

Dimitrios P. Louzis*

Steady-state priors and Bayesian variable selection in VAR forecasting

DOI 10.1515/snde-2015-0048

Abstract: This study proposes methods for estimating Bayesian vector autoregressions (VARs) with a (semi-) automatic variable selection and an informative prior on the unconditional mean or steady-state of the system. We show that extant Gibbs sampling methods for Bayesian variable selection can be efficiently extended to incorporate prior beliefs on the steady-state of the economy. Empirical analysis, based on three major US macroeconomic time series, indicates that the out-of-sample forecasting accuracy of a VAR model is considerably improved when it combines both variable selection and steady-state prior information.

Keywords: Bayesian VAR; macroeconomic forecasting; steadystates; variable selection.

1 Introduction

The seminal studies of Sims (1980), Doan, Litterman, and Sims (1984) and Litterman (1986) kick-started a flurry of research on Bayesian vector autoregressions (Bayesian VARs or BVARs hereafter) and their ability to generate accurate macroeconomic forecasts. Almost 30 years later, BVARs have been established as a standard forecasting tool in empirical macroeconomics (e.g. see Karlsson 2013 for a recent review and references therein).

A key element in the plethora of BVAR specifications is the shrinkage of dynamic parameters towards a specific representation of the data which reflects researchers' prior beliefs and deals with the over-parameterization problem. A popular shrinkage method is the Minnesota (or Litterman) prior (Doan, Litterman, and Sims 1984; Litterman 1986) and its variants (Banbura, Giannone, and Reichlin 2010; Koop 2013; Giannone, Lenza, and Primiceri 2015) which allow for different levels of shrinkage on VAR coefficients and in some cases lead to conjugate posterior densities eliminating the need of posterior simulations. Recently, Bayesian least absolute shrinkage and selection operator (Lasso) has also been proposed for VAR shrinkage (Korobilis 2013; Gefang 2014).

Another strand of the literature has proposed Bayesian variable selection as an alternative way of VAR shrinkage. In general, variable selection techniques refer to a statistical procedure that decides stochastically which of the variables enter the VAR equation and which not, based on information provided by the data. Variable selection can be performed either by imposing a tight prior around zero on some of the VAR coefficients (George, Sun, and Ni 2008; Korobilis 2008; Koop 2013) or via restricting coefficients to be exactly zero (Korobilis 2013). In the former case, all variables enter the VAR equations but some of them have coefficients very close, but not exactly, zero; whereas, in the latter case, some of the variables are actually excluded from the VAR system leading to a *restricted* VAR specification. The method of George, Sun, and Ni (2008) is generally faster and computationally more stable than Korobilis (2013) method (see also George and McCulloch 1997). On the other hand, the method of Korobilis (2013) is considered more flexible since it is fully automatic and independent of the prior specified on the dynamic parameters, a feature that simplifies posterior simulations and enables the adoption of the method in non-linear VAR models.

*Corresponding author: Dimitrios P. Louzis, Bank of Greece – Economic Analysis and Research Department, Athens, Greece, e-mail: dlouzis@aueb.gr; dlouzis@bankofgreece.gr

Although the main bulk of the BVAR literature concentrates on prior specification and shrinkage techniques about the dynamic VAR coefficients, Villani (2009) argues that the unconditional mean or the steady state of the system is an equally, or even more, important aspect of BVARs forecasting performance. The rationale is that longer-term forecasts of a stationary VAR converge to their unconditional means and thus, their estimated level plays a crucial role in forecasting accuracy. Moreover, economists usually have a more crystallized view on the steady-state level of an economy than on its short-term fluctuations, implying that a VAR model that accounts for this kind of prior information can improve its forecasting behavior. Therefore, the author proposes a mean-adjusted representation of a VAR model that enables the incorporation of prior beliefs on steady-states. The empirical evidence, so far, suggests that steady-state VAR models outperform their counterparts with uninformative priors on constant terms (e.g. see Beechey and Österholm 2008; Villani 2009; Clark 2011; Wright 2013). In Beechey and Österholm (2010), the authors also show that univariate steady-state AR models have better forecasting performance than AR models with traditional specifications.¹

Against this background, this study extends those of Villani (2009), George, Sun, and Ni (2008) and Korobilis (2013) and proposes methods of estimating a VAR model that incorporates prior beliefs on the steady-state and also adopts a Bayesian variable selection method. More specifically, we show that extant Gibbs sampling algorithms for Bayesian variable selection proposed by George, Sun, and Ni (2008) or Korobilis (2013) can be efficiently extended to allow for priors on the steady-state (Villani 2009). The essential steps include re-writing the VAR model in a mean-adjusted form, which allows prior elicitation for the unconditional mean, and adding an extra block to the Gibbs sampler of George, Sun, and Ni (2008) or Korobilis (2013) that draws from the full conditional posterior density of the steady-state parameters.

The proposed specification is evaluated in terms of an out-of-sample point forecasting exercise based on three US macroeconomic variables, namely, real gross domestic product (GDP) growth, inflation rate (consumer price index, CPI) and short-term interest rates (Federal funds rate). The out-of-sample period covers almost 46 years from 1969:Q4 to 2015:Q1 and the suggested model is compared to alternative established specifications such as the steady-state VAR model of Villani (2009) and the VAR model with variable selection methods either *à la* George, Sun, and Ni (2008) or *à la* Korobilis (2013). Finally, as a robustness check, we use three alternative prior specifications on the dynamic VAR parameters and variables in both ‘gap’ form, i.e. deviations from the long-run expectations, and non-gap form. Nonetheless, we should mention that since our main purpose is to evaluate the alternative specifications in terms of point forecasting, we do not consider steady-state VAR specifications enhanced with stochastic volatility that materially improve density forecasts (Clark 2011).

The rest of the paper is organized as follows. Section 2 develops the steady-state VAR with variable selection techniques and the Gibbs sampling algorithms. Section 3 presents the empirical results while Section 4 concludes this paper.

2 Steady-state VARs with variable selection

A reduced-form VAR is written as:

$$B(L)y_t = cd_t + \varepsilon_t \quad (1)$$

where y_t is a $m \times 1$ vector of time series at time t with $t=1, \dots, T$ observations, $B(L) = I_m - B_1L - \dots - B_pL^p$ with $Ly_t = y_{t-1}$, ε_t are the errors distributed as $N(0, \Sigma)$ with Σ being the $m \times m$ covariance matrix and d_t is a q -dimensional vector of exogenous deterministic variables such as constants, dummies or time trends. Assuming stationarity for y_t , the unconditional mean or steady-state of the VAR process in Eq. (1) is defined as $E(y_t) = \mu_t = B(L)^{-1}cd_t$. From the steady-state definition it is clear that it is hard to encapsulate prior opinions with respect to μ_t into Eq. (1).²

¹ Other studies that employ steady-state VARs are Adolfson et al. (2007), Jarociński and Smets (2008), Van Roye (2011), Österholm (2012).

² Nonetheless, it is obvious that if one specifies a prior on c he also specifies implicitly a prior on μ_t .

To circumvent this problem, Villani (2009) proposes a *steady-state* representation of the VAR model that is practically a *deviations-from-mean* parameterization, i.e.:

$$B(L)(y_t - \varphi d_t) = \varepsilon_t \quad (2)$$

where $\varphi = B(L)^{-1}c$ and the long-run mean is $\mu_t = \varphi d_t$. Therefore, a researcher can incorporate his prior beliefs on μ_t by directly specifying priors on φ . The *steady-state* VAR in Eq. (2) can be rewritten as:

$$\tilde{y}_t = B_1 \tilde{y}_{t-1} + \dots + B_p \tilde{y}_{t-p} + \varepsilon_t \quad (3)$$

with $\tilde{y}_t = y_t - \varphi d_t$ being the *mean-adjusted* times series of y_t . It is straightforward to extend the steady-state VAR in Eq. (3) to allow for Bayesian variable selection. The alternative variable selection approaches are discussed in Sections 2.1 and 2.2 below.

2.1 Stochastic search variable selection of George, Sun, and Ni (2008)

Before proceeding to the description of the *stochastic search variable selection* (SSVS) method proposed by George, Sun, and Ni (2008) it is useful to we rewrite the steady-state VAR using matrix notation, i.e.:

$$\tilde{Y} = \tilde{X}B + E \quad (4)$$

where \tilde{Y} and E are defined as a $T \times m$ matrices with their t -th row being $\tilde{Y}_t = \tilde{y}'_t = (\tilde{y}_{1t}, \tilde{y}_{2t}, \dots, \tilde{y}_{mt})$ and $E_t = \varepsilon'_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{mt})$, respectively. \tilde{X} is a $T \times k$ matrix with its t -th row being $\tilde{X}_t = \tilde{x}'_t = (\tilde{y}'_{t-1}, \dots, \tilde{y}'_{t-p})$ and $k = mp$ being the number of explanatory variables in each VAR equation. The matrix of VAR coefficients is defined as $B = (B_1, \dots, B_p)'$ with $\beta = \text{vec}(B)$ being the $n \times 1$ vector of regression coefficients and $n = mk = m^2 p$ being the total number of VAR dynamic coefficients.

George, Sun, and Ni (2008) uses the insights of the hierarchical Bayesian modeling so as to specify the prior distribution on β as a mixture of two Normal distributions. More specifically:

$$\beta | \gamma \sim N_n(\underline{b}, DD) \quad (5)$$

with $\gamma = (\gamma_1, \dots, \gamma_n)'$ being a $n \times 1$ vector of dummy variables which take the values of 1 and 0 and $D = \text{diag}(h_1, \dots, h_n)$ being a $n \times n$ diagonal matrix where:

$$h_j = \begin{cases} \tau_{0j} & \text{if } \gamma_j = 0 \\ \tau_{1j} & \text{if } \gamma_j = 1 \end{cases} \quad (6)$$

Eqs. (5) and (6) imply the following prior on each of the $j=1, \dots, n$ elements of β

$$\beta_j | \gamma_j \sim (1 - \gamma_j) N(\underline{b}_j, \tau_{0j}^2) + \gamma_j N(\underline{b}_j, \tau_{1j}^2) \quad (7)$$

The prior in Eq. (7) is hierarchical since it depends on the unknown dummy parameter γ_j , for which we specify a prior and we estimate its posterior distribution using the available information. Obviously, the degree of shrinkage towards the prior means, \underline{b}_j , depends on the hyperparameters τ_{0j}^2 and τ_{1j}^2 which are pre-selected by the researcher so that $\tau_{0j}^2 < \tau_{1j}^2$. Thus, if γ_j equals zero then β_j is forced to take a value very close to its prior mean because it is drawn from the first 'tight' Normal with a very small prior variance. Obviously, if the prior mean is zero the dynamic coefficient is shrunk towards zero (but not exactly zero) and the variable is almost excluded from the model. Whereas, if γ_j equals one then β_j is drawn from the second 'loose' Normal which is relatively uninformative due to its large prior variance.

The three alternative specifications used for hyperparameters \underline{b} and D are discussed in Section 3 and in more detail in Appendix B. Moreover, the γ_j dummy variables are assumed to be independent of each other $\forall j$ and their prior is defined as a Bernoulli density with π_j prior probability, i.e.:

$$\gamma_j | \gamma_{-j} \sim \text{Bernoulli}(\pi_j), \quad (8)$$

with $\Pr(\gamma_j = 1) = \pi_j$ and $\Pr(\gamma_j = 0) = 1 - \pi_j$ and γ_{-j} denoting all elements of γ except for the j -th.

We also follow George, Sun, and Ni (2008) and we impose restrictions on the inverse of the error covariance matrix, Σ^{-1} , via its Cholesky factor, Ψ . In particular, let $\Sigma^{-1} = \Psi\Psi'$ with Ψ being an upper triangular $m \times m$ matrix. The authors propose a Gamma prior for the squared elements of the main diagonal of Ψ , $\psi = (\psi_{11}, \dots, \psi_{mm})'$, i.e. $\psi_{jj}^2 \sim \text{gamma}(a_j, b_j)$ (with ψ_{jj} being independent of each other), and a mixture of two Normal distributions for the non-zero off diagonal elements of Ψ . In particular, we define $\eta_i = (\psi_{1i}, \dots, \psi_{i-1i})'$ for $i=2, \dots, m$ and we assume that:

$$\eta_i | \omega_i \sim N_{i-1}(0, D_i D_i) \quad (9)$$

where $\omega_i = (\omega_{1i}, \dots, \omega_{i-1i})'$ is a vector of dummy variables and $D_i = \text{diag}(h_{1i}, \dots, h_{i-1i})$ is a $(i-1) \times (i-1)$ matrix with:

$$h_{ji} = \begin{cases} \kappa_{0ji} & \text{if } \omega_{ji} = 0 \\ \kappa_{1ji} & \text{if } \omega_{ji} = 1 \end{cases} \quad (10)$$

Hyperparameters κ_{0ji} and κ_{1ji} are predetermined constants with $\kappa_{0ji} < \kappa_{1ji}$. As in Eq. (7), Eqs. (8) and (9) imply the following mixture of Normals for each element of η_i :

$$\psi_{ji} | \omega_{ji} \sim (1 - \omega_{ji}) N(0, \kappa_{0ji}^2) + \omega_{ji} N(0, \kappa_{1ji}^2) \quad (11)$$

Again each element of vector $\omega = (\omega'_2, \dots, \omega'_m)'$ is independent of each other and has a Bernoulli prior density, i.e.:

$$\omega_{ji} | \omega_{\setminus ji} \sim \text{Bernoulli}(\underline{q}_{ji}) \quad (12)$$

with $\Pr(\omega_{ji}=1) = \underline{q}_{ji}$ and $\Pr(\omega_{ji}=0) = 1 - \underline{q}_{ji}$.

Finally, the prior on the steady-state parameter, φ , is $\varphi \sim N_{mq}(\underline{b}_\varphi, \underline{V}_\varphi)$ assuming independence between φ and β as in Villani (2009).

2.1.1 Posterior inference for the steady-state VAR with stochastic search variable selection

The posterior inference is based on the idea that conditional on the steady-state parameters, φ , the VAR specification in Eq. (4) is a standard SSVS VAR for the mean-adjusted time series \tilde{y}_t , and therefore the Gibbs sampler framework of George, Sun, and Ni (2008) can be implemented.³ In fact, estimation of Eq. (4) requires only an extra block which samples the steady-state parameters from a normal density described in Villani (2009). A general Gibbs sampler algorithm for the steady-state VAR with stochastic search variable selection, which draws sequentially from the full conditional posterior density of the parameters, contains the following steps:⁴

1. Draw $\psi_{jj}^2 | \beta, \varphi, \eta; \mathcal{D}$ from $\text{gamma}(\bar{a}_j, \bar{b}_j)$;
2. Draw $\eta_i | \beta, \varphi, \psi; \mathcal{D}$ from $N_{i-1}(\bar{h}_i, \bar{\Delta}_i)$;
3. Draw $\omega_{ji} | \omega_{\setminus ji}, \beta, \varphi, \psi; \mathcal{D}$ from $\text{Bernoulli}(\bar{q}_{ji})$;
4. Draw $\beta | \gamma, \varphi, \Psi; \mathcal{D}$ from $N_n(\bar{\beta}, \bar{V})$;
5. Draw $\gamma_j | \gamma_{\setminus j}, \beta, \varphi, \Psi; \mathcal{D}$ from $\text{Bernoulli}(\bar{\pi}_j)$;
6. Draw $\varphi | \beta, \Psi; \mathcal{D}$ from $N_{mq}(\bar{b}_\varphi, \bar{V}_\varphi)$.

where $\mathcal{D} = \{y_1, \dots, y_T, d_1, \dots, d_T\}$ denote the data. The quantities $\bar{a}_j, \bar{b}_j, \bar{h}_i, \bar{\Delta}_i, \bar{q}_{ji}, \bar{\beta}, \bar{V}, \bar{\pi}_j, \bar{b}_\varphi$ and \bar{V}_φ as well as the posterior sampling procedure are described in detail in Appendix A.

³ Alternatively, one can argue that conditional on γ and ω the model in (4) is a standard steady-state VAR and full conditional posterior distributions of Villani (2009) can be implemented. In both cases the posterior analysis is identical.

⁴ A full conditional posterior density of a parameter is the posterior density conditional on all other parameters.

2.2 Variable selection of Korobilis (2013)

The *steady-state* VAR in Eq. (3) can also be extended to allow for *variable selection* (VS) in the form of Korobilis (2013). In contrast to George, Sun, and Ni (2008), the Korobilis (2013) method does exclude variables, in a sense that the dynamic coefficients take a value *exactly* zero and not close to zero. This means that each of the VAR equations may have different lagged variables and the steady-state VAR in Eq. (3) should be re-written as a system of seemingly unrelated equations (SUR) i.e.:

$$\tilde{y}_t = \tilde{z}_t \beta + \varepsilon_t \quad (13)$$

Where $\tilde{z}_t = I_m \otimes \tilde{x}_t$ is a $m \times n$ matrix with $n = m^2 p$ being, again, the total number of coefficients in Eq. (13), $\tilde{x}_t = (\tilde{y}'_{t-1}, \dots, \tilde{y}'_{t-p})$ is a $1 \times k$ ($k = mp$) vector containing all regressors at time t and $\beta = \text{vec}(B)$ is the $n \times 1$ vector of the VAR dynamic coefficients. The model in Eq. (13) is an *unrestricted steady-state* VAR since no restrictions are incorporated in the $[\beta_j]_{j=1}^n$ elements of β . By contrast, the Bayesian variable selection method proposed by Korobilis (2013) restricts some of the β_j coefficients to be zero as follows

$$\begin{cases} \beta_j = 0 & \text{if } \gamma_j = 0 \\ \beta_j \neq 0 & \text{if } \gamma_j = 1 \end{cases} \quad (14)$$

where γ_j is an indicator variable and the j -th element of the vector $\gamma = (\gamma_1, \dots, \gamma_n)'$. We define the *steady-state* VAR with *variable selection* as

$$\tilde{y}_t = \tilde{z}_t \theta + \varepsilon_t \quad (15)$$

where $\theta = \Gamma \beta$ and Γ is a $n \times n$ diagonal matrix with the elements of γ on its main diagonal, i.e. $\Gamma_{jj} = \gamma_j$. The specification in Eq. (15) implies that for $\gamma_j = 1 \forall j$ we get the unrestricted steady-state VAR of Eq. (13).

Bayesian estimation and inference on β , γ , φ and Σ requires the specification of prior distributions which are generally based on the propositions of Korobilis (2013) and Villani (2009). Accordingly, we define the prior density for the dynamic parameters, β , as a multivariate normal distribution, i.e. $\beta \sim N_n(\underline{b}, \underline{V})$, with hyperparameters \underline{b} and \underline{V} being further specified in Section 3 and Appendix B. The γ_j dummy variables are assumed to be independent of each other $\forall j$ and their prior is defined as a Bernoulli density with $\underline{\pi}_j$ prior probability, i.e. $\gamma_j | \gamma_{\setminus j} \sim \text{Bernoulli}(\underline{\pi}_j)$, where $\gamma_{\setminus j}$ denotes all elements of γ except for the j -th. Finally, the prior on Σ is taken to be the scale invariant improper Jeffreys prior, i.e. $\Sigma \propto |\Sigma|^{-(m+1)/2}$, while the prior on φ is $\varphi \sim N_{mq}(\underline{b}_\varphi, \underline{V}_\varphi)$ assuming independence between φ and β .

2.2.1 Posterior inference for the steady-state VAR with variable selection

The posterior inference is based on the same idea already presented in Section 2.1.1 for the steady-state VAR with stochastic search variable selection. This means that we implement the Gibbs sampler framework of Korobilis (2013) for the mean-adjusted time series \tilde{y}_t and then we add an extra fourth block which samples steady-state parameters from a normal density described in Villani (2009). A general Gibbs sampler algorithm for the steady-state VAR with variable selection, contains the following four steps:

1. Draw $\beta | \gamma, \varphi, \Sigma; \mathcal{D}$ from $N_n(\bar{\beta}, \bar{V})$;
2. Draw $\gamma_j | \gamma_{\setminus j}, \beta, \varphi, \Sigma; \mathcal{D}$ from $\text{Bernoulli}(\bar{\pi}_j)$;
3. Draw $\Sigma^{-1} | \gamma, \beta, \varphi; \mathcal{D}$ from $\text{Wishart}(T, S^{-1})$;
4. Draw $\varphi | \gamma, \beta, \Sigma; \mathcal{D}$ from $N_{mq}(\bar{b}_\varphi, \bar{V}_\varphi)$.

Appendix A describes in detail the formulation of $\bar{\beta}$, \bar{V} , $\bar{\pi}_j$, S , \bar{b}_φ and \bar{V}_φ as well as the posterior sampling procedure.

In general, the extra block which draws the steady-state parameters requires only a very small portion of the total computing time, for both Gibbs sampling algorithms described in Sections 2.1.1 and 2.2.1, as also

pointed out in Villani (2009). In practice, the Gibbs sampler is proved to be very efficient and convergence problems may arise only under certain conditions. More specifically, the unconditional mean is not identified for a non-stationary VAR process and this may lead to convergence difficulties *only if* it is combined with an uninformative (i.e. very large prior variance) steady-state prior (Villani 2005; Villani 2009, 646–647). Villani (2005, 2009) shows theoretically how an informative prior on steady-state stabilizes the Gibbs sampler even if the VAR system approaches a unit root process. The author also uses a simulation exercise and shows that even moderately informative steady-state priors can produce acceptable posterior simulations. We also confirm the abovementioned findings for the proposed models by implementing a similar simulation exercise.⁵

3 Empirical analysis

We use three major US macroeconomic series, i.e. real GDP growth rate, CPI inflation rate and effective Federal funds rate, in order to estimate the proposed model and evaluate its forecasting performance (henceforth, we refer to these variables as GDP (r_t), inflation (π_t) and interest rate (i_t), respectively). The data series cover the period from 1954:Q2 to 2015:Q1 and were obtained from the St. Louis' FRED database on a quarterly basis.⁶ Real GDP growth and CPI inflation rate are calculated as annualized quarter-on-quarter percentage changes, i.e. $r_t = 400 \times \ln(x_t/x_{t-1})$ and $\pi_t = 400 \times \ln(p_t/p_{t-1})$ where p and x are quarterly CPI inflation index and real GDP, respectively, while we use the interest rate in levels. Figure 1 shows the evolution of the variables for the full sample period.

The inflation graph reveals the well-known slowly moving trend component which rises during the 1970s (Great Inflation) and starts falling after a decade (Great Moderation). The main bulk of the macroeconomic forecasting literature argues that this trend has to be taken into consideration in terms of inflation forecasting, while others argue that it represents the agent's perceptions of the Fed's time varying long-run inflation target (see Faust and Wright 2013 and references therein). In the same vein, Chan and Koop (2014) find a structural break on the steady-state of inflation and interest rates during the 1970s, while Clark (2011) and Wright (2013) account for this time-varying local mean for inflation and interest rates by employing the long-term inflation expectations based on Blue Chip surveys.⁷ Moreover, a number of studies show that the survey based long-run inflation expectations can be conveniently approximated by using a simple exponential smoothing (e.g. see Cogley 2002; Clark 2011; Faust and Wright 2013 among others).

Here, we consider a twofold approach. First, we estimate the alternative specifications using as endogenous variables the raw figures, i.e. the vector of the endogenous variables is given by $y_t = (r_t, \pi_t, i_t)'$. Second, we also use inflation and interest rates in a gap form, i.e. deviations from the long-run inflation expectations, where the long-run expectations are roughly proxied by the exponential smoothed counterpart of inflation, π_t^{es} . More specifically, we estimate the exponential smoothed inflation as $\pi_t^{es} = \alpha \pi_{t-1}^{es} + (1-\alpha) \pi_t$ where exponential smoothing parameter is set equal to 0.95. Then, inflation and interest variables in a gap form are defined as $\pi_t^{gap} = \pi_t - \pi_t^{es}$ and $i_t^{gap} = i_t - \pi_t^{es}$, respectively, which are also depicted in Figure 1 along with the π_t^{es} variable. Therefore, now, the vector of endogenous variables is given by $y_t = (r_t, \pi_t^{gap}, i_t^{gap})'$ and the steady-state priors are set on the gap variables. Since the alternative VAR models provide forecasts only for the π_t^{gap} and i_t^{gap} variables, we also assume a random walk process for the low frequency component of inflation, π_t^{es} , in order

⁵ Simulation results are available upon request.

⁶ More specifically, the series were downloaded from <http://research.stlouisfed.org/fred2> and are defined in more detail as follows: *Real gross domestic product* (code: CDPC96), 3 decimal, billions of chained 2009 dollars, seasonally adjusted annual rate. *Consumer Price Index* (code: CPIAUCSL_PC1) for all urban consumers: all Items, percentage change from a year ago, seasonally adjusted. *Effective Federal Funds Rate* (code: FEDFUNDS), percent, not seasonally adjusted.

⁷ In particular, Clark (2011) uses the inflation and interest rate variables in a gap form taking the difference between the raw variable and the long-run inflation expectations from the Blue Chip Consensus and sets steady-state priors on the gap variables. By contrast, Wright (2013) uses the raw variables and employs the Blue Chip long-term forecasts to set the steady-state priors on the endogenous variables. Both alternatives account for the trend uncertainty of inflation and interest rate.

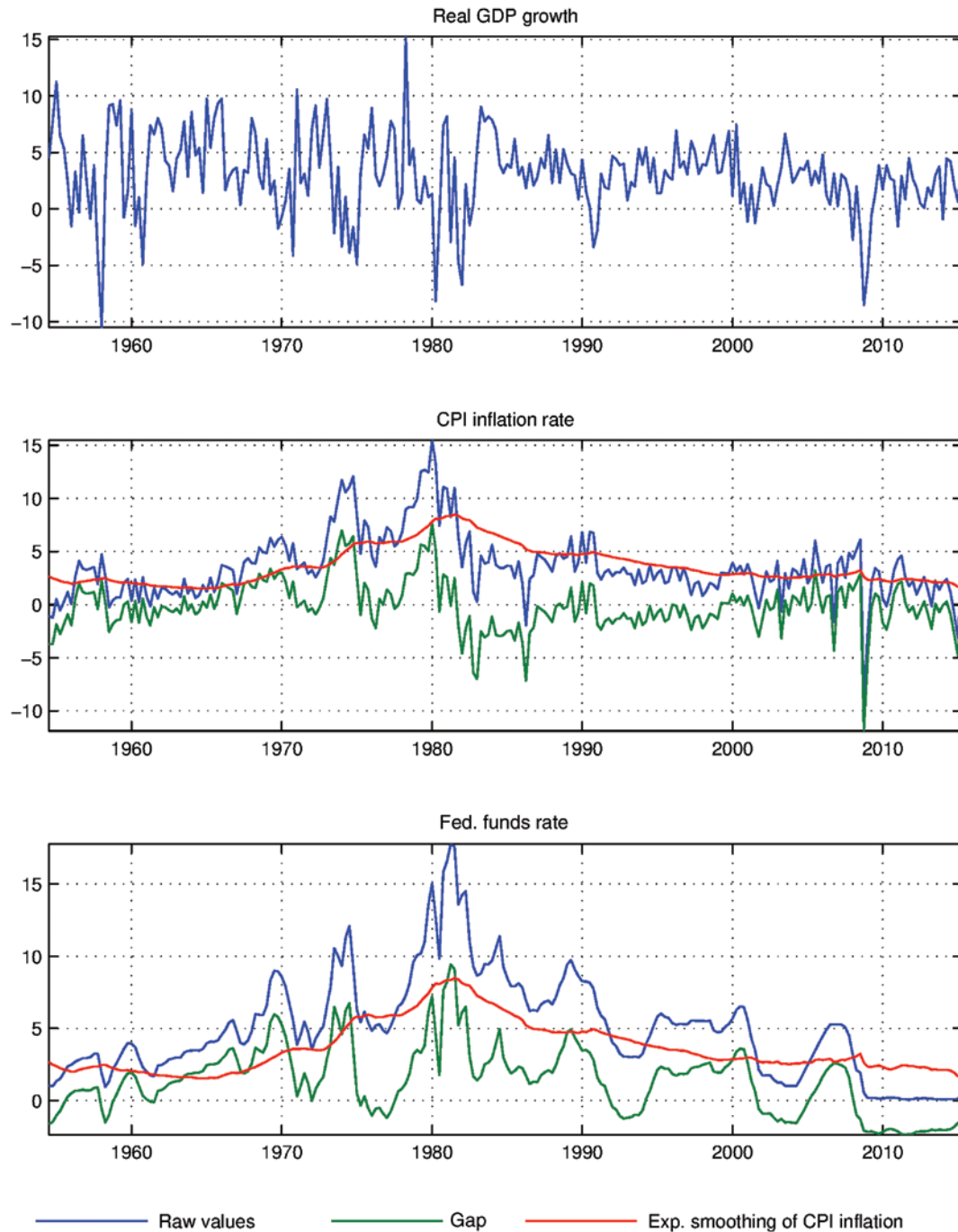


Figure 1: Annualized real GDP growth rate, CPI inflation rate and Federal Funds rate in (%).

to get forecasts for the raw figures of inflation and interest rate, π_t and i_t , respectively (Faust and Wright 2013). Although the two alternative approaches are mainly used as a robustness check for the alternative specifications, a comparison between gap and non-gap models is always interesting.

3.1 Prior specifications and alternative models

An initial first step in Bayesian estimation is to specify the hyperparameters in prior distributions described in Sections 2.1 and 2.2. In particular, we follow George, Sun, and Ni (2008), Koop (2013) and Korobilis (2013)

and we use three distinct prior specifications on the dynamic parameters vector, β . In particular for the SSVS VAR models we use (i) the default ‘*semi-automatic*’ approach of George, Sun, and Ni (2008) (standard prior hereafter), (ii) the combination of the SSVS method with the Minnesota prior of Koop (2013) and (iii) the combination of the SSVS method with the Ridge prior (see also Korobilis 2013). Similarly for the VS VARs we use (i) the ridge regression prior (ridge prior hereafter), (ii) the Minnesota prior and (iii) the hierarchical Bayes shrinkage prior (shrink prior hereafter); all three of them are based on the Normal distribution and are briefly described in Appendix B.⁸ As regards the VS specifications, we set the prior probability of the Bernoulli density equal to 0.8, i.e. $\pi_j=0.8$, while the standard non-informative prior $|\Sigma|^{-(m+1)/2}$ is used for Σ . For the SSVS specifications we set $\pi_j=0.5$, $q_j=0.7$, $\kappa_{0ji}=1$, $\kappa_{1ji}=5$, $a_j=0.01$ and $b_j=0.01$.

In the steady-state version of VAR models we also have to specify the prior mean and standard deviation on steady-states coefficients, φ . Following the recent literature (e.g. see Clark 2011; Österholm 2012), we set the steady-state GDP growth and inflation rate equal to 3% and 2%, respectively, and the nominal interest rate equal to 4%, while we assume a 0.5 standard deviation for GDP and inflation and 0.7 for interest rate. Finally, the steady-state levels for inflation and interest rate variables in a gap form are set equal to 0% and 2%, respectively (see Clark 2011).

The primary scope of the empirical analysis is to examine whether the forecasting ability of a Bayesian VAR model is improved when we incorporate prior beliefs with respect to the steady-state of the economy and we also allow for a (semi-) automatic Bayesian variable selection. To that end, we compare the forecasting ability of the following six alternative BVAR specifications (models’ names in parenthesis):

- (i) a baseline BVAR without variable selection and steady-state priors (Standard),
- (ii) a steady-state BVAR (Villani 2009) (SSP),
- (iii) a BVAR with variable selection (Korobilis 2013) (VS),
- (iv) a BVAR with both variable selection and steady-state priors (VS-SSP),
- (v) a BVAR with stochastic search variable selection (George, Sun, and Ni 2008) (SSVS),
- (vi) a BVAR with stochastic search variable selection and steady-state priors (SSVS-SSP).

These six alternative specifications are estimated using each of the three prior specifications on β meaning that we have in total 18 distinct BVAR models. For all models we use a lag length of 4. For comparison reasons we also estimate a standard VAR model using ordinary least squares (OLS) with one lag.

3.2 In-sample analysis

In this section we present some estimation results using the full data sample (1954:Q2–2015:Q1). We use the Gibbs samplers in Appendix A to sample 10,000 draws with a thinning of 5, i.e. we keep one every five draws, after discarding the first 2000 draws used for initial convergence (burn-in period). We evaluate the convergence of the proposed Markov Chain Monte Carlo (MCMC) algorithm by implementing the inefficiency factors (IFs) metric proposed by Primiceri (2005). Table 1 presents a summary of the distribution for the inefficiency factors of the parameters posterior draws, while Figure 2 shows the IFs across all parameters. The empirical evidence reveals that the convergence of the Gibbs samplers is more than satisfactory with the majority of the IFs being well below the threshold value of 20 (see Primiceri 2005).⁹

Table 2 presents the posterior means and standard deviations for the steady-states across different priors and VAR specifications. Overall, the results show that neither different priors nor parameter restrictions (i.e. models with variable selection) have a significant impact on steady-states estimation. In particular, the steady-state growth rate of GDP is estimated to be close to 3% ranging between 2.99% and 3.06%. The average

⁸ As pointed out in Korobilis (2013) all these specifications offer some kind of shrinkage but no exact zero shrinkage as in the variable selection methodology.

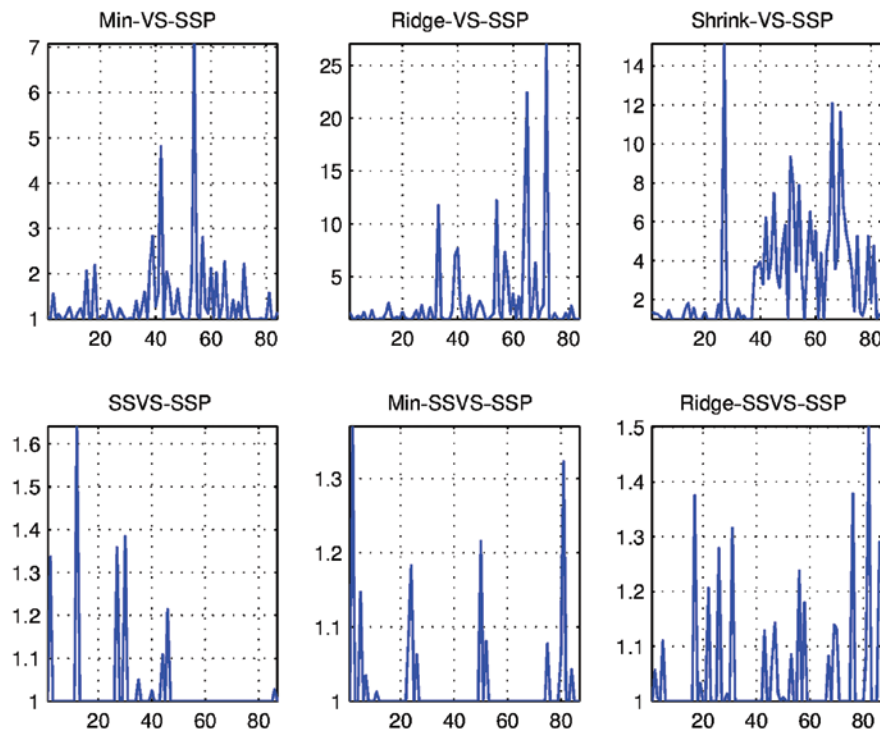
⁹ According to Primiceri (2005) a value of IF around or below 20 indicates that the convergence of the MCMC algorithm is satisfactory.

Table 1: MCMC convergence diagnostics: summary of distribution of inefficiency factors.

	Median	Mean	Min	Max	10-th Percentile	90-th percentile
Min-VS-SSP	1.09	1.38	1.00	7.08	1.00	2.13
Ridge-VS-SSP	1.15	2.63	1.00	27.06	1.00	6.39
Shrink-VS-SSP	1.55	3.06	1.00	15.15	1.00	6.55
SSVS-SSP	1.00	1.02	1.00	1.64	1.00	1.02
Min-SSVS-SSP	1.00	1.02	1.00	1.37	1.00	1.06
Ridge-SSVS-SSP	1.00	1.05	1.00	1.50	1.00	1.20

This table presents a summary of the distribution of inefficiency factors (IFs) for the posterior estimates of the parameters. Models are estimated using the full sample and variables in a non-gap form. Inefficiency factors are defined as the inverse of the relative numerical efficiency measure of Geweke (1992) and are calculated as $\left(1+2\sum_{k=1}^{\infty}\rho_k\right)$, with ρ_k being the k -th

autocorrelation of the posterior estimates chain. Following Primiceri (2005), we calculate IFs using a 4% tapered window for the estimation of spectral density at frequency zero. Min, Ridge and Shrink are the Minnesota, Ridge and Shrink priors for the dynamic coefficients, respectively. (prior)-VS-SSP is a steady-state BVAR(4) model with variable selection *à la* Korobilis (2013). (prior)-SSVS-SSP is a steady-state BVAR(4) model with stochastic search variable selection *à la* George, Sun, and Ni (2008). In SSVS-SSP we use the standard prior specification for the regression coefficients proposed in George, Sun, and Ni (2008).

**Figure 2:** MCMC convergence diagnostics: inefficiency factors.

Notes: The figures show the inefficiency factors for all parameters across the proposed specifications using the full sample and variables in a non-gap form. Min, Ridge and Shrink are the Minnesota, Ridge and Shrink priors for the dynamic coefficients, respectively. (prior)-VS-SSP is a steady-state BVAR(4) model with variable selection *à la* Korobilis (2013). (prior)-SSVS-SSP is a steady-state BVAR(4) model with stochastic search variable selection *à la* George, Sun, and Ni (2008). In SSVS-SSP we use the standard prior specification for the regression coefficients proposed in George, Sun, and Ni (2008).

CPI inflation across priors and specifications is close to 2.4% and 0.02% for the inflation in a gap form, while the steady-state of the Fed funds rate is estimated to be close to 3.85% (1.7% for the gap form).

In Table 3 we also present estimation results regarding the restriction parameters γ_j using the VS method. As pointed out in Korobilis (2013) the posterior mean of γ_j can be seen as an average probability of including

Table 2: Posterior means and standard deviations of the steady-states.

							VAR dynamic coefficients prior		
	Ridge		Minnesota		Shrink		Standard SSVS	Minnesota SSVS	Ridge SSVS
	no VS	with VS	no VS	with VS	no VS	with VS			
Real GDP growth									
Posterior mean	3.03 (3.02)	3.03 (3.00)	3.02 (3.02)	3.03 (3.01)	3.04 (2.99)	3.04 (3.01)	3.06 (2.99)	3.06 (2.99)	3.02 (3.04)
St. deviation	0.30 (0.26)	0.29 (0.27)	0.29 (0.27)	0.29 (0.27)	0.27 (0.25)	0.27 (0.25)	0.30 (0.27)	0.29 (0.26)	0.27 (0.16)
CPI Inflation									
Posterior mean	2.40 (0.01)	2.40 (0.00)	2.42 (0.00)	2.42 (0.03)	2.39 (0.03)	2.39 (0.02)	2.42 (0.03)	2.39 (0.03)	2.42 (0.02)
St. deviation	0.43 (0.30)	0.43 (0.31)	0.44 (0.30)	0.42 (0.30)	0.42 (0.32)	0.44 (0.31)	0.42 (0.36)	0.45 (0.34)	0.43 (0.17)
Federal funds rates									
Posterior mean	3.85 (1.69)	3.86 (1.71)	3.83 (1.68)	3.84 (1.69)	3.84 (1.70)	3.85 (1.71)	3.83 (1.61)	3.89 (1.65)	3.84 (1.89)
St. deviation	0.58 (0.51)	0.57 (0.50)	0.58 (0.51)	0.56 (0.50)	0.57 (0.52)	0.58 (0.53)	0.55 (0.46)	0.55 (0.46)	0.57 (0.44)

Models are estimated using the full sample period. Numbers in parenthesis show the steady-states estimates using inflation and Federal funds rate in a gap form. ‘no VS’ denotes the models without variable selection while ‘with VS’ denotes the models with variable selection using the method of Korobilis (2013). SSVS denotes the stochastic search variable selection method of George, Sun, and Ni (2008).

the respective β_j parameter to the true model. Our intention is to examine whether the incorporation of steady-state priors in a restricted VAR model has a significant impact on γ_j 's estimations. Indeed, the results in Table 3 reveal that the differences in γ_j estimations between models with and without steady-state priors are negligible with the greatest divergences been observed among the models with different priors on the regression coefficients. We also see that the first own lag is always included in the model across variables and priors, as expected. Moreover, the estimation results with respect to the γ_j 's and ω_{ji} 's parameters of the SSVS method (not shown here due to space considerations) also show that the inclusion of the steady-state priors in a SSVS VAR model has an insignificant impact.

3.3 Out-of-sample forecasting analysis

We evaluate the alternative VAR specifications using a forecasting horizon of 12 quarters (3 years), $h=1, \dots, 12$, and an out-of-sample period spanning from 1969:Q4 to 2015:Q1. The forecasts across all horizons are produced using a recursive forecasting scheme. This means that we use an initial sample (1954:Q2–1969:Q3) to generate forecasts from 1969:Q4 to 1972:Q3, i.e. 12 quarters ahead. Next, we allow the sample to expand and include one more period, i.e. 1954:Q2–1969:Q4, and generate forecasts from 1970:Q1 to 1972:Q4. This procedure is continued till the end of the sample period. As mentioned in Section 3.2 estimation and forecasting results are based on 10,000 posterior simulations (after a burn-in period of 2000 simulations), keeping one every five draws, and forecasts are generated iteratively following Korobilis (2013, 215–216).

Before proceeding to the forecasting evaluation, we present the sequential GDP, inflation and interest rate out-of-sample forecasts along with the actual variables in Figures 3–5. The forecasts has been generated by five alternative models: a VAR(1) estimated with OLS, two VAR models with variable selection (VS and SSVS) and a Minnesota prior, and two steady-state VAR models with variable selection (VS and SSVS) and a Minnesota prior. Overall, the forecasts generated by the steady-state VARs tend to converge to the steady-state of each variable much faster than its counterparts, as expected. This phenomenon is particularly evident

Table 3: Posterior means of the γ vector elements using the method of Korobilis (2013).

	Ridge		Minnesota		Shrink	
	Uninformative	Steady-state	Uninformative	Steady-state	Uninformative	Steady-state
Dependent variable: r_t						
r_{t-1}	1.00	1.00	1.00	1.00	1.00	1.00
π_{t-1}	0.13	0.12	0.24	0.28	0.74	0.68
i_{t-1}	0.26	0.34	0.26	0.28	0.68	0.71
r_{t-2}	0.76	0.77	0.97	0.96	0.78	0.75
π_{t-2}	0.18	0.22	0.43	0.48	0.74	0.74
i_{t-2}	1.00	1.00	0.98	0.99	0.81	0.78
r_{t-3}	0.09	0.09	0.64	0.64	0.67	0.68
π_{t-3}	0.31	0.21	0.64	0.57	0.70	0.70
i_{t-3}	1.00	0.99	0.98	0.97	0.80	0.79
r_{t-4}	0.06	0.07	0.65	0.66	0.62	0.67
π_{t-4}	0.13	0.15	0.59	0.57	0.72	0.69
i_{t-4}	0.34	0.30	0.58	0.59	0.68	0.67
Dependent variable: π_t						
r_{t-1}	0.06	0.07	0.34	0.37	0.62	0.70
π_{t-1}	1.00	1.00	1.00	1.00	1.00	1.00
i_{t-1}	1.00	1.00	1.00	1.00	0.98	0.96
r_{t-2}	0.14	0.07	0.69	0.57	0.67	0.66
π_{t-2}	0.09	0.10	0.61	0.66	0.68	0.65
i_{t-2}	0.86	0.96	0.88	0.96	0.78	0.87
r_{t-3}	0.05	0.06	0.55	0.59	0.62	0.67
π_{t-3}	1.00	1.00	1.00	1.00	1.00	1.00
i_{t-3}	0.40	0.23	0.71	0.63	0.83	0.76
r_{t-4}	0.16	0.27	0.83	0.86	0.69	0.72
π_{t-4}	0.08	0.07	0.69	0.71	0.61	0.64
i_{t-4}	0.21	0.20	0.61	0.59	0.71	0.69
Dependent variable: i_t						
r_{t-1}	1.00	1.00	1.00	1.00	1.00	1.00
π_{t-1}	0.05	0.05	0.26	0.35	0.62	0.67
i_{t-1}	1.00	1.00	1.00	1.00	1.00	1.00
r_{t-2}	0.22	0.19	0.80	0.84	0.71	0.71
π_{t-2}	0.92	0.70	0.99	0.94	0.98	0.92
i_{t-2}	1.00	1.00	0.99	1.00	0.94	0.89
r_{t-3}	0.02	0.03	0.61	0.59	0.61	0.61
π_{t-3}	0.03	0.05	0.55	0.55	0.67	0.59
i_{t-3}	1.00	0.98	1.00	1.00	0.79	0.79
r_{t-4}	0.02	0.03	0.64	0.60	0.64	0.58
π_{t-4}	0.04	0.04	0.57	0.58	0.60	0.61
i_{t-4}	0.92	0.86	0.82	0.81	0.77	0.70

Models are estimated using the full sample and variables in a non-gap form. ‘Uninformative’ denotes models with uninformative priors on the steady-state.

during the Great Inflation period, i.e. 1970–1980; take as an example the inflation forecasts generated by the models at the end of 1979, during which the variable takes its maximum value. The models without a steady state prior seem to over-estimate the steady-state levels and their forecasts drift away in even higher levels. Another example is the GDP forecasts during the recession periods of the 1970s and 1980s during which the non steady-state models seem to under-estimate the unconditional means of the processes. Another interesting point is that the differences in the forecast patterns between steady-state and non-steady-state models are mitigated when we use variables in a gap form (see the right columns of Figures 3–5) implying that these two distinct categories of models may also be closer in terms of forecasting evaluation when we use variables in a gap form (see also the discussion in Section 3.3.2).

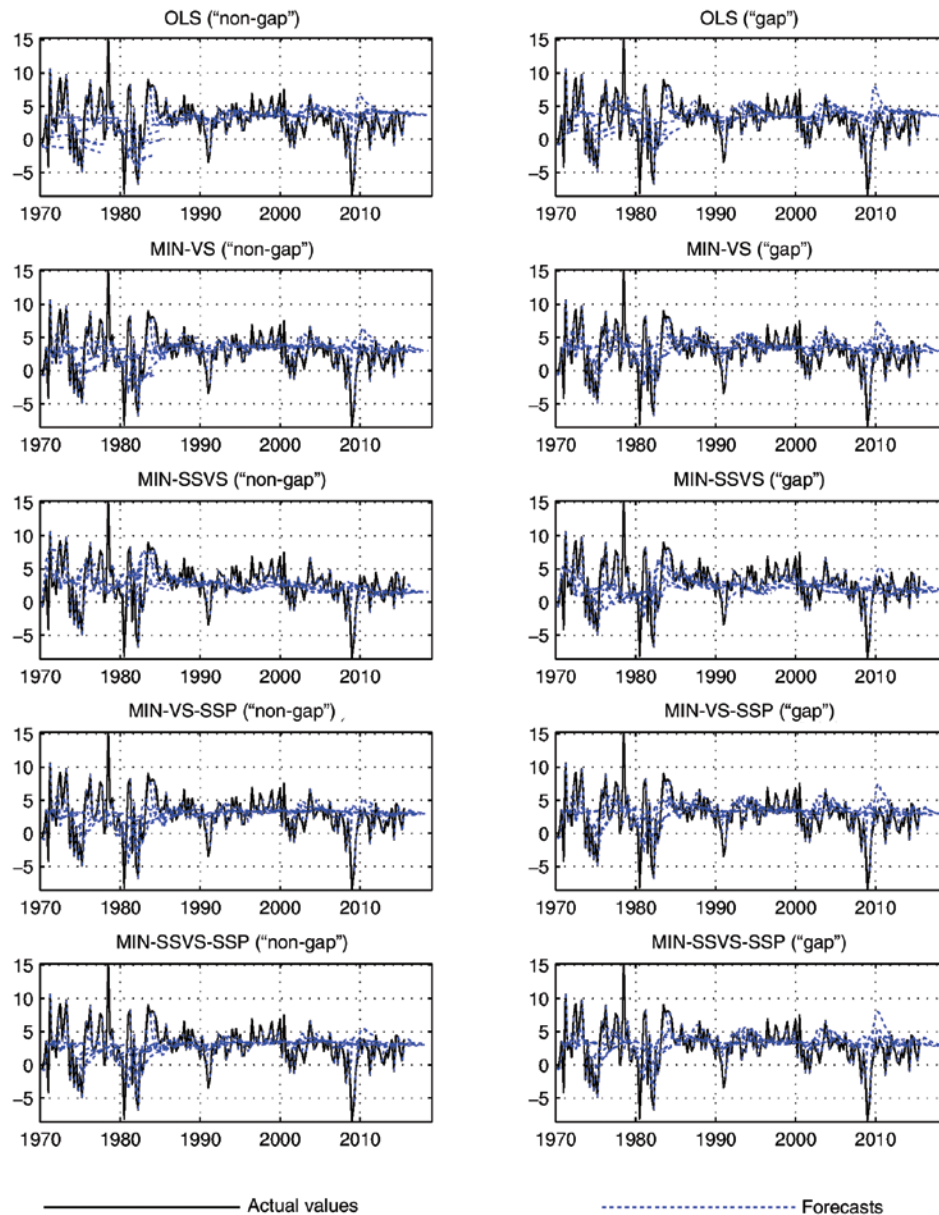


Figure 3: Out-of-sample forecasts for real GDP growth (%).

Notes: This figure presents the annualized GDP growth out-of-sample forecasts for 12 quarters ahead (dotted lines) along with the actual values (solid line). The left (right) column presents the forecasts generated by models with inflation and interest rates variables specified in “non-gap” (“gap”) form. MIN denotes the Minnesota prior, VS and SSVS denotes the variable selection and stochastic search variable selection as employed by Korobilis (2013) and George, Sun, and Ni (2008), respectively and SSP denotes the steady state prior. The forecasts are plotted every other quarter for clarity in presentation.

Our main research interest is to examine whether the abovementioned forecasting behavior leads to superior forecasting ability for the models that incorporate both steady-state priors and variable selection techniques. To that end, we use two popular forecasting evaluation metrics, namely, the *root mean squared forecast error* (RMSFE) and the *mean absolute forecast error* (MAFE). Following the usual convention, we use relative forecasting evaluation metrics, i.e. we present the RMSFE and MAFE metrics of the various competing models as proportion of the corresponding RMSFE and MAFE metrics of a *random walk with drift* (RW) process for the levels of the variables.

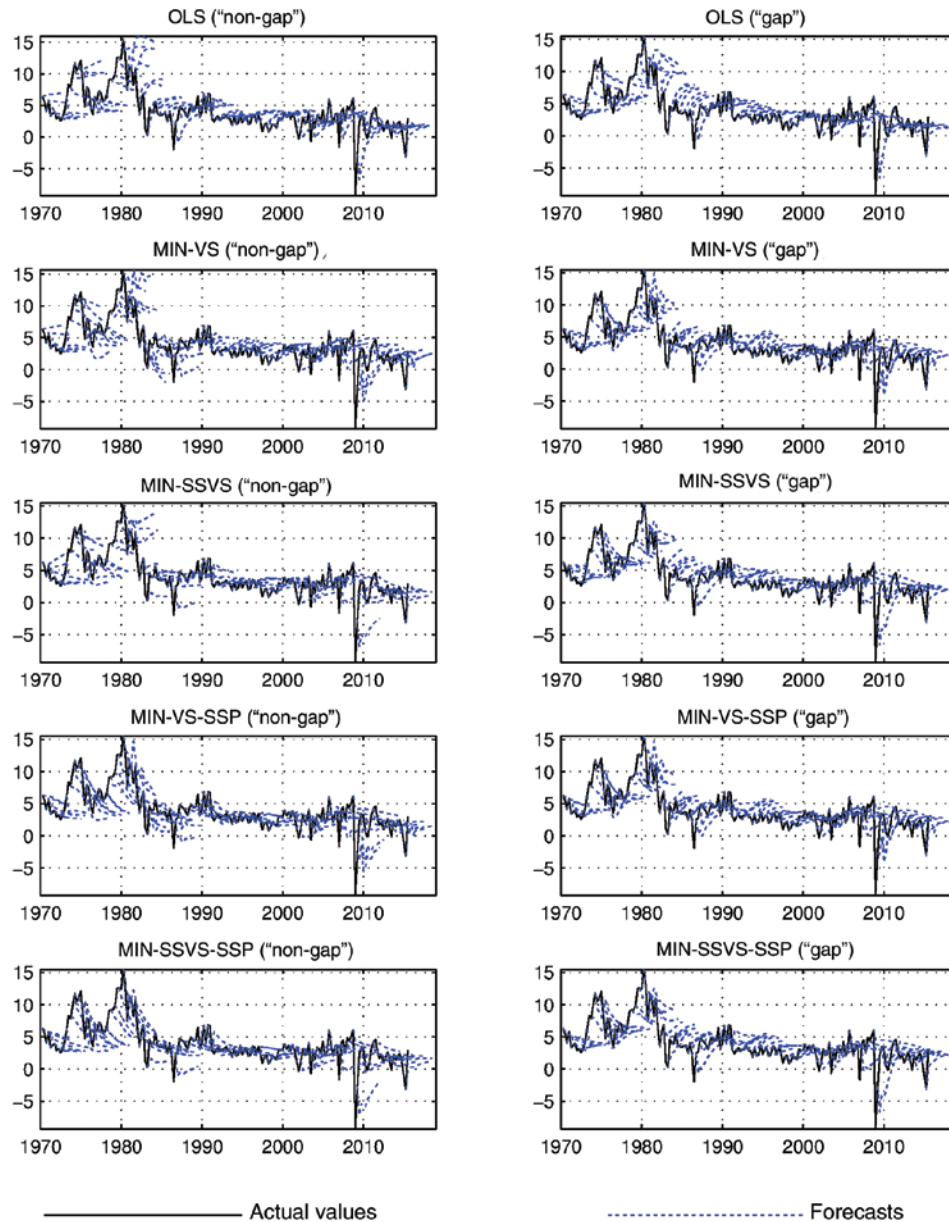


Figure 4: Out-of-sample forecasts for CPI inflation (%).

Notes: This figure presents the annualized CPI inflation out-of-sample forecasts for 12 quarters ahead (dotted lines) along with the actual values (solid line). The left (right) column presents the forecasts generated by models with inflation and interest rates variables specified in “non-gap” (“gap”) form. MIN denotes the Minnesota prior, VS and SSVS denotes the variable selection and stochastic search variable selection as employed by Korobilis (2013) and George, Sun, and Ni (2008), respectively and SSP denotes the steady state prior. The forecasts are plotted every other quarter for clarity in presentation.

$$\text{Relative RMSFE}_{ij}^h = \frac{\text{RMSFE}_{ij}^h}{\text{RMSFE}_{i,RW}^h}, \quad \text{Relative MAFE}_{ij}^h = \frac{\text{MAFE}_{ij}^h}{\text{MAFE}_{i,RW}^h}$$

where $h=1, \dots, 12$ is the forecasting horizon, i is the variable of interest, i.e. GDP, inflation and interest rate, and j are the various competing models outlined in Section 3.1.. Values below one indicate that the corresponding model outperforms the random walk process and vice versa.

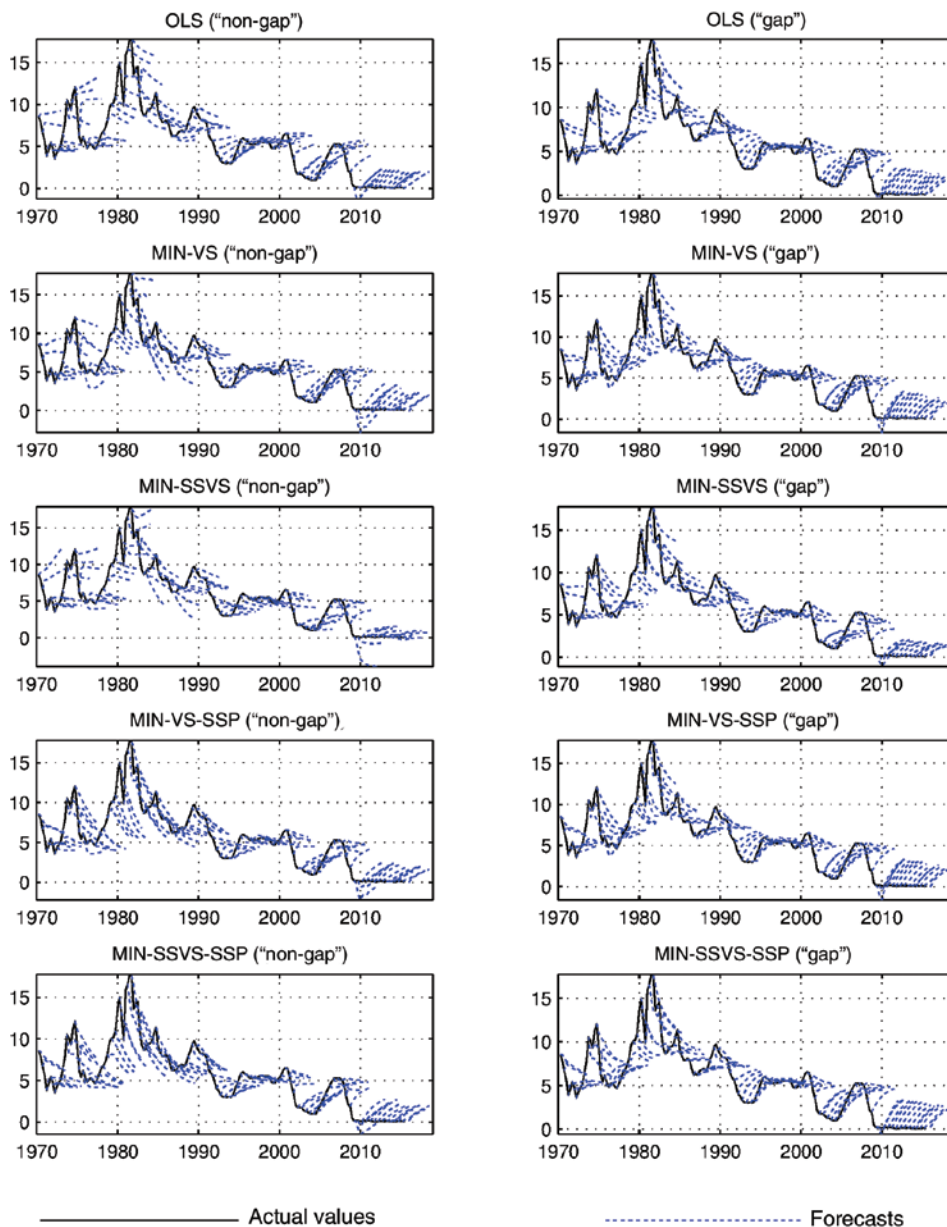


Figure 5: Out-of-sample forecasts for Federal Funds rate (%).

Notes: This figure presents the Federal Funds rate out-of-sample forecasts for 12 quarters ahead (dotted lines) along with the actual values (solid line). The left (right) column presents the forecasts generated by models with inflation and interest rates variables specified in “non-gap” (“gap”) form. MIN denotes the Minnesota prior, VS and SSVS denotes the variable selection and stochastic search variable selection as employed by Korobilis (2013) and George, Sun, and Ni (2008), respectively and SSP denotes the steady state prior. The forecasts are plotted every other quarter for clarity in presentation.

3.3.1 Forecasting results using variables in a non-gap form

Figures 6 and 7 present the relative RMSFE and MAFE out-of sample results for the (i) Standard, (ii) VS, (iii) SSP and (iv) VS-SSP models using variables in a non-gap form. Thus, Figures 6 and 7 examine the forecasting ability of the models that adopt variable selection technique as implemented in Korobilis (2013). On the other hand, Figures 8 and 9 compare models that employ the stochastic search variable selection of George, Sun, and Ni (2008). In particular, Figures 8 and 9 present the relative RMSFE and MAFE out-of sample results for

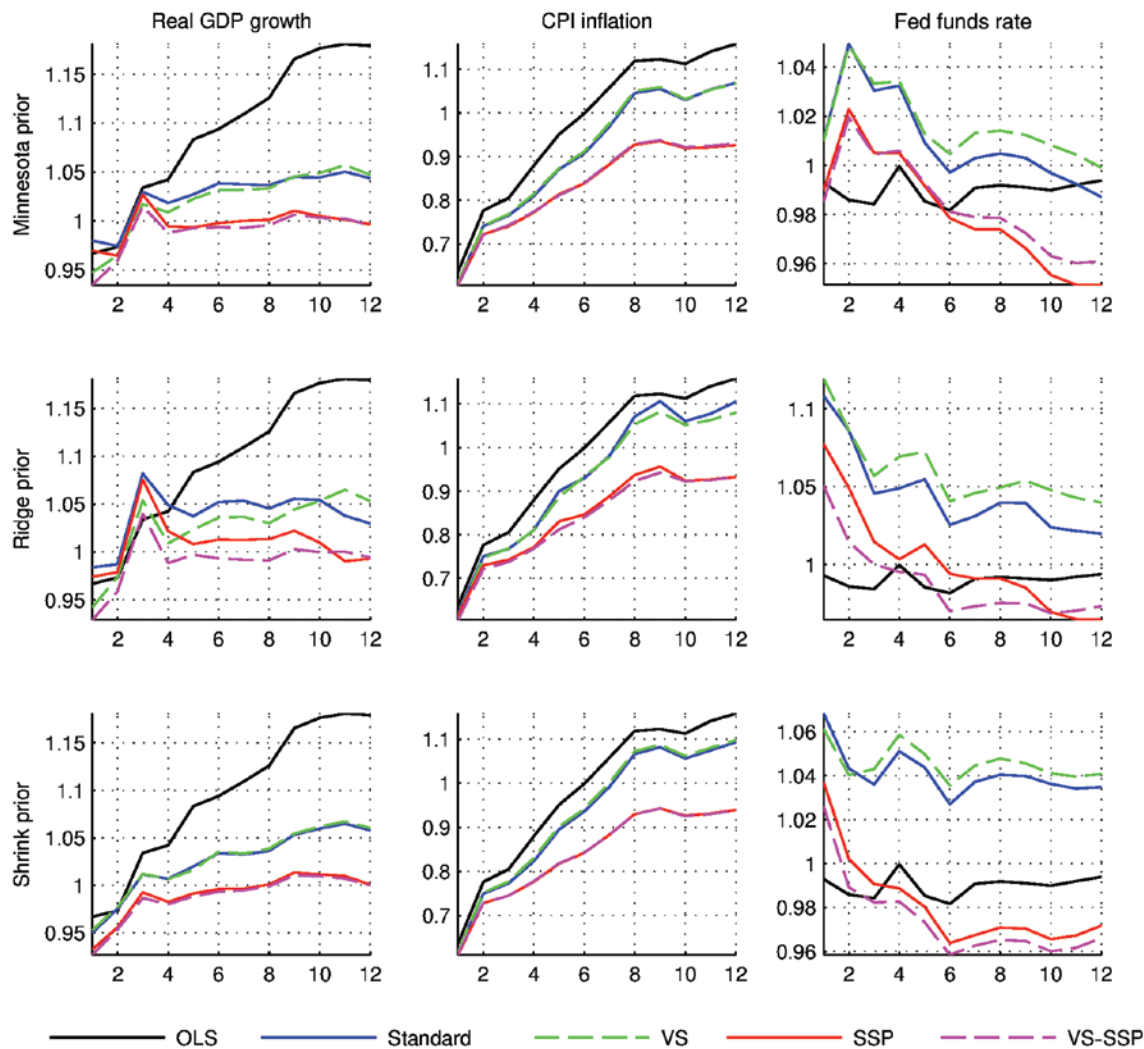


Figure 6: Relative RMSFE results for models using variable selection and variables in a non-gap form.

Notes: The figures show the relative RMSFE metric as a function of the forecasting horizon for three alternative priors for the VAR dynamic coefficients, i.e. Minnesota, Ridge and Shrink priors. OLS is a standard VAR(1) estimated using ordinary least squares. Standard is a BVAR(4) model, VS is BVAR(4) with variable selection *à la* Korobilis (2013), SSP is a BVAR(4) with steady-state prior and VS-SSP is a steady-state BVAR(4) model with variable selection. Inflation and Federal Funds rate are specified in a non-gap form.

the (i) Standard, (iii) SSP, (v) SSVS and (vi) SSVS-SSP models.¹⁰ The forecasting results from an OLS VAR(1) are always included in the figures for comparison reasons, while each row of the figure corresponds to a different prior specification with regard to the dynamic coefficients.¹¹

Beginning from Figures 6 and 7 we see that, overall, the empirical findings indicate that the proposed model, i.e. the steady-state VAR model with variable selection (VS-SSP) (purple dashed line) is the best performing model in the majority of the cases examined here, followed closely by the steady-state VAR (SSP) (red solid line). In some cases the differences between the two models are actually negligible, e.g. see the Inflation

¹⁰ See also Section 3.1 for a description of the models.

¹¹ The rationale behind this grouping is twofold: First, it facilitates the assessment of the proposed specification, i.e. steady state VARs with variable selection, against its counterparts. Recall that our main objective is to investigate whether the combination of steady-state priors and variables selection techniques can lead to superior forecasting behavior and not the comparison of the alternative variable selection techniques per se. Second, from a practical point of view it improves the readability of the figures.

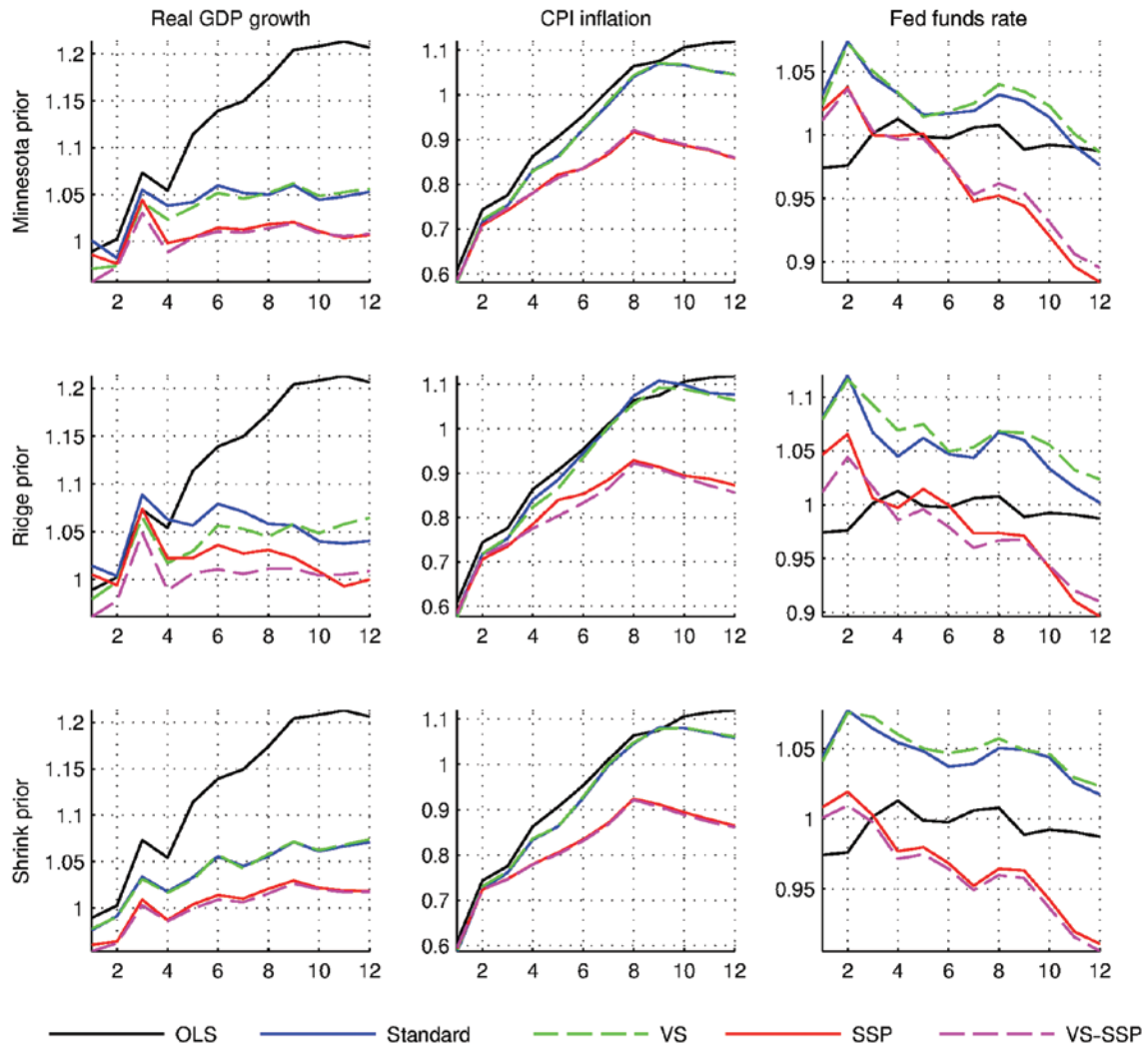


Figure 7: Relative MAPE results for models using variable selection and variables in a non-gap form.

Notes: The figures show the relative MAPE metric as a function of the forecasting horizon for three alternative priors for the VAR dynamic coefficients, i.e. Minnesota, Ridge and Shrink priors. OLS is a standard VAR(1) estimated using ordinary least squares. Standard is a BVAR(4) model, VS is BVAR(4) with variable selection *à la* Korobilis (2013), SSP is a BVAR(4) with steady-state prior and VS-SSP is a steady-state BVAR(4) model with variable selection. Inflation and Federal Funds rate are specified in a non-gap form.

forecasts for the Minnesota and Shrink priors; however, we should note that the SSP model clearly outperforms the proposed model only in the case of the interest rate variable for the Minnesota and Ridge priors and for horizons greater than six (6) and ten (10), respectively. These results also align with other studies that highlight the contribution of steady-state priors to the accuracy of macroeconomic forecasts (Beechey and Österholm 2008; Villani 2009; Clark 2011; Wright 2013).

Moreover, it is also evident that the degree of improvement for the VS-SSP over the SSP model clearly depends on the prior used for the dynamic coefficients and the level of information it carries (see also Korobilis 2013). Thus, the improvement is substantial when we use the relative uninformative ridge prior (the figures in the middle row) and marginal or even negligible when we use priors that are more informative (Shrink or Minnesota priors), as expected.

As regards the rest of the models, their ranking for the GDP and Inflation variables is qualitatively similar with the VS-SSP and SSP models being followed by the VS (green dashed line) and Standard (blue solid line) Bayesian VAR models. Again, we find that, overall, the variable selection improves forecasting

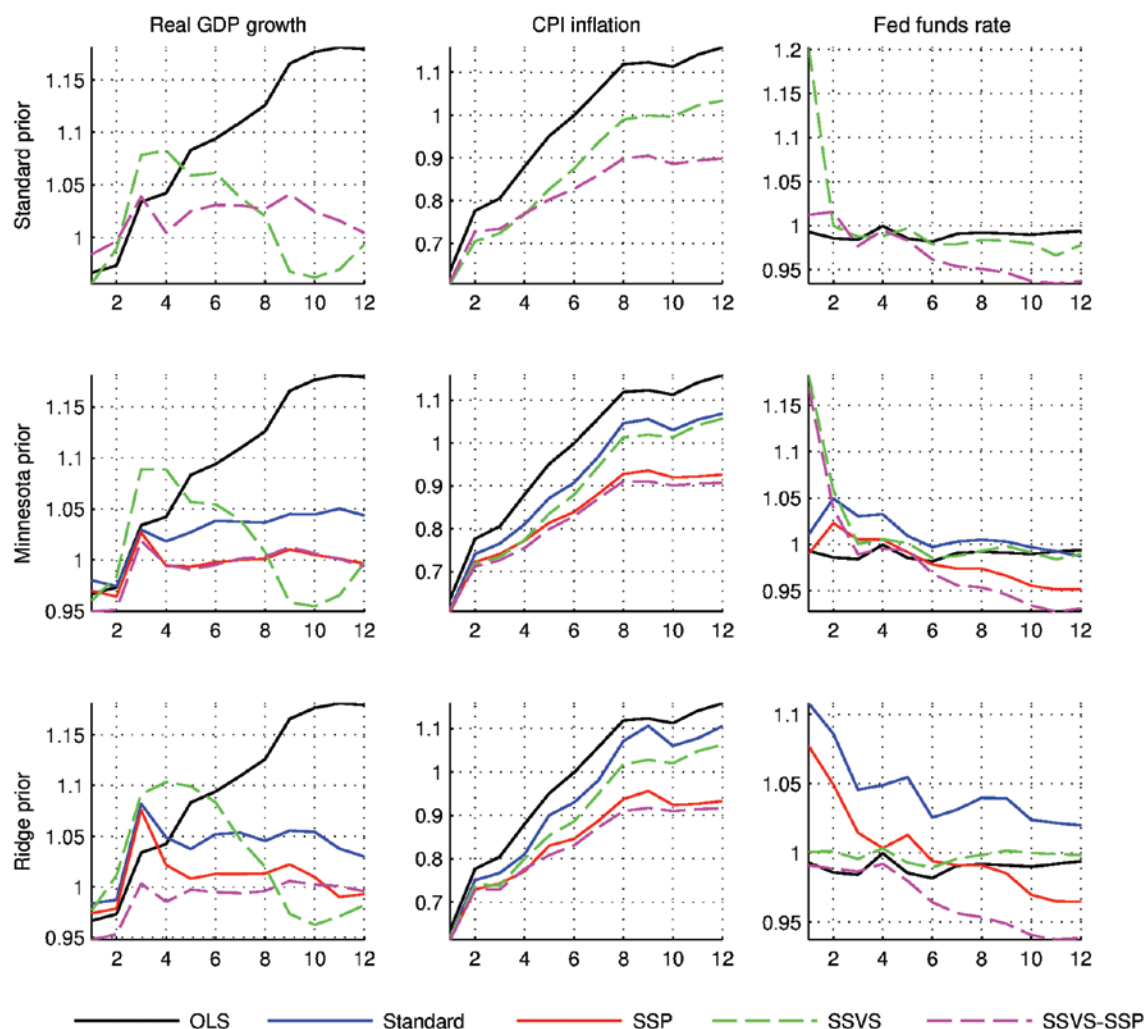


Figure 8: Relative RMSFE results for models using stochastic search variable selection and variables in a non-gap form. Notes: The figures show the relative RMSFE metric as a function of the forecasting horizon for three alternative priors for the VAR dynamic coefficients, i.e. Standard, Minnesota and Ridge priors. Standard prior is specified as in George, Sun, and Ni (2008). OLS is a standard VAR(1) estimated using ordinary least squares. SSVS is a BVAR(4) model with stochastic search variable selection *à la* George, Sun, and Ni (2008) and SSVS-SSP is a BVAR(4) with steady-state prior and stochastic search variable selection. Inflation and Federal Funds rate are specified in a non-gap form.

accuracy mainly when it is combined with less informative priors, while the OLS model is constantly the worst performer. By contrast, the forecasting results for the interest rates reveal that steady-state VAR models with or without variable selection (SSP or VS-SSP) and the OLS model are the main competitors, since they produce overall the best forecasts. In particular, Bayesian models (SSP or VS-SSP) outperform the OLS model for forecasting horizons longer than four or five quarters, across evaluation metrics. In line with the literature, a RW process also provides good short-term forecasts for the interest rates (e.g. see Villani 2009; Banbura, Giannone, and Reichlin 2010; Korobilis 2013; Giannone, Lenza, and Primiceri 2015).¹²

¹² In general, as regards interest rates, BVARs tend to produce better forecasts than the RW benchmark or the OLS model when the priors assume a random walk process for the raw variable (e.g. see Giannone, Lenza, and Primiceri 2015) or a white noise process for the first differences (e.g. see Koop 2013). Using the raw interest rate and specifying a prior that assumes stationarity (white noise or an AR(1) with an 0.8 autoregression parameter as in our case) may lead to inferior forecasts. Nevertheless, as pointed out in Section 2 the stationarity assumption is necessary for the mean-adjusted VAR models.

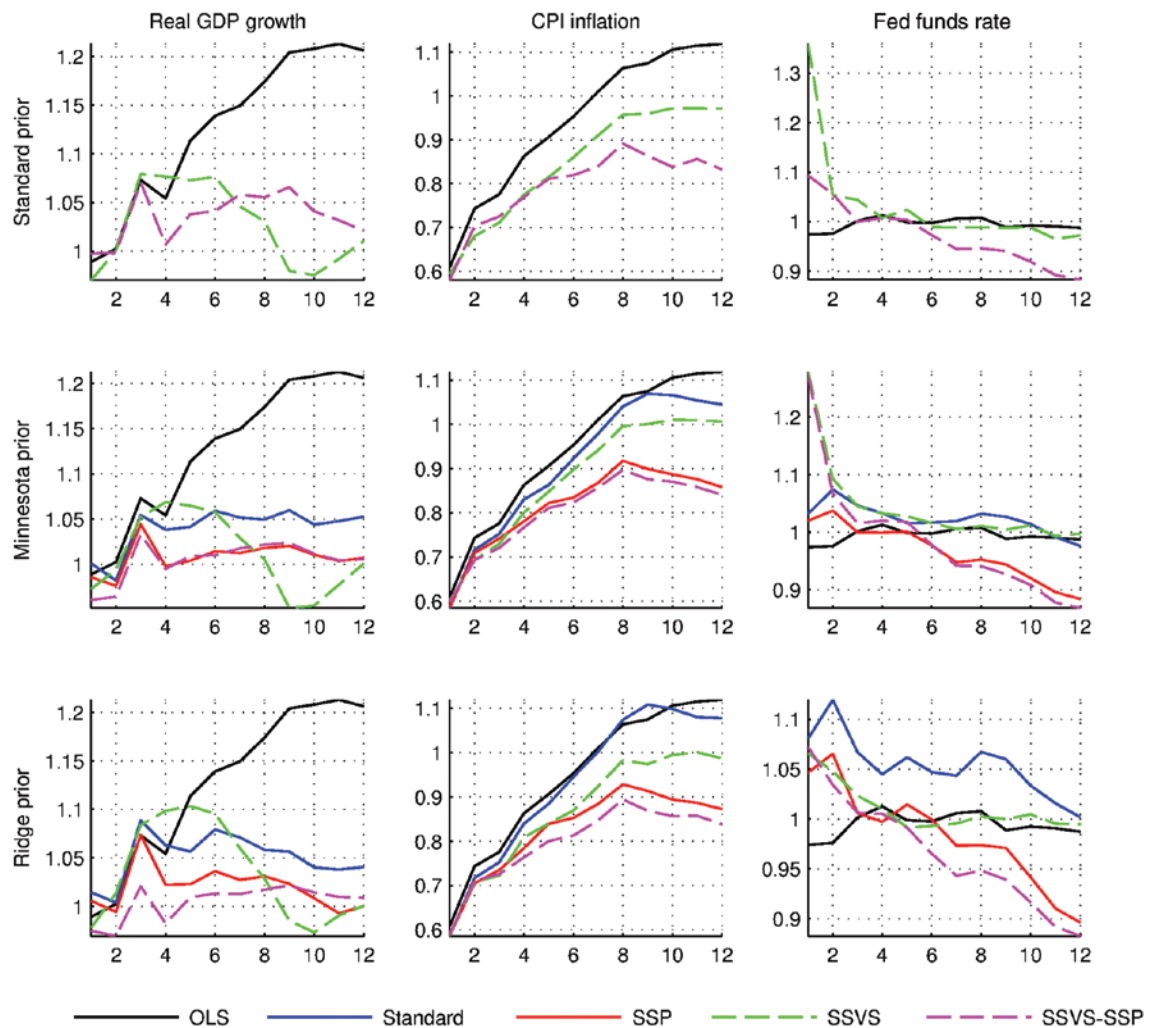


Figure 9: Relative MAFE results for models using stochastic search variable selection and variables in a non-gap form.

Notes: The figures show the relative MAFE metric as a function of the forecasting horizon for three alternative priors for the VAR dynamic coefficients, i.e. Standard, Minnesota and Ridge priors. Standard prior is specified as in George, Sun, and Ni (2008). OLS is a standard VAR(1) estimated using ordinary least squares. SSVS is a BVAR(4) model with stochastic search variable selection *à la* George, Sun, and Ni (2008) and SSVS-SSP is a BVAR(4) with steady-state prior and stochastic search variable selection. Inflation and Federal Funds rate are specified in a non-gap form.

We now turn to the evaluation of the models that adopt the stochastic search variable selection of George, Sun, and Ni (2008) and are presented in Figures 8 and 9. The overall picture is more or less the same with the one presented in Figures 6 and 7, with the proposed specification, i.e. the SSVS-SSP model, being the overall best performing model. In particular, the SSVS-SSP is rarely outranked by some of its competitors, with the most striking examples being the GDP and the interest rate forecasts for horizons greater than eight (8) and smaller than four (4) quarters, where the proposed model is usually beaten by the SSVS and the OLS models respectively.

Moreover, in line with the findings of Koop (2013), we also find that the SSVS model tends to improve forecasting accuracy over the OLS and Standard Bayesian models for most of the cases examined here. Furthermore, the general ranking of the models is not that different from the one reported for Figures 6 and 7 with the most remarkable exception being the interest rate variable, where the SSVS model consistently outranks the Standard BVAR model (while the corresponding VS model in Figures 6 and 7 does not). Finally, a visual inspection of Figures 6–9 also reveal that, in general, SSVS models offer greater forecasting improvements compared to the VS models. A formal analysis for this issue as well as a comparison among all alternative specifications is provided in Section 3.3.3.

3.3.2 Forecasting results using variables in a gap form

This section discusses the forecasting results produced by VAR models that employ the inflation and interest rate variables in gap form as described in Section 3. Figures 10 and 11 present the forecasting results for the relative RMSFE and MAFE evaluation metrics for the models using the variable selection technique *à la* Korobilis (2013), while Figures 12 and 13 present the forecasting results for the models using the variable selection technique *à la* George, Sun, and Ni (2008).

The overall picture remains relatively unchanged with respect to the results presented in Figures 6–9 for the models using variables in a non-gap form. Again the steady-state VARs combined with both types of variable selection can lead to forecasting improvements across variables, evaluation metrics and priors. In addition, we observe that the adoption of variable selection techniques lead to significant forecasting improvements only in the case of relatively uninformative priors (e.g. the ridge prior), as expected, while the incorporation of steady-state priors is also usually beneficial for forecasting accuracy.

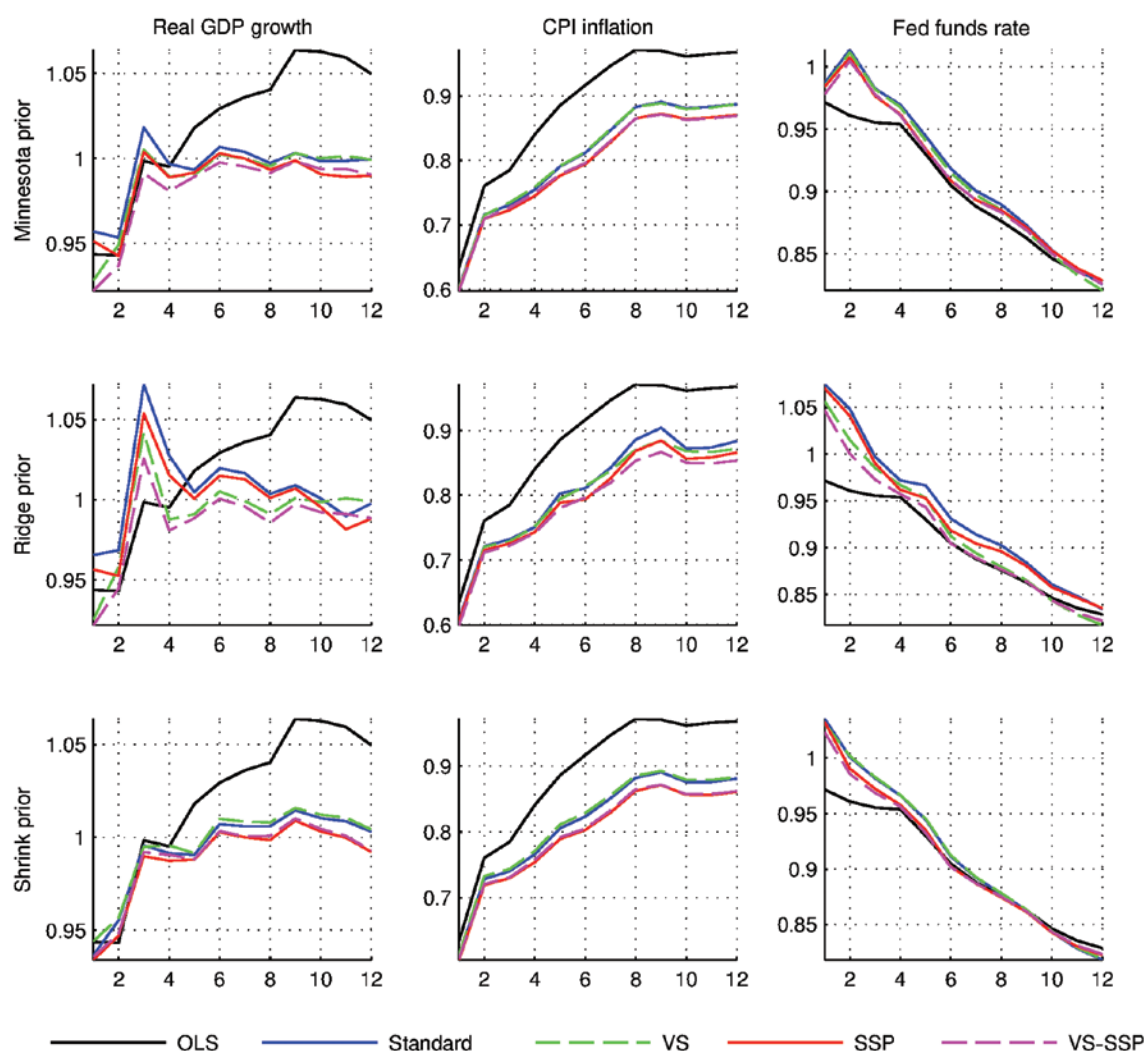


Figure 10: Relative RMSFE results for models using variable selection and variables in a gap form.

Notes: The figures show the relative RMSFE metric as a function of the forecasting horizon for three alternative priors for the VAR dynamic coefficients, i.e. Minnesota, Ridge and Shrink priors. OLS is a standard VAR(1) estimated using ordinary least squares. Standard is a BVAR(4) model, VS is BVAR(4) with variable selection *à la* Korobilis (2013), SSP is a BVAR(4) with steady-state prior and VS-SSP is a steady-state BVAR(4) model with variable selection. Inflation and Federal Funds rate are specified in a gap form.

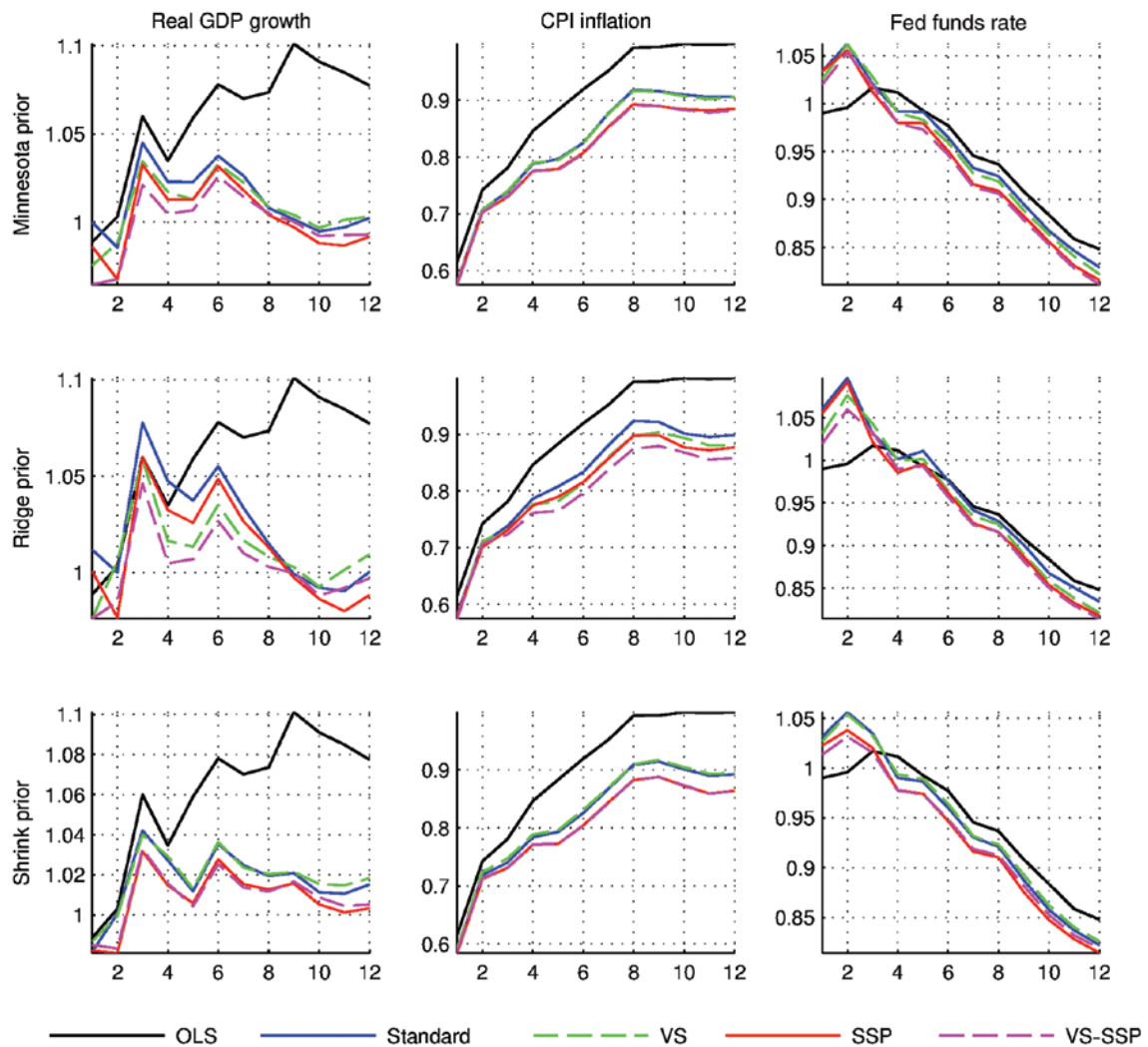


Figure 11: Relative MAFE results for models using variable selection and variables in a gap form.

Notes: The figures show the relative MAFE metric as a function of the forecasting horizon for three alternative priors for the VAR dynamic coefficients, i.e. Minnesota, Ridge and Shrink priors. OLS is a standard VAR(1) estimated using ordinary least squares. Standard is a BVAR(4) model, VS is BVAR(4) with variable selection *à la* Korobilis (2013), SSP is a BVAR(4) with steady-state prior and VS-SSP is a steady-state BVAR(4) model with variable selection. Inflation and Federal Funds rate are specified in a gap form.

Nonetheless, we should also note that the difference in the forecasting accuracy between the steady-state models and the models with uninformative steady-state priors has been mitigated, especially for the inflation and interest rate variables, compared to the models using non-gap variables (see Figures 6–9). This probably reflects the forecasting patterns presented in the right column of Figures 3–5 and show that the misestimation of steady-state levels from the Bayesian models with uninformative steady-state priors is not that severe compared to the models using non-gap variables (left column of Figures 3–5).

A possible explanation is that in the case of models with variables in a gap form, the two components of the variables, i.e. the smoothed and the gap component, are modeled separately. In particular, since the low-frequency component forecasts are common across all models, the potential improvement for the steady-state models over the models with uninformative steady-state comes exclusively from the gap component of the variable. This is obviously a smaller portion of the final number to be forecasted and consequently, the prospective improvement would also be analogously smaller.

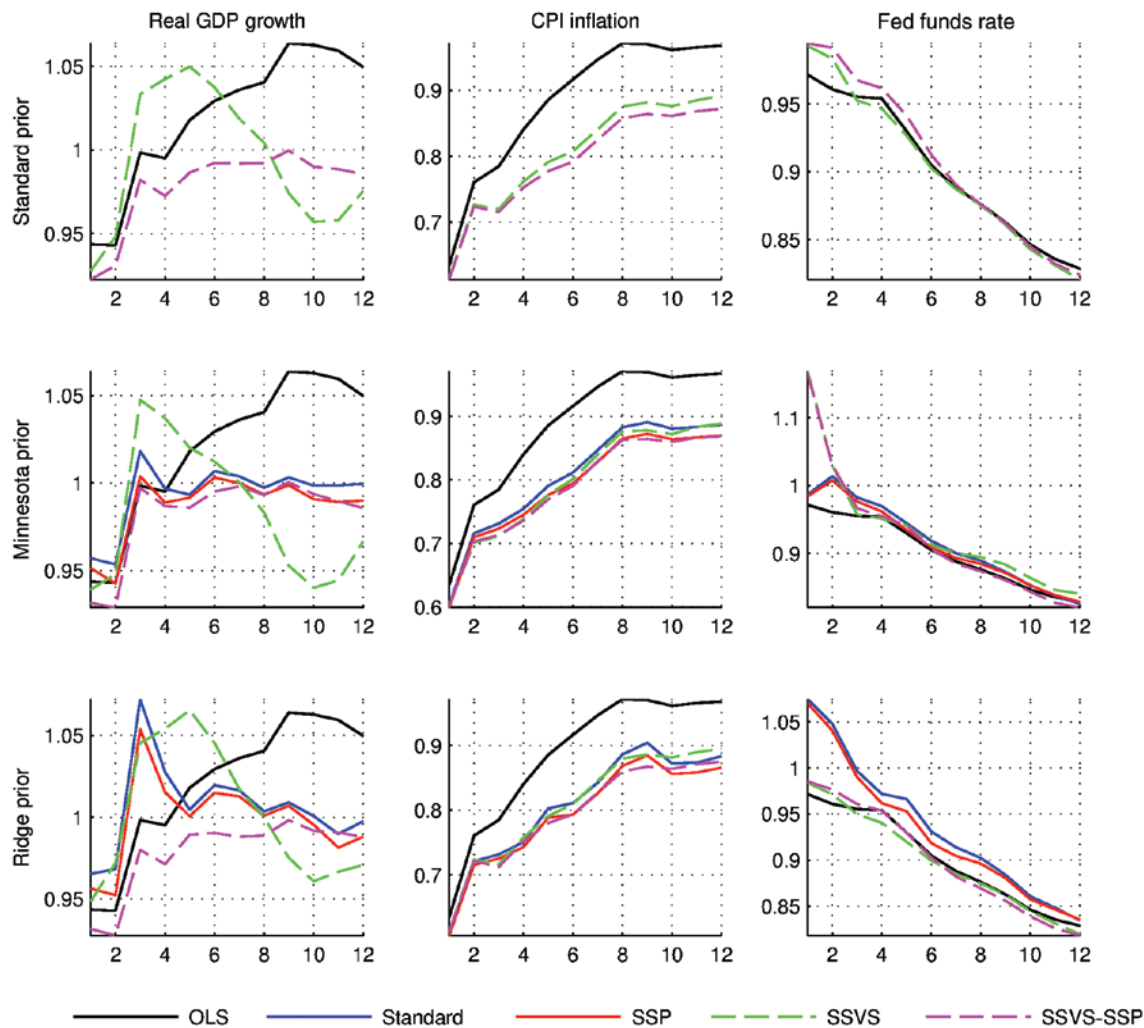


Figure 12: Relative RMSFE results for models using stochastic search variable selection and variables in a gap form. Notes: The figures show the relative RMSFE metric as a function of the forecasting horizon for three alternative priors for the VAR dynamic coefficients, i.e. Standard, Minnesota and Ridge priors. Standard prior is specified as in George, Sun, and Ni (2008). OLS is a standard VAR(1) estimated using ordinary least squares. SSVS is a BVAR(4) model with stochastic search variable selection *à la* George, Sun, and Ni (2008) and SSVS-SSP is a BVAR(4) with steady-state prior and stochastic search variable selection. Inflation and Federal Funds rate are specified in a gap form.

3.3.3 Model confidence set results

In this section, we employ the Model Confidence Set (MCS) method of Hansen, Lunde, and Nason (2003, 2011) and we construct a set of models, $M_{1-\alpha}^* \subseteq M_0$, that present statistically superior predictive ability at a given confidence level. The scope of this analysis is twofold: first, based on the RMSFE and MAFE metrics we discern that set of models that generate statistically significant superior forecasts and second, we evaluate the forecasting ability of the full set of models across priors and specifications. Next, we briefly describe the MCS methodology.

Assuming an initial set of $M=M_0$ models, the MCS method is based on a specific loss function, $L_{m,t}$ with $m=1, \dots, M$, and applies an iterative process of sequential Equal Predictive Ability (EPA) tests of the form:

$$H_{0,M_0}: E(d_{mk,t})=0 \text{ for all } m, k \in M \quad (16)$$

where $d_{mk,t}=L_{m,t}-L_{k,t}$ is the loss differential between models m and k and $L_{\bullet,t}$ is one of the RMSFE or MAFE at each point in time, t . A rejection of the null hypothesis indicates that a model has inferior predictive

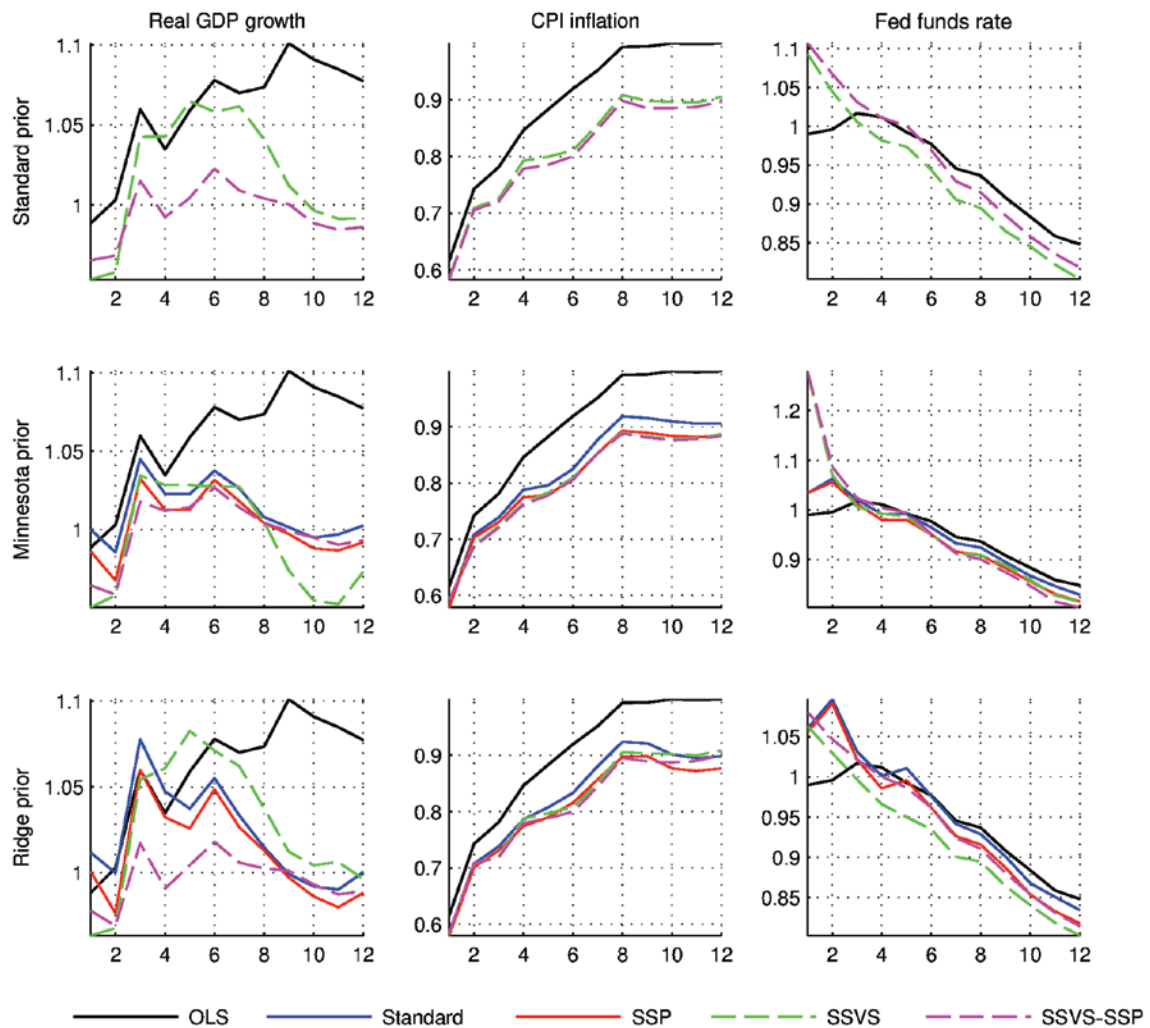


Figure 13: Relative MAFE results for models using stochastic search variable selection and variables in a gap form.

Notes: The figures show the relative MAFE metric as a function of the forecasting horizon for three alternative priors for the VAR dynamic coefficients, i.e. Standard, Minnesota and Ridge priors. Standard prior is specified as in George, Sun, and Ni (2008). OLS is a standard VAR(1) estimated using ordinary least squares. SSVS is a BVAR(4) model with stochastic search variable selection *à la* George, Sun, and Ni (2008) and SSVS-SSP is a BVAR(4) with steady-state prior and stochastic search variable selection. Inflation and Federal Funds rate are specified in a gap form.

ability and should not be included in the MCS at an α significance level. The EPA test in (16) is repeated for the remaining $M_{1-\alpha}$ models, with $M_{1-\alpha} \subset M$, and this procedure continues until the null hypothesis cannot be rejected. The final set of surviving models forms the MCS at a $1-\alpha$ confidence level, denoted by $M_{1-\alpha}^*$. The models included in the MCS have equal predictive ability, but they outperform the eliminated models, while the MCS p -values indicate the probability of a model being a member of the MCS.¹³

Table 4 presents only a synopsis of the main MCS results for four selected forecasting horizons, i.e. $h=1, 4, 8$ and 12 , while we refer the interested reader to Tables C.1–C.4 of Appendix C for a more detailed analysis. The first and second column for each of the variables in Table 4 presents the percentage of times a model is included in the MCS at 10% significance level across the selected forecasting horizons and evaluation metrics for models using variables in a gap and non-gap form, respectively. The third column repeats the MCS results

¹³ For details on MCS technique and its implementation see Hansen, Lunde, and Nason (2003, 2011). The MCS is implemented using MULCOM 2.00 package for Ox, kindly provided by the authors. The MULCOM 2.00 package is available at http://mit.econ.au.dk/vip_hm/alunde/mulcom/mulcom.htm.

Table 4: Synopsis of the Model confidence set (MCS) results.

Models\variables	GDP			CPI			FFR			Avg.		
	Gap	Non-gap	Total	Gap	Non-gap	Total	Gap	Non-gap	Total	Gap	Non-gap	Total
1. RW (benchmark)	50	75	63	0	0	0	25	50	38	25	42	34
2. OLS	25	25	25	13	0	06	50	38	44	29	21	25
3. Min-Standard	38	13	25	25	13	19	25	38	31	29	21	25
4. Min-VS	50	38	44	25	25	25	38	38	38	38	34	36
5. Min-SSP	63	63	63	50	50	50	50	50	50	54	54	54
6. Min-VS-SSP	88	100	94	38	50	44	75	50	63	67	67	67
7. Ridge-standard	25	13	19	25	13	19	0	38	19	17	21	19
8. Ridge-VS	63	38	50	25	25	25	25	0	13	38	21	29
9. Ridge-SSP	50	25	38	50	38	44	25	38	31	42	34	38
10. Ridge-VS-SSP	88	100	94	100	63	81	50	50	50	79	71	75
11. Shrink-standard	25	38	31	13	13	13	25	13	19	21	21	21
12. Shrink-VS	25	38	31	13	13	13	38	13	25	25	21	23
13. Shrink-SSP	50	63	56	50	38	44	75	50	63	58	50	54
14. Shrink-VS-SSP	38	100	69	63	38	50	75	75	75	59	71	65
15. SSVS	50	50	50	13	50	31	75	63	69	46	54	50
16. SSVS-SSP	100	25	63	50	100	75	0	88	44	50	71	61
17. Min-SSVS	75	63	69	63	13	38	38	25	31	59	34	46
18. Min-SSVS-SSP	100	75	88	63	100	81	63	75	69	75	83	79
19. Ridge-SSVS	63	50	56	13	13	13	88	38	63	55	34	44
20. Ridge-SSVS-SSP	88	100	94	38	88	63	50	88	69	59	92	75

This table shows the percentage (%) of times a model is included in the MCS at 10% significance level across evaluation metrics and forecasting horizons. Bold faced numbers indicate the models with the highest (%) of times in the MCS. RW is random walk model with drift in levels. OLS is a standard VAR(1) estimated using ordinary least squares. Min, Ridge and Shrink are the Minnesota, Ridge and Shrink priors for the dynamic coefficients, respectively. Standard is a BVAR(4) model, VS is BVAR(4) with variable selection *à la* Korobilis (2013), SSP is a BVAR(4) with a steady-state prior, VS-SSP is a steady-state BVAR(4) model with variable selection, SSVS is a BVAR(4) with stochastic search variable selection *à la* George, Sun, and Ni (2008) and SSVS-SSP is a BVAR(4) with steady-state prior and stochastic search variable selection. For the SSVS (15) and SSVS-SSP (16) models the prior on the dynamic coefficients is that of George, Sun, and Ni (2008).

across all (gap and non-gap) models.¹⁴ The last three columns of Table 4, named Avg., show the average percentage of times a model is included in the MCS across all three variables. The initial set includes all 20 models: the RW and OLS models, and all 18 BVARs.

In general, the synopsis of the MCS results in Table 4 confirms the graphical evidence presented in Sections 3.3.1 and 3.3.2, but also allows for several other comparisons among models that were not feasible before. The main conclusions emerge from Table 4 are discussed below.¹⁵

First, except for the interest rate variable for models with gap variables, in all other cases the best performing model (denoted with bold faced entries), i.e. the model that maximizes the percentage of times included in the MCS, is a steady-state model with some kind of variable selection (VS or SSVS). Thus, we can state that, in general, the proposed specification which incorporates steady-state prior beliefs in a VAR model with Bayesian variable selection enhances the macroeconomic forecasting accuracy of small-scale VARs irrespective of the prior used on the dynamic parameters.

Second, the empirical evidence also confirms that, on average, there is an unambiguous ranking of the various Bayesian VAR models (see Avg. columns).¹⁶ Particularly, Standard BVARs are usually outranked by models with variable selection which, in turn, are consistently outperformed by BVARs with an informative

¹⁴ This is practically the average of Gap and Non-gap columns.

¹⁵ The discussion of the results is based mainly on Avg. columns which effectively summarize the key findings of the paper.

¹⁶ Here, the ranking of the models is based on the percentage of times they are included in the MCS. Obviously, higher percentages indicate better forecasting models.

steady-state prior. Of course, as already mentioned, models that combine both steady-state priors and variable selection outrank all the above mentioned specifications since the percentage of times included in the MCS is consistently higher compared to other specifications.

Third, the synopsis of the results presented in Table 4 give us the opportunity to compare the two alternative variable selection techniques, i.e. the variable selection as implemented in Korobilis (2013) and the stochastic search variable selection of George, Sun, and Ni (2008). Overall, the SSVS models tend to outperform the VS models irrespective of whether they incorporate a steady-state prior or not. More specifically, the average percentage of times a VS or a VS-SSP model is part of the MCS is 29% and 69%, while the corresponding averages for the SSVS and SSVS-SSP models are 47% and 72%, respectively.

Fourth, evidence presented in Tables C.1–C.4 shows that controlling for the slowly moving local mean for inflation is beneficial for forecasting purposes as it has recurrently mentioned in the extant literature (e.g. see Faust and Wright 2013 and references therein). Indeed, the average relative RMSFE and MAFE metrics for inflation and interest rate are lower for models with variables in a gap form relative to models with raw variables. However, the difference in the forecasting accuracy for gap and non-gap models is greater when we use models with uninformative steady-state priors.

Finally, the overall best performing model is the SSVS-SSP model combined with the Minnesota prior which is included in the MCS in 79% of cases (see column Avg., Total) followed closely by the SSVS-SSP and VS-SSP models with a Ridge prior, which are both part of the MCS in 75% of cases. Therefore, empirical evidence also suggests that the simple and relatively uninformative ridge prior can also provide accurate forecasts when combined with variable selection techniques, compared to the more informative Minnesota or Shrink priors.

4 Conclusions

Empirical evidence in extant literature has highlighted the importance of steady-state prior beliefs in Bayesian VAR forecasting. Moreover, Bayesian variable selection techniques have also been suggested as an efficient and automatic way of VAR shrinkage and model parsimony with beneficial effects on forecasting accuracy. In this paper, we propose Gibbs sampler algorithms for estimating Bayesian VAR models that efficiently combine this two promising VAR specifications: informative priors on the steady state of the system and parameter restrictions based on Bayesian variable selection methods. We evaluate the proposed specification in terms of an out-of-sample forecasting exercise using three major US macroeconomic variables and we find that it clearly outperforms alternative VAR models that encapsulate only one (or none) of the above-mentioned specifications. Empirical evidence also suggests that these results are robust against alternative priors on dynamic parameters that carry different degree of information regarding VAR shrinkage.

Acknowledgments: I gratefully acknowledge two anonymous reviewers, Bruce Mizrach (the Editor-in-Chief), Gary Koop (the Associated Editor) and Heather Gibson for their constructive and insightful comments and suggestions that considerably improved this article.

Disclaimer: The views expressed in this article do not necessarily represent Bank of Greece.

Appendix A: Posterior inference in the steady-state VARs with Bayesian variable selection

As already mentioned, conditional on the steady-state parameters, φ , the VAR models described in Section 2 are standard VAR models with (stochastic search) variable selection for the mean-adjusted series \tilde{y}_t . Therefore, the first five and three blocks of the Gibbs samplers presented here, reproduce the George,

Sun, and Ni (2008) and Korobilis (2013) algorithm for the mean-adjusted series, respectively. The last extra block samples the steady-state parameters, φ , conditional on all other parameters, from a normal posterior density. Conditional on β and γ (or equivalently θ for the variable selection method) the model is a standard steady-state VAR model and can be written in a form so that we can apply Villani's (2009) methodology.

Stochastic search variable selection à la George, Sun, and Ni (2008)

Next, we derive the full conditional posteriors of the six-block Gibbs sampler algorithm for the steady-state VAR with stochastic search variable selection which employ the priors of Section 2.1. The notation used in this appendix is identical to the one used in Section 2.1. More specifically, the Markov Chain Monte Carlo (MCMC) algorithm comprises sequential drawing from the posterior distribution using the following steps.

Step 1. Draw ψ_{jj}^2 , $j=1, \dots, m$, independently, from the following Gamma density:

$$\psi_{jj}^2 | \beta, \varphi, \eta; Y \sim \text{gamma}(\bar{a}_j, \bar{b}_j) \quad (\text{A.1})$$

where $\bar{a}_j = \underline{a}_j + \frac{1}{2}T$ and \bar{b}_j is given by:

$$\bar{b}_j = \begin{cases} b_1 + \frac{1}{2}v_{11} & \text{if } j=1 \\ b_{jj} + \frac{1}{2}\{v_{jj} - v_j'[V_{j-1} + (D_j D_j)^{-1}]^{-1}v_j\} & \text{if } j=2, \dots, m \end{cases}$$

where $V = (\tilde{Y} - \tilde{X}B)'(\tilde{Y} - \tilde{X}B)$ with v_{ij} elements, V_j is the upper left $j \times j$ block of V and $v_j = (v_{1j}, \dots, v_{j-1,j})'$.

Step 2. Draw η_i , $i=2, \dots, m$, independently of one another, from the following $(i-1)$ -dimensional multivariate Normal distribution

$$\eta_i | \beta, \omega, \varphi, \psi; Y \sim N_{i-1}(\bar{h}_i, \bar{\Delta}_i) \quad (\text{A.2})$$

where $\bar{\Delta}_i = [V_{i-1} + (D_i D_i)^{-1}]^{-1}$ and $\bar{h}_i = -\psi_{ii} \bar{\Delta}_i v_i$.

Step 3. Draw ω_{ji} , $i=2, \dots, m$ and $j=1, \dots, i-1$, independently of one another, from the following Bernoulli distribution:

$$\omega_{ji} | \omega_{\setminus ji}, \beta, \varphi, \psi; Y \sim \text{Bernoulli}(\bar{q}_{ji}) \quad (\text{A.3})$$

$$\text{where } \bar{q}_{ji} = \left[\frac{1}{\kappa_{0ji}} \exp\left(-\frac{\eta_{ji}^2}{2\kappa_{0ji}^2}\right) \underline{q}_{ji} \right] / \left[\frac{1}{\kappa_{0ji}} \exp\left(-\frac{\eta_{ji}^2}{2\kappa_{0ji}^2}\right) \underline{q}_{ji} + \frac{1}{\kappa_{1ji}} \exp\left(-\frac{\eta_{ji}^2}{2\kappa_{1ji}^2}\right) (1 - \underline{q}_{ji}) \right].$$

Step 4. Draw β from the following n -dimensional multivariate Normal distribution:

$$\beta | \gamma, \varphi, \Psi; Y \sim N_n(\bar{\beta}, \bar{V}) \quad (\text{A.4})$$

where $\bar{\beta} = \bar{V}[(DD)^{-1}\underline{b} + \text{vec}(\tilde{X}'\tilde{Y}\Sigma^{-1})]$ and $\bar{V} = [\Sigma^{-1} \otimes (\tilde{X}'\tilde{X}) + (DD)^{-1}]^{-1}$.

Step 5. Draw γ_j , $j=1, \dots, n$, independently of one another, from the following Bernoulli distribution:

$$\gamma_j | \gamma_{\setminus j}, \beta, \varphi, \Psi; Y \sim \text{Bernoulli}(\bar{\pi}_j) \quad (\text{A.5})$$

$$\text{where } \bar{\pi}_j = \left[\frac{1}{\tau_{0j}} \exp\left(-\frac{\beta_j^2}{2\tau_{0j}^2}\right) \underline{\pi}_j \right] / \left[\frac{1}{\tau_{0j}} \exp\left(-\frac{\beta_j^2}{2\tau_{0j}^2}\right) \underline{\pi}_j + \frac{1}{\tau_{1j}} \exp\left(-\frac{\beta_j^2}{2\tau_{1j}^2}\right) (1 - \underline{\pi}_j) \right].$$

Step 6. Draw the steady-state coefficients φ from the following mq -dimensional multivariate Normal density

$$\varphi | \beta, \Psi; Y \sim N_{mq}(\bar{b}_\varphi, \bar{V}_\varphi) \quad (\text{A.6})$$

For this step the VAR model conditional on β is written as follows:

$$(y_t - \varphi d_t) = B_1(y_{t-1} - \varphi d_{t-1}) + \dots + B_p(y_{t-p} - \varphi d_{t-p}) + \varepsilon_t \quad (\text{A.7})$$

Rearranging the terms in Eq. (A.7) we get:

$$B(L)y_t = B(L)\varphi d_t + \varepsilon_t = \varphi d_t - B_1\varphi d_{t-1} - \dots - B_p\varphi d_{t-p} + \varepsilon_t \quad (\text{A.8})$$

where $B(L) = I_m - B_1L - \dots - B_pL^p$. Following Villani (2009) we define Y_φ as a $T \times m$ matrix with its t -th row being $Y_{\varphi t} = (B(L)y_t)'$, D as a $T \times (p+1)q$ matrix which t -th row is given by $D_t = (d_t', -d_{t-1}', \dots, -d_{t-p}')$ and $\Lambda = [\varphi, B_1\varphi, \dots, B_p\varphi]'$ as the $(p+1)q \times m$ matrix of coefficients. Therefore, the model in (A.8) can be written as $Y_\varphi = D\Lambda + E$ and standard results for multivariate regressions can be applied (see Villani 2009, 646–647). Thus, given that $\text{vec}(\Lambda') = \text{Uvec}(\varphi)$ with

$$U = \begin{pmatrix} I_{mq} \\ I_q \otimes B_1 \\ \vdots \\ I_q \otimes B_p \end{pmatrix}$$

we define the mean, \bar{b}_φ , and variance, \bar{V}_φ , of the posterior distribution in (A.6) as $\bar{b}_\varphi = \bar{V}_\varphi (U' \text{vec}(\Sigma^{-1} Y' D) + \bar{V}_\varphi^{-1} \bar{b}_\varphi)$ and $\bar{V}_\varphi^{-1} = U' (D' D \otimes \Sigma^{-1}) U + \bar{V}_\varphi^{-1}$ respectively.

Variable selection à la Korobilis (2013)

We follow Korobilis (2013) and we re-write the VAR model in a vectorized form to exploit the computational efficiency of matrix multiplications. To this end, we define \tilde{y} as a $T \times m$ matrix with its t -th row being $\tilde{Y}_t = \tilde{y}'_t = (\tilde{y}_{1t}, \tilde{y}_{2t}, \dots, \tilde{y}_{mt})$, \tilde{X} as an $T \times k$ matrix with its t -th row being $\tilde{X}_t = \tilde{x}'_t$ and E as a $T \times m$ with its t -th row being $E_t = \varepsilon'_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{mt})$. Then, $\tilde{y} = \text{vec}(\tilde{y})$ and $\varepsilon = \text{vec}(E)$ are $Tm \times 1$ vectors with $\text{vec}(\cdot)$ being the standard operator which stacks the columns of a matrix. Then, the VAR model can be written as

$$\tilde{y} = \tilde{Z} \theta + \varepsilon \quad (\text{A.9})$$

$(Tm \times 1) \quad (Tm \times n) \quad (n \times 1) \quad (Tm \times 1)$

where $\tilde{Z} = I_m \otimes \tilde{X}$ is a $Tm \times n$ block diagonal matrix with \tilde{X} being repeatedly on its main diagonal, $\theta = \Gamma\beta$ is the $n \times 1$ vector of coefficients with $\beta = \text{vec}(B)$. To facilitate reading, we remind that m is the number of variables, $k = mp$ is the number of explanatory variables in each VAR equation, and $n = mk = m^2p$ is the total number of VAR coefficients. Also recall that $\tilde{y}_t = y_t - \varphi d_t$ and $\tilde{x}_t = (\tilde{y}'_{t-1}, \dots, \tilde{y}'_{t-p})$ are the mean-adjusted time-series of y_t and x_t , respectively.

The full conditional posteriors for the steady-state VAR with variable selection using the priors specified in Section 2.2 are given below. More specifically, the four-block Gibbs sampler algorithm includes the following steps:

Step 1. Draw the slope coefficients β from the following n -dimensional multivariate Normal density

$$\beta | \gamma, \varphi, \Sigma, y; Z \sim N_n(\bar{\beta}, \bar{V}) \quad (\text{A.10})$$

where $\bar{V} = (\bar{V}^{-1} + \tilde{Z}'(\Sigma^{-1} \otimes I_T)\tilde{Z})^{-1}$, $\bar{\beta} = \bar{V}(\bar{V}^{-1}\bar{b} + \tilde{Z}'(\Sigma^{-1} \otimes I_T)\tilde{y})$ and $\tilde{Z}^* = \tilde{Z}\Gamma$.

Step 2. Draw γ_j in random order j with $j = 1, \dots, n$ from

$$\gamma_j | \gamma_{\setminus j}, \beta, \phi, \Sigma; y, Z \sim \text{Bernoulli}(\bar{\pi}_j) \quad (\text{A.11})$$

where $\bar{\pi}_j = l_{0j} / (l_{0j} + l_{1j})$, $l_{0j} = \exp(-0.5(\tilde{y} - \tilde{Z}\theta^*)(\Sigma^{-1} \otimes I_T)(\tilde{y} - \tilde{Z}\theta^*))\pi_j$, $l_{1j} = \exp(-0.5(\tilde{y} - \tilde{Z}\theta^{**})'(\Sigma^{-1} \otimes I_T)(\tilde{y} - \tilde{Z}\theta^{**}))(1 - \pi_j)$ with θ^* and θ^{**} being equal to θ but with their j -th element being equal to β_j and 0, respectively.

Step 3. Draw Σ^{-1} from

$$\Sigma^{-1} | \beta, \gamma, \varphi; y, Z \sim \text{Wishart}(T, S^{-1}) \quad (\text{A.12})$$

For this step, re-write the VAR as

$$\tilde{Y} = \tilde{X} \Theta + E \quad (\text{A.13})$$

$(T \times m) \quad (T \times k)(k \times m) \quad (T \times m)$

where Θ is a $k \times m$ matrix which ij -th element is given by $\Theta_{ij} = \theta_{(j-1)k+i}$ for $i=1, \dots, k$ and $j=1, \dots, m$. Then, S is given by $S = E'E = (\tilde{Y} - \tilde{X}\Theta)'(\tilde{Y} - \tilde{X}\Theta)$.

Step 4. Draw the steady-state coefficients φ from the following m -dimensional multivariate Normal density

$$\varphi | \beta, \gamma, \Sigma; y, Z \sim N_{mq}(\bar{b}_\varphi, \bar{V}_\varphi) \quad (\text{A.14})$$

For this step the VAR model conditional on β and γ (or equivalently θ) is written as follows

$$(y_t - \varphi d_t) = \Theta_1(y_{t-1} - \varphi d_{t-1}) + \dots + \Theta_p(y_{t-p} - \varphi d_{t-p}) + \varepsilon_t \quad (\text{A.15})$$

where Θ_p , $p=1, \dots, p$ is a $m \times m$ matrix of coefficients which ij -th element is given by $\Theta_{l,ij} = \Theta_{(l-1)m+j, i}$ for $l=1, \dots, p$ and $i, j=1, \dots, m$. Rearranging the terms in Eq. (A.15) we get

$$\Theta(L)y_t = \Theta(L)\varphi d_t + \varepsilon_t = \varphi d_t - \Theta_1\varphi d_{t-1} - \dots - \Theta_p\varphi d_{t-p} + \varepsilon_t \quad (\text{A.16})$$

where $\Theta(L) = I_m - \Theta_1 L - \dots - \Theta_p L^p$. Following Villani (2009) I define Y_φ as a $T \times m$ matrix with its t -th row being $Y_{\varphi t} = (\Theta(L)y_t)'$, D as a $T \times (p+1)q$ matrix which t -th row is given by $D_t = (d_t', -d_{t-1}', \dots, -d_{t-p}')'$ and $\Lambda = [\varphi, \Theta_1\varphi, \dots, \Theta_p\varphi]'$ as the $(p+1)q \times m$ matrix of coefficients. Therefore, the model in Eq. (A.16) can be written as $Y_\varphi = D\Lambda + E$ and standard results for multivariate regressions can be applied (see Villani 2009, 646–647). Thus, given that $\text{vec}(\Lambda') = U\text{vec}(\varphi)$ with

$$U = \begin{pmatrix} I_{mq} \\ I_q \otimes \Theta_1 \\ \vdots \\ I_q \otimes \Theta_p \end{pmatrix}$$

we define the mean, \bar{b}_φ , and variance, \bar{V}_φ , of the posterior distribution in (A.14) as $\bar{b}_\varphi = \bar{V}_\varphi (U' \text{vec}(\Sigma^{-1} Y_\varphi' D) + V_\varphi^{-1} \underline{b}_\varphi)$ and $\bar{V}_\varphi^{-1} = U'(D'D \otimes \Sigma^{-1})U + V_\varphi^{-1}$ respectively.

The Gibbs samplers described in this Appendix were implemented in Matlab and extend the Matlab codes for VAR models with variable selection and stochastic search variable selection kindly provided by Gary Koop and Dimitris Korobilis in their websites: personal.strath.ac.uk/gary.koop/bayes_matlab_code_by_koop_and_korobilis.html and <https://sites.google.com/site/dimitriskorobilis/matlab>.

Appendix B: Specifying priors on β

This appendix briefly describes the three different types of prior distribution on β used in the VS and SSVS models (for more details see George, Sun, and Ni (2008), Koop (2013), Korobilis (2013) and references therein). More specifically, for the VS models we use:

- The *Ridge regression prior* which defines $\underline{b} = 0_{n \times 1}$ and $V = \lambda I_n$ with the hyperparameter λ determining the degree of shrinkage on β . We choose $\lambda = 100$ for the intercepts (diffuse prior) and $\lambda = 9$ for the dynamic parameters.
- The popular *Minnesota prior* which assumes that variables follow a AR(1) process implying that all elements of \underline{b} are zero except for the parameter of the first own lag of each of the variables which is equal to δ_i .¹⁷ Here, we set the autoregressive parameter equal to 0.25 for the GDP and 0.8 for the inflation and

¹⁷ Doan, Litterman, and Sims (1984) and Litterman (1986) originally proposed that all variables follow a random walk process meaning that $\delta = 1$. However, if we work with stationary variables we can set $\delta < 1$.

interest rate variables. The prior covariance matrix \underline{V} is assumed to be a $n \times n$ diagonal matrix with each of its diagonal elements being defined as

$$\underline{v}_{ij}^l = \begin{cases} 100\sigma_i^2 & \text{for the intercepts} \\ \lambda_1/l^2 & \text{for own lag parameters} \\ \lambda_1\lambda_2\sigma_i^2/l^2\sigma_j^2 & \text{for the } j\text{th lagged variable parameter } (i \neq j) \end{cases} \quad (\text{B.1})$$

where $l=1, \dots, p$ denotes the lag. We define σ_i^2 as the residual variance from a univariate AR(p) for variable i estimated with OLS. We choose $\lambda_1=0.2$ and $\lambda_2=0.5$ as in Clark (2011).

- c) The *hierarchical Bayes shrinkage prior (Shrink)* proposed by Korobilis (2013) is an hierarchical Normal-Jeffreys prior which defines $\underline{b}=0_{n \times 1}$ and $\underline{V}_{jj}=\lambda_j, j=1, \dots, n$ with

$$\lambda_j = \begin{cases} \delta_{100}(\lambda) & \text{for the intercepts} \\ 1/\lambda_j & \text{otherwise} \end{cases} \quad (\text{B.2})$$

where $\delta_{100}(\lambda)$ is the Dirac delta function. Assuming a scale invariant Jeffreys prior on λ_j , its posterior value is solely data driven. This is in contrast to the previous two approaches where the shrinkage parameter λ is an *ad hoc* selection of the researchers.

For the SSVS models we use the following three priors on regression coefficients:

- d) The *Standard ‘semi-automatic’* prior of George, Sun, and Ni (2008) which defines $\tau_{0j}=c_0\sqrt{\text{var}(\beta_j)}$ and $\tau_{1j}=c_1\sqrt{\text{var}(\beta_j)}$, where $\text{var}(\beta_j)$ is the OLS estimate of the variance of β_j parameters, $c_0=1/10$ and $c_1=10$.
- e) The *conjunction of the Minnesota prior* and the SSVS method proposed by Koop (2013) where $\tau_{0j}=c_0\sqrt{\text{var}_{\text{Min}}(\beta_j)}$ and $\tau_{1j}=c_1\sqrt{\text{var}_{\text{Min}}(\beta_j)}$, with $\text{var}_{\text{Min}}(\beta_j)$ being the variance of the β_j parameters as defined by the Minnesota prior in (b) with $c_0=1/10$ and $c_1=1$. This prior specification implies that if $\gamma_j=1$ then we impose the same degree of shrinkage on β_j as in the Minnesota prior, whereas if $\gamma_j=0$ we impose additional shrinkage.
- f) The *conjunction of the ridge prior* and the SSVS method where $\tau_{0j}=c_0\sqrt{\text{var}_{\text{Ridge}}(\beta_j)}$ and $\tau_{1j}=c_1\sqrt{\text{var}_{\text{Ridge}}(\beta_j)}$, with $\text{var}_{\text{Ridge}}(\beta_j)$ being the variance of the β_j parameters as defined by the Ridge prior above in (a) with $c_0=1/10$ and $c_1=1$. Again the prior suggests that if $\gamma_j=1$ then we impose the same degree of shrinkage on β_j as in the Ridge prior, whereas if $\gamma_j=0$ we impose additional shrinkage.

We should note that the mean-adjusted form used in the steady-state VARs does not contain constant terms and the priors presented above are implemented only for the dynamic coefficients.

Appendix C: Additional forecasting results

Table C.1: Relative RMSFE out-of-sample results (variables in a non-gap form).

Models\horizon (quarters)	Real GDP growth				CPI inflation				Federal funds rate			
	h=1	h=4	h=8	h=12	h=1	h=4	h=8	h=12	h=1	h=4	h=8	h=12
1. RW (benchmark)	3.36	3.38•	3.33•	3.32•	3.38	3.45	3.55	3.64	0.99•	2.22	3.33	3.86
2. OLS	0.97•	1.04	1.13	1.18	0.64	0.88	1.12	1.16	0.99•	1.00	0.99	0.99
3. Min-standard	0.98	1.02	1.04	1.04	0.61	0.81	1.05	1.07	1.01•	1.03	1.00	0.99
4. Min-VS	0.95•	1.01	1.03	1.05	0.61•	0.81	1.05	1.07	1.01•	1.03	1.01	1.00
5. Min-SSP	0.97	0.99•	1.00	1.00•	0.61•	0.77•	0.93	0.93	0.99•	1.00	0.97	0.95
6. Min-VS-SSP	0.93•	0.99•	1.00•	1.00•	0.61•	0.77•	0.93	0.93	0.99•	1.01	0.98	0.96
7. Ridge-standard	0.98	1.05	1.05	1.03	0.62	0.81	1.07	1.11	1.11•	1.05	1.04	1.02
8. Ridge-VS	0.94•	1.01	1.03	1.05	0.61•	0.81	1.05	1.08	1.12	1.07	1.05	1.04
9. Ridge-SSP	0.97	1.02	1.01	0.99•	0.61	0.77•	0.94	0.93	1.08•	1.00	0.99	0.96
10. Ridge-VS-SSP	0.93•	0.99•	0.99•	0.99•	0.60•	0.77•	0.92•	0.93	1.05•	1.00	0.98	0.97
11. Shrink-standard	0.95•	1.01	1.04	1.06	0.62	0.82	1.07	1.09	1.07•	1.05	1.04	1.03
12. Shrink-VS	0.95•	1.01	1.04	1.06	0.62	0.83	1.07	1.10	1.06•	1.06	1.05	1.04
13. Shrink-SSP	0.93•	0.98•	1.00	1.00	0.61•	0.78	0.93	0.94	1.04•	0.99	0.97	0.97
14. Shrink-VS-SSP	0.93•	0.98•	1.00•	1.00•	0.61	0.78	0.93•	0.94	1.03•	0.98•	0.97•	0.97
15. SSVS	0.96•	1.08	1.02	0.99•	0.61•	0.77•	0.99	1.03	1.20	0.99•	0.98•	0.98
16. SSVS-SSP	0.98	1.00	1.03	1.00	0.61•	0.77•	0.90•	0.90•	1.01•	0.99	0.95•	0.94•
17. Min-SSVS	0.96•	1.09	1.01	1.00•	0.61•	0.77	1.01	1.06	1.18	1.01	0.99	0.99
18. Min-SSVS-SSP	0.95•	1.00•	1.00	0.99•	0.61•	0.75•	0.91•	0.91•	1.17	1.00	0.95•	0.93•
19. Ridge-SSVS	0.98	1.10	1.02	0.98•	0.62	0.80	1.02	1.06	1.00•	1.00	1.00	1.00
20. Ridge-SSVS-SSP	0.95•	0.99•	1.00•	1.00•	0.61	0.78•	0.91•	0.92•	0.99•	0.99	0.95•	0.94•

The first row shows the root mean squared forecast error (RMSFE) for the benchmark random walk (RW) model. For the rest of the models the table shows the relative RMSFE. The symbol, •, indicates that a model belongs to the MCS because its p -value is greater than the prespecified significance level, α , where $\alpha=0.10$. The MCS p -values are calculated using the *quadratic* test statistic. RW is random walk model with drift in levels. OLS is a standard VAR(1) estimated using ordinary least squares. Min, Ridge and Shrink are the Minnesota, Ridge and Shrink priors for the dynamic coefficients, respectively. Standard is a BVAR(4) model, VS is BVAR(4) with variable selection *à la* Korobilis (2013), SSP is a BVAR(4) with a steady-state prior, VS-SSP is a steady-state BVAR(4) model with variable selection, SSVS is a BVAR(4) with stochastic search variable selection *à la* George, Sun, and Ni (2008) and SSVS-SSP is a BVAR(4) with steady-state prior and stochastic search variable selection. For the SSVS (17) and SSVS-SSP (18) models the prior on the dynamic coefficients is that of George, Sun, and Ni (2008). The out-of-sample period is 1969:Q4–2015:Q1.

Table C.2: Relative MAFE out-of-sample results (variables in a non-gap form).

Models\horizon (quarters)	Real GDP growth				CPI inflation				Federal funds rate			
	h=1	h=4	h=8	h=12	h=1	h=4	h=8	h=12	h=1	h=4	h=8	h=12
1. RW (benchmark)	2.39	2.40•	2.36•	2.34•	2.42	2.46	2.53	2.62	0.58•	1.59•	2.52•	3.03
2. OLS	0.99•	1.05	1.17	1.21	0.61	0.86	1.06	1.12	0.97•	1.01•	1.01	0.99
3. Min-standard	1.00	1.04	1.05	1.05•	0.59•	0.83	1.04	1.05	1.03•	1.03•	1.03	0.98
4. Min-VS	0.97•	1.02	1.05	1.06•	0.58•	0.83	1.04	1.05	1.02•	1.03•	1.04	0.99
5. Min-SSP	0.99	1.00•	1.02•	1.01•	0.59•	0.78•	0.92	0.86	1.02•	1.00•	0.95•	0.88
6. Min-VS-SSP	0.96•	0.99•	1.01•	1.01•	0.59•	0.78•	0.92	0.86	1.01•	1.00•	0.96•	0.90
7. Ridge-standard	1.01	1.06	1.06	1.04•	0.59•	0.84	1.07	1.08	1.08•	1.04•	1.07	1.00
8. Ridge-VS	0.98•	1.02	1.05	1.06•	0.58•	0.82	1.06	1.06	1.08	1.07	1.07	1.02
9. Ridge-SSP	1.01	1.02	1.03	1.00•	0.59•	0.78•	0.93	0.87	1.05•	1.00•	0.97	0.90
10. Ridge-VS-SSP	0.96•	0.99•	1.01•	1.01•	0.58•	0.78•	0.92	0.86•	1.01•	0.99•	0.97•	0.91
11. Shrink-standard	0.98•	1.02	1.06	1.07•	0.59•	0.83	1.05	1.06	1.04•	1.05	1.05	1.02
12. Shrink-VS	0.98•	1.02	1.06	1.07•	0.59•	0.84	1.05	1.06	1.04•	1.06	1.06	1.02
13. Shrink-SSP	0.96•	0.99•	1.02	1.02•	0.59•	0.78•	0.92	0.86	1.01•	0.98•	0.96•	0.91
14. Shrink-VS-SSP	0.95•	0.99•	1.02•	1.02•	0.60•	0.78•	0.92	0.86	1.00•	0.97•	0.96•	0.91
15. SSVS	0.97•	1.08	1.03	1.01•	0.59•	0.77•	0.96	0.97	1.36	1.01•	0.99•	0.97
16. SSVS-SSP	1.00	1.01•	1.06	1.02•	0.58•	0.77•	0.89•	0.83•	1.09	1.01•	0.95•	0.88•
17. Min-SSVS	0.97•	1.07	1.01•	1.00•	0.60	0.80	1.00	1.01	1.28	1.03•	1.01	1.00
18. Min-SSVS-SSP	0.96•	1.00•	1.02	1.01•	0.60•	0.77•	0.90•	0.84•	1.27	1.02•	0.94•	0.87•
19. Ridge-SSVS	0.98•	1.10	1.03•	1.00•	0.59•	0.81	0.98	0.99	1.07•	1.01•	1.00	0.99
20. Ridge-SSVS-SSP	0.97•	0.98•	1.02•	1.01•	0.59•	0.76•	0.89•	0.84•	1.07	1.01•	0.95•	0.88•

The first row shows the root mean absolute forecast error (MAFE) for the benchmark random walk (RW) model. For the rest of the models the table shows the relative RMSFE. The symbol, •, indicates that a model belongs to the MCS because its p -value is greater than the prespecified significance level, α , where $\alpha=0.10$. The MCS p -values are calculated using the *quadratic* test statistic. RW is random walk model with drift in levels. OLS is a standard VAR(1) estimated using ordinary least squares. Min, Ridge and Shrink are the Minnesota, Ridge and Shrink priors for the dynamic coefficients, respectively. Standard is a BVAR(4) model, VS is BVAR(4) with variable selection \hat{a} la Korobilis (2013), SSP is a BVAR(4) with a steady-state prior, VS-SSP is a steady-state BVAR(4) model with variable selection, SSVS is a BVAR(4) with stochastic search variable selection \hat{a} la George, Sun, and Ni (2008) and SSVS-SSP is a BVAR(4) with steady-state prior and stochastic search variable selection. For the SSVS (17) and SSVS-SSP (18) models the prior on the dynamic coefficients is that of George, Sun, and Ni (2008). The out-of-sample period is 1969:Q4–2015:Q1.

Table C.3: Relative RMSFE results (variables in a gap form).

Models\horizon (quarters)	Real GDP growth			CPI inflation			Federal funds rate		
	h=1	h=4	h=8	h=12	h=1	h=4	h=8	h=12	h=12
1. RW (benchmark)	3.36	3.38	3.33•	3.32	3.38	3.45	3.55	3.64	0.99
2. OLS	0.94•	1.00	1.04	1.05	0.63	0.84	0.97	0.97	0.97•
3. Min-standard	0.96	1.00	1.00•	1.00	0.60•	0.75	0.88	0.89	0.99
4. Min-VS	0.93•	0.99	1.00•	1.00	0.60•	0.76	0.88	0.89	0.98
5. Min-SSP	0.95	0.99	0.99•	0.99	0.60•	0.75•	0.87•	0.87	0.98•
6. Min-VS-SSP	0.92•	0.98•	0.99•	0.99	0.60•	0.75	0.86•	0.87	0.98•
7. Ridge-standard	0.97	1.03	1.00	1.00	0.61•	0.75	0.89	0.88	1.07
8. Ridge-VS	0.93•	0.99•	0.99•	1.00	0.60•	0.75	0.87	0.87	1.06
9. Ridge-SSP	0.96	1.02	1.00	0.99•	0.61•	0.74•	0.87	0.87	1.07
10. Ridge-VS-SSP	0.92•	0.98•	0.99•	0.99	0.60•	0.74•	0.85•	0.85•	1.05
11. Shrink-standard	0.94•	0.99	1.01	1.00	0.61	0.77	0.88	0.88	1.04
12. Shrink-VS	0.94•	1.00	1.01	1.00	0.61	0.77	0.89	0.88	1.03
13. Shrink-SSP	0.93•	0.99•	1.00	0.99	0.61	0.75	0.86•	0.86	1.03
14. Shrink-VS-SSP	0.94•	0.99	1.00	0.99	0.61	0.76	0.86•	0.86	1.02
15. SSVS	0.93•	1.04	1.00	0.98•	0.62	0.76	0.88	0.89	0.99
16. SSVS-SSP	0.92•	0.97•	0.99•	0.99•	0.61	0.75•	0.86•	0.87	0.99
17. Min-SSVS	0.94•	1.04	0.98•	0.97•	0.60•	0.74•	0.88	0.89	1.17
18. Min-SSVS-SSP	0.93•	0.99•	0.99•	0.99•	0.60•	0.74•	0.86•	0.87	1.17
19. Ridge-SSVS	0.95•	1.05	1.00•	0.97•	0.61	0.76	0.88	0.90	0.98•
20. Ridge-SSVS-SSP	0.93•	0.97•	0.99•	0.99	0.61•	0.75	0.86•	0.88	0.99

The first row shows the root mean squared forecast error (RMSFE) for the benchmark random walk (RW) model. For the rest of the models the table shows the relative RMSFE. The symbol, •, indicates that a model belongs to the MCS because its p -value is greater than the prespecified significance level, α , where $\alpha=0.10$. The MCS p -values are calculated using the *quadratic* test statistic. RW is random walk model with drift in levels. OLS is a standard VAR(1) estimated using ordinary least squares. Min, Ridge and Shrink are the Minnesota, Ridge and Shrink priors for the dynamic coefficients, respectively. Standard is a BVAR(4) model, VS is BVAR(4) with variable selection *à la* Korobilis (2013), SSP is a BVAR(4) with a steady-state prior, VS-SSP is a steady-state BVAR(4) model with variable selection, SSVS is a BVAR(4) with stochastic search variable selection *à la* George, Sun, and Ni (2008) and SSVS-SSP is a BVAR(4) with steady-state prior and stochastic search variable selection. For the SSVS (17) and SSVS-SSP (18) models the prior on the dynamic coefficients is that of George, Sun, and Ni (2008). The out-of-sample period is 1969:Q4–2015:Q1.

Table C.4: Relative MAFE out-of-sample results (variables in a gap form).

Models\horizon (quarters)	Real GDP growth				CPI inflation				Federal funds rate			
	h=1	h=4	h=8	h=12	h=1	h=4	h=8	h=12	h=1	h=4	h=8	h=12
1. RW (benchmark)	2.39•	2.40•	2.36•	2.34	2.42	2.46	2.53	2.62	0.58•	1.59•	2.52	3.03
2. OLS	0.99•	1.03	1.07	1.08	0.62•	0.85	0.99	1.00	0.99•	1.01	0.94	0.85
3. Min-standard	1.00•	1.02	1.01•	1.00	0.58•	0.79	0.92	0.91	1.03•	0.99•	0.92	0.83
4. Min-VS	0.98•	1.02	1.01•	1.00	0.58•	0.79	0.92	0.91	1.03•	0.99•	0.92	0.82
5. Min-SSP	0.99•	1.01•	1.00•	0.99•	0.58•	0.77•	0.89	0.88	1.03	0.98•	0.91•	0.82
6. Min-VS-SSP	0.96•	1.00•	1.00•	0.99•	0.58•	0.78	0.89	0.88	1.02•	0.98•	0.91•	0.81•
7. Ridge-standard	1.01•	1.05	1.01•	1.00	0.58•	0.79	0.92	0.90	1.06	1.00	0.93	0.83
8. Ridge-VS	0.98•	1.02	1.01•	1.01	0.58•	0.78	0.90	0.88	1.03•	1.00	0.93	0.82
9. Ridge-SSP	1.00•	1.03	1.01•	0.99•	0.58•	0.77•	0.90	0.88	1.06	0.99•	0.92	0.82
10. Ridge-VS-SSP	0.98•	1.00•	1.00•	1.00•	0.57•	0.76•	0.87•	0.86•	1.02•	0.99•	0.92	0.81
11. Shrink-standard	0.98•	1.03	1.02	1.02	0.59•	0.78	0.91	0.89	1.03	0.99•	0.92	0.82
12. Shrink-VS	0.99•	1.03	1.02	1.02	0.59•	0.79	0.91	0.89	1.03•	0.99•	0.92	0.83
13. Shrink-SSP	0.98•	1.01	1.01•	1.00	0.58•	0.77•	0.88•	0.86	1.02•	0.98•	0.91•	0.81
14. Shrink-VS-SSP	0.99•	1.02	1.01•	1.01	0.59•	0.77•	0.88•	0.86•	1.01•	0.98•	0.91•	0.82
15. SSVS	0.95•	1.04	1.04	0.99•	0.59•	0.79	0.91	0.91	1.09	0.98•	0.90•	0.80•
16. SSVS-SSP	0.97•	0.99•	1.00•	0.99•	0.58•	0.78•	0.90	0.90	1.11	1.01	0.91	0.82
17. Min-SSVS	0.95•	1.03	1.00•	0.97•	0.59•	0.76•	0.89•	0.89	1.28	0.99•	0.91•	0.81
18. Min-SSVS-SSP	0.96•	1.01•	1.00•	0.99•	0.59•	0.76•	0.89	0.88	1.28	1.00	0.90•	0.80•
19. Ridge-SSVS	0.96•	1.06	1.04	1.00•	0.58•	0.79	0.91	0.91	1.06	0.97•	0.89•	0.80•
20. Ridge-SSVS-SSP	0.98•	0.99•	1.00•	0.99•	0.59•	0.78	0.90	0.90	1.08	1.00	0.91•	0.81

The first row shows the root mean absolute forecast error (MAFE) for the benchmark random walk (RW) model. For the rest of the models the table shows the relative RMSFE. The symbol, •, indicates that a model belongs to the MCS because its p -value is greater than the prespecified significance level, α , where $\alpha=0.10$. The MCS p -values are calculated using the *quadratic* test statistic. RW is random walk model with drift in levels. OLS is a standard VAR(1) estimated using ordinary least squares. Min, Ridge and Shrink are the Minnesota, Ridge and Shrink priors for the dynamic coefficients, respectively. Standard is a BVAR(4) model, VS is BVAR(4) with variable selection *à la* Korobilis (2013), SSP is a BVAR(4) with a steady-state prior, VS-SSP is a steady-state BVAR(4) model with variable selection, SSVS is a BVAR(4) with stochastic search variable selection *à la* George, Sun, and Ni (2008) and SSVS-SSP is a BVAR(4) with steady-state prior and stochastic search variable selection. For the SSVS (17) and SSVS-SSP (18) models the prior on the dynamic coefficients is that of George, Sun, and Ni (2008). The out-of-sample period is 1969:Q4–2015:Q1.

References

- Adolfson, M., M. K. Andersson, J. Lindé, M. Villani, and A. Vredin. 2007. "Modern Forecasting Models in Action: Improving Macroeconomic Analyses at Central Banks." *International Journal of Central Banking* 3: 111–144.
- Banbura, M., D. Giannone, and L. Reichlin. 2010. "Large Bayesian Vector Auto Regressions." *Journal of Applied Econometrics* 25: 71–92.
- Beechey, M., and P. Österholm. 2008. "A Bayesian Vector Autoregressive Model with Informative Steady-State priors for the Australian Economy." *The Economic Record* 84 (267): 449–465.
- Beechey, M., and P. Österholm. 2010. "Forecasting Inflation in an Inflation-Targeting Regime: A Role for Informative Steady-State Priors." *International Journal of Forecasting* 26 (2): 248–264.
- Chan, J. C. C., and G. Koop. 2014. "Modelling Breaks and Clusters in Steady States of Macroeconomic Variables." *Computational Statistics & Data Analysis* 76: 186–193.
- Clark, T. E. 2011. "Real-Time Density Forecasts from Bayesian Vector Autoregressions with Stochastic Volatility." *Journal of Business and Economics Statistics* 29 (3): 327–341.
- Cogley, T. 2002. "A Simple Adaptive Measure of Core Inflation." *Journal of Money, Credit and Banking* 34 (1): 94–113.
- Doan, T., R. Litterman, and C. Sims. 1984. "Forecasting and Conditional Projection using Realistic Prior Distributions." *Econometric Reviews* 3 (1): 1–144.
- Faust, J., and J. H. Wright. 2013. "Forecasting Inflation." *Handbook of Economic Forecasting* 2(Part A): 3–56.
- Gefang, D. 2014. "Bayesian Doubly Adaptive Elastic-Net Lasso for VAR Shrinkage." *International Journal of Forecasting* 30 (1): 1–11.
- George, E. I., and R. E. McCulloch. 1997. "Approaches for Bayesian Variable Selection." *Statistica Sinica* 7 (2): 339–373.
- George, E. I., D. Sun, and S. Ni. 2008. "Bayesian Stochastic Search for VAR Models Restrictions." *Journal of Econometrics* 142 (1): 553–580.
- Geweke, J. 1992. "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments." In *Bayesian Statistics*, edited by J. M. Bernardo, J. Berger, A. P. Dawid and A. F. M. Smith, 169–193. Oxford: Oxford University Press.
- Giannone, D., M. Lenza, and G. Primiceri. 2015. "Prior selection for Vector Autoregressions." *Review of Economics and Statistics* 97 (2): 436–451.
- Hansen, P. R., A. Lunde, and J. M. Nason. 2003. "Choosing the Best Volatility Models: The Model Confidence Set Approach." *Oxford Bulletin of Economics and Statistics* 65 (S1): 839–861.
- Hansen, P. R., A. Lunde, and J. M. Nason. 2011. "The Model Confidence Set." *Econometrica* 79 (2), 453–479.
- Jarociński, M., and J. F. Smets. 2008. "House Prices and the Stance of Monetary Policy." *Federal Reserve Bank of St. Louis Review* 90: 339–365.
- Karlsson, S. 2013. "Forecasting with Bayesian Vector Autoregressions." *Handbook of Economic Forecasting* 2 (Part B): 791–897.
- Koop, G. 2013. "Forecasting with Medium and Large Bayesian VARs." *Journal of Applied Econometrics* 28 (2): 177–203.
- Korobilis, D. 2008. "Forecasting with Vector Autoregressions with Many Predictors." *Advances in Econometrics* 23, 403–431.
- Korobilis, D. 2009. "VAR Forecasting Using Bayesian Variable Selection." *MPRA Paper No. 21124*.
- Korobilis, D. 2013. "VAR Forecasting Using Bayesian Variable Selection." *Journal of Applied Econometrics* 28: 204–230.
- Litterman, R. B. 1986. "Forecasting with Bayesian Vector Autoregressions – Five Years of Experience." *Journal of Business and Economic Statistics* 4 (1): 25–38.
- Österholm, P. 2012. "The Limited Usefulness of Macroeconomic Bayesian VARs when Estimating the Probability of a US Recession." *Journal of Macroeconomics* 34 (1): 76–86.
- Primiceri, G. E. 2005. "Time Varying Structural Vector Autoregressions and Monetary Policy." *The Review of Economic Studies* 72 (3): 821–852.
- Sims, C. 1980. "Macroeconomics and Reality." *Econometrica* 48 (1): 1–48.
- Van Roye, B. 2011. "Financial Stress and Economic Activity in Germany and the Euro area." *Kiel Working Papers* No 1743.
- Villani, M. 2005. "Inference in Vector Autoregressive Models with Informative Prior on the Steady State." *Sveriges Riksbank working paper series* No 181.
- Villani, M. 2009. "Steady-State Priors for Vector Autoregressions." *Journal of Applied Econometrics* 24: 630–650.
- Wright, J. H. 2013. "Evaluating Real-Time VAR Forecasts with an Informative Democratic Prior." *Journal of Applied Econometrics* 28: 762–776.