Fast, Order-Invariant Bayesian Inference in VARs using the Eigendecomposition of the Error Covariance Matrix

Ping Wu*
Department of Economics, University of Strathclyde
and
Gary Koop
Department of Economics, University of Strathclyde

September 2, 2025

Abstract

Bayesian inference in Vector Autoregressions (VARs) involves manipulating large matrices which appear in the posterior (or conditional posterior) of the VAR coefficients. For large VARs, the computational burden of these manipulations can render empirical work impractical. In response to this, many researchers transform their VARs so as to allow for Bayesian estimation to proceed one equation at a time. This leads to a massive reduction in the computational burden. This transformation involves taking the Cholesky decomposition for the error covariance matrix. However, this strategy implies that posterior inference depends on the order the variables enter the VAR. In this paper we develop an alternative transformation, based on the eigendecomposition, which does not lead to order dependence. Beginning with an inverse-Wishart prior on the error covariance matrix, we derive and discuss the properties of the prior it implies on the eigenmatrix and eigenvalues. We then show how an extension of the prior on the eigenmatrix can allow for greater flexibility while maintaining many of the benefits of conjugacy. We leverage this flexibility to extend the prior on the eigenvalues to allow for stochastic volatility. The properties of the eigendecomposition approach are investigated in a macroeconomic forecasting exercise involving VARs with 20 variables.

Keywords: Eigendecomposition, order invariance, large vector autoregression

*Both authors are research fellows in the Economic Statistics Centre of Excellence (ESCoE). We would like to thank Joshua Chan, Herman van Dijk, Luis Gruber, Niko Hauzenberger, Gregor Kastner, Aubrey Poon, Mike West, Dan Zhu, participants at the 13th European Seminar on Bayesian Econometrics and Örebro University 6th Annual Workshop on Financial Econometricsr, as well as Editor Atsushi Inoue, an anonymous Associate Editor and two anonymous reviewers, for many constructive comments and valuable suggestions.

1 Introduction

Vector Autoregressions (VARs) have shown their usefulness in a range of applications in macroeconomics (e.g., Sims, 1980; Cogley and Sargent, 2005; Primiceri, 2005; Koop, 2013; Korobilis, 2013). Starting with Banbura et al. (2010), macroeconomists have increasingly worked with large VARs, containing tens or even hundreds of dependent variables. Bayesian inference and prediction with such large VARs present an enormous computational challenge. The main computational bottleneck relates to the posterior for the matrix of VAR coefficients, $\mathbf{A}$. For VARs involving n variables and p lags this will contain $n^2 \times p$ coefficients which is huge for empirically reasonable choices for n and p. Manipulation involving features such as the posterior covariance matrix of $\mathbf{A}$ will involve working with matrices of dimension $pn^2 \times pn^2$. However, working with the VAR one equation at a time leads to manipulations involving $pn \times pn$ matrices. Even though such manipulations must be repeated n times, this equation-by-equation strategy has been shown, e.g. by Carriero et al. (2019), to lead to an $O(n^2)$ reduction in the computational burden. Even for $n = 10$ the benefits of equation by equation estimation can be seen to be substantial and for $n = 100$ they are enormous.

In the conventional (reduced form) VAR, with $\mathbf{y}_t = (y_{1,t}, \ldots, y_{n,t})'$ being the vector of dependent variables and Σ being the error covariance matrix, equation-by-equation estimation is not possible since the off-diagonal elements of Σ would be ignored. However, if the VAR is transformed so that the error covariance matrix becomes diagonal, equation-by-equation estimation is valid. These considerations motivate most of the recent large VAR literature which uses the Cholesky decomposition to do this transformation. In particular, let $\Sigma^{-1} = \mathbf{L}'\mathbf{C}^{-1}\mathbf{L}$, where $\mathbf{C}$ is a diagonal matrix and $\mathbf{L}$ is a lower triangular matrix with ones on the diagonal, then the Cholesky-transformed VAR has dependent variables $\mathbf{L}\mathbf{y}_t$ and a diagonal error covariance matrix and equation-by-equation estimation is possible.

Several researchers have pointed out that the use of the Cholesky decomposition leads to order dependence. Order dependence occurs if the posterior for any VAR parameter depends on the way the variables are ordered in the VAR. Order dependence also implies that the predictive distribution for a variable changes if its ordering in the VAR changes. In the case of Bayesian VAR analysis, ordering issues can arise due to the prior used for the error covariance matrix. Carriero et al. (2019) theoretically demonstrate that the posterior of the VAR coefficients, conditional on the error covariance matrix Σ remains invariant to ordering. The lack of order invariance arises due to the fact that the implied prior on Σ is not order invariant. Thus, most of the discussion in this literature, including the present paper, centers on the error covariance matrix.

Another important recent paper is Arias et al. (2023) which demonstrates the importance of the ordering issue in Cholesky-transformed VARs both theoretically and empirically. The authors show that, while point forecasts are not sensitive to the way variables are ordered, predictive standard deviations can be substantially affected. These points are reinforced in Chan et al. (2023) who also show that the problems with order invariance become more pronounced as the dimension of the VAR grows. In other words, the problem is most acute precisely where the Cholesky transformation is most necessary.[1]

There are some papers which retain the Cholesky decomposition, but seek to choose the optimal order or average over orders. Papers such as Levy and Lopes (2021) and Wu and Koop (2023) develop methods for computing the probability of each ordering, then do order selection or order averaging. However, the limitation of these approaches is that they are only practical in relatively small VARs. When n is large, the number of possible orderings becomes enormous, dramatically increasing the computational burden of these approaches.

[1] Arias et al. (2023) and Chan et al. (2023) use VARs with stochastic volatility as opposed to homoskedastic VARs which are the focus of the first half of this paper. But their theoretical derivations also hold with homoskedastic VARs.

These considerations motivate the present paper. In it, we develop a new approach to Bayesian VAR analysis which uses the eigendecomposition of the error covariance. We call this model BVAR-eig. We begin by considering the inverse-Wishart prior for Σ which is known to be order invariant. We show that this implies priors for the eigenmatrix (i.e. the matrix of eigenvectors) and eigenvalues which appear in BVAR-eig to have Bingham and inverse-Gamma distributions, respectively. We refer to this as the $\mathcal{BIG}$ prior. We develop an efficient MCMC algorithm that allows for Bayesian posterior and predictive inference based on the this prior.

Using our BVAR-eig with $\mathcal{BIG}$ prior, we carry out a substantial empirical exercise involving homoskedastic VARs with up to 20 macro variables. We compare the performance of our eigendecomposition-based approach to the Cholesky-based approach (labeled BVAR-chol) and to full system estimation of a conventional Bayesian VAR with inverse-Wishart prior (BVAR-IW). In relation to BVAR-IW we demonstrate large improvements in computation time. Relative to the BVAR-chol, we demonstrate some improvements in in-sample and forecast performance. We also demonstrate the BVAR-chol's high degree of sensitivity to the choice of variable order.

In the second half of the paper, we consider extensions of BVAR-eig which could be empirically useful. The main extension we consider is the addition of stochastic volatility (SV). Since SV is empirically important in many applications (e.g., Cogley and Sargent, 2005; Primiceri, 2005; Koop and Korobilis, 2013) we highlight this extension as being of substantive interest.

It is not straightforward to introduce SV to the VAR with $\mathcal{BIG}$ prior since it implies that the eigenvalues control both the scale of the error covariance and the orientation of the eigenmatrix. Hence, we follow a suggestion of Hoff (2009b) and break this dependence by introducing a separate prior parameter to control the orientation of the eigenmatrix. We call this the independent Bingham inverse-gamma or $\mathcal{IBIG}$ prior. The $\mathcal{IBIG}$ prior is potentially interesting in and of itself. However, we introduce it mainly since it allows us to introduce SV into the VAR in a simple way. Since the eigenvalues now only relate to the scale of the variances, we can assume they are SV processes without imposing restrictions on the orientation of the eigenmatrix leading to the model we call BVAR-eigSV.

We also consider a scale-adjusted version of BVAR-eigSV which can be used by researchers who want an approach which is not only order invariant, but also does not depend on the scale the variables are measured in. We end with discussion of methods of allowing for time-variation in the eigenmatrix.

2 Eigendecomposition and the Bayesian VAR

2.1 Transforming the VAR: BVAR-eig and BVAR-chol

To fix the basic ideas, consider a standard homoskedastic VAR of order p. Let $\mathbf{y}_t = (y_{1,t}, \ldots, y_{n,t})'$ be an $n \times 1$ vector of variables that is observed over the periods $t = 1, \ldots, T$. Then, the VAR(p) is given by:

$$\mathbf{y}_t = \mathbf{a} + \mathbf{A}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{A}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \Sigma), \tag{1}$$

where $\mathbf{a}$ is an $n \times 1$ vector of intercepts, $\mathbf{A}_1, \ldots, \mathbf{A}_p$ are $n \times n$ coefficient matrices. Note that there are n equations and each equation has $k = np + 1$ regressors which leads to a total of $nk = n^2 p + n$ coefficients. We will refer to this as the reduced form VAR.

The eigendecomposition of a positive definite covariance matrix Σ is given by $\Sigma = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}'$. The matrix $\mathbf{U}$ is an $n \times n$ orthogonal matrix of eigenvectors and is called the eigenmatrix. $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_n)$ is the matrix of non-negative eigenvalues.

The BVAR-eig is produced by multiplying the VAR by $\mathbf{U}'$ leading to a transformed VAR with $\mathbf{U}'\mathbf{y_t}$ on the left hand side and a diagonal error covariance matrix $\boldsymbol{\Lambda}$. The BVAR-chol uses the Cholesky-decomposition leading to a transformed VAR with $\mathbf{Ly_t}$ on the left hand side and a diagonal error covariance matrix $\mathbf{C}$. Thus, both have a diagonal error covariance matrix and it is this property which allows for equation-by-equation Bayesian estimation of both of them.

Note, however, that for the BVAR-chol the lower triangularity of $\mathbf{L}$ means that the equation for each variable will have different contemporaneous values of variables appearing in them. For instance, the first variable will appear in all equations, the second will appear in all remaining $(n-1)$ equations, etc.. In contrast, other than being orthogonal $\mathbf{U}$ is unrestricted which means each equation will have the same variables appearing in it. It is these properties which, through the choices of prior used in these models, that leads to the fact that the BVAR-eig is order-invariant while BVAR-chol is not.

2.2 Priors for BVAR-eig and BVAR-chol

For the BVAR-chol, it is standard to use a Normal prior on the elements of $\mathbf{L}$ and inverse-gamma priors on diagonal elements in $\mathbf{C}$. As noted above, it is well-established that this leads to a posterior which depends on the way the variables are ordered in the VAR. It is also worth stressing that this prior does not imply an inverse-Wishart prior on Σ use of which does imply order invariance in the reduced form VAR.[2]

To develop a prior for BVAR-eig, our strategy will be to begin with the commonly used inverse-Wishart prior on the reduced form VAR error covariance matrix Σ, then work out what it implies for $\mathbf{U}$ and $\mathbf{\Lambda}$. In particular, we assume the prior for Σ to have shape parameter $v > 0$ and scale matrix $\mathbf{S}_0$ and, thus, $\Sigma \sim \mathcal{IW}(v, \mathbf{S}_0)$.

Sub-section 2.4 of Hoff (2009b) shows what the inverse-Wishart prior on Σ implies for

[2]It is important to distinguish between the BVAR-chol representation as used, e.g., in Carriero et al. (2019) and the asymmetric conjugate prior approach of Chan (2022). The latter author uses a modified Cholesky decomposition (see footnote 5) and provides a different way of decomposing an inverse-Wishart covariance matrix which also leads to order-invariant inference (see Proposition 1).

its eigenmatrix and eigenvalues. It turns out the former has a Bingham distribution and the diagonal elements of the latter have inverse-Gamma distributions. We will not reproduce the full proofs of Hoff (2009b) but provide some details as they allow us to discuss some properties of the Bingham distribution.

A matrix $\mathbf{U}$ is said to have a generalised Bingham distribution if its density function is given by

$$p(\mathbf{U} \mid \mathbf{D}, \mathbf{B}, \mathbf{V}) = c(\mathbf{D}, \mathbf{B})\, \mathrm{e}^{\mathrm{tr}(\mathbf{B}\mathbf{U}'\mathbf{V}\mathbf{D}\mathbf{V}'\mathbf{U})}, \tag{2}$$

where $c(\mathbf{D}, \mathbf{B})$ is the integrating constant given in equation (8) of Hoff (2009b). $\mathbf{D}$ and $\mathbf{B}$ are diagonal matrices. It can be shown that, conditional on Λ, the prior for $\mathbf{U}$ has this form where $\mathbf{B} = -\frac{1}{2}\mathbf{\Lambda}^{-1}$. $\mathbf{V}$ and $\mathbf{D}$ are defined through the eigendecomposition of $\mathbf{S}_0$. That is $\mathbf{S}_0 = \mathbf{V}\mathbf{D}\mathbf{V}'$.

The Bingham distribution, proposed for a vector by Bingham (1974) and extended to a matrix variate version by Khatri and Mardia (1977) is the prior we use on the eigenmatrix. The conditional posterior of the eigenmatrix also turns out to be Bingham and a computationally efficient method for taking draws from it is developed in Hoff (2009b).

The inverse-Wishart prior for Σ can be shown to imply the prior for the eigenvalues in $\mathbf{\Lambda}$ to be independent inverse-gamma distributions with arguments $\alpha = \frac{v+n-1}{2}$, $\beta = \frac{1}{2}u_i'\left(\mathbf{V}\mathbf{D}\mathbf{V}'\right)u_i$ where u_i is the ith column of $\mathbf{U}$.

Thus, the inverse-Wishart prior for Σ implies a Bingham and inverse-Gamma, or $\mathcal{BIG}$, prior for the eigenmatrix and eigenvalues. We emphasize that this equivalence between the $\mathcal{BIG}$ and the inverse Wishart prior automatically implies order-invariance (and, thus, the proof of order invariance given in Appendix A is not needed for this case).

For both the eigendecomposition and Cholesky decomposition versions of the model, we make prior hyperparameter choices which, where possible, are the same. Remember that

$\mathbf{S}_0 = \mathbf{VDV}'$. To elicit prior hyperparameters for the $\mathcal{BIG}$ prior, we adopt the following strategy. We use a relatively non-informative Minnesota prior for the error covariance matrix involving a degrees of freedom of $v = n + 3$, and $\mathbf{S}_0 = \text{diag}\left(s_1^2, \ldots, s_n^2\right)$. We then set $\mathbf{V} = I_n$ resulting in $\mathbf{D} = \mathbf{S}_0$. Further details of prior hyperparameter choice are provided in Appendix B.

2.3 MCMC Algorithms for BVAR-eig and BVAR-chol

The MCMC algorithm for BVAR-chol we use is the one in Chan et al. (2022) to which the reader is referred for further details. The model of Chan et al. (2022) has SV. In this section, we consider only homoskedastic models and use the same inverse-Gamma priors on the error variances as we do for BVAR-eig.

For the BVAR-eig, we will show, drawing on results from Hoff (2009b), that the conditional posteriors of the eigenmatrix and eigenvalues are Bingham and inverse-Gamma and draws from them can easily be taken. But before providing details, it is worth mentioning two issues which arise in this algorithm which are specific to our use of the eigendecomposition. The first is that the $\mathcal{BIG}$ prior suffers from the label switching problem. One will get the same likelihood by permuting the columns of eigenmatrix and then the eigenvalues accordingly. This can be overcome using the random permutation sampler of Frühwirth-Schnatter (2001). The second is that, for reasons explained below, the MCMC algorithm proposed by Hoff (2009b) breaks down the full eigenmatrix $\mathbf{U}$ into draws of two columns at a time instead of simply drawing one column at a time. This feature is particularly important for high-dimensional VARs and we will show that the computation is as fast as when using the Cholesky decomposition.

Posterior draws can be obtained by sampling sequentially from:

Step1 : $p\left(\mathbf{U} \mid \mathbf{Y}, \mathbf{A}, \mathbf{\Lambda}\right)$

Step2 : $p\left(\mathbf{A} \mid \mathbf{Y}, \mathbf{U}, \mathbf{\Lambda}\right)$

Step3 : $p\left(\mathbf{\Lambda} \mid \mathbf{Y}, \mathbf{U}, \mathbf{A}\right)$

where $\mathbf{Y}$ denotes the data and $\mathbf{A}$ is the matrix containing all the VAR coefficients. The last two steps are standard and will only briefly be described here. Most of this sub-section will deal with Step 1 which is non-standard due to the orthogonality of $\mathbf{U}$.

To see why the last two steps are standard consider the reduced form VAR defined in (1). Since $\mathbf{U}$ is orthogonal, $\mathbf{U}^{-1} = \mathbf{U}'$ and we can left multiply the VAR by $\mathbf{U}'$ to get

$$\mathbf{U}'\mathbf{y}_t = \mathbf{U}'\mathbf{a} + \mathbf{U}'\mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{U}'\mathbf{A}_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\mathbf{0}, \mathbf{\Lambda}\right). \tag{3}$$

This is a VAR with diagonal error covariance matrix which allows for equation-by-equation estimation. The reduced form VAR coefficients, $\mathbf{A}$, are drawn using the algorithm of Chan et al. (2023). Note that the latter algorithm extends the algorithm of Chan et al. (2022), which is for the case where the Cholesky decomposition is used, to the case where the impact matrix is not restricted to be lower triangular and, thus, is relevant for our BVAR-eig. Chan et al. (2023) show that the computational complexity of the step where the reduced form VAR coefficients are drawn is of the same order as the Cholesky-based algorithm of Chan et al. (2022). Any of the standard VAR priors can be used for $\mathbf{A}$. In this paper, we use the Horseshoe prior (e.g., Carvalho et al., 2010; Cross et al., 2020). The eigenvalues play the role of error variances and standard formulae for their inverse-Gamma conditional posteriors apply when inverse-Gamma priors are used.

The conditional posterior distribution of $\mathbf{U}$ is:

$$p(\mathbf{U} \mid \mathbf{Y}, \mathbf{A}, \mathbf{\Lambda}, \mathbf{D}, \mathbf{V}) = \prod_{i=1}^{n} \mathrm{e}^{-\frac{1}{2}\lambda_i^{-1}u_i'(\mathbf{VDV}'+(\mathbf{Y}-\mathbf{XA})'(\mathbf{Y}-\mathbf{XA}))u_i}. \tag{4}$$

which has the form of a Bingham distribution and seems to suggest that one can simply draw the eigenvectors in the eigenmatrix one at a time from the Bingham distribution.

However, Hoff (2009b) shows that such a strategy will not work since the chain of draws will be reducible. This arises from the fact that $\mathbf{U}$ is orthogonal and, thus, the conditional distribution of $\mathbf{u}_i$ given the other columns of $\mathbf{U}$ has restricted support. Fortunately, Hoff (2009a) proposes a solution to this: draw the eigenvectors two columns at a time.

To see how this algorithm works, let us take the first and second columns of the eigenvectors, $\{\mathbf{u}_1, \mathbf{u}_2\}$, as an example. Conditional on the remaining columns, their distribution is equivalent to $\mathbf{RQ}$ where $\mathbf{R}$ is an orthonormal basis for the nullspace of the remaining columns and $\mathbf{Q}$ is orthogonal with density

$$p(\mathbf{Q}) \propto e^{\mathrm{tr}\left(-\frac{1}{2}\mathbf{\Lambda}_{1,2}^{-1}\mathbf{Q}'\mathbf{KQ}\right)} = \mathrm{e}^{-\frac{1}{2}\left(\lambda_1^{-1}q_1'\mathbf{K}q_1+\lambda_2^{-1}q_2'\mathbf{K}q_2\right)}, \tag{5}$$

where $\mathbf{K} = \mathbf{R}'\big(\mathbf{VDV}' + (\mathbf{Y}-\mathbf{XA})'(\mathbf{Y}-\mathbf{XA})\big)\mathbf{R}$, q_i is the ith column of $\mathbf{Q}$.

Since $\mathbf{Q}$ is orthogonal, it can be parameterized as $\mathbf{Q} = \begin{pmatrix} \cos\phi & q\sin\phi \\ \sin\phi & -q\cos\phi \end{pmatrix}$, for some $\phi \in (0, 2\pi)$ and $q = \pm 1$. The second column q_2 of $\mathbf{Q}$ is a linear function of the first column q_1, and the uniform density on the circle is constant in ϕ, so the joint density of (ϕ, s) is simply $p(\mathbf{Q}(\phi, q))$. Sampling from this distribution can be accomplished by first sampling $\phi \in (0, 2\pi)$ from a density proportional to $p(\mathbf{Q}(\phi, q))$, and then sampling q uniformly from {-1,1}. The density $p(\mathbf{Q}(\phi, q))$ can be obtained from Equation (5), by replacing q_1 and q_2 with $\begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}$ and $\begin{pmatrix} q\sin\phi \\ -q\cos\phi \end{pmatrix}$, respectively.

$$p(\phi) \propto \exp\left\{-\frac{1}{2}\left(\left(\frac{k_{1,1}}{\lambda_1}+\frac{k_{2,2}}{\lambda_2}\right)\cos^2(\phi)+\left(\frac{k_{2,2}}{\lambda_1}+\frac{k_{1,1}}{\lambda_2}\right)\sin^2(\phi)+\left(\frac{k_{1,2}}{\lambda_1}+\frac{k_{2,1}}{\lambda_1}-\frac{k_{1,2}}{\lambda_2}-\frac{k_{2,1}}{\lambda_2}\right)\cos(\phi)\sin(\phi)\right)\right\}$$

where $k_{i,j}$ is the (i,j)th element in matrix $\mathbf{K}$, which is 2×2.

To summarize, the Gibbs sampling scheme for the eigenvectors is as follows: Given $\mathbf{U}^{(j)} = \mathbf{U}$, perform steps $a - e$ for each pair $(n_1, n_2) \subset \{1, \ldots, n\}$ in random order:

Step a: let $\mathbf{R}$ be the null space of $\mathbf{U}_{[,-(n_1,n_2)]}$;

Step b: compute $\mathbf{K} = \mathbf{R}'\big(\mathbf{VDV}' + (\mathbf{Y}-\mathbf{XA})'(\mathbf{Y}-\mathbf{XA})\big)\mathbf{R}$;

Step c: sample $\phi \in (0, 2\pi)$ from the density proportional to $p(\phi)$;

Step d: sample s uniformly from {-1,1};

Step e: set $\mathbf{R} = \mathbf{E}(\phi, q)$ and $\mathbf{U}_{[,(n_1,n_2)]} = \mathbf{RQ}$.

Set $\mathbf{U}^{(j+1)} = \mathbf{U}$.

3 Empirical Illustration: Homoskedastic Models

3.1 Overview and Data Description

We now investigate the properties of our BVAR-eig approach as compared with other homoskedastic VARs. We use a dataset that consists of 20 quarterly US variables with a sample period from 1960Q1 to 2021Q3. It is constructed from the FRED-QD database of the Federal Reserve Bank of St. Louis as described in McCracken and Ng (2016). The dataset contains a range of standard macroeconomic and financial variables, such as real GDP, industrial production, inflation rates, labor market variables and interest rates. In it, the data is transformed to stationarity and there are four lags in all models. The complete list of variables and transformations is given in Appendix C. The variables are ordered in the same manner as they are listed there. Note that macroeconomic variables are ordered first and financial variables ordered last.

In this empirical exercise, we produce iterative forecasts for $h \in \{1, 2, 3, 4\}$-steps-ahead with evaluation period beginning in 1988Q1. To assess forecasting accuracy, we use root mean square forecast errors (RMSFEs) for point forecasts and sums of log predictive likelihoods (log PLs) or their averages (ALPLs) for density forecasts. The full set of forecasting results are available in Appendix D. We summarize these results here using the one-step ahead log PL which is an approximation to the marginal likelihood, see Geweke and Amisano (2010).

3.2 Comparison against Cholesky Approaches

The importance of the ordering of variables in Cholesky-decomposed VARs has been established in papers such as Arias et al. (2023) and Chan et al. (2023). Nevertheless it is useful to reinforce its importance by considering two orderings of the variables. The standard ordering described in the preceding sub-section and the reverse ordering (RO). We refer to these two models as BVAR-chol and BVAR-chol-RO. Theory tells us that the BVAR-eig approach is invariant to ordering so we only present results for the standard ordering for this model. We also include BVAR-IW. This is equivalent to BVAR-eig and, hence, its inclusion sheds light only on MCMC efficiency and computational time.

Table 1 presents log PLs for the four models. We consider two data periods: one is the full sample (1960Q1 to 2021Q3), another is pre-pandemic (1960Q1 to 2018Q4). Both periods establish an important finding: the log PL for the two Cholesky approaches are very different from one another and eigen-decomposed approaches have higher log PL than the Cholesky ones. Thus, at least for this data set, we have evidence that ordering can have a substantive effect in Cholesky-transformed models and that our approach is superior to a standard implementation of a Cholesky-based approach. In addition, the numerical standard errors (NSEs) provide evidence that during the pre-pandemic period, the BVAR-eig produces a log PL that is equivalent to that of the BVAR-IW model.[3] We note that estimating marginal likelihoods using posterior and predictive simulation methods (as we do in this paper for all models) is extremely difficult. In this context, we find the results in Table 1 encouraging where, even with $10,000$ MCMC draws, we are able to obtain results which clearly distinguish between the models.

It is worth stressing that the joint log PL we use to approximate our marginal likelihoods evaluates a 20-dimensional predictive density at each point in the forecast window. This

[3] To compute the NSE, we follow Chan and Eisenstat (2015). We calculate the log PL for $50,000$ draws. Then instead of one single chain, we run 10 parallel chains with each $5,000$ draws. The parallel chains are used to compute the NSE.

requires accurate estimation of the 20 by 20 covariance matrix (all pairwise dependencies) and precise simulation of the full multivariate density, where small errors in off-diagonal terms compound in high dimensions. In contrast, marginal ALPLs (which are presented in Table A-4 of Appendix D) rely on 20 separate 1-dimensional distributions, which are computationally simpler (only means and variances matter) and less sensitive to small errors.

Table 1: Log predictive likelihoods.

	With pandemic				Without pandemic (2018Q4)			
Model	BVAR-IW	BVAR-chol	BVAR-chol-RO	BVAR-eig	BVAR-IW	BVAR-chol	BVAR-chol-RO	BVAR-eig
Log PL	-6,739	-7,410	-7,129	-6,780	-3,282	-3,361	-3,337	-3,286
NSE	26.44	25.99	50.84	41.20	2.10	2.54	2.98	2.39

A complete set of forecasting results for these models for our 20 variables, 4 forecast horizons and two forecast metrics is available in Appendix D. The ALPL results for the individual variables reinforce the patterns found in the marginal likelihoods. In addition, the RMSFEs in Appendix D show that the point forecast performance of all approaches is quite similar. This is consistent with earlier findings in the literature, e.g. Arias et al. (2023), that ordering issues in VARs have little effect point forecasts.[4] Another interesting finding relates to BVAR-chol-RO, which orders the financial variables first and macro variables last and is found to forecast better than BVAR-chol. This superiority is largely due to the ability of the reverse-ordered model to provide better forecasts of the financial variables.

3.3 Further Results with Homoskedastic Models

We have discussed why computational time is an important issue with large VARs and this is confirmed in Table 2. The codes are run using MATLAB on a desktop with an Intel(R) Core(TM) i5-9500 CPU @ 3.00GHz processor and 6 cores. In Table 2 we compare the

[4]We note that our findings relate to unconditional forecasts. Arias et al. (2023) also consider conditional forecasts and find ordering issues to have a more substantial effect on conditional point forecasts.

computation time measured in efficiency units (i.e. by multiplying by the effective sample size so as to produce a measure of time necessary to achieve one i.i.d. draw)[5] of BVAR-eig and BVAR-chol as well as conventional (i.e. not equation-by-equation) estimation using an inverse-Wishart prior. It can be seen that the computation times of BVAR-eig and BVAR-chol are roughly the same. Using the inverse Wishart prior leads to computation time which is almost 10 times as high.

Table 2: The mean computation times (in seconds) to obtain one i.i.d. equivalent draw using the proposed Eigendecomposition method compared to the Cholesky decomposition method. All BVARs have $n = 20$ variables and $p = 4$ lags.

		Computation times	N_{para} [1]
	IW	1.11	5,224
Cholesky decomposition	chol	0.13	5,034
Eigendecomposition	eig-BIG	0.16	5,244

[1] N_{para} is the number of parameters (including hyperparameters) that are estimated in each model.

Appendix D provides additional empirical results illustrating the properties of our approach. First it provides scatter plots of estimates of the elements of the error covariance matrix using the $\mathcal{BIG}$ prior against the inverse-Wishart prior which confirm empirically what we already knew theoretically, that they are equivalent apart from MCMC approximation error. Second, it investigates the impact of the prior in the BVAR-chol model on the ordering issue. Intuitively, since the ordering issue arises due to properties of the prior, if a relatively noninformative prior is used the impact of the ordering issue should be lessened. In addition, the ordering issue should be more substantial in larger models. The results in Appendix D confirm these two conjectures. It considers small (5 variable) and large (20 variable) models, estimated using the standard and reverse ordering, using a BVAR-chol with a relatively non-informative prior. Figure A-1 shows that forecast differ-

[5] The total computational time and effective sample size are based on 10,000 posterior draws. In each model, there are different numbers of parameters to be estimated. So there will potentially be a different computational time in terms of efficiency unit for every parameter. We present the mean of them.

ences (as measured by RMSFE) between the two orders were small (but still present) in the smaller model, but substantial in the larger model. The appendix also presents ALPLs using the relatively non-informative prior and finds them to be lower that those produced using BVAR-eig.

4 Adding SV using the $\mathcal{IBIG}$ prior

In the preceding sections we have assumed the error covariance matrix to be homoskedastic. We showed how the $\mathcal{BIG}$ prior on the eigenvalues and eigenmatrix was equivalent to the inverted-Wishart prior on the reduced form VAR error covariance matrix but allowed for equation-by-equation estimation. Thus, we did not introduce a new model, but rather developed much improved computation for an existing model. In this section, we move beyond the homoskedastic reduced form VAR to allow for the time variation in volatilities which characterizes most macroeconomic and financial data sets and is so necessary for density forecasting.[6] In this section, we show how the eigendecomposition approach can be extended to allow for stochastic volatility in BVARs, leading to the BVAR-eigSV model which is order invariant and allows for equation-by-equation estimation. The introduction of SV is achieved through our development of the $\mathcal{IBIG}$ prior and then allowing for time variation in its eigenvalues.

To introduce the $\mathcal{IBIG}$ prior, note first some restrictive properties of the $\mathcal{BIG}$ prior. The assumption that $\mathbf{B} = -\frac{1}{2}\mathbf{\Lambda}^{-1}$ implies that the prior for $\mathbf{U}$ depends on $\mathbf{\Lambda}$. In other words, the eigenvalues don't simply control the scale of the error covariance matrix Σ, but also influence its orientation via its entry into the prior for the eigenmatrix. In this sense the $\mathcal{BIG}$ prior (or equivalently the inverse-Wishart distribution) is restrictive. But we can relax this assumption by following a suggestion of Hoff (2009b) which is to let $\mathbf{B}$ be a

[6]Note that allowing for time variation in the VAR coefficients is also possible. For instance, it is straightforward to extend the methods in the paper to allow for them to evolve according to random walks.

matrix of prior hyperparameters which are not linked to $\mathbf{\Lambda}$. The $\mathcal{IBIG}$ prior lets $\mathbf{B}$ be a diagonal matrix of parameters which controls the orientation of $\mathbf{U}$, no longer explicitly linked to $\mathbf{\Lambda}$. The role of the latter now solely relates to the scale of the error covariance matrix. Note that $\mathbf{B}$ can be treated as matrix of known prior hyperparameters, but it is also possible to treat it as unknown and estimate it. We adopt the latter strategy in this paper. More specifically, we use a Normal prior for $\mathbf{B}$ and develop Bayesian methods for its estimation.

It is useful to provide some more intuition by considering the two dimensional plane. Any point on it can be denoted by a straight line connecting the origin to the point. The angle between the line and the X-axis and the length of the line determine the location of the point. In the eigendecomposition, the eigenvalues give you the length (scale) of the line and the eigenvectors give you the angle (orientation). The $\mathcal{BIG}$ prior imposes a dependence between the length and the angle that may be undesirable, especially when it comes to building an econometric model where you want one aspect to change but not the other. We will introduce models of parameter change in a subsequent section of this paper. But note that when we use the Cholesky decomposition $\mathbf{C}$, which is the diagonal matrix of error variances, determines the scale. Papers such as Primiceri (2005) add separate, independent, random walk processes for $\mathbf{L}$ and $\mathbf{C}$. And many other papers, following Cogley and Sargent (2005), fix $\mathbf{L}$ and replace $\mathbf{C}$ with a diagonal matrix of SV processes. In this paper we want to do something similar and, thus, it is important to use the $\mathcal{IBIG}$ prior where the scale is independent of the orientation.[7]

The BVAR-eigSV model is

$$\mathbf{y}_t = \mathbf{a} + \mathbf{A}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{A}_p \mathbf{y}_{t-p} + \mathbf{U}\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{\Lambda}_t), \tag{6}$$

[7]It is also worth noting that, in a cross-sectional data context, Hoff (2009b) notes that the Wishart distribution is useful for modelling the covariance of a single population, but is too restrictive with multiple populations.

where $\mathbf{\Lambda}_t = \text{diag}\left(\mathrm{e}^{h_{1,t}}, \ldots, \mathrm{e}^{h_{n,t}}\right)$ is diagonal. Notice that the eigenvalues may vary over time. Each of the log-volatilities follows a stationary AR(1) process:[8]

$$h_{i,t} = \mu_i + \phi_i(h_{i,t-1} - \mu_i) + u^h_{i,t}, \quad u^h_{i,t} \sim \mathcal{N}\left(0, \omega_i^2\right), \tag{7}$$

for $t = 2, \ldots, T$, where $|\phi_i| < 1$ the initial states are specified as $h_{i,1} \sim \mathcal{N}\left(\mu_i, \omega_i^2/\left(1 - \phi_i^2\right)\right)$. This BVAR-eigSV model has the same advantage as BVAR-cholSV in that it allows for equation-by-equation estimation (e.g., Koop et al., 2019; Carriero et al., 2019).

The properties of BVAR-eigSV, in terms of the types of time variation in the error covariance it allows for, are similar to those of the Cholesky-decomposed version of the model which allows for AR(1) evolution of the diagonal elements $\mathbf{C}$. That is, the independence assumption of the $\mathcal{IBIG}$ means that each of the elements of $\mathbf{\Lambda}_t$ is interpreted as a scale (which evolves over time according to an AR process) and does not affect the orientation of the eigenvector. However, both BVAR-eigSV and BVAR-cholSV do have the same restrictive properties noted by Primiceri (2005) in his discussion of the model used in Cogley and Sargent (2005). That is, Cogley and Sargent (2005), and the present paper, use a BVAR-cholSV where $\mathbf{L}$ is constant over time. Primiceri (2005) points out some restrictions of this assumpion. For instance, it implies that the effect of a shock to the i^{th} variable on the j^{th} variable is constant over time. Primiceri (2005) relaxes this assumption by allowing for the free elements of $\mathbf{L}$ to have AR(1) (or random walk) behavior. Allowing for AR behavior for the elements of the eigenmatrix $\mathbf{U}$ is not sensible since it is an orthogonal matrix. In sub-section 5.6 we discuss how time variation can be introduced in $\mathbf{U}$.

We also develop methods for prior hyperparameter estimation for the prior on the eigenmatrix. In the Bayesian VAR literature, it is increasingly common to estimate prior

[8]Allowing the errors in the volatility processes to be correlated with one another would be a simple extension of our model. Alternatively, a factor structure could be used to obtain parsimony as in factor stochastic volatility models (e.g., Pitt and Shephard, 1999; Chib et al., 2006; Kastner and Huber, 2020; Chan, 2023).

hyperparameters. For instance, Giannone et al. (2015) develops methods for estimating the optimal degree of prior shrinkage in Bayesian VARs. This approach relies on the marginal likelihood, which has a closed form expression for the homoskedastic VAR with natural conjugate prior. Once SV has been added (or non-conjugate priors are used) as in the BVAR-cholSV and BVAR-eigSV, the marginal likelihood no longer has a closed form expression and can be very difficult to calculate. Another method of prior hyperparameter selection is to adopt a hierarchical prior structure which treats the hyperparameters as unknown parameters. A good example is Amir-Ahmadi et al. (2020) which uses this strategy in the context of the time-varying parameter VAR. Here we focus on hyperparameter estimation for the prior for the eigenmatrix $\mathbf{U}$. This matrix is constant over time but we adopt a similar hierarchical prior strategy. As we shall see, the conditional posteriors which result have simple forms which leads to a straightforward Gibbs sampling algorithm.

Relative to our earlier BVAR-eig, the new aspects of the MCMC algorithm relate to $\mathbf{\Lambda}_t$ as well as $\mathbf{U}$ and its prior hyperparameters: $\mathbf{B}$, $\mathbf{D}$ and $\mathbf{V}$. Accordingly we will only discuss the MCMC steps for these, with other steps being unchanged. $\mathbf{\Lambda}_t$ can also be easily dealt with since we use standard methods, in particular we implement the auxiliary mixture sampler of Kim et al. (1998) in conjunction with the precision sampler of Chan and Jeliazkov (2009).

The prior for $\mathbf{U}$, $p(\mathbf{U} \mid \mathbf{D}, \mathbf{B}, \mathbf{V})$, takes the Bingham form given in (2). The hyperparameters in this prior, $\mathbf{B}$, $\mathbf{D}$ and $\mathbf{V}$, are treated as unknown parameters. For reasons to be explained, we do not directly place a prior on these parameters, but work with transformations of them. The prior on the transformed parameters is given in the algorithm.

Noting that, conditional on $\mathbf{U}$, the data provides no additional information for $\mathbf{B}$, $\mathbf{D}$ and $\mathbf{V}$, the full conditional posterior distributions simplify and our MCMC algorithm involves the following distributions:

Step1 : $p\left(\mathbf{U} \mid \mathbf{Y}, \mathbf{A}, \mathbf{\Lambda}, \mathbf{D}, \mathbf{B}, \mathbf{V}\right)$

Step2 : $p(\mathbf{B}, \mathbf{D} \mid \mathbf{U}, \mathbf{V})$

Step3 : $p(\mathbf{V} \mid \mathbf{U}, \mathbf{D}, \mathbf{B})$

which we discuss in turn in this sub-section.

Step 1 requires only minor modification relative to the homoskedastic case.

Step1 : $p(\mathbf{U} \mid \mathbf{Y}, \mathbf{A}, \mathbf{\Lambda}, \mathbf{D}, \mathbf{B}, \mathbf{V})$

$$p(\mathbf{U} \mid \mathbf{Y}, \mathbf{A}, \mathbf{\Lambda}, \mathbf{D}, \mathbf{B}, \mathbf{V}) \propto \prod_{i=1}^{n} \mathrm{e}^{u_i'\left(b_i \mathbf{V}\mathbf{D}\mathbf{V}' + \sum_{t=1}^{T} -\frac{1}{2}\lambda_{i,t}^{-1}(\mathbf{y}_t' - \mathbf{x}_t'\mathbf{A})'(\mathbf{y}_t' - \mathbf{x}_t'\mathbf{A})\right)u_i},$$

which is a Bingham distribution. We use the same strategy of drawing two columns of $\mathbf{U}$ at a time that we used for the homoskedastic case with $\mathcal{IBIG}$ prior.

Step2 : $p(\mathbf{D}, \mathbf{B} \mid \mathbf{U}, \mathbf{V})$

Since $p(\mathbf{D}, \mathbf{B} \mid \mathbf{U}, \mathbf{V}) \propto p(\mathbf{U} \mid \mathbf{D}, \mathbf{B}, \mathbf{V})\, p(\mathbf{D}, \mathbf{B} \mid \mathbf{V})$ the Bingham prior for $\mathbf{U}$ is the key component of this conditional posterior. Note, however, that it involves an integrating constant, $c(\mathbf{D}, \mathbf{B})$ which could be ignored in our earlier derivations of the posterior for $\mathbf{U}$, but cannot be ignored in Step 2. The resulting conditional posterior is no longer of a convenient form. Accordingly, we follow an approximate strategy suggested in Hoff (2009b). One aspect of this strategy is to use an approximation to the integrating constant, $\widetilde{c}(\mathbf{D}, \mathbf{B})$ which is given in equation (8) of Hoff (2009b). But this approximation causes an identification issue since it implies the scales of $\mathbf{D}$ and $\mathbf{B}$ are not separately identifiable. Accordingly, Hoff (2009b) reparameterizes the diagonal matrices $\mathbf{D}$ and $\mathbf{B}$ as having diagonal elements ordered and bounded between zero and one times a scalar constant:

$$\mathrm{diag}(\mathbf{D}) = (d_1, \ldots, d_n) = \sqrt{w}\,(\theta_1, \ldots, \theta_n)\,,$$

$$\mathrm{diag}(\mathbf{B}) = (b_1, \ldots, b_n) = \sqrt{w}\,(\beta_1, \ldots, \beta_n)\,,$$

where $w > 0, 1 = \theta_1 > \theta_2 > \ldots > \theta_{n-1} > \theta_n = 0$ and $1 = \beta_1 > \beta_2 > \ldots > \beta_{n-1} > \beta_n = 0$. Note that these restrictions remove the label switching problem which occurred with the $\mathcal{BIG}$ prior since they provide a specific ordering for the columns of the eigenmatrix.

This leads to

$$p(\mathbf{U} \mid \theta, \beta, w, \mathbf{V}) \propto w^{m/2} e^{(-w\theta'(\mathbf{I}-\mathbf{M})\beta)} \prod_{i<j} (\theta_i - \theta_j)^{1/2} (\beta_i - \beta_j)^{1/2}, \tag{8}$$

where $m = \begin{pmatrix} n \\ 2 \end{pmatrix} = \frac{n(n-1)(n-2)\cdots(n-2+1)}{2!}$, the matrix $\mathbf{M} = (\mathbf{V}'\mathbf{U}) \circ (\mathbf{V}'\mathbf{U})$, and $\circ$ is the Hadamard product operator denoting elementwise multiplication.

It can be seen that we have a Gamma density for w (conditional on other parameters) which, combined with a Gamma prior leads to a conditional posterior which is Gamma. The Gamma prior we use is

$$w \sim \mathcal{G}\left(\eta_0/2, \tau_0^2)\right),$$

where $\mathcal{G}$ denotes the Gamma distribution. For the hyperparameters, we follow Hoff (2009b) and set $\eta_0 = 2$, $\tau_0^2 = 10$. This choice will result in an exponential prior distribution that has its mode at $w = 0$ but is very diffuse. Thus our prior has a mode at eigenvector homogeneity but the diffuse feature means we are allowing for a large range of possible values for eigenvector heterogeneity.

With regard to the prior on $p(\theta, \beta)$, it is taken to be such that $1 > \theta_2 > \ldots > \theta_{n-1} > 0$ and $1 > \beta_2 > \ldots > \beta_{n-1} > 0$ are two independent sets of order statistics of uniform random variables on $[0, 1]$. Multiplying this prior by (8) leads to a conditional posterior that can easily be sampled from on the grid $[0, 1]$.

To summarize the algorithm to update w, θ and β involves:

Step a: sample w from a Gamma distribution;

Step b: for each $i \in \{2, \ldots, n-1\}$ sample $\theta_i \in (\theta_{i-1}, \theta_{i+1})$ from the density proportional to

$$e^{-\theta_i\left(w\beta'\mathbf{M}_{[i,]}\right)}\prod_{j:j\neq i}\left|\theta_i-\theta_j\right|^{1/2}.$$

Step c: for each $j \in \{2,\ldots,n-1\}$ sample $\beta_j \in (\beta_{j-1}, \beta_{j+1})$ from the density proportional to

$$e^{-\beta_j\left(w\mathbf{M}'_{[,j]}\theta\right)}\prod_{i:i\neq j}\left|\beta_j-\beta_i\right|^{1/2}.$$

Step3 : $p\left(\mathbf{V} \mid \mathbf{U},\mathbf{D},\mathbf{B}\right)$

Since $p\left(\mathbf{V} \mid \mathbf{U},\mathbf{D},\mathbf{B}\right) \propto p\left(\mathbf{U} \mid \mathbf{D},\mathbf{B},\mathbf{V}\right)p\left(\mathbf{V} \mid \mathbf{D},\mathbf{B}\right)$ the Bingham prior for $\mathbf{U}$ again plays a key role. However, since the integrating constant, $c(\mathbf{D},\mathbf{B})$ does not depend on $\mathbf{V}$, the complications of Step 2 do not arise. Assuming the prior distribution for $\mathbf{V}$ to be the uniform (invariant) measure on $\mathcal{O}_n$ we obtain

$$\begin{aligned} p(\mathbf{V} \mid \mathbf{U},\mathbf{D},\mathbf{B}) &\propto \mathrm{e}^{\mathrm{tr}(\mathbf{BU'VDV'U})} \\ &= \mathrm{e}^{\mathrm{tr}(\mathbf{DV'UBU'V})}, \end{aligned} \tag{9}$$

which is a Bingham distribution. We use the same strategy of drawing two columns of $\mathbf{V}$ at a time as we do when drawing $\mathbf{U}$.

5 Macroeconomic Forecasting Using the BVAR-eigSV

5.1 Summary and Overview

In this section we carry out a forecasting exercise using the 20 variable data set described in Section 3. The forecast exercise is set up in the same way as in that section except that now all of our models include SV. In addition to our BVAR-eigSV and BVAR-cholSV, we consider two other order-invariant approaches. These are the Order Invariant SV (BVAR-OISV) model proposed in Chan et al. (2023) and the common stochastic volatility (BVAR-

CSV) model introduced in Carriero et al. (2016). All information about priors is available in Appendix B.

Before presenting forecasting results, we present evidence on computation times and on the flexibility of BVAR-eigSV in modeling the time varying error covariance matrix.

5.2 Computation Times

To our knowledge, the fastest available algorithm for the Cholesky based approach is that of Chan et al. (2022). It has computational complexity of $O(n^4)$ as opposed to the $O(n^6)$ of full system estimation of the reduced form VAR. Hence, we compare our methods to the fastest algorithm available instead of to full system estimation which has already been established as being very slow. Chan et al. (2023) develops an extension of this algorithm for use with BVAR-OISV which is slightly slower. Carriero et al. (2016) is a more restricted model, involving only one SV process, so it is expected to be much faster.

We report the computational times (in seconds) to obtain one i.i.d. equivalent draw in Table 3. The computational burden of BVAR-eigSV is larger than that of BVAR-cholSV and BVAR-CSV but faster than that of BVAR-OISV. The main reason for the increase in computational burden with the eigendecomposition approach is that the effective sample size for the volatilities is lower than that of the volatilities in BVAR-cholSV. If we use the effective sample size only for the constant parameters in the model, we find computation times for BVAR-eigSV and BVAR-cholSV to be approximately the same. Thus, the computational burden of BVAR-eigSV is not that great even in large models and is less than that of another main order-invariant approach, BVAR-OISV.

Table 3: The mean computation times (in seconds) to obtain one i.i.d. equivalent draw using the proposed Eigendecomposition method compared to the Cholesky decomposition method. All BVARs have $n = 20$ variables and $p = 4$ lags.

		Computation times	N_{para}[1]
Cholesky decomposition	cholSV	0.68	10,296
Eigendecomposition	eigSV	12.73	10,563
Non-decomposition	OISV	37.79	10,926
	CSV	0.03	2,286

[1] N_{para} is the number of parameters (including hyperparameters) that are estimated in each model.

5.3 Investigating the Properties of the Time-Varying Error Covariance Matrix

Stochastic volatility models such as BVAR-cholSV have well known properties that have been found empirically useful in macroeconomics. Since our model implies the elements of $\mathbf{\Lambda}_t$ are interpreted as scales, we would expect our BVAR-eigSV to be capable of modelling these properties. To investigate this empirically, Figure 1 plots the time-varying elements of an important block of the error covariance matrix for BVAR-eigSV, BVAR-cholSV and BVARcholSV-RO.[9] The block involves industrial production, the unemployment rate, PCE inflation, and the Federal funds rate.

The three lines in each panel of Figure 1 tend to be broadly similar to one another. However, there are appreciable differences particularly in the covariances and the way volatility spikes are handles. It is interesting to note that the differences between BVAR-cholSV and BVAR-cholSV-RO are just as substantive as the differences between either of the Cholesky-based models and BVAR-eigSV. In fact it is often the case that the estimates produced by BVAR-eigSV lie between the two Cholesky based approaches. This is yet more evidence that the choice of ordering matters in large VARs when the Cholesky decomposition is

[9] This figure excludes data for the pandemic period to make the comparison clearer. Figures based on the full sample and for small VARs are available in Appendix E.

used, thus reinforcing the empirical evidence of Hartwig (2020). The main difference is that BVAR-eigSV is often (but not always) finding larger changes in variances or covariances in unstable times such as the financial crisis or the late 1970s suggesting it is better at picking up abrupt changes in volatility than the Cholesky-based approach.

Posterior means are plotted in Figure 1. Appendix E2 produces versions of this figure which contain credible intervals. These suggest that the SV processes for BVAR-eigSV and BVAR-cholSV are estimated to a similar degree of accuracy.

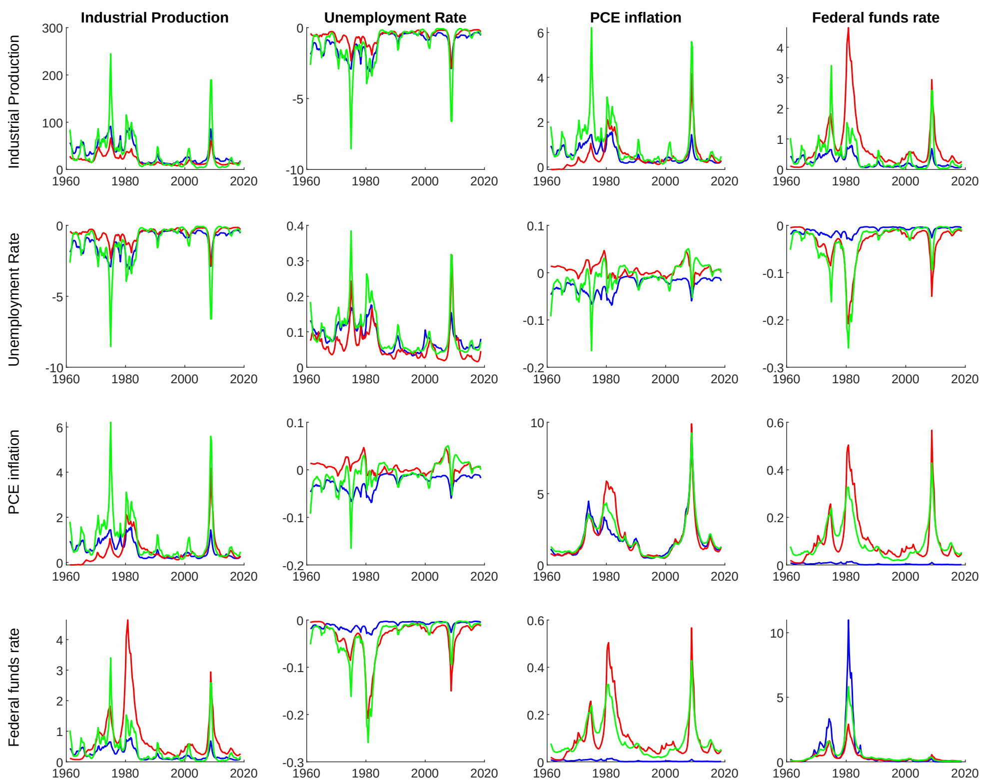

Figure 1: Estimates (posterior means) of elements of time-varying covariance matrix. BVAR-cholSV(blue),BVAR-cholSV-RO(red),BVAR-eigSV(green).

5.4 Results of the Forecasting Exercise

In this subsection, we report results for our four models for different forecast horizons. Table 4 contains the joint ALPL for all 20 variables. At all horizons, the parsimonious

BVAR-CSV approach consistently outperforms all other methods. But BVAR-eigSV is performing better than either of the Cholesky based approaches at short horizons and its forecast metrics lie between the two Cholesky based approaches at longer horizons.

A comparison of results for BVAR-cholSV and BVAR-cholSV-RO shows the sensitivity of ordering. It is interesting to note that at short horizons the original ordering of the variables forecasts better than the reverse ordering. But at longer horizons this finding is reversed. The eigendecomposition approach, in contrast, appears to offer an intermediate solution, mitigating some of the extreme sensitivity observed under alternative Cholesky orderings. The BVAR-OISV model typically exhibits the poorest forecast performance.

Table 4: Joint ALPL for 20 macroeconomic variables.

Models	$h = 1$	$h = 2$	$h = 3$	$h = 4$
BVAR-CSV	**-4,360**	**-8,533**	**-11,118**	**-11,254**
BVAR-cholSV	-4,834	-10,358	-12,997	-16,126
BVAR-cholSV-RO	-6,350	-10,571	-12,921	-14,964
BVAR-eigSV	-5,501	-10,198	-13,223	-15,260
BVAR-OISV	-5,640	-11,269	-16,448	-20,056

Figure 2 summarizes the forecasting performance for each individual variable. The left panel denotes the value of percentage gains in RMSFEs and ALPLs, while the right panel denotes the significance level according to the Diebold Mariano test.[10] The benchmark for comparison is always the BVAR-CSV model with 20 variables. Each row corresponds to a specific model: BVAR-cholSV, BVAR-cholSV-RO, and BVAR-eigSV. Green/red means a model provides more/less accurate forecasts. Intensity of color indicates the degree of difference. White/pale means the difference is not significantly different from zero.

By comparing the right panels, we further emphasize that point forecasts show relatively small differences across models, while density forecasts exhibit more pronounced variations. Within density forecasts, we are finding that the financial variables play an important role in

[10] A complete table is available in Appendix E.

generating the differences in forecasting performance across models. For the macroeconomic variables there are few statistically significant differences in forecast performance.

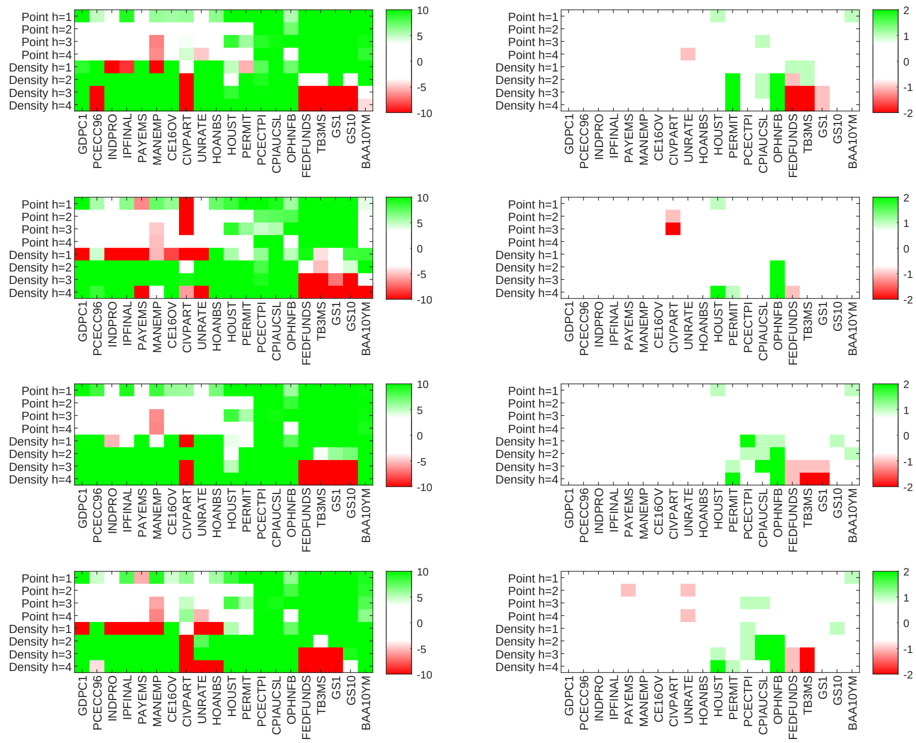

Figure 2: Forecasting results from the large BVAR-SV. From Top to Bottom: BVAR-cholSV against BVAR-CSV, BVAR-cholSV-RO against BVAR-CSV, BVAR-eigSV against BVAR-CSV, BVAR-OISV against BVAR-CSV. Left: values of percentage gains in RMSFEs and ALPLs. Right: significance level according to the Diebold Mariano test: Value 0 means not significant. Value 1 means 0.10 significance level for a two-sided Diebold and Mariano(1995) test. Value 2 means 0.05 significance level.

We are finding that the BVAR-CSV performs well jointly even if it does not always forecast well marginally for the individual variables. This is largely due to the pandemic period. In particular, the pandemic induced synchronized volatility spikes across macro and financial variables. BVAR-CSVs common volatility process captured this co-movement well in a parsimonious fashion, improving the joint ALPL. Pre-pandemic, volatility was less synchronized. In this period, the BVAR-CSV does not perform as well jointly because it forces all variables to share a single volatility process, missing cross-variable differences. However, it works reasonably well for macro variables (which have similar volatility patterns) but fails for financial variables (which have distinct, erratic volatility). These findings are supported in Appendix E.6 which presents forecasting results for the pre-pandemic period.

5.5 A scale-adjusted version of BVAR-eigSV (BVAR-SeigSV)

A potential challenge when using the eigendecomposition is that variables exhibiting high volatility (or variance) can disproportionately influence the covariance matrix, leading to their dominance in the principal components (eigenvectors). Rather than relying on data standardization, we propose a scale-adjusted BVAR-eigSV model (BVAR-SeigSV) to mitigate the issue of scale sensitivity.

The BVAR-SeigSV model is

$$\mathbf{y}_t = \mathbf{a} + \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{S}\mathbf{U}\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}_t), \tag{10}$$

where $\mathbf{S}$ is diagonal $\mathbf{S} = \text{diag}(S_1, \ldots, S_n)$. Instead of estimating the n parameters in $\mathbf{S}$, we use a standardization commonly used when doing principal components analysis. Specifically, let σ_i denote the sample standard deviation of the variable $i, i = 1, \ldots, n$. We set $S_i = \delta\sigma_i$ and use one scaling parameter δ. A similar idea has been used in Amir-Ahmadi et al. (2020) where they group parameters into several groups and estimate one parameter for each group. An inverse gamma prior $\mathcal{IG}(10, 9)$ is imposed on parameter δ such that the prior mean is 1 with 0.1 variance.

Table 5 and Figure 3 present results from our forecasting exercise including this scale-adjusted model. Table 1 shows that overall, and at all horizons, the scale adjusted version of eigendecomposition approach consistently outperforms all other methods including BVAR-CSV. Figure 3 shows that these forecast improvements are mostly for the financial variables and for the density forecasts. This illustrates the potential importance of adjusting for scale.

Appendix E3 presents additional evidence on the importance of scale adjustment. It compares volatility estimates from two BVAR-eigSV models, both not adjusted for scale, for several key variables. One of the models uses all 20 variables, the other uses 19 variables dropping the most volatile one. The figures reveal non-negligible differences between the volatility estimates showing how the inclusion or exclusion of a highly volatile variable can impact the estimated volatilities for all variables.

Table 5: Joint ALPL for 20 macroeconomic variables.

Models	$h = 1$	$h = 2$	$h = 3$	$h = 4$
BVAR-CSV	-4,360	-8,533	-11,118	-11,254
BVAR-eigSV	-5,501	-10,198	-13,223	-15,260
BVAR-SeigSV	**-4,312**	**-7,191**	**-9,380**	**-10,884**

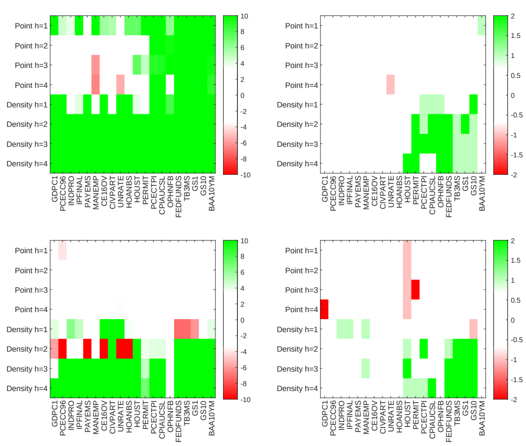

Figure 3: Forecasting results from the large BVAR-SV. Left: percentage gains in point forecast (RMSFEs). Right: percentage gains in density forecast (ALPLs). Upper: BVAR-SeigSV against BVAR-CSV. Lower: BVAR-SeigSV against BVAR-eigSV. Left: values of percentage gains in RMSFEs and ALPLs. Right: significance level according to the Diebold Mariano test: Value 0 means not significant. Value 1 means 0.10 significance level for a two-sided Diebold and Mariano(1995) test. Value 2 means 0.05 significance level.

5.6 Adding Time Variation in the Eigenmatrix U

So far all the models in this paper have assumed **U** to be constant over time. There are various ways we could allow for time-variation in **U** as well. Simply allowing for random walk behavior of its elements is undesirable given that **U** is orthogonal. However, the model and algorithm developed in Hoff (2009b) does allow for heterogeneous groups of error covariances. It is a cross-sectional paper and the groups are simply different known groups of individuals. In a time series context such groups could be different known regimes each having a constant eigenmatrix. For instance, if a structural break was known to occur at a particular point in time the two groups would be observations before and after the break. The extension to unknown groups (e.g. a structural break model with unknown break date) would involve adding another block to the MCMC algorithm for drawing the break date. Thus, handling various structural break or regime change models would be straightforward.

But there are also ways of adding gradual time variation that are similar in spirit to the random walk or AR time variation used by Primiceri (2005) in the Cholesky-based version of the model. These draw on theory from the directional statistics literature motivated by the relationship between the eigenmatrix and orientations and angles. For instance, Daniels and Kass (1999) and Pourahmadi (2005) reparameterize the $n \times n$ orthogonal matrix by its $\frac{n(n-1)}{2}$ Givens angles $\theta_t = (\theta_{21t}, \theta_{31,t}, \cdots, \theta_{n,n-1,t})$ and then write a first-order difference equation for $\{\theta_t\}$. Related to this, a similar issue arises in the time varying cointegration literature. In cointegrated models, the matrix of long run multipliers in the Vector Error Correction model involves a reduced rank structure and is decomposed into two parts. One of these is an orthogonal matrix. Allowing for time varying cointegration implies this orthogonal matrix is time varying. Koop et al. (2011) has proposed a way to add gradual time variation in this orthogonal matrix while still retaining its orthogonality. This approach could also be used to model gradual time variation in **U**.

Our BVAR-eigSV imposes a constant impact matrix which as noted, e.g., by Primiceri (2005) imposes restrictions on the manner which the error covariance matrix evolves. This raises the empirical question: are these restrictions supported by the data? To address this,

Figure 5 compares the same block of the error covariance matrix for two order-invariant models: one featuring a constant impact matrix but time varying volatilities (BVAR-SeigSV, in blue) and the other incorporating the time-varying impact version of the BVAR-OISV proposed by Chan et al. (2023) (in red). The shaded regions represent the 84% credible intervals.[11]

The two models are producing broadly similar estimates of the the error variances and covariances. For instance, for the error variances, they both exhibit increases at the expected places (e.g. during the financial crisis and the inflationary 1970s and inflation fighting early 1980s) and are all fairly flat during the Great Moderation period. The error covariances also tend to exhibit changes at these places. But the BVAR-eigSV is better at picking up abrupt changes. Another finding is that the model with constant impact matrix produces tighter credible intervals. This suggests that the SV processes in BVAR-eigSV are quite flexible and adding the flexibility of a time varying impact matrix adds little.

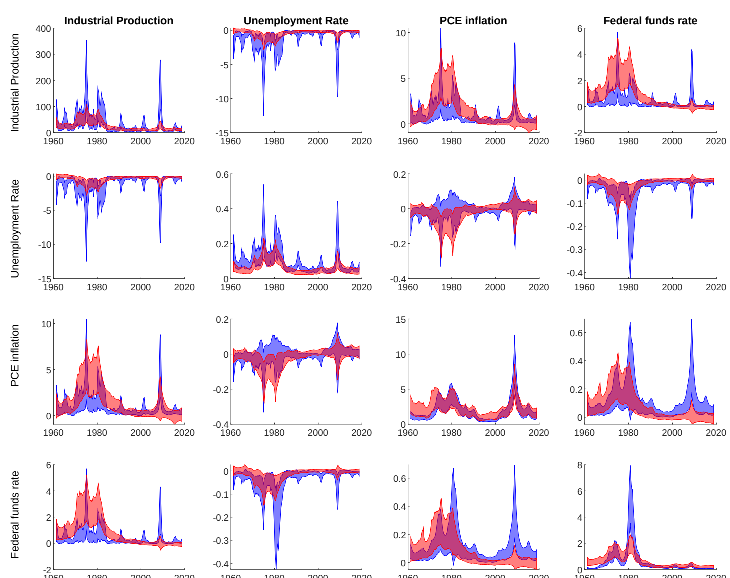

Figure 4: 84% credible intervals of elements of time-varying covariance matrix (excluding data for the pandemic period). Blue: BVAR-eigSV. Red: BVAR-OISV with TVP impact.

[11]In Appendix E5, we also consider the BVAR-cholSV with a time-varying impact matrix and compare large models against small models.

Table 6 compares the forecasting performance without/with time-varying impact matrix. We find models without time-varying impact matrix tend to forecast better at shorter horizons, although at longer horizons BVAR-OISV(TVP) does forecast best.

Table 6: Joint ALPL for 20 macroeconomic variables

Models	$h = 1$	$h = 2$	$h = 3$	$h = 4$
BVAR-eigSV	**-5,501**	**-10,198**	-13,223	-15,260
BVAR-OISV	-5,640	-11,269	-16,448	-20,056
BVAR-OISV(TVP)	-6,228	-10,867	**-12,115**	**-14,580**

6 Summary and Conclusions

In this paper, we have developed an order-invariant Bayesian approach to VAR estimation based on the eigendecomposition of the error covariance matrix. We have also extended this model to alow for stochastic volatility and developed a scale-adjusted version it. There are many other Bayesian VAR approaches and, of course, the choice between these should depend on empirical and computational considerations. But we note that, relative to full system approaches (e.g. based on an inverse Wishart prior), our approach has substantial computational advantages since it allows for equation by equation estimation. Relative to approaches which are not order invariant (e.g. those based on the Cholesky decomposition) our approach has the advantage that it is order invariant. Relative to some of the other order invariant approaches discussed in Arias et al. (2023), our approach has computational advantages in that posterior simulation is done using a straightforward Gibbs sampler and does not involve more sophisticated MCMC methods (e.g. slice sampling) which may fail to scale well. Relative to the order invariant approach of Chan et al. (2023), the difference with our approach lies mainly in identification. That is, our BVAR-eigSV is always identified whereas Chan et al. (2023) is only identified if there $n - 1$ distinct volatility processes in the model.

References

P. Amir-Ahmadi, C. Matthes, and M.-C. Wang. Choosing prior hyperparameters: With applications to time-varying parameter models. *Journal of Business & Economic Statistics*, 38(1):124–136, 2020.

J. E. Arias, J. F. Rubio-Ramirez, and M. Shin. Macroeconomic forecasting and variable ordering in multivariate stochastic volatility models. *Journal of Econometrics*, 235(2): 1054–1086, 2023.

M. Banbura, D. Giannone, and L. Reichlin. Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92, 2010.

C. Bingham. An antipodally symmetric distribution on the sphere. *The Annals of Statistics*, pages 1201–1225, 1974.

A. Carriero, T. E. Clark, and M. Marcellino. Common drifting volatility in large Bayesian VARs. *Journal of Business & Economic Statistics*, 34(3):375–390, 2016.

A. Carriero, T. E. Clark, and M. Marcellino. Large Bayesian Vector Autoregressions with stochastic volatility and non-conjugate priors. *Journal of Econometrics*, 212(1):137–154, 2019.

C. M. Carvalho, N. G. Polson, and J. G. Scott. The horseshoe estimator for sparse signals. *Biometrika*, 97(2):465–480, 2010.

J. Chan, A. Carriero, T. E. Clark, and M. Marcellino. Corrigendum to: Large Bayesian Vector Autoregressions with stochastic volatility and non-conjugate priors. *Journal of Econometrics*, 227(2):506–512, 2022.

J. C. Chan. Minnesota-type adaptive hierarchical priors for large Bayesian VARs. *International Journal of Forecasting*, 37(3):1212–1226, 2021.

J. C. Chan. Comparing stochastic volatility specifications for large Bayesian VARs. *Journal of Econometrics*, 235(2):1419–1446, 2023.

J. C. Chan and E. Eisenstat. Marginal likelihood estimation with the cross-entropy method. *Econometric Reviews*, 34(3):256–285, 2015.

J. C. Chan and I. Jeliazkov. Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(1-2):101–120, 2009.

J. C. Chan, G. Koop, and X. Yu. Large order-invariant Bayesian VARs with stochastic volatility. *Journal of Business & Economic Statistics*, pages 1–13, 2023.

J. C. C. Chan. Asymmetric conjugate priors for large bayesian vars. *Quantitative Economics*, 13(3):–, 2022. URL `https://onlinelibrary.wiley.com/doi/abs/`.

S. Chib, F. Nardari, and N. Shephard. Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics*, 134(2):341–371, 2006.

T. Cogley and T. J. Sargent. Drifts and volatilities: monetary policies and outcomes in the post wwii us. *Review of Economic Dynamics*, 8(2):262–302, 2005.

J. L. Cross, C. Hou, and A. Poon. Macroeconomic forecasting with large Bayesian VARs: Global-local priors and the illusion of sparsity. *International Journal of Forecasting*, 36 (3):899–915, 2020.

M. J. Daniels and R. E. Kass. Nonconjugate Bayesian estimation of covariance matrices and its use in hierarchical models. *Journal of the American Statistical Association*, 94 (448):1254–1263, 1999.

S. Frühwirth-Schnatter. Markov Chain Monte Carlo estimation of classical and dynamic switching and mixture models. *Journal of the American Statistical Association*, 96(453): 194–209, 2001.

J. Geweke and G. Amisano. Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting*, 26(2):216–230, 2010. ISSN 0169-2070. doi: https://doi.org/10.1016/j.ijforecast.2009.10.007. URL `https://www.sciencedirect.com/science/article/pii/S0169207009001757`. Special Issue: Bayesian Forecasting in Economics.

D. Giannone, M. Lenza, and G. E. Primiceri. Prior selection for Vector Autoregressions. *The Review of Economics and Statistics*, 97(2):436–451, 2015.

B. Hartwig. Robust inference in time-varying structural VAR models: The dc-cholesky multivariate stochastic volatility model. *Deutsche Bundesbank Discussion Paper*, (34/2020), 2020.

P. D. Hoff. Simulation of the matrix Bingham-von Mises-Fisher distribution, with applications to multivariate and relational data. *Journal of Computational Graphics and Statistics*, 18(2):438–456, 2009a.

P. D. Hoff. A hierarchical eigenmodel for pooled covariance estimation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(5):971–992, 2009b.

G. Kastner and F. Huber. Sparse Bayesian Vector Autoregressions in huge dimensions. *Journal of Forecasting*, 39(7):1142–1165, 2020.

C. Khatri and K. V. Mardia. The von Mises–Fisher matrix distribution in orientation statistics. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1): 95–106, 1977.

S. Kim, N. Shephard, and S. Chib. Stochastic volatility: likelihood inference and comparison with ARCH models. *The review of economic studies*, 65(3):361–393, 1998.

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

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

G. Koop, R. Leon-Gonzalez, and R. W. Strachan. Bayesian inference in a time varying cointegration model. *Journal of Econometrics*, 165(2):210–220, 2011.

G. Koop, D. Korobilis, and D. Pettenuzzo. Bayesian compressed Vector Autoregressions. *Journal of Econometrics*, 210(1):135–154, 2019.

D. Korobilis. VAR forecasting using Bayesian variable selection. *Journal of Applied Econometrics*, 28(2):204–230, 2013.

B. P. Levy and H. F. Lopes. Dynamic ordering learning in multivariate forecasting. *arXiv preprint arXiv:2101.04164*, 2021.

M. K. Pitt and N. Shephard. Filtering via simulation: Auxiliary particle filters. *Journal of the American statistical association*, 94(446):590–599, 1999.

M. Pourahmadi. Simultaneous modelling of covariance matrices: glm, Bayesian and non-parametric perspectives, 2005.

G. E. Primiceri. Time varying structural Vector Autoregressions and monetary policy. *The Review of Economic Studies*, 72(3):821–852, 2005.

C. A. Sims. Macroeconomics and reality. *Econometrica*, pages 1–48, 1980.

P. Wu and G. Koop. Estimating the ordering of variables in a VAR using a Plackett–Luce prior. *Economics Letters*, 230:111247, 2023. ISSN 0165-1765. doi: https://doi.org/10.1016/j.econlet.2023.111247. URL https://www.sciencedirect.com/science/article/pii/S0165176523002720.

Appendices - for online publication only

A Proof of Order Invariance

It is well-known that Bayesian inference in the VAR with inverse-Wishart prior on the error covariance matrix is order-invariant. Given the equivalence between our $\mathcal{BIG}$ prior and the inverse-Wishart prior it would have sufficed to simply state this fact to establish order-invariance of our approach. However, in the interest of theoretical rigor and for other priors involving the eigendecomposition such as the $\mathcal{IBIG}$ prior, in this appendix we provide a proof of order invariance.

A.1 Original and New parameters

Consider Equation 1, which is a standard homoscedastic VAR of order p:

$$\mathbf{y}_t = \mathbf{a} + \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{U}\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}).$$

Suppose we permute the order of the dependent variables. More precisely, let $\mathbf{P}$ denote an arbitrary permutation matrix of dimension n. By left multiplying matrix $\mathbf{P}$, we get:

$$\mathbf{P}\mathbf{y}_t = \mathbf{P}\mathbf{b} + \mathbf{P}\mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{P}\mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{P}\mathbf{U}\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}).$$

We define $\widetilde{\mathbf{y}}_t = \mathbf{P}\mathbf{y}_t$, $\widetilde{\mathbf{b}} = \mathbf{P}\mathbf{b}$, $\widetilde{\mathbf{A}}_1 = \mathbf{P}\mathbf{A}_1, \ldots, \widetilde{\mathbf{A}}_p = \mathbf{P}\mathbf{A}_p$, then we can get the new VAR(p)
:

$$\widetilde{\mathbf{y}}_t = \widetilde{\mathbf{b}} + \widetilde{\mathbf{A}}_1\mathbf{y}_{t-1} + \cdots + \widetilde{\mathbf{A}}_p\mathbf{y}_{t-p} + \mathbf{P}\mathbf{U}\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}). \tag{11}$$

It is clear that if we permute the order of the dependent variables via $\widetilde{\mathbf{y}}_t = \mathbf{P}\mathbf{y}_t$, new coefficient matrices ($\widetilde{\mathbf{b}}$ and $\widetilde{\mathbf{A}}_j$) can be obtained by permuting the rows and columns of original coefficient matrices accordingly.

Next, we show evidence that the error covariance matrix is also permuted. The error covariance matrix in Equation 1 is

$$\Sigma = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}' = \left(\mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\right)\left(\mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\right)'.$$

The error covariance matrix in Equation (11) is

$$\widetilde{\Sigma} = \mathbf{P}\mathbf{U}\boldsymbol{\Lambda}(\mathbf{P}\mathbf{U})' = \left(\mathbf{P}\mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{P}'\right)\left(\mathbf{P}\mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{P}'\right)' = \left(\widetilde{\mathbf{U}}\widetilde{\boldsymbol{\Lambda}}^{\frac{1}{2}}\right)\left(\widetilde{\mathbf{U}}\widetilde{\boldsymbol{\Lambda}}^{\frac{1}{2}}\right)'.$$

The second equality is because all permutation matrices are orthogonal matrices, we have $\mathbf{P}^{-1} = \mathbf{P}'$. New matrices $\widetilde{\mathbf{U}} = \mathbf{P}\mathbf{U}$, $\widetilde{\boldsymbol{\Lambda}}^{\frac{1}{2}} = \boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{P}'$ mean that the old eigenvector and eigenvalue matrices are permuted according to $\mathbf{P}$.

Thus far, we have shown that if we permute the order of the dependent variables via $\widetilde{\mathbf{y}}_t = \mathbf{P}\mathbf{y}_t$, new parameters can be obtained by permuting the old parameters accordingly. This is applicable to both coefficient matrices ($\widetilde{\mathbf{b}}$ and $\widetilde{\mathbf{A}}_j$) and the error covariance matrix $\widetilde{\Sigma}$. We summarize them in Table A-1.

Table A-1: Original and New parameters

	Original	New	Relationship
Intercepts	$\mathbf{b}$	$\widetilde{\mathbf{b}}$	$\widetilde{\mathbf{b}} = \mathbf{P}\mathbf{b}$
Coefficients	$\mathbf{A}_j$	$\widetilde{\mathbf{A}}_j$	$\widetilde{\mathbf{A}}_j = \mathbf{P}\mathbf{A}_j$
Eigenvectors	$\mathbf{U}$	$\widetilde{\mathbf{U}}$	$\widetilde{\mathbf{U}} = \mathbf{P}\mathbf{U}$
Eigenvalues	$\boldsymbol{\Lambda}^{\frac{1}{2}}$	$\widetilde{\boldsymbol{\Lambda}}^{\frac{1}{2}}$	$\widetilde{\boldsymbol{\Lambda}}^{\frac{1}{2}} = \boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{P}'$

Next, we prove that the conditional likelihood function implied by this new ordering is the same as that of the original ordering.

A.2 Original and New conditional likelihoods

We first stack the original dependent variables $\mathbf{y}_t$ and new dependent variables $\widetilde{\mathbf{y}}_t$ over time.

We stack the original dependent variables into a $T \times n$ matrix $\mathbf{Y}$ so that its t-th row is $\mathbf{y}_t'$. Now, let $\mathbf{X}$ be a $T \times k$ matrix of regressors, where the t-th row is $\mathbf{x}_t' = \left(1, \mathbf{y}_{t-1}', \ldots, \mathbf{y}_{t-p}'\right)$. Next, let $\mathbf{A} = (\mathbf{b}, \mathbf{A}_1, \ldots, \mathbf{A}_p)'$ denote the $k \times n$ matrix of VAR coefficients. Then, we can write the original VAR(p) as follows:

$$\mathbf{Y} = \mathbf{X}\mathbf{A} + \mathbf{E}\mathbf{U}', \tag{12}$$

where $\mathbf{E}$ is a $T \times n$ matrix of innovations in which the t-th row is $\boldsymbol{\varepsilon}_t'$. It follows that

$$\mathrm{vec}(\mathbf{E}\mathbf{U}') \sim \mathcal{N}\left(\mathbf{0}, (\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}') \otimes \mathbf{I}_T\right).$$

Stacking the new dependent variables $\widetilde{\mathbf{y}}_t$ into $\widetilde{\mathbf{Y}} = (\widetilde{\mathbf{y}}_1, \ldots, \widetilde{\mathbf{y}}_T)'$, Next, let $\widetilde{\mathbf{A}} = \left(\widetilde{\mathbf{b}}, \widetilde{\mathbf{A}}_1, \ldots, \widetilde{\mathbf{A}}_p\right)'$ denote the $k \times n$ matrix of VAR coefficients. Then, we can write the new VAR(p) as follows:

$$\widetilde{\mathbf{Y}} = \mathbf{X}\widetilde{\mathbf{A}} + \mathbf{E}\widetilde{\mathbf{U}}', \tag{13}$$

where $\mathbf{E}$ is a $T \times n$ matrix of innovations in which the t-th row is $\boldsymbol{\varepsilon}_t'$. It follows that

$$\mathrm{vec}(\mathbf{E}\widetilde{\mathbf{U}}') \sim \mathcal{N}\left(\mathbf{0}, (\widetilde{\mathbf{U}}\widetilde{\boldsymbol{\Lambda}}\widetilde{\mathbf{U}}') \otimes \mathbf{I}_T\right).$$

Table A-2 summarizes their relationships. Next, we prove that $p(\widetilde{\mathbf{Y}} \mid \widetilde{\mathbf{A}}, \widetilde{\mathbf{U}}, \widetilde{\boldsymbol{\Lambda}}) = p(\mathbf{Y} \mid \mathbf{A}, \mathbf{U}, \boldsymbol{\Lambda})$. We first show the conditional likelihood of the original ordering, then we show the conditional likelihood of the new ordering.

Table A-2: Original and New matrices

	Original	New	Relationship
Dependent variables	$\mathbf{Y}$	$\widetilde{\mathbf{Y}}$	$\widetilde{\mathbf{Y}} = \mathbf{Y}\mathbf{P}'$
Coefficients	$\mathbf{A}$	$\widetilde{\mathbf{A}}$	$\widetilde{\mathbf{A}} = \mathbf{A}\mathbf{P}'$
Eigenvectors	$\mathbf{U}$	$\widetilde{\mathbf{U}}$	$\widetilde{\mathbf{U}} = \mathbf{P}\mathbf{U}$
Eigenvalues	$\mathbf{\Lambda}^{\frac{1}{2}}$	$\widetilde{\mathbf{\Lambda}}^{\frac{1}{2}}$	$\widetilde{\mathbf{\Lambda}}^{\frac{1}{2}} = \mathbf{\Lambda}^{\frac{1}{2}}\mathbf{P}'$
Conditional Likelihoods	$p(\mathbf{Y} \mid \mathbf{A}, \mathbf{U}, \mathbf{\Lambda})$	$p(\widetilde{\mathbf{Y}} \mid \widetilde{\mathbf{A}}, \widetilde{\mathbf{U}}, \widetilde{\mathbf{\Lambda}})$	same

A.3 Likelihoods of the original ordering

Likelihoods of the original ordering are given by

$$p(\mathbf{Y} \mid \mathbf{A}, \mathbf{U}, \mathbf{\Lambda}) = (2\pi)^{-\frac{Tn}{2}} |\mathbf{U\Lambda U}'|^{-\frac{T}{2}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\left((\mathbf{Y}-\mathbf{XA})(\mathbf{U\Lambda U}')^{-1}(\mathbf{Y}-\mathbf{XA})'\right)}, \tag{14}$$

$$= (2\pi)^{-\frac{Tn}{2}} |\mathbf{\Lambda}|^{-\frac{T}{2}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\left((\mathbf{Y}-\mathbf{XA})(\mathbf{U\Lambda U}')^{-1}(\mathbf{Y}-\mathbf{XA})'\right)}, \tag{15}$$

$$= (2\pi)^{-\frac{Tn}{2}} \prod_{j=1}^{n} \lambda_j^{-T/2} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\left((\mathbf{Y}-\mathbf{XA})(\mathbf{U\Lambda U}')^{-1}(\mathbf{Y}-\mathbf{XA})'\right)}, \tag{16}$$

$$= (2\pi)^{-\frac{Tn}{2}} \prod_{j=1}^{n} \lambda_j^{-T/2} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\left((\mathbf{Y}-\mathbf{XA})\mathbf{U\Lambda}^{-1}\mathbf{U}'(\mathbf{Y}-\mathbf{XA})'\right)}, \tag{17}$$

$$= (2\pi)^{-\frac{Tn}{2}} \prod_{j=1}^{n} \lambda_j^{-T/2} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\left(\mathbf{U\Lambda}^{-1}\mathbf{U}'(\mathbf{Y}-\mathbf{XA})'(\mathbf{Y}-\mathbf{XA})\right)}. \tag{18}$$

Proof from (14) to (15):

Since $\det(\mathbf{BC}) = \det(\mathbf{B})\det(\mathbf{C})$, and $\det\left(\mathbf{B}^{-1}\right) = \frac{1}{\det(\mathbf{B})}$, we can get $|\mathbf{U}\Lambda\mathbf{U}'|^{-\frac{T}{2}} = (|\mathbf{U}||\Lambda||\mathbf{U}'|)^{-\frac{T}{2}} = (|\mathbf{U}||\Lambda||\mathbf{U}^{-1}|)^{-\frac{T}{2}} = \left(|\mathbf{U}||\Lambda|\frac{1}{|\mathbf{U}|}\right)^{-\frac{T}{2}} = |\Lambda|^{-\frac{T}{2}}$.

Proof from (15) to (16):

Matrix Λ is diagonal, and for a diagonal matrix B, $\det(B) = b_{11}b_{22}\cdots b_{nn} = \prod_{i=1}^{n} b_{ii}$, so $|\mathbf{\Lambda}|^{-\frac{T}{2}} = \left(\prod_{j=1}^{n} \lambda_j\right)^{-\frac{T}{2}} = \prod_{j=1}^{n} \lambda_j^{-T/2}$.

Proof from (16) to (17):

Matrix $\mathbf{U}$ is orthogonal, which means that $\mathbf{U}^{-1} = \mathbf{U}'$.

Then $(\mathbf{U\Lambda U}')^{-1} = (\mathbf{U}')^{-1}(\mathbf{\Lambda})^{-1}(\mathbf{U})^{-1} = \mathbf{U\Lambda}^{-1}\mathbf{U}'$.

Proof from (17) to (18): $\mathrm{tr}(\mathbf{BCD}) = \mathrm{tr}(\mathbf{CDB}) = \mathrm{tr}(\mathbf{DBC})$

A.4 Likelihoods of the new ordering

Likelihoods of the new ordering are given by

$$
\begin{aligned}
p(\widetilde{\mathbf{Y}} \mid \widetilde{\mathbf{A}}, \widetilde{\mathbf{U}}, \widetilde{\mathbf{\Lambda}}) &= (2\pi)^{-\frac{Tn}{2}} |(\widetilde{\mathbf{U}}\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{U}}')|^{-\frac{T}{2}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\left((\widetilde{\mathbf{Y}}-\mathbf{X}\widetilde{\mathbf{A}})(\widetilde{\mathbf{U}}\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{U}}')^{-1}(\widetilde{\mathbf{Y}}-\mathbf{X}\widetilde{\mathbf{A}})'\right)}, && (19)\\
&= (2\pi)^{-\frac{Tn}{2}} |\mathbf{\Lambda}|^{-\frac{T}{2}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\left((\widetilde{\mathbf{Y}}-\mathbf{X}\widetilde{\mathbf{A}})(\widetilde{\mathbf{U}}\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{U}}')^{-1}(\widetilde{\mathbf{Y}}-\mathbf{X}\widetilde{\mathbf{A}})'\right)}, && (20)\\
&= (2\pi)^{-\frac{Tn}{2}} \prod_{j=1}^{n} \lambda_j^{-T/2} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\left((\widetilde{\mathbf{Y}}-\mathbf{X}\widetilde{\mathbf{A}})(\mathbf{PU\Lambda U'P'})^{-1}(\widetilde{\mathbf{Y}}-\mathbf{X}\widetilde{\mathbf{A}})'\right)}, && (21)\\
&= (2\pi)^{-\frac{Tn}{2}} \prod_{j=1}^{n} \lambda_j^{-T/2} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\left((\widetilde{\mathbf{Y}}-\mathbf{X}\widetilde{\mathbf{A}})\mathbf{PU\Lambda}^{-1}\mathbf{U'P'}(\widetilde{\mathbf{Y}}-\mathbf{X}\widetilde{\mathbf{A}})'\right)}, && (22)\\
&= (2\pi)^{-\frac{Tn}{2}} \prod_{j=1}^{n} \lambda_j^{-T/2} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\left((\mathbf{Y}-\mathbf{XA})\mathbf{U\Lambda}^{-1}\mathbf{U'}(\mathbf{Y}-\mathbf{XA})'\right)}, && (23)\\
&= (2\pi)^{-\frac{Tn}{2}} \prod_{j=1}^{n} \lambda_j^{-T/2} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\left(\mathbf{U\Lambda}^{-1}\mathbf{U'}(\mathbf{Y}-\mathbf{XA})'(\mathbf{Y}-\mathbf{XA})\right)}, && (24)\\
&= p(\mathbf{Y} \mid \mathbf{A}, \mathbf{U}, \mathbf{\Lambda}).
\end{aligned}
$$

Proof from (19) to (20):

Since $\det(\mathbf{BC}) = \det(\mathbf{B})\det(\mathbf{C})$, and $\det(\mathbf{B}^{-1}) = \frac{1}{\det(\mathbf{B})}$, we can get $|(\widetilde{\mathbf{U}}\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{U}}')|^{-\frac{T}{2}} = |\mathbf{PU\Lambda}^{\frac{1}{2}}\mathbf{P'P}(\mathbf{\Lambda}^{\frac{1}{2}})'\mathbf{U'P'}|^{-\frac{T}{2}} = |\mathbf{PU\Lambda U'P'}|^{-\frac{T}{2}} = (|\mathbf{P}||\mathbf{U}||\mathbf{\Lambda}||\mathbf{U}'||\mathbf{P}'|)^{-\frac{T}{2}} = (|\mathbf{P}||\mathbf{U}||\mathbf{\Lambda}||\mathbf{U}^{-1}||\mathbf{P}^{-1}|)^{-\frac{T}{2}} = \left(|\mathbf{P}||\mathbf{U}||\mathbf{\Lambda}||\frac{1}{|\mathbf{U}|}|\frac{1}{|\mathbf{P}|}\right)^{-\frac{T}{2}} = |\mathbf{\Lambda}|^{-\frac{T}{2}}$.

Proof from (20) to (21): $\widetilde{\mathbf{U}}\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{U}}' = \mathbf{PU\Lambda}^{\frac{1}{2}}\mathbf{P'P}(\mathbf{\Lambda}^{\frac{1}{2}})'\mathbf{U'P'} = \mathbf{PU\Lambda U'P'}$.

Proof from (21) to (22): Same as Proof from (16) to (17).

Proof from (22) to (23):

From Table A-2, $\widetilde{\mathbf{Y}} = \mathbf{YP}'$, $\widetilde{\mathbf{A}} = \mathbf{AP}'$.

So $\widetilde{\mathbf{Y}} - \mathbf{X}\widetilde{\mathbf{A}} = (\mathbf{Y} - \mathbf{XA})\mathbf{P}'$, $(\widetilde{\mathbf{Y}} - \mathbf{X}\widetilde{\mathbf{A}})' = \mathbf{P}(\mathbf{Y} - \mathbf{XA})'$.

Then

$(\widetilde{\mathbf{Y}} - \mathbf{X}\widetilde{\mathbf{A}})\mathbf{PU\Lambda}^{-1}\mathbf{U'P'}(\widetilde{\mathbf{Y}} - \mathbf{X}\widetilde{\mathbf{A}})' = (\mathbf{Y} - \mathbf{XA})\mathbf{P'PU\Lambda}^{-1}\mathbf{U'P'P}(\mathbf{Y} - \mathbf{XA})' = (\mathbf{Y} -$

$\mathbf{XA})\mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}'(\mathbf{Y}-\mathbf{XA})'$.

Proof from (23) to (24): Same as Proof from (17) to (18).

Combining Section A.1 and Section A.2, we provide evidence that if we permute the order of the dependent variables via $\widetilde{\mathbf{y}}_t = \mathbf{P}\mathbf{y}_t$, new parameters can be obtained by permuting the old parameters accordingly. This is applicable to both coefficient matrices ($\widetilde{\mathbf{b}}$ and $\widetilde{\mathbf{A}}_j$) and the error covariance matrix $\widetilde{\Sigma}$. And the conditional likelihood function implied by this new ordering is the same as that of the original ordering.

B Priors

For all models, we use the same prior for the VAR coefficients. Any choice could be made for this without altering the main messages of the paper. We choose the Minnesota-type Horseshoe prior of Chan et al. (2023). Please note that this prior is put on the parameters of the reduced form VAR coefficients $\mathbf{A}_i$, not the Eigen- or Cholesky-transformed coefficients.

Let $\boldsymbol{\alpha}_i$ denote the VAR coefficients in the i-th equation, $i = 1, \ldots, n$. We consider the Minnesota-type horseshoe prior proposed in Chan et al. (2023), which can impose cross-variable shrinkage and have substantial mass around 0. For $\alpha_{i,j}$, the j-th coefficient in the i-th equation, the prior is:

$$(\alpha_{i,j} \mid \kappa_1, \kappa_2, \psi_{i,j}) \sim \mathcal{N}\left(0, \kappa_{i,j}\psi_{i,j}C_{i,j}\right),$$
$$\sqrt{\psi}_{i,j} \sim \mathcal{C}^+(0,1),$$
$$\sqrt{\kappa}_1, \sqrt{\kappa}_2 \sim \mathcal{C}^+(0,1),$$

where $\mathcal{C}^+(0,1)$ denotes the standard half-Cauchy distribution. κ_1 and κ_2 are the global variance components that are common to, respectively, coefficients of own and other lags, whereas each $\psi_{i,j}$ is a local variance component specific to the coefficient $\alpha_{i,j}$. The constants

$C_{i,j}$ are obtained as

$$C_{i,j} = \begin{cases} \frac{1}{l^2}, & \text{for the coefficient on the } l\text{-th lag of variable } i, \\ \frac{s_i^2}{l^2 s_j^2}, & \text{for the coefficient on the } l\text{-th lag of variable } j, j \neq i, \end{cases}$$

where s_r^2 denotes the sample variance of the residuals from an AR(4) model for the variable $r, r = 1, \dots, n$.

For priors on the error covariance matrix (or its related parameters), the associated parameters vary across models. To ensure clarity, we discuss them on a model-by-model basis. In the following, we define

$$e_t = \mathbf{y}_t - (\mathbf{a} + \mathbf{A}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{A}_p \mathbf{y}_{t-p}).$$

B.1 BVAR-IW

In the BVAR-IW model where $e_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$, the prior is directly imposed on the error covariance matrix Σ. We assume an inverse-Wishart distribution $\Sigma \sim \mathcal{IW}(v, \mathbf{S}_0)$, where $\mathcal{IW}$ denotes the inverse-Wishart distribution with degrees of freedom v and a positive definite scale matrix $\mathbf{S}_0$. We set $v = n + 3$ where n is the number of variables (20 in our main analysis), and set $\mathbf{S}_0 = \text{diag}(s_1^2, ..., s_n^2)$ where s_r^2 is defined above.

B.2 BVAR-chol

In the BVAR-chol model where $e_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$, the prior is imposed on the Cholesky-transformed parameters. Rewrite $e_t = L^{-1} \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, C)$ where L is a lower triangular matrix with ones on the diagonal and free elements $L_{i,j}$, and C is a diagonal matrix denoted by σ_i^2. The priors are

$$\sigma_i^2 \sim \mathcal{IG}\left(\frac{v+i-n}{2}, \frac{s_i^2}{2}\right), \quad i = 1, \ldots, n,$$

$$L_{i,j} \sim \mathcal{N}\left(0, \frac{s_i^2}{s_j^2}\right), \quad 1 \leqslant j < i \leqslant n, i = 2, \ldots, n.$$

B.3 BVAR-eig

In the BVAR-eig model, the prior is imposed on the eigen-transformed parameters which we call the $\mathcal{BIG}$ prior. Rewrite $e_t = \mathbf{U}\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda})$ where $\mathbf{U}$ is the eigenmatrix, and $\boldsymbol{\Lambda}$ is a diagonal matrix which contains eigenvalues λ_i. The priors are

$$\mathbf{U} \sim \mathcal{B}(\mathbf{D}, \mathbf{B}, \mathbf{V}),$$

where $\mathcal{B}$ denotes the Bingham distribution. $\mathbf{V}$ and $\mathbf{D}$ are obtained as the eigenmatrix and eigenvalues from the eigendecomposition of $\mathbf{S}_0$. Since $\mathbf{S}_0$ is diagonal, this implies that we can set $\mathbf{V}$ to be an identity matrix $\mathbf{V} = I_n$, and set $\mathbf{D} = \mathbf{S}_0$. Therefore, in the $\mathcal{BIG}$ prior, we do not estimate the hyperparameters $\mathbf{V}$ and $\mathbf{D}$. The $\mathcal{BIG}$ prior sets $\mathbf{B} = -\frac{1}{2}\boldsymbol{\Lambda}^{-1}$.

For the eigenvalues, $\mathcal{BIG}$ uses the inverse-Gamma distribution

$$\lambda_i \sim \mathcal{IG}\left(\frac{v+n-1}{2}, \frac{u_i'(\mathbf{VDV}')u_i}{2}\right), \quad i = 1, \ldots, n.$$

B.4 Treatment of SV

In the next part of this appendix, we consider several models which have SV. In all such models, each of the log-volatilities follows a stationary AR(1) process:

$$h_{i,t} = \mu_i + \phi_i(h_{i,t-1} - \mu_i) + u_{i,t}^h, \quad u_{i,t}^h \sim \mathcal{N}\left(0, \omega_i^2\right), \tag{25}$$

For the coefficient ϕ_i, we assume a normal prior $\mu_i \sim \mathcal{N}(0.95, 0.05^2)$ and impose $|\phi_i| < 1$ to ensure stationarity. The stationary distributions have nonzero means μ_i and we assume a normal prior $\mu_i \sim \mathcal{N}(0, 1)$. The initial conditions follow the stationary distributions $h_{i,1} \sim \mathcal{N}(\mu_i, \omega_i^2/(1 - \phi_i^2))$. The error variances follow an independent inverse gamma distribution $\omega_i^2 \sim \mathcal{IG}(3, 0.05)$ such that the prior mean is 0.025.

B.5 BVAR-CSV

In the BVAR-CSV model, we write $e_t \sim \mathcal{N}\left(\mathbf{0}, e^{h_t}\Sigma\right)$ where e^{h_t} is the latent stochastic volatility and Σ is an $n \times n$ matrix.

For Σ, it follows an inverse-Wishart distribution $\Sigma \sim \mathcal{IW}(v, \mathbf{S}_1)$. We set $v = n+3$. For $\mathbf{S}_1$, we follow Amir-Ahmadi et al. (2020) by adopting a hierarchical modeling approach: $\mathbf{S}_1 = \text{diag}(\kappa)v\mathbf{S}_0\text{diag}(\kappa)$ where $\mathbf{S}_0 = \text{diag}(s_1^2, ..., s_n^2)$. The hyperparameter κ is an $n \times 1$ vector. We estimate the vector κ. We assume it has an independent inverse-Gamma prior $\kappa \sim \mathcal{IG}(1, 0.1)$. We draw κ with a random walk Metropolis-Hasting step.

B.6 BVAR-cholSV

In the BVAR-cholSV model, the error term is specified as:

$$e_t = L\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\mathbf{0}, C_t\right).$$

The time-varying matrix $C_t = \text{diag}(e^{h_{1,t}}, ..., e^{h_{n,t}})$. The constant matrix L is lower-triangular with ones on the diagonal. Let $L_{i,j}$ denote the free elements. Following Chan (2021), we employ a Minnesota-type horseshoe prior for these elements. Specifically, the prior structure is defined as:

$$
\begin{aligned}
(L_{i,j} \mid \kappa, \psi_{i,j}) &\sim \mathcal{N}\left(0, \kappa\psi_{i,j}C_{i,j}\right), \\
\sqrt{\psi}_{i,j} &\sim \mathcal{C}^{+}(0,1), \\
\sqrt{\kappa} &\sim \mathcal{C}^{+}(0,1),
\end{aligned}
$$

where $C_{i,j} = \frac{s_i^2}{s_j^2}$ and we estimate κ and $\psi_{i,j}$.

B.7 BVAR-eigSV

In the BVAR-eigSV model, we write

$$
e_t = \mathbf{U}\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Lambda}_t\right).
$$

The time-varying matrix $\boldsymbol{\Lambda}_t = \text{diag}(e^{h_{1,t}}, ..., e^{h_{n,t}})$. The constant matrix $\mathbf{U}$ is the eigenmatrix. The prior on $\mathbf{U}$ is

$$
\mathbf{U} \sim \mathcal{B}\left(\mathbf{D}, \mathbf{B}, \mathbf{V}\right),
$$

where the hyperparameters $\mathbf{D}$, $\mathbf{B}$, and $\mathbf{V}$ are estimated as described in Section 4 of the main paper.

B.8 BVAR-SeigSV

In the BVAR-SeigSV model, the error term is specified as:

$$
e_t = \mathbf{S}\mathbf{U}\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Lambda}_t\right).
$$

The matrix $\mathbf{S}$ is defined as $\mathbf{S} = \delta\text{diag}\left(\sigma_1, \ldots, \sigma_n\right)$ where σ_i denotes the standard deviation of variable $y_{i,1:T}$. δ is a scaling parameter with an inverse Gamma prior $\mathcal{IG}(10,9)$. The time-varying matrix $\mathbf{\Lambda}_t$ and the constant matrix $\mathbf{U}$ are the same as that in BVAR-eigSV model above.

B.9 BVAR-OISV

In the BVAR-OISV model as proposed in Chan et al. (2023), the error term is specified as

$$e_t = F\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\mathbf{0}, C_t\right).$$

The time-varying matrix $C_t = \mathrm{diag}(e^{h_{1,t}}, ..., e^{h_{n,t}})$. The constant matrix F is a full matrix. Following Chan (2021), we use a Minnesota-type horseshoe prior. Specifically, the prior structure is defined as:

$$\begin{aligned}
(F_{i,j} \mid \kappa, \psi_{i,j}) &\sim \mathcal{N}\left(0, \kappa\psi_{i,j}C_{i,j}\right), \\
\sqrt{\psi}_{i,j} &\sim \mathcal{C}^{+}(0,1), \\
\sqrt{\kappa} &\sim \mathcal{C}^{+}(0,1),
\end{aligned}$$

where $C_{i,j} = \frac{s_i^2}{s_j^2}$ and we estimate κ and $\psi_{i,j}$.

B.10 BVAR-OISV(TVP)

In the BVAR-OISV(TVP) model as proposed in Chan et al. (2023), it is the same as BVAR-OISV except the full matrix F. Instead of having a constant F, the model allows for time variation in its elements. Following their approach, each element $f_{i,j,t}$ evolves as a random walk

$$f_{i,j,t} = f_{i,j,t-1} + e_{i,j,t}, \quad e_{i,j,t} \sim \mathcal{N}\left(0, \omega_{i,j}^2\right).$$

We estimate the error variance $\omega_{i,j}^2$ and assume an inverse-gamma prior $\omega_{i,j}^2 \sim \mathcal{IG}(0, s_i^2)$. For macroeconomic variables (the first ten variables in the dataset), we set $s_i^2 = 0.001$, whereas for financial variables (the remaining ten variables), we impose a tighter prior with $s_i^2 = 0.0005$. This distinction reflects the differing dynamics between the two groups: macroeconomic variables typically exhibit higher persistence and smoother trends, justifying a more flexible prior to accommodate larger adjustments in the impact matrix. In contrast, financial variables are often subject to more erratic volatility and short-lived shocks, necessitating stricter constraints on their evolution to avoid overfitting transient fluctuations.

C Data

Table A-3: Description of variables used in the forecasting application

Variable	Mnemonic	Transformation
Real Gross Domestic Product	GDPC1	400Δ log
Personal Consumption Expenditures	PCECC96	400Δ log
Industrial Production Index	INDPRO	400Δ log
Industrial Production: Final Products	IPFINAL	400Δ log
All Employees: Total nonfarm	PAYEMS	400Δ log
All Employees: Manufacturing	MANEMP	400Δ log
Civilian Employment	CE16OV	400Δ log
Civilian Labor Force Participation Rate	CIVPART	no transformation
Civilian Unemployment Rate	UNRATE	no transformation
Nonfarm Business Section: Hours of All Persons	HOANBS	400Δ log
Housing Starts: Total	HOUST	400Δ log
New Private Housing Units Authorized by Building Permits	PERMIT	400Δ log
Personal Consumption Expenditures: Chain-type Price index	PCECTPI	400Δ log
Consumer Price Index for All Urban Consumers: All Items	CPIAUCSL	400Δ log
Nonfarm Business Section: Real Output Per Hour of All Persons	OPHNFB	400Δ log
Effective Federal Funds Rate	FEDFUNDS	no transformation
3-Month Treasury Bill: Secondary Market Rate	TB3MS	no transformation
1-Year Treasury Constant Maturity Rate	GS1	no transformation
10-Year Treasury Constant Maturity Rate	GS10	no transformation
Moody's Seasoned Baa Corporate Bond Yield Relative to Yield on 10-Year Treasury Constant Maturity	BAA10YM	no transformation

D Homoskedastic Models: More comparisons

D.1 Homoskedastic Models: Forecast Performance in absolute values

Table A-4: RMSFE and ALPL of 20 macroeconomic time series.

Variables	Models	RMSFE				ALPL			
		$h=1$	$h=2$	$h=3$	$h=4$	$h=1$	$h=2$	$h=3$	$h=4$
GDPC1	BVAR-IW	**4.958**	4.946	4.952	4.941	**-3.097**	**-3.255**	**-3.239**	**-3.254**
	BVAR-chol	5.581	4.954	4.951	4.953	-3.354	-3.472	-3.482	-3.497
	BVAR-chol-RO	4.986	**4.928**	**4.928**	**4.925**	-3.188	-3.375	-3.365	-3.391
	BVAR-eig	4.987	4.980	4.993*	4.963	-3.126	-3.369	-3.363	-3.327
PCECC96	BVAR-IW	5.154	5.348	**5.329**	5.296	**-3.310**	**-3.813**	**-3.858**	**-3.856**
	BVAR-chol	5.781	**5.323**	5.387	5.343	-3.625	-3.973	-4.057	-4.056
	BVAR-chol-RO	5.222	5.363	5.350	**5.289**	-3.420	-3.966	-4.062	-4.040
	BVAR-eig	**5.150**	5.350	5.331	5.303	-3.474	-4.052	-4.098	-4.046
INDPRO	BVAR-IW	7.949	7.902	7.904	7.910	**-3.503**	**-3.657**	**-3.690**	**-3.698**
	BVAR-chol	8.650	7.952	7.933	7.924	-3.636	-3.867	-3.876	-3.874
	BVAR-chol-RO	8.112	**7.893**	**7.884**	**7.870**	-3.600	-3.776	-3.790	-3.796
	BVAR-eig	**7.925**	7.921	7.940	7.926	-3.508	-3.782	-3.798	-3.811
IPFINAL	BVAR-IW	8.289	**8.020**	**8.062**	8.079	**-3.595**	**-3.749**	**-3.786**	**-3.785**
	BVAR-chol	9.393	8.082	8.148	8.107	-3.847	-3.953	-4.009	-4.006
	BVAR-chol-RO	8.353	8.033	8.075	**8.066**	-3.709	-3.885	-3.934	-3.915
	BVAR-eig	**8.285**	8.041	8.093	8.096	-3.624	-3.856	-3.884	-3.901
PAYEMS	BVAR-IW	**5.476**	5.226	5.242	5.266	**-7.305**	**-7.972**	**-8.860**	**-9.112**
	BVAR-chol	6.630	5.486	5.288	5.305	-9.636	-11.087	-11.056	-11.594
	BVAR-chol-RO	5.557	5.238	5.259	5.287	-8.426	-9.897	-10.615	-10.047
	BVAR-eig	5.476	**5.220**	**5.235**	**5.264**	-8.033	-9.174	-9.406	-9.471
MANEMP	BVAR-IW	4.792	4.708	4.834	4.876	**-3.052**	**-3.206**	**-3.354**	**-3.387**
	BVAR-chol	5.391	5.006	5.007	5.021	-3.266	-3.450	-3.687	-3.723
	BVAR-chol-RO	4.902	4.807	4.949*	4.975**	-3.172	-3.364	-3.533	-3.520
	BVAR-eig	**4.749**	**4.689**	**4.829**	**4.866**	-3.060	-3.235	-3.447	-3.505
CE16OV	BVAR-IW	5.562	5.645	5.637	5.639	-6.105	**-7.095**	**-7.199**	**-7.315**
	BVAR-chol	6.371	5.727	**5.632**	5.641	-7.533	-8.693	-8.787	-8.232
	BVAR-chol-RO	5.548	5.639	5.649	5.636	-6.834	-8.542	-8.587	-8.968
	BVAR-eig	**5.538****	**5.631**	5.635	**5.634**	**-6.057**	-7.546	-7.874	-7.953
CIVPART	BVAR-IW	0.256	0.339	0.396	0.456	-0.201	-1.468	-2.472	-3.455
	BVAR-chol	0.274	0.355	0.411	0.481	-0.318	-1.471	-2.516	-3.585
	BVAR-chol-RO	0.255	0.339	0.401	0.463	-0.247	-1.407	-2.153	-2.838
	BVAR-eig	**0.255**	**0.334**	**0.388**	**0.444**	**0.112**	**-0.797**	**-1.516**	**-2.189**
UNRATE	BVAR-IW	**0.934**	**1.199**	1.354	1.475	-4.505	-7.650	-10.633	-13.079
	BVAR-chol	1.065	1.392	1.570	1.686	-6.179	-10.142	-14.525	-17.799
	BVAR-chol-RO	0.943	1.221	1.388	1.514	-5.968	-9.854	-12.594	-14.512
	BVAR-eig	0.934	1.202	**1.352**	**1.469**	**-2.733**	**-4.899**	**-6.727**	**-8.891**
HOANBS	BVAR-IW	6.542	6.438	6.436	6.450	**-4.106**	**-4.504**	**-4.626**	**-4.695**
	BVAR-chol	7.325	6.515	6.452	6.491	-4.762	-5.117	-5.231	-5.535
	BVAR-chol-RO	6.622	6.467	6.473	6.478	-4.335	-4.939	-5.176	-5.001
	BVAR-eig	**6.481***	**6.404**	**6.413**	**6.423****	-4.163	-4.768	-5.052	-5.084

[1] Notes: The bold figure indicates the best model in each case. $*$, $**$ and $***$ denote, respectively, 0.10, 0.05 and 0.01 significance level for a two-sided Diebold and Mariano(1995) test.The benchmark model in the Diebold Mariano test is BVAR-IW with 20 variables.

Table A-4: Continued: RMSFE and ALPL of 20 macroeconomic time series.

Variables	Models	RMSFE				ALPL			
		$h=1$	$h=2$	$h=3$	$h=4$	$h=1$	$h=2$	$h=3$	$h=4$
HOUST	BVAR-IW	27.521	31.746	**31.371**	**31.407**	-4.741	-5.064	-5.090	-5.108
	BVAR-chol	**27.285**	32.040	32.011*	31.910	-4.737	-5.051	-5.064	-5.064
	BVAR-chol-RO	27.304	32.038	31.877*	31.836	**-4.725**	**-4.999**	**-5.000***	**-5.001****
	BVAR-eig	27.877	**31.410**	31.392	31.727	-4.759	-5.262**	-5.323**	-5.381**
PERMIT	BVAR-IW	**29.156**	29.656	29.433	**29.417**	-4.804	-4.933	-4.940	-4.952
	BVAR-chol	29.576	29.931	30.105	30.054	-4.827	-4.877	-4.888	-4.892*
	BVAR-chol-RO	29.183	29.889	29.911	29.925	-4.796	**-4.845***	**-4.849****	**-4.851****
	BVAR-eig	29.742	**29.321**	**29.191**	29.689	**-4.793**	-5.039*	-5.079**	-5.131**
PCECTPI	BVAR-IW	**1.620**	**1.817**	**1.820**	**1.879**	**-1.934**	**-2.307**	**-2.415**	**-2.589**
	BVAR-chol	1.666	1.908	1.927*	1.951	-1.968	-2.383	-2.494*	-2.650
	BVAR-chol-RO	1.625	1.841	1.866	1.921	-2.012	-2.370	-2.447	-2.600
	BVAR-eig	1.634	1.857	1.878*	1.941**	-1.997	-2.386	-2.492	-2.682
CPIAUCSL	BVAR-IW	2.101	**2.299**	2.317	2.337	**-2.177**	**-2.522**	-2.621	-2.773
	BVAR-chol	2.142	2.415	2.390	2.413	-2.192	-2.571	**-2.611**	**-2.768**
	BVAR-chol-RO	**2.080**	2.309	**2.304**	**2.329**	-2.264	-2.607	-2.668	-2.796
	BVAR-eig	2.116	2.359*	2.382*	2.417**	-2.264	-2.656	-2.747	-2.936
OPHNFB	BVAR-IW	2.804	2.856	2.830	2.839	-2.468	-2.492	-2.493	-2.499
	BVAR-chol	2.848	2.847	2.795	2.811	-2.470	-2.488	-2.484	-2.494
	BVAR-chol-RO	**2.788**	**2.795***	**2.769***	**2.790**	**-2.431**	**-2.451**	**-2.452**	**-2.462**
	BVAR-eig	2.820	2.865	2.859	2.881	-2.457	-2.485	-2.495	-2.514*
FEDFUNDS	BVAR-IW	0.380	0.737	1.064	1.382	-0.934	-1.720	-2.482	-3.236
	BVAR-chol	0.439	0.811	1.139	1.453	-0.957	-1.762	-2.535	-3.299
	BVAR-chol-RO	**0.363**	**0.716**	**1.040**	**1.348**	**-0.834*****	**-1.607*****	**-2.231****	**-2.764****
	BVAR-eig	0.366	0.739	1.073	1.408	-0.863**	-1.705	-2.440	-3.140
TB3MS	BVAR-IW	0.380	0.695	0.980	1.258	-0.800	-1.559	-2.329	-3.109
	BVAR-chol	0.419**	0.741	1.022	1.300	-0.834**	-1.586	-2.330	-3.113
	BVAR-chol-RO	**0.367**	**0.672**	**0.951**	**1.230**	**-0.716*****	**-1.426*****	**-2.043****	**-2.616****
	BVAR-eig	0.375	0.688	0.979	1.267	-0.753**	-1.507	-2.230**	-2.926***
GS1	BVAR-IW	0.436	0.767	**1.052**	**1.331**	-0.843	-1.635	-2.440	-3.255
	BVAR-chol	0.463**	0.802*	1.090	1.368	-0.879**	-1.641	-2.414	-3.225
	BVAR-chol-RO	0.431	0.767	1.060	1.343	**-0.769****	**-1.534*****	**-2.198****	**-2.795****
	BVAR-eig	**0.430**	**0.764**	1.056	1.349	-0.798**	-1.600	-2.360**	-3.089***
GS10	BVAR-IW	0.382	**0.622**	**0.787**	**0.944**	-0.516	-1.516	-2.425	-3.286
	BVAR-chol	0.399**	0.652*	0.817	0.965	-0.568**	-1.516	-2.353	-3.125*
	BVAR-chol-RO	0.384	0.630	0.798	0.951	**-0.487**	**-1.428**	**-2.110*****	**-2.671*****
	BVAR-eig	**0.382**	0.627	0.798	0.964	-0.501	-1.487	-2.307**	-3.048**
BAA10YM	BVAR-IW	**0.347**	**0.529**	0.639	0.709	-0.371	-1.472	-2.288	-2.879
	BVAR-chol	0.356	0.550	0.667	0.734*	-0.399	-1.568	-2.432	-3.086
	BVAR-chol-RO	0.352	0.534	0.644	0.716	-0.448	-1.559	-2.261	-2.810
	BVAR-eig	0.350	0.531	**0.639**	**0.708**	**-0.324**	**-1.336**	**-2.000**	**-2.504**

[1] Notes: The bold figure indicates the best model in each case. $*$, $**$ and $***$ denote, respectively, 0.10, 0.05 and 0.01 significance level for a two-sided Diebold and Mariano(1995) test.The benchmark model in the Diebold Mariano test is BVAR-IW with 20 variables.

D.2 A Cholesky Non-informative prior

The ordering issue is due to the prior on the lower-triangular matrix L. This suggests that a relatively non-informative prior on the lower-triangular matrix L might be close to being ordering invariant at least in small models where the need for prior information is less.[12] Here we show that this is true to some extent in small VARs, but as the dimension increases, the ordering issue becomes more severe. This can be seen in the following figure which contains two panels. Both are from a Cholesky-based BVAR using a relatively noninformative prior (BVAR-chol-Noninfo) but with the variables ordered in two different ways. The relatively non-informative prior we use is $p\left(L_{i,j}\right) \propto N\left(0, 10^4\right)$. The only difference is

[12]We are grateful to the referee for pointing this out.

the dimension n. In Panel A, $n = 5$. In Panel B, $n = 20$. Within each panel, point forecast (in terms of percentage gains in RMSFE of one ordering relative to the other) and density forecast (in terms of percentage gains in ALPL of one ordering relative to the other) are reported. Note that, if the two different ways of ordering the data were giving the same results then each panel would be white. For point forecast, they are. However, for density forecast, they are not. So clearly ordering issues matter even with a noninformative prior. But they matter even more in larger models.

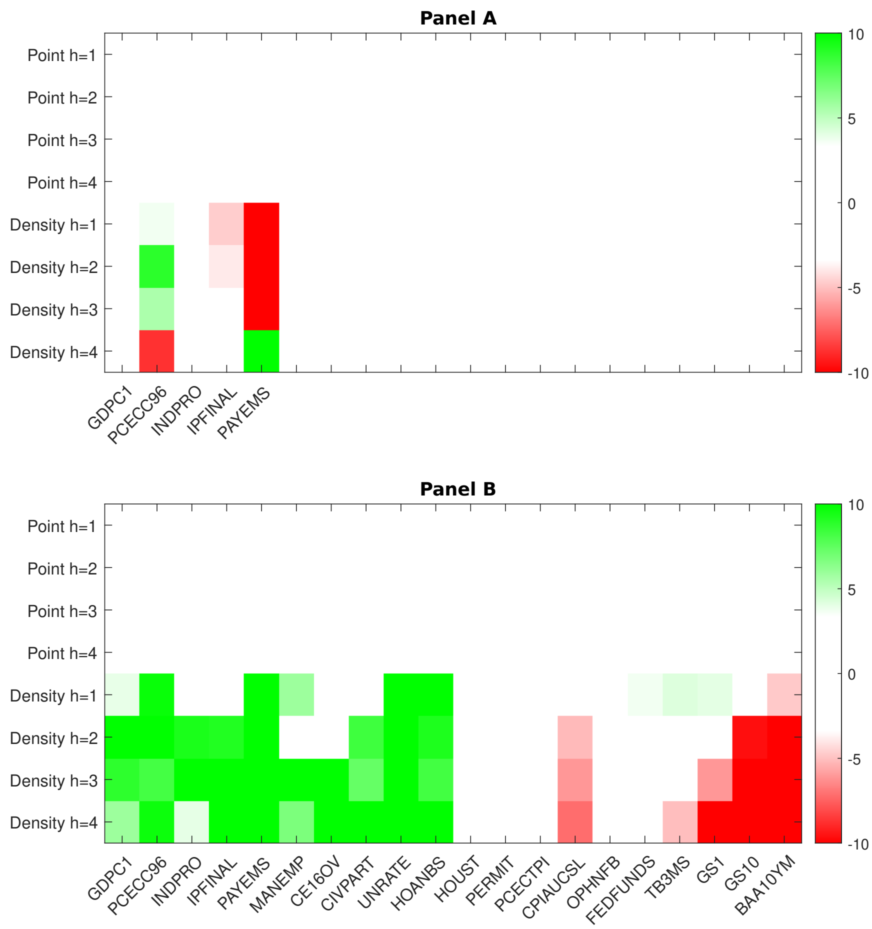

Figure A-1: Forecasting results from the large BVAR with Non-informative prior on the impact matrix. Panel A: $n = 5$. Panel B: $n = 20$. Within each panel, point forecast (in terms of percentage gains in RMSFE of one ordering relative to the other) and density forecast (in terms of percentage gains in ALPL of one ordering relative to the other) are reported.

Next we compare the Log marginal likelihood values of BIG prior and Cholesky Non-informative prior in Table A-5. As is shown, the data supports the BVAR-eig model with the use of the noninformative leading to substantial deterioration in this metric.

Table A-5: Log marginal likelihood values.

Models	log ML
BVAR-chol-Noninfo	-7,454
BVAR-chol-Noninfo-RO	-7,093
BVAR-eig	-6,780

D.3 A comparison between BVAR-eig and BVAR-chol

Here we compare the Eigendecomposition and Cholesky decomposition using a small BVAR with the four core macroeconomic time series. The two decompositions provide similar forecast results.

Table A-6: RMSFE and ALPL of four core macroeconomic time series.

Variables	Models	RMSFE				ALPL			
		$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
Industrial Production	BVAR-chol	8.752	8.586	8.286	**8.215**[1]	-3.582	**-3.933**	-3.972	**-3.973**
	BVAR-chol-RO	8.733	8.576	**8.283**	8.230*	-3.582	-3.936	-3.979	-3.995
	BVAR-eig	8.573	8.585	8.342	8.329	-3.563	-3.937	-3.978	-3.993
Unemployment Rate	BVAR-chol	1.215	1.526	1.760	1.900	-4.663	-8.324	-11.146	-13.114
	BVAR-chol-RO	1.214	1.525	1.753	1.887	-4.714	-8.127	-10.801	-13.104
	BVAR-eig	1.198	**1.482**	**1.684**	**1.797**	-4.511	**-7.907**	-10.532	**-12.316**
PCE inflation	BVAR-chol	1.734	1.988	2.002	2.085	-1.959	-2.440	-2.611	-2.868
	BVAR-chol-RO	1.736	1.991	2.005	2.082	-1.960	-2.456	-2.617	-2.868
	BVAR-eig	**1.716**	**1.952**	1.973	2.047	-1.943	-2.428	**-2.578**	-2.818
Federal funds rate	BVAR-chol	**0.523**	**0.937**	**1.249**	**1.542**	-0.976	**-1.915**	-2.718	-3.544
	BVAR-chol-RO	0.525	0.939	1.250	1.543	**-0.974**	-1.918*	-2.724*	-3.553**
	BVAR-eig	0.545	0.993	1.305	1.589	-0.989	-1.918	-2.682	-3.474

[1] Results are based on small models where $n = 4$. Notes: The bold figure indicates the best model in each case. $*$, $**$ and $***$ denote, respectively, 0.10, 0.05 and 0.01 significance level for a two-sided Diebold and Mariano(1995) test. The benchmark model in the Diebold Mariano test is BVAR-chol.

D.4 A comparison between IW prior and eigen priors

Here, we illustrate the equivalence of parameter estimates using scatter plots. Figure A-2 presents the posterior mean of the error covariance matrix, comparing BVAR-eig against the BVAR-IW model.

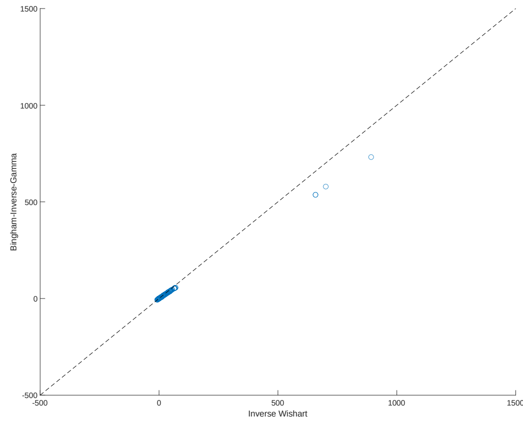

Figure A-2: Scatter plots of the posterior mean estimates of VAR error covariance matrix. The Dashed line is the diagonal line.

E Adding Stochastic Volatility

E.1 Effective Sample Size

To figure out what has slowed the computational time of eigSV and OISV down relative to the Cholesky based approach, we plot a figure for effective sample size (ESS). The figure presents ESS comparisons across different parameter groups: the overall ESS (Panel A) which includes all parameters, the ESS for volatilities $h_{1:T,i}$ (Panel B) and the ESS for other parameters (Panel C). For each panel, we compare four models: the BVAR-cholSV (Column (i)), the BVAR-eigSV (Column (ii)), the BVAR-SeigSV (Column (iii)) and the BVAR-OISV (Column (iv)). Each column is a boxplot of the MCMC effective sample size in 12,000 posterior draws. The results indicate that the discrepancy primarily arises from the volatility parameters, with eigSV exhibiting a higher ESS than OISV.

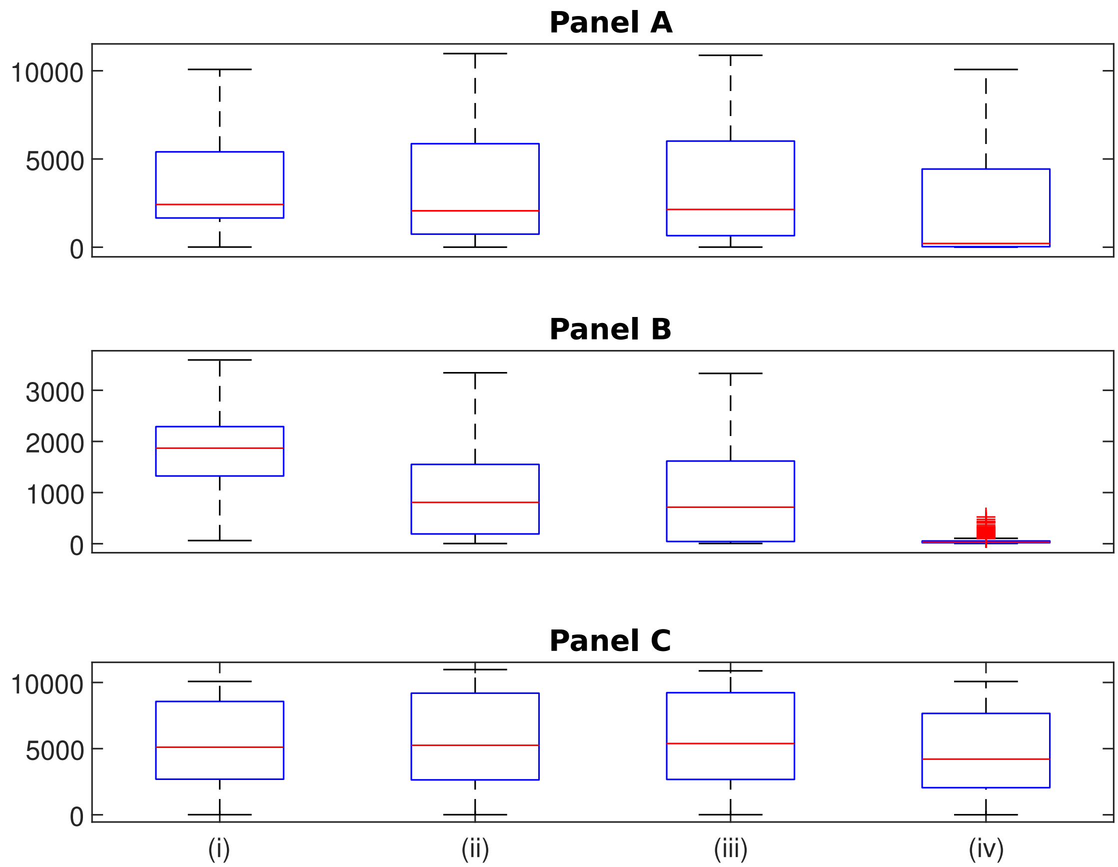

Figure A-3: Panel A: Overall ESS. Panel B: ESS for volatilities. Panel C: ESS for other parameters. Column (i): the BVAR-cholSV. Column (ii): the BVAR-eigSV. Column (iii): the BVAR-SeigSV. Column (iv): the BVAR-OISV.

E.2 Investigating the Properties of the Time-Varying Error Covariance Matrix

E.2.1 Complementing Figure 2

Here we report on comparing models with a constant impact matrix. The four models we compare are: a BVAR-eigSV, a BVAR-SeigSV, a BVAR-cholSV, and a BAR-cholSV with reverse orderings. We first look at large models with 20 variables, then small variables with the four key variables, and finally a large and a small model. The block involves industrial production, the unemployment rate, PCE inflation, and the Federal funds rate.

Comparing large models

The following figure shows the full sample estimates of elements of time-varying covariance matrix as sequence: full sample and credible intervals.

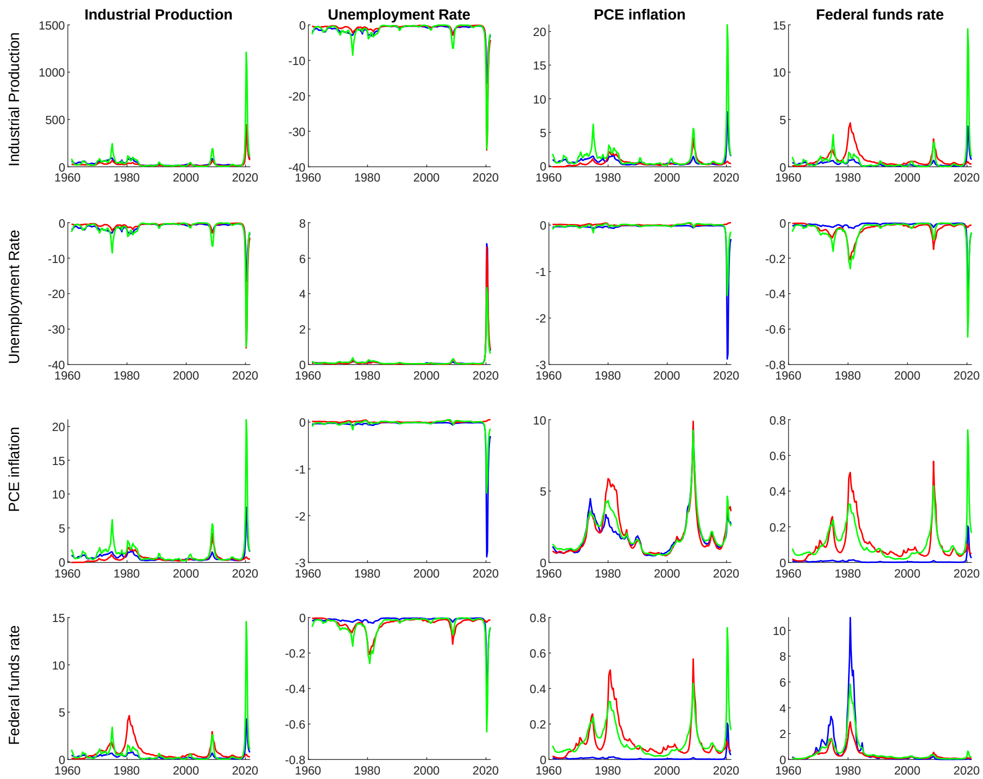

Figure A-4: Estimates of elements of time-varying covariance matrix (full sample). BVAR-cholSV(blue), BVAR-cholSV-RO(red), BVAR-eigSV(green).

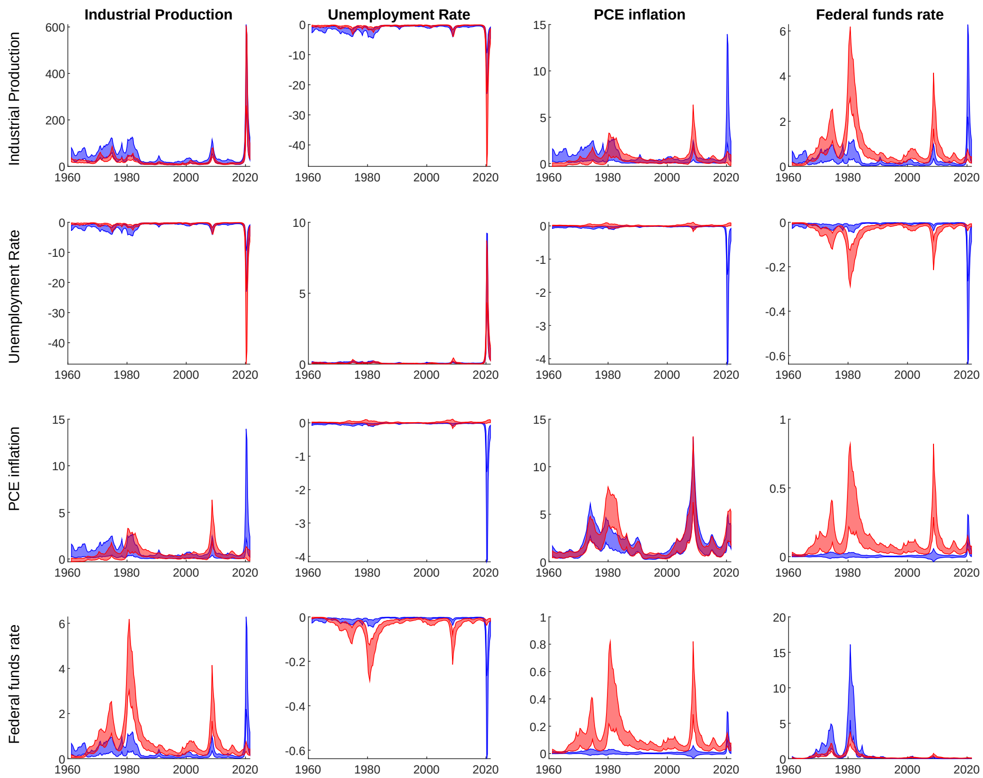

Figure A-5: Credible intervals of elements of time-varying covariance matrix. BVAR-cholSV (blue), BVAR-cholSV-RO (red).

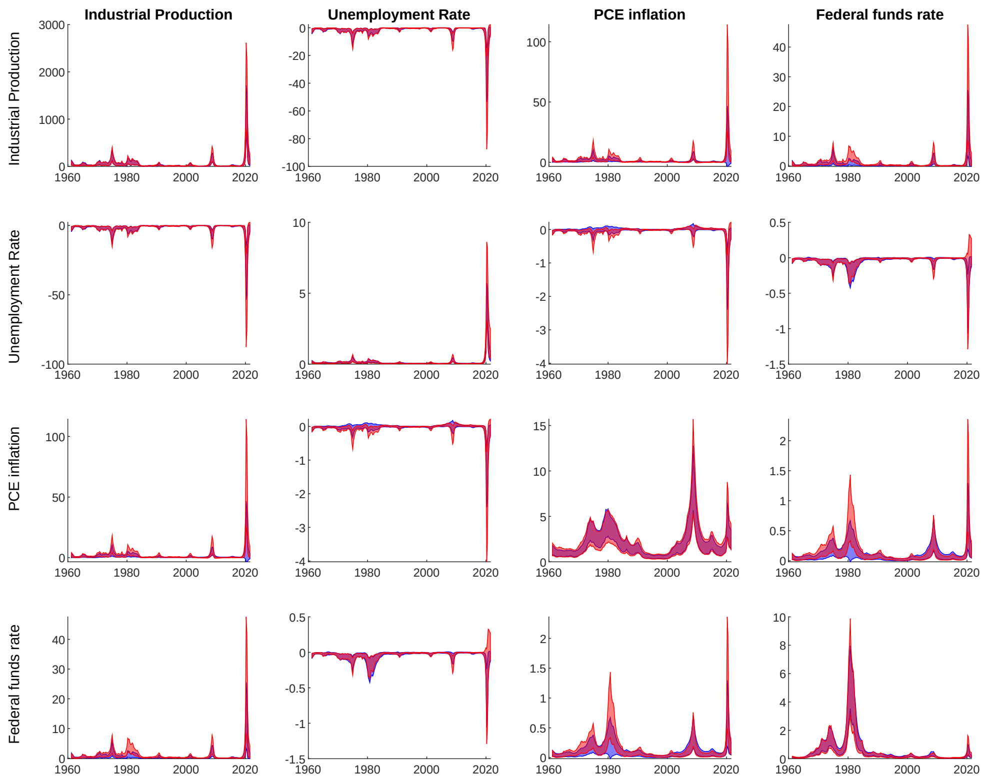

Figure A-6: Credible intervals of elements of time-varying covariance matrix. BVAR-eigSV (blue), BVAR-SeigSV (red).

Comparing small models

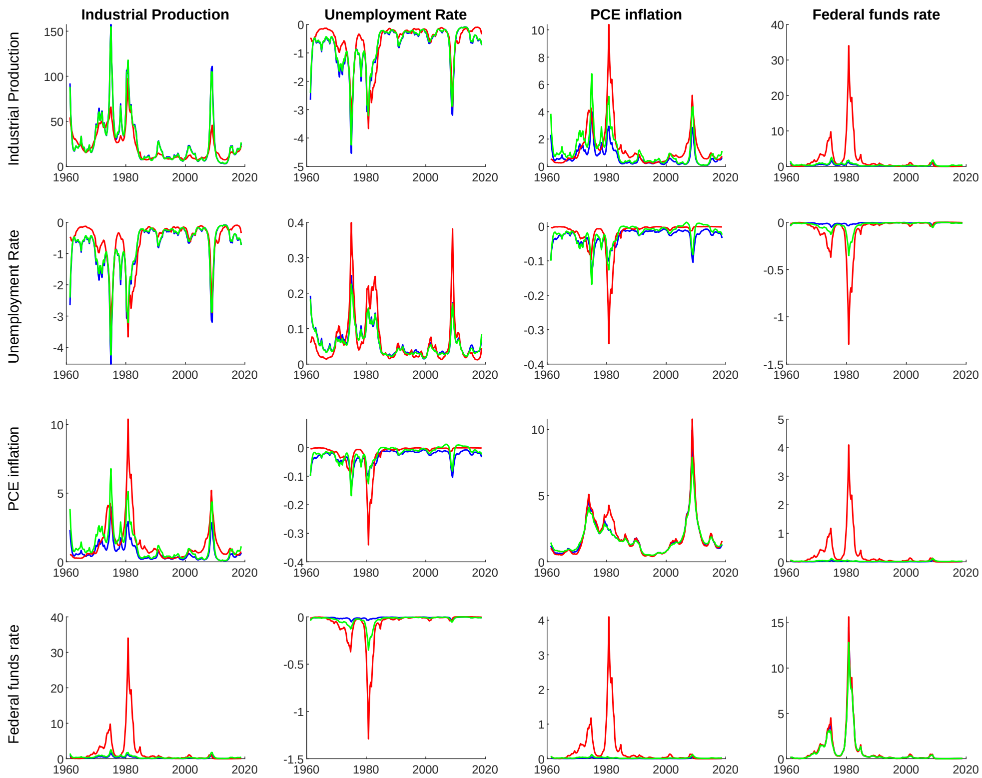

Figure A-7: Estimates of elements of time-varying covariance matrix (excluding data for the pandemic period). BVAR-cholSV(blue), BVAR-cholSV-RO(red), BVAR-eigSV(green).

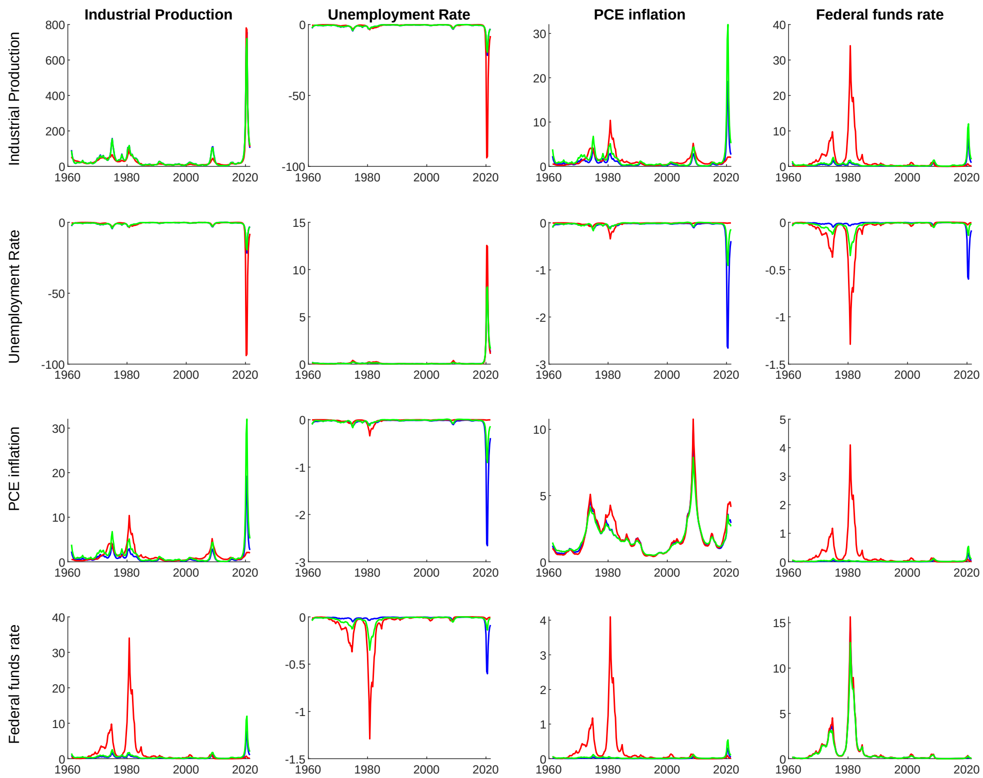

Figure A-8: Estimates of elements of time-varying covariance matrix (full sample). BVAR-cholSV(blue), BVAR-cholSV-RO(red), BVAR-eigSV(green).

Comparing a large BVAR-eigSV and a small BVAR-cholSV

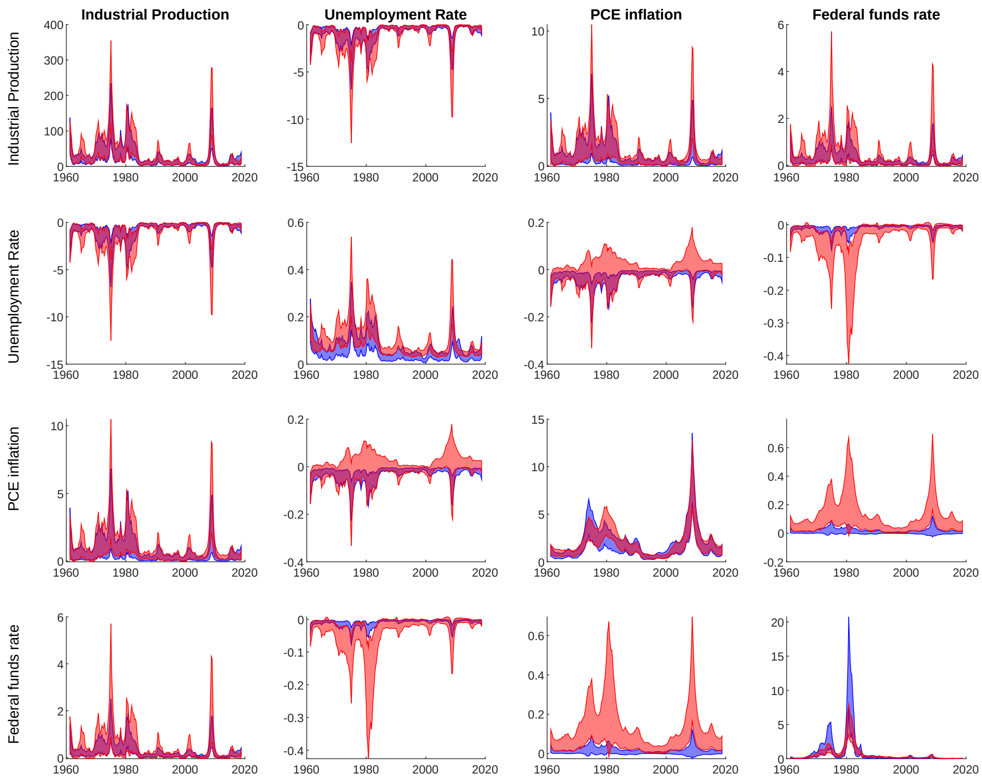

Figure A-9: Credible intervals of elements of time-varying covariance matrix (excluding data for the pandemic period). A small BVAR-cholSV (blue) and a large BVAR-eigSV (red).

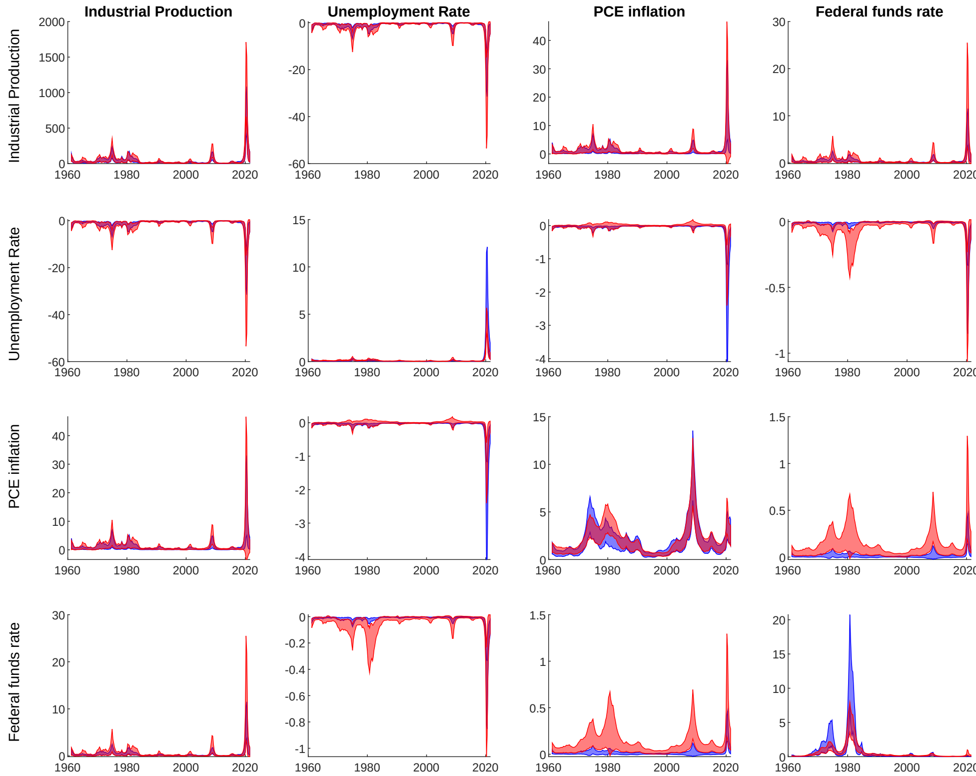

Figure A-10: Credible intervals of elements of time-varying covariance matrix. A small BVAR-cholSV (blue) and a large BVAR-eigSV (red).

E.3 Impact of highly volatile series

From a large BVAR-eigSV:

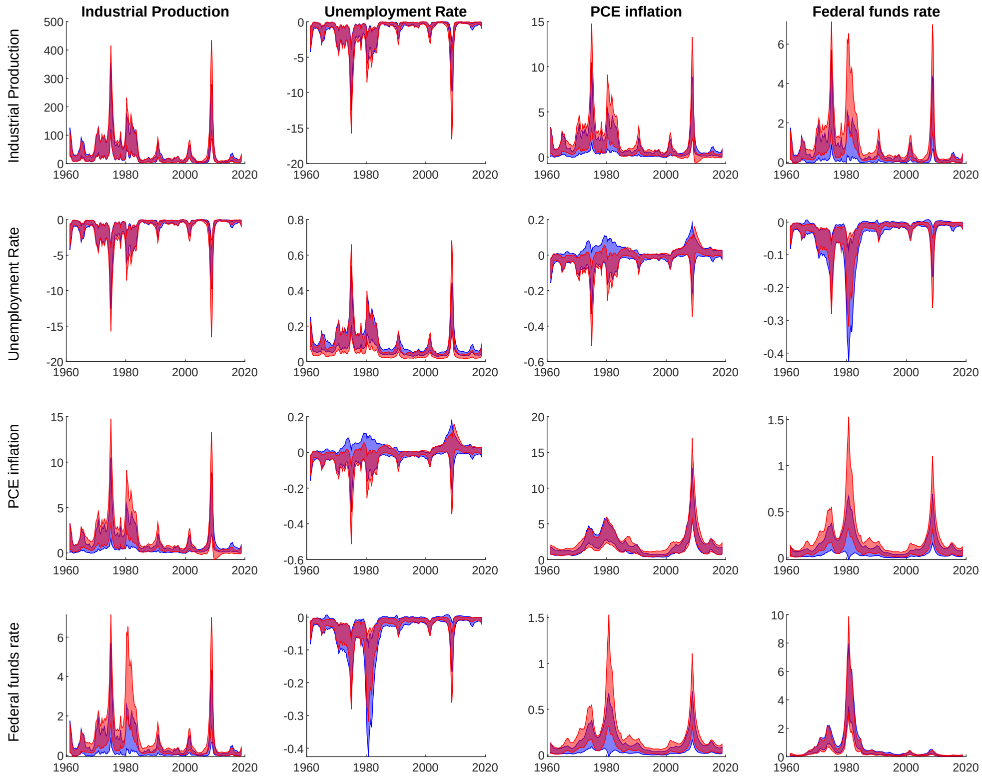

Figure A-11: 84% credible intervals of elements of time-varying covariance matrix (excluding data for the pandemic period). Without S&P 500 price index (blue) and with S&P 500 price index (red).

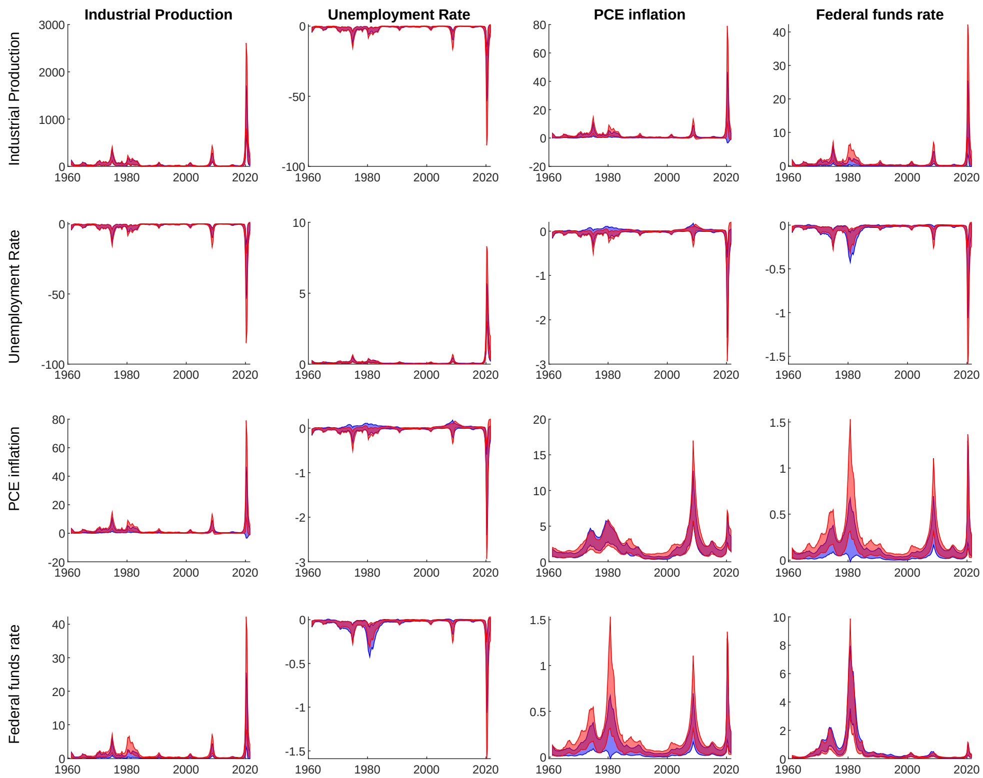

Figure A-12: 84% credible intervals of elements of time-varying covariance matrix. Without S&P 500 price index (blue) and with S&P 500 price index (red).

To see what will happen with a small VAR, the following figure plots two estimates: one is the small BVAR-SeigSV with four core variables. Another is the small BVAR-eigSV with five variables (four core variables plus the transformed S&P 500 price index). To show the patterns before pandemic more clearly, we plot the following figure.

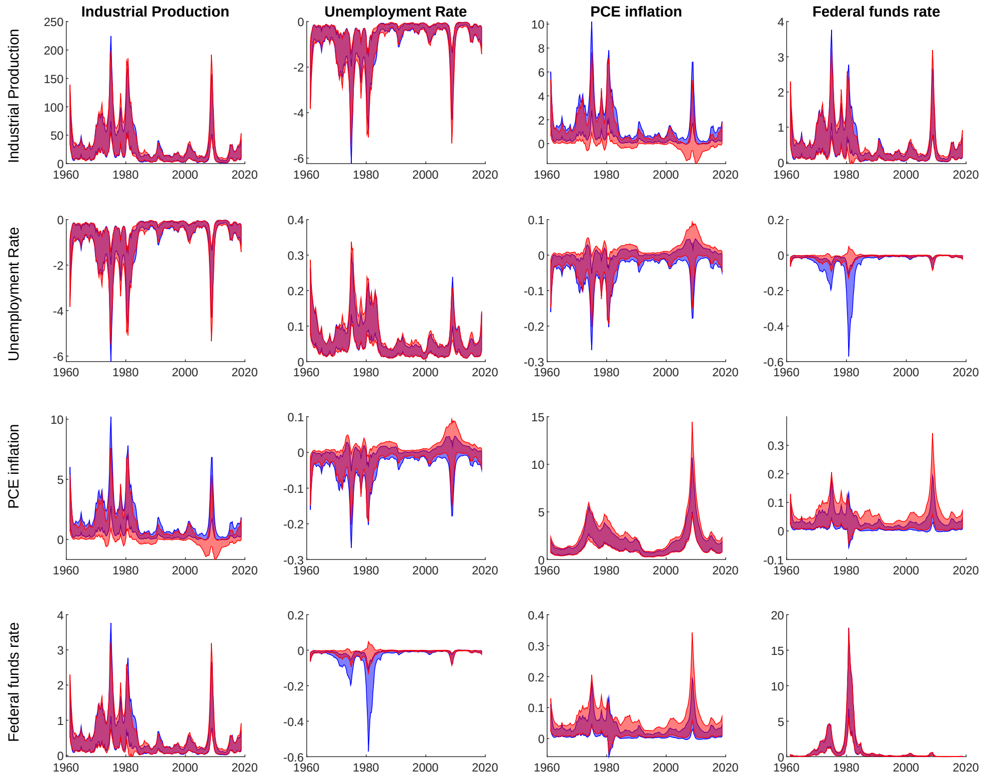

Figure A-13: 84% credible intervals of elements of time-varying covariance matrix (excluding data for the pandemic period). Without S&P 500 price index (blue) and with S&P 500 price index (red).

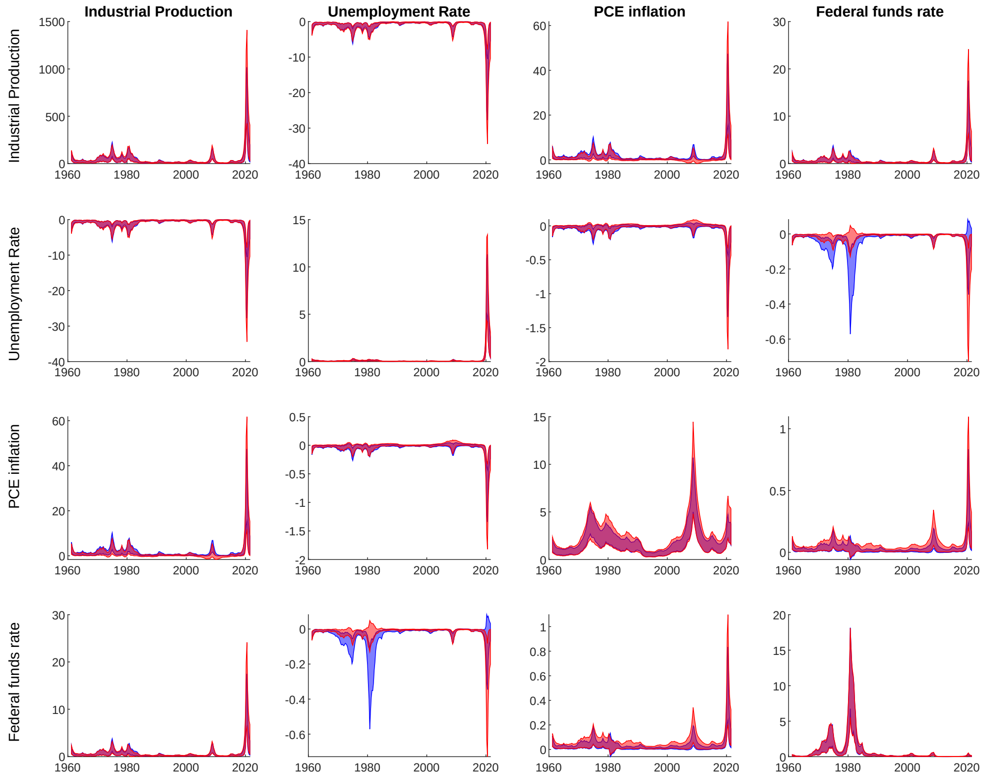

Figure A-14: 84% credible intervals of elements of time-varying covariance matrix. Without S&P 500 price index (blue) and with S&P 500 price index (red).

E.4 Forecasting performance

Table A-7: RMSFE and ALPL of 20 macroeconomic time series.

Variables	Models	RMSFE				ALPL			
		$h=1$	$h=2$	$h=3$	$h=4$	$h=1$	$h=2$	$h=3$	$h=4$
GDPC1	BVAR-CSV	5.901	5.010	4.970	4.936	-2.542	-3.527	-3.644	-3.581
	BVAR-cholSV	**4.921**	**4.910**	**4.913**	**4.911**	-2.454	-2.887	-3.287	-3.145
	BVAR-cholSV-RO	5.043	4.947	4.938	4.943	-2.674	-3.125	-3.352	-3.364
	BVAR-eigSV	4.980	4.922	4.914	4.918	-2.444	**-2.672**	-2.872	-2.928
	BVAR-SeigSV	5.000	4.937	4.928	4.932	**-2.402**	-2.730	**-2.847**	**-2.813**
	BVAR-OISV	5.031	4.949	4.951	4.951	-2.724	-3.269	-3.382	-3.426
	BVAR-OISV(TVP)	6.610	5.483	5.151	5.116	-2.471	-3.032	-3.262	-3.200
PCECC96	BVAR-CSV	5.602	5.389	**5.335**	**5.283**	-2.560	-3.853	-4.108	-3.850
	BVAR-cholSV	5.327	**5.370**	5.373	5.302	-2.387	-3.466	-4.207	-4.069
	BVAR-cholSV-RO	5.290	5.384	5.373	5.294	-2.513	-3.289	-3.751	-3.763
	BVAR-eigSV	**5.132**	5.395	5.377	5.285	-2.370	**-2.901**	-3.350	-3.490
	BVAR-SeigSV	5.342	5.421	5.424	5.329	**-2.341**	-3.003	**-3.193**	**-3.070**
	BVAR-OISV	5.343	5.405	5.411	5.320	-2.419	-3.134	-3.647	-3.891
	BVAR-OISV(TVP)	7.325	5.763	5.388	5.433	-2.386	-3.047	-3.258	-3.467
INDPRO	BVAR-CSV	8.264	**7.782**	**7.892**	8.000	-2.973	-3.951	-4.350	-4.282
	BVAR-cholSV	7.976	7.908	7.904	**7.894**	-3.081	-3.463	-3.581	-3.615
	BVAR-cholSV-RO	8.185	7.968	7.993	8.016	-3.180	-3.743	-3.896	-3.967
	BVAR-eigSV	8.064	7.930	7.901	7.929	-3.025	-3.342	-3.537	-3.649
	BVAR-SeigSV	**7.936**	7.896	7.901	7.930	**-2.961**	**-3.318**	**-3.414**	**-3.444**
	BVAR-OISV	8.118	7.964	7.954	7.978	-3.258	-3.692	-3.813	-3.886
	BVAR-OISV(TVP)	9.352	8.032	7.900	8.092	-2.988	**-3.306**	-3.463	-3.612
IPFINAL	BVAR-CSV	9.077	8.087	8.056	8.237	-3.005	-3.937	-4.401	-4.348
	BVAR-cholSV	8.248	**8.027**	**8.047**	**8.051**	-3.092	-3.457	-3.623	-3.662
	BVAR-cholSV-RO	8.499	8.080	8.131	8.143	-3.171	-3.578	-3.758	-3.773
	BVAR-eigSV	8.247	8.043	8.059	8.073	-3.013	-3.293	-3.516	-3.627
	BVAR-SeigSV	**8.096**	8.037	8.069	8.088	**-2.963**	**-3.279**	**-3.410**	**-3.407**
	BVAR-OISV	8.351	8.087	8.114	8.140	-3.245	-3.668	-3.795	-3.720
	BVAR-OISV(TVP)	12.631	8.575	8.137	8.244	**-2.950**	-3.475	-3.736	-3.863
PAYEMS	BVAR-CSV	**6.000**	**5.383**	**5.201**	**5.192**	-3.962	-13.481	-13.019	-11.084
	BVAR-cholSV	6.144	5.410	5.314	5.335	-2.799	-5.726	-9.216	-8.406
	BVAR-cholSV-RO	6.386	5.504	5.323	5.334	-4.570	-9.245	-10.300	-14.145
	BVAR-eigSV	6.210	5.411	5.298	5.327	-1.829	**-3.847**	-5.063	-5.076
	BVAR-SeigSV	6.146	5.436	5.326	5.350	**-1.798**	-5.004	**-4.799**	**-4.537**
	BVAR-OISV	6.328	5.483*	5.352	5.369	-7.000	-12.029	-8.374	-10.247
	BVAR-OISV(TVP)	7.221	5.792	5.652	5.466	-3.423	-6.717	-7.180	-5.390
MANEMP	BVAR-CSV	5.346	4.911	**4.629**	**4.654**	-2.223	-3.535	-3.976	-3.985
	BVAR-cholSV	5.021	4.868	4.935	4.951	-2.435	-3.038	-3.148	-3.372
	BVAR-cholSV-RO	4.953	4.840	4.848	4.888	-2.274	-3.431	-3.566	-3.985
	BVAR-eigSV	4.897	4.817	4.920	4.955	-2.249	**-2.754**	-3.095	-3.219
	BVAR-SeigSV	**4.767**	**4.777**	4.911	4.958	**-2.218**	-2.764	**-2.943**	**-3.040**
	BVAR-OISV	4.852	4.778	4.890	4.958	-2.633	-3.344	-3.398	-3.771
	BVAR-OISV(TVP)	6.154	5.756	5.105	4.808	**-2.207**	-3.008	-3.559	-3.646
CE16OV	BVAR-CSV	6.151	5.772	5.665	**5.593**	-3.536	-7.694	-8.198	-6.272
	BVAR-cholSV	5.799	5.721	5.658	5.635	-3.102	-4.938	-4.713	-5.779
	BVAR-cholSV-RO	**5.777**	**5.668**	5.668	5.630	-3.618	-4.998	-4.802	-4.597
	BVAR-eigSV	5.784	5.685	**5.650**	5.643	-2.423	**-3.827**	-4.707	-4.851
	BVAR-SeigSV	5.790	5.713	5.655	5.648	**-2.242**	-4.243	**-4.452**	**-4.036**
	BVAR-OISV	5.871	5.734	5.651	5.641	-3.068	-4.934	-4.956	-5.806
	BVAR-OISV(TVP)	7.110	6.166	5.939	5.689	-2.515	-4.012	**-4.081**	-4.955
CIVPART	BVAR-CSV	0.280	0.362	0.434	0.511	0.173	-1.005	-1.726	-2.248
	BVAR-cholSV	0.263	0.353	0.419	0.488	0.141	-1.153	-2.086	-3.040
	BVAR-cholSV-RO	0.618	0.485*	0.495**	0.507	-0.234	-1.010	-1.632	-2.307
	BVAR-eigSV	0.264	0.355	0.422	0.493	0.063	-0.889	-1.833	-2.694
	BVAR-SeigSV	0.265	0.358	0.428	0.498	0.203	**-0.694**	**-1.362**	**-1.975**
	BVAR-OISV	**0.263**	**0.352**	**0.413**	**0.479**	0.139	-1.111	-1.882	-2.861
	BVAR-OISV(TVP)	0.372	0.460	0.614	0.685	0.150	**-0.432**	**-0.922**	**-1.330**
UNRATE	BVAR-CSV	**0.971**	1.284	1.447	1.530	-1.510	-8.460	-11.670	-10.328
	BVAR-cholSV	0.975	1.290	1.476	1.602*	-1.284	-4.776	-6.291	-8.011
	BVAR-cholSV-RO	0.986	**1.265**	**1.411**	**1.522**	-2.709	-6.465	-8.429	-10.513
	BVAR-eigSV	0.971	1.269	1.435	1.560	-0.385	**-2.722**	-5.446	-7.071
	BVAR-SeigSV	0.976	1.298	1.479	1.613*	**-0.230**	-3.277	**-4.132**	**-5.179**
	BVAR-OISV	0.992	1.307*	1.485	1.611*	-4.121	-8.383	-9.210	-12.722
	BVAR-OISV(TVP)	1.132	1.551	1.642	1.726	-0.808	**-2.544**	-4.532	**-4.547**
HOANBS	BVAR-CSV	7.252	6.588	**6.382**	**6.332**	-2.912	-5.155	-5.680	-4.998
	BVAR-cholSV	6.793	6.538	6.482	6.467	-2.801	-4.104	-4.686	-4.663
	BVAR-cholSV-RO	6.725	**6.503**	6.448	6.463	-2.614	-3.526	-3.986	-4.410
	BVAR-eigSV	6.754	6.525	6.467	6.466	-2.490	**-2.968**	-3.522	-3.584
	BVAR-SeigSV	**6.716**	6.554	6.483	6.482	**-2.454**	-3.338	**-3.333**	**-3.289**
	BVAR-OISV	6.841	6.554	6.482	6.484	-3.371	-4.990	-4.974	-5.205
	BVAR-OISV(TVP)	8.102	7.112	6.814	6.382	**-2.389**	-3.054	-3.860	-4.181

Table A-7: RMSFE and ALPL of 20 macroeconomic time series.

Variables	Models	RMSFE				ALPL			
		$h=1$	$h=2$	$h=3$	$h=4$	$h=1$	$h=2$	$h=3$	$h=4$
HOUST	BVAR-CSV	30.351	32.727	34.549	32.377	-4.713	-5.098	-5.219	-5.325
	BVAR-cholSV	27.276*	31.793	31.552	31.688	-4.663	-5.090	-5.132	-5.174
	BVAR-cholSV-RO	27.780*	**31.787**	**31.500**	**31.534**	**-4.658**	-4.977	-5.012	-5.058**
	BVAR-eigSV	**27.227***	31.829	31.627	31.712	-4.674	-5.112	-5.168	-5.223
	BVAR-SeigSV	28.143	32.201	31.958	32.045	-4.673	-4.979	-5.012	-5.043**
	BVAR-OISV	27.804	32.008	31.677	31.672	-4.660	**-4.969**	**-4.998***	**-5.036****
	BVAR-OISV(TVP)	34.489	53.415	63.443	38.171*	-4.796	-5.413**	-5.833***	-6.015**
PERMIT	BVAR-CSV	33.263	30.360	31.479	30.528	-4.660	-4.981	-5.073	-5.244
	BVAR-cholSV	29.134	**29.587**	29.651	29.776	-4.713	-4.804**	-4.838**	-4.850**
	BVAR-cholSV-RO	29.332	29.633	**29.556**	**29.682**	**-4.653**	-4.804	-4.866	-4.938*
	BVAR-eigSV	**28.850**	29.799	29.730	29.835	-4.675	-4.811	-4.864*	-4.908**
	BVAR-SeigSV	29.588	29.965	29.955	29.981	-4.686	**-4.773****	**-4.815****	**-4.838****
	BVAR-OISV	29.431	29.927	29.796	29.825	-4.663	-4.817	-4.861	-4.922*
	BVAR-OISV(TVP)	34.580	40.020	45.796	34.815	-4.806**	-5.225**	-5.545**	-5.816*
PCECTPI	BVAR-CSV	1.896	1.958	1.922	2.151	-1.836	-2.215	-2.340	-2.712
	BVAR-cholSV	1.551	1.754	1.749	1.800	-1.759	-2.123	-2.242	-2.465
	BVAR-cholSV-RO	1.626	1.813	1.831	1.864	-1.775	-2.134	-2.247	-2.496
	BVAR-eigSV	**1.530**	**1.717**	**1.728**	1.774	-1.719**	-2.035*	-2.188	-2.385
	BVAR-SeigSV	1.557	1.741	1.749	1.792	**-1.718***	**-1.994***	**-2.076****	**-2.239**
	BVAR-OISV	1.562	1.728	1.732*	**1.761**	-1.733*	-2.037*	-2.145*	-2.337
	BVAR-OISV(TVP)	2.580	2.631	2.441	2.599	-1.881	-2.234	-2.355	-2.620
CPIAUCSL	BVAR-CSV	2.332	2.512	2.452	2.723	-2.119	-2.471	-2.596	-2.918
	BVAR-cholSV	2.041	2.249	**2.216***	2.252	-2.013	-2.310*	-2.379*	-2.555
	BVAR-cholSV-RO	2.115	2.319	2.321	2.354	-2.004	-2.302	-2.397	-2.617
	BVAR-eigSV	**2.039**	**2.230**	2.216	2.247	-1.996*	-2.238*	-2.364**	-2.511
	BVAR-SeigSV	2.066	2.237	2.226	2.252	**-1.985***	**-2.197****	**-2.251****	**-2.376**
	BVAR-OISV	2.077	2.235	2.220*	2.235	-1.993	-2.246**	-2.324**	-2.476
	BVAR-OISV(TVP)	2.844	3.505	2.966	3.287	-2.086	-2.403	-2.576	-2.845
OPHNFB	BVAR-CSV	2.913	3.115	3.142	2.864	-2.507	-2.629	-2.655	-2.699
	BVAR-cholSV	2.756	2.894	**2.768**	**2.771**	-2.435	-2.476**	-2.489**	-2.502***
	BVAR-cholSV-RO	2.743	2.868	2.799	2.812	-2.455	-2.483**	-2.503**	-2.516**
	BVAR-eigSV	2.742	2.848	2.777	2.786	-2.430*	-2.463**	**-2.471****	**-2.480*****
	BVAR-SeigSV	2.732	**2.818**	2.775	2.798	**-2.426***	**-2.461****	-2.472**	-2.488***
	BVAR-OISV	**2.731**	2.860	2.788	2.792	-2.433	-2.475**	-2.481**	-2.489***
	BVAR-OISV(TVP)	3.286	3.156	3.385	3.139	-2.489	-2.611	-2.687	-2.749
FEDFUNDS	BVAR-CSV	0.594	1.116	1.519	1.842	-0.625	-1.697	-2.610	-3.446
	BVAR-cholSV	**0.321**	**0.623**	**0.928**	**1.236**	**-0.025***	-1.710*	-3.306**	-4.814**
	BVAR-cholSV-RO	0.393	0.753	1.094	1.429	-0.529	-1.706	-2.866	-3.937*
	BVAR-eigSV	0.351	0.680	1.004	1.317	-0.381	-1.539	-2.776*	-3.973*
	BVAR-SeigSV	0.348	0.668	0.985	1.294	-0.454	**-1.388****	**-2.188****	**-2.884****
	BVAR-OISV	0.337	0.647	0.956	1.265	-0.104	-1.573	-2.993*	-4.319*
	BVAR-OISV(TVP)	0.792	1.672	2.248	2.488	-0.806**	-1.968**	-2.726	-3.282
TB3MS	BVAR-CSV	0.520	0.955	1.313	1.606	-0.553	-1.679	-2.603	-3.445
	BVAR-cholSV	**0.333**	**0.625**	**0.903**	**1.179**	**-0.098***	-1.657*	-3.138**	-4.523**
	BVAR-cholSV-RO	0.421	0.759	1.061	1.357	-0.594	-1.729	-2.807	-3.840
	BVAR-eigSV	0.366	0.681	0.971	1.249	-0.334	-1.643	-3.015*	-4.326**
	BVAR-SeigSV	0.356	0.662	0.948	1.223	-0.406	**-1.393***	**-2.244***	**-2.963***
	BVAR-OISV	0.353	0.653	0.935	1.214	-0.160	-1.679	-3.089**	-4.372**
	BVAR-OISV(TVP)	0.650	1.352	1.760	2.035	-0.718	-1.735	-2.347*	**-2.893****
GS1	BVAR-CSV	0.662	1.130	1.477	1.740	-0.726	-1.866	-2.743	-3.536
	BVAR-cholSV	**0.406**	**0.719**	**0.989**	**1.256**	**-0.356**	-1.749	-2.996*	-4.137*
	BVAR-cholSV-RO	0.479	0.838	1.137	1.432	-0.745	-1.833	-2.812	-3.761
	BVAR-eigSV	0.421	0.744	1.023	1.295	-0.494	-1.806	-3.107*	-4.349**
	BVAR-SeigSV	0.413	0.732	1.004	1.272	-0.555	**-1.523****	**-2.312***	**-2.980***
	BVAR-OISV	0.418	0.736	1.010	1.282	-0.467	-1.732	-2.855	-3.878
	BVAR-OISV(TVP)	0.924	1.784	2.191	2.301	-0.835	-1.797	-2.375*	**-2.894****
GS10	BVAR-CSV	0.469	0.774	0.998	1.162	-0.641	-1.756	-2.501	-3.124
	BVAR-cholSV	0.367	0.592	0.735	0.866	-0.450	-1.728	-2.622	-3.358
	BVAR-cholSV-RO	0.391	0.636	0.798	0.948	-0.557	-1.711	-2.602	-3.433
	BVAR-eigSV	0.367	0.593	0.737	0.870	**-0.448***	-1.687	-2.651	-3.489
	BVAR-SeigSV	**0.365**	0.591	0.731	0.861	-0.471**	**-1.406***	**-2.021***	**-2.499***
	BVAR-OISV	0.371	0.596	0.742	0.876	-0.485*	-1.592	-2.408	-3.102
	BVAR-OISV(TVP)	0.618	1.126	1.355	1.454	-0.667	-1.745*	-2.489	-3.041
BAA10YM	BVAR-CSV	0.374	0.577	0.693	0.771	-0.301	-1.568	-2.384	-3.156
	BVAR-cholSV	0.336*	0.519	0.627	0.704	-0.188	-1.451	-2.387	-3.198
	BVAR-cholSV-RO	0.359	0.557	0.687	0.785	-0.216	-1.452	-2.375	-3.289
	BVAR-eigSV	0.337*	0.520	0.626	**0.699**	-0.152	-1.234*	-2.069	-2.728
	BVAR-SeigSV	**0.333***	0.518	0.627	0.702	**-0.111**	**-1.056**	**-1.648**	**-2.110**
	BVAR-OISV	0.339*	0.526	0.639	0.722	-0.159	-1.304	-2.186	-2.922
	BVAR-OISV(TVP)	0.480	0.716	0.742	0.769	-0.233	-1.257*	-1.921	-2.472

[1] Notes: The bold figure indicates the best model in each case. $*$, $**$ and $***$ denote, respectively, 0.10, 0.05 and 0.01 significance level for a two-sided Diebold and Mariano(1995) test.The benchmark model in the Diebold Mariano test is BVAR-CSV with 20 variables.

E.5 Adding Time Variation in the Eigenmatrix U: Complementing Figure 5

Below are more comparisons against models with a time-varying impact matrix. First, we compare large models with 20 variables: BVAR-eigSV and BVAR-cholSV(TVP) with 20 variables. Then, we compare a large and a small model: BVAR-eigSV with 20 variables and BVAR-OISV(TVP) with 4 variables, and BVAR-eigSV with 20 variables and BVAR-cholSV(TVP) with 4 variables.

Comparing large models

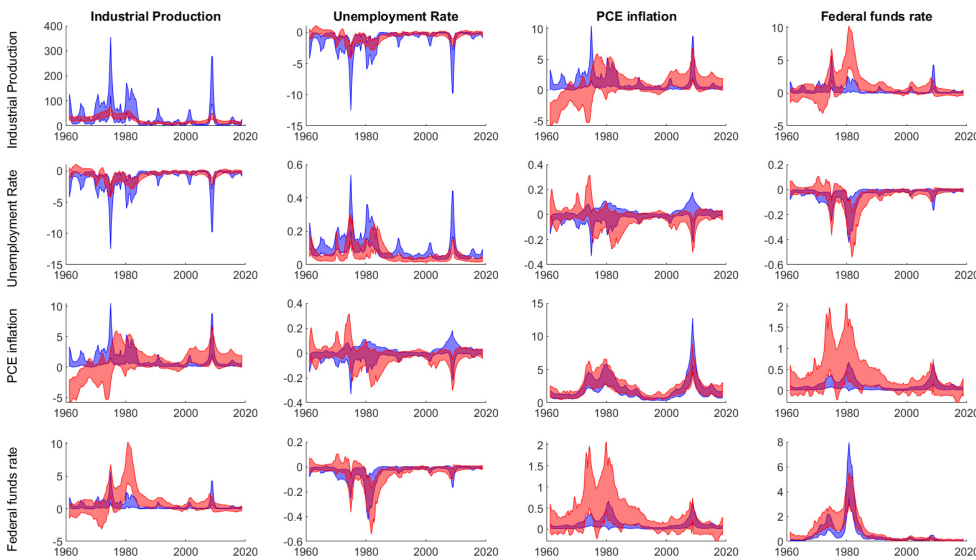

Figure A-15: 84% credible intervals of elements of time-varying covariance matrix (excluding data for the pandemic period). Blue: BVAR-eigSV. Red: BVAR-cholSV with TVP impact.

Comparing a large model and a small model

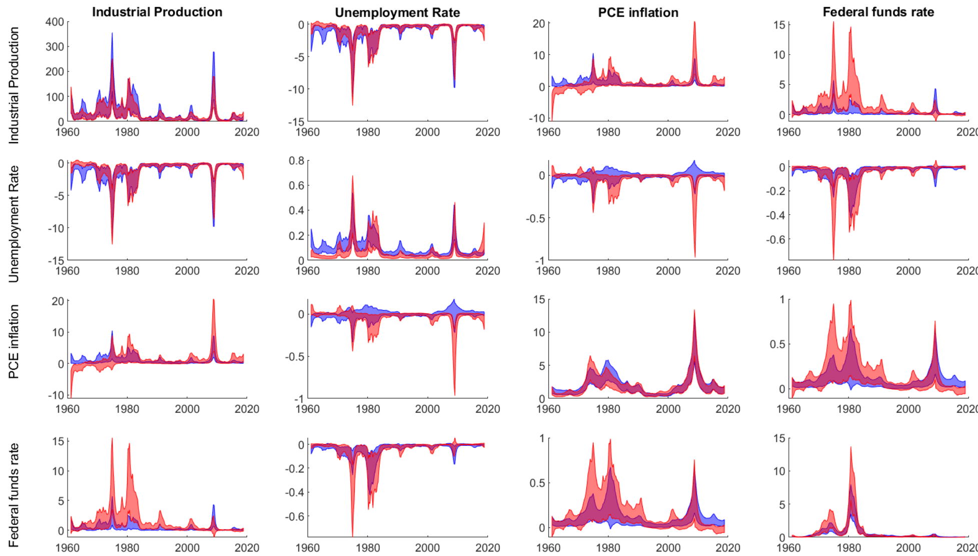

Figure A-16: 84% credible intervals of elements of time-varying covariance matrix (excluding data for the pandemic period). Blue: BVAR-eigSV with 20 variables. Red: BVAR-OISV with TVP impact and 4 variables.

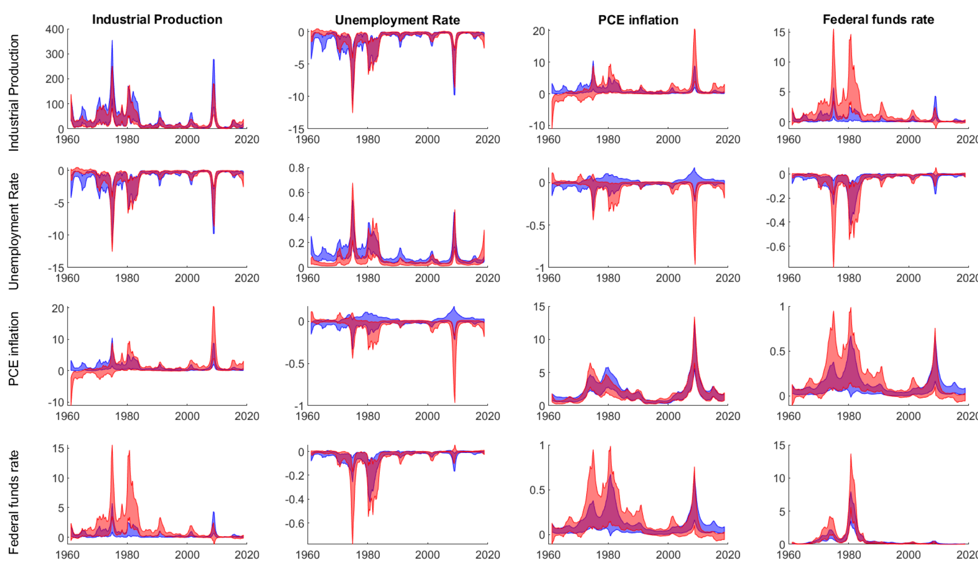

Figure A-17: 84% credible intervals of elements of time-varying covariance matrix (excluding data for the pandemic period). Blue: BVAR-eigSV with 20 variables. Red: BVAR-cholSV with TVP impact and 4 variables.

E.6 Pre-pandemic forecasting results

In this subsection, we report forecasting performance using data before pandemic, more specifically, the forecasting period is up to 2018Q4. We first provide the joint forecasting performance, followed by forecasting individual variables.

Table A-8: Joint ALPL for 20 macroeconomic variables (up to 2018Q4).

Models	$h=1$	$h=2$	$h=3$	$h=4$
BVAR-CSV	-3,166	**-4,191**	**-4,971**	**-5,726**
BVAR-cholSV	-2,801	-4,632	-6,255	-7,991
BVAR-cholSV-RO	-2,984	-4,397	-5,679	-7,190
BVAR-eigSV	-2,908	-4,600	-6,043	-7,583
BVAR-OISV	**-2,756**	-4,703	-6,714	-9,125

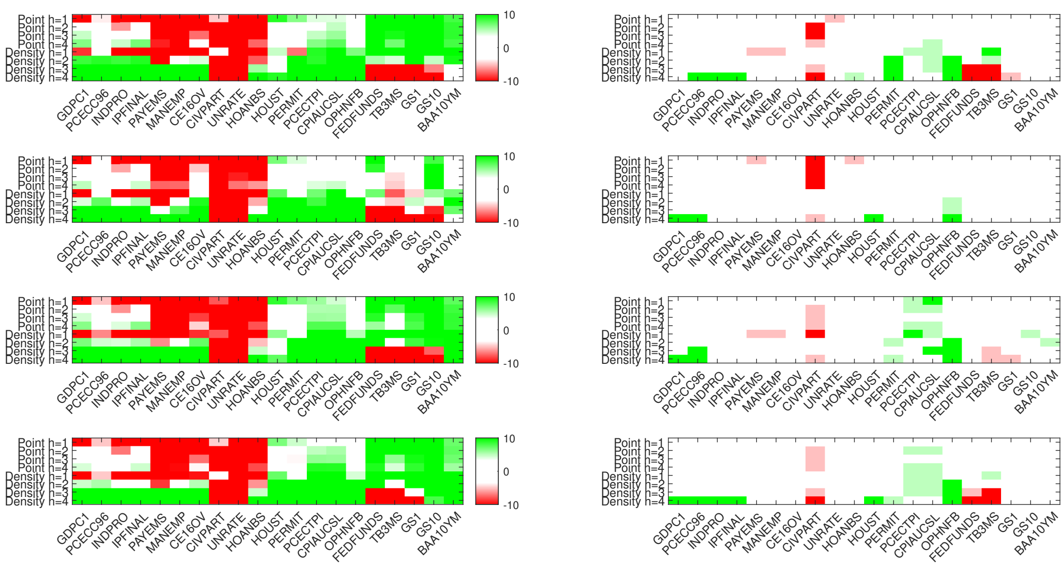

Figure A-18: Forecasting results from the large BVAR-SV (up to 2018Q4). From Top to Bottom: BVAR-cholSV against BVAR-CSV, BVAR-cholSV-RO against BVAR-CSV, BVAR-eigSV against BVAR-CSV, BVAR-OISV against BVAR-CSV. Left: values of percentage gains in RMSFEs and ALPLs. Right: significance level according to the Diebold Mariano test: Value 0 means not significant. Value 1 means 0.10 significance level for a two-sided Diebold and Mariano(1995) test. Value 2 means 0.05 significance level.

Table A-9: RMSFE and ALPL of 20 macroeconomic time series.

Variables	Models	RMSFE				ALPL			
		$h=1$	$h=2$	$h=3$	$h=4$	$h=1$	$h=2$	$h=3$	$h=4$
GDPC1	BVAR-CSV	**2.086**	**2.449**	2.631	2.686	**-2.149**	-2.385	-2.514	-2.658
	BVAR-cholSV	2.424	2.450	**2.505**	**2.511**	-2.236	**-2.300**	**-2.364**	**-2.408**
	BVAR-cholSV-RO	2.360	2.464	2.545	2.551	-2.280	-2.343	-2.388	-2.414**
	BVAR-eigSV	2.425	2.455	2.513	2.513	-2.259	-2.318	-2.388	-2.415**
	BVAR-OISV	2.441	2.489	2.562	2.568	-2.268	-2.338	-2.391	-2.418**
PCECC96	BVAR-CSV	**1.886**	1.940	2.009	2.199	**-2.005**	-2.140	-2.245	-2.440
	BVAR-cholSV	1.958	1.916	**1.963**	**2.129**	-2.024	**-2.050**	**-2.099**	**-2.215****
	BVAR-cholSV-RO	1.914	**1.906**	1.966	2.141	-2.028	-2.065	-2.120	-2.240**
	BVAR-eigSV	1.978	1.943	1.989	2.161	-2.069	-2.094	-2.141**	-2.238**
	BVAR-OISV	1.977	1.954	2.018	2.189	-2.053	-2.084	-2.135	-2.240**
INDPRO	BVAR-CSV	**3.843**	**5.029**	5.417	5.616	**-2.685**	-3.110	-3.396	-3.582
	BVAR-cholSV	4.987	5.333	5.343	**5.350**	-2.871	**-3.025**	**-3.104**	**-3.161****
	BVAR-cholSV-RO	4.637	5.314	5.421	5.472	-2.847	-3.081	-3.188	-3.263
	BVAR-eigSV	5.098	5.342	**5.336**	5.378	-2.855	-3.067	-3.196	-3.318
	BVAR-OISV	4.992	5.377	5.392	5.436	-2.897	-3.101	-3.195	-3.225**
IPFINAL	BVAR-CSV	**3.740**	**4.703**	5.013	5.262	**-2.677**	-3.023	-3.256	-3.447
	BVAR-cholSV	4.680	4.736	**4.851**	**4.898**	-2.848	**-2.932**	**-3.017**	**-3.067****
	BVAR-cholSV-RO	4.458	4.784	4.943	5.018	-2.834	-2.971	-3.074	-3.132
	BVAR-eigSV	4.711	4.749	4.875	4.925	-2.825	-2.968	-3.107	-3.221
	BVAR-OISV	4.693	4.822	4.950	5.019	-2.882	-2.993	-3.089	-3.122**
PAYEMS	BVAR-CSV	**0.776**	**1.231**	**1.524**	**1.713**	**-1.081**	**-1.787**	-2.327	-2.714
	BVAR-cholSV	1.175	1.581	1.749	1.869	-1.426*	-1.908	**-2.163**	**-2.380**
	BVAR-cholSV-RO	1.021*	1.454	1.688	1.838	-1.316	-1.912	-2.314	-2.616
	BVAR-eigSV	1.152	1.575	1.764	1.890	-1.365*	-1.850	-2.166	-2.433
	BVAR-OISV	1.153	1.574	1.767	1.903	-1.336	-1.872	-2.183	-2.433
MANEMP	BVAR-CSV	**1.609**	**2.607**	**3.172**	**3.470**	**-1.891**	-2.511	-2.944	-3.226
	BVAR-cholSV	2.535	3.333	3.692	3.819	-2.167*	-2.531	**-2.760**	**-2.880**
	BVAR-cholSV-RO	2.197	3.111	3.553	3.722	-2.028	-2.542	-2.866	-3.020
	BVAR-eigSV	2.556	3.327	3.719	3.822	-2.103*	**-2.481**	-2.763	-2.949
	BVAR-OISV	2.511	3.283	3.674	3.809	-2.035	-2.512	-2.825	-2.976
CE16OV	BVAR-CSV	**1.331**	**1.595**	**1.700**	**1.810**	**-1.721**	-2.026	-2.246	-2.397
	BVAR-cholSV	1.646	1.770	1.821	1.870	-1.848	-1.984	-2.069	-2.131
	BVAR-cholSV-RO	1.462	1.671	1.747	1.826	-1.734	**-1.934**	-2.050	-2.159
	BVAR-eigSV	1.646	1.811	1.842	1.891	-1.809	-1.953	**-2.032**	**-2.121**
	BVAR-OISV	1.586	1.773	1.819	1.874	-1.816	-1.974	-2.057	-2.133
CIVPART	BVAR-CSV	**0.149**	**0.205**	**0.239**	**0.284**	**0.452**	**-0.248**	**-0.791**	**-1.326**
	BVAR-cholSV	0.156	0.237**	0.301**	0.370*	0.452	-0.534	-1.408*	-2.258**
	BVAR-cholSV-RO	0.250**	0.322**	0.356**	0.393**	-0.004	-0.777	-1.361	-1.994*
	BVAR-eigSV	0.160	0.246*	0.315*	0.390*	0.376***	-0.475	-1.279	-2.091*
	BVAR-OISV	0.156	0.234*	0.295*	0.361*	0.430	-0.549	-1.388*	-2.189**
UNRATE	BVAR-CSV	**0.200**	**0.390**	**0.580**	**0.779**	**0.396**	**-0.992**	**-2.439**	**-4.000**
	BVAR-cholSV	0.239*	0.459	0.678	0.887	0.135	-1.188	-2.662	-4.284
	BVAR-cholSV-RO	0.221	0.429	0.634	0.833	0.237	-1.212	-2.830	-4.641
	BVAR-eigSV	0.235	0.453	0.671	0.879	0.161	-1.147	-2.618	-4.264
	BVAR-OISV	0.229	0.448	0.669	0.877	0.219	-1.229	-2.787	-4.555
HOANBS	BVAR-CSV	**1.714**	**2.260**	**2.519**	**2.693**	**-2.007**	**-2.296**	-2.511	-2.718
	BVAR-cholSV	2.636	2.803	2.883	2.891	-2.249	-2.372	**-2.454**	**-2.499***
	BVAR-cholSV-RO	2.372*	2.626	2.788	2.856	-2.175	-2.351	-2.488	-2.560
	BVAR-eigSV	2.705	2.828	2.893	2.904	-2.227	-2.346	-2.464	-2.543
	BVAR-OISV	2.588	2.799	2.899	2.909	-2.203	-2.357	-2.465	-2.517

Table A-9: Continued: RMSFE and ALPL of 20 macroeconomic time series.

Variables	Models	RMSFE				ALPL			
		$h=1$	$h=2$	$h=3$	$h=4$	$h=1$	$h=2$	$h=3$	$h=4$
HOUST	BVAR-CSV	27.211	29.426	28.829	**28.378**	-4.647	-4.948	-5.012	-5.131
	BVAR-cholSV	**24.750**	**29.215**	28.753	28.857	-4.599	-4.980	-5.004	-5.041
	BVAR-cholSV-RO	25.347	29.341	**28.738**	28.710	**-4.572**	-4.846	-4.871	-4.904**
	BVAR-eigSV	24.833	29.250	28.930	28.914	-4.585	-4.969	-5.015	-5.057
	BVAR-OISV	25.358	29.485	28.957	28.887	-4.575	**-4.836**	**-4.851**	**-4.878****
PERMIT	BVAR-CSV	29.506	28.791	**27.604**	**27.810**	-4.616	-4.910	-4.946	-5.125
	BVAR-cholSV	27.597	**28.289**	28.371	28.404	-4.687	-4.755**	**-4.783****	**-4.793****
	BVAR-cholSV-RO	28.240	28.454	28.320	28.328	**-4.603**	**-4.726**	-4.791	-4.846
	BVAR-eigSV	**27.562**	28.512	28.522	28.523	-4.629	-4.731*	-4.791	-4.824*
	BVAR-OISV	28.113	28.757	28.577	28.567	-4.610	-4.743	-4.786	-4.822*
PCECTPI	BVAR-CSV	1.561	1.779	1.779	1.848	-1.788	-2.169	-2.296	-2.704
	BVAR-cholSV	1.496	1.721	1.693	1.729	-1.697*	-2.080	-2.181	-2.406
	BVAR-cholSV-RO	1.575	1.775	1.774	1.770	-1.718	-2.086	-2.193	-2.423
	BVAR-eigSV	**1.470***	**1.683***	1.675	1.703*	**-1.658****	-1.991	-2.137	-2.326
	BVAR-OISV	1.509	1.690*	**1.669**	**1.678***	-1.670*	**-1.986***	**-2.075***	**-2.266**
CPIAUCSL	BVAR-CSV	2.043	2.280	2.251	2.301	-2.069	-2.406	-2.534	-2.899
	BVAR-cholSV	1.959	2.169	2.117	2.124*	-1.944*	-2.247*	-2.304*	-2.484
	BVAR-cholSV-RO	2.029	2.236	2.219	2.193	-1.929	-2.234	-2.319	-2.533
	BVAR-eigSV	**1.952****	2.156*	2.123	2.119*	-1.924*	**-2.170**	-2.290**	-2.435
	BVAR-OISV	1.998	**2.153***	**2.114**	**2.091***	**-1.923***	-2.178*	**-2.242***	**-2.395**
OPHNFB	BVAR-CSV	2.630	2.624	2.662	2.654	-2.437	-2.534	-2.567	-2.613
	BVAR-cholSV	2.587	**2.602**	**2.623**	**2.636**	**-2.371**	**-2.384****	**-2.401****	**-2.417*****
	BVAR-cholSV-RO	2.601	2.627	2.645	2.671	-2.404	-2.410*	-2.432*	-2.454**
	BVAR-eigSV	2.603	2.614	2.633	2.644	-2.385	-2.400**	-2.412**	-2.425***
	BVAR-OISV	**2.583**	2.627	2.646	2.650	-2.393	-2.407**	-2.423**	-2.431***
FEDFUNDS	BVAR-CSV	0.439	0.816	1.109	1.406	-0.580	-1.618	**-2.503**	**-3.332**
	BVAR-cholSV	**0.320**	**0.628**	**0.938**	**1.249**	**0.000**	-1.564	-3.085**	-4.406**
	BVAR-cholSV-RO	0.386	0.744	1.083	1.411	-0.509	-1.608	-2.690	-3.684
	BVAR-eigSV	0.349	0.683	1.009	1.324	-0.370	-1.461	-2.610	-3.720
	BVAR-OISV	0.336	0.653	0.966	1.277	-0.078	**-1.433**	-2.773*	-3.942**
TB3MS	BVAR-CSV	0.411	0.746	1.011	1.282	-0.509	-1.608	**-2.507**	**-3.339**
	BVAR-cholSV	**0.339**	**0.636**	**0.916**	**1.192**	**-0.092****	**-1.564***	-2.964**	-4.215**
	BVAR-cholSV-RO	0.420	0.756	1.054	1.342	-0.585	-1.653	-2.661	-3.623
	BVAR-eigSV	0.371	0.690	0.982	1.258	-0.326	-1.578	-2.862*	-4.077*
	BVAR-OISV	0.359	0.664	0.948	1.226	-0.155*	-1.589	-2.925**	-4.123**
GS1	BVAR-CSV	0.483	0.845	1.123	1.407	-0.682	-1.798	**-2.649**	**-3.428**
	BVAR-cholSV	**0.413**	**0.728**	**0.996**	**1.259**	**-0.345**	-1.647	-2.800	-3.836*
	BVAR-cholSV-RO	0.471	0.823	1.114	1.396	-0.727	-1.750	-2.657	-3.527
	BVAR-eigSV	0.427	0.750	1.027	1.295	-0.477	-1.715	-2.917	-4.056*
	BVAR-OISV	0.426	0.744	1.016	1.284	-0.467	**-1.640**	-2.673	-3.612
GS10	BVAR-CSV	0.414	0.686	0.858	1.016	-0.583	-1.651	-2.318	-2.891
	BVAR-cholSV	0.364	0.580	0.705	**0.821**	-0.420	-1.602	-2.384	-3.016
	BVAR-cholSV-RO	0.380	0.612	0.750	0.874	-0.522	-1.612	-2.409	-3.141
	BVAR-eigSV	**0.364**	**0.578**	**0.704**	0.821	**-0.412***	-1.553	-2.393	-3.128
	BVAR-OISV	0.368	0.582	0.710	0.829	-0.461	**-1.476**	**-2.192**	**-2.810**
BAA10YM	BVAR-CSV	0.368	0.575	0.697	0.781	-0.269	-1.529	-2.389	-3.233
	BVAR-cholSV	**0.338**	**0.524**	**0.636**	0.717	-0.163	-1.434	-2.360	-3.200
	BVAR-cholSV-RO	0.364	0.567	0.702	0.803	-0.200	-1.428	-2.317	-3.257
	BVAR-eigSV	0.340	0.526	0.636	**0.713**	**-0.122**	**-1.201***	**-2.041**	**-2.724**
	BVAR-OISV	0.342	0.532	0.648	0.735	-0.137	-1.281	-2.142	-2.903

[1] Notes: The bold figure indicates the best model in each case. $*$, $**$ and $***$ denote, respectively, 0.10, 0.05 and 0.01 significance level for a two-sided Diebold and Mariano(1995) test.The benchmark model in the Diebold Mariano test is BVAR-CSV with 20 variables.