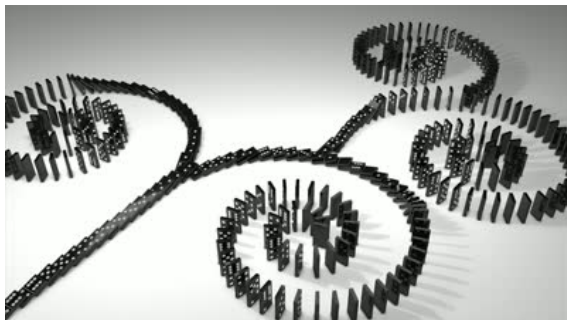
Overview

Research Objectives

- *Propose a framework that is flexible and can be used to do forecasting or structural analysis*
 - ▶ *Flexible prior distributions account for the heterogeneity observed in the world economy*
 - ▶ *Hierarchical modelling framework helps to automatically select appropriate country-specific models*
- *Do Bayesian methods help to cure the curse of dimensionality and increase the forecasting performance?*

This is achieved by combining the literature on Bayesian VARs with the literature on global VARs

The Global Vector Autoregressive model



- The GVAR model is a compact representation of the world economy
- Estimation of an encompassing VAR is not feasible \rightarrow even for moderate N
- GVAR short-cut: Estimate N country models in first stage, then stack them together to yield a global model.

Global VARs in a nutshell

- For each country $i \in 1, \dots, N$, a VARX* model is estimated:

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

where $x_{it}^* := \sum_{j \neq i}^N \omega_{ij} x_{jt}$ and $\varepsilon_{it} \sim \mathcal{N}(0, \Sigma_i)$

- After some straightforward algebra it is possible to rewrite the GVAR in a standard VAR form

$$x_t = b_0 + b_1 t + F x_{t-1} + \Gamma_0 d_t + \Gamma_1 d_{t-1} + e_t,$$

$x_t = (x_{0,t}, x_{1,t}, \dots, x_{N,t})$ denotes the global vector and b_0, b_1, Γ_1 contain the corresponding stacked vectors containing the parameter vectors of the country-specific specifications

Prior Specification

- For prior implementation, convenient to work with the parameter vector $\Psi_i = (a'_{i0} \ a'_{i1} \ \text{vec}(\Phi_i)' \ \text{vec}(\Lambda_{i0})' \ \text{vec}(\Lambda_{i1})' \ \text{vec}(\pi_{i0})' \ \text{vec}(\pi_{i1})')'$
- We assume the following conjugate prior setup on the coefficients of the local models:

Prior Setup

$$\begin{aligned}\Psi_i | \Sigma_i^{-1} &\sim \mathcal{N}(\underline{\mu}_\Psi, \underline{V}_\Psi) \\ \Sigma_i^{-1} &\sim \mathcal{W}(\underline{v}, \underline{S})\end{aligned}$$

- Several choices for $\underline{\mu}_\Psi$ and \underline{V}_Ψ possible
- Several Choices implemented: Minnesota prior, Single Unit Root prior, simpler variants
- Note that the prior dependence of Ψ_i on Σ_i^{-1} could be dropped
- Finally, adding additional hierarchy leads to the Stochastic Search Variable Selection Prior

Stochastic Search Variable Selection Prior

Impose a mixture prior on the coefficients:

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

where $\delta_{i,j}$ is a dummy random variable which corresponds to coefficient j in country i . $\tau_{1j}^2 \gg \tau_{0j}^2$ implies that if $\delta_{i,j} = 0$, the prior for $\Psi_{i,j}$ is centered around zero

- Estimation of the model using this prior setups requires MCMC \rightarrow Gibbs sampling with data augmentation
- The SSVS prior tackles model uncertainty, which implies that coefficients on "unimportant" variables are shrunk towards zero
- This prior facilitates different individual country model structures \rightarrow perfect for the present application

Computation of the Global Predictive Density

- Define $\Xi := (b'_0, b'_1, \text{vec}(F)', \text{vec}(\Gamma_0), \text{vec}(\Gamma_1))'$ and let $p(\Omega|\mathcal{D})$ be the posterior of the global VC-matrix
- Interest centers on $p(\Xi|\mathcal{D})$, not on the country individual $p(\Psi_i|\mathcal{D})$
- Interesting problem: How to estimate the global variance - covariance matrix Ω ?
- Using the usual GVAR algebra, it is possible to transform draws from $p(\Psi_i|\mathcal{D})$ for all countries to get a valid draw from $p(\Xi|\mathcal{D})$
- The predictive density is given by

$$p(x_{T+n}|\mathcal{D}) = \int \int p(x_{T+n}|\mathcal{D}, \Xi, \Omega) p(\Xi, \Omega|\mathcal{D}) d\Xi d\Omega$$

Data & Forecasting Design

Variable	Description, country coverage in brackets
y	real GDP (100%)
Δp	CPI inflation (100%)
e	Nominal exchange rate vis-à-vis the US dollar, deflated by national price levels (CPI) (98.10%)
i_S	Short-term interest rates (90.40%)
i_L	Long-term interest rates (32.70%)
p_{oil}	Price of oil (100%)

- *Weights* to construct foreign variables (and stack the model) based on average trade flows ($\omega_{ij,t}$)
- Together 45 countries + EA (in nominal terms, 92% of global output)
- Recursive one-quarter-ahead and one-year-ahead (four-quarters-ahead) predictions obtained by reestimating the models over a rolling window
- The initial estimation period ranges from 1995Q1 to 2009Q4, we use the period 2010Q1-2012Q4 as out-of-sample hold-out observations for the comparison of predictive ability across specifications
- Forecast comparison exercise based on root mean square error (RMSE) and the continuous rank probability score (CRPS)

Results: One Quarter Ahead

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 -

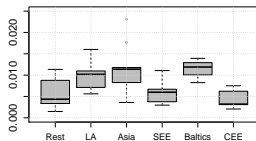
Results: One Year Ahead

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

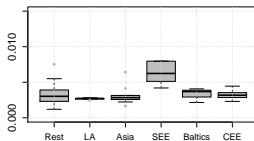
	NC	M	SUR	IW	M-IW	SSVS	Diffuse	Pesaran	AR
y	1.0090 (0.0399)	0.4857 (0.0187)	0.9999 (0.0229)	0.8847 (0.0423)	0.5799 (0.0360)	0.3651 (0.0148)	0.7192 (0.0224)	0.9332 -	0.5198 -
Δp	0.9204 (0.0356)	0.8538 (0.0084)	1.2064 (0.0178)	0.9126 (0.0393)	0.9172 (0.0346)	0.7442 (0.0081)	1.2205 (0.0118)	1.1408 -	1.0895 -
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 -

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

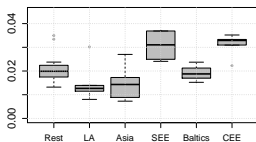
Real GDP



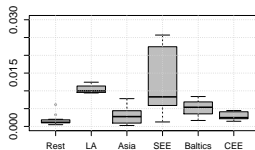
Inflation



Real Exchange Rate



Short-term interest rates



What about time varying Σ_i ?

- To allow for time changing variance covariance structures at the country level we have two possibilities:
 - 1 Follow Primiceri (2005) and allow the variances and covariances in the different equations to change over time
 - 2 Take the simpler route and follow Clark, Carriero and Marcellino (2012) and use a scalar factor to drive the volatility for our macro aggregates
- From now on we assume that the individual country variance-covariance structure is time changing

$$\begin{aligned}\Sigma_{i,t} &= \exp(h_{i,t}/2) \times \Sigma_i \\ h_{i,t} &= \eta_i + \rho_i(h_{i,t-1} - \eta_i) + \sigma_i e_{i,t} \\ e_{i,t} &\sim \mathcal{N}(0, 1)\end{aligned}$$

- Rescaling all observations by $\exp(h_{i,t}/2)$ allows us to use the same computations as in the standard case
- Sampling the latent log volatilities is done following Kastner & Fruehwirth-Schnatter (2013): Use **Ancillarity-sufficiency interweaving strategy (ASIS)**

Prior Setup for the SV Case

Prior Setup

$$\Psi_i | \Sigma_i \sim \mathcal{N}(\underline{\mu}_\Psi, \Sigma_i \otimes \underline{V}_\Psi)$$

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

$$\eta_i \sim \mathcal{N}(\underline{\mu}_\eta, \underline{V}_\eta)$$

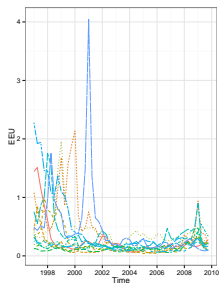
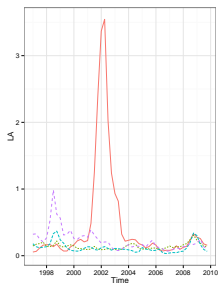
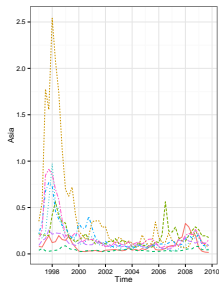
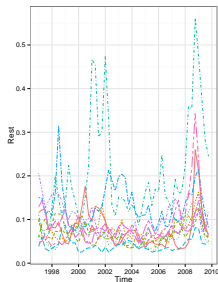
$$\frac{\rho_i + 1}{2} \sim \mathcal{B}(a_0, b_0)$$

$$\sigma_i \sim \mathcal{G}(1/2, 1/2B_\sigma)$$

- \underline{V}_Ψ is set to implement a symmetric Minnesota prior specification
- Choice of hyperparameters a_0 and b_0 for ρ_i can be quite influential, especially in our case
- We set $a_0 = 5$ and $b_0 = 1.5$, which translates into a prior mean of 0.54 and a prior standard deviation of 0.31 for ρ_i
- For all other components the hyperparameters are set such that the prior is effectively rendered non-influential

Filtered Volatilities for the World Economy

Posterior mean of $\exp(h_{i,t}/2)$



Results: One Quarter Ahead

Forecasting Performance, One-Quarter-Ahead: Log Predictive Score

	Diffuse	M	NC	SUR	SSVS	CSV
y	18.8837	29.0401	10.3133	18.2054	29.5344	35.6668
Δp	24.1329	37.8102	10.7397	21.9634	31.8327	39.8977
e	-3.8832	16.2315	8.3951	-3.8468	13.7503	11.3237
i_S	17.3429	34.1630	9.7453	15.5587	26.8580	38.2122
i_L	28.6069	42.3675	10.5221	28.5819	29.9277	42.0357

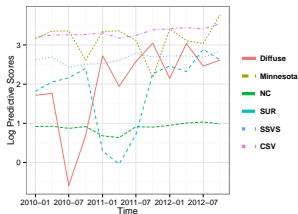
- Large outperformance of CSV specification for GDP, Inflation and short-term interest rates

Evolution of Log Predictive Scores over time - 1-step ahead predictive density.

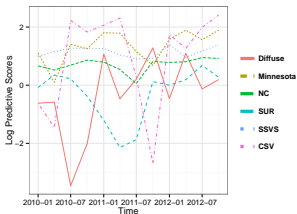
(a) Real GDP



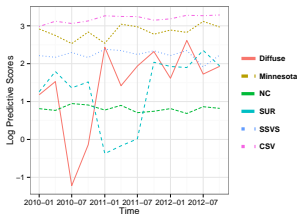
(b) Inflation



(c) Real Exchange Rate



(d) Short-term interest rates



Conclusions & Further Remarks

- Strong performance of Bayesian GVARs, both in terms of precise point forecasts and density forecasts
- SSVS and the simple Minnesota prior show the strongest performance in terms of density and point forecasting
- Performance of natural conjugate prior specifications could be increased by using hyperpriors (Giannone, Lenza & Primiceri, 2012). However, this would imply giving up flexibility as compared to the SSVS prior
- Preliminary results of a B-GVAR with SV indicate strong performance in terms of density forecasting (at little additional cost in terms of computing)

Selected Literature



Neil R. Ericsson & Erica L. Reisman. 2012.

Evaluating a Global Vector Autoregression for Forecasting
International Advances in Economic Research, Vol. 18, 247-258



Robert B. Litterman. 2009.

Forecasting with Bayesian Vector Autoregressions: Five Years of Experience
Journal of Business & Economic Statistics , Vol. 4, 25-38



M. Hashem Pesaran, Til Schuermann & L. Vanessa Smith. 2009.

Forecasting economic and financial variables with global VARs
International Journal of Forecasting, Vol. 25, 642-675



M. Hashem Pesaran, Til Schuermann & Scott M. Weiner. 2004.

Modelling Interdependencies Using a Global Error-Correcting Macroeconometric Model
Journal of Business & Economic Statistics, Vol. 22, 129-162.