BGVAR: Bayesian Global Vector Autoregressions with shrinkage priors in **R**

Maximilian Böck
Vienna School of
International Studies
Vienna University of
Economics and Business

Martin Feldkircher
Vienna School of
International Studies
Oesterreichische Nationalbank

Florian Huber
University of Salzburg

Abstract

This document introduces the R package **BGVAR** to estimate Bayesian global vector autoregressions (GVAR) with shrinkage priors and stochastic volatility. The Bayesian treatment of GVARs allows to include large information sets by mitigating issues related to overfitting. This often improves inference as well as out-of-sample forecasts. Computational efficiency is achieved by using C++ to considerably speed up time-consuming functions. To maximize usability, the package includes numerous functions for carrying out structural inference and forecasting. These include generalized and structural impulse response functions, forecast error variance, and historical decompositions as well as conditional forecasts.

Keywords: Global Vector Autoregressions, Bayesian Inference, Time Series Analysis, R.

1. Introduction

Today's world is characterized by high degrees of trade and financial integration: Firms diversify production and investment across domestic and foreign locations, capital flows are unrestricted, and business cycles strongly correlated. As a consequence, econometric models that treat countries as isolated entities may miss important information for forecasting and counterfactual analysis. Such concerns do not arise with global vector autoregressive (GVAR) models since these estimate the intensity of cross-unit correlation using a global system of equations approach. This for estimating the effects of a local shock to the remaining units in the system. In fact, this kind of spillover analysis is the predominant application of the GVAR framework (Chudik and Pesaran 2016).

The GVAR framework was originally proposed in Pesaran, Schuermann, and Weiner (2004) and consists of two steps: In the first step, small-scale unit-specific models are estimated conditional on the remaining units in the sample. These units can be countries, regions, companies or individuals. In the second step, the estimated parameters of these local models are stacked and combined to yield a global representation of the whole system. To this end, some exogenous measure of connectivity of the cross-sectional units is needed. The global representation can then be used to study the propagation and dynamics of an exogenous shock to the system. Examples cover applications to monetary policy (e.g., Burriel and

Galesi 2018; Benecká, Fadejeva, and Feldkircher 2020; Böck, Feldkircher, and Siklos 2021; Colabella 2021; Georgiadis 2015; Hájek and Horváth 2018), credit shocks (e.g., Eickmeier and Ng 2015; Fadejeva, Feldkircher, and Reininger 2017), oil price shocks (e.g., Cashin, Mohaddes, Raissi, and Raissi 2014), and house price shocks (Cesa-Bianchi 2013). Other applications feature conditional forecasts / scenario analysis (e.g., Pesaran *et al.* 2004; Greenwood-Nimmo, Nguyen, and Shin 2012), and unconditional forecasting (Crespo Cuaresma, Feldkircher, and Huber 2016; Dovern, Feldkircher, and Huber 2016; Chudik, Grossman, and Pesaran 2016; Pesaran, Schuermann, and Smith 2009).[1]

One advantage of the GVAR approach is that it is not necessarily restricted to a single type of unit (such as countries) but it can be used to model the interaction between different layers of aggregation of the data. Georgiadis (2015) and Feldkircher, Gruber, and Huber (2020b) model single countries within a monetary union and the monetary authority as a special entity; Cashin *et al.* (2014) set up a satellite model that determines oil prices based on world demand which in turns comes from the local (country-) models from the GVAR. The power of GVARs lies in the simplicity of the two-step approach that is based on estimating small, local models (with small dimensions of matrices to invert) while still allowing for parameter heterogeneity across units.

The general increase in data availability offers the researcher the possibility to analyze and model more and more complex dynamics. The cost, however, is the large number of parameters in the local models to be estimated and – related to that – issues of over-fitting and numerical instability. A potential solution to this is the use of Bayesian shrinkage priors, which put discipline on the coefficients. The usefulness of Bayesian methods in this context has been demonstrated by a range of recent studies (see, e.g., Carriero, Clark, and Marcellino 2019; Koop and Korobilis 2010). Following Crespo Cuaresma *et al.* (2016) and Feldkircher and Huber (2016), in **BGVAR**, the local models are hence estimated using Bayesian shrinkage priors. These include the well-known Minnesota prior (Litterman 1986), the Stochastic Search Variable Selection (SSVS) prior (George, Sun, and Ni 2008), the Normal-Gamma (NG) prior (Huber and Feldkircher 2019), and the horseshoe prior (Carvalho, Polson, and Scott 2010).

Despite its popularity, there is no implementation of GVARs in R (R Core Team 2019) to date.[2] There are, however, some Bayesian VAR packages which can be used to estimate large-scale models and thus be regarded as competitors to **BGVAR**: **bvarsv** by Krueger (2015), **bvartools** by Mohr (2019), **mfbvar** by Ankargren and Yang (2020), **BVAR** by Kuschnig and Vashold (2021), and **BHSBVAR** by Richardson (2020).[3] Importantly though, most of the priors employed in these packages are variations of Minnesota type shrinkage priors, which are well suited to deal with large data but ignore the cross-sectional data structure typically present in GVAR applications. Moreover, most packages assume a constant error variance over time, an assumption that is often not fulfilled when dealing with macroeconomic data. Allowing for heteroskedasticity in the error variances can be particularly useful when the time period under study is volatile, leading to considerable improvements in terms of forecasting

---

[1]For a recent survey on GVARs, see Chudik and Pesaran (2016). Alternative methods to capture cross-sectional dependence in large-dimensional datasets are discussed in Feldkircher, Gruber, and Huber (2020a).

[2]The only programming language for which a thorough treatment of GVARs is available is MATLAB with the toolbox of Smith, L.V. and Galesi, A. (2014) which follows closely the methodology outlined in Pesaran *et al.* (2004). But even in MATLAB there is no toolbox to estimate GVARs in a Bayesian fashion.

[3]The most established VAR implementation in R is **vars** (Pfaff 2008). The package uses a classical estimation framework which is probably not well suited for the estimation of unrestricted, large-scale VARs, which is why it is not listed among the other, more direct competitors above.

(see, e.g., Clark 2011).

The **BGVAR** package contributes to existing R packages in several dimensions: First, it fills an important gap by providing a GVAR implementation in R. By taking a Bayesian stance, **BGVAR** provides a GVAR framework with local models estimated using different, prominent shrinkage priors and stochastic volatility. Second, most of the computationally costly functions are coded in C++ making use of the **Rcpp** environment (Eddelbuettel and François 2011). This enables estimating large-scale Bayesian GVAR models on a standard computer. Third, the package includes convenience functions such as print and plot functions that allow to summarize the output, carry out residual and model diagnostics, and analyze the in-sample fit of the model. To enhance user-friendliness, method and argument names follow the **vars** package where applicable. Last, **BGVAR** includes several functions essential to perform inference in multivariate time series models, generalized to the GVAR case. These cover forecast error variance and historical decompositions, unconditional and conditional forecasts, and impulse response analysis. The latter can be specified without restrictions on the variance-covariance matrix, by using orthogonalization, or by introducing zero- and sign-restrictions.

The remainder of the paper is structured as follows. The next section introduces the econometric framework and the available prior choices. Section 3 describes the full extent of features available and describes its dependencies. In Section 4, we showcase the main features of **BGVAR** (model estimation, validation, impulse response analysis, and forecast error variance decompositions) with an example to US equity prices. Finally, Section 5 concludes.

2. Econometric Framework

2.1. Model Framework and Estimation

The GVAR modeling approach consists of two stages. The first stage deals with specifying unit-specific models. The second stage combines these models to obtain a global representation of the model.

In this discussion, we remain very general and assume a panel structure where $i = 1, \ldots, N$ indicates unit i and the total number of units is N. For each i, we assume that a k_i-dimensional vector of endogenous variables $\boldsymbol{y}_{it}$ follows a vector-autoregressive model with P lags. These are complemented with contemporaneous and lagged exogenous regressors $\boldsymbol{y}_{it}^*$ of dimension k_i^*. For a pre-specified set of weights w_{ij} (which satisfy $\sum_{j=1}^{N} w_{ij} = 1$ and $w_{ii} = 0$), the vector $\boldsymbol{y}_{it}^*$ represents a weighted average of the other units:

$$\boldsymbol{y}_{it}^* = \sum_{j=1}^{N} w_{ij} \boldsymbol{y}_{jt}.$$

Note that for the ease of notation, we assume equal weights for each element in $\boldsymbol{y}_{it}^*$. The package allows to relax this restriction by using separate weights for e.g., different variable types.

The resulting unit-specific model is given by:

$$\boldsymbol{y}_{it} = \boldsymbol{a}_{i0} + \boldsymbol{a}_{i1} t + \sum_{p=1}^{P} \boldsymbol{\Phi}_{ip} \boldsymbol{y}_{it-j} + \sum_{q=0}^{Q} \boldsymbol{\Lambda}_{iq} \boldsymbol{y}_{it-j}^* + \boldsymbol{\varepsilon}_{it}, \quad \boldsymbol{\varepsilon}_{it} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{it}), \tag{1}$$

where $\boldsymbol{a}_{i0}$ and $\boldsymbol{a}_{i1}$ denote k_i-dimensional coefficient vectors. These relate to the deterministic component of the model which we assume to consist of an intercept and a trend term.[4] Next, $\boldsymbol{\Phi}_{ij}$ $(j = 1, \ldots, P)$ denotes a $k_i \times k_i$-dimensional coefficient matrix associated with the lagged dependent variables, $\boldsymbol{\Lambda}_{ij}$ $(j = 0, \ldots, Q)$ denotes a $k_i \times k_i^*$ matrix of coefficients associated with the (weakly) exogenous regressors in $\boldsymbol{y}_{it}^*$, and $\boldsymbol{\varepsilon}_{it}$ is a Gaussian error vector with zero mean and $k_i \times k_i$-dimensional variance-covariance matrix $\boldsymbol{\Sigma}_{it}$. To allow for unit-by-unit estimation, we assume that $\boldsymbol{\varepsilon}_{it}$ and $\boldsymbol{\varepsilon}_{jt}$ are independent for all i,j and hence $\boldsymbol{\Sigma}_t = \mathrm{diag}(\boldsymbol{\Sigma}_{1t}, \ldots, \boldsymbol{\Sigma}_{Nt})$ is a block-diagonal matrix. This also implies that any co-movement between $\boldsymbol{y}_{it}$ and $\boldsymbol{y}_{jt}$ is exclusively driven by $\boldsymbol{y}_{it}^*$. The specification of the unit-specific models is completed by decomposing $\boldsymbol{\Sigma}_{it} = \boldsymbol{V}_i \boldsymbol{D}_{it} \boldsymbol{V}_i^\top$, where $\boldsymbol{V}_i$ denotes a lower uni-triangular matrix and $\boldsymbol{D}_{it} = \mathrm{diag}(e^{d_{i1,t}}, \ldots, e^{d_{ik_i,t}})$ with $d_{ij,t}$ following an AR(1) process.

The second step consists of stacking the estimated coefficients of the unit-specific models into a global system of equations. This step is straightforward and involves only algebraic transformations of the coefficients. For the sake of brevity, we outline the derivation of the global model by assuming $P = Q = 1$. This implies that

$$\boldsymbol{y}_{it} = \boldsymbol{a}_{i0} + \boldsymbol{a}_{i1}t + \boldsymbol{\Phi}_{i1}\boldsymbol{y}_{it-1} + \boldsymbol{\Lambda}_{i0}\boldsymbol{y}_{it}^* + \boldsymbol{\Lambda}_{i1}\boldsymbol{y}_{it-1}^* + \boldsymbol{\varepsilon}_{it}.$$

In what follows, we re-write the unit-specific models in terms of the global vector of endogenous variables $\boldsymbol{y}_t = \left(\boldsymbol{y}_{1t}^\top, \ldots, \boldsymbol{y}_{Nt}^\top\right)^\top$ of dimension $k = \sum_{j=1}^N k_j$. For that purpose, we need a unit-specific link matrix $\mathcal{W}_i$ $(k_i \times k)$ and a selection matrix $\boldsymbol{S}_i$ $(k_i \times k)$. The former stores the unit-specific weights to construct the (weakly) exogenous regressors $\boldsymbol{y}_{it}^*$, while the latter selects the unit-specific variables from the global vector. This results into

$$\boldsymbol{S}_i\boldsymbol{y}_t = \boldsymbol{a}_{i0} + \boldsymbol{a}_{i1}t + \boldsymbol{\Phi}_{i1}\boldsymbol{S}_i\boldsymbol{y}_{t-1} + \boldsymbol{\Lambda}_{i0}\mathcal{W}_i\boldsymbol{y}_t + \boldsymbol{\Lambda}_{i1}\mathcal{W}_i\boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_{it},$$
$$\underbrace{(\boldsymbol{S}_i - \boldsymbol{\Lambda}_{i0}\mathcal{W}_i)}_{=\boldsymbol{G}_i}\boldsymbol{y}_t = \boldsymbol{a}_{i0} + \boldsymbol{a}_{i1}t + \underbrace{(\boldsymbol{\Phi}_{i1}\boldsymbol{S}_i + \boldsymbol{\Lambda}_{i1}\mathcal{W}_i)}_{=\boldsymbol{H}_{i1}}\boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_{it},$$

Stacking this expression for all countries leads to

$$\boldsymbol{G}\boldsymbol{y}_t = \boldsymbol{a}_0 + \boldsymbol{a}_1t + \boldsymbol{H}_1\boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t,$$

with $\boldsymbol{a}_0 = \left(\boldsymbol{a}_{10}^\top, \ldots, \boldsymbol{a}_{N0}^\top\right)^\top$, $\boldsymbol{a}_1 = \left(\boldsymbol{a}_{11}^\top, \ldots, \boldsymbol{a}_{N1}^\top\right)^\top$, $\boldsymbol{G} = \left(\boldsymbol{G}_1^\top, \ldots, \boldsymbol{G}_N^\top\right)^\top$, $\boldsymbol{H}_1 = \left(\boldsymbol{H}_{11}^\top, \ldots, \boldsymbol{H}_{N1}^\top\right)^\top$, and $\boldsymbol{\varepsilon}_t = \left(\boldsymbol{\varepsilon}_{1t}^\top, \ldots, \boldsymbol{\varepsilon}_{Nt}^\top\right)^\top$. Pre-multiplying the expression with $\boldsymbol{G}^{-1}$ yields the global vector autoregressive model

$$\begin{aligned}\boldsymbol{y}_t &= \boldsymbol{G}^{-1}\boldsymbol{a}_0 + \boldsymbol{G}^{-1}\boldsymbol{a}_1t + \boldsymbol{G}^{-1}\boldsymbol{H}_1\boldsymbol{y}_{t-1} + \boldsymbol{G}^{-1}\boldsymbol{\varepsilon}_t \\ &= \boldsymbol{b}_0 + \boldsymbol{b}_1t + \boldsymbol{F}_1\boldsymbol{y}_{t-1} + \boldsymbol{e}_t,\end{aligned}$$

with $\boldsymbol{b}_0$ and $\boldsymbol{b}_1$ denoting $k \times 1$ global vectors consisting of deterministic terms, $\boldsymbol{F}_1$ denoting a $k \times k$-dimensional global coefficient matrix, and $\boldsymbol{e}_t$ the $k \times 1$-dimensional vector of global residuals. $\boldsymbol{e}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{et})$ is a Gaussian shock with $\boldsymbol{\Sigma}_{et}$ denoting the $k \times k$-dimensional global variance-covariance matrix. Note that $\boldsymbol{\Sigma}_{et} = \boldsymbol{G}^{-1}\boldsymbol{\Sigma}_t\boldsymbol{G}^{-1\top}$ is a full matrix. This implies that contemporaneous relations in the model are effectively driven by the components that determine $\boldsymbol{G}$, namely $\boldsymbol{\Lambda}_{i0}$ and the weights $\boldsymbol{w}_i = (w_{i1}, \ldots, \ldots, w_{iN})^\top$.

---

[4] **BGVAR** allows for including additional deterministic components and exogenous variables.

The same algebraic transformations as above can be used to derive the global representation of the model for the more general case with $P^* = \max(P, Q)$ lags:

$$\boldsymbol{y}_t = \boldsymbol{b}_0 + \boldsymbol{b}_1 t + \sum_{j=1}^{P^*} \boldsymbol{F}_j \boldsymbol{y}_{t-j} + \boldsymbol{e}_t, \quad \boldsymbol{e}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{et}). \tag{2}$$

Here, $\boldsymbol{F}_j$ $(j = 1, \ldots, P^*)$ denotes a $k \times k$-dimensional global coefficient matrix. Note that the corresponding parameters depend on the estimates of the submodels as well as the exogenously specified weights. Equation (2) resembles a standard large-scale VAR model with P^* lags.

2.2. Prior Specification

Since the GVAR model features a large number of free parameters, we adopt a Bayesian approach to estimation and inference. The package allows for estimating the GVAR model using three different priors. All these priors capture the notion that non-stationary variables follow a random walk process a priori. However, they differ with respect to how shrinkage is introduced. In this paper, we only summarize the main features of these priors and refer to the relevant literature for more information.

To simplify prior implementation, we collect all unit-specific coefficients in a $N_i = k_i K_i$ vector $\boldsymbol{\Psi}_i = (\boldsymbol{a}_{i0}, \boldsymbol{a}_{i1}, \boldsymbol{\Phi}_{i1}, \ldots, \boldsymbol{\Phi}_{iP}, \boldsymbol{\Lambda}_{i0}, \ldots, \boldsymbol{\Lambda}_{iQ}^*)^\top$ (with $K_i = k_i P + k_i^*(Q + 1) + 2$). Depending on the transformations of the time series in $\boldsymbol{y}_{it}$, we assume that the prior mean of $\boldsymbol{\Phi}_{i1}$ is equal to an identity matrix (if $\boldsymbol{y}_{it}$ is integrated of order one) or equal to a zero matrix (if $\boldsymbol{y}_{it}$ is stationary). To capture a situation where some of the endogenous variables are stationary but persistent, the prior mean can also be centered on an AR(1) process. For the remaining coefficients, the prior mean is always set equal to zero. For further convenience, we capture these prior assumptions on the mean in a vector $\underline{\boldsymbol{\Psi}}_i$.

All priors on the unit-specific coefficients in $\boldsymbol{\Psi}_i$ are Gaussian and have a prior mean equal to $\underline{\boldsymbol{\Psi}}_i$. Differences in the prior arise through the treatment of the prior variance $\underline{\boldsymbol{V}}_i$. More precisely, the prior on $\boldsymbol{\Psi}_i$ is given by

$$\boldsymbol{\Psi}_i \sim \mathcal{N}(\underline{\boldsymbol{\Psi}}_i, \underline{\boldsymbol{V}}_i). \tag{3}$$

Again, as stated above, we assume that $\underline{\boldsymbol{\Psi}}_i$ is the same for all priors. The different priors result from different choices of $\underline{\boldsymbol{V}}_i$.

The first prior we consider is a variant of the well-known non-conjugate Minnesota prior (for an overview, see Koop and Korobilis 2010, 2013). The prior variance of this setup differs across the different coefficient types in our model. First, for some equation j in unit i we assume that the prior variance on the kth lag of $y_{ij,t}$, the jth element of $\boldsymbol{y}_{it}$, is equal to λ_1^2/k^2. λ_1 is a hyperparameter that is either set by the researcher or estimated. This specification implies that higher lag orders will be forced towards zero since the prior variance is decreasing in the lag length. The intuition is that the more recent past is likely to explain more variation in $x_{ij,t}$ as compared to the more distant past.

Lags of the other endogenous variables $y_{is,t}$ for $s \neq j$ are treated differently. For those, the prior variance equals $\sigma_j^2 \sigma_s^{-2} \lambda_1^2 \lambda_2^2 / k^2$. The first term serves to control for scaling differences with σ_j, σ_s denoting the OLS residual standard deviations obtained by estimating univariate autoregressive models of order P and Q, respectively. The second term is similar to the prior variance we use on its "own" lags of a given variable but assume that $\lambda_1 > \lambda_2$. This controls

for the fact that lags of other variables tend to be less important than own lags of a given variable. Higher lag orders are, again, increasingly forced to zero.

The third block we consider are the coefficients on the weakly exogenous variables. For them, we specify the prior variance analogously to the case described in the previous paragraph but replace λ_2 with λ_3. This results in a prior variance of $\sigma_j^2 \sigma_s^{-2} \lambda_1^2 \lambda_3^2 / (k+1)^2$. Finally, we set the prior variance on the deterministic part of the model equal to λ_4, with λ_4 being a large value to render the prior effectively non-informative.

Since λ_1, λ_2, and λ_3 can be considered to be influential hyperparameters, we prefer inferring them from the data. Similar to Giannone, Lenza, and Primiceri (2015), we specify Gamma priors on $\lambda_s (s = 1, \ldots, 3)$. These Gamma priors are set to be weakly informative. This procedure allows us to estimate the shrinkage hyperparameters alongside the remaining coefficients of the model.

The main advantage of the prior discussed above is that it effectively depends only on a low number of hyperparameters. The downside of this, however, is a probably too restrictive structure since it implies the same degree of shrinkage on, for example, the coefficients related to the contemporaneous values of $\boldsymbol{y}_{it}^*$. To allow for more flexibility and control for model uncertainty, we use the Stochastic Search Variable Selection (SSVS) prior put forward in George and McCulloch (1993) and subsequently introduced to the VAR literature by George *et al.* (2008). This prior imposes a scale mixture of two Gaussian distributions on each element in $\boldsymbol{\Psi}_i$, Ψ_{ij}

$$\Psi_{ij} \mid \delta_{ij} \sim (1 - \delta_{ij}) \mathcal{N}(\underline{\Psi}_{ij}, \tau_{0,ij}^2) + \delta_{ij} \mathcal{N}(\underline{\Psi}_{ij}, \tau_{1,ij}^2), \quad \delta_{ij} \sim \text{Bernoulli}(\underline{p}), \tag{4}$$

where δ_{ij} is a binary random variable, $\tau_{1,ij}^2 \gg \tau_{0,ij}^2$ denote prior scaling parameters and $\underline{p}$ is the prior inclusion probability. If $\delta_{ij} = 1$, the prior variance is $\tau_{1,ij}^2$. This will be set to a large value, implying that relatively little information is introduced through the prior. By contrast, if $\delta_{ij} = 1$, the prior variance equals $\tau_{0,ij}^2$ which will be very close to zero. In this case, the prior dominates the likelihood and the resulting posterior estimate will be close to zero (i.e., the jth regressor will be excluded). George *et al.* (2008) modify this approach by setting $\tau_{0,ij}^2 = c_0 \hat{\sigma}_{\psi_{ij}}^2$ and $\tau_{1,ij}^2 = c_1 \hat{\sigma}_{\psi_{ij}}^2$. Here, we let $\hat{\sigma}_{\psi_{ij}}^2$ denote the OLS variance of ψ_{ij}, c_0 is set to a number close to zero, and c_1 will be large. Standard choices from the literature are $c_0 = 0.1$ and $c_1 = 10$. Finally, we set the prior inclusion probability $\underline{p} = 0.5$. This implies that all covariates are equally likely to enter the model a priori.

Spike and slab priors such as the SSVS prior have excellent theoretical properties (Bhattacharya, Pati, Pillai, and Dunson 2015). However, in large models, estimating the discrete indicators is difficult since vast model space of dimension 2^{K_i} needs to be explored. For moderately-sized datasets, K_i is typically large and even carefully constructed MCMC algorithms might run into mixing issues.

As a solution, one can approximate the behavior of a spike and slab prior by specifying an absolutely continuous shrinkage prior. Here, we use the Normal-Gamma (NG) prior as a variant of the global-local shrinkage priors originally proposed in Griffin, Brown *et al.* (2010) and applied to the VAR case in Huber and Feldkircher (2019). The NG prior is given by

$$\Psi_{ij} \mid \lambda_i^2, \theta_{ij} \sim \mathcal{N}(\underline{\Psi}_{ij}, 2\lambda_i^{-2}\theta_{ij}), \qquad \theta_{ij} \sim \mathcal{G}(\tau_\theta, \tau_\theta), \quad \lambda_i^2 \sim \mathcal{G}(d_\lambda, e_\lambda). \tag{5}$$

θ_{ij} denotes the *local* shrinkage parameter that is coefficient specific and λ_i is a *global* shrinkage term that pulls all elements in $\boldsymbol{\Psi}_i$ towards zero. This can be viewed as a common scaling factor

with the θ_{ij} allowing for coefficient-specific deviations in light of a large value of λ_i^2. On both the global and local parameters, we impose Gamma distributed priors with hyperparameters $\tau_\theta, d_\lambda, e_\lambda$. τ_θ controls the tail behavior of the prior with small values placing more prior mass on zero and leading to heavier tails. The remaining two hyperparameters d_λ and e_λ control the amount of global shrinkage with small values (i.e., of order 0.01) leading to heavy shrinkage towards the origin. To allow for more flexibility, it is possible to specify a hyperprior on $\tau_\theta \sim Exp(1)$. This prior setup leads to the Bayesian LASSO (Park and Casella 2008) and imposes heavier tails for smaller values of τ_θ. Since the conditional posterior distribution is not available in closed form, this hyperprior is implemented through an additional adaptive random-walk Metropolis-Hastings step. Since this is computationally more demanding than a simple Gibbs step, we allow for its depreciation and – unless the parameter is not user-specified – set it to $\tau_\theta = log(k_i)^{-1}$.

The key feature of this prior is that if the local scaling parameters are integrated out, the corresponding marginal prior is heavy tailed and effectively allows to separate signals from noise. Huber and Feldkircher (2019) show that this prior works well for US macroeconomic data and modify it to also capture features of the Minnesota prior such as increased penalty for higher lag lengths.

One potential drawback of the NG prior is that one has to specify hyperparameters. A possible remedy is the horseshoe (HS) prior (Carvalho *et al.* 2010), which is free of user-chosen hyperparameters but still highly adaptive and robust to a variety of situations. The HS prior is given by

$$\Psi_{ij} \mid \lambda_{ij} \sim \mathcal{N}(\underline{\Psi}_{ij}, \lambda_{ij}^2 \tau_i^2), \qquad \lambda_{ij} \sim C^+(0,1), \quad \tau_i \sim C^+(0,1) \tag{6}$$

where $C^+(0,a)$ is a standard half-Cauchy distribution on the positive reals with scale parameter a. λ_{ij} denotes the *local* shrinkage parameter that is coefficient specific and τ_i is a *global* shrinkage term that pulls all elements in $\boldsymbol{\Psi}_i$ towards zero. Makalic and Schmidt (2015) provide a simple and efficient sampling scheme based on auxiliary variables that lead to conjugate conditional posterior distribution of all parameters.

This concludes the prior setup for the coefficients. What remains is the prior set-up for the variance. We follow Carriero *et al.* (2019) and Carriero, Chan, Clark, and Marcellino (2022) and decompose the unit-specific covariance matrix into $\boldsymbol{\Sigma}_{it} = \boldsymbol{V}_i \boldsymbol{D}_{it} \boldsymbol{V}_i^\top$, where $\boldsymbol{V}_i$ is a lower uni-triangular matrix and $\boldsymbol{D}_{it}$ is a diagonal matrix with the log-volatilities on its diagonal. Here, we rely on the prior setup proposed in Kastner and Frühwirth-Schnatter (2014).[5] Since we sample the free elements in $\boldsymbol{V}_i$ using the algorithm outlined in Carriero *et al.* (2019) and Carriero *et al.* (2022), we can use a shrinkage prior similar to the ones discussed above.

2.3. Structural Inference

One of the main applications of a GVAR is to assess the domestic and international impulse responses to a unit-specific shock. **BGVAR** allows for calculating generalized impulse response functions (GIRFs, Pesaran and Shin 1998), orthogonalized impulse response functions using a Cholesky decomposition, and impulse response functions given a set of user-specified sign restrictions.

---

[5]This corresponds to using a Gaussian prior with zero mean and variance 10 on the unconditional mean of the log-volatility, a Beta prior on the persistence parameter ρ, $(\rho + 1)/2 \sim \mathcal{B}(25, 5)$ and a non-conjugate Gamma prior on the error variance of the log-volatility process $\sigma_h^2 \sim \mathcal{G}(1/2, 1/2)$. In case we do not use SV, we drop the subindex t and use weakly informative inverted Gamma priors on the innovation variances.

In principle, well-known identification procedures of the VAR literature can be applied to the global solution of the GVAR. Since the GVAR depends on an exogenous measure of connectivity of the cross-sectional units, there is no need for a full identification of the global model. The bulk of the GVAR literature consequently identifies the model only *locally* (Dees, di Mauro, Pesaran, and Smith 2007; Eickmeier and Ng 2015; Feldkircher and Huber 2016) which is also the way identification is implemented in **BGVAR**.

Local identification implies applying the Cholesky decomposition or sign restrictions only in the unit (country) of interest. For example, to identify a shock in the first unit model, it suffices to apply the identification scheme under consideration to Σ_{1t} (with $i = 1$ denoting the first unit model). More specifically, this boils down to

$$Q_1^{-1} y_{1t} = Q_1^{-1} a_{10} + Q_1^{-1} a_{11} t + Q_1^{-1} \sum_{j=1}^{P} \Phi_{1j} y_{1,t-j} + Q_1^{-1} \sum_{j=0}^{Q} \Lambda_{ij} y_{i,t-j}^* + \underbrace{Q_1^{-1} \varepsilon_{1t}}_{:=v_{1t}}, \quad (7)$$

with Q_1^{-1} being the lower Cholesky factor of Σ_{1t}. To construct the corresponding global representation of the GVAR with a locally identified shock, we have to set up a matrix Q

$$Q = \begin{pmatrix} Q_1 & 0 & \dots & 0 \\ 0 & I_{k_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & I_{k_N} \end{pmatrix},$$

with I_{k_i} denoting an identity matrix of dimension k_i. The global vector of structural errors is then obtained by $Q^{-1} \varepsilon_t = (v_{1t}', \varepsilon_{2t}', \dots, \varepsilon_{1t}')'$. Hence, the local identification yields orthogonal residuals in country model i, whereas the remaining residuals are potentially correlated. As a result, impulse responses to shocks in unit 1 are structural whereas responses to shocks outside the first unit model are of the generalized type.

Identification via sign restrictions follows in the same fashion with the exception that instead of Q_1 we use $Q_1 R_1$ with R_1 denoting a $k_1 \times k_1$ rotation matrix. The rotation matrix has to be sampled either by using the algorithm of Rubio-Ramirez, Waggoner, and Zha (2010) or – in case we couple zero with sign restrictions – the algorithm of Arias, Rubio-Ramírez, and Waggoner (2018). **BGVAR** also allows to place sign restrictions on the cross-section or to examine responses to a joint / regional shock. This implies that the identity matrices on the diagonal of Q are replaced by suitable $Q_i R_i$ matrices. It is worth noting that this increases computational costs considerably.

3. The BGVAR Package

The **BGVAR** package provides a user-friendly implementation of Bayesian Global Vector Autoregressions in R (R Core Team 2019).

Its main function is the `bgvar` method:

```
R> model.1 <- bgvar(Data, W, plag=1, draws=5000, burnin=5000, prior="NG",
+                   SV=TRUE, hold.out=0, thin=1, hyperpara=NULL, eigen=TRUE,
+                   Ex=NULL, trend=FALSE, expert=NULL, verbose=TRUE)
```

At a minimum, the user has to submit a `Data` argument and a $N \times N$ weighting matrix. The `Data` can be submitted either as a $T \times k$ matrix. The column-names of the matrix object have to follow a specific naming convention: The string should consist of two parts, with the first indicating the unit model and the second referring to the name of the variable. These two parts should be separated by a dot.[6] The package also provides a function to read data from `Excel`. Note that the order of the unit models should be the same as indicated by the column and row names of the weight matrix `W`.

The local models in **BGVAR** can be estimated using four shrinkage priors: The Minnesota prior (`prior="MN"`) of Litterman (1986), the Stochastic Search Variable Selection prior (`prior="SSVS"`) of George *et al.* (2008), the Normal-Gamma prior (`prior="NG"`) of Huber and Feldkircher (2019), and the Horseshoe prior (`prior="HS"`) of Carvalho *et al.* (2010). For each of the implemented priors, all hyperparameters and -priors (if available) are adjustable. This can be done by submitting a list object to `hyperpara`.

Other options regarding estimation involve that we use all observations in the sample (`hold.out=0`), no trend and no truly exogenous variables (`trend=FALSE, Ex=NULL`), 5,000 burn-ins and retained posterior draws (`burnin=5000, draws=5000`), and a lag length of one (`plag=1`). Additionally, we allow to estimate the local models with stochastic volatility to control for possible heteroskedasticity (`SV=TRUE`).

In order to ensure stability of the global model, explosive draws of the posterior are discarded. We follow recommendations in Dees *et al.* (2007) and Smith, L.V. and Galesi, A. (2014) and discard draws for which the eigenvalues of the companion matrix of the global solution of the model exceed 1.05. Other cut-off values can be specified via `eigen`. For applications that require a large number of draws from the joint posterior, saving the full set of posterior draws can become unfeasible. To address this, we can use a thinning factor to store only each thinth draw from the posterior. The default option uses a thinning of 1 (`thin=1`), which implies that each posterior draw will be saved.

BGVAR allows to modify the standard GVAR framework in several ways. This can be done by submitting a list to `expert`. Expert settings comprise the treatment of specific unit models (e.g., a unit to model oil prices or common monetary policy, see, Cashin *et al.* 2014; Feldkircher *et al.* 2020b), distinct weights to compute weakly exogenous variables (Eickmeier and Ng 2015), or exclusion restrictions on the set of variables for which weakly exogenous variables should be calculated.

Other settings relate to storage and computational aspects. Per default, **BGVAR** makes use of **Rccp** which yields substantial computational benefits. To side-step the C++ environment (e.g., for more in-depth error-handling), one can specify `expert = list(use_R = TRUE)`.

The package is available cross-platform and depends only on well established R packages. Most of the package is implemented in native R for transparency and is also available on Github[7] in order to make it easier for users to create adaptions, suggest new features, or report bugs. Since the estimation and parts of the provided tools are computationally expensive, the package relies on C++. This implies a main dependency on **Rcpp** (Eddelbuettel and François 2011) and **RcppArmadillo** (Eddelbuettel and Sanderson 2014). To ensure an easy

---

[6]Alternatively, the user can submit data as a list object of length N, with each element in the list containing the $T \times k_i$ data of one specific unit model. The names of the list should refer to the unit models. The package provides a converter function between matrix- and list-style input data.

[7]There is a development version of the package available on GitHub: `https://github.com/mboeck11/BGVAR`.

way to modify the package, every line of code written in C++ also has a fall-back option written in R (see previous paragraph). To further speed up computation, **BGVAR** allows the use of parallel computation in the estimation function. This implies that **BGVAR** relies on **parallel** (R Core Team 2019). Stochastic volatility is implemented using the **stochvol** package of Kastner (2016). Further dependencies are **mvtnorm** (Genz, Bretz, Miwa, Mi, Leisch, Scheipl, and Hothorn 2020) and **GIGrvg** (Leydold and Hormann 2017) to simulate from the conditional posterior distributions, the packages **abind** (Plate and Heiberger 2016), **MASS** (Venables and Ripley 2002), **Matrix** (Bates and Maechler 2019), **readxl** (Wickham and Bryan 2019), **xts** (Ryan and Ulrich 2020), and **zoo** (Zeileis and Grothendieck 2005) for data manipulation, and **coda** (Plummer, Best, Cowles, and Vines 2006) and **bayesm** (Rossi 2019) to inspect posterior output.

BGVAR includes several further methods that are worth discussing. As such, forecasting methods are of particular interest, since GVARs possess excellent forecasting properties (Crespo Cuaresma *et al.* 2016; Dovern *et al.* 2016). More specifically, **BGVAR** allows for computing unconditional as well as conditional forecast distributions. Conditional forecasts can be used to perform scenario analysis, i.e., examining the effects of fixing the future path of one variable on the remaining variables. The implementation of conditional forecasts in **BGVAR** follows closely the work of Jarociński (2010).

The main application of GVARs is to estimate spillover effects between cross-sectional units or more generally analyze cross-unit dependence. This can be achieved through impulse response functions (IRFs), forecast error variance decompositions (FEVDs) and historical decompositions (HDs). Impulse response functions show how a unit model reacts to exogenous shocks. These shocks can both originate in the same or another unit model. The propagation of these shocks through the system is of great interest, but relies on proper identification. **BGVAR** offers different identification schemes: recursive identification based on the Cholesky decomposition, sign restrictions, or a combination of sign- and zero-restrictions. While the implementation of the Cholesky decomposition is computationally cheap, sign-restrictions require a computationally costly algorithm. In particular, we implement the algorithm of Rubio-Ramirez *et al.* (2010) and Arias *et al.* (2018). To speed up computation, impulse response functions are implemented in C++. Forecast error variance decompositions indicate the amount of information each variable contributes to the other variables in the system by examining how much of the forecast error variance of each of the variables can be explained by exogenous shocks. A historical decomposition allows to examine the relative importance of (structural) shocks in explaining deviations of a time series from its unconditional mean. This can be used to assess the hypothetical question of how data would have looked like if it was driven only by a particular shock (e.g., monetary policy shock) or a combination of shocks.

BGVAR also includes different and ready to use datasets. A very popular data set with an extensive country and time period coverage is `pesaranData`, which has been used in Dees *et al.* (2007) and updated by Mohaddes and Raissi (2020). We will use this data set for the illustrative example and describe it in more detail. The data set `eerData` covers 43 countries and the euro area as a regional aggregate. A subset of this data set has been used in Feldkircher and Huber (2016). `testdata` is a further subset of `eerData` which contains only three countries and is used for illustrational purposes of some – time-consuming – functions. All of these data sets feature a quarterly frequency and are available either in levels or as first differenced data. The data sets also include example weight matrices, which are typically

Table 1: Functionality of **BGVAR**.

Function	Class	Description
Estimation function		
bgvar	bgvar	Estimation of the Bayesian global vector autoregression.
Tool functions		
irf	bgvar.irf	Computation of impulse response functions.
predict	bgvar.pred	Computation of (conditional) predictions.
fevd	bgvar.fevd	Computation of a forecast error variance decomposition.
gfevd	bgvar.fevd	Computation of a generalized forecast error variance decomposition.
hd	bgvar.hd	Computation of a historical decomposition.
*Helper functions - **bgvar***		
coef		S3 method of bgvar: Extracts coefficients.
coefficients		S3 method of bgvar: Extracts coefficients.
dic		S3 method of bgvar: Computes Deviance information criterion.
fitted		S3 method of bgvar: Extracts fitted values.
logLik	logLik	S3 method of bgvar: Compututation of the log-likelihood.
print		S3 method of bgvar, bgar.summary, and bgvar.CD.
plot		S3 method of bgvar and bgvar.resid.
resid	bgvar.resid	S3 method of bgvar: Extracts residuals.
residuals	bgvar.resid	S3 method of bgvar: Extracts residuals.
summary	bgvar.summary	S3 method of bgvar: Computes summary statistics.
vcov		S3 Method of bgvar: Extracts the covariance matrix.
avg.pair.cc		Average pairwise cross-sectional correlations.
conv.diag	bgvar.CD	Computes Geweke's convergence diagnostic.
resid.corr.test		Residual autocorrelation test.
matrix_to_list		Converts an input matrix to a suitable list for estimation.
list_to_matrix		Converts an input list to a suitable matrix for estimation.
excel_to_list		Reads Data from Excel in a suitable list for estimation.
*Helper functions - **irf***		
print		S3 method of bgvar.irf.
plot		S3 method of bgvar.irf.
get_shockinfo		Creates shockinfo argument.
add_shockinfo		Adapts shockinfo argument.
*Helper functions - **predict***		
plot		S3 method of bgvar.pred.
print		S3 method of bgvar.pred, bgvar.lps, and bgvar.rmse.
lps	bgvar.lps	Log-predictive scores of predictions.
rmse	bgvar.rmse	Root mean squared errors of predictions.
*Helper functions - **fevd / hd***		
plot		S3 method of bgvar.fevd and bgvar.hd.
print		S3 method of bgvar.fevd and bgvar.hd.

based on annual, bilateral trade flows. Last, `monthlyData` contains monthly macroeconomic and financial observations for 28 countries. The corresponding weight matrix is based on bilateral input-output flows from the world input-output data base (Timmer, Dietzenbacher, Los, Stehrer, and de Vries 2015).

To sum up, **BGVAR** allows for the estimation of Bayesian global vector autoregressions in a user-friendly manner. Furthermore, state-of-the-art methods to conduct predictions and structural inference in time series analysis are available. The package offers various tools for structural analysis, such as impulse response analysis, forecast error variance decompositions, and historical decompositions. In addition, the package includes a comprehensive documentation as well as several helper functions to ensure user-friendliness. These functions comply with those of related R packages and include `coef`, `fitted`, `logLik`, `plot`, `predict`, `print`, `resid`, `summary`, as well as `vcov`. An overview of all functions, classes, and available methods is provided in Table 1.

4. Application: A Shock to US Equity Prices

In this section, we showcase the features of **BGVAR** by investigating the international effects of a negative shock to US equity prices. We use a built-in dataset and go through a typical application: (1) loading the data, (2) estimating the model and examining its statistical properties, (3) evaluating alternative specifications, and finally (4) performing structural analysis.

We start with loading the **BGVAR** package and the built-in dataset `pesaranData`. We also set a seed for reproducibility of the results.

```
R> library("BGVAR")
R> data("pesaranData")
R> set.seed(413)
```

`pesaranData` contains quarterly observations from 1979Q2 to 2019Q4 and for 29 countries. The countries are listed in Table 2 and comprise a good balance between advanced and emerging economies as well as oil producers and consumers. Most of the advanced economies are located in Europe, either as part of the euro area or with an independent monetary policy authority. Other advanced economies comprise major industrial countries from the rest of the world, such as the United States (US). With the exception of Japan, one country from Latin America (LatAm), the Middle East and African (MEA) region and Asia can be considered to be emerging economies. Taken together, the countries account for about 90% of global output. The variables covered in the data set comprise real GDP (y_{it}), consumer price inflation (Dp_{it}), short- and long-term interest rates (r_{it} and lr_{it}), the real exchange rate against the US dollar (ep_{it}), and equity prices (eq_{it}). All variables except interest rates are in logarithmic form. The exchange rate is deflated using consumer prices with an increase implying a real depreciation of the local currency against the dollar. Short-term interest rates are transformed as $0.25 \times ln(1 + R_{it}/100)$, with R_{it} denoting 3-months money market rates. A similar transformation is used for long-term rates, with the underlying interest rates corresponding to 10-year government bond yields.[8] All country models at least contain data

---

[8]Global control variables, such as oil prices, a price index for agricultural raw materials, as well as metal prices are contained in the object `dominant` but are by default not included in any of the country models.

Table 2: Country Overview

EA	Non-EA	RoW	LatAm	MEA	Asia
AT	CH	AU	CL	MY	CN
BE	SE	CA		ZA	IN
FI	NO	NZ		TR	ID
FR	GB	US			JP
DE					KR
IT					MY
ES					PH
NL					SG
					TH

Notes: Euro Area (EA), Rest of World (RoW), Latin America (LatAm), Middle-East & Africa (MEA)

on output and inflation. The remaining variables, especially data on long-term yields and equity prices, are missing for some emerging economies. Last, we need to consider a weight matrix that links the country models. Here, we use `W.8016` which automatically loads with `pesaranData`. The matrix is based on annual bilateral trade flows averaged over the period 1980 to 2016. Other weight matrices can be constructed using trade flows. For an in-depth description of the data and underlying sources, see Mohaddes and Raissi (2020).

4.1. Estimation and Model Diagnostics

To estimate a model with the default options, we have to submit as a minimum a data set and a weight matrix:

```
R> model.1 <- bgvar(Data = pesaranData, W = W.8016)
```

The default options imply that we use the NG prior (`prior="NG"`), 5,000 burn-ins and retained posterior draws (`burnin=5000`, `draws=5000`), no trend (`trend=FALSE`), a lag length of one (`plag=1`), and stochastic volatility (`SV=TRUE`).

There are several ways of summarizing the model's output. A first impression can be gathered by considering the information printed in the console. The `print` method provides an overview about the variables contained in each country model, where an asterisk denotes weakly exogenous variables, a double asterisks denotes exogenous variables, and variables without asterisks show the endogenous variables per unit.

A more elaborate way of inspecting the model's properties can be obtained by invoking the `summary` method:

```
R> summary(model.1)
```

```
--------------------------------------------------------------------------------
Model Info:
```

```
Prior: Normal-Gamma prior (NG)
Number of lags for endogenous variables: 1
Number of lags for weakly exogenous variables: 1
Number of posterior draws: 5000/1=5000
Number of stable posterior draws: 1035
Number of cross-sectional units: 28
--------------------------------------------------------------------------
Convergence diagnostics
Geweke statistic:
2176 out of 23870 variables' z-values exceed the 1.96 threshold (9.12%).
--------------------------------------------------------------------------
F-test, first order serial autocorrelation of cross-unit residuals
Summary statistics:
========= ========== ======
\          # p-values in %
========= ========== ======
>0.1       65         42.21%
0.05-0.1   10         6.49%
0.01-0.05  24         15.58%
<0.01      55         35.71%
========= ========== ======

--------------------------------------------------------------------------
Average pairwise cross-unit correlation of unit-model residuals
Summary statistics:
======= ========= ========= ========= ========= ========= =========
\        y         Dp        r         lr        ep        eq
======= ========= ========= ========= ========= ========= =========
<0.1     28 (100%) 28 (100%) 28 (100%) 18 (100%) 27 (100%) 25 (100%)
0.1-0.2  0 (0%)    0 (0%)    0 (0%)    0 (0%)    0 (0%)    0 (0%)
0.2-0.5  0 (0%)    0 (0%)    0 (0%)    0 (0%)    0 (0%)    0 (0%)
>0.5     0 (0%)    0 (0%)    0 (0%)    0 (0%)    0 (0%)    0 (0%)
======= ========= ========= ========= ========= ========= =========

--------------------------------------------------------------------------
```

We first investigate whether the posterior distribution has converged to its target distribution. **BGVAR** draws on Geweke's convergence diagnostic (CD), implemented in the **coda** package. In a nutshell, the diagnostic is based on a test for equality of the means of the first and last part of a Markov chain (by default we use the first 10% and the last 50%). If the samples are drawn from the stationary distribution of the chain, the two means are equal and Geweke's statistic has an asymptotically standard normal distribution. The statistic indicates practical convergence for the majority of the variables.

The next statistic is an F-test of the first-order serial autocorrelation in the residuals of each equation in the GVAR. By default and in compliance with the literature, we consider the residuals of each unit model as opposed to the residuals of the global solution of the GVAR. The null hypothesis of the F-test corresponds to the case of no serial correlation. The table shows the share of p-values that fall into different significance categories. Changing the number of lags and switching stochastic volatility on / off both can impact residual serial

```
R> yfit <- fitted(model.1)
R> plot(model.1, global=FALSE, resp="US")
```
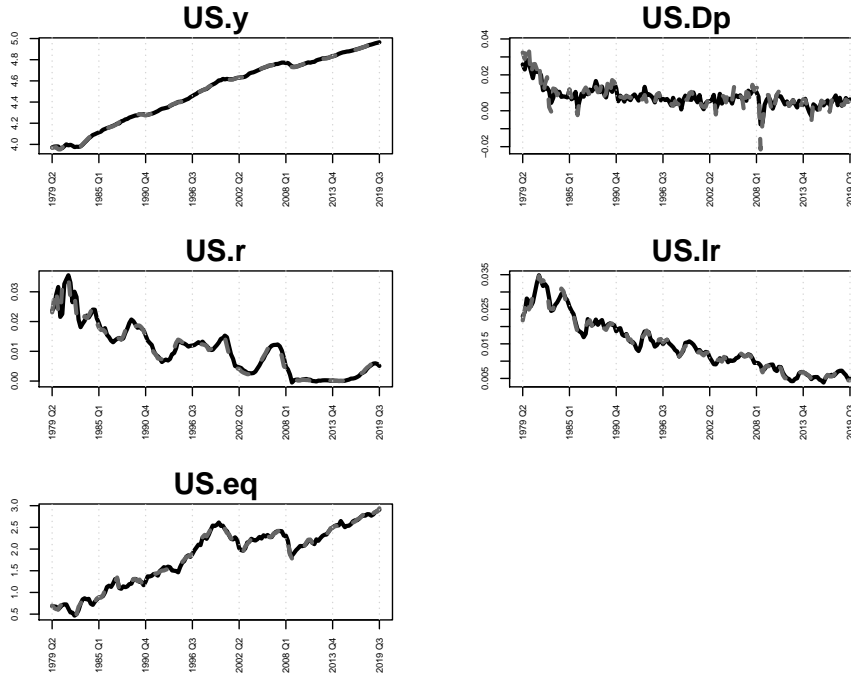

Figure 1: In-sample fit for US variables.

correlation. The GVAR framework relies on the assumption that the inclusion of weakly exogenous variables soaks up most of the cross-unit correlation in the residuals. Significant correlations prohibit a structural and spillover analysis (Dees *et al.* 2007). This can be inspected by looking at simple pair-wise correlations, which are provided in the bottom part of the summary output. More specifically, we provide a statistic on the cross-unit correlation of (posterior median) residuals of the unit models. In our example, correlation is reasonably small.

Besides the `summary` method, the package provides additional ways of inspecting the properties of the model. For example, the in-sample fit can be obtained by using `fitted` and the residuals using `residuals`. For both methods, the default option is to look at the residuals of the global solution of the model. Setting `global=FALSE`, provides results of the unit models' residuals. In Figure 1 we show the in-sample fit for the US model.

For all variables considered, the model provides a reasonable in-sample fit. In the next subsection, we further examine the out-of-sample properties of different model specifications.

4.2. Model Validation

In this sub-section, we briefly outline ways to validate the specification for the estimated model. One way to do this is to use a forecasting exercise. In particular, we are interested in whether other prior-choices and lag lengths would improve the out-of-sample properties of the model. We do this for all four available priors and `plag=1` and `plag=2` lags. For the sake

of brevity, we only use a small number of posterior draws. By specifying `hold.out=8`, we tell **BGVAR** to reserve eight observations as a hold-out sample that we can use to compute predictive performance measures. The following lines estimate the models for the different specifications:

```
R> model1.ssvs <- bgvar(Data=pesaranData,W=W.8016,plag=1,hold.out=8,thin=2,
+                 draws=1000,burnin=1000, prior="SSVS")
R> model1.ng <- bgvar(Data=pesaranData,W=W.8016,plag=1,hold.out=8,thin=2,
+                 draws=1000,burnin=1000, prior="NG")
R> model1.mn <- bgvar(Data=pesaranData,W=W.8016,plag=1,hold.out=8,thin=2,
+                 draws=1000,burnin=1000, prior="MN")
R> model1.hs <- bgvar(Data=pesaranData,W=W.8016,plag=1,hold.out=8,thin=2,
+                 draws=1000,burnin=1000,prior="HS")
R> model2.ssvs <- bgvar(Data=pesaranData,W=W.8016,plag=2,hold.out=8,thin=2,
+                 draws=1000,burnin=1000,prior="SSVS")
R> model2.ng <- bgvar(Data=pesaranData,W=W.8016,plag=2,hold.out=8,thin=2,
+                 draws=1000,burnin=1000, prior="NG")
R> model2.mn <- bgvar(Data=pesaranData,W=W.8016,plag=2,hold.out=8,thin=2,
+                 draws=1000,burnin=1000, prior="MN")
R> model2.hs <- bgvar(Data=pesaranData,W=W.8016,plag=2,hold.out=8,thin=2,
+                 draws=1000,burnin=1000,prior="HS")
```

We use the `predict` method to compute dynamic predictions based on the estimated coefficients. We set the forecast horizon to `n.ahead=8`. As a summary measure of the predictive performance, we calculate log-predictive scores (LPS) using the `lps` method. We could do this for selected variables of interest. Below, we simply calculate the sum of all LPS to get an overall impression of the forecasting performance under a particular specification:

```
R> fcast <- predict(model1.ssvs, n.ahead=8)
R> lps.ssvs<-sum(lps(fcast))
```

We repeat the same exercise for the other specifications and summarize the results in Table 3 below.

	SSVS	NG	MN	HS
p=1	3631.88	3473.59	3623.63	3477.40
p=2	3647.38	3413.72	3792.75	3455.45

Table 3: LPS scores

This forecasting exercise is not decisive whether using `plag=2` lags pays off for predictive accuracy. An alternative way of model selection is the use of information criteria. In **BGVAR**, we have implemented the deviance information criterion (DIC) of Spiegelhalter, Best, Carlin, and Van Der Linde (2002), which can be calculated using the `dic` command. In Table 4, we show the DIC values relative to the one of the default specification using the NG prior and one lag.

	SSVS	NG	MN	HS
p=1	1.00	1.00	17.90	0.00
p=2	0.08	0.00	0.01	1.50

Table 4: DIC relative to NG (p=1)

Together with the assessment of LPS values, the DIC points to a more parsimonious model. Taken together, the MN prior with `plag=1` lags seems to yield a good compromise between out-of-sample performance and model complexity with the data used in the example.

```
          used  (Mb) gc trigger  (Mb) limit (Mb)   max used    (Mb)
Ncells   1591509 85.0   2529041  135.1        NA    2529041   135.1
Vcells 101526239 774.6 404716196 3087.8    102400 1327618215 10129.0
```

4.3. Structural Analysis

In this section, we proceed by estimating the cross-unit propagation of a local shock. More specifically, we assess the international effect of a 1% decrease in US equity prices. This kind of spillover analysis is the most frequent application of the GVAR framework (see Chudik and Pesaran 2016).

The application of the GVAR, however, can also be motivated from a second angle: In case of strong cross-unit correlation in the data, even if the focus does not lie on spillovers but on the domestic response in a unit model, the omission of foreign factors can cause misleading inference (for a recent application in that context, see Feldkircher and Siklos 2019). One way of assessing cross-unit interlinkages that goes beyond a simple correlation analysis is by using a (generalized) forecast error variance decomposition (Diebold and Yılmaz 2014).

The **BGVAR** package offers to estimate both, structural and generalized forecast error variance decompositions (GFEVD). The following computes a GFEVD with a forecast horizon of 24 periods (`n.ahead=24`):

```
R> gfevd.1=gfevd(model.1,n.ahead=24,running=TRUE)$FEVD
```

Here, we have chosen to calculate the posterior mean of the GFEVD as a running mean. The alternative would be to store the full set of posterior draws and to calculate posterior statistics of interest ex post.

To assess the degree of cross-unit dependence, we can calculate the shares of forecast error variance explained by domestic and foreign factors. In the following, we focus on Canada, an important trading partner of the US. Figure 2 shows the relative importance of domestic and foreign factors in explaining Canadian output:

The findings reveal a typical pattern commonly observed in international macroeconomics. First, a large degree of forecast error variance can be accounted for by own, domestic variables. However, as the forecast horizon increases, other (foreign) factors become increasingly important. This implies that only short-run fluctuations in output can be more or less accurately modeled by only considering domestic variables.
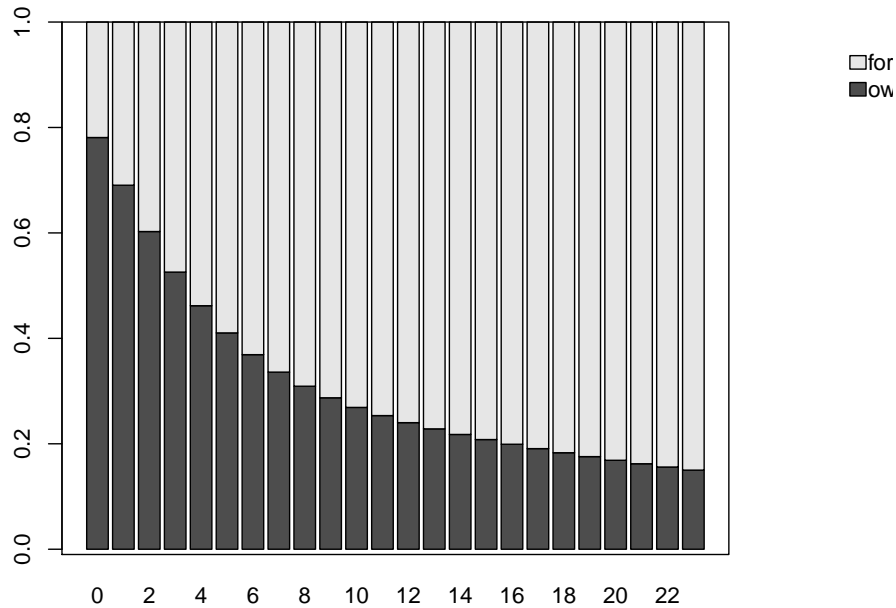
Figure 2: GFEVD of Canadian output.

To see whether these results are specific to Canadian output, we can examine the share of output variance accounted for by domestic factors across countries. For the ease of illustration, we focus on two forecast horizons, $h = 0$ (impact) and $h = 24$ (long-run). The results are depicted in Figure 3.

The top panel of Figure 3, shows the fraction of output variance explained by domestic factors at the $h = 0$ forecast horizon. We see that for all economies, the share is above 0.6. This implies that short-run fluctuations in output tend to be largely explained by domestic policy. Domestic factors tend to be even more important for emerging economies (e.g., Brazil, Indonesia, China, etc.) and some selected advanced economies (Japan, Australia, and Norway). The bottom panel of Figure 3 shows the same statistic for the $h = 24$ forecast horizon. Here, we see a pronounced decrease of the importance of domestic factors in explaining forecast error variance of output. More specifically, at a longer horizon, global factors tend to be as important as domestic factors for most economies. An exception to this general pattern can be seen for output in Japan, the Philippines, Brazil, and Indonesia.

Having established the importance of taking cross-unit linkages into account, we proceed by investigating spillovers from a negative shock to US equity prices. Following the bulk of the GVAR literature (Dees *et al.* 2007; Eickmeier and Ng 2015; Feldkircher and Huber 2016), we identify the US equity price *locally*. This translates into applying the identifying restrictions only in the unit (country) of interest.

We start with using a Cholesky decomposition of the residuals in the US country model to identify the equity price shock. It is worth stressing that a recursive identification scheme is not invariant to different orderings of the variables in the unit model. In our example, the ordering is output, inflation, short- and long-term interest rates followed by equity prices. In
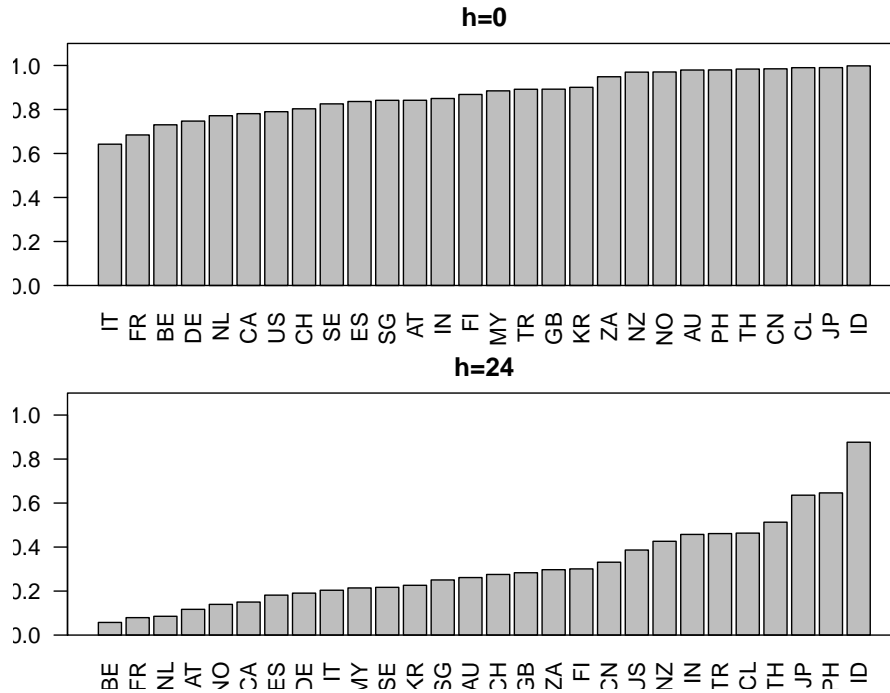
Figure 3: Share of output variance explained by domestic factors.

the following, we set up the information needed to compute impulse responses using the `irf` function:

```
R> shockinfo_chol <- get_shockinfo("chol")
R> shockinfo_chol$shock <- "US.eq"
R> shockinfo_chol$scale <- -1
R> irf_chol <- irf(model.1, n.ahead = 24, shockinfo = shockinfo_chol)
```

`get_shockinfo` sets up a list with elements `shock` (i.e., the name of the variable to shock) and `scale`. By default, the shock is a *positive* innovation to the equity price residual. In the example, we have re-scaled the innovation to −1. In a final step, the list is then passed onto the `irf` function, which demands the impulse response horizon (`n.ahead=24`) as a further argument.

We can plot the computed impulse response function by using the `plot` command. The plot shows the posterior median response (solid, black line) along 68% (dark grey) and 80% (light grey) credible intervals. We see that the initial -1% decrease in the equity price index translates into a persistently negative response up to about 15 quarters. This has negative effects on output, which also declines significantly and up to 20 quarters.

The GVAR framework now allows us to investigate spillovers of the equity price shock to all other variables and countries in the system. For illustration, we focus on spillovers to output in important trading partners of the US, namely Canada (CA) and Germany (DE), as well as countries that are considerably less integrated with the US, namely Switzerland (CH), Turkey (TR). The results are shown in Figure 5.
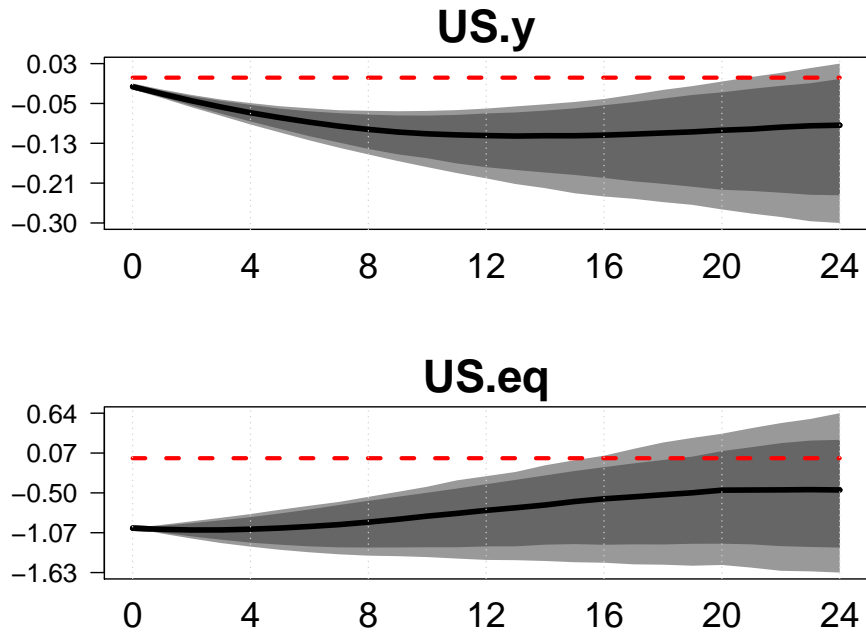
```
R> plot(irf_chol, resp=c("US.eq","US.y"))
```


Figure 4: Impulse response functions in the US country model (Cholesky identification).

```
R> plot(irf_chol, resp=c("CH.y","TR.y","CA.y","DE.y"))
```
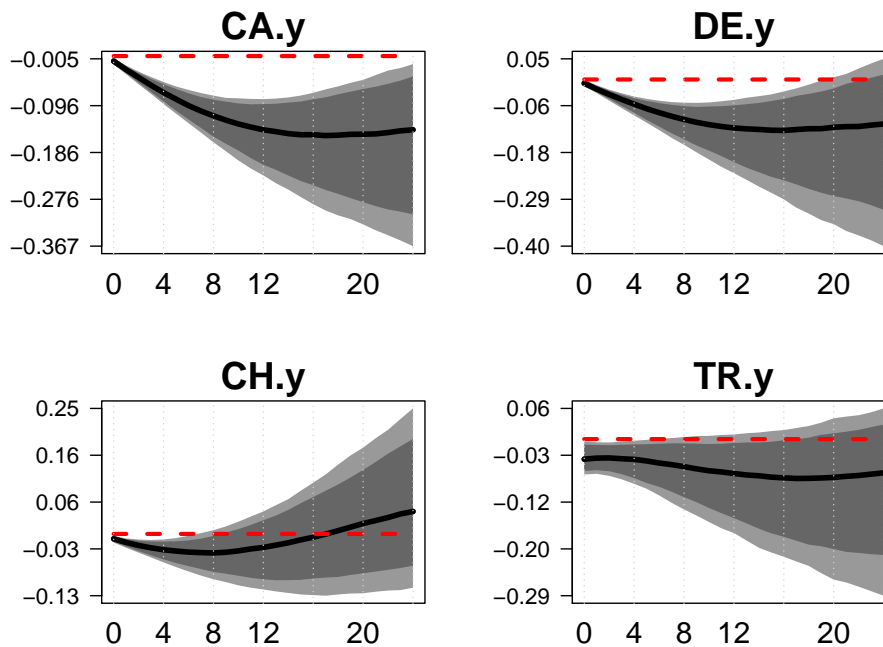

Figure 5: Output spillovers for selected countries (Cholesky identification).

The figure shows that the countries that are to a higher degree interlinked with the US witness a larger, negative impact on output. In that case, spillovers also tend to be more persistent. By contrast, effects on output in Switzerland and Turkey are far less pronounced.

We contrast these results by using an alternative way to identify the US equity shock, namely sign restrictions. More specifically, we use a very loose identification which only consists of the condition that both equity prices and output have to move in the same direction.

```
R> shockinfo_sign <- get_shockinfo("sign")
R> shockinfo_sign <- add_shockinfo(shockinfo_sign, shock="US.eq",
+                                  restriction=c("US.eq","US.y"),
+                                  sign=c(">",">"),
+                                  horizon=c(4,4), prob=1, scale=-1)
R> irf_sign <- irf(model.1, n.ahead = 24, shockinfo=shockinfo_sign)
```

As with the Cholesky identification, we first have to set up an auxiliary object using `get_shockinfo`. The necessary information regarding the sign restrictions can be added in a second step using `add_shockinfo`. We always consider a positive innovation / shock, hence the first element in `sign` has to be positive ($>$). The second element reflects the restriction of a parallel increase in output and is hence positive as well. Both restrictions have to hold for four periods (impact and three quarters) and have to be exactly binding (`prob=1`). Since we are interested in a negative shock to equity prices, we use `scale=-1`. The results are shown in Figure 6.

```
R> plot(irf_sign, resp=c("US.eq","US.y"), shock="US.eq")
```
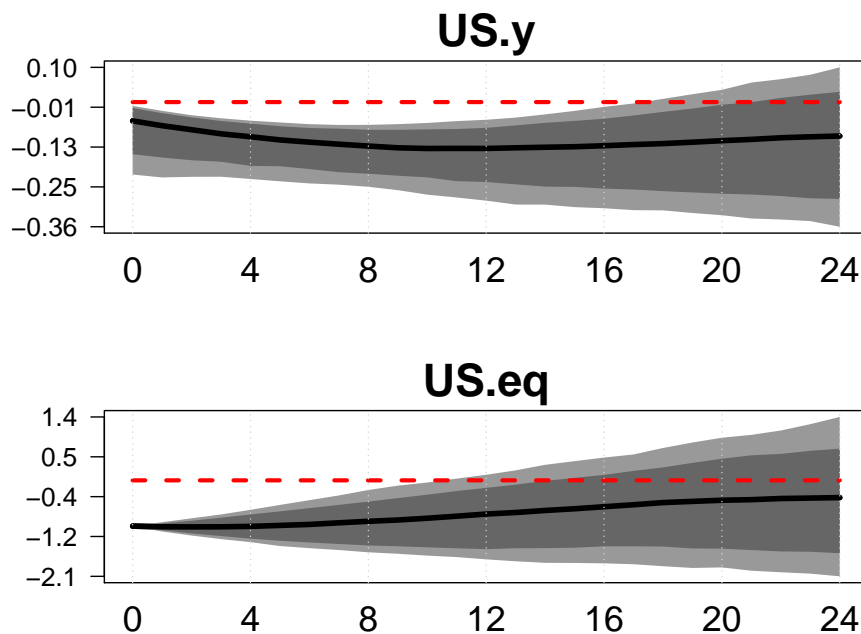

Figure 6: Output spillovers for selected countries (sign restrictions).

The figure demonstrates a negative response of both US equity prices and output. Negative responses to output prevail up to 17 quarters and hence considerably longer than the horizon

of the restrictions imply. Responses are hence qualitatively similar to the ones obtained using an identification strategy based on the Cholesky decomposition.

5. Summary and Discussion

This article describes the functionality of the R package **BGVAR** to estimate Bayesian global vector autoregressions with shrinkage priors and stochastic volatility. The package includes convenience functions such as `print` and `plot` functions that allow to summarize the output, carry out residual and model diagnostics, and analyze the in-sample fit of the model. In addition, the package offers a fully fledged toolkit to compute impulse response functions, forecast error variance, and historical error variance decompositions. Structural shocks can be identified by either using a combination of zero and sign restrictions or by applying a simple Cholesky decomposition. Sign restrictions and shocks can be applied on single units or on the cross-section. Generalized impulse response functions are implemented as well. The package also provides functions to conduct and evaluate out-of-sample forecasts as well as conditional forecasts that allow to set a future path for a particular variable of interest.

Further applications of the **BGVAR** package are contained in the vignette that accompanies the package. These feature estimation of the models with different weights, including additional units to model e.g., the joint monetary policy in the euro area (Feldkircher *et al.* 2020b), or the oil price (Cashin *et al.* 2014), applications of zero-and sign restrictions coupled with rationality conditions of D'Amico and King (2015), and conditional and unconditional forecasts. Naturally, all the necessary commands are explained in the package documentation that also include detailed examples.

The Bayesian treatment of the unit-specific models is a unique feature of **BGVAR**. This allows to include large information sets and in turn, typically results in considerable gains in forecast performance over standard GVAR models. Another stand-alone feature of **BGVAR** is its efficient implementation: large parts of the estimation function are written in C++ utilizing the **Rcpp** and **RcppArmadillo** packages. In addition, both estimation and structural analysis can be carried out using parallel computing drawing on the **parallel** package. This implies that most applications can be carried out on a standard computer and in a reasonable amount of time, which greatly enhances the usability of the **BGVAR** package.

Computational details

The results in this paper were obtained using R 4.0.3 with the **BGVAR** 2.4.6 package. R itself and all packages used are available from the Comprehensive R Archive Network (CRAN) at https://CRAN.R-project.org/.

Acknowledgements

References

Ankargren S, Yang Y (2020). **mfbvar**: *Mixed-Frequency Bayesian VAR Models*. R package version 0.5.3, URL https://CRAN.R-project.org/package=mfbvar.

Arias JE, Rubio-Ramírez JF, Waggoner DF (2018). "Inference Based on Structural Vector Autoregressions Identified with Sign and Zero Restrictions: Theory and Applications." *Econometrica*, **86**(2), 685–720.

Bates D, Maechler M (2019). **Matrix**: *Sparse and Dense Matrix Classes and Methods*. R package version 1.2-18, URL https://CRAN.R-project.org/package=Matrix.

Benecká S, Fadejeva L, Feldkircher M (2020). "The impact of euro Area monetary policy on Central and Eastern Europe." *Journal of Policy Modeling*, **42**(6), 1310–1333.

Bhattacharya A, Pati D, Pillai NS, Dunson DB (2015). "Dirichlet-Laplace Priors for Optimal Shrinkage." *Journal of the American Statistical Association*, **110**(512), 1479–1490.

Böck M, Feldkircher M, Siklos PL (2021). "International effects of euro area forward guidance." *Oxford Bulletin of Economics and Statistics*, **83**(5), 1066–1110.

Burriel P, Galesi A (2018). "Uncovering the heterogeneous effects of ecb unconventional monetary policies across euro area countries." *European Economic Review*, **101**, 201–229.

Carriero A, Chan J, Clark TE, Marcellino M (2022). "Corrigendum to "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors"[J. Econometrics 212 (1)(2019) 137–154]." *Journal of Econometrics*, **227**(2), 506–512.

Carriero A, Clark TE, Marcellino M (2019). "Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-conjugate Priors." *Journal of Econometrics*, **212**(1), 137–154.

Carvalho CM, Polson NG, Scott JG (2010). "The horseshoe estimator for sparse signals." *Biometrika*, **97**(2), 465–480.

Cashin P, Mohaddes K, Raissi M, Raissi M (2014). "The Differential Effects of Oil Demand and Supply Shocks on the Global Economy." *Energy Economics*, **44**(C), 113–134.

Cesa-Bianchi A (2013). "Housing cycles and macroeconomic fluctuations: A global perspective." *Journal of International Money and Finance*, **37**(C), 215–238.

Chudik A, Grossman V, Pesaran MH (2016). "A multi-country approach to forecasting output growth using PMIs." *Journal of Econometrics*, **192**(2), 349–365.

Chudik A, Pesaran MH (2016). "Theory and Practice of GVAR Modelling." *Journal of Economic Surveys*, **30**(1), 165–197.

Clark TE (2011). "Real-time Density Forecasts from Bayesian Vector Autoregressions with Stochastic Volatility." *Journal of Business & Economic Statistics*, **29**(3), 327–341.

Colabella A (2021). "Do ECB's Monetary Policies Benefit EMEs? A GVAR Analysis on the Global Financial and Sovereign Debt Crises and Postcrises Period." *Oxford Bulletin of Economics and Statistics*, **83**(2), 472–494.

Crespo Cuaresma J, Feldkircher M, Huber F (2016). "Forecasting with Global Vector Autoregressive Models: A Bayesian Approach." *Journal of Applied Econometrics*, **31**(7), 1371–1391.

D'Amico S, King TB (2015). "What Does Anticipated Monetary Policy Do?" *Working Paper Series WP-2015-10*, Federal Reserve Bank of Chicago.

Dees S, di Mauro F, Pesaran HM, Smith LV (2007). "Exploring the International Linkages of the Euro Area: a Global VAR Analysis." *Journal of Applied Econometrics*, **22**(1).

Diebold FX, Yılmaz K (2014). "On the network topology of variance decompositions: Measuring the connectedness of financial firms." *Journal of Econometrics*, **182**(1), 119–134.

Dovern J, Feldkircher M, Huber F (2016). "Does Joint Modelling of the World Economy Pay Off? Evaluating Global Forecasts from a Bayesian GVAR." *Journal of Economic Dynamics and Control*, **70**, 86 – 100.

Eddelbuettel D, François R (2011). "**Rcpp**: Seamless R and C++ Integration." *Journal of Statistical Software*, **40**(8), 1–18.

Eddelbuettel D, Sanderson C (2014). "**RcppArmadillo**: Accelerating R with High-performance C++ Linear Algebra." *Computational Statistics and Data Analysis*, **71**, 1054–1063.

Eickmeier S, Ng T (2015). "How do US credit supply shocks propagate internationally? A GVAR approach." *European Economic Review*, **74**, 128 – 145. ISSN 0014-2921. doi:http://dx.doi.org/10.1016/j.euroecorev.2014.11.011.

Fadejeva L, Feldkircher M, Reininger T (2017). "International Spillovers from Euro Area and US Credit and Demand Shocks: A focus on Emerging Europe." *Journal of International Money and Finance*, **70**, 1–25.

Feldkircher M, Gruber T, Huber F (2020a). "International effects of a compression of euro area yield curves." *Journal of Banking & Finance*, **113**, 11–14.

Feldkircher M, Gruber T, Huber F (2020b). "International Effects of a Compression of Euro Area Yield Curves." *Journal of Banking & Finance*, **113**, 11–14.

Feldkircher M, Huber F (2016). "The International Transmission of US Shocks – Evidence from Bayesian Global Vector Autoregressions." *European Economic Review*, **81**, 167–188.

Feldkircher M, Siklos P (2019). "Global inflation dynamics and inflation expectations." *International Review of Economics & Finance*, **64**, 217–241.

Genz A, Bretz F, Miwa T, Mi X, Leisch F, Scheipl F, Hothorn T (2020). **mvtnorm**: *Multivariate Normal and t Distributions*. R package version 1.0-12, URL https://CRAN.R-project.org/package=mvtnorm.

George EI, McCulloch RE (1993). "Variable Selection via Gibbs Sampling." *Journal of the American Statistical Association*, **88**(423), 881–889.

George EI, Sun D, Ni S (2008). "Bayesian Stochastic Search for VAR Model Restrictions." *Journal of Econometrics*, **142**(1), 553–580.

Georgiadis G (2015). "Examining asymmetries in the transmission of monetary policy in the euro area: Evidence from a mixed cross-section global VAR model." *European Economic Review*, **75**(C), 195–215.

Giannone D, Lenza M, Primiceri GE (2015). "Prior Selection for Vector Autoregressions." *Review of Economics and Statistics*, **97**(2), 436–451.

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.

Griffin JE, Brown PJ, *et al.* (2010). "Inference with Normal-Gamma Prior Distributions in Regression Problems." *Bayesian Analysis*, **5**(1), 171–188.

Hájek J, Horváth R (2018). "International spillovers of (un)conventional monetary policy: The effect of the ECB and the US Fed on non-euro EU countries." *Economic Systems*, **42**, 91–105.

Huber F, Feldkircher M (2019). "Adaptive Shrinkage in Bayesian Vector Autoregressive Models." *Journal of Business & Economic Statistics*, **37**(1), 27–39.

Jarociński M (2010). "Conditional Forecasts and Uncertainty about Forecast Revisions in Vector Autoregressions." *Economics Letters*, **108**(3), 257–259.

Kastner G (2016). "Dealing with Stochastic Volatility in Time Series Using the R Package **stochvol**." *Journal of Statistical Software*, **69**.

Kastner G, Frühwirth-Schnatter S (2014). "Ancillarity-sufficiency Interweaving Strategy (ASIS) for Boosting MCMC Estimation of Stochastic Volatility Models." *Computational Statistics & Data Analysis*, **76**, 408–423.

Koop G, Korobilis D (2010). "Bayesian Multivariate Time Series Methods for Empirical Macroeconomics." *Foundations and Trends(R) in Econometrics*, **3**(4), 267–358. doi:10.1561/0800000013.

Koop G, Korobilis D (2013). "Large time-varying parameter VARs." *Journal of Econometrics*, **177**(2), 185–198.

Krueger F (2015). **bvarsv***: Bayesian Analysis of a Vector Autoregressive Model with Stochastic Volatility and Time-Varying Parameters.* R package version 1.1, URL https://CRAN.R-project.org/package=bvarsv.

Kuschnig N, Vashold L (2021). "BVAR: Bayesian vector autoregressions with hierarchical prior selection in R." *Journal of Statistical Software*, **100**, 1–27.

Leydold J, Hormann W (2017). **GIGrvg***: Random Variate Generator for the GIG Distribution.* R package version 0.5, URL https://CRAN.R-project.org/package=GIGrvg.

Litterman RB (1986). "Forecasting with Bayesian Vector Autoregressions – Five Years of Experience." *Journal of Business & Economic Statistics*, **4**(1), 25–38.

Makalic E, Schmidt DF (2015). "A simple sampler for the horseshoe estimator." *IEEE Signal Processing Letters*, **23**(1), 179–182.

Mohaddes K, Raissi M (2020). "Compilation, Revision and Updating of the Global VAR (GVAR) Database, 1979Q2-2019Q4." *University of Cambridge: Faculty of Economics (mimeo)*.

Mohr FX (2019). **bvartools**: *Functions for Bayesian Inference of Vector Autoregressive Models*. R package version 0.0.2, URL https://CRAN.R-project.org/package=bvartools.

Park T, Casella G (2008). "The Bayesian Lasso." *Journal of the American Statistical Association*, **103**(482), 681–686.

Pesaran HH, Shin Y (1998). "Generalized Impulse Response Analysis in Linear Multivariate Models." *Economics Letters*, **58**(1), 17–29.

Pesaran MH, Schuermann T, Smith LV (2009). "Forecasting Economic and Financial Variables with Global VARs." *International Journal of Forecasting*, **25**(4), 642–675.

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

Pfaff B (2008). "VAR, SVAR and SVEC Models: Implementation Within R Package **vars**." *Journal of Statistical Software*, **27**(4), 1–32.

Plate T, Heiberger R (2016). **abind**: *Combine Multidimensional Arrays*. R package version 1.4-5, URL https://CRAN.R-project.org/package=abind.

Plummer M, Best N, Cowles K, Vines K (2006). "CODA: Convergence Diagnosis and Output Analysis for MCMC." *R News*, **6**(1), 7–11. URL https://journal.r-project.org/archive/.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Richardson P (2020). **BHSBVAR**: *Structural Bayesian Vector Autoregression Models*. R package version 2.0.2, URL https://CRAN.R-project.org/package=BHSBVAR.

Rossi P (2019). **bayesm**: *Bayesian Inference for Marketing/Micro-Econometrics*. R package version 3.1-4, URL https://CRAN.R-project.org/package=bayesm.

Rubio-Ramirez JF, Waggoner DF, Zha T (2010). "Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference." *The Review of Economic Studies*, **77**(2), 665–696.

Ryan JA, Ulrich JM (2020). **xts**: *eXtensible Time Series*. R package version 0.12-0, URL https://CRAN.R-project.org/package=xts.

Smith, LV and Galesi, A (2014). "GVAR Toolbox." URL https://sites.google.com/site/gvarmodelling/gvar-toolbox.

Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002). "Bayesian measures of model complexity and fit." *Journal of the royal statistical society: Series b (statistical methodology)*, **64**(4), 583–639.

Timmer MP, Dietzenbacher E, Los B, Stehrer R, de Vries GJ (2015). "An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production." *Review of International Economics*, **23**, 575–605.

Venables WN, Ripley BD (2002). *Modern Applied Statistics with S.* Fourth edition. Springer-Verlag, New York.

Wickham H, Bryan J (2019). *readxl: Read Excel Files.* R package version 1.3.1, URL `https://CRAN.R-project.org/package=readxl`.

Zeileis A, Grothendieck G (2005). "**zoo**: S3 Infrastructure for Regular and Irregular Time Series." *Journal of Statistical Software*, **14**(6), 1–27.

Affiliation:

Maximilian Böck
Vienna School of International Studies
Favoritenstraße 15a
1040 Vienna, Austria
E-mail: `maximilian.boeck@da-vienna.ac.at`

and

Vienna University of Economics and Business
Welthandelsplatz 1
1020 Vienna, Austria
E-mail: `maximilian.boeck@wu.ac.at`

Martin Feldkircher
Vienna School of International Studies
Favoritenstraße 15a
1040 Vienna, Austria
E-mail: `martin.feldkircher@da-vienna.ac.at`

and

Oesterreichische Nationalbank
Otto-Wagner-Platz 3
1090 Vienna, Austria
E-mail: `martin.feldkircher@oenb.at`

Florian Huber
University of Salzburg
Mönchsberg 2a
5020 Salzburg, Austria
E-mail: `florian.huber@sbg.ac.at`