BANK OF GREECE
EUROSYSTEM

Working Paper

Steady-state priors and Bayesian variable selection in VAR forecasting

Dimitrios P. Louzis

195

JULY 2015

WORKINGPAPERWORKINGPAPERWORKINGPAPERWO

BANK OF GREECE
Economic Analysis and Research Department – Special Studies Division
21, E. Venizelos Avenue
GR-102 50 Athens
Tel: +30210-320 3610
Fax:+30210-320 2432

www.bankofgreece.gr

Printed in Athens, Greece
at the Bank of Greece Printing Works.

ISSN 1109-6691

STEADY-STATE PRIORS AND BAYESIAN VARIABLE SELECTION IN VAR FORECASTING

Dimitrios P. Louzis
Bank of Greece

Abstract

This study proposes methods for estimating Bayesian vector autoregressions (VARs) with an 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, Steady states, Variable selection, Macroeconomic forecasting

JEL Classifications: C32

Acknowledgement: The author would like to thank Heather Gibson for her constructive comments and suggestions. The views expressed in this article do not necessarily represent those of the Bank of Greece.

Correspondence:
Dimitrios P. Louzis
Bank of Greece,
Economic Analysis and Research Department
3 Amerikis Str., 10564
Athens, Greece.
Email: dlouzis@aueb.gr ; dlouzis@bankofgreece.gr

1. Introduction

The seminal studies of Sims (1980), Doan et al. (1984) and Linterman (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 thirty 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 Linterman) prior (Doan et al. 1984; Linterman, 1986) and its variants (Banbura et al, 2010; Koop 2013) 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 stochastically decides 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 et al., 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 variable selection 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.

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 sate 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 Villani, 2009; Beechey and Österholm, 2008; Clark 2011). In Beechhey 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) 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 algorithm for Bayesian variable selection proposed by 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 fourth block to the Gibbs sampler of 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 forecasting exercise based on three US macroeconomic variables, namely, real gross domestic product (GDP) growth, inflation (consumer price index, CPI) and short-term interest rates (Federal funds rate). The out-of-sample period covers 20 years from 1993:Q1 to 2013:Q4 and the suggested model is compared to alternative established specifications such us the steady-state VAR model of Villani (2009) and the VAR model with variable selection of Korobilis (2013). Finally, as a robustness check we also use three alternative prior specifications on the dynamic VAR parameters (see

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

also Korobilis, 2013) and investigate whether we reach different conclusions in terms of estimation results and forecasting performance.

The rest of the paper is organized as follows. Section 2 develops the steady-state VAR with variable selection and the Gibbs sampling algorithm. 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

$$B(L)y_t = cd_t + \varepsilon_t \tag{1}$$

where y_t is a $m \times 1$ vector of time series at time t with $t = 1,...,T$ observations, $B(L) = I_m - B_1 L - ... - B_p L^p$ 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] 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.:

$$B(L)(y_t - \varphi d_t) = \varepsilon_t \tag{2}$$

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 extended to allow for Bayesian variable selection in the form of Korobilis (2013). This means that each of the VAR equations may have different lagged variables and Eq. (2) should be re-written as a system of seemingly unrelated equations (SUR). To see this, first write the steady-state VAR as

$$\tilde{y}_t = B_1 \tilde{y}_{t-1} + ... + B_p \tilde{y}_{t-p} + \varepsilon_t \tag{3}$$

with $\tilde{y}_t = y_t - \varphi d_t$ being the *mean-adjusted* times series of y_t. Then, the SUR representation of Eq. (3) is

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

$$\tilde{y}_t = \tilde{z}_t \beta + \varepsilon_t \tag{4}$$

where $\tilde{z}_t = I_m \otimes \tilde{x}_t$ is a $m \times n$ matrix with $n = m^2 p$ being the total number of coefficients in (4), $\tilde{x}_t = (\tilde{y}'_{t-1}, ..., \tilde{y}'_{t-p})$ is a $1 \times k$ $(k = mp)$ vector containing all dependent variables at time t, $\beta = vec(B)$ is a $n \times 1$ vector of the VAR dynamic parameters with $B = (B_1, ..., B_p)'$. The model in Eq. (4) 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

$$\begin{cases} \beta_j = 0 \text{ if } \gamma_j = 0 \\ \beta_j \neq 0 \text{ if } \gamma_j = 1 \end{cases} \tag{5}$$

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

$$\tilde{y}_t = \tilde{z}_t \theta + \varepsilon_t \tag{6}$$

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. (6) implies that for $\gamma_j = 1$ for $\forall j$ we get the unrestricted steady-state VAR of Eq. (4).

2.1. Prior distributions

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 for $\forall j$ and their prior density is defined as a Bernoulli density with $\underline{\pi}_j$ prior probability, i.e. $\gamma_j | \gamma_{\setminus -j} \sim Bernoulli(\underline{\pi}_j)$, 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 .*Posterior distributions*

The posterior inference is based on the idea that conditional on the steady-state parameters, φ, the VAR specification in Eq. (6) is a standard restricted VAR for the mean-adjusted time series $\tilde{y}_t$ and $\tilde{z}_t$, and therefore the Gibbs sampler framework of Korobilis (2013) can be implemented.[3] In fact, estimation of (6) requires only 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, which draws sequentially from the full conditional posterior density of the parameters, contains the following four steps:[4]

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

where $\mathcal{D} = \{y_1, ..., y_T, d_1, ..., d_T\}$ denote the data. The quantities $\overline{\beta}$, $\overline{V}$, $\overline{\pi}_j$, S, $\overline{b}_{\varphi}$ and $\overline{V}_{\varphi}$ as well as the posterior sampling procedure are described in detail in Appendix A.

Compared to the three-block structure of the standard restricted VAR model, the extra block requires only a very small portion of the total computing time 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, 2009, Appendix A, p. 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

[3] Alternatively, one can argue that conditional on γ the model in Eq. (6) 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.

simulations. We also confirm the abovementioned findings for the proposed model by implementing a similar simulation exercise.[5] Moreover, in Appendix C, we use simulated datasets and we show that the Bayesian variable selection technique of Korobilis (2013) works equally well under the new steady-state framework.

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 form 1974:Q1 to 2013:Q4 and were obtained from the St. Louis' FRED database on a quarterly basis.[6] Real GDP growth and CPI inflation rates are calculated as year-on-year percentage changes (i.e. the percentage change of the current quarter over the corresponding quarter of the previous year) while we use the interest rate in levels. Thus, the vector of the endogenous variables is given by $y_t = (r_t, \pi_t, i_t)'$ and all variables are expressed in yearly percentages.

3.1. Prior specifications and alternative models

An initial first step in Bayesian estimation is to specify the hyperparameters in prior distributions described in Section 2.1. We follow closely Korobillis (2013) and we use three distinct prior specifications on the dynamic parameters vector, β. These are: (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.[7] We also set the prior probability of the Bernoulli density equal to 0.8, i.e. $\underline{\pi}_j = 0.8$, while the standard non-informative prior $|\Sigma|^{-(m+1)/2}$ is used for Σ.

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

In the steady-state version of VAR models we also have to specify the prior mean and standard deviation on steady-states coefficients, φ. We follow Österholm (2012) and we set both steady-state GDP growth and inflation rate equal to 2% and nominal interest rate equal to 4%, while we assume a 0.5 standard deviation for all three variables. These values reflect researcher's perception regarding the long-run level of the variables and are also in accordance with the contribution of Jaronciski and Smets (2008) who use economic theory, i.e. the Fisher equation, to specify steady-state values for the US economy.

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 an automatic Bayesian variable selection. To that end, we compare the forecasting ability of the following four alternative specifications:

(i) a baseline BVAR without variable selection and steady-state priors,

(ii) a steady-state BVAR (Villani, 2009)

(iii) a BVAR with variable selection (Kororbillis, 2013) and

(iv) a BVAR with both variable selection and steady-state priors.

Each of these four specifications is estimated using each of the three prior specifications on β (see Appendix B for details), resulting in twelve 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 (1974:Q1 - 2014:Q1). We use the Gibbs sampler in Appendix A to sample 30,000 draws of models' parameters after discarding the first 20,000 draws used for initial convergence (burn-in period). Convergence of the Gibbs sample was excellent in all cases.

Table 1 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 2.68% ranging between 2.63 and 2.75.

The average CPI inflation across priors and specifications is close to 2.71% with its minimum (maximum) value being 2.49 (2.91) while the Fed funds rate is estimated to be close to 3.78% (3.65 and 3.95 are the minimum and maximum values respectively)

[Insert Table 1 here]

In Table 2 we also present estimation results regarding the restriction parameters γ_j. As pointed out in Korobilis (2013) the posterior mean of γ_j can be seen as an average probability of including 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 2 reveal that the differences in γ_j estimations between models with and without steady-state priors are non-negligible. More specifically, except for the first own lag of each dependent variable which is always one across models, meaning that the first lag should always be included, all other γ_j estimates differentiate substantially. In particular, steady-state models tend to produce, overall, much lower γ_j posterior means with the majority of these estimates being lower than 50%. For instance, the second lag of inflation (π_{t-2}) and the fourth lag of interest rates (i_{t-4}) in models with uninformative steady-state seem to affect the current level of inflation with the average γ_j being close to 95% and 87%, respectively, across priors. Nonetheless, these estimates are close to 21% and 13% for the steady-state models.

[Insert Table 2 here]

However, as underlined in Korobilis (2013) what really matters is not the γ_j posterior mean *per se* but the combined posterior mean of $\theta_j = \beta_j \gamma_j$. The intuition is that a lagged dependent variable will enter the equation with a θ_j coefficient which implies that if $\theta_j \approx 0$ this variable is actually excluded from the VAR system. Therefore, in practice, the differences between θ_j's of alternative specifications may not always be as acute as implied by the posterior probabilities of inclusion. An indicative example is the i_{t-4} variable in the inflation equation with a Shrink prior. In particular, the posterior mean of $\gamma_{i_{t-4}}$ is 85% and 19% for the uninformative and the steady-state models respectively, but we get posterior means of $\theta_{i_{t-4}}$ that are very close to each other (-0.20 and -0.22 respectively).

3.3. Out-of-sample forecasting analysis

We evaluate the alternative VAR specifications using a forecasting horizon of twelve quarters (3 years), $h = 1,...,12$, and an out-of-sample period spanning from 1993:Q1 to 2013:Q4. The forecasts across all horizons are produced using a recursive forecasting scheme. This means that we use an initial sample (1974:Q1-1992:Q4) to generate forecasts from 1993:Q1 to 1995:Q4, i.e. 12 quarters ahead. Next, we allow the sample to expand and include one more period, i.e. 1974:Q1-1993:Q1, and generate forecasts from 1993:Q2 to 1996:Q1. This procedure is continued till the end of the sample period. As mentioned in Section 3.2 estimation and forecasting results are based on 30,000 posterior simulations (after a burn-in period of 20,000 simulations) and forecasts are generated iteratively following Korobilis (2013, p. 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 1–3. The forecasts has been generated by three alternative models: a VAR(1) estimated with OLS, a VAR model with variable selection and a Minnesota prior and a steady-state VAR model with variable selection and a Minnesota prior. As expected, the forecasts generated by the steady-state VAR converge to the steady-state of each variable much faster than its counterparts. By contrast, the other two models seem to over- or under- estimate the steady-state levels and their forecasts drift away even when the level of the variables is too high (or too low) to justify an upward (or a downwar) movement. 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 square forecast error* (RMSFE) and the *mean absolute forecast error* (MAFE).

[Insert Figures 1–3 here]

Figures 4, 5 and 6 present the out-of sample forecasting results for the ridge, Minnesota and shrink prior on β, respectively. 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 RMAFE and MAFE metrics of a *random walk* (RW) process

$$\text{Realtive } RMSFE_{ij}^{h} = \frac{RMSFE_{ij}^{h}}{RMSFE_{i,RW}^{h}}, \text{ Realtive } MAFE_{ij}^{h} = \frac{RMAFE_{ij}^{h}}{RMAFE_{i,RW}^{h}}$$

where h = 1,…,12 is the forecasting horizon, i is the variable of interest, i.e. GDP, inflation and interest rate, and j are the following competing models: a VAR(1) estimated with OLS (VAR-OLS), a BVAR(4) model (VAR), a steady-state BVAR(4) model (VAR-STEADY), a BVAR(4) model with variable selection (VAR-VS) and finally a steady-state BVAR(4) model with variable selection (VAR-VS-STEADY). Values below one indicate that the corresponding model outperforms the random walk process and vice versa. The left column in Figures presents the relative RMSFE of the various competing models as a function of the forecasting horizon while the right column presents the relative MAFE.

[Insert Figures 4–6 here]

Overall, the results presented in Figures 4–6 indicate that the proposed model, i.e. the steady-state VAR model with variable selection (red dotted lined), is the best performing model across variables, forecasting horizons, priors on β and evaluation metrics. Moreover, the proposed model is the only Bayesian specification that consistently outperforms the VAR-OLS model and the RW process. These results indicate that the incorporation of steady-state beliefs in a VAR model with variable selection enhances its forecasting ability and leads to more accurate macroeconomic forecasting.

More specifically, as regards the GDP forecasts the proposed model almost always outperforms its competitors. Only at very long forecasting horizons ($h \geq 10$) the rest of the BVARs produce RMSFE and MAFE metrics which are very close to the ones produced by the proposed model. This evidence can be explained on the grounds of Figure 1 which shows that a BVAR with variable selection, but uninformative steady-state, usually overestimates steady-state at short horizons but eventually tends to generate long-term forecasts that are not that far from the steady-state levels presented in Table 1.

The picture is almost the same with respect to inflation forecasts with the proposed model outperforming all its competitors. Only the steady-state VAR in the case of Minnesota prior (Figure 5) generates forecasts of comparable forecasting accuracy at very long forecasting horizons ($h \geq 10$).

Moreover, the forecasting results for the interest rates reveal that steady-state VAR models with or without variable selection and the VAR-OLS are the main competitors, since they produce overall the best forecasts. In particular, Bayesian models outperform the VAR-OLS model for forecasting horizons longer than five quarters, across evaluation metrics. The proposed model also outperforms the steady-state VAR at shorter horizons with the latter being the overall best forecasting model for horizons longer than six and ten quarters for the RMSFE and MAFE metrics, respectively. In line with the literature (e.g. see Villani, 2009; Korobilis, 2013) a RW process also provides good short-term forecasts for the interest rates.

In addition, variable selection (red solid line) tends to improve GDP and inflation forecasts over unrestricted VARs (blue solid line) across all forecasting horizons. Nonetheless, as also evidenced in Korobilis (2013), the degree of improvement depends on the prior used and the level of information it carries. Thus, the improvement is substantial when we use the relative uninformative ridge prior and marginal or even negligible when we use priors that are more informative (Shrink and Minnesota prior respectively). The results for the interest rates are mixed and mostly depend on the forecasting horizon.

Finally, the empirical findings underline the importance of forcing the endogenous variables to converge to a specific steady-state level. In particular, steady-state VAR outperforms both OLS and BVARs with or without variable selection in the majority of cases examined here. Therefore, in practice, steady-state VAR is the second best performing model after the steady-state VAR with variable selection. These results also align with other studies that highlight the contribution of steady-state priors to the accuracy of macroeconomic forecasts (Villani, 2009; Beechey and Österholm, 2008; Clark, 2011).

3.3.1. Model confidence set results

In this section, we employ the Model Confidence Set (MCS) method of Hansen et al. (2003, 2011) and we construct a set of models, $\mathrm{M}^{*}_{1-a} \subseteq \mathrm{M}_0$, 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 $\mathrm{M} = \mathrm{M}_0$ models, the MCS method is based on a specific loss function, $L_{m,t}$ with $m = 1,...,\mathrm{M}$, and applies an iterative process of sequential Equal Predictive Ability (EPA) tests of the form:

$$H_{0,\mathrm{M}_0} : E(d_{mk,t}) = 0 \text{ for all } m, k \in \mathrm{M} \tag{15}$$

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 ability and should not be included in the MCS at an a significance level. The EPA test in Eq. (10) is repeated for the remaining M_{1-a} models, with $\mathrm{M}_{1-a} \subset \mathrm{M}$, and this procedure continues until the null hypothesis cannot be rejected. The final set of surviving models forms the MCS at a $1-a$ confidence level, denoted by M^*_{1-a}. 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.[8]

Table 3 presents only a synopsis of the main MCS results, while we refer the interested reader to Tables D.1-D.3 of Appendix D for a more detailed analysis. The first column for each of the variables in Table 3 presents the percentage of times a model is included in the MCS at 10% significance level across forecasting horizons and evaluation metrics. The second column presents the percentage of times a model also ranks first, i.e. minimizes RMSFE or MAFE, across forecasting horizons and evaluation metrics. The initial set includes all 13 models: the VAR-OLS and all twelve BVARS.

[Insert Table 3 here]

The MCS results in Table 3 confirm the evidence presented in Section 3.3. In particular, with regard to GDP forecasts, we find that the steady-state VAR with variable selection (VAR-VS-STEADY) and a shrink prior is always included in the MCS across forecasting horizons and evaluation metrics followed by the VAR-VS-STEADY with the Minnesota and ridge prior which are part of the MCS in 95.8% of cases. The steady-state VAR (VAR-STEADY) models are included in the MCS in approximately 50% of cases while the percentages for the rest of the models are much

[8] For details on MCS technique and its implementation see Hansen et al. (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_htm/alunde/mulcom/mulcom.htm.

lower. The VAR-VS-STEADY model with the Minnesota prior ranks first in 50% of cases while the VAR-VS-STEADY models with ridge and shrink priors follow.

The results with respect to inflation are clear cut since the VAR-VS-STEADY with the ridge prior is always part of the MCS and ranks first across forecasting horizons and evaluation metrics. The VAR-VS-STEADY model with Minnesota prior ranks first in 42% of cases as regards the interest rate, while it is part of the MCS in 96% of cases followed by VAR-VS-STEADY with ridge and shrink prior (88% of cases). VAR-OLS and steady-state VARs (VAR-STEADY) also present high percentages confirming the evidence presented in Section 3.3.

In general, the MCS results confirm that steady-state VARs with variable selection can considerably enhance the macroeconomic forecasting accuracy of small-scale VARs irrespective of the prior used on the dynamic parameters. The empirical evidence also suggest that the simple and relatively uninformative ridge prior can also provide accurate forecasts compared to the more informative and sophisticated Minnesota and 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 a Gibbs sampler algorithm for estimating a Bayesian VAR model that efficiently combines 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 abovementioned 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.

References

Adolfson, M., Andersson, M.K., Lindé. J., Villani, M., Vredin A. (2007). Modern forecasting models in action: Improving macroeconomic analyses at central banks. International Journal of Central Banking. 3, 111-144.

Banbura, M., Giannone, D., Reichlin, L. (2010). Large Bayesian vector auto regressions. Journal of Applied Econometrics. 25, 71-92.

Beechey, M. & Österholm. P. (2008). A Bayesian vector autoregressive model with informative steady-state priors for the Australian economy. The Economic Record. 84, 449-465.

Beechey, M. & Österholm. P. (2010). Forecasting inflation in an inflation-targeting regime: A role for informative steady-state priors. International Journal of forecasting. 26, 248-264.

Clark, T.E. (2011). Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility. Journal of Business and Economics Statistics. 29, 327-341.

Doan, T., Litterman, R., Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. Econometric Reviews. 3, 1-144.

Gefang, D. (2014). Bayesian doubly adaptive elastic-net Lasso for VAR shrinkage. International Journal of forecasting. 30, 1-11.

George, E.I., Sun, D., Ni, S. (2008). Bayesian stochastic search for VAR models restrictions. Journal of Econometrics. 142, 553-580.

Hansen, P.R., Lunde, A., Nason, J.M. (2003). Choosing the best volatility models: The model confidence set approach. Oxford Bulletin of Economics and Statistics. 65, 839-861.

Hansen, P.R., Lunde, A., Nason, J.M. (2011). The model confidence set. Econometrica. 79, 453-479.

Jarociński. M. & Smets. J.F. (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 autoregression. Handbook of Economic Forecasting. 2, 791-897.

Koop, G. (2013). Forecasting with medium and large Bayesian VARs. Journal of Applied Econometrics. 28, 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, 25-38.

Österholm. P. (2012). The limited usefulness of macroeconomic Bayesin VARs when estimating the probability of a US recession. Journal of Macroeconomics. 34, 76-86.

Sims, C. (1980). Macroeconomics and reality. Econometrica. 48, 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 Rirkbank working paper series, No 181.

Villani, M. (2009). Steady-state priors for vector autoregressions. Journal of Applied Econometrics, 24, 630-650.

Appendix A: Posterior inference in the linear steady-state VAR with variable selection

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 $\widetilde{Y}$ as a $T \times m$ matrix with its t^{th} row being $\widetilde{Y}_t = \widetilde{y}'_t = (\widetilde{y}_{1t}, \widetilde{y}_{2t}, \ldots, \widetilde{y}_{mt})$, $\widetilde{X}$ as an $T \times k$ matrix with its t^{th} row being $\widetilde{X}_t = \widetilde{x}_t$ and E as an $T \times m$ with its t^{th} row being $E_t = \varepsilon'_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{mt})$. Then, $\widetilde{y} = vec(\widetilde{Y})$ and $\varepsilon = vec(E)$ are $Tm \times 1$ vectors with $vec(\circ)$ being the standard operator which stacks the columns of a matrix. Then, the VAR model can be written as

$$\underset{(Tm\times 1)}{\widetilde{y}} = \underset{(Tm\times n)}{\widetilde{Z}}\underset{(n\times 1)}{\theta} + \underset{(Tm\times 1)}{\varepsilon} \qquad \text{(A. 1)}$$

where $\widetilde{Z} = I_m \otimes \widetilde{X}$ is a $Tm \times n$ block diagonal matrix with $\widetilde{X}$ being repeatedly on its main diagonal, $\theta = \Gamma\beta$ is the $n \times 1$ vector of coefficients with $\beta = vec(B)$. 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 $n = mk = m^2 p$ is the total number of VAR coefficients. Also recall that $\widetilde{y}_t = y_t - \varphi d_t$ and $\widetilde{x}_t = (\widetilde{y}'_{t-1}, \ldots, \widetilde{y}'_{t-p})$ are the mean-adjusted time-series of y_t and x_t respectively.

Next, we derive the full conditional posteriors of the four-block Gibbs sampler algorithm which employ the priors of Section 2.1. As already mentioned, conditional on steady-state parameters, φ, the VAR model in (A.1) is a standard VAR model with variable selection for the mean-adjusted series $\widetilde{y}_t$. Therefore, the first three blocks of the Gibbs sampler reproduce the Korobilis (2013) algorithm for the mean-adjusted series. The fourth 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) the model in (A.1) is a standard steady-state VAR model and can be written in a form so that we can apply Villani's (2009) methodology.

More specifically, the algorithm and the conditional posteriors are given below:

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

$$\beta | \gamma, \varphi, \Sigma, y, Z \sim N_n(\overline{\beta}, \overline{V}) \qquad \text{(A.2)}$$

where $\overline{V} = \left(\underline{V}^{-1} + \widetilde{Z}^{*\prime} \left(\Sigma^{-1} \otimes I_T \right) \widetilde{Z}^{*} \right)^{-1}$, $\overline{\beta} = \overline{V} \left(\underline{V}^{-1} \underline{b} + \widetilde{Z}^{*\prime} \left(\Sigma^{-1} \otimes I_T \right) \widetilde{y} \right)$ and $\widetilde{Z}^{*} = \widetilde{Z} \Gamma$.

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

$$\gamma_j \left| \gamma_{\setminus -j}, \beta, \varphi, \Sigma, y, Z \sim Bernoulli\left(\overline{\pi}_j \right) \right. \quad \text{(A.3)}$$

Where $\overline{\pi}_j = l_{0j} / \left(l_{0j} + l_{1j} \right)$, $l_{0j} = \exp\left(-0.5 \left(\widetilde{y} - \widetilde{Z}\theta^{*} \right)' \left(\Sigma^{-1} \otimes I_T \right) \left(\widetilde{y} - \widetilde{Z}\theta^{*} \right) \right) \pi_j$,

$l_{1j} = \exp\left(-0.5 \left(\widetilde{y} - \widetilde{Z}\theta^{**} \right)' \left(\Sigma^{-1} \otimes I_T \right) \left(\widetilde{y} - \widetilde{Z}\theta^{**} \right) \right) \left(1 - \pi_j \right)$ 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

$$\Sigma^{-1} \left| \beta, \gamma, \varphi, y, z \sim Wishart\left(T, S^{-1} \right) \right. \quad \text{(A.4)}$$

For this step, re-write the VAR as

$$\underset{(T \times m)}{\widetilde{Y}} = \underset{(T \times k)}{\widetilde{X}} \underset{(k \times m)}{\Theta} + \underset{(T \times m)}{E} \quad \text{(A.5)}$$

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

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

$$\varphi \left| \beta, \gamma, \Sigma, y, Z \sim N_{mq}\left(\overline{b}_{\varphi}, \overline{V}_{\varphi} \right) \right. \quad \text{(A. 6)}$$

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

$$\left(y_t - \varphi d_t \right) = \Theta_1 \left(y_{t-1} - \varphi d_{t-1} \right) + ... + \Theta_p \left(y_{t-p} - \varphi d_{t-p} \right) + \varepsilon_t \quad \text{(A.7)}$$

where Θ_l, $l = 1,...,p$ 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 $l = 1,...,p$ and $i, j = 1,...,m$. Rearranging the terms in (A.6) we get

$$\Theta(L) y_t = \Theta(L) \varphi d_t + \varepsilon_t = \varphi d_t - \Theta_1 \varphi d_{t-1} - ... - \Theta_p \varphi d_{t-p} + \varepsilon_t \quad \text{(A.8)}$$

where $\Theta(L) = I_m - \Theta_1 L - ... - \Theta_p L^p$. 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 $D_t = (d_t', -d_{t-1}', -..., -d_{t-p}')$ and $\Lambda = [\varphi, \Theta_1\varphi, ..., \Theta_p\varphi]'$ 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, p. 646-647). Thus, given that $vec(\Lambda') = U vec(\varphi)$ with

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

we define the mean, $\overline{b}_\varphi$, and variance, $\overline{V}_\varphi$, of the posterior distribution in (A.6) as

$$\overline{b}_\varphi = \overline{V}_\varphi \left(U' vec\left(\Sigma^{-1} Y_\varphi' D \right) + \underline{V}_\varphi^{-1} \underline{b}_\varphi \right) \qquad \text{and} \qquad \overline{V}_\varphi^{-1} = U' \left(D'D \otimes \Sigma^{-1} \right) U + \underline{V}_\varphi^{-1}$$

respectively.

The Gibbs sampler described above was implemented in Matlab and extends the Matlab code for VAR models with variable selection kindly provided by Dimitris Korobilis in his website: https://sites.google.com/site/dimitriskorobilis/matlab. The Matlab code is avalaible upon request.

Appendix B: Specifying priors on β

This appendix briefly describes the three different types of prior distribution on β (for more details see Kororbillis (2013) and references therein):

The *ridge regression prior* which defines $\underline{b} = 0_{n\times 1}$ and $\underline{V} = \lambda I_n$ 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.[9] Here, we set the autoregressive parameter equal to 0.5 for the GDP and 0.8 for the inflation and interest rate variables.

[9] Doan et al. (1984) and Linterman (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 $\delta < 1$.

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

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

where $l = 1,...,p$ 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 $\lambda = 0.1$ for the three variable VARs.

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

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

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.

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: Simulation analysis

In this Appendix we evaluate the performance of the proposed specification using simulated data sets as in Korobilis (2009). In particular, we generate 100 samples of six variables of length T = 50 using a stationary mean-adjusted VAR process with one lag ($p = 1$) and $d_t = 1$ for all t. The first lag parameter matrix, B, is given by

$$B = \begin{bmatrix} 0.95 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.95 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.95 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.95 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.95 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.95 \end{bmatrix}.$$

The matrix Ψ , where $\Sigma = (\Psi\Psi')^{-1}$ is given by

$$\Psi = \begin{bmatrix} 0.1 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0 & 0.1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.1 \end{bmatrix}$$

while the steady-states vector is specified as $\varphi' = (2,2,2,4,4,4)$. For each of the 100 samples we estimate two BVAR specifications: a steady-state VAR and a steady-state VAR with variable selection. For both specifications we use 30,000 posterior simulations after discarding the first 20,000 draws, while we use a ridge prior on the dynamic coefficients as described in Appendix B and a prior Bernoulli probability equal to 0.8, i.e. $\underline{\pi}_j = 0.8$. We also use informative steady-state priors with prior mean given by $\overline{b'_\varphi} = (2,2,2,4,4,4)$ and prior standard deviations being equal to 0.5 for all variables. The average over the 100 posterior means (along with the average posterior standard deviations in parenthesis) are presented below.

<u>*Specification I*</u>: Steady-state VAR

$$\hat{\varphi}' = \begin{bmatrix} \underset{(0.13)}{1.98} & \underset{(0.39)}{2.03} & \underset{(0.39)}{2.00} & \underset{(0.39)}{3.96} & \underset{(0.39)}{3.96} & \underset{(0.41)}{3.97} \end{bmatrix}$$

$$\hat{B}_{UN} = \begin{bmatrix} \underset{(0.18)}{0.98} & -0.05 & -0.03 & 0.44 & 0.38 & 0.32 \\ 0.08 & \underset{(0.41)}{0.65} & -0.06 & -0.09 & -0.10 & -0.10 \\ 0.08 & -0.09 & \underset{(0.41)}{0.64} & -0.11 & -0.10 & -0.12 \\ 0.09 & 0.11 & 0.12 & \underset{(0.47)}{0.79} & 0.12 & 0.10 \\ 0.09 & 0.04 & 0.04 & 0.03 & \underset{(0.47)}{0.73} & 0.04 \\ 0.09 & 0.04 & 0.01 & 0.03 & 0.02 & \underset{(0.45)}{0.77} \end{bmatrix}$$

<u>*Specification II*</u>**:** Steady state VAR with variable selection

$$\hat{\varphi}'_{VS} = \begin{bmatrix} \underset{(0.11)}{1.99} & \underset{(0.46)}{2.02} & \underset{(0.46)}{1.96} & \underset{(0.45)}{3.93} & \underset{(0.46)}{3.95} & \underset{(0.45)}{3.97} \end{bmatrix}$$

$$\hat{B}_{VS} = \begin{bmatrix} \underset{(0.01)}{0.95} & -0.04 & -0.06 & 0.07 & 0.12 & 0.12 \\ 0.00 & \underset{(0.08)}{0.88} & 0.02 & 0.01 & 0.00 & 0.01 \\ 0.00 & 0.03 & \underset{(0.07)}{0.89} & 0.01 & 0.02 & 0.01 \\ 0.00 & 0.01 & 0.01 & \underset{(0.09)}{0.87} & 0.01 & 0.02 \\ 0.00 & 0.01 & 0.01 & 0.02 & \underset{(0.09)}{0.85} & 0.03 \\ 0.00 & 0.01 & 0.01 & 0.02 & 0.03 & \underset{(0.09)}{0.86} \end{bmatrix}$$

$$\hat{\Gamma} = \begin{bmatrix} 1.00 & 0.22 & 0.23 & 0.29 & 0.31 & 0.30 \\ 0.02 & 1.00 & 0.18 & 0.13 & 0.14 & 0.13 \\ 0.02 & 0.21 & 1.00 & 0.13 & 0.18 & 0.14 \\ 0.02 & 0.13 & 0.15 & 1.00 & 0.19 & 0.19 \\ 0.02 & 0.15 & 0.15 & 0.18 & 1.00 & 0.20 \\ 0.02 & 0.15 & 0.14 & 0.17 & 0.21 & 1.00 \end{bmatrix}$$

where matrix $\hat{\Gamma}$ is the average posterior mean of γ vector in a matrix form. Simulation results reveal that the proposed specification works well in small samples. In particular, we show that the Bayesian variable selection technique of Korobilis (2013) leads to more accurate estimates of dynamic coefficients under the new steady-state framework. The $\hat{B}_{VS}$ average estimates are much closer to the data generating matrix, B, with smaller posterior standard deviations compared to the $\hat{B}_{UN}$ estimates produced by the unrestricted steady-state VAR. Finally, the average posterior means of steady-states are almost identical for both specifications.

Appendix D: Forecasting evaluation using Model Confidence Set

[Insert Table D.1 here]

[Insert Table D.2 here]

[Insert Table D.3 here]

Table 1 Posterior means and standard deviations of the steady-states

	VAR dynamic coefficients prior					
	Ridge		Minnesota		Shrink	
	no VS	with VS	no VS	with VS	no VS	with VS
Real GDP						
Posterior mean	2.66	2.69	2.69	2.63	2.75	2.68
St. deviation	0.34	0.38	0.33	0.49	0.31	0.42
Inflation						
Posterior mean	2.91	2.49	2.87	2.66	2.81	2.54
St. deviation	0.49	0.65	0.49	0.70	0.51	0.66
Federal funds rates						
Posterior mean	3.67	3.86	3.65	3.95	3.68	3.88
St. deviation	0.63	0.69	0.63	0.73	0.63	0.69

***Notes*:** 'no VS' denotes the models without variable selection while 'with VS' denotes the models with variable selection.

Table 2 Posterior means of γ vector elements using the full sample						
	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}	0.16	0.03	0.64	0.09	0.70	0.22
i_{t-1}	0.06	0.02	0.62	0.08	0.62	0.24
r_{t-2}	0.36	0.45	0.67	0.29	0.76	0.45
π_{t-2}	0.11	0.03	0.69	0.14	0.68	0.24
i_{t-2}	0.10	0.02	0.88	0.13	0.65	0.24
r_{t-3}	0.86	0.30	0.86	0.32	0.93	0.38
π_{t-3}	0.07	0.03	0.71	0.16	0.65	0.24
i_{t-3}	0.08	0.03	0.75	0.17	0.64	0.25
r_{t-4}	0.15	0.09	0.63	0.22	0.70	0.26
π_{t-4}	0.07	0.03	0.73	0.18	0.66	0.26
i_{t-4}	0.06	0.02	0.83	0.19	0.65	0.24
Dependent variable: π_t						
r_{t-1}	0.10	0.05	0.58	0.16	0.68	0.23
π_{t-1}	1.00	1.00	1.00	1.00	1.00	1.00
i_{t-1}	0.67	0.03	0.93	0.11	0.83	0.23
r_{t-2}	0.06	0.04	0.69	0.20	0.65	0.24
π_{t-2}	0.93	0.17	0.98	0.19	0.96	0.26
i_{t-2}	0.18	0.02	0.74	0.15	0.68	0.25
r_{t-3}	0.06	0.04	0.71	0.22	0.67	0.24
π_{t-3}	0.16	0.02	0.64	0.09	0.67	0.19
i_{t-3}	0.18	0.02	0.75	0.16	0.65	0.21
r_{t-4}	0.05	0.04	0.74	0.23	0.65	0.24
π_{t-4}	0.09	0.02	0.60	0.09	0.67	0.19
i_{t-4}	0.81	0.02	0.96	0.17	0.85	0.19
Dependent variable: i_t						
r_{t-1}	0.44	0.04	0.67	0.14	0.77	0.23
π_{t-1}	0.35	0.14	0.50	0.16	0.76	0.34
i_{t-1}	1.00	1.00	1.00	1.00	1.00	1.00
r_{t-2}	0.35	0.05	0.75	0.19	0.75	0.23
π_{t-2}	0.67	0.21	0.89	0.39	0.88	0.42
i_{t-2}	0.29	0.01	0.89	0.06	0.71	0.20
r_{t-3}	0.15	0.05	0.72	0.22	0.71	0.24
π_{t-3}	0.21	0.09	0.73	0.25	0.68	0.29
i_{t-3}	0.13	0.02	0.71	0.08	0.65	0.24
r_{t-4}	0.06	0.04	0.71	0.22	0.66	0.24
π_{t-4}	0.11	0.04	0.71	0.20	0.63	0.24
i_{t-4}	0.07	0.02	0.53	0.09	0.67	0.23

***Notes*: '**Uninformative' denotes models with uninformative priors on steady-state.

Table 3 Synopsis of the Model Confidence Set (MCS) results

		GDP		Inflation		Interest rates	
		% of times in the MCS	% of times a model ranks first	% of times in the MCS	% of times a model ranks first	% of times in the MCS	% of times a model ranks first
-	VAR-OLS	13	0	21	0	75	21
Ridge	VAR	13	0	0	0	0	0
	VAR-STEADY	54	0	17	0	67	25
	VAR-VS	38	0	4	0	0	0
	VAR-VS-STEADY	96	46	**100**	**100**	88	0
Minnesota	VAR	13	0	8	0	0	0
	VAR-STEADY	50	0	17	0	71	13
	VAR-VS	8	0	8	0	0	0
	VAR-VS-STEADY	96	**50**	21	0	**96**	**42**
Shrink	VAR	8	0	4	0	0	0
	VAR-STEADY	58	0	21	0	67	0
	VAR-VS	17	0	4	0	0	0
	VAR-VS-STEADY	**100**	4	25	0	88	0

***Notes*:** For each of the variables the table shows the percentage (%) of times a model is included in the MCS at 10% significance level as well as the percentage (%) of times a model ranks first across evaluation metrics (RMSFE and MAFE) and forecasting horizons. VAR-OLS is a standard VAR(1) estimated using ordinary least squares (OLS). VAR is a Bayesian VAR(4) model and the suffices '-STEADY', '-VS' and '-VS-STEADY' denote a steady-state VAR model, a VAR model with variable selection and a steady-state VAR model with variable selection, respectively. Bold faced number indicate the best performing model.

Table D.1 Model confidence set (MCS) results for GDP

Panel A. RMSFE results

Prior	Specification	Forecasting horizon (quarters ahead) 1	2	3	4	5	6	7	8	9	10	11	12
-	VAR-OLS	✓											
Ridge	VAR	✓											
	VAR-STEADY	✓											✓
	VAR-VS	✓											
	VAR-VS-STEADY	✓✓	✓✓	✓	✓	✓	✓	✓	✓		✓	✓✓	✓✓
Minnesota	VAR	✓											
	VAR-STEADY	✓											✓
	VAR-VS	✓											
	VAR-VS-STEADY	✓	✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓	✓
Shrink	VAR	✓											
	VAR-STEADY	✓											✓
	VAR-VS	✓											
	VAR-VS-STEADY	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Panel B. MAFE results

Prior	Specification	1	2	3	4	5	6	7	8	9	10	11	12
-	VAR-OLS	✓			✓								
Ridge	VAR											✓	✓
	VAR-STEADY	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓
	VAR-VS	✓	✓		✓				✓	✓	✓	✓	✓
	VAR-VS-STEADY	✓✓	✓✓	✓	✓	✓	✓	✓✓	✓✓	✓	✓✓	✓✓	✓✓
Minnesota	VAR	✓											✓
	VAR-STEADY	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓
	VAR-VS	✓											
	VAR-VS-STEADY	✓	✓	✓✓	✓✓	✓✓	✓✓	✓	✓	✓	✓	✓	
Shrink	VAR	✓											
	VAR-STEADY	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	VAR-VS	✓	✓		✓								
	VAR-VS-STEADY	✓	✓	✓	✓	✓	✓	✓	✓	✓✓	✓	✓	✓

***Notes*:** The MCS p-values are calculated using the *quadratic* test statistic. The symbol ✓ denotes that a model belongs to the MCS because its p-value is greater than the prespecified significance level, a, where $a = 0.10$. The symbol ✓✓ denotes that the respective model also ranks first among its counterparts, i.e. minimizes the RMSFE or MAFE metrics or maximizes the MCS p-values. VAR-OLS is a standard VAR(1) estimated using ordinary least squares (OLS). VAR is a Bayesian VAR(4) model and the suffices '-STEADY', '-VS' and '-VS-STEADY' denote a steady-state VAR model, a VAR model with variable selection and a steady-state VAR model with variable selection, respectively.

Table D.2 Model confidence set (MCS) results for inflation

Prior	Specification	1	2	3	4	5	6	7	8	9	10	11	12
Panel A. RMSFE results													
		Forecasting horizon (quarters ahead)											
-	VAR-OLS	✓	✓	✓									
Ridge	VAR												
	VAR-STEADY	✓	✓	✓									
	VAR-VS												
	VAR-VS-STEADY	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓
Minnesota	VAR	✓											
	VAR-STEADY	✓	✓	✓									
	VAR-VS	✓											
	VAR-VS-STEADY	✓	✓	✓									
Shrink	VAR												
	VAR-STEADY	✓	✓	✓									
	VAR-VS												
	VAR-VS-STEADY	✓	✓	✓									
Panel B. MAFE results													
-	VAR-OLS	✓	✓										
Ridge	VAR												
	VAR-STEADY	✓											
	VAR-VS	✓											
	VAR-VS-STEADY	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓
Minnesota	VAR	✓											
	VAR-STEADY	✓											
	VAR-VS	✓											
	VAR-VS-STEADY	✓	✓										
Shrink	VAR	✓											
	VAR-STEADY	✓	✓										
	VAR-VS	✓											
	VAR-VS-STEADY	✓	✓	✓									

Notes: See notes in Table D.1

Table D.3 Model confidence set (MCS) results for Federal funds rate

Panel A. RMSFE results

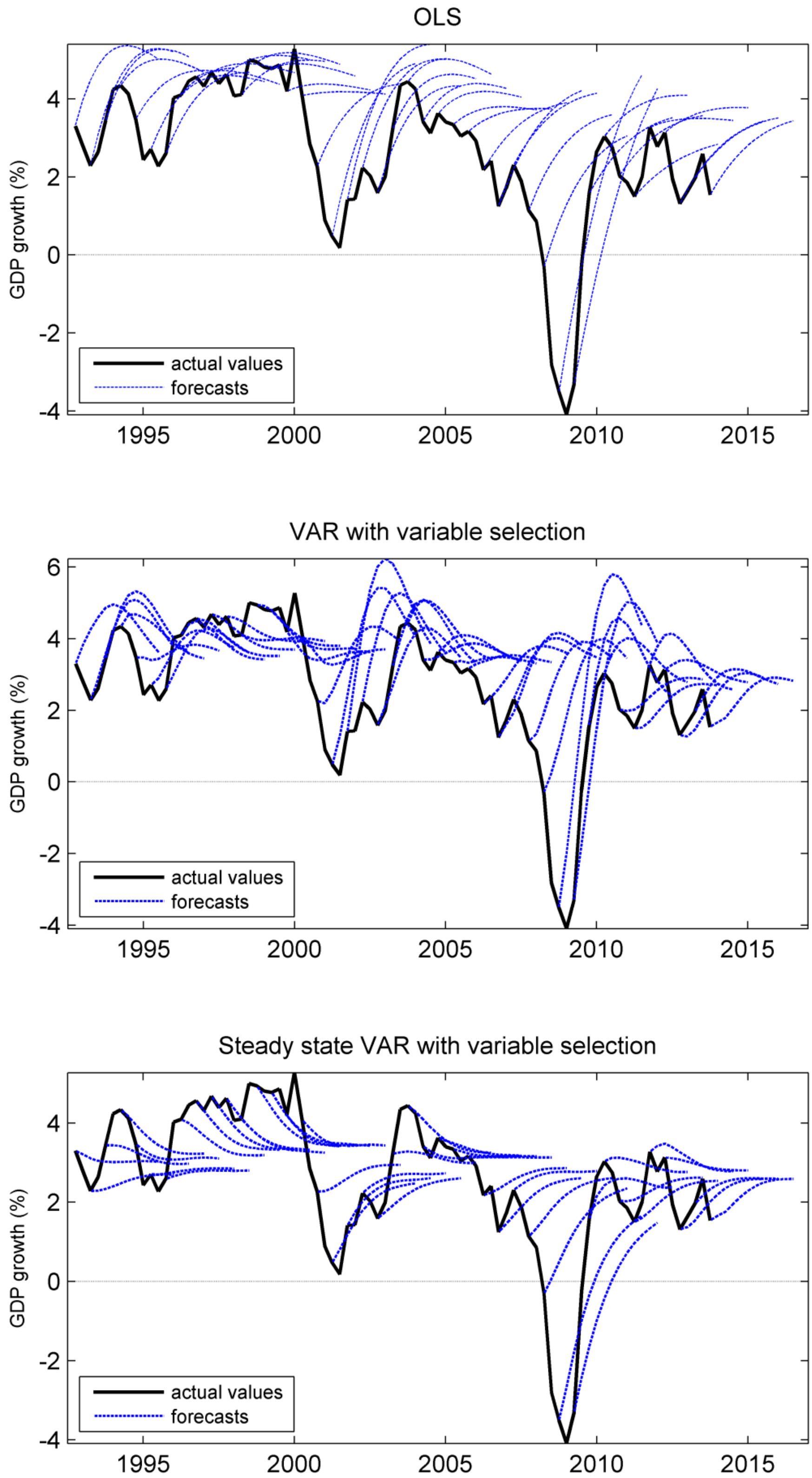
Prior	Specification	Forecasting horizon (quarters ahead)											
		1	2	3	4	5	6	7	8	9	10	11	12
-	VAR-OLS	✓✓	✓✓	✓✓	✓✓	✓✓	✓	✓	✓	✓			
Ridge	VAR												
	VAR-STEADY				✓	✓	✓	✓	✓	✓✓	✓✓	✓✓	✓✓
	VAR-VS												
	VAR-VS-STEADY	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Minnesota	VAR												
	VAR-STEADY	✓			✓	✓	✓✓	✓✓	✓✓	✓	✓	✓	✓
	VAR-VS												
	VAR-VS-STEADY	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Shrink	VAR												
	VAR-STEADY	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	VAR-VS												
	VAR-VS-STEADY	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Panel B. MAFE results

Prior	Specification	1	2	3	4	5	6	7	8	9	10	11	12
-	VAR-OLS	✓	✓	✓	✓	✓	✓	✓	✓	✓			
Ridge	VAR												
	VAR-STEADY						✓	✓	✓	✓	✓	✓✓	✓✓
	VAR-VS												
	VAR-VS-STEADY	✓	✓				✓	✓	✓	✓	✓	✓	✓
Minnesota	VAR												
	VAR-STEADY						✓	✓	✓	✓	✓	✓	✓
	VAR-VS												
	VAR-VS-STEADY	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓	✓
Shrink	VAR												
	VAR-STEADY						✓	✓	✓			✓	✓
	VAR-VS												
	VAR-VS-STEADY	✓	✓				✓	✓	✓	✓	✓	✓	✓

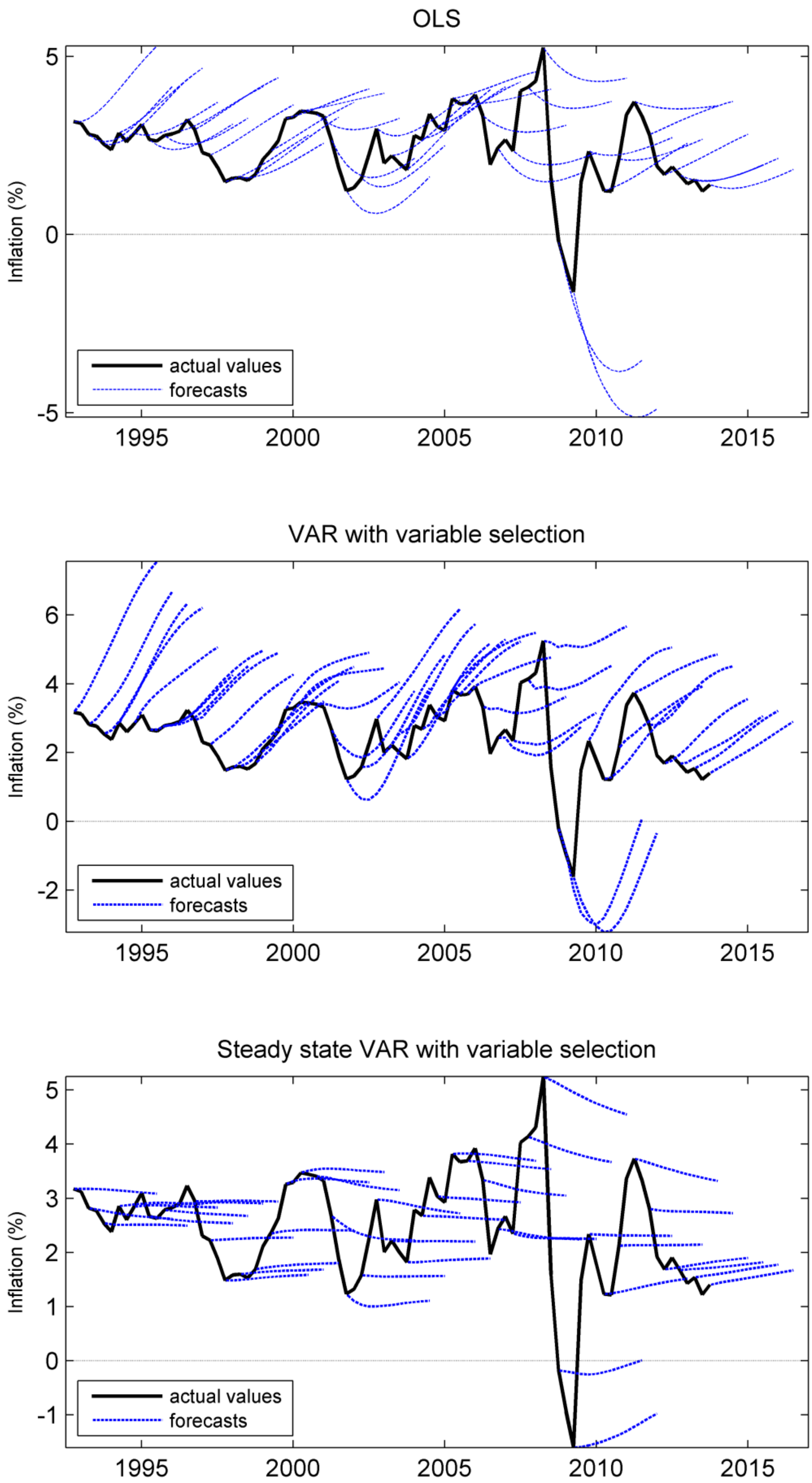
***Notes*:** See notes in Table D.1

Figure 1 Out-of-sample point forecasts for GDP growth



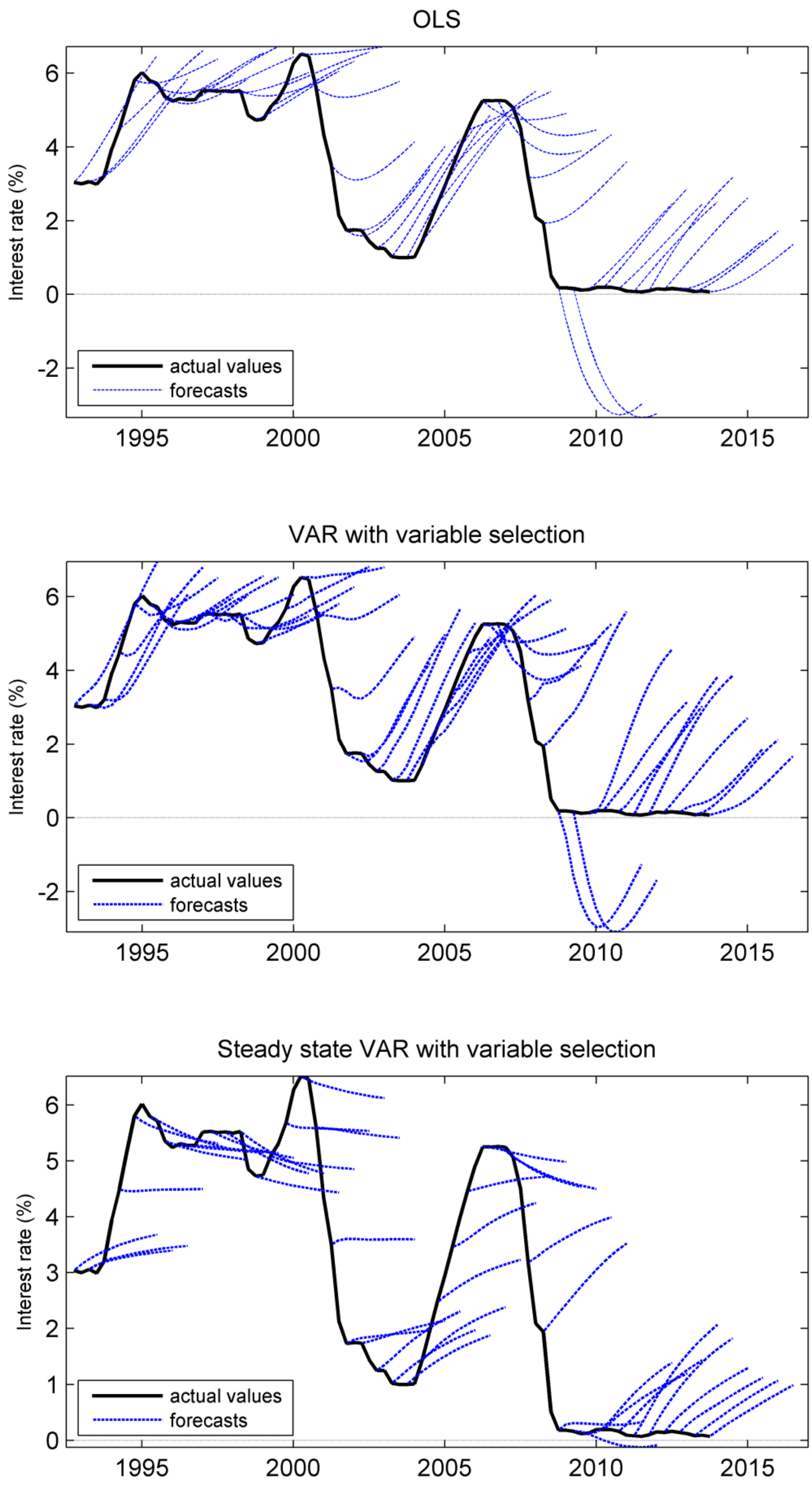
***Notes*:** The Figure presents the GDP growth (year-on-year) out-of-sample forecasts for 12 quarters ahead (dotted lines) along with the actual values (solid line). The forecasts are plotted every other quarter for clarity in presentation. For the BVAR models we use the Minnesota prior for the dynamic coefficients.

Figure 2 Out-of-sample point forecasts for inflation



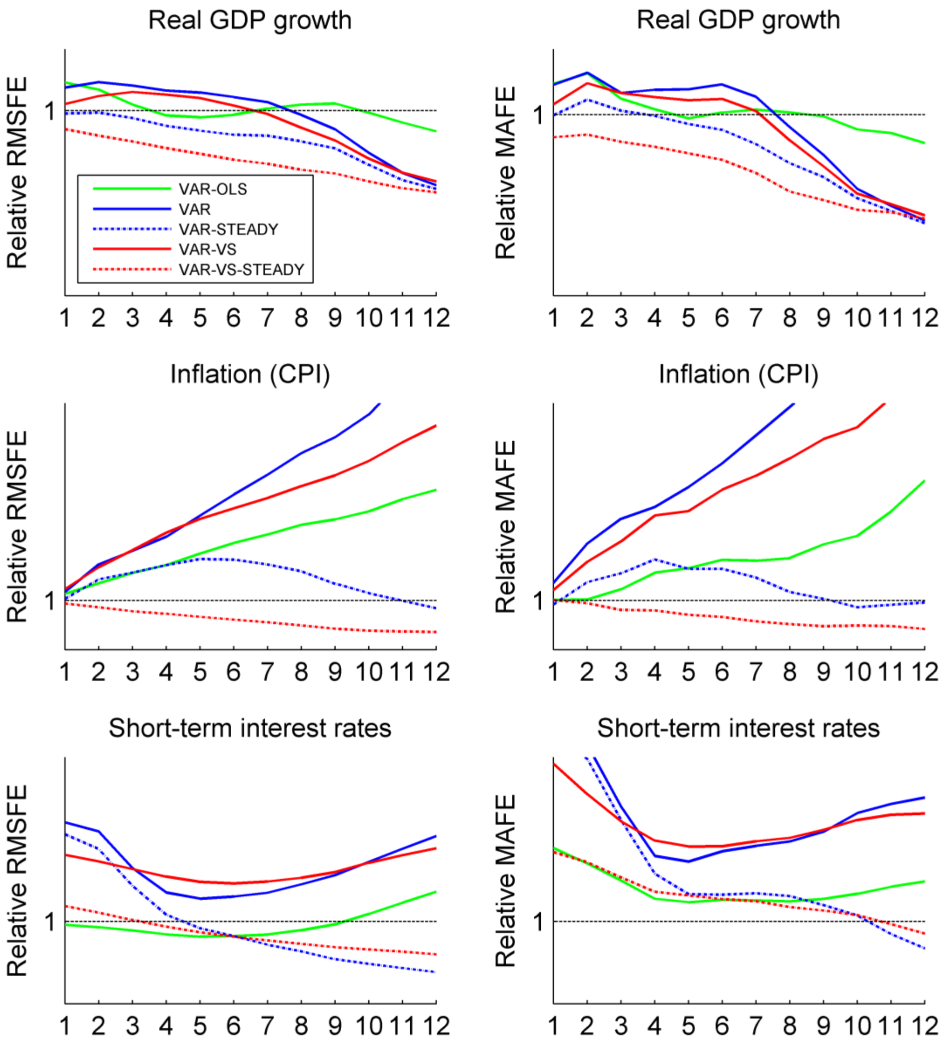
***Notes*:** The Figure presents the CPI inflation (year-on-year change) out-of-sample forecasts for 12 quarters ahead (dotted lines) along with the actual values (solid line). The forecasts are plotted every other quarter for clarity in presentation. For the BVAR models we use the Minnesota prior for the dynamic coefficients.

Figure 3 Out-of-sample point forecasts for Federal funds rate



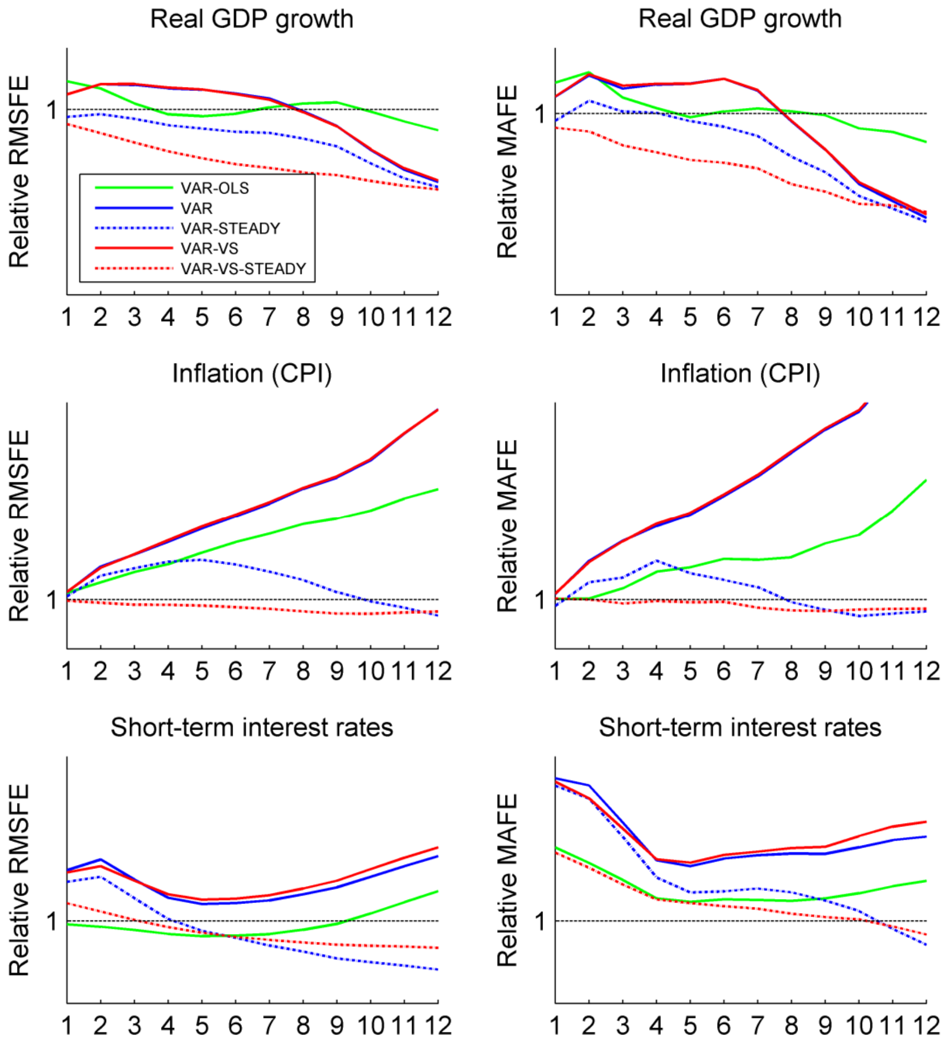
***Notes*:** The Figure presents the Federal funds rate out-of-sample forecasts for 12 quarters ahead (dotted lines) along with the actual values (solid line). The forecasts are plotted every other quarter for clarity in presentation. For the BVAR models we use the Minnesota prior for the dynamic coefficients.

Figure 4 Forecasting results using VAR models with ridge prior



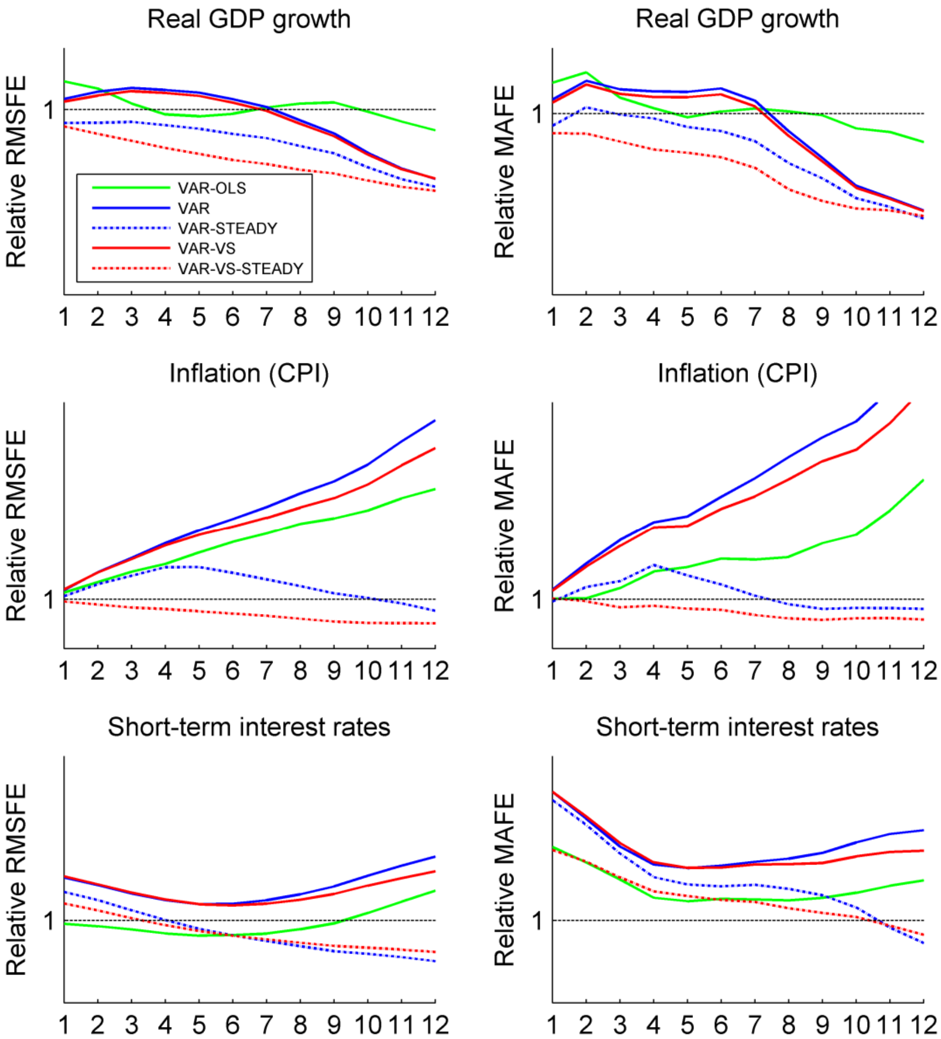
***Notes*:** The figures show the relative RMSFE and MAFE as a function of the forecasting horizon. VAR-OLS is a standard VAR(1) estimated using ordinary least squares (OLS). VAR is the Bayesian VAR(4) with a Ridge regression prior. The suffices '-STEADY', '-VS' and '-VS-STEADY' denote a steady-state VAR model, a VAR model with variable selection and a steady-state VAR model with variable selection, respectively.

Figure 5 Forecasting results using VAR models with Minnesota prior



Notes**:** The figures show the relative RMSFE and MAFE as a function of the forecasting horizon. VAR-OLS is a standard VAR(1) estimated using ordinary least squares (OLS). VAR is the Bayesian VAR(4) with a Minnesota regression prior. The suffices '-STEADY', '-VS' and '-VS-STEADY' denote a steady-state VAR model, a VAR model with variable selection and a steady-state VAR model with variable selection, respectively.

Figure 6 Forecasting results using VAR models with shrink prior



***Notes*:** The figures show the relative RMSFE and MAFE as a function of the forecasting horizon. VAR-OLS is a standard VAR(1) estimated using ordinary least squares (OLS). VAR is the Bayesian VAR(4) with a shrink regression prior. The suffices '-STEADY', '-VS' and '-VS-STEADY' denote a steady-state VAR model, a VAR model with variable selection and a steady-state VAR model with variable selection, respectively.

BANK OF GREECE WORKING PAPERS

185. Adam, A., and T., Moutos, "Industry-Level Labour Demand Elasticities Across the Eurozone: Will There Be Any Gain After the Pain of Internal Devaluation?" July, 2014.

186. Tagkalakis, O.A., "Fiscal Policy, Net Exports, and the Sectoral Composition of Output in Greece", September 2014.

187. Hondroyiannis, G. and D., Papaoikonomou, "When Does it Pay To Tax? Evidence from State-Dependent Fiscal Multipliers in the Euro Area", October 2014.

188. Charalambakis, C. E., "On Corporate Financial Distress Prediction: What Can we Learn From Private Firms in a Small Open Economy?, November 2014.

189. Pagratis, S., E., Karakatsani and E. Louri, "Bank Leverage and Return on Equity Targeting: Intrinsic Procyclicality of Short-Term Choices", November 2014.

190. Evgenidis, A. and C., Siriopoulos, "What are the International Channels Through Which a US Policy Shock is Transmitted to the World Economies? Evidence from a Time Varying Favar, January 2015.

191. Louzis, D. P., and A.T., Vouldis, "Profitability in the Greek Banking System: a Dual Investigation of Net Interest and Non-Interest Income", February 2015.

192. Papaspyrou, S.T, "EMU 2.0 - Drawing Lessons From the Crisis - a New Framework For Stability and Growth", March 2015.

193. Litina, A and T, Palivos, "Corruption and Tax Evasion: Reflections on Greek Tragedy", June 2015.

194. Balfoussia, H. and H.D. Gibson, "Financial Conditions and Economic Activity: The Potential Impact of the Targeted Longer-Term Refinancing Operations (TLTROS)", July 2015.