Macroeconomic Forecasting and Variable Ordering in Multivariate Stochastic Volatility Models

Jonas E. Arias* Juan F. Rubio-Ramírez† Minchul Shin‡

May 25, 2021

Abstract

We document five novel empirical findings on the well-known potential ordering drawback associated with the time-varying parameter vector autoregression with stochastic volatility developed by Cogley and Sargent (2005) and Primiceri (2005), CSP-SV. First, the ordering does not affect point prediction. Second, the standard deviation of the predictive densities implied by different orderings can differ substantially. Third, the average length of the prediction intervals is also sensitive to the ordering. Fourth, the best ordering for one variable in terms of log-predictive scores does not necessarily imply the best ordering for another variable under the same metric. Fifth, the best ordering for variable x in terms of log-predictive scores tends to put the variable x first while the worst ordering for variable x tends to put the variable x last. Then, we consider two alternative ordering invariant time-varying parameter VAR-SV models: the discounted Wishart SV model (DW-SV) and the dynamic stochastic correlation SV model (DSC-SV). The DW-SV underperforms relative to each ordering of the CSP-SV. The DSC-SV has an out-of-sample forecasting performance comparable to the median outcomes across orderings of the CSP-SV.

JEL classification: C8; C11; C32; C53

Keywords: Vector Autoregressions; Time-Varying Parameters; Stochastic Volatility; Variable Ordering; Cholesky Decomposition; Wishart Process; Dynamic Conditional Correlation; Out-of-sample Forecasting Evaluation

We are grateful to Manabu Asai, Joshua Chan, Todd Clark, Daniel Lewis, Christian Matthes, Mikkel Plagborg-Møller, Frank Schorfheide, Keith Sill, Jonathan Wright, and participants at the Philadelphia Fed brown-bag seminar for helpful comments. The views expressed in this paper are solely those of uthors and do not necessarily reflect the views of the Federal Reserve Bank of Atlanta, the Federal Reserve Bank of Philadelphia, or the Federal Reserve System. Any errors or omissions are the responsibility of the authors. No statements here should be treated as legal advice.

*Federal Reserve Bank of Philadelphia. Email: jonas.arias@phil.frb.org

†Emory University and Federal Reserve Bank of Atlanta. Email: juan.rubio-ramirez@emory.edu

‡Federal Reserve Bank of Philadelphia. Email: minchul.shin@phil.frb.org

1 Introduction

Several studies have shown the benefits of using the time-varying parameter vector autoregression with stochastic volatility developed by Cogley and Sargent (2005) and Primiceri (2005)—henceforth, CSP-SV—in forecasting exercises as well as for obtaining stylized facts of the U.S. economy.[1] To date, this model has become a workhorse framework for reduced-form and structural analysis. Furthermore, its popularity is likely to increase due to the existence of Bayesian methods for inference implemented and tested in widespread computer languages such as MATLAB, e.g., Del Negro and Primiceri (2015).

While the CSP-SV has reached a canonical status, it is well-known that it is not order invariant: the order of the variables affects the posterior distribution of the model parameters.[2] Yet, one important practical question remains unexplored. Is the ordering issue really a problem for point, density, and interval prediction in macroeconomics? Somewhat surprisingly, such a question has not been addressed in the macroeconomic forecasting literature where researchers generally estimate the CSP-SV or variant thereof using only one or a negligible subset of all possible orderings available.

This paper aims to fill this gap by assessing the pseudo out-of-sample forecasting performance of a four-variable CSP-SV under all of its orderings, and by contrasting it with two ordering invariant approaches for modeling stochastic volatility. The former will make clear that there are important differences across orderings that one can exploit to improve forecasts. The latter is crucial to highlight that not all ordering invariant models can fit the data as well as the CSP-SV.

We conduct our evaluation using U.S. data for four core macroeconomic variables: output growth, inflation, the 3-Month T-Bill rate, and the unemployment rate. We document five novel findings on the well-known potential ordering drawback intrinsic to the CSP-SV. First, the ordering does not affect point prediction. Second, the standard deviation of the predictive densities implied by different orderings can differ substantially. Third, the average length of the prediction intervals is also sensitive to the ordering. Fourth, the best ordering for one variable in terms of log-predictive scores does not necessarily imply the best ordering for another variable under the same metric. Fifth, the best ordering for variable x tends to put the variable x first while the worst ordering for variable x tends to put the variable x last.

Our results imply that the ordering of the variables in the CSP-SV should be justified *even* in reduced-form analysis such as macroeconomic forecasting. This may become computationally intractable as the number of possible orderings in a k-variable CSP-SV is $k!$. For example, Carriero, Clark and Marcellino (2019) compute and compare predictive densities based on 1,000

[1] For example, Clark (2011), D'Agostino, Gambetti and Giannone (2013), Baumeister and Peersman (2013), and Galí and Gambetti (2015).

[2] See e.g., Cogley and Sargent (2005), Primiceri (2005), Carriero, Clark and Marcellino (2019), Bognanni (2018), Hartwig (2020), and Chan et al. (2020).

randomly selected different variable orderings for a *single time period* (see Section C of their Supplementary Appendix). Even so, their large VAR includes 20 variables, and hence 1,000 orderings reflect only about $4 \times 10^{-14}\%$ of all possible orderings (i.e., $20! = 2.43 \times 10^{18}$).

Given the ordering dependence and the computational cost of checking all possible orderings, one may wonder if ordering invariant models can forecast as well as some of the orderings in the CSP-SV. We consider two classes of such models. The first is the ordering invariant dynamic linear model with discounted Wishart stochastic volatility model (DW-SV) developed by West and Harrison (1997), Uhlig (1997), Prado and West (2010), and Bognanni (2018). The second is an approach based on the decomposition of the time-varying reduced-form covariance matrix introduced by Engle (2002). To place an ordering invariant prior on the time-varying covariance matrix of a time-varying parameters vector autoregression model (TVP-VAR), we follow Asai and McAleer (2009) and impose a Wishart process on the correlation dynamics. This results in a TVP-VAR with dynamic and stochastic correlation-based multivariate stochastic volatility model, which we label DSC-SV. We sample from this model using the elliptical sampling approach developed by Murray, Adams and Mackay (2010). The incorporation of *theoretically* ordering invariant correlation-based multivariate stochastic volatility into a TVP-VAR complements the work of Hartwig (2020) who proposes an *almost empirically* ordering invariant methodology. The application of Murray, Adams and Mackay's (2010) approach to models with stochastic correlation-based multivariate stochastic volatility is new to the literature.

We find that the DW-SV underperforms in terms of point, density, and interval prediction relative to the other models under analysis. In all but one case, the root mean square error (RMSE) of the DW-SV is higher than all the RMSEs associated with all the possible ordering of the CSP-SV. In terms of joint density prediction, the sum of one-quarter-ahead log predictive score of the DW-SV is about 70 log units lower than that of the median CSP-SV and the DSC-SV, respectively. This large difference is also a feature of marginal log predictive scores for each variable. The empirical coverage rates based on the DW-SV model are much higher than those of other models for all variables and all horizons. Similar results are obtained when looking at the four- and eight-quarter-ahead forecast horizon. In contrast, in our application, the DSC-SV has a predictive performance comparable to the CSP-SV in terms of point, density, and interval prediction.

The rest of the paper is organized as follows. Section 2 briefly describes the CSP-SV. Section 3 gauges the role played by the ordering of the variables in the out-of-sample properties of the CSP-SV. Section 4 describes the DW-SV and the DSC-SV as well as their out of sample predictive performance. Section 5 concludes.

2 The CSP-SV

In this section we present the CSP-SV model and the priors. We also illustrate analytically how the ordering issue inherent to this model can affect its predictive density.

2.1 Model and Bayesian Inference

The most popular representation of the CSP-SV takes the form

$$\boldsymbol{y}_t' = vec(\boldsymbol{B}_t)'\boldsymbol{X}_t + \boldsymbol{\varepsilon}_t'\boldsymbol{\Sigma}_t\boldsymbol{A}_t'^{-1}, \quad \boldsymbol{\varepsilon}_t \sim \mathrm{N}(\boldsymbol{0}_{n\times 1}, \boldsymbol{I}_n), \quad \text{for } t = 1, \ldots, T, \tag{1}$$

where $\boldsymbol{y}_t$ is an $n \times 1$ vector, $\boldsymbol{X}_t' = \boldsymbol{I}_n \otimes \left[1, \boldsymbol{y}_{t-1}', \ldots, \boldsymbol{y}_{t-p}'\right]$ is an $n \times nm$ matrix with $m = np + 1$, $\boldsymbol{B}_t$ is an $m \times n$ matrix, $\boldsymbol{A}_t$ is an $n \times n$ lower triangular matrix with ones along the diagonal and $\boldsymbol{\Sigma}_t$ is a diagonal matrix. The matrices $\boldsymbol{A}_t$ and $\boldsymbol{\Sigma}_t$ are parameterized as

$$\boldsymbol{A}_t = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ \boldsymbol{\alpha}_{21,t} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \boldsymbol{\alpha}_{n1,t} & \cdots & \boldsymbol{\alpha}_{nn-1,t} & 1 \end{bmatrix}, \quad \boldsymbol{\Sigma}_t = \begin{bmatrix} \boldsymbol{\sigma}_{1,t} & 0 & \ldots & 0 \\ 0 & \boldsymbol{\sigma}_{2,t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \boldsymbol{\sigma}_{n,t} \end{bmatrix},$$

where $\boldsymbol{\alpha}_t = (\boldsymbol{\alpha}_{21,t}, \boldsymbol{\alpha}_{n1,t}, \ldots, \boldsymbol{\alpha}_{nn-1,t})'$ and $\boldsymbol{\sigma}_t = (\boldsymbol{\sigma}_{1,t}, ..., \boldsymbol{\sigma}_{n,t})'$ are the time-varying parameters governing the unrestricted entries of these matrices. The time-varying parameters of $\boldsymbol{B}_t$, $\boldsymbol{A}_t$, and $\boldsymbol{\Sigma}_t$ evolve according to random walks

$$\begin{aligned} vec(\boldsymbol{B}_t) &= vec(\boldsymbol{B}_{t-1}) + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim \mathrm{N}(\boldsymbol{0}_{mn\times 1}, \boldsymbol{Q}), & (2) \\ \boldsymbol{\alpha}_t &= \boldsymbol{\alpha}_{t-1} + \boldsymbol{\zeta}_t, \quad \boldsymbol{\zeta}_t \sim \mathrm{N}(\boldsymbol{0}_{n(n-1)/2\times 1}, \boldsymbol{S}), & (3) \\ \log \boldsymbol{\sigma}_t &= \log \boldsymbol{\sigma}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathrm{N}(\boldsymbol{0}_n, \boldsymbol{W}), & (4) \end{aligned}$$

where $\boldsymbol{Q}$ and $\boldsymbol{W}$ are unrestricted positive definite matrices, $\boldsymbol{S}$ is a block diagonal positive definite matrix with each block corresponding to the variance matrix of each j-th row of $\boldsymbol{A}_t$ for $j = 2, \ldots, n$, and $\log \boldsymbol{\sigma}_t = (\log \boldsymbol{\sigma}_{1,t}, ..., \log \boldsymbol{\sigma}_{n,t})'$.

In this paper, we use the same priors and simulation method implemented in the companion MATLAB code of Del Negro and Primiceri (2015). Thus, the initial states $\boldsymbol{B}_0$, $\boldsymbol{\alpha}_0$, and $\boldsymbol{\sigma}_0$, and the hyperparameters $\boldsymbol{Q}$, $\boldsymbol{S}$, and $\boldsymbol{W}$ are assumed to be independent of each other. The former are distributed according to the normal distribution and the latter are distributed according to the inverse-Wishart distribution.

Prior for $\boldsymbol{B}_0$. More specifically, $\text{vec}(\boldsymbol{B}_0) \sim \text{N}\left(vec(\hat{\boldsymbol{B}}), 4 \cdot V(vec(\hat{\boldsymbol{B}}))\right)$ where $vec(\hat{\boldsymbol{B}})$ and $V(vec(\hat{\boldsymbol{B}}))$ are the mean and variance OLS point estimates based on a time-invariant VAR estimated with a pre-sample of $T_0 = 40$ observations. That is, consider the VAR: $\boldsymbol{y}'_\ell = \boldsymbol{x}'_\ell \boldsymbol{B} + \boldsymbol{e}'_t$ for $\ell \in [-T_0+1, 0]$ with $\boldsymbol{e}'_t = \boldsymbol{\epsilon}'_t(\boldsymbol{A}^{-1})'$, $\boldsymbol{\epsilon}_t \sim \text{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$, $\boldsymbol{x}'_\ell = [1, \boldsymbol{y}_{\ell-1}, \ldots, \boldsymbol{y}_{\ell-p}]$, and note that $\hat{\boldsymbol{B}} = (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{X}'\boldsymbol{Y}$ and $V(vec(\hat{\boldsymbol{B}})) = \frac{\hat{\boldsymbol{e}}'\hat{\boldsymbol{e}}}{T_0} \otimes (\boldsymbol{X}'\boldsymbol{X})^{-1}$, where $\hat{\boldsymbol{e}} = \boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{B}}$, $\boldsymbol{Y}' = (\boldsymbol{y}_{-T_0+1}, \ldots, \boldsymbol{y}_0)$, $\boldsymbol{X}' = (\boldsymbol{x}_{-T_0+1}, \ldots, \boldsymbol{x}_0)$.

Prior for $\boldsymbol{\alpha}_0$. The matrix $\boldsymbol{\alpha}_0 \sim \text{N}(\hat{\boldsymbol{a}}, 4 \cdot V(\hat{\boldsymbol{a}}))$ where $\hat{\boldsymbol{a}}$ and $V(vec(\hat{\boldsymbol{a}}))$ are obtained using a pre-sample of T_0 observations. In particular, let vech_d be an operator that extracts the elements below the main diagonal of a matrix. Since $\boldsymbol{A}_0\boldsymbol{e}_t = \epsilon_t$, it follows that an estimate of $\text{vech}_d(\boldsymbol{A})$ can be obtained by projecting $vec(\hat{\boldsymbol{e}})$ onto $\boldsymbol{I}_n \otimes \hat{\boldsymbol{e}}$, where $\hat{\boldsymbol{e}}' = (\hat{\boldsymbol{e}}_{-T_0+1}, \ldots, \hat{\boldsymbol{e}}_0)$. Then, $\hat{\boldsymbol{a}}$ is set equal to the resulting estimate. The variance of $\boldsymbol{\alpha}_0$ is defined by setting $\hat{\boldsymbol{V}} = (\hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}/T_0) \otimes \boldsymbol{I}_{T_0}$, where $\hat{\boldsymbol{u}} = vec^{-1}_{T_0,n}\left(vec(\hat{\boldsymbol{e}}) - \hat{\boldsymbol{Z}}\hat{\boldsymbol{a}}\right)$, $\hat{\boldsymbol{Z}} = (\boldsymbol{I}_n \otimes \hat{\boldsymbol{e}})$, $\breve{V}(\hat{\boldsymbol{a}}) = (\hat{\boldsymbol{Z}}'\hat{\boldsymbol{Z}})^{-1}\hat{\boldsymbol{Z}}'\hat{\boldsymbol{V}}\hat{\boldsymbol{Z}}(\hat{\boldsymbol{Z}}'\hat{\boldsymbol{Z}})^{-1}$, and $V(\hat{\boldsymbol{a}})$ is a $\frac{n(n-1)}{2} \times \frac{n(n-1)}{2}$ matrix such that given $i = 1$, for j $= 1, n-1$, we have $V(\hat{\boldsymbol{a}}_j) = V(\hat{\boldsymbol{a}})_{i:i+j-1,i:i+j-1} = \breve{V}(\hat{\boldsymbol{a}})_{i:i+j-1,i:i+j-1}$ with i=i+j, and 0 otherwise.

Prior for $\log(\boldsymbol{\sigma}_0)$. The vector $\log(\boldsymbol{\sigma}_0) \sim \text{N}(\log(\hat{\boldsymbol{\sigma}}_0), \boldsymbol{I}_n)$, with $\hat{\boldsymbol{\sigma}}_0 = \text{diag}(vecd(\hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}/T_0)))^{-0.5}$, where $vecd(\boldsymbol{X})$ creates a vector from the diagonal elements of a matrix $\boldsymbol{M}$, and $diag(\boldsymbol{x})$ builds a diagonal matrix whose diagonal elements are given by $\boldsymbol{x}$.

Prior for the Hyperparameters. Turning to the prior for the hyperparameters, $\boldsymbol{Q} \sim \text{IW}(k^2_{\boldsymbol{Q}} \cdot 40 \cdot V(vec(\hat{\boldsymbol{B}})), 40)$, where $k_{\boldsymbol{Q}} = 0.01$, and $\boldsymbol{W} \sim \text{IW}(k^2_{\boldsymbol{W}} \cdot 4 \cdot \boldsymbol{I}_n, 4)$, where $\boldsymbol{k_W} = 0.01$.[3] As mentioned $\boldsymbol{S}$ is a block diagonal matrix partitioned with $n-1$ blocks where the j-th block is $\boldsymbol{S}_j \sim \text{IW}(k^2_{\boldsymbol{S}} \cdot (j+1) \cdot V(\hat{\boldsymbol{a}}_j), j+1)$ $j \in \{1, \ldots, n-1\}$ and $k_{\boldsymbol{S}} = 0.1$.

Equipped with this prior, Algorithm 2 in Del Negro and Primiceri (2015) simulates the posterior distribution of the history of volatitities $(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_T)$, the histories of coefficients $(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_T)$ and $(\boldsymbol{B}_1, \ldots, \boldsymbol{B}_T)$, and the parameters $\boldsymbol{Q}$, $\boldsymbol{S}$, and $\boldsymbol{W}$. Since the derivation of Algorithm 2 and its implementation is carefully documented in Del Negro and Primiceri (2015) and its companion code we refer the reader to their paper and code for additional details.

2.2 The Ordering Issue

As highlighted by Primiceri (2005), in this model the ordering of the variables affects the posterior distribution of the parameters. In this section, we will first replicate his two-variable

[3] The inverse-Wishart density is parameterized as follows: $\text{IW}_{(\nu,\boldsymbol{\Psi})}(\boldsymbol{\Sigma}) = c_{\text{IW}} \mid \boldsymbol{\Sigma} \mid^{-(\nu+n+1)/2} e^{-0.5\text{tr}(\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1})}$, where $c^{-1}_{\text{IW}} = \mid \boldsymbol{\Psi} \mid^{-\nu/2} 2^{\nu n}\Gamma_n(\nu/2)$.

example and then we will make some further assumptions to illustrate analytically how the ordering affects the predictive density implied by the model.

Let $\boldsymbol{\Omega}_t$ denote the reduced-form covariance matrix $\boldsymbol{A}_t^{-1}\boldsymbol{\Sigma}_t\boldsymbol{\Sigma}_t'\boldsymbol{A}_t^{-1\prime}$ and notice that

$$\boldsymbol{\Omega}_t = \begin{bmatrix} \left(e^{(\log \boldsymbol{\sigma}_{1,t-1}+\boldsymbol{\eta}_{1,t})}\right)^2 & -\boldsymbol{\alpha}_{21,t}\left(e^{(\log \boldsymbol{\sigma}_{1,t-1}+\boldsymbol{\eta}_{1,t})}\right)^2 \\ -\boldsymbol{\alpha}_{21,t}\left(e^{(\log \boldsymbol{\sigma}_{1,t-1}+\boldsymbol{\eta}_{1,t})}\right)^2 & (\boldsymbol{\alpha}_{21,t-1}+\boldsymbol{\zeta}_t)^2\left(e^{(\log \boldsymbol{\sigma}_{1,t-1}+\boldsymbol{\eta}_{1,t})}\right)^2 + \left(e^{(\log \boldsymbol{\sigma}_{2,t-1}+\boldsymbol{\eta}_{2,t})}\right)^2 \end{bmatrix}.$$

The expression above makes clear that, given the states in period $t-1$, the distribution of the first element of the diagonal of $\boldsymbol{\Omega}_t$ is proportional to a log-normal distribution. In contrast, the distribution of the second element of the diagonal of $\boldsymbol{\Omega}_t$ is not proportional to a log-normal distribution. Hence, inference under different orderings will imply different distributions for the entries of the reduced-form covariance matrix which could affect the model's predictive performance.

To see the latter analytically, assume that in our two-variable example $\boldsymbol{\sigma}_{1,t} = \boldsymbol{\sigma}_{2,t} = 1$, $\boldsymbol{\alpha}_{21,t} \overset{i.i.d}{\sim} \mathrm{N}(0,1)$, and that there are neither lags nor constant terms. Then, the predictive density of $\boldsymbol{y}_{1,t}$ is Gaussian while the predictive density of $\boldsymbol{y}_{2,t}$ is non-Gaussian.[4] In fact, the latter has a fatter tail than the former (e.g, Haldane, 1942). This makes clear that the ordering of the variables affects the predictive performance of the model. The crux of the matter is that we are placing a prior on the variance of the one-quarter forecast errors, $\boldsymbol{\Omega}_t$, *after* decomposing it via a Cholesky-decomposition so that $\boldsymbol{\Omega}_t = \boldsymbol{A}_t^{-1}\boldsymbol{\Sigma}_t\boldsymbol{\Sigma}_t\boldsymbol{A}_t^{\prime -1}$ and $\boldsymbol{A}_t^{-1}$ is the lower triangular matrix. Putting an independent prior on each element of $\boldsymbol{A}_t^{-1}$ does not lead to symmetric prior in terms of the marginal distribution of $\boldsymbol{y}_t$. Instead, imposing an inverse-Wishart prior on $\boldsymbol{\Omega}_t$ as in Section 4.1 and using an alternative decomposition of $\boldsymbol{\Omega}_t$ as in Section 4.2 are order invariant procedures.

In the next section, we will assess whether the actual predictive performance is an empirical issue in a standard setting.

3 Out-of-Sample Prediction for the CSP-SV Model

In this section we analyze the out-of-sample prediction of the CSP-SV model. We define the setup and then analyze point, density, and prediction intervals.

3.1 Setup

We estimate a four-variable quarterly frequency CSP-SV using U.S. data.[5] The four variables included in the model are output growth (real GDP growth), inflation (based on the Core PCE

[4]Under our simplifying assumptions $\boldsymbol{y}_{1,t} = \boldsymbol{\varepsilon}_{1,t} \sim \mathrm{N}(0,1)$ and $\boldsymbol{y}_{2,t} = -\boldsymbol{\alpha}_{21,t}\boldsymbol{\varepsilon}_{1,t} + \boldsymbol{\varepsilon}_{2,t} \sim \mathrm{N}(0,1)^2 + \mathrm{N}(0,1)$.

[5]The data was obtained from the FRED-QD Quarterly Database for Macroeconomic Research.

Price Index), the 3-Month T-bill rate, and the unemployment rate for the period 1970Q1:2016Q4. Output growth and inflation are computed using annualized % log-differences, and the 3-Month T-bill rate and the unemployment rate are expressed in %. We use data for the period 1960Q1:1969Q4 to construct our prior distribution. The model is estimated including two lags.

In a four-variable CSP-SV there are 24 different orderings. For each ordering, we recursively estimate and generate one-, four-, and eight-quarter-ahead predictions during 120 quarters starting in 1987Q1, i.e., when generating our first forecast we assume that we have data up to 1987Q1. Thus, our evaluation sample runs from 1987Q1 to 2016Q4. We index the quarters in which forecasts are made by $\tau \in \{1, \ldots, 120\}$ and we index the forecast horizon by $h \in \{1, 4, 8\}$. Accordingly, our first forecast is for 1987Q2 (when $\tau = 1$ and $h = 1$) and our latest forecast is for 2018Q4 (when $\tau = 120$ and $h = 8$). We evaluate the predictive performance under the 24 orderings through the lens of the RMSE for point prediction, the log predictive score for density prediction, and the empirical coverage and average length for interval prediction.

Notice that our exercise is non-trivial. We compute 23,040 predictive densities ($24 \times 120 \times 8$) based on 2,880 posterior distributions (24×120) of all the possible orderings of the four-variable CSP-SV. Each predictive distribution and posterior distribution is constructed based on 50,000 Markov chain Monte Carlo (MCMC) draws.

3.2 Point Prediction

Panel (a) in Table 1 shows the range of RMSEs and the median RMSE across the 24 orderings at one-, four-, and eight-quarter-ahead, where the point estimates are computed using the posterior mean of the predictive density. The RMSEs are computed over the evaluation sample.[6] The gist of these point prediction outcomes is that, although there are differences in performance, from the perspective of macroeconomic forecasting the differences in RMSE are not affected by the ordering in an economically meaningful manner.

Panel (b) in Table 1 reports the results from Diebold-Mariano (Diebold and Mariano, 1995) tests for equal predictive ability over the evaluation sample. With four variables, for each variable and horizon we have $\binom{24}{2} = 276$ possible orderings to compare. Consequently, we focus on testing the two orderings with the largest MSE difference. As can be seen, the null hypothesis of equal predictive ability is rejected at a 5% significance level in only 3 out of 12 cases: the p-value for four- and eight-quarter-ahead inflation forecasts and the p-value for the four-quarter-ahead 3-Month T-Bill forecast are below 0.05. And, even when rejected, these differences are very small. For example, largest MSE difference for four-quarter-ahead forecasts of inflation across all orderings is 0.03 percentage point.

While Table 1 shows that for all variables and horizons all the orderings perform similarly,

[6] Appendix A.3 describes the RMSE for all variables, orderings, and horizons.

Table 1: RMSE

(a) RMSE

	One-Quarter-Ahead		Four-Quarter-Ahead		Eight-Quarter-Ahead	
	Range	Median	Range	Median	Range	Median
Output Growth	[2.46,2.57]	2.51	[2.67,2.73]	2.70	[2.56,2.61]	2.59
Inflation	[0.60,0.60]	0.60	[0.77,0.80]	0.78	[0.86,0.92]	0.88
3-Month T-Bill	[0.33,0.34]	0.33	[1.14,1.18]	1.16	[1.92,2.01]	1.95
Unemployment	[0.20,0.20]	0.20	[0.80,0.83]	0.82	[1.42,1.47]	1.44

(b) Diebold-Mariano Equal Predictive Ability Test

	One-Quarter-Ahead	Four-Quarter-Ahead	Eight-Quarter-Ahead
Output Growth	0.11 (0.10)	0.06 (0.07)	0.05 (0.17)
Inflation	0.01 (0.23)	0.03 (0.02)	0.06 (0.00)
3-Month T-Bill	0.02 (0.06)	0.05 (0.03)	0.09 (0.20)
Unemployment	0.01 (0.10)	0.03 (0.07)	0.05 (0.16)

Note. Panel (a): Range indicates the minimum and maximum RMSE. Panel (b): Numbers are the MSE difference. Numbers in parentheses are p-values of Diebold-Mariano equal predictive ability tests.

Figure 1: Relative RMSE

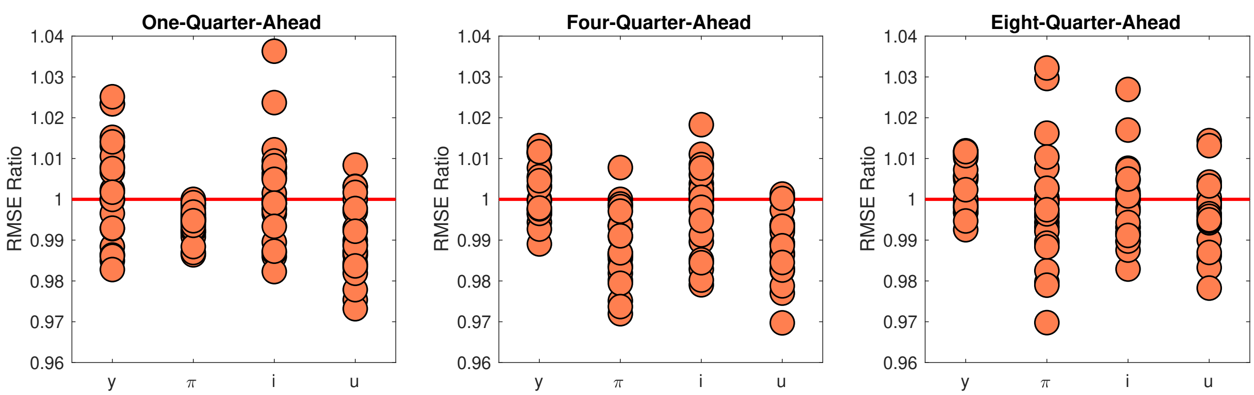


Figure 1 presents the results of the table in a different format to facilitate a comparison of the relative magnitude of the differences. In particular, the figure presents the RMSE for each model at the horizons under analysis relative to the reference ordering. There is a 4% upper bound and lower bound difference in RMSEs. To see the implication of this number, notice that the RMSE can be interpreted as the standard deviation of the forecast error. Combining this interpretation with the RMSEs associated with the first ordering (2.51 for output growth, 0.6 for inflation, 0.33 for the 3-Month T-Bill rate, and 0.2 for the unemployment rate), it follows that a 5% reduction of the RMSE is equivalent to a reduction of about 0.13 percentage point in terms of RMSE for annualized output growth, which is modest from a macroeconomic forecasting perspective. Similarly, the gains that could be obtained in terms of RMSE for inflation, the 3-Month T-Bill rate, and the unemployment rate are small: 0.03 percentage point, 0.02 percentage point, and 0.01 percentage point respectively. Thus, Figure 1 reaffirms the message that in terms of point estimates the observed differences in terms of RMSE do not translate into relevant economic discrepancies.

3.3 Density Prediction

In general, macroeconomic forecasters are interested not only in point prediction but also in density prediction. We evaluate the density prediction performance using the sum of log predictive scores (LPSs) over the evaluation sample for each of the three horizons under analysis. We consider the joint predictive density as well as the predictive density for each of the variables.

Table 2: Log Predictive Scores

(a) Sum of Log Predictive Scores

	One-Quarter-Ahead		Four-Quarter-Ahead		Eight-Quarter-Ahead	
	Range	Median	Range	Median	Range	Median
Joint	[-371.86,-346.23]	-354.31	[-739.69,-716.06]	-729.74	[-948.09,-910.97]	-931.17
Output Growth	[-279.41,-274.29]	-276.56	[-290.48,-284.50]	-287.11	[-292.80,-283.26]	-287.03
Inflation	[-111.03,-106.76]	-108.68	[-149.13,-142.56]	-144.98	[-179.46,-169.80]	-174.75
3-Month T-Bill	[-33.10,-15.74]	-21.22	[-192.47,-183.55]	-187.70	[-267.13,-260.09]	-262.18
Unemployment	[26.94,33.53]	31.84	[-138.68,-123.95]	-129.53	[-227.22,-205.30]	-212.01

(b) Amisano-Giacomini Equal Predictive Ability Test

	One-Quarter-Ahead	Four-Quarter-Ahead	Eight-Quarter-Ahead
Joint	25.63 (0.01)	23.63 (0.17)	37.12 (0.31)
Output Growth	5.12 (0.17)	5.99 (0.10)	9.54 (0.04)
Inflation	4.27 (0.03)	6.57 (0.00)	9.65 (0.00)
3-Month T-Bill	17.36 (0.06)	8.92 (0.27)	7.04 (0.23)
Unemployment	6.59 (0.28)	14.73 (0.22)	21.92 (0.34)

Note. Panel (a): Range indicates the minimum and maximum LPS. Panel (b): Numbers are the difference in the sum of LPSs. Numbers in parentheses are p-values of Amisano-Giacomini equal predictive ability test.

Panel (a) in Table 2 shows that the sum of one-quarter-ahead joint LPSs for the best ordering and for the worst ordering under this metric are -346.23 and -371.86, respectively.[7] Their difference is about 26, which implies that the LPSs differ by 0.21 every quarter, on average. When looking at four- and eight-quarter-ahead predictive densities the differences between the best and worst ordering in terms of the sum of LPSs at each respective horizon are about 35 and 40, therefore in each quarter the LPSs will differ by even more than in the case of one-quarter-ahead densities.

[7] Appendix A.3 describes the LPSs for all variables, orderings, and horizons under analysis.

Panel (b) in Table 2 shows the Amisano-Giacomini equal predictive ability test (Amisano and Giacomini, 2007) for the sum of joint LPSs and for the sum of the marginal LPSs of each variable. Similar to the case of point prediction, for each variable specification (i.e., joint LPS or marginal LPSs) and horizon we have 276 possible orderings to compare. Hence, for ease of exposition, we only test the difference between the best and worst ordering in terms of the sum of LPSs associated with each variable specification and horizon.

Let's begin by examining the Amisano-Giacomini tests for the sum of joint LPSs. The difference between the best and worst ordering is statistically significant in the case of the one-quarter-ahead densities, and statistically not different from zero in the case of the four- and eight-quarter-ahead densities. A roughly similar pattern emerges when looking at the sum of marginal LPSs for each variable: there is a heterogeneity in the scores and in some cases the Amisano-Giacomini test tells us that some differences are statistically significant. Altogether, the null hypothesis of equal predictive ability is rejected at a 5% significance level in only 5 out of 15 cases.

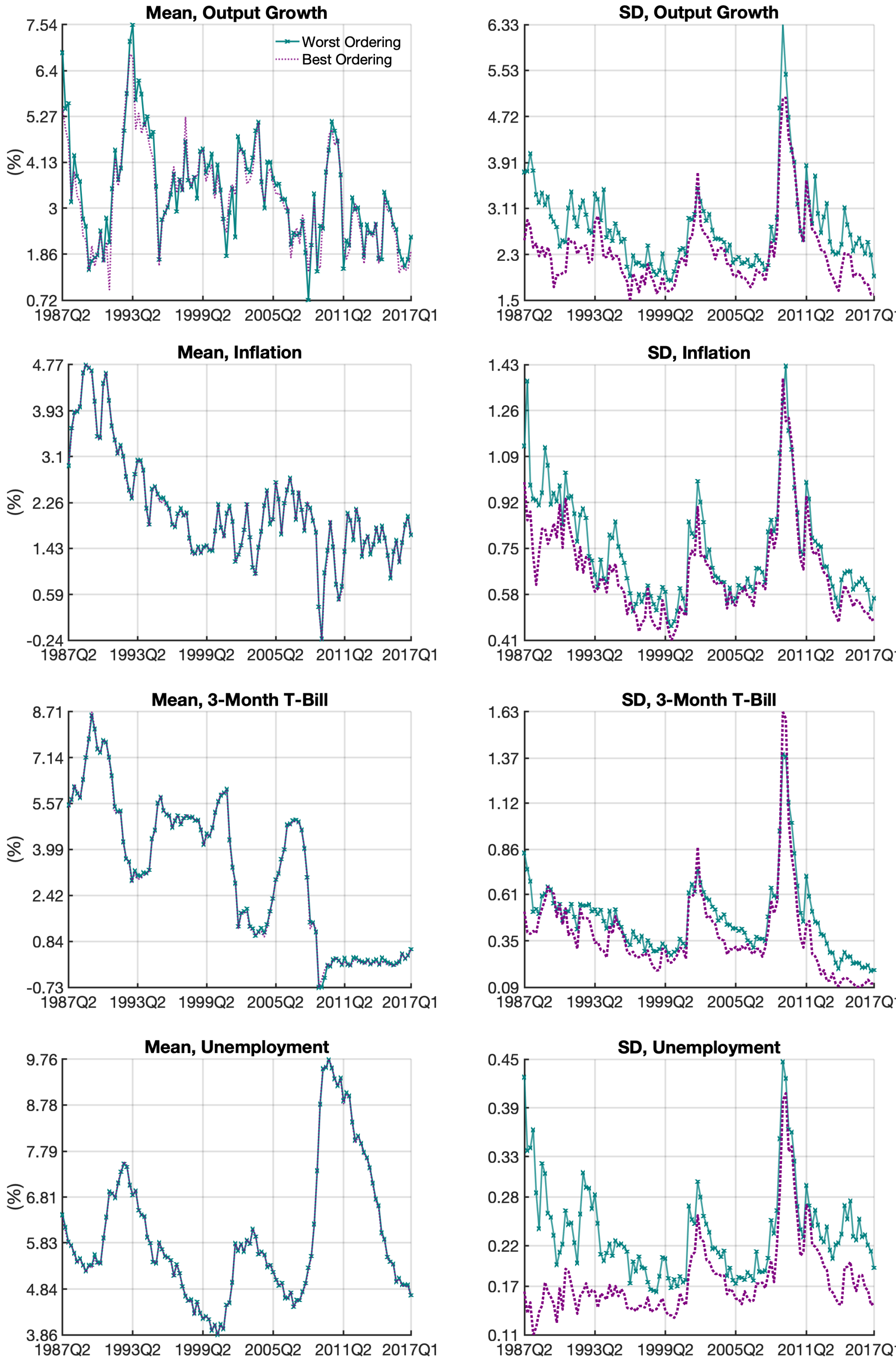
In contrast to the case of point prediction, we now show that when analyzing predictive densities the differences are important from an economic point of view. Figure 2 presents the mean and the standard deviation of the one-quarter-ahead predictive densities computed recursively over the evaluation sample for the best and worst ordering. The best and worst ordering are chosen in terms of the sum of marginal LPSs over the evaluation sample of the corresponding variable under analysis. Hence, the best and worst ordering are kept constant when producing the figure.

The second moments of the predictive densities implied by each of these orderings portray a different picture regarding the uncertainty associated with the economic outlook—an important aspect of macroeconomic forecasting as emphasized by Clark (2011). The green solid lines with markers represent the predictive densities associated with the worst orderings. The purple dotted lines represent the predictive densities associated with the best orderings.

The results are striking. The mean predictions are almost identical but the standard deviations of the predictive densities are quite different.[8] As can be seen, for each of the variables the worst ordering in terms of the sum of marginal LPSs paints a more uncertain outlook than the best ordering under the same metric. The fact that the mean predictions are almost identical is not surprising given that as shown in Section 3.2 the point forecasts are almost identical. Consequently, what it is new here is the extent to which the standard deviations can differ across orderings. Had the worst ordering been used at a policy institution such as a central bank for a span of 10 years, it would have offered policymakers a more uncertain outlook for the unemployment rate on the single basis of a seemingly arbitrary ordering choice.

[8]As shown in Appendix A.4, the same holds when looking at four- and eight-quarter-ahead predictive densities.

Figure 2: Predictive Densities and Ordering



Note. Mean and standard deviation (SD) of the one-quarter-ahead predictive density.

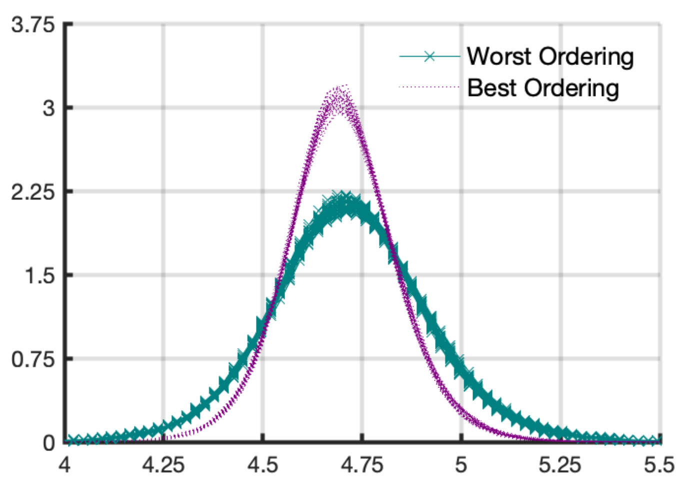
All told, the analysis above suggests that the differences in the sum of LPSs reported in Table 2 are driven by the distributional characteristics beyond the mean and they are large enough to paint a different economic outlook.

3.3.1 Robustness of the Results

The findings just described raise two questions. First, given that we construct the predictive densities relying on simulation-based methods, one could wonder if the observed differences between the standard deviations of the best and worst ordering are driven by the numerical error. Second, Figure 2 only describes first and second moments, but to what extent does the ordering affect the entire shape of the predictive density?

To answer these questions, we compute the one-quarter-ahead predictive density for the unemployment rate for the last period of the evaluation sample based on 30 independent MCMC chains, each chain consisting of 20,000 draws from the posterior distribution of the model parameters. We focus on unemployment because it is the variable for which the difference between the standard deviation of the one-quarter-ahead predictive density of the best and worst ordering in terms of the sum of one-quarter-ahead marginal LPSs over the evaluation sample is the largest. Figure 3 shows the results. The green solid lines with markers represent the predictive densities associated with the worst ordering in each MCMC chain. The purple dotted lines represent the predictive densities associated with the best ordering in each MCMC chain. As can be seen, it is unlikely that numerical error drives our results. Second, the observed difference in the uncertainty in the predictive densities leads to a noticeable difference in tail probabilities.

Figure 3: One-Quarter-Ahead Predictive Densities and Ordering



NOTE. One-quarter-ahead predictive densities based on 30 MCMC chains, each chain consisting of 20,000 posterior draws.

3.3.2 Deeper Dive into the LPS

This section highlights how the best ordering depends on the variable and the forecast horizon under analysis: the best ordering for a given variable-horizon pair does not necessarily imply the best ordering for another variable-horizon pair. Let's start with the one-quarter-ahead forecast horizon. Table 3 shows that the best ordering for predicting output growth in terms of the sum of one-quarter-ahead marginal LPSs is the worst ordering for predicting the 3-Month T-Bill rate. In addition, it shows that the best (worst) ordering for any variable tends to have the variable in question ordered first (last).

Table 3: Best and Worst Ordering

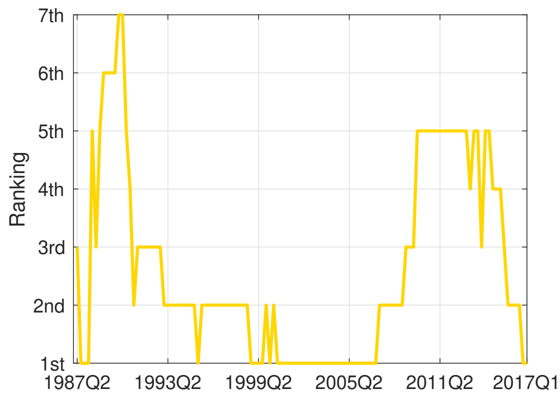
Variable	Ordering	First	Second	Third	Fourth	LPS
Output Growth	Best	y	u	π	i	-274.29
	Worst	i	π	u	y	-279.41
Inflation	Best	π	u	i	y	-106.76
	Worst	u	y	π	i	-111.03
3-Month T-Bill	Best	i	u	y	π	-15.74
	Worst	y	u	π	i	-33.10
Unemployment	Best	u	y	i	π	33.53
	Worst	π	i	y	u	26.94

To further scrutinize the punchline of Table 3, we compute the Spearman's rank correlation coefficients. Table 4 shows the Spearman's coefficients and the p-values (the null hypothesis is no-correlation) for the rank correlation between the ranking of orderings in terms of the sum of one-quarter-ahead marginal LPSs for output growth and the ranking of orderings in terms of the sum of LPSs for each of the remaining variable specifications and horizons. Hence, this table shows rankings differ not only across variables but also across horizons. For example, conditional on output growth, a ranking based in terms of the sum of one-quarter-ahead marginal LPSs has an about 0.3 correlation with a ranking based on the sum of eight-quarter-ahead marginal LPSs.

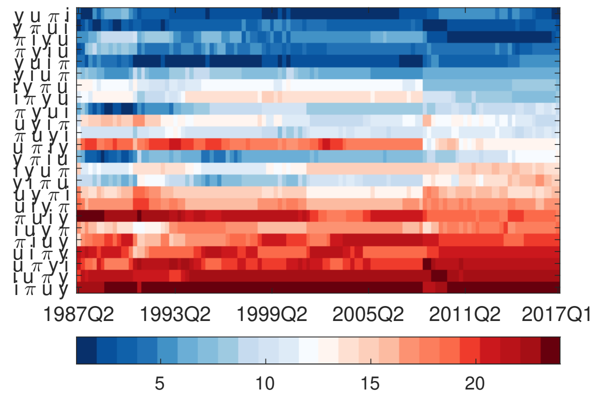
Table 4: Spearman Rank Correlation

	Correlation			p-values		
	h=1	h=4	h=8	h=1	h=4	h=8
Output Growth	1	0.47	0.29	0	0.02	0.17
Inflation	-0.06	-0.02	0.06	0.78	0.93	0.76
3-Month T-Bill	-0.27	-0.30	0.01	0.21	0.16	0.95
Unemployment	-0.38	-0.16	-0.02	0.07	0.45	0.93
Joint	0.17	-0.25	0.24	0.42	0.24	0.26

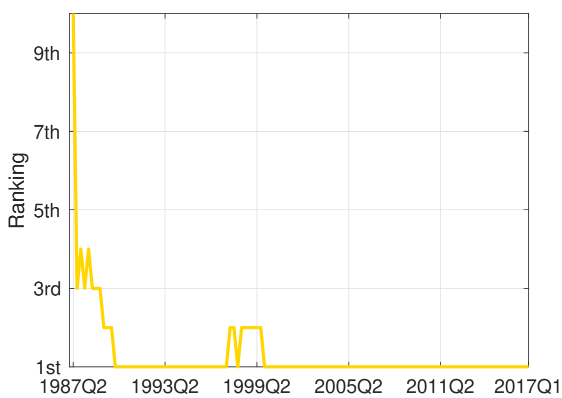
Figure 4: Time-Varying Ranking of one-quarter-ahead LPSs



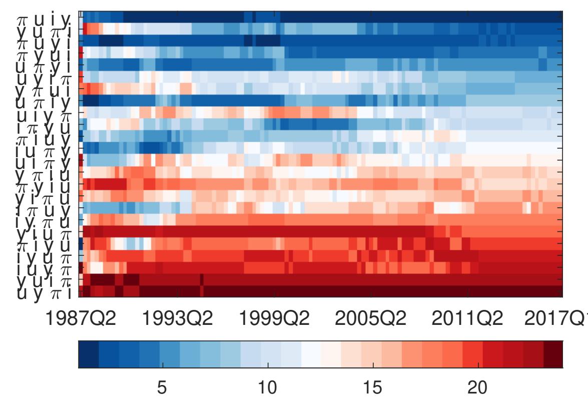
(a) Output Growth, Best Ordering



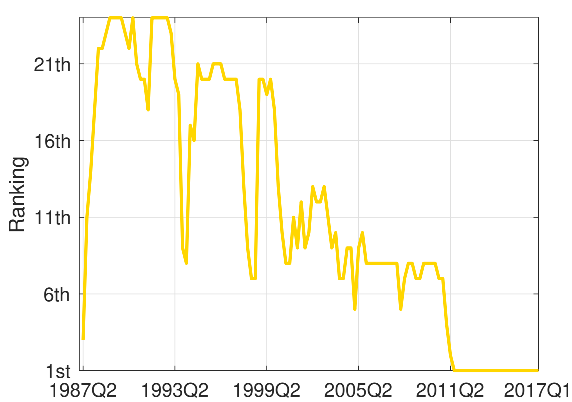
(b) Output Growth



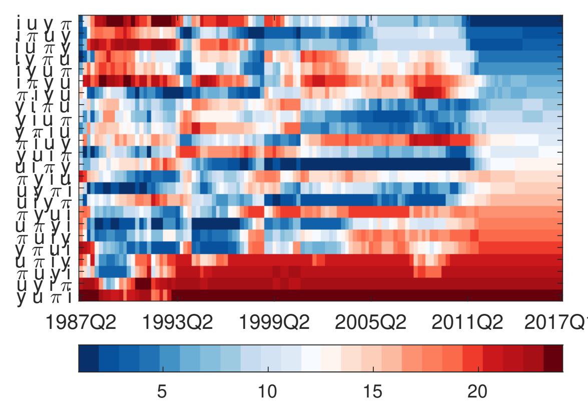
(c) Inflation, Best Ordering



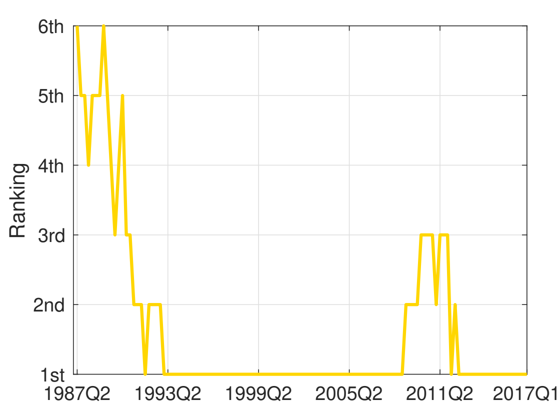
(d) Inflation



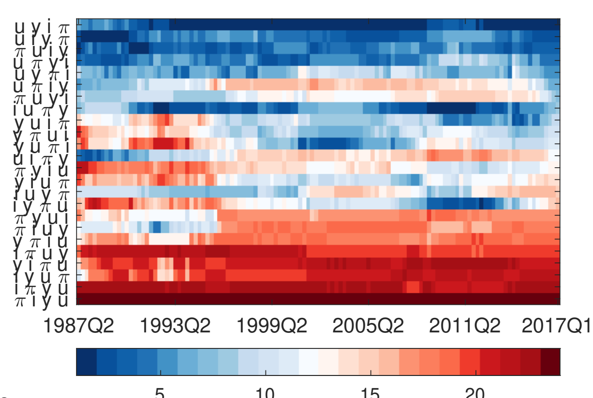
(e) 3-Month T-Bill, Best Ordering



(f) 3-Month T-Bill



(g) Unemployment, Best Ordering



(h) Unemployment

While Table 4 is informative about the correlation across variables and horizons, it is silent on how the ranking of the orderings varies over the evaluation sample. The latter is important because if the ranking changes frequently researchers would need to rank the orderings often, which is time consuming. Hence, to conclude this section, we assess the degree of serial correlation across rankings.

Figure 4 shows the degree of serial correlation in rankings computed recursively since the beginning of the evaluation sample. In particular, at each quarter of our evaluation sample we compute a ranking based on the sum of marginal LPS up to such quarter. In Panel (a) we report the evolution of the best performing ordering in terms of the sum of one-quarter-ahead marginal LPSs for output growth over the entire evaluation sample, i.e., (y, u, π, i). The panel plots how this particular ordering ranked throughout the evaluation sample. For example, in 1987Q2 the (y, u, π, i) ordering is the third best ordering while in 2011Q2 it is the fifth. Panel (b) summarizes this information for the 24 possible orderings using a colormap. The darkest blue corresponds to the best ordering and the darkest red corresponds to the worst ordering. The panel pairs (c,d), (d,f), and (g,h) do the same for the rest of the variables.

Clearly, the ranking changes throughout the evaluation sample. For example, in the case of output growth the (y, u, π, i) ordering ranks in the top 5 during the first 7 quarters, then its ranking drops to 7th before returning to the top 5 for the remainder of the evaluation sample. Similar results are obtained for the rest of the variables. Appendix A.4 shows that analogous results hold when looking at four-quarter-ahead and eight-quarter-ahead forecast horizons. The main difference is that the ranking in terms of the sum of eight-quarter-ahead LPSs for unemployment exhibits larger swings.

3.4 Interval Prediction

In addition to the point prediction and density prediction performance based on RMSEs and LPSs, macroeconomic forecasters are commonly interested in analyzing prediction intervals constructed using tail quantiles. For each variable, the $x\%$ prediction interval is an interval that covers an outcome with $x\%$ posterior probability. Based on this definition, we construct 70% symmetric probability intervals for each predictive density and we evaluate these intervals by means of their coverage rate and their average length over the evaluation sample.

Panel (a) in Table 5 presents the empirical coverage rate of 70% prediction intervals for each variable and horizon under study. Theoretically, we expect the prediction intervals to cover the realized outcome 70% of the times over our evaluation sample, nevertheless in practice there is substantial variation across orderings. In addition to the coverage rates, shorter intervals offer sharper prediction and hence it is important to assess their average length (see, for example, Askanazi et al., 2018). To see this, Panel (b) in Table 5 shows the average length of the 70%

prediction intervals. As it was the case in Panel (a), there is heterogeneity across orderings.

Table 5: Interval Prediction Evaluation

(a) Empirical coverage rate of 70% prediction intervals						
	One-Quarter-Ahead		Four-Quarter-Ahead		Eight-Quarter-Ahead	
	Range	Median	Range	Median	Range	Median
Output Growth	[0.68,0.74]	0.70	[0.71,0.79]	0.74	[0.78,0.86]	0.81
Inflation	[0.67,0.75]	0.69	[0.83,0.88]	0.86	[0.89,0.94]	0.93
3-Month T-Bill	[0.72,0.84]	0.77	[0.67,0.80]	0.72	[0.62,0.72]	0.66
Unemployment	[0.64,0.78]	0.69	[0.61, 0.78]	0.68	[0.54,0.65]	0.57

(b) Average length of 70% prediction intervals						
	One-Quarter-Ahead		Four-Quarter-Ahead		Eight-Quarter-Ahead	
	Range	Median	Range	Median	Range	Median
Output Growth	[4.62,5.51]	4.94	[5.11,6.18]	5.51	[5.55,6.82]	5.98
Inflation	[1.32,1.47]	1.35	[2.10,2.34]	2.18	[2.83,3.17]	2.99
3-Month T-Bill	[0.66,0.87]	0.71	[2.14,2.67]	2.29	[3.45,4.29]	3.68
Unemployment	[0.35,0.45]	0.37	[1.13,1.40]	1.22	[1.83,2.22]	1.93

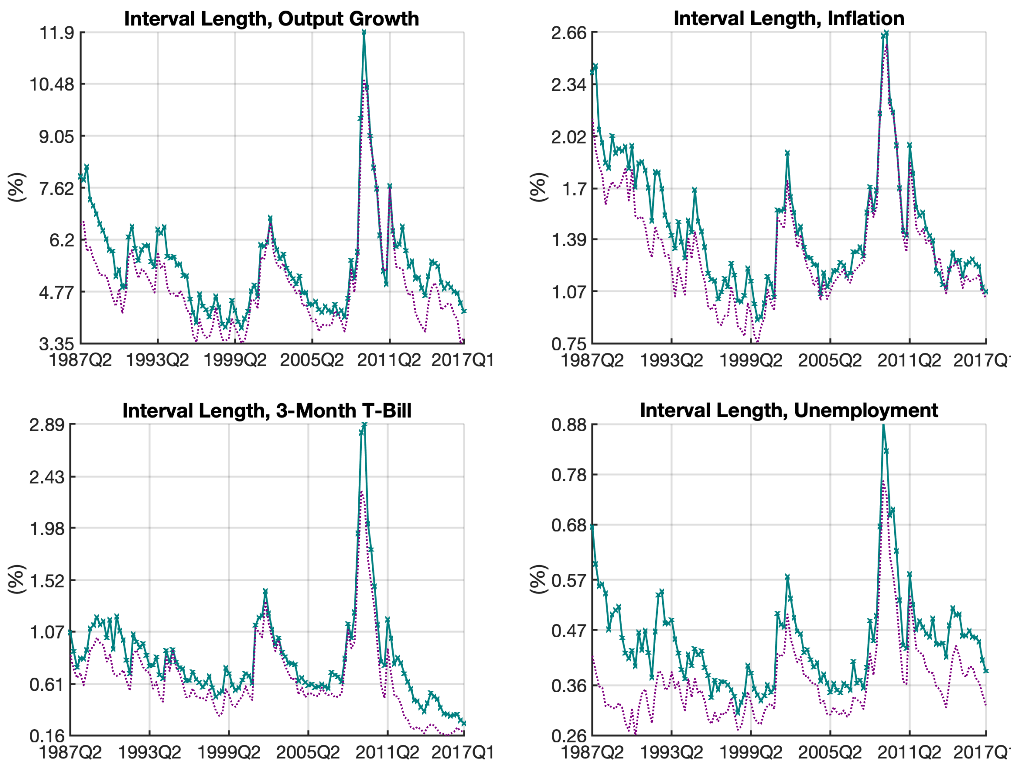
NOTE. Panel (a): Range indicates the smallest and largest empirical coverage rate of the 70% prediction interval across the 24 possible orderings. Panel (b): Range indicates the narrowest and widest prediction interval across the 24 possible orderings.

In parallel to the case of density prediction, the differences across orderings are economically relevant. Figure 5 shows the average length of the one-quarter-ahead 70% predictive intervals computed recursively over the evaluation sample. For simplicity, we focus on the prediction intervals associated with the best and worst ordering, where the best (worst) ordering is the one with the smallest (largest) difference between the empirical coverage and the nominal coverage rate over the evaluation sample. Had the worst ordering instead of the best ordering been systematically used at a central bank for a span of 10 years, it would have persistently offered policymakers a less sharp prediction. Appendix A.5 shows that the same holds when looking at four- and eight-quarter-ahead prediction intervals.

3.5 Summary for the CSP-SV model

In this section we have shown that (1) the order of the variables is important for forecasting performance, (2) if we care about more than point prediction the difference in performance is economically relevant, (3) the best ordering depends on the variable and forecast horizon of interest, and (4) the best ordering varies over time. For these reasons as well as due to the potential infeasibility of checking all possible orderings in larger models, it is interesting to

Figure 5: One-Quarter-Ahead Prediction Interval and Ordering



NOTE. Length of the corresponding intervals. Intervals are computed based on the one-quarter-ahead predictive density throughout the evaluation sample. The difference between the empirical coverage and the nominal coverage is largest for the worst ordering and smallest for the best ordering.

compare the performance of the CSP-SV model with ordering invariant models. We do that in the next section.

4 Ordering Invariant Models

The ordering dependent forecasting performance documented in Section 3 motivates us to consider two ordering invariant strategies for modeling stochastic volatility. To be ordering invariant, one has to start with a prior for the reduced-form covariance matrix that is ordering invariant. The first modeling approach places a Wishart prior on $\boldsymbol{\Omega}_t$ and it is known as dynamic linear model with discounted Wishart stochastic volatility (DW-SV). The second modeling approach decomposes the reduced-form covariance matrix into $\boldsymbol{\Omega}_t = \boldsymbol{D}_t\boldsymbol{P}_t\boldsymbol{D}_t'$, where $\boldsymbol{D}_t$ is a diagonal matrix and $\boldsymbol{P}_t$ is a correlation matrix, and it imposes an ordering invariant prior on $\boldsymbol{D}_t$ and $\boldsymbol{P}_t$. We adapt such decomposition, inspired by Engle (2002), into a time-varying parameter VAR and we label the resulting model a time-varying parameter VAR with dynamic and stochastic correlation-based multivariate stochastic volatility model (DSC-SV).[9]

The rationale behind our choice of these two ordering invariant approaches is as follows. We choose the DW-SV model for three reasons. First, it is widely used and a practical choice in financial time series modelling, e.g., Prado and West (2010). Since stochastic volatility is prevalent in this area it is natural to ask whether the utility of this framework can extrapolate to macroeconomic forecasting. Second, it was proposed as a tool for vector autoregression with stochastic volatility by the seminal work of Uhlig (1997). Third, it offers tractable and convenient filtering formulas, which facilitates likelihood evaluation.

We propose the DSC-SV model for analogous reasons. First, like the DW-SV, the type of approach introduced by Engle (2002) has been widely adopted in financial econometrics suggesting it could also prove useful for macroeconomic forecasting. Second, just like the discounting Wishart process, the approach of decomposing the reduced-form covariance matrix into $\boldsymbol{\Omega}_t = \boldsymbol{D}_t\boldsymbol{P}_t\boldsymbol{D}_t'$ can be integrated in an ordering invariant time-varying parameters VAR. Third, while not straightforward, we develop a feasible MCMC algorithm to estimate the proposed model based on the elliptical slice sampler of Murray, Adams and Mackay (2010).[10]

We are particularly interested in assessing whether these models can have equal or superior forecasting performance than the CSP-SV under any of its the orderings. Thus, we estimate the ordering invariant models on the same data (and training sample) as the CSP-SV. Likewise, we include two-lags.

4.1 The DW-SV

This DW-SV was developed by West and Harrison (1997), Uhlig (1997), Prado and West (2010), and Bognanni (2018). Since the model is well documented in the literature let us provide a succinct summary of its structure and how to conduct Bayesian inference with it. Let the vector of endogenous variables $\boldsymbol{y}_t$ evolve as follows:

$$\begin{aligned}
\boldsymbol{y}_t' &= \boldsymbol{x}_t'\boldsymbol{B}_t + \boldsymbol{u}_t', \quad \boldsymbol{u}_t \sim \mathrm{N}(\boldsymbol{0}_{n\times 1}, \boldsymbol{H}_t^{-1}) && (5)\\
\boldsymbol{B}_t &= \boldsymbol{B}_{t-1} + \boldsymbol{\Omega}_t, \quad \boldsymbol{\Omega}_t \sim \mathrm{N}(\boldsymbol{0}_{m\times n}, \boldsymbol{W}, \boldsymbol{H}_t^{-1}) && (6)\\
\boldsymbol{H}_{t+1} &= \frac{\mathcal{U}(\boldsymbol{H}_t)'\boldsymbol{\Gamma}_{t+1}\mathcal{U}(\boldsymbol{H}_t)}{\beta}, \quad \beta \in (0,1), \quad \boldsymbol{\Gamma}_{t+1} \sim \mathcal{B}e\left(\frac{\beta h}{2}, \frac{1}{2}\right), \quad \beta h \geq n && (7)
\end{aligned}$$

[9] While Hartwig (2020) relies on a decomposition similar to the $\boldsymbol{\Omega}_t = \boldsymbol{D}_t\boldsymbol{P}_t\boldsymbol{D}_t'$ decomposition, he models $\boldsymbol{P}_t$ based on a Cholesky factorization, and therefore the resulting model is not ordering invariant.

[10] There are alternative models for the time-varying reduced-form covariance matrix that are ordering invariant. Some of them have been tested on macroeconomic data, and have been shown to produce a predictive distribution comparable to some orderings of the CSP-SV model without time-varying parameters, (Karapanagiotidis, 2014; Chan et al., 2020). In Section 4.4, we discuss these and other potential modeling strategies including those that have never been applied to macroeconomic data.

where $\boldsymbol{y}_t'$ is $1 \times n$ and $\boldsymbol{x}_t' = \left[\boldsymbol{y}_{t-1}', \ldots, \boldsymbol{y}_{t-p}', 1\right]$ is $1 \times m$. Let $\mathcal{D}_t = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t\}$ for $t = 1, \ldots, T$ and $\mathcal{D}_0 = \varnothing$. Hence, $\boldsymbol{B}_t$ is $m \times n$ and $\boldsymbol{H}_t$ is $n \times n$.

Given a prior distribution $(\boldsymbol{B}_0, \boldsymbol{H}_1) \mid \mathcal{D}_0 \sim \text{NW}\left(\boldsymbol{M}_{0|0}, \boldsymbol{C}_{0|0}, \boldsymbol{S}_{1|0}, \beta h\right)$, and conditioning on h, β, and $\boldsymbol{W}$, the posterior distribution of the DW-SV can be evaluated recursively using the results shown in Appendix A.1. Following Bognanni (2018), we set $h = 1/(1-\beta)$ and take into account the fact that β and $\boldsymbol{W}$ are unknown by imposing a prior distribution over these parameters. We assume that β has a four-parameter beta distribution. The parameters characterising the support of the distribution are $\beta_{\min} = n(n+1)^{-1}$ and $\beta_{\max} = 1$. The shape parameters are $a = 323.33$ and $b = 30$ so that the expected value of β is 0.92. We set $\boldsymbol{W} \sim \text{IW}(\boldsymbol{S}_0, \nu_0 - m + 1)$, where $\boldsymbol{S}_0 = \delta^2(\nu_0 - m - 1)(\boldsymbol{X}_0'\boldsymbol{X}_0)^{-1}$.[11] It is common to inform the selection of $\boldsymbol{S}_0$ using a pre-sample of ν_0 observations where $\boldsymbol{X}_0' = [\boldsymbol{x}_{-\nu_0+1}, \ldots, \boldsymbol{x}_0]$. This implies that the draws of $\boldsymbol{W}$ will be centered around $\frac{\boldsymbol{S}_0}{\nu_0 - m - 1} = \delta^2(\boldsymbol{X}_0'\boldsymbol{X}_0)^{-1}$. Inspired by Primiceri's (2005) approach, we will set $\nu_0 = 40$ and $\delta = 0.01$.

The remaining parameters of the distribution for $(\boldsymbol{B}_0, \boldsymbol{H}_1) \mid \mathcal{D}_0$, that is $(\boldsymbol{M}_{0|0}, \boldsymbol{C}_{0|0}, \boldsymbol{S}_{1|0}, \beta h)$. We follow the literature and use a pre-sample of ν_0 observations to set $\boldsymbol{M}_{0|0} = \left(\boldsymbol{X}_0'\boldsymbol{X}_0\right)^{-1}\boldsymbol{X}_0'\boldsymbol{Y}_0$, $\boldsymbol{C}_{0|0} = \kappa\left(\boldsymbol{X}_0'\boldsymbol{X}_0\right)^{-1}$, and $\boldsymbol{S}_{1|0} = \gamma\left(\frac{1}{\nu_0}\sum_{t=-\nu_0+1}^{0}\boldsymbol{u}_t'\boldsymbol{u}_t\right)^{-1}$, where $\boldsymbol{Y}_0' = [\boldsymbol{y}_{-\nu_0+1}, \ldots, \boldsymbol{y}_0]$. We set $\kappa = 4$ so that at the OLS estimates our prior for $\boldsymbol{B}_0$ given $\boldsymbol{H}_1$ is equivalent to the prior imposed by Primiceri (2005), and we set $\gamma \approx 1/n$ so that the expected value of $\boldsymbol{H}_1$ is in about the same order of magnitude as the inverse of the ordinary least squares (OLS) estimate of the variance matrix of the residuals based on the pre-sample. The DW-SV is estimated using the Gibbs Sampling algorithm proposed by Bognanni (2018). Appendix A.1 summarizes the algorithm.

4.2 The DSC-SV

The approach that decomposes the reduced-form covariance matrix $\boldsymbol{\Omega_t}$ into $\boldsymbol{D_t P_t D_t'}$ was first introduced into econometrics by Engle (2002). Since then, several econometric models rely on this decomposition to model time-varying covariance matrices; see e.g., the literature review by Chib, Omori and Asai (2009). To place an ordering invariant prior on the time-varying covariance matrix of a TVP-VAR-SV model, we follow Asai and McAleer (2009) and impose a Wishart process-based prior on the dynamics of the matrix $\boldsymbol{P_t}$. We label the resulting model a time-varying parameters VAR with dynamic and stochastic correlation-based multivariate stochastic volatility model.

The main difference relative to the CSP-SV is in the decomposition of the reduced-form covariance matrix. Formally, the DSC-SV model is defined as follows

$$\boldsymbol{y}_t' = vec(\boldsymbol{B}_t)'\boldsymbol{x}_t + \boldsymbol{u}_t', \quad \boldsymbol{u}_t \sim \text{N}(\boldsymbol{0}_{n\times 1}, \boldsymbol{D}_t\boldsymbol{P}_t\boldsymbol{D}_t), \text{for } t = 1, \ldots, T, \tag{8}$$

[11] When working with the DW-SV, the inverse-Wishart is parameterized as in Prado and West (2010).

where $\boldsymbol{B}_t$ is modelled as in Equation (2) of the CSP-SV, which we reproduce below

$$vec(\boldsymbol{B}_t) = vec(\boldsymbol{B}_{t-1}) + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim \mathrm{N}(\mathbf{0}_{mn\times 1}, \boldsymbol{Q}). \tag{9}$$

Turning to the decomposition of the covariance matrix of the reduced form shocks, i.e., $\boldsymbol{D}_t\boldsymbol{P}_t\boldsymbol{D}_t$, $\boldsymbol{D}_t$ is a diagonal matrix that contains the standard deviations of the reduced form shocks and $\boldsymbol{P}_t$ is a correlation matrix. The diagonal elements of $\boldsymbol{D}_t$ are modelled analogously to how $\boldsymbol{\Sigma}_t$ is modelled in the CSP-SV. Accordingly, we let $\boldsymbol{D}_t = diag(\sqrt{\boldsymbol{\delta}_t})$, where $\boldsymbol{\delta}_t = (\boldsymbol{\delta}_{1,t}, \boldsymbol{\delta}_{2,t}, ..., \boldsymbol{\delta}_{n,t})'$, and assume that $\log \boldsymbol{\delta}_t$ evolves analogously to $\log \boldsymbol{\sigma}_t$ in Equation (4), that is,

$$\log \boldsymbol{\delta}_t = \log \boldsymbol{\delta}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}_{n\times 1}, \boldsymbol{W}). \tag{10}$$

As mentioned, $\boldsymbol{P}_t$ is assumed to be a function of a Wishart process. More specifically, we start with the standardization suggested by Engle (2002), that transforms a positive definite matrix $\boldsymbol{Q}_t$ into a correlation matrix $\boldsymbol{P}_t$,

$$\boldsymbol{P}_t = (\boldsymbol{Q}_t^*)^{-1}\boldsymbol{Q}_t(\boldsymbol{Q}_t^*)^{-1}, \quad \text{where } \boldsymbol{Q}_t^* = (diag(vecd(\boldsymbol{Q}_t))^{1/2}. \tag{11}$$

Then, we model the dynamic evolution of $\boldsymbol{Q}_t$ based on the following Wishart process,

$$(\boldsymbol{Q}_{t+1})^{-1}|k, \boldsymbol{S}_t \sim \mathrm{W}(k, \boldsymbol{S}_t), \quad \text{where } \boldsymbol{S}_t^{-1} = k(\boldsymbol{Q}_t)^{d/2}\boldsymbol{A}^{-1}(\boldsymbol{Q}_t)^{d/2}, \tag{12}$$

and k is the degrees of freedom parameter to be estimated. The time-dependent scale parameter of the Wishart distribution $\boldsymbol{S}_t$ is a function of $\boldsymbol{Q}_t$, a degrees of freedom parameter k, another scalar parameter d that governs the general persistence of $\boldsymbol{Q}_t$, and a $n \times n$ positive definite symmetric matrix $\boldsymbol{A}$. The fractional power $(\boldsymbol{Q}_t)^{-d/2}$ is defined by using a singular value decomposition.[12]

Equations (8)-(12) summarize the DSC-SV model. Let us now discuss the priors that we use to conduct Bayesian inference. We impose the same exact prior on $\boldsymbol{Q}$ as in the CSP-SV model. For the parameters governing the Wishart process $\boldsymbol{Q}_t$ (i.e., d, k, and $\boldsymbol{A}^{-1}$), we assume the same prior distribution as in Asai and McAleer (2009), that is,

$$d \sim \mathrm{U}(-1, 1), \quad k \sim \mathrm{EXP}(\lambda_0)I_{(n,\infty)}, \quad \boldsymbol{A}^{-1} \sim \mathrm{W}(\gamma_0, \boldsymbol{C}_0) \tag{13}$$

where $\mathrm{EXP}(\lambda_0)$ denotes an exponential distribution with the following density, $p(k) = \lambda_0 e^{-\lambda_0 k}$, and $I_{(n,\infty)}$ is an indicator function that takes the value of one when $k \in (n, \infty)$. Our choice for hyperparameters, $\lambda_0 = 5$, $\gamma_0 = n$, and $\boldsymbol{C}_0^{-1} = \gamma_0 \boldsymbol{I}_n$, implies a quite loose prior over $\boldsymbol{P}_t$ dynamics.

We assume that $\boldsymbol{W}$ is a diagonal matrix and let w_i denote its (i,i)-th element. Each $w_i \sim \mathrm{IG}(\underline{\mathrm{k}}_{w,i}, \underline{\mathrm{s}}_{w,i})$. We set $\underline{\mathrm{k}}_{w,i} = 2$ and $\underline{\mathrm{s}}_{w,i} = (\underline{\mathrm{k}}_{w,i} - 1)var(\widehat{e}_i)/T_0$, where $var(\widehat{e}_i)$ is a variance

[12]Suppose $\boldsymbol{X} = \boldsymbol{SVD}$, where $\boldsymbol{V}$ is a diagonal matrix. Then, $\boldsymbol{X}^d = \boldsymbol{SV}^d\boldsymbol{D}$.

of the OLS residual based on the training sample. Note also that $\boldsymbol{Q}_0$ is assumed to be fixed and known. Let $\widehat{\boldsymbol{R}}_0$ be the correlation matrix of the OLS residuals from the training sample, and let $\widehat{\boldsymbol{D}}_0$ be the diagonal matrix with diagonal elements being the standard deviation of the OLS residuals from the same sample. Then, we set $\boldsymbol{Q}_0 = \widehat{\boldsymbol{D}}_0 \widehat{\boldsymbol{P}}_0 \widehat{\boldsymbol{D}}_0$ so that $\widehat{\boldsymbol{P}}_0 = (\boldsymbol{Q}_0^*)^{-1} \boldsymbol{Q}_0 (\boldsymbol{Q}_0^*)^{-1}$. Appendix A.2 describes the MCMC algorithm that we use to generate a sequence of draws from the posterior distribution. Importantly, our algorithm is different from that of Asai and McAleer (2009). While they implement a two-quarter algorithm where the dynamic correlation matrices and their related parameters are drawn conditional on the posterior mean of variances (i.e., the posterior mean of $\{\boldsymbol{D}_1, \boldsymbol{D}_2, ..., \boldsymbol{D}_T\}$), we propose and implement a novel algorithm that generates draws from the full joint posterior distribution of unknowns using the elliptical sampling proposed by Murray, Adams and Mackay (2010).

Last but not least, let's highlight that as it was the case with the CSP-SV and the DLM-SV our choice of priors for the DSC-SV is in line with common choices in the literature.

4.3 Forecasting Performance

In this section we assess the out-of-sample prediction performance of the DW-SV and DSC-SV and we contrast it with the performance of the CSP-SV. As in Section 3, we focus on point prediction, density prediction, and interval prediction.

4.3.1 Point Prediction

Table 6 reproduces Table 1 and compares the CSP-SV model to the two ordering invariant models. The columns labeled DW-SV and DSC-SV denote the RMSE for the DW-SV and DSC-SV, respectively.

The table offers two main lessons. First, the DW-SV underperforms the other models under analysis. For all but one case, the RMSE of the DW-SV is higher than the RMSE associated with all the possible ordering of the CSP-SV. The exception is the eight-quarter-ahead RMSE of output growth where the DW-SV performs almost as well as the best CSP-SV. Second, the DSC-SV model produces point predictions roughly equal to the median outcomes of the CSP-SV. This is expected because the conditional mean in these two models is identical. The small differences in RMSEs are mainly due to different heteroscedasticity assumptions, which indirectly affects the conditional mean estimates and their point forecasts.

In any case, it could be argued that the differences in point prediction performance between the DW-SV and either of the two remaining models are tolerable from a macroeconomic forecasting perspective. For example, the one-quarter-ahead RMSE of output growth obtained when using the DW-SV is only 25 basis points larger than the CSP-SV with the smallest RMSE, and the average errors in the one-quarter-ahead RMSE for inflation, the 3-Month T-Bill rate,

Table 6: RMSE

One-Quarter-Ahead	CSP-SV Range	CSP-SV Median	DW-SV	DSC-SV
Output Growth	[2.46, 2.57]	2.51	2.70	2.57
Inflation	[0.60, 0.60]	0.60	0.62	0.60
3-Month T-bill	[0.33, 0.34]	0.33	0.37	0.34
Unemployment	[0.20, 0.20]	0.20	0.21	0.20
Four-Quarter-Ahead	**CSP-SV Range**	**CSP-SV Median**	**DW-SV**	**DSC-SV**
Output Growth	[2.67, 2.73]	2.70	3.02	2.57
Inflation	[0.77, 0.80]	0.78	0.88	0.79
3-Month T-bill	[1.14, 1.18]	1.16	1.27	1.15
Unemployment	[0.80, 0.83]	0.82	0.92	0.83
Eight-Quarter-Ahead	**CSP-SV Range**	**CSP-SV Median**	**DW-SV**	**DSC-SV**
Output Growth	[2.56, 2.61]	2.59	2.57	2.49
Inflation	[0.86, 0.92]	0.88	1.14	0.88
3-Month T-bill	[1.92, 2.01]	1.95	2.13	1.89
Unemployment	[1.42, 1.47]	1.44	1.61	1.44

NOTE. DW-SV denotes the RMSE for the DW-SV model, and DSC-SV denotes the RMSE for the DSC-SV model.

and the unemployment rate are roughly equivalent.

4.3.2 Density Prediction

When comparing the performance in terms of predictive densities, it is evident that there are large and economically meaningful discrepancies across the three models. Table 7 reproduces Table 2 and compares the CSP-SV model to the two ordering invariant models.

The table offers three main results. First, the DW-SV underperforms the CSP-SV under all orderings as well as the DSC-SV. Notice that in terms of joint density prediction, the sum of one-quarter-ahead LPSs of the DW-SV is about 70 log units lower than that of the median CSP-SV and the DSC-SV. This large difference is also a feature of the sum of one-quarter-ahead marginal LPSs for the each variable. The same results are obtained when looking at the four-quarter- and eight-quarter-ahead forecast horizons. Second, the one-quarter-ahead predictive performance of the DSC-SV is competitive relative to CSP-SV. In most cases it is within the CSP-SV range, and only in five cases it performs worse than the worst ordering for CSP-SV. Third, at the four-quarter- and eight-quarter-ahead forecast horizons, the joint density prediction based on the DSC-SV is slightly worse than the one based on the CSP-SV under any of its orderings. Even so, the marginal predictive densities based on the DSC-SV and the median outcomes of the CSP-SV are of quite similar quality.

Table 7: Log Predictive Score

One-Quarter-Ahead	CSP-SV Range	CSP-SV Median	DW-SV	DSC-SV
Joint	[-371.86,-346.23]	-354.31	-425.88	-359.90
Output Growth	[-279.41,-274.29]	-276.56	-382.17	-279.72
Inflation	[-111.03,-106.76]	-108.68	-125.84	-110.39
3-Month T-bill	[-33.10,-15.74]	-21.22	-51.42	-10.33
Unemployment	[26.94,33.53]	31.84	-25.24	20.80
Four-Quarter-Ahead	CSP-SV Range	CSP-SV Median	DW-SV	DSC-SV
Joint	[-739.69,-716.06]	-729.74	-813.50	-748.34
Output Growth	[-290.48,-284.50]	-287.11	-400.50	-287.96
Inflation	[-149.13,-142.56]	-144.98	-175.35	-147.35
3-Month T-bill	[-192.47,-183.55]	-187.70	-208.89	-180.38
Unemployment	[-138.68,-123.95]	-129.53	-151.43	-144.42
Eight-Quarter-Ahead	CSP-SV Range	CSP-SV Median	DW-SV	DSC-SV
Joint	[-948.09,-910.97]	-931.17	-996.95	-950.44
Output Growth	[-292.80,-283.26]	-287.03	-418.27	-287.66
Inflation	[-179.46,-169.80]	-174.75	-217.19	-176.06
3-Month T-bill	[-267.13,-260.09]	-262.18	-274.95	-256.58
Unemployment	[-227.22,-205.30]	-212.01	-222.61	-215.87

Note. DW-SV denotes the LPS for the DW-SV model and DSC-SV denotes the LPS for the DSC-SV model.

4.3.3 Interval Prediction

Finally, we turn to contrasting the empirical coverage rates and the length of 70% prediction intervals. Table 8 reproduces Table 5 and compares the CSP-SV model to the two ordering invariant models. The most salient finding that emerges from Panel (a) is that the empirical coverage rates based on the DW-SV are much higher than those of the models for all variables and all horizons. Consequently, the predictive density based on the DW-SV is much wider than what it should be based on the desired nominal coverage rate, which could explain a low LPS of the DW-SV relative to other models described above. In contrast, empirical coverage rates for one-quarter-ahead forecasts based on DSC-SV are very close to the desired nominal coverage rate. Notice that while under some orderings the CSP-SV produces one-quarter-ahead prediction interval coverage rates significantly above (i.e., 84%) or below (i.e., 64%) the nominal rate, all one-quarter-ahead prediction intervals implied by the DSC-SV are at most 3 percentage points away from nominal coverage rate. Turning to the four-quarter-ahead prediction intervals, the DSC-SV has similar empirical coverage rates to the median implied by the DSC-SV orderings. For the eight-quarter-ahead prediction intervals, neither model produces a well-calibrated prediction interval.

Table 8: Interval prediction comparisons

(a) Empirical coverage rate of 70% prediction interval

One-Quarter-Ahead	CSP-SV Range	CSP-SV Median	DW-SV	DSC-SV
Output Growth	[0.68,0.74]	0.70	0.86	0.70
Inflation	[0.67,0.75]	0.69	0.84	0.69
3-Month T-bill	[0.72,0.84]	0.77	0.87	0.73
Unemployment	[0.64,0.78]	0.69	0.79	0.68
Four-Quarter-Ahead	**CSP-SV Range**	**CSP-SV Median**	**DW-SV**	**DSC-SV**
Output Growth	[0.71,0.79]	0.74	0.86	0.78
Inflation	[0.83,0.88]	0.86	0.89	0.83
3-Month T-bill	[0.67,0.80]	0.72	0.82	0.68
Unemployment	[0.61, 0.78]	0.68	0.76	0.65
Eight-Quarter-Ahead	**CSP-SV Range**	**CSP-SV Median**	**DW-SV**	**DSC-SV**
Output Growth	[0.78,0.86]	0.81	0.93	0.87
Inflation	[0.89,0.94]	0.93	0.88	0.88
3-Month T-bill	[0.62,0.72]	0.66	0.72	0.63
Unemployment	[0.54,0.65]	0.57	0.72	0.60

(b) Average length of 70% prediction interval

One-Quarter-Ahead	CSP-SV Range	CSP-SV Median	DW-SV	DSC-SV
Output Growth	[4.62,5.51]	4.94	7.06	5.01
Inflation	[1.32,1.47]	1.35	1.77	1.33
3-Month T-bill	[0.66,0.87]	0.71	1.15	0.62
Unemployment	[0.35,0.45]	0.37	0.53	0.40
Four-Quarter-Ahead	**CSP-SV Range**	**CSP-SV Median**	**DW-SV**	**DSC-SV**
Output Growth	[5.11,6.18]	5.51	8.53	5.81
Inflation	[2.10,2.34]	2.18	2.92	2.19
3-Month T-bill	[2.14,2.67]	2.29	3.14	2.00
Unemployment	[1.13,1.40]	1.22	1.77	1.31
Eight-Quarter-Ahead	**CSP-SV Range**	**CSP-SV Median**	**DW-SV**	**DSC-SV**
Output Growth	[5.55,6.82]	5.98	10.39	6.54
Inflation	[2.83,3.17]	2.99	4.06	3.06
3-Month T-bill	[3.45,4.29]	3.68	5.03	3.30
Unemployment	[1.83,2.22]	1.93	2.88	2.11

NOTE. Panel (a): DW-SV indicates the empirical coverage rate of the DW-SV. DSC-SV indicates the empirical coverage rate of the DCS-SV. Panel (b): DW-SV indicates the average length of the 70% prediction interval in the DW-SV. DSC-SV indicates the average length of the 70% prediction interval in the DSC-SV.

Panel (b) confirms the insights obtained from Panel (a). The DW-SV tends to have wider intervals relative to the prediction intervals implied by the CSP-SV and the DSC-SV for all variables and at all horizons, which is in line with the higher empirical coverage rates documented above. Overall, the average length of the prediction intervals based on the CSP-SV under all of its orderings and the DSC-SV are comparable: the length based on the DSC-SV falls into the CSP-SV range. The few exceptions are the prediction intervals for the 3-Month T-Bill. In this case, the intervals based on the DSC-SV are shorter than the other intervals, which explains why the DSC-SV produces better predictive density, measured by the log predictive score, for the 3-Month T-Bill relative to the other models.

4.4 Discussion

Features Underlying the Forecasting Performance. Our analysis shows that the DW-SV presents excessively wide predictive densities for all variables at all horizons relative to the other models under analysis. This is related to two restrictive assumptions that make the DW-SV analytically tractable.

First, the shocks to the time-varying parameters, $\boldsymbol{B}_t$, are scaled by the time-varying reduced-form covariance matrix, $\boldsymbol{H}_t^{-1}$. While in some cases this can be a reasonable assumption, it restricts the variance of the parameters governing the conditional mean to be an increasing function of the covariance matrix of the reduced-form innovations. The CSP-SV and the DSC-SV are not subject to such a restriction and a consequence the variance of the predictive density can be smaller than in the presence of the restriction as in our application.

Second, the DW-SV imposes a discounting stochastic process driven by a singular multivariate Beta distribution, which when combined with a Wishart prior distribution on the time-varying reduced-form covariance matrices induces a Wishart posterior distribution. Consequently, there are at most two tightness parameters (β and h) that govern the properties of the shocks underlying the stochastic process for the reduced-form covariance matrix. Hence, even though the Wishart-based modeling is a parsimonious approach, it is too restrictive relative to the CSP-SV and the DSC-SV.[13]

Turning to the DSC-SV, notice that it can be viewed as a hybrid approach between the DW-SV and CSP-SV modeling approaches. This is because it decomposes the time-varying reduced-form covariance matrices into two pieces: a time-varying conditional variance and a time-varying conditional correlation. The former is modeled similarly to the CSP-SV (i.e., by means of a random-walk process) and the latter is modeled similarly to the DW-SV model (i.e., by means of a Wishart-based process). Our forecasting performance evaluation shows that by assuming a random-walk process on the logarithm of the conditional variances, the

[13]See Lopes and Polson (2014) for a general comparison between Wishart priors and the prior induced by the Cholesky decomposition for the Bayesian estimation of a non-time-varying covariance matrix.

marginal predictive densities are comparable to those based on the CSP-SV. Thus, imposing a random-walk process either on the standard deviation of the structural shocks (i.e., $\boldsymbol{\Sigma}_t$) as in the CSP-SV or on the standard deviation of the reduced-form shock (i.e, $\boldsymbol{D}_t$) as in the DSC-SV leads to superior out-of-sample forecasting performance relative to the DW-SV.

Importantly, the DSC-SV is ordering-invariant because the time-varying correlation matrix is modelled via a Wishart process. This is an appealing feature because it opens the door to structural analysis, however our empirical exercise shows that the multi-step *joint* predictive density produced by this model underperforms the CSP-SV. This implies that the correlation dynamics of the DSC-SV model could be misspecified relative to the CSP-SV: the Wishart distribution-based approach for the time-varying correlation matrices may be too restrictive as the single scalar parameter (k) controls the tightness of the distribution.

On Alternative Approaches. Although we argue that the DW-SV is too tightly parameterized to fit macroeconomic data, there are more flexible Wishart or inverted Wishart processes for multivariate stochastic volatility models in exchange for higher computational complexity. Some of these models have been applied to macroeconomic forecasting problems. For example, Karapanagiotidis (2014) compares the predictive performance of the inverse Wishart stochastic volatility model with some models based on the Cholesky decomposition using four U.S. macroeconomic variables. And, Chan et al. (2020) develop a VAR model with a multivariate stochastic volatility inverse Wishart process. They compare its predictive performance with other VAR models with stochastic volatility based on the Cholesky decomposition using twenty U.S. macroeconomic variables. Related VARs with Wishart processes are also employed in structural economic analysis, see e.g., Rondina (2013) and Shin and Zhong (2020).

Unlike our DSC-SV, it is possible to model the time-varying correlation matrix, $\boldsymbol{P}_t$, using random-walk processes rather than (inverse) Wishart processes. In particular, Archakov and Hansen (2020) introduce a numerically invertible mapping from the space of non-singular $n \times n$ correlation matrices to a $n(n-1)/2 \times 1$ real vector, $\boldsymbol{\gamma}(\cdot) : \mathbb{C}^{n\times n} \to \mathbb{R}^{n(n-1)/2}$ and show that the mapping is ordering invariant. Using their mapping, it is possible to model $\boldsymbol{P}_t$ as $\boldsymbol{\gamma}(\boldsymbol{P}_t) = \boldsymbol{\gamma}(\boldsymbol{P}_{t-1}) + \zeta_t, \quad \zeta_t \sim \mathrm{N}(0, \boldsymbol{S})$, where $\boldsymbol{S}$ is a $n(n-1)/2 \times n(n-1)/2$ positive definite matrix. Our preliminary computations reveal that such an approach is on par with the DSC-SV in terms of predictive performance.

Another yet interesting approach is to assume a common stochastic volatility so that the reduced-form covariance matrix can be written as $\boldsymbol{\Omega}_t = exp(h_t)\boldsymbol{\Omega}$, where h_t is a scalar log stochastic volatility process and $\boldsymbol{\Omega}$ is a $n \times n$ positive definite matrix that is constant over time. As long as the prior distribution of $\boldsymbol{\Omega}$ is ordering invariant (e.g., an inverse Wishart distribution), the resulting multivariate stochastic volatility model is robust to variable ordering. Carriero, Clark and Marcellino (2016) and Chan (2020) integrate this type of common stochastic volatility models into VARs to fit several macroeconomic variables.

Finally, even though it is less popular in macroeconomics, it may be also possible to model the time-varying reduced-form covariance matrix based on observation-driven approaches. This type of models includes the multivariate generalized autoregressive conditional heteroskedasticity (GARCH) models surveyed in Bauwens, Laurent and Rombouts (2006), the dynamic conditional correlation (DCC) model of Engle (2002), and the multivariate generalized autoregressive score (GAS) model of Creal, Koopman and Lucas (2011).

5 Conclusion

This paper shows that the out-of-sample forecasting performance of the CSP-SV depends on the ordering of the variables. When the object of interest is density and interval prediction, the differences are noticeable and persistent. Hence, our results offer useful guidance for policymakers and forecasters at central banks, who have been increasingly interested in density forecasts. In addition, our paper proposes an ordering invariant DSC-SV approach that features an out-of-sample forecasting performance comparable to the CSP-SV.

Finally, let us highlight that the priors used in each model are based on standard specifications. The results in Giannone, Lenza and Primiceri (2015) and Amir-Ahmadi, Matthes and Wang (2020) show that additional forecasting gains for each model could be obtained by optimally choosing the prior hyperparameters controlling the informativeness of the priors and the smoothness of the time-varying parameters.

References

Amir-Ahmadi, P., C. Matthes, and M.-C. Wang (2020). Choosing Prior Hyperparameters: With Applications to Time-Varying Parameter Models. *Journal of Business & Economic Statistics 38*(1), 124–136.

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

Archakov, I. and P. R. Hansen (2020). A New Parametrization of Correlation Matrices. *arXiv preprint arXiv:2012.02395*.

Asai, M. and M. McAleer (2009). The Structure of Dynamic Correlations in Multivariate Stochastic Volatility Models. *Journal of Econometrics 150*(2), 182–192.

Askanazi, R., F. X. Diebold, F. Schorfheide, and M. Shin (2018). On the Comparison of Interval Forecasts. *Journal of Time Series Analysis 39*(6), 953–965.

Atchadé, Y. F. and J. S. Rosenthal (2005). On Adaptive Markov Chain Monte Carlo Algorithms. *Bernoulli 11*(5), 815–828.

Baumeister, C. and G. Peersman (2013). Time-Varying Effects of Oil Supply Shocks on the US Economy. *American Economic Journal: Macroeconomics 5*(4), 1–28.

Bauwens, L., S. Laurent, and J. V. Rombouts (2006). Multivariate GARCH Models: A Survey. *Journal of Applied Econometrics 21*(1), 79–109.

Bognanni, M. (2018). A Class of Time-Varying Parameter Structural VARs for Inference under Exact or Set Identification. *Federal Reserve Bank of Cleveland Working Paper 1*(18-11), 1–61.

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

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

Chan, J. C. (2020). Large Bayesian VARs: A Flexible Kronecker Error Covariance Structure. *Journal of Business & Economic Statistics 38*(1), 68–79.

Chan, J. C., A. Doucet, R. León-González, and R. W. Strachan (2020). Multivariate Stochastic Volatility with Co-heteroscedasticity. *GRIPS Discussion Papers 20-09*.

Chib, S., Y. Omori, and M. Asai (2009). Multivariate Stochastic Volatility. In *Handbook of Financial Time Series*, pp. 365–400. Springer.

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

Cogley, T. and T. J. Sargent (2005). Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US. *Review of Economic Dynamics 8*(2), 262–302.

Creal, D., S. J. Koopman, and A. Lucas (2011). A Dynamic Multivariate Heavy-Tailed Model for Time-Varying Volatilities and Correlations. *Journal of Business & Economic Statistics 29*(4), 552–563.

D'Agostino, A., L. Gambetti, and D. Giannone (2013). Macroeconomic Forecasting and Structural Change. *Journal of Applied Econometrics 28*(1), 82–101.

Del Negro, M. and G. E. Primiceri (2015). Time Varying Structural Vector Autoregressions and Monetary Policy: A Corrigendum. *Review of Economic Studies 82*(4), 1342–1345.

Diebold, F. X. and R. S. Mariano (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics 13*(3).

Engle, R. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics 20*(3), 339–350.

Galí, J. and L. Gambetti (2015). The Effects of Monetary Policy on Stock Market Bubbles: Some Evidence. *American Economic Journal: Macroeconomics 7*(1), 233–57.

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

Haldane, J. (1942). Moments of the Distributions of Powers and Products of Normal Variates. *Biometrika 32*(3/4), 226–242.

Hartwig, B. (2020). Robust Inference in Time-Varying Structural VAR Models: The DC-Cholesky Multivariate Stochastic Volatility Model. *Deutsche Bundesbank Discussion Paper 34/2020*.

Karapanagiotidis, P. (2014). Improving Bayesian VAR Density Forecasts through Autoregressive Wishart Stochastic Volatility. *Working Paper*.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *Review of Economic Studies 65*(3), 361–393.

Lopes, H. F. and N. G. Polson (2014). Bayesian Instrumental Variables: Priors and Likelihoods. *Econometric Reviews 33*(1-4), 100–121.

Murray, I., R. Adams, and D. Mackay (2010). Elliptical Slice Sampling. *Journal of Machine Learning Research: W&CP 9*, 541–548.

Prado, R. and M. West (2010). *Time Series: Modeling, Computation, and Inference*. CRC Press.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions and Monetary Policy. *Review of Economic Studies 72*(3), 821–852.

Rondina, F. (2013). Time Varying SVARs, Parameter Histories, and the Changing Impact of Oil Prices on the US Economy. *Working Paper*.

Shin, M. and M. Zhong (2020). A New Approach to Identifying the Real Effects of Uncertainty Shocks. *Journal of Business & Economic Statistics 38*(2), 367–379.

Uhlig, H. (1997). Bayesian Vector Autoregressions with Stochastic Volatility. *Econometrica: Journal of the Econometric Society 65*(1), 59–73.

West, M. and J. Harrison (1997). *Bayesian Forecasting and Dynamic Models.* Springer Science & Business Media.

A Appendix

A.1 Inference in the DW-SV

Bognanni's (2018) Gibbs Sampler relies on two steps to sample from $p(\mathcal{B}_T, \mathcal{H}_T, \beta, \boldsymbol{W}|\mathcal{D}_T)$, where $\mathcal{B}_t = \{\boldsymbol{B}_0, \ldots, \boldsymbol{B}_t\}$ and $\mathcal{H}_t = \{\boldsymbol{H}_1, \ldots, \boldsymbol{H}_t\}$ for $t = 0, \ldots, T$. The first step consists of drawing from $p(\boldsymbol{W} \mid \beta, \mathcal{B}_T, \mathcal{H}_T, \mathcal{D}_T)$, which is straightforward given that $p(\boldsymbol{W} \mid \beta, \mathcal{B}_T, \mathcal{H}_T, \mathcal{D}_T)$ is an inverse-Wishart distribution. The second step consists of drawing from $p(\beta, \mathcal{B}_T, \mathcal{H}_T \mid \boldsymbol{W}, \mathcal{D}_T)$. The key to obtaining draws from such distribution is to notice that we can rewrite

$$p(\beta, \mathcal{B}_T, \mathcal{H}_T \mid \boldsymbol{W}, \mathcal{D}_T) = p(\mathcal{B}_T, \mathcal{H}_T \mid \beta, \boldsymbol{W}, \mathcal{D}_T) p(\beta \mid \boldsymbol{W}, \mathcal{D}_T).$$

The reader should notice that it is straightforward to draw from $p(\mathcal{B}_T, \mathcal{H}_T \mid \beta, \boldsymbol{W}, \mathcal{D}_T)$ based on the work of Uhlig (1997) and Prado and West (2010) summarized by Algorithm 1 and Table A.1. Hence, all that is left is to draw from $p(\beta \mid \boldsymbol{W}, \mathcal{D}_T)$. We accomplish this using a Metropolis-within-Gibbs step as in Bognanni (2018).

Table A.1: Summary for $t = 1, \ldots, T$

Distribution of Interest	Distributional Family	Parameters
Step 1– Prior at time t		
$(\boldsymbol{B}_{t-1}, \boldsymbol{H}_t) \mid \mathcal{D}_{t-1}$	$\mathrm{NW}\left(\boldsymbol{M}_{t-1\|t-1}, \boldsymbol{C}_{t-1\|t-1}, \boldsymbol{S}_{t\|t-1}, \beta h\right)$	$\boldsymbol{M}_{t-1\|t-1}, \boldsymbol{C}_{t-1\|t-1}, \boldsymbol{S}_{t\|t-1}, \beta h$
$(\boldsymbol{B}_t, \boldsymbol{H}_t) \mid \mathcal{D}_{t-1}$	$\mathrm{NW}\left(\boldsymbol{M}_{t\|t-1}, \boldsymbol{C}_{t\|t-1}, \boldsymbol{S}_{t\|t-1}, \beta h\right)$	$\boldsymbol{M}_{t\|t-1} = \boldsymbol{M}_{t-1\|t-1}$ $\boldsymbol{C}_{t\|t-1} = \boldsymbol{C}_{t-1\|t-1} + \boldsymbol{W}$
$\boldsymbol{H}_t \mid \mathcal{D}_{t-1}$	$\mathrm{W}(\boldsymbol{S}_{t\|t-1}, \beta h)$	
$\boldsymbol{B}_t \mid \boldsymbol{H}_t, \mathcal{D}_{t-1}$	$\mathrm{N}(\boldsymbol{M}_{t\|t-1}, \boldsymbol{C}_{t\|t-1}, \boldsymbol{H}_t^{-1})$	
Step 2– Posterior at time t		
$(\boldsymbol{B}_t, \boldsymbol{H}_t) \mid \mathcal{D}_t$	$\mathrm{NW}\left(\boldsymbol{M}_{t\|t}, \boldsymbol{C}_{t\|t}, \boldsymbol{S}_{t\|t}, \beta h + 1\right)$	$\boldsymbol{M}_{t\|t} = \boldsymbol{C}_{t\|t}(\boldsymbol{C}_{t\|t-1}^{-1}\boldsymbol{M}_{t\|t-1} + \boldsymbol{x}_t\boldsymbol{y}_t')$ $\boldsymbol{C}_{t\|t}^{-1} = \boldsymbol{C}_{t\|t-1}^{-1} + \boldsymbol{x}_t\boldsymbol{x}_t'$ $\boldsymbol{S}_{t\|t}^{-1} = \boldsymbol{S}_{t\|t-1}^{-1} + \boldsymbol{e}_t\left(1 - \boldsymbol{x}_t'\boldsymbol{C}_{t\|t}\boldsymbol{x}_t\right)\boldsymbol{e}_t'$ where $\boldsymbol{e}_t = \boldsymbol{y}_t - \boldsymbol{M}_{t\|t-1}'\boldsymbol{x}_t$
$\boldsymbol{H}_t \mid \mathcal{D}_t$	$\mathrm{W}(\boldsymbol{S}_{t\|t}, \beta h + 1)$	
$\boldsymbol{B}_t \mid \boldsymbol{H}_t, \mathcal{D}_t$	$\mathrm{N}(\boldsymbol{M}_{t\|t}, \boldsymbol{C}_{t\|t}, \boldsymbol{H}_t^{-1})$	
Step 3– Prior at time $t+1$		
$(\boldsymbol{B}_t, \boldsymbol{H}_{t+1}) \mid \mathcal{D}_t$	$\mathrm{NW}\left(\boldsymbol{M}_{t\|t}, \boldsymbol{C}_{t\|t}, \boldsymbol{S}_{t+1\|t}, \beta h\right)$	$\boldsymbol{S}_{t+1\|t} = \frac{1}{\beta}\boldsymbol{S}_{t\|t}$

NOTE Filtering formulas for the DW-SV.

The posterior parameters are simulated using Algorithm 1:

Algorithm 1. *The following algorithm draws from $p(\mathcal{B}_T, \mathcal{H}_T|\mathcal{D}_T)$ given h, β, and $\boldsymbol{W}$.*

1. *Draw* $\boldsymbol{H}_T \mid \mathcal{D}_T \sim \mathrm{W}(\boldsymbol{S}_{T|T}, \beta h + 1)$.
2. *Draw* $\boldsymbol{B}_T \mid \boldsymbol{H}_T \sim \mathrm{N}(\boldsymbol{M}_{T|T}, \boldsymbol{C}_{T|T}, \boldsymbol{H}_T^{-1})$.

3. Let $t = T - 1$.

4. Draw $\boldsymbol{H}_t \mid \boldsymbol{H}_{t+1}, \mathcal{D}_t$ using equation $\boldsymbol{H}_t = \beta \boldsymbol{H}_{t+1} + \boldsymbol{\Upsilon}_t$, where $\boldsymbol{\Upsilon}_t \mid \mathcal{D}_t \sim \mathrm{W}(\boldsymbol{S}_{t|t}, 1)$.

5. Draw $\boldsymbol{B}_t \mid \boldsymbol{B}_{t+1}, \boldsymbol{H}_{t+1}, \mathcal{D}_t$ from $\boldsymbol{B}_t \mid \boldsymbol{B}_{t+1}, \boldsymbol{H}_{t+1}, \mathcal{D}_t \sim \mathrm{N}\left(\boldsymbol{M}_{t|t+1}, \boldsymbol{C}_{t|t+1}, \boldsymbol{H}_{t+1}^{-1}\right)$.

6. If $t \geq 2$, let $t \leftarrow t - 1$ and go to Step 4.

7. Draw $\boldsymbol{B}_0 \mid \boldsymbol{B}_1, \boldsymbol{H}_1, \mathcal{D}_0$ using the distribution described in Step 5.

A.2 Inference in the DSC-SV

We develop an algorithm that generates posterior draws of the unknown parameters in the DSC-SV. The algorithm generates draws that can be used to approximate the following posterior density

$$p(\mathcal{B}_T, \mathcal{D}_T, \mathcal{P}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{Q}, \boldsymbol{W} | \mathcal{D}_T) \tag{A.1}$$

where $\mathcal{B}_t = \{\boldsymbol{B}_1, \boldsymbol{B}_2, ..., \boldsymbol{B}_t\}$, $\mathcal{V}_t = \{\boldsymbol{D}_1, \boldsymbol{D}_2, ..., \boldsymbol{D}_t\}$, $\mathcal{P}_t = \{\boldsymbol{P}_1, \boldsymbol{P}_2, ..., \boldsymbol{P}_t\}$, and $\mathcal{D}_t = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t\}$ for $t = 1, \ldots, T$. Our proposed algorithm (i.e., Algorithm 2) is a Metropolis-Hastings within Gibbs sampling algorithm that iterates over multiple blocks. For ease of exposition, we first present the general algorithm and then we discuss the details of each step.

Algorithm 2. *The following draws from a density that approximates $p(\mathcal{B}_T, \mathcal{V}_T, \mathcal{P}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{Q}, \boldsymbol{W} | \mathcal{D}_T)$,*

1. Draw $\mathcal{B}_T$ from $p(\mathcal{B}_T | \mathcal{V}_T, \mathcal{P}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{Q}, \boldsymbol{W}, \mathcal{D}_T)$.

2. Draw $\boldsymbol{Q}$ from $p(\boldsymbol{Q}|, \mathcal{B}_T, \mathcal{V}_T, \mathcal{P}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{W}, \mathcal{D}_T)$.

3. Draw $\mathcal{P}_T$ from $p(\mathcal{P}_T | \mathcal{B}_T, \mathcal{V}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{Q}, \boldsymbol{W}, \mathcal{D}_T)$.

4. Draw $\boldsymbol{A}^{-1}$ from $p(\boldsymbol{A}^{-1} | \mathcal{B}_T, \mathcal{V}_T, \mathcal{P}_T, d, k, \boldsymbol{Q}, \boldsymbol{W}, \mathcal{D}_T)$.

5. Draw d from $p(d | \mathcal{B}_T, \mathcal{V}_T, \mathcal{P}_T, \boldsymbol{A}^{-1}, k, \boldsymbol{Q}, \boldsymbol{W}, \mathcal{D}_T)$.

6. Draw k from $p(k | \mathcal{B}_T, \mathcal{V}_T, \mathcal{P}_T, \boldsymbol{A}^{-1}, d, \boldsymbol{Q}, \boldsymbol{W}, \mathcal{D}_T)$.

7. Draw $\mathcal{V}_T$ from $p(\mathcal{V}_T | \mathcal{B}_T, \mathcal{P}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{Q}, \boldsymbol{W}, \mathcal{D}_T)$.

8. Draw $\boldsymbol{W}$ from $p(\boldsymbol{W} | \mathcal{B}_T, \mathcal{V}_T, \mathcal{P}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{Q}, \mathcal{D}_T)$.

In Steps 1 and 2, that is when drawing the time-varying parameter coefficients and the parameter governing their law of motion, we exactly follow Primiceri (2005). This is possible because we can recover the reduced form variance-covariance matrix using $\mathcal{V}_T$ and $\mathcal{P}_T$, $(\boldsymbol{A}_t^{-1})\boldsymbol{\Sigma}_t\boldsymbol{\Sigma}_t(\boldsymbol{A}_t^{-1})' = \boldsymbol{D}_t\boldsymbol{P}_t\boldsymbol{D}_t'$ for all t. In Steps 3 to 6, that is when drawing the time-varying correlation parameters, we follow Asai and McAleer (2009), who propose an MCMC algorithm that generates posterior draws of $(\mathcal{P}_T, \boldsymbol{A}^{-1}, k, d)$ from the following model

$$\boldsymbol{y}_t^{*'} \sim \mathrm{N}(\boldsymbol{0}_{1\times n}, 1, \boldsymbol{P}_t). \tag{A.2}$$

Conditional on $\mathcal{B}_T$ and $\mathcal{V}_T$, the DSC-SV model can be transformed into the above model by letting

$$\boldsymbol{y}_t^{*'} = (\boldsymbol{y}_t' - vec(\boldsymbol{B}_t)'\boldsymbol{x}_t)\boldsymbol{D}_t^{-1}. \tag{A.3}$$

We implement Step 7 (i.e., drawing $\mathcal{V}_T$) differently than in the methods based on the standard mixture approximation developed by Kim, Shephard and Chib (1998). We apply the elliptical slice sampling of Murray, Adams and Mackay (2010) to sample $\mathcal{V}_T$ from its conditional posterior distribution to deal with the time-varying correlation of reduced-form shocks introduced by $\boldsymbol{P}_t$. Step 8 is a standard inverse gamma posterior updating because we impose a conjugate prior on each non-zero entry of $\boldsymbol{W}$ (i.e., w_i for $i = 1, ..., n$).

While the papers mentioned above provide the details relevant to implement each step of Algorithm 2, below we discuss those that are new and essential to reproducing our results. In particular, we introduce a correction to one of the formulas in Asai and McAleer (2009) (note on step 3) and we illustrate how the novel elliptical slice sampler can be applied to sample the log stochastic volatilities (note on Step 7).

Note on Step 3. As discussed above, Asai and McAleer (2009) turn Step 3 into the problem of drawing $\{\boldsymbol{Q}_1^{-1}, ..., \boldsymbol{Q}_T^{-1}\}$ from the auxiliary model (A.2). This is possible because there is a well-defined mapping from $\{\boldsymbol{Q}_1^{-1}, ..., \boldsymbol{Q}_T^{-1}\}$ to $\mathcal{P}_T$,

$$\boldsymbol{P}_t = (\boldsymbol{Q}_t^*)^{-1}\boldsymbol{Q}_t(\boldsymbol{Q}_t^*)^{-1}, \quad \boldsymbol{Q}_t^* = (diag(vecd(\boldsymbol{Q}_t))^{1/2}, \text{ for } t = 1, \dots, T. \tag{A.4}$$

Then, we sample from

$$p(\{\boldsymbol{Q}_1^{-1}, ..., \boldsymbol{Q}_T^{-1}\}|\mathcal{B}_T, \mathcal{V}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{Q}, \boldsymbol{W}, \mathcal{D}_T)$$

by drawing $\boldsymbol{Q}_t^{-1}$ from the density

$$p(\boldsymbol{Q}_t^{-1}|\boldsymbol{Q}_{1:(t-1)}^{-1}, \boldsymbol{Q}_{(t+1):T}^{-1}, \mathcal{B}_T, \mathcal{V}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{Q}, \boldsymbol{W}, \mathcal{D}_T), \tag{A.5}$$

for $t = 1, 2, ..., T$, where $\boldsymbol{Q}_{s:t} = \{\boldsymbol{Q}_s, \dots, \boldsymbol{Q}_t\}$ with $t \geq s$ and $\boldsymbol{Q}_{T+1:T} = \emptyset$. Importantly, the conditional posterior density (A.5) can be simplified. Notice that for $t = 1, 2, ..., (T-1)$,

$$\begin{aligned}
&p(\boldsymbol{Q}_t^{-1}|\boldsymbol{Q}_{1:(t-1)}^{-1}, \boldsymbol{Q}_{(t+1):T}^{-1}, \mathcal{B}_T, \mathcal{V}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{Q}, \boldsymbol{W}, \mathcal{D}_T) \\
&\qquad\qquad\qquad \propto W_n(\boldsymbol{Q}_t^{-1}|k, \boldsymbol{S}_{t-1}) \times \mathrm{N}(0, \boldsymbol{P}_t) \times W_n(\boldsymbol{Q}_{t+1}^{-1}|k, \boldsymbol{S}_t) \\
&\qquad\qquad\qquad \propto \underbrace{e^{tr\left(-\frac{1}{2}(\boldsymbol{S}_{t-1}^{-1} + \boldsymbol{z}_t\boldsymbol{z}_t')\boldsymbol{Q}_t^{-1}\right)} \times |\boldsymbol{Q}_t^{-1}|^{(k+1-n-1)/2}}_{\propto W_n\left(\boldsymbol{Q}_t^{-1}|(k+1),\, (\boldsymbol{S}_{t-1}^{-1} + \boldsymbol{z}_t\boldsymbol{z}_t')^{-1}\right)} \\
&\qquad\qquad\qquad\quad \times \underbrace{|\boldsymbol{Q}_t^{-1}|^{(-1-dk)/2}|\boldsymbol{P}_t^{-1}|^{1/2} e^{tr\left(-\frac{1}{2}\boldsymbol{z}_t\boldsymbol{z}_t'(\boldsymbol{P}_t^{-1} - \boldsymbol{Q}_t^{-1})\right)} e^{tr\left(-\frac{1}{2}\boldsymbol{S}_t^{-1}\boldsymbol{Q}_{t+1}^{-1}\right)}}_{= f(\boldsymbol{Q}_t^{-1})}.
\end{aligned} \tag{A.6}$$

And, for $t = T$, we have[14]

$$p(\boldsymbol{Q}_T^{-1}|\boldsymbol{Q}_{1:(T-1)}^{-1}, \mathcal{B}_T, \mathcal{V}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{Q}, \boldsymbol{W}, \mathcal{D}_T) \propto \underbrace{e^{tr\left(-\frac{1}{2}(\boldsymbol{S}_{T-1}^{-1}+\boldsymbol{z}_T\boldsymbol{z}_T')\boldsymbol{Q}_T^{-1}\right)} \times |\boldsymbol{Q}_T^{-1}|^{(k+1-n-1)/2}}_{\propto W_n\left(\boldsymbol{Q}_T^{-1}|(k+1),(\boldsymbol{S}_{T-1}^{-1}+\boldsymbol{z}_T\boldsymbol{z}_T')^{-1}\right)}$$
$$\times \underbrace{e^{\left(-\frac{1}{2}tr(\boldsymbol{P}_T^{-1}-\boldsymbol{Q}_T^{-1})\boldsymbol{z}_T\boldsymbol{z}_T'\right)} \times \left|\prod_{i=1}^{n} q_{ii,t}^{1/2}\right|}_{=f(\boldsymbol{Q}_T^{-1})}. \tag{A.7}$$

Then, for $t = 1, 2, ..., T$, we employ a Metropolis-Hastings algorithm by generating a candidate draw $\boldsymbol{Q}_{t,*}^{-1}$ from the Wishart proposal density, $\mathrm{W}(\boldsymbol{Q}_t^{-1}|(k+1), (\boldsymbol{S}_{t-1}^{-1} + \boldsymbol{z}_t\boldsymbol{z}_t')^{-1})$ and accept it with probability $\min\left(\frac{f(\boldsymbol{Q}_{t,*}^{-1})}{f(\boldsymbol{Q}_{t,c}^{-1})}, 1\right)$, where $\boldsymbol{Q}_{t,c}^{-1}$ is the current state value. If the proposal draw is rejected, we set $\boldsymbol{Q}_t^{-1} = \boldsymbol{Q}_{t,c}^{-1}$.

Note on Step 5 and Step 6. We generate k and d based on a Random-Walk Metropolis-Hastings algorithm. Conditional posterior distributions of d and k are derived in Appendix A.3 of Asai and McAleer (2009). We adaptively tune and select the random-walk proposal densities so that we have 30% acceptance rate (Atchadé and Rosenthal, 2005).

Note on Step 7. We sample $\boldsymbol{\delta}_{i,1:T}$ from its conditional posterior density for each $i = 1, 2, ..., m$. These conditional posterior densities are derived from the following auxiliary model,

$$\begin{aligned}(\boldsymbol{y}_t' - vec(\boldsymbol{B}_t)'\boldsymbol{x}_t)' = \boldsymbol{u}_t \sim \mathrm{N}(0, \boldsymbol{D}_t\boldsymbol{P}_t\boldsymbol{D}_t')\\ \log \boldsymbol{\delta}_{i,t} = \log \boldsymbol{\delta}_{i,t-1} + \eta_{i,t}, \quad \eta_{i,t} \sim \mathrm{N}(0, w_i)\end{aligned} \tag{A.8}$$

with $\boldsymbol{D}_t = diag(\sqrt{\boldsymbol{\delta}_t})$ and $\boldsymbol{\delta}_t = (\boldsymbol{\delta}_{1,t}, \boldsymbol{\delta}_{2,t}, ..., \boldsymbol{\delta}_{n,t})'$. The likelihood function is then Gaussian and we have,

$$p(\boldsymbol{u}_1, ..., \boldsymbol{u}_T|\mathcal{V}_T, \mathcal{P}_T) \propto \prod_{t=1}^{T} \left|(\boldsymbol{D}_t\boldsymbol{P}_t\boldsymbol{D}_t')^{-1}\right|^{1/2} \exp\left(-\frac{1}{2}\boldsymbol{u}_t'(\boldsymbol{D}_t\boldsymbol{P}_t\boldsymbol{D}_t')^{-1}\boldsymbol{u}_t\right). \tag{A.9}$$

[14] In Appendix A.1 of Asai and McAleer (2009), one term is missing in their derivation of $f(\boldsymbol{Q}_T^{-1})$. We thank Manabu Asai for a helpful discussion.

We assume $\boldsymbol{\delta}_{i,0} \sim \mathrm{N}(m_{\delta,i,0}, w_i V_{\delta,i,0})$ to obtain convenient forms for the first two moments of the prior distribution for each $\boldsymbol{\delta}_{i,t}$,

$$\begin{aligned} E[\boldsymbol{\delta}_{i,t}] &= m_{\delta,i,0} \\ Var(\boldsymbol{\delta}_{i,t}) &= w_i V_{\delta,i,0} + t w_i = (V_{\delta,i,0} + t) w_i \\ Cov(\boldsymbol{\delta}_{i,t}, \boldsymbol{\delta}_{i,s}) &= (V_{\delta,i,0} + min(t,s)) w_i. \end{aligned} \tag{A.10}$$

As a consequence, $\boldsymbol{\delta}_i = (\boldsymbol{\delta}_{i,1}, \boldsymbol{\delta}_{i,2}, ..., \boldsymbol{\delta}_{i,T})' \sim \mathrm{N}(\boldsymbol{m}_{\delta,i}, \boldsymbol{V}_{\delta,i})$. In our implementation, we place a loose prior on $\boldsymbol{\delta}_i$ and set $\boldsymbol{m}_{\delta,i} = 0$ and $\boldsymbol{V}_{\delta,i} = 10$ for all $i = 1, 2, ..., n$.

Algorithm 3, described below, draws $\boldsymbol{\delta}_i$ from its conditional posterior distribution

$$p(\boldsymbol{\delta}_i | \boldsymbol{\delta}_{1:(i-1)}, \boldsymbol{\delta}_{(i+1):n}, \mathcal{B}_T, \mathcal{P}_T, \boldsymbol{A}^{-1}, d, k, \boldsymbol{Q}, \boldsymbol{W}, \mathcal{D}_T). \tag{A.11}$$

Notice that the algorithm works with the vector of the demeaned log volatility,

$$\widetilde{\boldsymbol{\delta}}_{\boldsymbol{i}} = \boldsymbol{\delta}_i - \boldsymbol{m}_{\delta,i}. \tag{A.12}$$

We recover $\boldsymbol{\delta}_i$ by adding the prior mean back to the demeaned log volatility, $\boldsymbol{\delta}_i = \widetilde{\boldsymbol{\delta}}_i + \boldsymbol{m}_{\delta,i}$. To simplify the notation, we define the following density,

$$p(\boldsymbol{u}_{1:T} | \widetilde{\boldsymbol{\delta}}_i', Others) \triangleq p(\boldsymbol{u}_1, ..., \boldsymbol{u}_T | \boldsymbol{\delta}^{(c)}_{1:(i-1)}, \boldsymbol{\delta}_i', \boldsymbol{\delta}^{(c)}_{(i+1):n}, \mathcal{P}_T), \tag{A.13}$$

with an understanding that $\boldsymbol{\delta}_i^{(c)}$ is the current state value and $\boldsymbol{\delta}_i'$ is the proposed state value.

Algorithm 3. ***Elliptical slice sampler for*** $\widetilde{\boldsymbol{\delta}}_i$. *Enter the following steps with the current state value,* $\widetilde{\boldsymbol{\delta}}_i^{(c)}$, *and*

1. *Generate* $\mathbf{v} \sim \mathrm{N}(0, \boldsymbol{V}_{\delta,i})$ *and* $u \sim \mathrm{U}[0,1]$.
2. *Generate* $\theta \sim \mathrm{U}[0, 2\pi]$. *Let* $[\theta_{min}, \theta_{max}] = [\theta - 2\pi, \theta]$.
   (a) *(Proposal)* $\widetilde{\boldsymbol{\delta}}_i' = \widetilde{\boldsymbol{\delta}}_i^{(c)} \cos(\theta) + \mathbf{v} \sin(\theta)$
   (b) *(Accept/Reject) If* $p(\boldsymbol{u}_{1:T} | \widetilde{\boldsymbol{\delta}}_i', Others)/p(\boldsymbol{u}_{1:T} | \widetilde{\boldsymbol{\delta}}_i^{(c)}, Others) > u$, *exit (i.e., go to step 3). Otherwise, move on to (c).*
   (c) *(Adaptation) If* $\theta < 0$, *then* $\theta_{min} = \theta$. *Otherwise,* $\theta_{max} = \theta$.
   (d) *Update* $\theta \sim \mathrm{U}[\theta_{min}, \theta_{max}]$, *and go to (a).*
3. *Update* $\widetilde{\boldsymbol{\delta}}_i^{(c)} = \widetilde{\boldsymbol{\delta}}_i'$, *and* $\boldsymbol{\delta}_i^{(c)} = \widetilde{\boldsymbol{\delta}}_i^{(c)} + \boldsymbol{m}_{\delta,i}$

We complete Step 7 by iterating this algorithm for all $\boldsymbol{\delta}_i$, $i = 1, 2, ..., n$.

A.3 RMSEs and LPSs of the CSP-SV

Tables A.2 to A.7 describe the RMSE and LPS for all possible orderings of the CSP-SV. Each table has four sections determined by the name of the variables. In each section, the first column shows the rank of the corresponding model specification based on either RMSE or LPS. The second column describes the ordering of the variables (from first to last) in the CSP-SV. For example, for the first row of Table A.2, the second column indicates that the unemployment rate is ordered first, the 3-Month T-Bill second, inflation third, and output growth fourth. The third column describes the RMSE. The fourth column presents the p-value of Diebold-Mariano (or, Amisano-Giacomini) test for equal predictive ability (two-sided) between the best ordering and the corresponding model.

A.4 Predictive Densities four- and eight-quarter-ahead

Figures A.1 and A.2 present the mean and the standard deviation of the four-quarter-ahead and eight-quarter-ahead predictive densities over the forecasting sample for the best and worst ordering, respectively. For each panel, we pick the ex-post best and worst predictive densities based on the sum of the log predictive score of the corresponding individual variable.

A.5 Deeper Dive into LPS four- and eight-quarter-ahead

Figures A.3 and A.4 show the time-varying rankings of four-quarter-ahead and eight-quarter-ahead log predictive scores. Figures A.5 and A.6 show the time-varying average length of the 70% prediction intervals for four-quarter-ahead and eight-quarter-ahead log predictive scores.

Table A.2: RMSE Ranking, $h = 1$

Output Growth				Inflation				3-Month T-Bill				Unemployment			
R	Order	RMSE	pval	Rank	Order	RMSE	pval	R	Order	RMSE	pval	R	Order	RMSE	pval
1	u, i, π, y	2.57	NaN	1	y, π, i, u	0.604	NaN	1	y, u, π, i	0.345	NaN	1	y, i, π, u	0.204	NaN
2	u, π, i, y	2.565	0.36	2	i, y, u, π	0.604	0.56	2	u, y, i, π	0.341	0.04	2	π, i, y, u	0.203	0.75
3	i, u, y, π	2.565	0.65	3	π, y, u, i	0.603	0.43	3	y, π, u, i	0.337	0.05	3	i, y, u, π	0.203	0.56
4	i, π, u, y	2.545	0.26	4	π, i, u, y	0.603	0.58	4	π, y, u, i	0.336	0.13	4	i, π, u, y	0.202	0.34
5	i, u, π, y	2.542	0.31	5	u, π, i, y	0.602	0.25	5	i, π, y, u	0.336	0.08	5	y, π, i, u	0.202	0.33
6	π, u, y, i	2.539	0.32	6	π, i, y, u	0.602	0.74	6	π, i, u, y	0.335	0.06	6	i, u, y, π	0.202	0.35
7	i, y, u, π	2.533	0.29	7	y, i, π, u	0.602	0.13	7	π, u, i, y	0.335	0.07	7	π, y, u, i	0.202	0.23
8	y, i, π, u	2.527	0.38	8	i, y, π, u	0.601	0.15	8	π, u, y, i	0.335	0.08	8	π, u, y, i	0.201	0.07
9	u, π, y, i	2.526	0.17	9	i, π, y, u	0.601	0.09	9	π, y, i, u	0.335	0.06	9	u, i, π, y	0.201	0.14
10	π, u, i, y	2.523	0.14	10	u, y, i, π	0.601	0.19	10	u, π, y, i	0.335	0.03	10	i, y, π, u	0.2	0.37
11	π, i, u, y	2.513	0.14	11	i, u, y, π	0.601	0.12	11	y, u, i, π	0.333	0.07	11	i, u, π, y	0.2	0.23
12	u, i, y, π	2.511	0.12	12	π, u, y, i	0.601	0.16	12	y, π, i, u	0.333	0.09	12	u, π, i, y	0.2	0.17
13	y, π, i, u	2.507	0.21	13	u, i, π, y	0.601	0.07	13	π, i, y, u	0.333	0.08	13	y, π, u, i	0.2	0.19
14	u, y, π, i	2.507	0.13	14	π, y, i, u	0.601	0.21	14	u, i, y, π	0.333	0.04	14	π, u, i, y	0.199	0.09
15	π, y, u, i	2.506	0.17	15	u, π, y, i	0.601	0.16	15	y, i, π, u	0.333	0.09	15	y, i, u, π	0.199	0.24
16	i, π, y, u	2.498	0.07	16	u, i, y, π	0.6	0.08	16	y, i, u, π	0.332	0.07	16	π, y, i, u	0.199	0.25
17	u, y, i, π	2.489	0.09	17	i, u, π, y	0.6	0.16	17	u, π, i, y	0.332	0.05	17	y, u, π, i	0.199	0.21
18	i, y, π, u	2.489	0.07	18	π, u, i, y	0.599	0.08	18	u, y, π, i	0.332	0.03	18	π, i, u, y	0.199	0.25
19	y, i, u, π	2.478	0.09	19	y, π, u, i	0.599	0.07	19	u, i, π, y	0.331	0.04	19	u, i, y, π	0.199	0.08
20	y, π, u, i	2.473	0.09	20	y, i, u, π	0.599	0.18	20	i, π, u, y	0.329	0.05	20	i, π, y, u	0.199	0.44
21	y, u, i, π	2.472	0.10	21	y, u, π, i	0.598	0.12	21	i, u, π, y	0.329	0.04	21	u, π, y, i	0.198	0.09
22	π, y, i, u	2.472	0.09	22	u, y, π, i	0.597	0.23	22	i, y, u, π	0.329	0.08	22	u, y, i, π	0.198	0.08
23	y, u, π, i	2.469	0.08	23	i, π, u, y	0.596	0.49	23	i, y, π, u	0.328	0.05	23	y, u, i, π	0.197	0.21
24	π, i, y, u	2.464	0.10	24	y, u, i, π	0.596	0.23	24	i, u, y, π	0.327	0.06	24	u, y, π, i	0.197	0.10

Note. The table reports the one-quarter-ahead RMSE for the four variables included in the CSP-SV: output growth, inflation, the 3-Month T-Bill, and unemployment. For each variable, the column labeled R denotes the ranking and the column labeled Order denotes the variable order specification associated with each of the 24 possible orderings. The column labeled RMSE shows the RMSE error and the column labeled pval presents the p-value of the Diebold-Mariano test for equal predictive ability (two-sided) between the best ordering and each of the remaining orderings.

Table A.3: RMSE Ranking, $h = 4$

Output Growth				Inflation				3-Month T-Bill				Unemployment			
R	Order	RMSE	pval	Rank	Order	RMSE	pval	R	Order	RMSE	pval	R	Order	RMSE	pval
1	u , π , y , i	2.731	NaN	1	π , i , y , u	0.796	NaN	1	y , u , π , i	1.182	NaN	1	i , y , π , u	0.827	NaN
2	i , u , π , y	2.729	0.90	2	y , π , i , u	0.79	0.45	2	π , y , i , u	1.174	0.44	2	y , i , π , u	0.826	0.97
3	u , i , y , π	2.727	0.29	3	y , i , π , u	0.789	0.36	3	u , y , π , i	1.171	0.54	3	i , y , u , π	0.826	0.92
4	i , π , u , y	2.727	0.76	4	i , y , u , π	0.788	0.35	4	u , y , i , π	1.17	0.18	4	y , π , i , u	0.825	0.93
5	π , i , u , y	2.727	0.82	5	i , π , u , y	0.788	0.55	5	π , u , y , i	1.168	0.49	5	π , i , y , u	0.823	0.62
6	π , u , i , y	2.726	0.46	6	π , i , u , y	0.785	0.25	6	y , π , u , i	1.165	0.19	6	π , y , u , i	0.82	0.64
7	π , u , y , i	2.726	0.84	7	π , y , u , i	0.785	0.29	7	π , i , u , y	1.164	0.34	7	y , π , u , i	0.82	0.55
8	i , u , y , π	2.717	0.80	8	i , π , y , u	0.785	0.22	8	y , u , i , π	1.163	0.29	8	i , π , y , u	0.82	0.59
9	u , y , i , π	2.711	0.02	9	u , i , y , π	0.783	0.20	9	y , i , π , u	1.162	0.22	9	i , u , y , π	0.82	0.44
10	i , π , y , u	2.71	0.62	10	i , y , π , u	0.783	0.22	10	π , i , y , u	1.162	0.29	10	π , y , i , u	0.819	0.22
11	u , i , π , y	2.708	0.69	11	u , π , y , i	0.78	0.18	11	i , π , y , u	1.161	0.24	11	i , π , u , y	0.819	0.47
12	i , y , π , u	2.705	0.49	12	i , u , π , y	0.779	0.15	12	y , π , i , u	1.161	0.17	12	y , i , u , π	0.817	0.09
13	u , y , π , i	2.704	0.13	13	u , y , π , i	0.779	0.29	13	y , i , u , π	1.161	0.15	13	π , u , y , i	0.817	0.60
14	y , π , i , u	2.696	0.11	14	π , u , i , y	0.779	0.17	14	π , y , u , i	1.159	0.21	14	i , u , π , y	0.816	0.20
15	π , y , u , i	2.695	0.12	15	π , y , i , u	0.779	0.21	15	u , π , y , i	1.158	0.19	15	y , u , π , i	0.815	0.11
16	u , π , i , y	2.69	0.60	16	y , i , u , π	0.778	0.27	16	π , u , i , y	1.157	0.20	16	π , i , u , y	0.814	0.11
17	i , y , u , π	2.69	0.34	17	y , π , u , i	0.777	0.15	17	u , π , i , y	1.155	0.25	17	π , u , i , y	0.814	0.43
18	y , π , u , i	2.689	0.08	18	u , y , i , π	0.777	0.23	18	i , y , u , π	1.151	0.12	18	u , i , π , y	0.813	0.29
19	π , y , i , u	2.689	0.11	19	y , u , i , π	0.776	0.24	19	i , y , π , u	1.149	0.03	19	u , π , y , i	0.812	0.34
20	π , i , y , u	2.687	0.23	20	y , u , π , i	0.774	0.20	20	i , u , y , π	1.144	0.12	20	u , i , y , π	0.811	0.33
21	y , i , u , π	2.686	0.15	21	u , π , i , y	0.774	0.05	21	u , i , π , y	1.143	0.12	21	u , y , i , π	0.811	0.25
22	y , i , π , u	2.681	0.15	22	π , u , y , i	0.77	0.06	22	i , π , u , y	1.141	0.06	22	u , π , i , y	0.808	0.18
23	y , u , i , π	2.676	0.16	23	u , i , π , y	0.769	0.04	23	u , i , y , π	1.138	0.06	23	y , u , i , π	0.806	0.04
24	y , u , π , i	2.666	0.07	24	i , u , y , π	0.768	0.02	24	i , u , π , y	1.136	0.03	24	u , y , π , i	0.8	0.07

NOTE. The table reports the four-quarter-ahead RMSE for the four variables included in the CSP-SV: output growth, inflation, the 3-Month T-Bill, and unemployment. For each variable, the column labeled R denotes the ranking and the column labeled Order denotes the variable order specification associated with each of the 24 possible orderings. The column labeled RMSE shows the RMSE error and the column labeled pval presents the p-value of the Diebold-Mariano test for equal predictive ability (two-sided) between the best ordering and each of the remaining orderings.

Table A.4: RMSE Ranking, $h = 8$

Output Growth				Inflation				3-Month T-Bill				Unemployment			
R	Order	RMSE	pval	Rank	Order	RMSE	pval	R	Order	RMSE	pval	R	Order	RMSE	pval
1	i, π, u, y	2.607	NaN	1	i, π, u, y	0.915	NaN	1	u, y, π, i	2.01	NaN	1	i, y, π, u	1.47	NaN
2	π, i, y, u	2.607	0.98	2	π, i, y, u	0.913	0.88	2	π, u, y, i	1.991	0.34	2	i, π, y, u	1.468	0.91
3	u, i, y, π	2.607	0.96	3	u, y, π, i	0.901	0.36	3	π, y, i, u	1.972	0.20	3	π, i, u, y	1.455	0.35
4	π, u, i, y	2.607	0.94	4	u, i, y, π	0.896	0.08	4	u, y, i, π	1.972	0.34	4	i, u, π, y	1.454	0.39
5	π, i, u, y	2.606	0.91	5	i, y, u, π	0.894	0.15	5	u, π, y, i	1.967	0.11	5	π, y, i, u	1.454	0.22
6	π, u, y, i	2.605	0.93	6	i, u, π, y	0.889	0.07	6	y, u, i, π	1.967	0.13	6	i, y, u, π	1.45	0.38
7	u, π, y, i	2.604	0.75	7	i, y, π, u	0.889	0.27	7	y, i, π, u	1.961	0.15	7	y, π, i, u	1.449	0.37
8	i, u, π, y	2.603	0.67	8	y, u, i, π	0.887	0.25	8	y, u, π, i	1.959	0.27	8	π, i, y, u	1.448	0.17
9	u, y, i, π	2.595	0.49	9	y, π, i, u	0.887	0.05	9	π, u, i, y	1.959	0.07	9	i, u, y, π	1.448	0.32
10	u, y, π, i	2.592	0.14	10	y, i, π, u	0.887	0.05	10	y, π, i, u	1.957	0.12	10	y, π, u, i	1.447	0.22
11	i, π, y, u	2.591	0.51	11	π, i, u, y	0.886	0.11	11	y, i, u, π	1.957	0.10	11	y, i, u, π	1.446	0.07
12	π, y, u, i	2.591	0.60	12	u, π, i, y	0.885	0.10	12	y, π, u, i	1.953	0.16	12	u, y, i, π	1.444	0.33
13	i, y, π, u	2.589	0.40	13	u, π, y, i	0.884	0.00	13	π, i, u, y	1.952	0.21	13	π, u, i, y	1.443	0.41
14	u, i, π, y	2.583	0.53	14	π, u, i, y	0.883	0.00	14	u, π, i, y	1.946	0.14	14	u, π, y, i	1.443	0.44
15	y, π, i, u	2.577	0.36	15	y, i, u, π	0.882	0.20	15	π, i, y, u	1.944	0.19	15	y, i, π, u	1.443	0.32
16	y, i, π, u	2.576	0.44	16	π, y, u, i	0.881	0.00	16	i, y, u, π	1.944	0.17	16	u, i, y, π	1.442	0.42
17	y, π, u, i	2.574	0.18	17	i, π, y, u	0.88	0.18	17	i, y, π, u	1.942	0.12	17	π, y, u, i	1.442	0.27
18	π, y, i, u	2.572	0.16	18	π, y, i, u	0.878	0.12	18	i, u, y, π	1.941	0.22	18	i, π, u, y	1.442	0.19
19	i, u, y, π	2.572	0.47	19	u, y, i, π	0.877	0.02	19	π, y, u, i	1.94	0.11	19	π, u, y, i	1.441	0.44
20	i, y, u, π	2.571	0.46	20	u, i, π, y	0.877	0.00	20	u, i, π, y	1.94	0.12	20	y, u, π, i	1.435	0.01
21	y, u, i, π	2.568	0.24	21	y, u, π, i	0.871	0.05	21	u, i, y, π	1.937	0.06	21	u, i, π, y	1.43	0.19
22	y, i, u, π	2.568	0.18	22	y, π, u, i	0.869	0.03	22	i, π, u, y	1.937	0.07	22	y, u, i, π	1.429	0.04
23	u, π, i, y	2.563	0.54	23	i, u, y, π	0.868	0.00	23	i, u, π, y	1.933	0.03	23	u, y, π, i	1.425	0.17
24	y, u, π, i	2.558	0.17	24	π, u, y, i	0.86	0.00	24	i, π, y, u	1.924	0.20	24	u, π, i, y	1.418	0.16

Note. The table reports the eight-quarter-ahead RMSE for the four variables included in the CSP-SV: output growth, inflation, the 3-Month T-Bill, and unemployment. For each variable, the column labeled R denotes the ranking and the column labeled Order denotes the variable order specification associated with each of the 24 possible orderings. The column labeled RMSE shows the RMSE error and the column labeled pval presents the p-value of the Diebold-Mariano test for equal predictive ability (two-sided) between the best ordering and each of the remaining orderings.

Table A.5: LPS Ranking, $h = 1$

Joint				Output Growth				Inflation				3-Month T-Bill				Unemployment			
R	Order	LPS	pval	R	Order	LPS	pval	R	Order	LPS	pval	R	Order	LPS	pval	R	Order	LPS	pval
1	i , y , π , u	-346.23	NaN	1	y , u , π , i	-274.29	NaN	1	π , u , i , y	-106.76	NaN	1	i , u , y , π	-15.74	NaN	1	u , y , i , π	33.53	NaN
2	y , i , u , π	-347.84	0.63	2	y , π , u , i	-274.39	0.92	2	y , u , π , i	-107.07	0.70	2	i , π , u , y	-15.98	0.67	2	u , i , y , π	33.14	0.65
3	i , u , π , y	-348.54	0.39	3	π , i , y , u	-274.48	0.91	3	π , u , y , i	-107.2	0.48	3	i , u , π , y	-16	0.33	3	π , u , i , y	33.06	0.56
4	i , π , y , u	-348.99	0.21	4	π , y , i , u	-274.79	0.57	4	π , y , u , i	-107.32	0.45	4	i , y , π , u	-16.35	0.35	4	u , π , y , i	32.98	0.45
5	π , y , i , u	-349.19	0.54	5	y , u , i , π	-274.9	0.49	5	u , π , y , i	-107.36	0.09	5	i , y , u , π	-16.76	0.29	5	u , y , π , i	32.89	0.58
6	y , π , u , i	-349.9	0.65	6	y , i , u , π	-275.55	0.15	6	u , y , i , π	-107.52	0.28	6	i , π , y , u	-17.77	0.05	6	u , π , i , y	32.72	0.53
7	π , i , u , y	-351.59	0.09	7	i , y , π , u	-275.81	0.46	7	y , π , u , i	-107.56	0.36	7	π , i , y , u	-17.9	0.17	7	π , u , y , i	32.54	0.29
8	y , π , i , u	-352.42	0.25	8	i , π , y , u	-275.83	0.50	8	u , π , i , y	-107.7	0.09	8	y , i , π , u	-19.37	0.29	8	i , u , π , y	32.46	0.54
9	u , i , y , π	-352.56	0.38	9	π , y , u , i	-276.39	0.05	9	u , i , y , π	-108.28	0.06	9	y , i , u , π	-19.42	0.26	9	y , u , i , π	32.34	0.58
10	π , u , i , y	-352.96	0.30	10	u , y , i , π	-276.39	0.33	10	i , π , y , u	-108.4	0.16	10	y , π , i , u	-19.89	0.27	10	y , π , u , i	32.26	0.55
11	y , u , i , π	-353.17	0.23	11	π , u , y , i	-276.45	0.21	11	π , i , u , y	-108.53	0.03	11	π , i , u , y	-20.42	0.10	11	y , u , π , i	32.25	0.55
12	π , y , u , i	-354.18	0.31	12	u , π , i , y	-276.54	0.45	12	i , u , π , y	-108.65	0.03	12	y , u , i , π	-20.71	0.23	12	u , i , π , y	32	0.10
13	u , π , y , i	-354.44	0.27	13	y , π , i , u	-276.58	0.04	13	u , i , π , y	-108.7	0.02	13	u , i , π , y	-21.73	0.28	13	π , y , i , u	31.67	0.52
14	i , y , u , π	-354.53	0.02	14	i , y , u , π	-277.27	0.07	14	y , π , i , u	-108.82	0.05	14	π , y , i , u	-22.08	0.18	14	y , i , u , π	31.54	0.44
15	y , i , π , u	-354.78	0.12	15	y , i , π , u	-277.27	0.03	15	π , y , i , u	-108.86	0.07	15	u , y , π , i	-22.81	0.22	15	i , u , y , π	31.42	0.29
16	i , u , y , π	-354.96	0.05	16	u , y , π , i	-277.39	0.14	16	y , i , π , u	-108.89	0.04	16	u , i , y , π	-23.08	0.22	16	i , y , π , u	31.38	0.54
17	i , π , u , y	-356.14	0.00	17	u , i , y , π	-277.54	0.17	17	i , π , u , y	-108.94	0.04	17	π , y , u , i	-24.23	0.16	17	π , y , u , i	31.26	0.36
18	u , y , π , i	-358.21	0.10	18	π , u , i , y	-277.67	0.28	18	i , y , π , u	-109.17	0.06	18	u , π , y , i	-25.46	0.17	18	π , i , u , y	30.77	0.35
19	y , u , π , i	-359.21	0.16	19	i , u , y , π	-277.71	0.26	19	y , i , u , π	-109.79	0.05	19	π , u , i , y	-25.46	0.15	19	y , π , i , u	30.6	0.31
20	u , y , i , π	-359.74	0.11	20	π , i , u , y	-278.15	0.24	20	π , i , y , u	-110.12	0.02	20	y , π , u , i	-25.69	0.16	20	i , π , u , y	29.86	0.33
21	u , i , π , y	-361.34	0.03	21	u , i , π , y	-278.17	0.20	21	i , y , u , π	-110.3	0.01	21	u , π , i , y	-28.82	0.14	21	y , i , π , u	29.55	0.23
22	π , i , y , u	-362.87	0.00	22	u , π , y , i	-278.33	0.10	22	i , u , y , π	-110.42	0.01	22	π , u , y , i	-30.27	0.13	22	i , y , u , π	29.4	0.24
23	π , u , y , i	-370.36	0.04	23	i , u , π , y	-279.21	0.16	23	y , u , i , π	-110.65	0.06	23	u , y , i , π	-31.6	0.07	23	i , π , y , u	28.67	0.28
24	u , π , i , y	-371.86	0.01	24	i , π , u , y	-279.41	0.17	24	u , y , π , i	-111.03	0.03	24	y , u , π , i	-33.1	0.06	24	π , i , y , u	26.94	0.28

NOTE. The table reports the one-quarter-ahead LPSs for the four variables included in the CSP-SV: output growth, inflation, the 3-Month T-Bill, and unemployment. For each variable, the column labeled R denotes the ranking and the column labeled Order denotes the variable order specification associated with each of the 24 possible orderings. The column labeled LPS shows the one-quarter-ahead log predictive score and the column labeled pval presents the p-value of the Amisano-Giacomini test for equal predictive ability (two-sided) between the best ordering and each of the remaining orderings.

Table A.6: LPS Ranking, $h = 4$

Joint				Output Growth				Inflation				3-Month T-Bill				Unemployment			
R	Order	LPS	pval	R	Order	LPS	pval	R	Order	LPS	pval	R	Order	LPS	pval	R	Order	LPS	pval
1	y , i , π , u	-716.06	NaN	1	y , u , i , π	-284.5	NaN	1	π , u , y , i	-142.56	NaN	1	i , u , π , y	-183.55	NaN	1	i , u , π , y	-123.95	NaN
2	π , y , u , i	-718.25	0.74	2	y , π , u , i	-285.17	0.40	2	u , π , i , y	-143.21	0.18	2	i , π , u , y	-184.08	0.58	2	u , y , π , i	-124.1	0.97
3	u , π , y , i	-718.6	0.64	3	π , y , u , i	-285.24	0.42	3	π , u , i , y	-143.26	0.21	3	i , y , u , π	-184.42	0.63	3	i , π , u , y	-126.17	0.59
4	u , i , y , π	-719.14	0.63	4	y , i , u , π	-285.27	0.30	4	y , u , π , i	-143.4	0.52	4	i , u , y , π	-184.96	0.35	4	i , u , y , π	-126.35	0.18
5	i , π , u , y	-720.69	0.23	5	y , u , π , i	-285.33	0.37	5	y , π , u , i	-143.96	0.22	5	i , y , π , u	-185.28	0.34	5	π , u , y , i	-126.61	0.57
6	i , y , u , π	-721.46	0.27	6	π , y , i , u	-285.39	0.01	6	u , π , y , i	-144.09	0.01	6	i , π , y , u	-185.51	0.37	6	u , π , y , i	-126.79	0.58
7	π , u , i , y	-723.81	0.27	7	y , i , π , u	-285.5	0.51	7	π , y , u , i	-144.12	0.11	7	π , i , y , u	-185.73	0.38	7	i , y , u , π	-126.83	0.43
8	y , π , i , u	-727.97	0.34	8	π , i , y , u	-285.53	0.36	8	i , π , y , u	-144.31	0.14	8	y , i , π , u	-186.76	0.43	8	u , i , y , π	-126.92	0.59
9	y , u , i , π	-728.37	0.05	9	y , π , i , u	-285.9	0.31	9	u , y , i , π	-144.35	0.04	9	u , i , y , π	-187.21	0.49	9	π , u , i , y	-127.21	0.48
10	y , u , π , i	-728.99	0.12	10	i , y , u , π	-286.85	0.24	10	π , y , i , u	-144.41	0.17	10	y , π , i , u	-187.4	0.35	10	π , i , u , y	-128.34	0.19
11	π , u , y , i	-729.18	0.10	11	u , y , i , π	-286.97	0.24	11	y , π , i , u	-144.6	0.03	11	y , i , u , π	-187.57	0.27	11	π , y , u , i	-128.85	0.24
12	π , i , u , y	-729.33	0.17	12	i , y , π , u	-287.08	0.09	12	π , i , u , y	-144.96	0.02	12	π , i , u , y	-187.64	0.06	12	u , i , π , y	-129.43	0.25
13	i , u , π , y	-730.14	0.42	13	π , u , y , i	-287.14	0.09	13	y , i , π , u	-145	0.02	13	u , i , π , y	-187.76	0.47	13	i , y , π , u	-129.64	0.09
14	π , i , y , u	-731.41	0.05	14	u , y , π , i	-287.21	0.17	14	i , π , u , y	-145.29	0.02	14	π , y , u , i	-188.15	0.46	14	u , π , i , y	-129.66	0.38
15	i , y , π , u	-731.54	0.27	15	i , π , y , u	-287.63	0.09	15	i , u , π , y	-145.37	0.00	15	π , y , i , u	-188.92	0.21	15	y , i , π , u	-129.83	0.24
16	i , π , y , u	-731.66	0.21	16	u , i , y , π	-287.88	0.10	16	i , y , π , u	-145.44	0.02	16	y , u , i , π	-189.72	0.16	16	y , π , i , u	-130.51	0.22
17	y , π , u , i	-733.17	0.24	17	u , π , y , i	-288.11	0.11	17	u , i , π , y	-145.77	0.00	17	π , u , i , y	-189.85	0.31	17	y , u , i , π	-131.34	0.25
18	u , y , π , i	-734.4	0.01	18	u , i , π , y	-288.52	0.15	18	u , i , y , π	-146.27	0.00	18	y , π , u , i	-190.06	0.33	18	π , y , i , u	-131.38	0.12
19	u , π , i , y	-735.06	0.06	19	π , u , i , y	-288.57	0.15	19	y , i , u , π	-146.5	0.01	19	u , π , y , i	-190.07	0.33	19	y , i , u , π	-131.65	0.17
20	u , y , i , π	-735.65	0.19	20	u , π , i , y	-288.62	0.14	20	i , u , y , π	-146.95	0.00	20	u , π , i , y	-190.99	0.41	20	π , i , y , u	-131.72	0.18
21	π , y , i , u	-736.19	0.24	21	π , i , u , y	-289.53	0.08	21	i , y , u , π	-147.16	0.00	21	u , y , π , i	-191.25	0.18	21	i , π , y , u	-132.91	0.10
22	y , i , u , π	-737.42	0.27	22	i , u , y , π	-289.62	0.04	22	π , i , y , u	-147.39	0.02	22	π , u , y , i	-191.75	0.41	22	y , u , π , i	-133.24	0.20
23	u , i , π , y	-739.64	0.11	23	i , u , π , y	-289.81	0.05	23	y , u , i , π	-148.29	0.00	23	u , y , i , π	-191.75	0.28	23	u , y , i , π	-135.32	0.41
24	i , u , y , π	-739.69	0.17	24	i , π , u , y	-290.48	0.10	24	u , y , π , i	-149.13	0.00	24	y , u , π , i	-192.47	0.27	24	y , π , u , i	-138.68	0.22

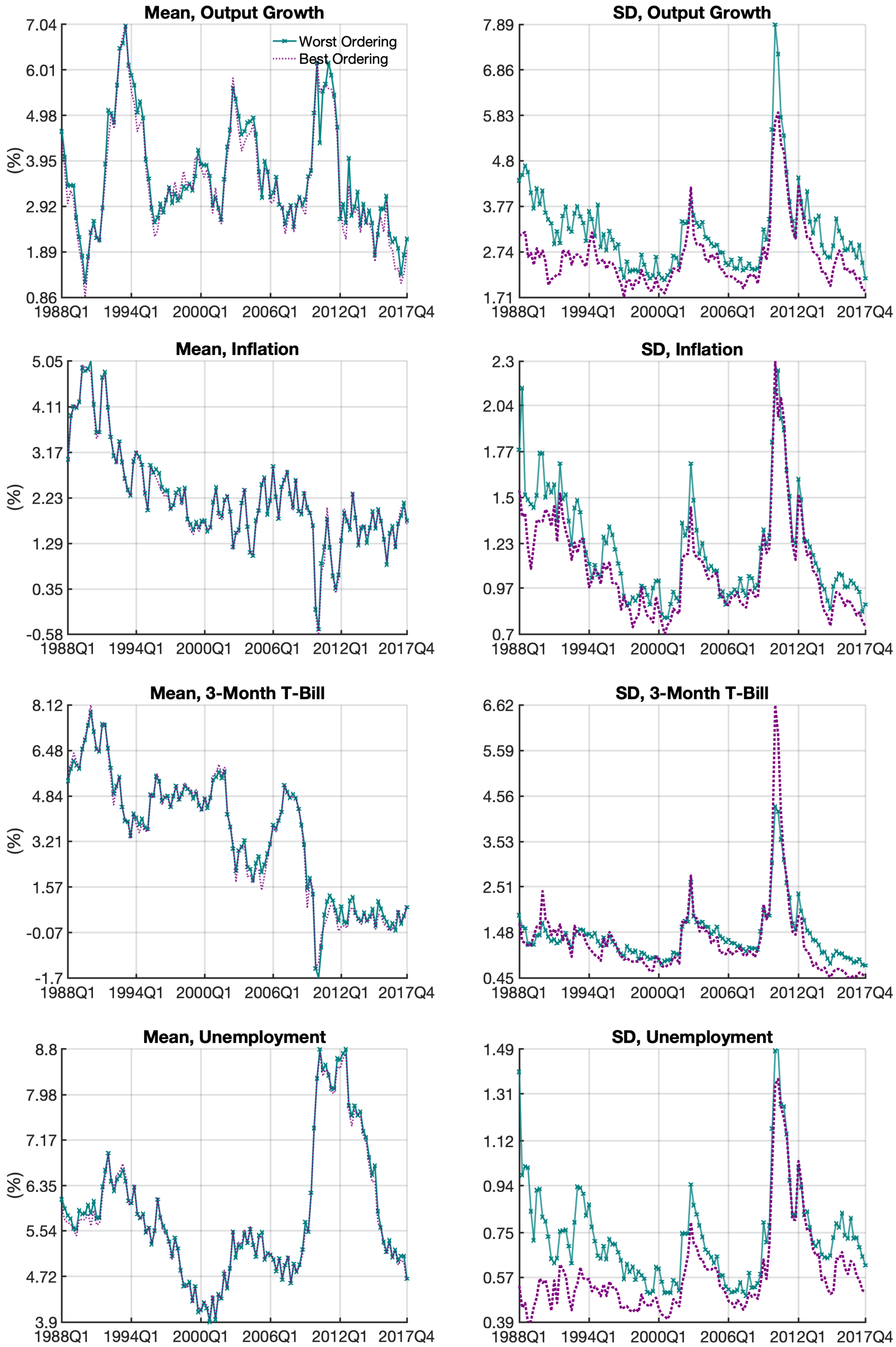
NOTE. The table reports the four-quarter-ahead LPS for the four variables included in the CSP-SV: output growth, inflation, the 3-Month T-Bill, and unemployment. For each variable, the column labeled R denotes the ranking and the column labeled Order denotes the variable order specification associated with each of the 24 possible orderings. The column labeled LPS shows the four-quarter-ahead log predictive score and the column labeled pval presents the p-value of the Amisano-Giacomini test for equal predictive ability (two-sided) between the best ordering and each of the remaining orderings.

Table A.7: LPS Ranking, $h = 8$

Joint				Output Growth				Inflation				3-Month T-Bill				Unemployment			
R	Order	LPS	pval	R	Order	LPS	pval	R	Order	LPS	pval	R	Order	LPS	pval	R	Order	LPS	pval
1	y, π, i, u	-910.97	NaN	1	y, u, π, i	-283.26	NaN	1	π, u, y, i	-169.8	NaN	1	π, y, u, i	-260.09	NaN	1	i, π, u, y	-205.3	NaN
2	y, i, π, u	-915.53	0.40	2	y, π, u, i	-283.31	0.90	2	u, π, i, y	-170.56	0.48	2	i, π, y, u	-260.2	0.99	2	i, y, u, π	-208	0.53
3	i, π, u, y	-917.45	0.51	3	y, u, i, π	-283.7	0.59	3	y, u, π, i	-171.71	0.24	3	i, u, π, y	-260.21	0.98	3	u, y, π, i	-208.18	0.74
4	u, π, y, i	-919.19	0.25	4	y, i, u, π	-284.17	0.31	4	y, π, u, i	-172.19	0.08	4	i, π, u, y	-260.26	0.98	4	i, u, π, y	-209.15	0.52
5	i, y, u, π	-919.72	0.21	5	π, y, i, u	-284.25	0.04	5	π, u, i, y	-172.23	0.03	5	i, y, u, π	-260.5	0.94	5	y, i, π, u	-209.54	0.56
6	π, u, i, y	-924.32	0.17	6	y, i, π, u	-285.15	0.27	6	u, π, y, i	-172.81	0.01	6	π, i, y, u	-260.83	0.91	6	π, i, u, y	-210.66	0.24
7	u, i, y, π	-924.9	0.15	7	y, π, i, u	-285.18	0.22	7	π, y, u, i	-172.93	0.02	7	y, π, i, u	-261.44	0.71	7	π, i, y, u	-210.77	0.24
8	i, y, π, u	-925.43	0.06	8	π, y, u, i	-285.24	0.13	8	π, y, i, u	-173.13	0.05	8	y, u, π, i	-261.5	0.54	8	u, i, y, π	-211.58	0.52
9	y, π, u, i	-926.97	0.21	9	u, y, i, π	-285.41	0.29	9	y, π, i, u	-173.66	0.01	9	u, i, y, π	-261.68	0.53	9	u, i, π, y	-211.59	0.39
10	π, y, u, i	-927.03	0.36	10	u, y, π, i	-286.65	0.10	10	u, y, i, π	-174.25	0.01	10	y, i, π, u	-261.76	0.69	10	π, u, i, y	-211.62	0.52
11	π, i, u, y	-928.01	0.06	11	u, π, y, i	-287	0.08	11	y, i, π, u	-174.27	0.01	11	i, u, y, π	-262.03	0.76	11	y, π, i, u	-211.8	0.46
12	π, i, y, u	-929.63	0.10	12	i, y, π, u	-287.03	0.01	12	i, π, y, u	-174.6	0.03	12	i, y, π, u	-262.09	0.77	12	y, u, i, π	-211.89	0.44
13	u, y, i, π	-932.7	0.14	13	π, u, y, i	-287.04	0.08	13	π, i, u, y	-174.9	0.01	13	y, π, u, i	-262.28	0.22	13	i, u, y, π	-212.12	0.26
14	i, π, y, u	-933.36	0.13	14	i, y, u, π	-287.15	0.09	14	u, i, π, y	-174.96	0.00	14	u, y, i, π	-262.3	0.38	14	π, y, u, i	-212.85	0.40
15	π, u, y, i	-934.38	0.06	15	u, i, y, π	-287.28	0.03	15	i, y, π, u	-175.24	0.01	15	π, u, i, y	-262.67	0.38	15	π, u, y, i	-213.41	0.48
16	π, y, i, u	-934.56	0.19	16	π, i, y, u	-287.8	0.01	16	y, i, u, π	-175.58	0.00	16	π, i, u, y	-262.75	0.56	16	u, π, y, i	-214.5	0.46
17	u, π, i, y	-935.52	0.07	17	π, u, i, y	-287.97	0.10	17	i, u, π, y	-175.63	0.00	17	u, π, y, i	-262.89	0.32	17	u, π, i, y	-214.56	0.39
18	i, u, π, y	-935.69	0.15	18	i, π, y, u	-288.25	0.03	18	u, i, y, π	-176.17	0.00	18	u, π, i, y	-263.04	0.44	18	i, y, π, u	-214.73	0.21
19	y, u, i, π	-937.86	0.03	19	u, i, π, y	-289.01	0.08	19	i, π, u, y	-177.01	0.00	19	π, y, i, u	-263.37	0.32	19	i, π, y, u	-215.12	0.13
20	y, i, u, π	-938	0.18	20	u, π, i, y	-289.99	0.06	20	i, u, y, π	-177.02	0.00	20	π, u, y, i	-263.52	0.39	20	π, y, i, u	-215.13	0.33
21	u, i, π, y	-940.86	0.04	21	i, u, y, π	-290.29	0.02	21	i, y, u, π	-177.99	0.00	21	y, i, u, π	-263.95	0.43	21	u, y, i, π	-216.36	0.45
22	i, u, y, π	-942.95	0.08	22	π, i, u, y	-290.41	0.04	22	π, i, y, u	-178.13	0.01	22	u, i, π, y	-264.15	0.36	22	y, i, u, π	-216.69	0.33
23	u, y, π, i	-945	0.02	23	i, u, π, y	-290.93	0.01	23	y, u, i, π	-178.17	0.00	23	y, u, i, π	-266.6	0.24	23	y, π, u, i	-219.41	0.32
24	y, u, π, i	-948.09	0.31	24	i, π, u, y	-292.8	0.04	24	u, y, π, i	-179.46	0.00	24	u, y, π, i	-267.13	0.23	24	y, u, π, i	-227.22	0.34

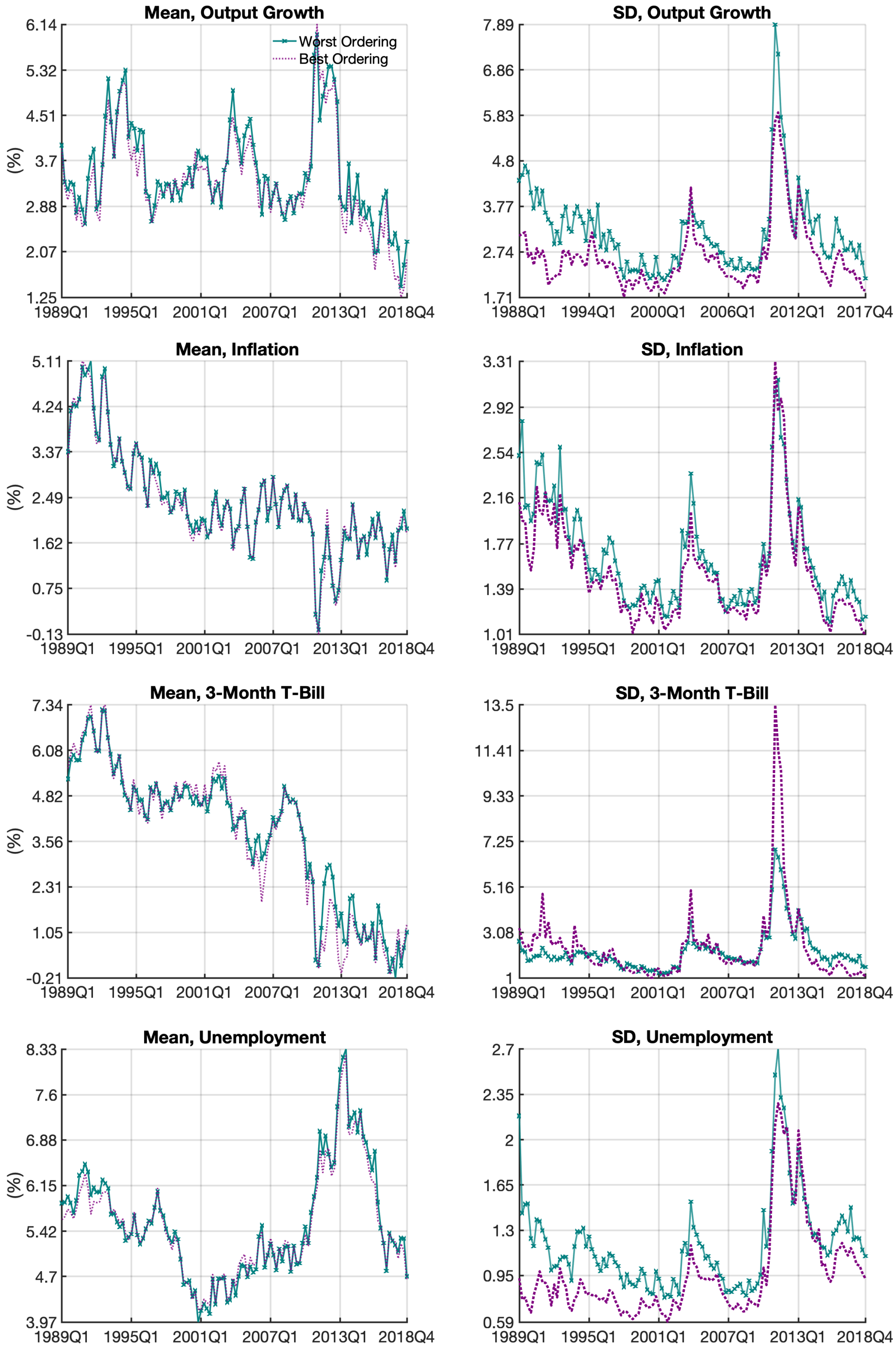
NOTE. The table reports the eight-quarter-ahead LPS for the four variables included in the CSP-SV: output growth, inflation, the 3-Month T-Bill, and unemployment. For each variable, the column labeled R denotes the ranking and the column labeled Order denotes the variable order specification associated with each of the 24 possible orderings. The column labeled LPS shows the eight-quarter-ahead log predictive score and the column labeled pval presents the p-value of the Amisano-Giacomini test for equal predictive ability (two-sided) between the best ordering and each of the remaining orderings.

Figure A.1: Four-Quarter-Ahead Predictive Density and Ordering



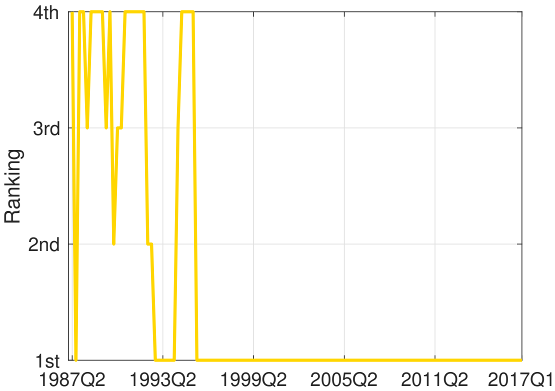
Note. Mean and standard deviation (SD) of the four-quarter-ahead predictive density throughout the forecasting sample.

Figure A.2: Eight-Quarter-Ahead Predictive Density and Ordering

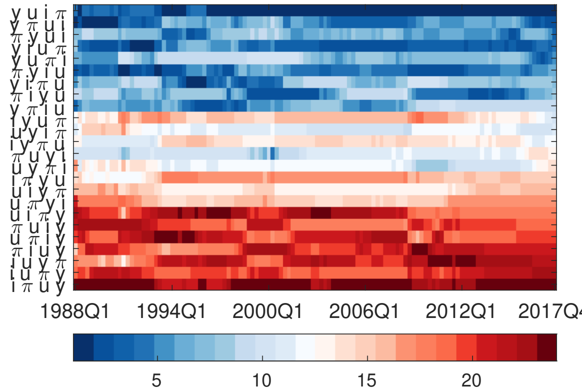


NOTE. Mean and standard deviation (SD) of the eight-quarter-ahead predictive density throughout the forecasting sample.

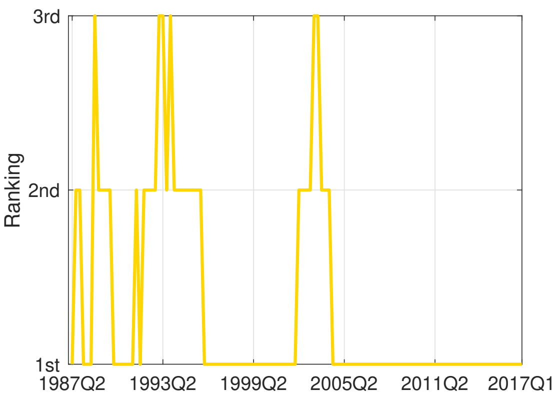
Figure A.3: Time-Varying Ranking of Four-Quarter-Ahead Log Predictive Scores



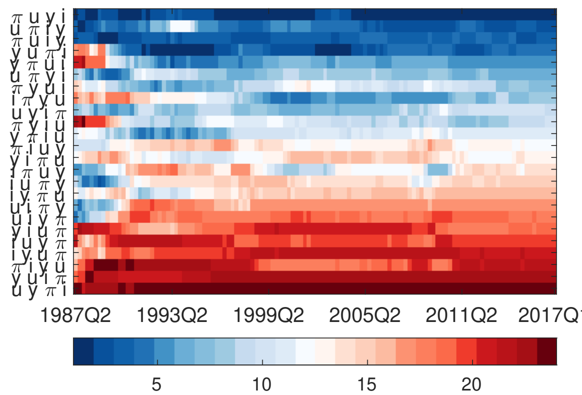
(a) Output Growth, Best Ordering



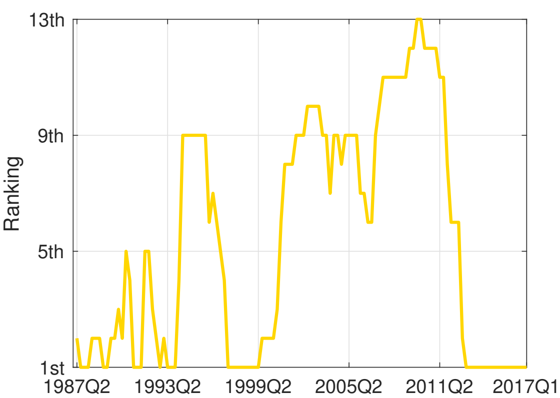
(b) Output Growth



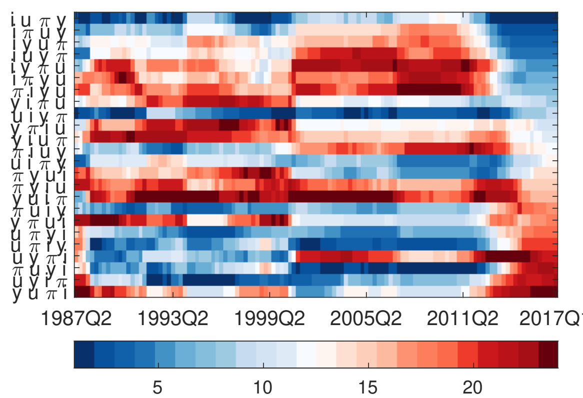
(c) Inflation, Best Ordering



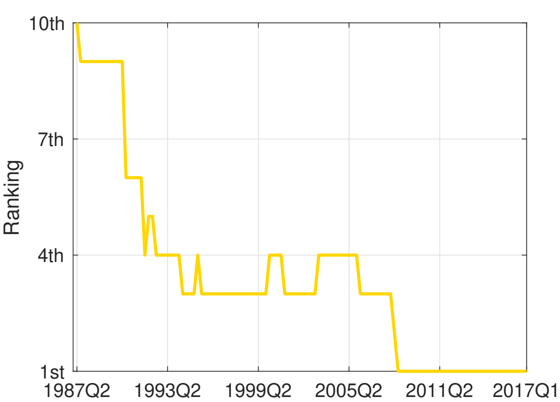
(d) Inflation



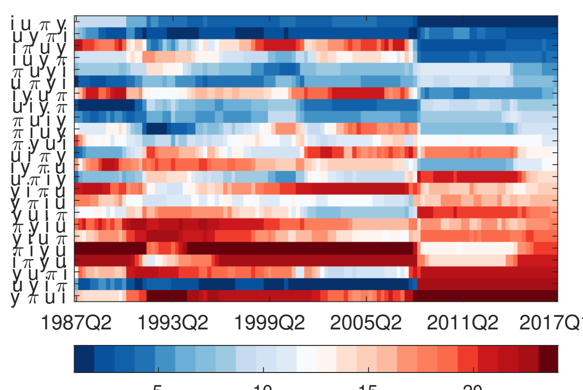
(e) 3-Month T-Bill, Best Ordering



(f) 3-Month T-Bill

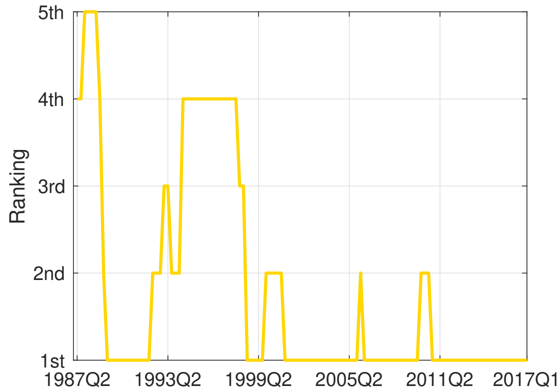


(g) Unemployment, Best Ordering

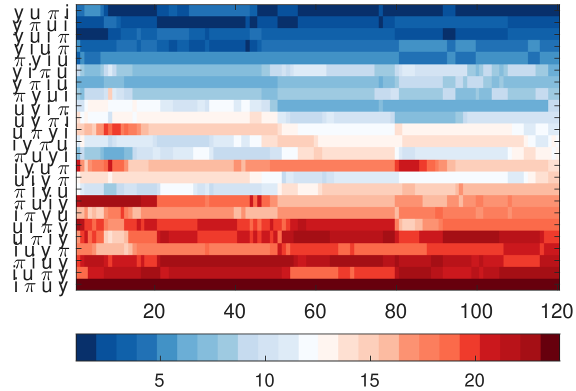


(h) Unemployment

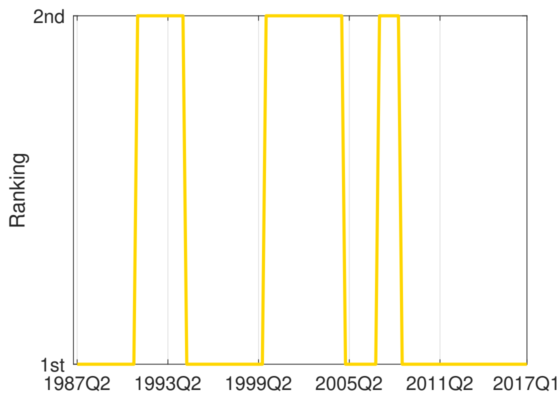
Figure A.4: Time-Varying Ranking of Eight-Quarter-Ahead Log Predictive Scores



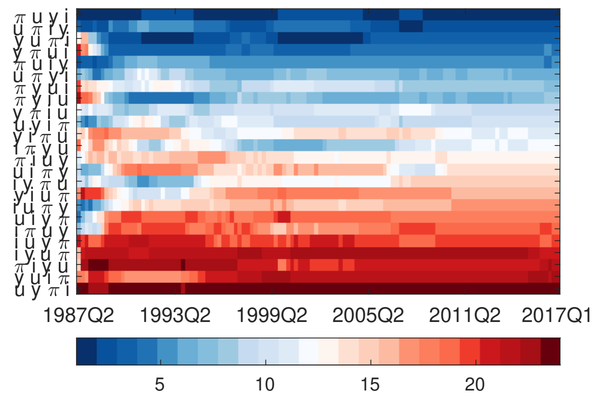
(a) Output Growth, Best Ordering



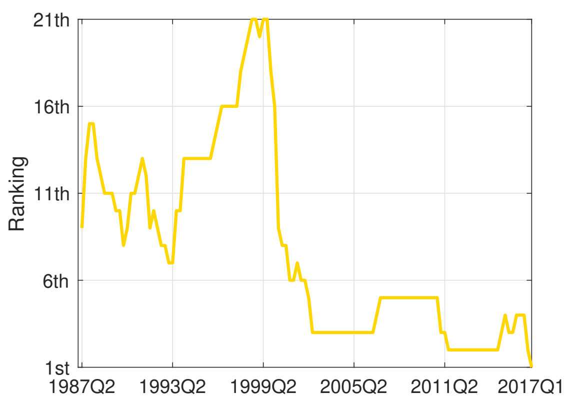
(b) Output Growth



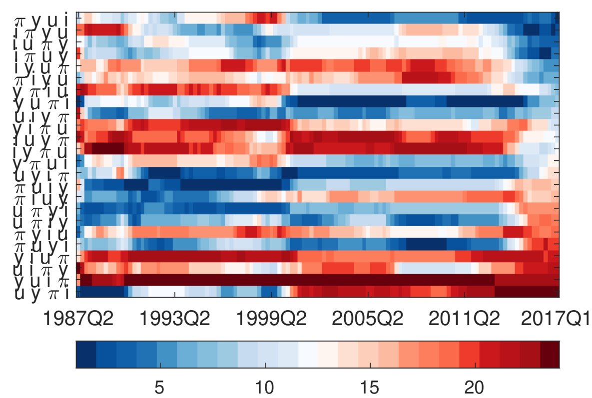
(c) Inflation, Best Ordering



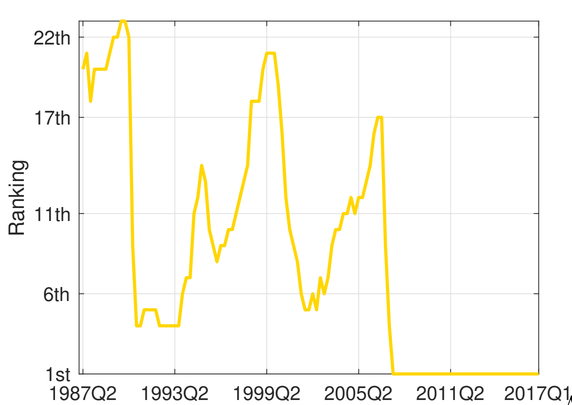
(d) Inflation



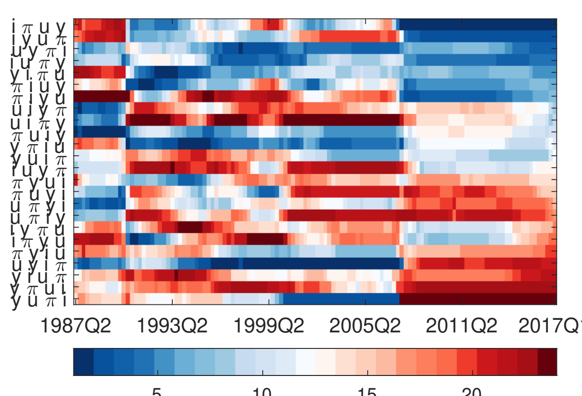
(e) 3-Month T-Bill, Best Ordering



(f) 3-Month T-Bill

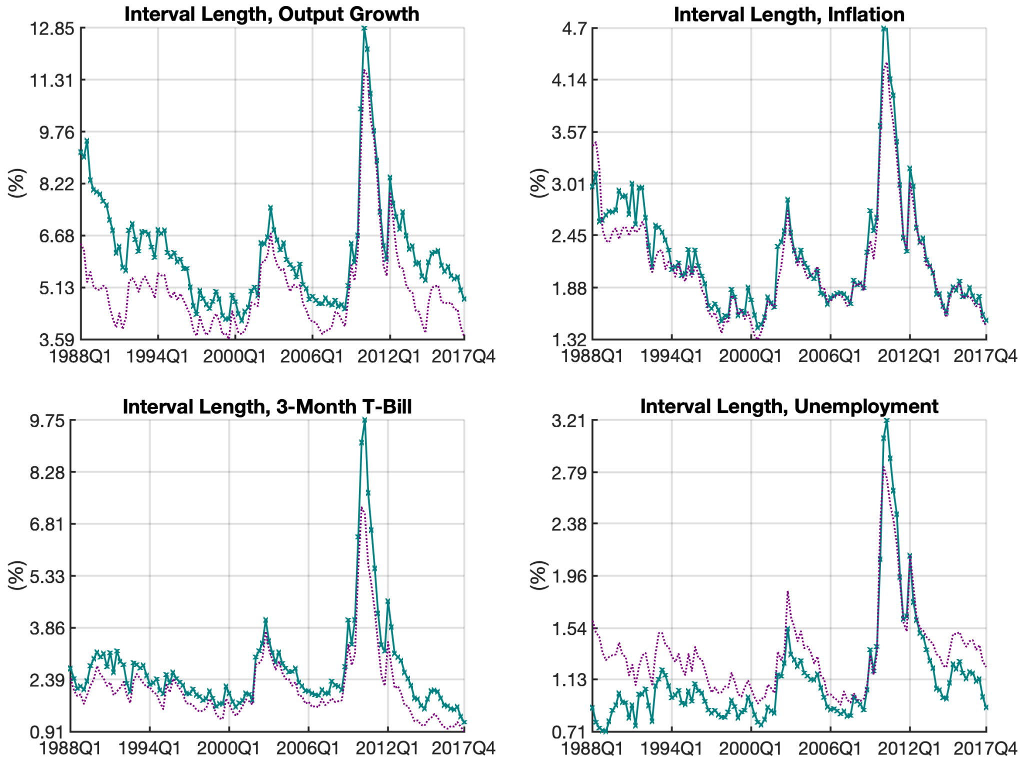


(g) Unemployment, Best Ordering



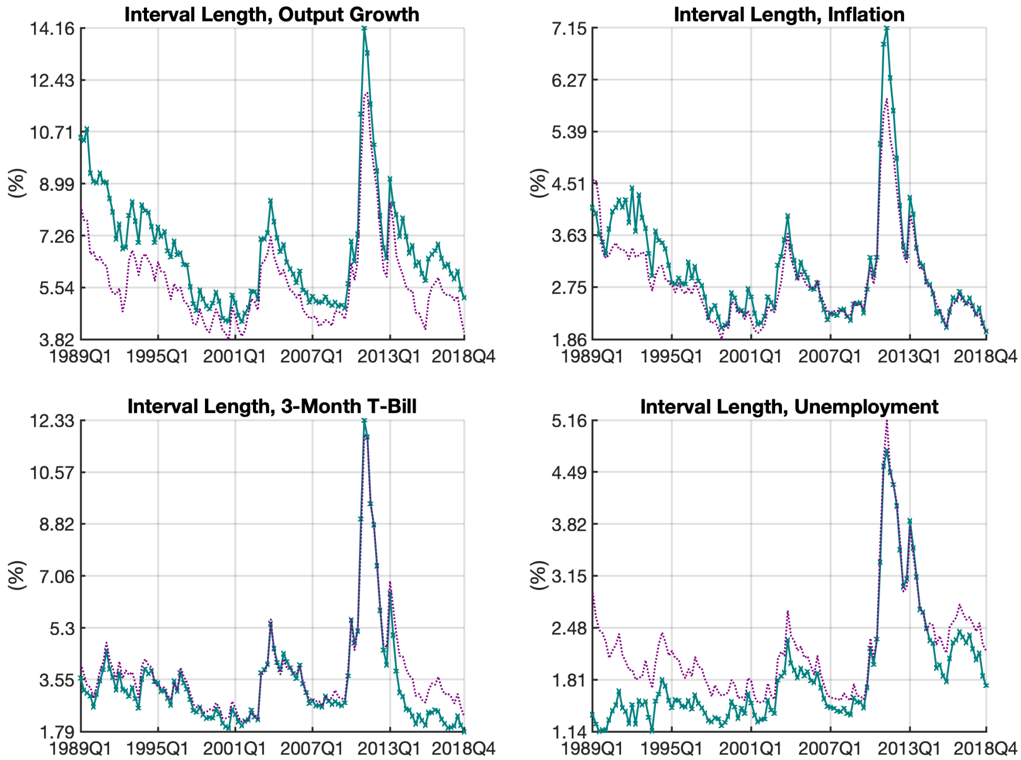
(h) Unemployment

Figure A.5: Four-Quarter-Ahead Prediction Interval and Ordering



Note. Length of the corresponding intervals. Intervals are computed based on the four-quarter-ahead predictive density throughout the evaluation sample. The difference between the empirical coverage and the nominal coverage is largest for the worst ordering and smallest for the best ordering.

Figure A.6: Eight-Quarter-Ahead Prediction Interval and Ordering



NOTE. Length of the corresponding intervals. Intervals are computed based on the four-quarter-ahead predictive density throughout the evaluation sample. The difference between the empirical coverage and the nominal coverage is largest for the worst ordering and smallest for the best ordering.