Macroeconomic forecasting and structural changes in steady states

Dimitrios P. Louzis
MACROECONOMIC FORECASTING AND STRUCTURAL CHANGES IN STEADY STATES

Dimitrios P. Louzis
Bank of Greece

Abstract
This article proposes methods for estimating a Bayesian vector autoregression (VAR) model with an informative steady state prior which also accounts for possible structural changes in the long-term trend of the macroeconomic variables. I show that, overall, the proposed time-varying steady state VAR model can lead to superior point and density macroeconomic forecasting compared to constant steady state VAR specifications.

JEL classification numbers: C32

Keywords: Steady states; time-varying parameters; macroeconomic forecasting

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
21 E.Venizelos Avenue,
GR 10250, Athens, Greece
e-mail: dlouzis@bankofgreece.gr; dlouzis@aueb.gr
1. Introduction

Bayesian vector autoregressions (VARs) have been established as a standard forecasting tool for macroeconomists who have, now, plenty of choices in terms of priors and model specification. In the seminal study of Villani (2009), the author highlights the importance of steady state priors in macroeconomic forecasting and proposes Bayesian methods for estimating VAR models with informative steady state priors. These models have been shown to materially improve macroeconomic forecasting accuracy (Beechey and Österholm 2010; Clark, 2011; Wright, 2013).

However, as shown in Chan and Koop (2014), steady states may undergo significant structural changes. The authors, based on Bayesian clustering methods, find that the steady states of inflation and the interest rate in the US economy changed during the 1970s. Structural breaks in the steady states have also been implicitly taken into account by Clark (2011) and Wright (2013), who employed the long-term macroeconomic expectations of Blue Chip surveys, in order to impose steady state priors on the endogenous variables.

The purpose of this study is to examine whether macroeconomic forecasting can be substantially improved when we explicitly account for the structural changes in the steady states. In particular, this note proposes a time-varying steady state VAR (TVSS-VAR) model which is estimated using Bayesian methods employed in the time-varying parameter VAR (TVP-VAR) models (Primiceri, 2005). Recent empirical evidence suggests that TVP-VARs can adequately capture structural changes in terms of macroeconomic forecasting (D’Agostino et al., 2013; Carriero et al. 2015; Bauwens et al. 2015). Here, I focus solely on the time-varying nature of the unconditional mean and the ability of the TVSS-VAR to deliver superior macroeconomic forecasts compared to a standard (constant) steady state VAR.

The paper is structured as follows: Section 2 presents the model; Section 3 describes the dataset, estimation and out-of-sample forecasting results; and Section 4 concludes.
2. The model

Villani (2009) proposes the following steady state representation of a VAR model (SS-VAR):

\[
(y_t - \mu d_t) = B_1(y_{t-1} - \mu d_{t-1}) + \ldots + B_p(y_{t-p} - \mu d_{t-p}) + \varepsilon_t \quad \text{with} \quad \varepsilon_t \sim i.i.d.N(0, \Sigma)
\]

(1)

where \( y_t \) is a \( m \times 1 \) vector of endogenous variables, \( \mu d_t \) is the unconditional mean of the process assuming stationarity for \( y_t \), with \( d_t \) being a \( q \times 1 \) vector of exogenous deterministic variables such as constants, dummies or trends. Obviously, this representation allows the economist to incorporate his steady state prior beliefs by directly specifying priors on \( \mu \).

Here, I extend the SS-VAR model in order to account for possible structural changes in the steady states. In particular, I define a time-varying steady state VAR (TVSS-VAR) model as follows:

\[
(y_t - \mu_t d_t) = B_1(y_{t-1} - \mu_{t-1} d_{t-1}) + \ldots + B_p(y_{t-p} - \mu_{t-p} d_{t-p}) + \varepsilon_t
\]

(2)

Following the TVP-VAR literature, I specify the dynamics of the steady states parameters, \( \mu_t \), as driftless random walks (e.g. see Primiceri, 2005):

\[
vec(\mu_t) = vec(\mu_{t-1}) + \eta_t \quad \text{with} \quad \eta_t \sim i.i.d.N(0, Q_\mu)
\]

(3)

where \( vec(\cdot) \) is the standard operator which stacks the columns of a matrix. The error terms, \( (\varepsilon_t, \eta_t) \), are assumed to be jointly normally distributed with zero mean and covariance matrix defined as:

\[
\text{Var}\left[\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix}\right] = \begin{bmatrix} \Sigma & 0 \\ 0 & Q_\mu \end{bmatrix}
\]

(4)

2.1. Bayesian inference

The TVSS-VAR model is estimated using Bayesian methods. The exact Gibbs sampling algorithm and the conditional posterior distributions are presented in the
Technical Appendix. Here, I concentrate more on the specification of the priors which is an essential first step in Bayesian inference.

Typically, the prior for the time invariant dynamic coefficients and the errors covariance matrix is a multivariate normal distribution and a scale-invariant improper Jeffreys prior, respectively

$$\beta \sim N_{m^2p} \left( \underline{\beta}, \underline{\Sigma} \right)$$

$$\Sigma \propto \left| \Sigma \right|^{-(m+1)/2}$$

where $\beta = \text{vec}(B)$ and $B = \left[ B_1, \ldots, B_p \right]$. Moreover, for the dynamic coefficients, $\beta$, I specify a Minnesota prior with overall and cross-equation tightness being set to 0.2 and 0.5, respectively, while the prior mean on the first own lag is set to 0.25 for GDP and 0.8 for inflation and the interest rate (Clark, 2011).

The priors for the initial state (i.e. initial condition) of the time-varying steady states and the covariance matrix $Q_\mu$ are a multivariate normal and an inverse-Wishart (IW) distribution, respectively

$$\text{vec}(\mu_0) \sim N_{mq} \left( \mu, \underline{\Omega} \right)$$

$$Q_\mu \sim IW \left( \underline{\Gamma}, \xi \right)$$

where $\xi$ is the prior degrees of freedom and $\underline{\Gamma}$ is the $mq \times mq$ prior scale matrix. The prior degrees of freedom are chosen to be equal to $mq+1$, i.e. the minimum necessary for reassuring a proper prior, while the prior scale matrix is defined as $\underline{\Gamma} = k^2 \xi \underline{\Omega}$ (Primiceri, 2005). In practice, prior hyperparameters $\mu$ and $k$ encapsulate economist’s prior beliefs with regard to the level and the degree of time-variation of the steady states; therefore, they should be sensibly specified in order to avoid implausible behaviour and optimize models’ forecasting performance.

\[1\] Notice that the prior for each $\mu_t$, $t = 1, \ldots, T$ is implicitly defined recursively as $\text{vec}(\mu_t) \sim N_{mq} \left( \text{vec}(\mu_{t-1}), Q_\mu \right)$. 

5
More specifically, for the initial conditions of the steady states I rely on the recent literature (Clark, 2011; Österholm, 2012) and economic theory (Fisher equation) and I set \( \mu = [3,2,5] \), expressed in annualized percentage points, and
\[
\Omega = \text{diag} \left( [0.5^2,0.5^2,0.7^2] \right),
\]
where \( \text{diag} \) is a diagonal matrix. As regards the prior scale matrix, the TVP-VAR literature, typically, chooses a very tight prior (i.e. small \( k \) ) in order to avoid implausible behaviour and allow for a smooth variation of the time-varying dynamic coefficients. Here, I set \( k = 0.4 \), a relatively large value, in order to capture possibly large structural changes in the unconditional mean. Nonetheless, empirical results are also presented for other values of \( k \) (see Section 3).

3. Empirical results

I evaluate the forecasting performance of the proposed model using three US macroeconomic series, i.e. real GDP growth rate, the CPI inflation rate and the effective Federal funds rate (GDP, inflation and the interest rate hereafter), spanning the period 1954:Q2 to 2015:Q1 on a quarterly basis.\(^2\) GDP and inflation are calculated as annualized quarter-on-quarter percentage changes while the interest rate is used in levels. For all models considered in this article I use four lags and I run the Gibbs sampler for 12,000 iterations discarding the first 10,000 iteration and keeping one in every five draws (D’Agostino et al. 2013).

Next, I present some estimation results for the steady states. In particular, Figure 1 depicts the posterior median of the time-varying steady states for the full sample using different values for the \( k \) hyperparmeter, \( k = [0.01,0.1,0.2,0.4,0.8,1] \), along with a constant steady state VAR estimated as in Villani (2009). The larger the \( k \) hyperparmeter, the larger are the changes in the steady-states, as expected. For a very small value of \( k \), i.e. for \( k = 0.01 \) (black dashed line), steady states are almost time-invariant as in the case of the SS-VAR model (blue line).

Estimation results are also in line with the stylized facts regarding the US economy. More specifically, inflation and the interest rate have an upward trend in the 1970s, i.e.

\(^2\) The series were downloaded from http://research.stlouisfed.org/fred2.
the Great Inflation period, and a subsequent de-escalation in the 1980s during the Great Moderation period $\forall k > 0.01$. Moreover, the time-varying steady states for both variables fall well below the constant steady state (blue line) during the recent Great Recession (2007-2009). Regarding GDP, the time-varying steady state fluctuates around the constant steady state (blue line) for most of the time, with the exception of the Great Recession, during which falls permanently in a lower level. Therefore, it seems that following the Great Recession the US economy has probably entered into a new regime characterized by lower growth, inflation and interest rates.

3.1. Out-of-sample forecasting analysis

The primary scope of the study is to examine whether the forecasting ability of a standard steady state VAR model is improved when we account for structural changes in the steady state. To that end, I compare the forecasting ability of the proposed TVSS-VAR model with the SS-VAR of Villani (2009) and a baseline Bayesian VAR (BVAR) with uninformative steady state priors. Finally, following standard practice, I use as a benchmark model a random walk with drift in levels (RW).

I consider a forecasting horizon of twelve quarters, $h = 1, \ldots, 12$, and an out-of-sample period spanning from 1969:Q4 to 2015:Q1. The forecasts across all horizons are generated iteratively using a recursive forecasting scheme. Lastly, following D’Agostino et al. (2013) I assume that the drifting steady states remain constant at their current level when I consider multi-step forecasts.

Following the recent contributions in the field, I evaluate the forecasting performance of the models in terms of both point and density forecasts (e.g see Carriero et al., 2015). More specifically, I evaluate the point forecasts using the following metric:

$$\text{Relative \ RMSFE}_{i,h}^m = \left(1 - \frac{\text{RMSFE}_{i,h}^m}{\text{RMSFE}_{i,h}^{RW}}\right) \times 100$$

where $\text{RMSFE}_{i,h}^m = \sqrt{N^{-1} \sum_{j=1}^{N} \left(f_{i,j+h}^m - f_{i,j+h}^m \right)^2}$ with $f_{i,j+h}^m$ being the forecast error for model $m$, variable $i$, and forecast horizon $h$, while $N$ is the number of out-of-sample observations. For the density forecasts I use:
\[ \Delta \text{SCORE}^m_{t,h} = (\text{SCORE}^m_{t,h} - \text{SCORE}^{RW}_{t,h}) \times 100 \] (6)

where \( \text{SCORE}^m_{t,h} = N^{-1} \sum_{i=1}^N \log \tilde{p}(y_{i,t+h} | y_t, m) \), with \( \tilde{p}(y_{i,t+h} | y_t, m) \) being the predictive density produced by model \( m \) and evaluated at the realized value of \( y_{i,t+h} \). The predictive density is obtained using a univariate kernel estimation based on the Gibbs sampling output.

[Insert Figure 1]

In practice, both metrics in Eqs. (5) and (6) show the percentage gains or losses over the benchmark (RW) model in terms of RMSFE and \( \text{SCORE} \). Obviously the highest the value of the metric, the better the model is, while positive values indicate that the model outperforms the benchmark.

[Insert Figure 2]

Figure 2 presents the values for both evaluation measures as a function of the forecasting horizon. Overall, the results suggest that the TVSS-VAR (with \( k = 0.4 \)) does outperform its counterparts in terms of both point and density forecasting evaluation. More specifically, regarding the relative RMSFE metric, the proposed model ranks first across almost all forecasting horizons for GDP and inflation followed by the constant steady state counterpart and the BVAR. The picture is different regarding the interest rate where the SS-VAR specification outperforms the TVSS-VAR and ranks first across all horizons. Nonetheless, acknowledging time-variation in the steady states is much more beneficial in terms of density forecasting since the improvement of the proposed model over its competitors is impressive. In particular, the TVSS-VAR outranks its competitors across all horizons and variables. It is also worth noting that TVSS-VAR and SS-VAR outperform the RW model in all instances examined here.

[Insert Figure 3]

I also investigate the robustness of the forecasting results by performing two sensitivity checks. The first is with respect to the forecasting period: Figure 3 presents the results for the 1985:Q1-2015:Q1 period for which macroeconomic forecasting is generally more difficult (Faust and Wright, 2013). Forecasting results are qualitatively
similar to those presented for the full forecasting period. The second check concerns the hyperaparameter $k$: Figure 4 presents the forecasting results for $k = \{0.01, 0.1, 0.2, 0.4, 0.8, 1\}$ along with the constant steady state VAR. Overall, the results are robust to the different choices of $k$ with the only exceptions being for point forecasts of GDP and inflation and for $k \leq 0.2$.

[Insert Figure 4]

4. Conclusions

This article proposes a time-varying steady state VAR model which accounts for structural changes in the steady state level of macroeconomic variables. Using three US macroeconomic variables I show that, overall, the proposed model materially improves both point and density forecast accuracy compared to a constant steady state VAR.
Figure 1 Time-varying and constant steady states estimates for the full sample.
Figure 2 Forecasting results over the sample: 1969:Q4-2015:Q1

Figure 3 Forecasting results over the sample: 1985:Q1-2015:Q1
Figure 4 Forecasting results with different values of $k$ hyperparameter
Appendix

A. Technical Appendix: Posterior sampling for the time-varying steady state VAR model

This Appendix presents the Gibbs sampler algorithm for the time-varying steady state VAR (TVSS-VAR) model. For convenience, I rewrite the model and the priors as specified in the main text. Thus, the TVSS-VAR model is defined as:

\[
(y_t - \mu_t d_t) = B_1 (y_{t-1} - \mu_{t-1} d_{t-1}) + \ldots + B_p (y_{t-p} - \mu_{t-p} d_{t-p}) + \varepsilon_t \quad \text{with} \quad \varepsilon_t \sim N(0, \Sigma) \quad (A.7)
\]

\[
\text{vec}(\mu_t) = \text{vec}(\mu_{t-1}) + \eta_t \quad \text{with} \quad \eta_t \sim N(0, Q_\mu) \quad (A.8)
\]

and

\[
\begin{bmatrix} e_t \\ \eta_t \end{bmatrix} \sim i.i.d. N \begin{pmatrix} 0 & \Sigma \\ 0 & 0 \\ 0 & Q_\mu \end{pmatrix} \quad (A.9)
\]

where \( y_t \) is a \( m \times 1 \) vector of endogenous variables, \( B_i, i = 1, \ldots, p \) are \( m \times m \) dynamic coefficient matrices and \( d_t \) is a \( q \times 1 \) vector of exogenous deterministic variables such as constants, dummies or trends. The priors for the time invariant parameters and the initial state of the state equation are:

\[
\beta \sim N_{m \times p} (\beta, \Sigma) \quad (A.10)
\]

\[
\Sigma \propto |\Sigma|^{-(m+1)/2} \quad (A.11)
\]

\[
\text{vec}(\mu_0) \sim N_{mq} (\mu, \Omega) \quad (A.12)
\]

\[
Q_\mu \sim IW(\Gamma, \xi) \quad (A.13)
\]

where \( \beta = \text{vec}(B) \) and \( B = [B_1, \ldots, B_p] \), \( IW \) is the inverse-Wishart distribution, \( \xi = mq + 1 \) and \( \Gamma = k^\frac{\xi}{2} \Omega \). All priors are assumed to be independent from each other.

For computational reasons it is also useful to rewrite the TVSS-VAR model using matrix notation, i.e.:

\[
\tilde{Y} = \tilde{X} B + E \quad (A.14)
\]
where $\tilde{Y}$ and $E$ are defined as a $T \times m$ matrices with their $t^{th}$ row being $\tilde{Y}_t = \tilde{y}_t = [\tilde{y}_{1t}, ..., \tilde{y}_{mt}]$ and $E_t = \varepsilon_t' = [\varepsilon_{1t}, ..., \varepsilon_{mt}]$, respectively, while $\tilde{y}_t$ is the mean-adjusted time-series of $y_t$ defined as $\tilde{y}_t = y_t - \mu_t d_t$. Finally, the $t^{th}$ row of the $T \times k$ matrix $\tilde{X}_t$ is defined as $\tilde{X}_t = \tilde{x}_t = [\tilde{x}_{1t}, ..., \tilde{x}_{kt}]$.

The TVSS-VAR model is estimated using the Gibbs sampling algorithm outlined below. The basic idea is that conditional on steady states, $\mu_t \forall t$, the model is a standard Bayesian VAR model for the mean-adjusted variables and thus standard results can be applied (e.g. see Villani, 2009). On the other hand, conditional on other parameters, I show that the TVSS-VAR model can be written in a form of a standard TVP-VAR model of Primiceri (2005) and thus I can estimate the unobserved time-varying steady states by applying the methods proposed by the author. More specifically, the Gibbs sampler involves drawing sequentially from the following conditional posteriors:

1. Sample $\beta$ conditional on other model parameters and data from

$$\beta | \mu^T, Q_\mu, \Sigma, D \sim N(\beta, V)$$

with $D = \{y_t, d_t, ..., d_T\}$, $\mu^T = \{\mu_t \}_{t=1}^T$, $\beta = \bar{V} = \text{vec}(\tilde{X}' \tilde{Y} \Sigma^{-1}) + V^{-1} \beta$ and $V = [\Sigma^{-1} \otimes (\tilde{X}' \tilde{X}) + V^{-1}]^{-1}$ and $\otimes$ denoting the kronecker product.

2. Sample $\Sigma$ conditional on other model parameters and data from

$$\Sigma | \beta, \mu^T, Q_\mu, D \sim IW(\bar{E}' \bar{E}, T)$$

where $E = (\bar{Y} - \tilde{X}B)$.


For this step the TVSS-VAR model should be rewritten in a form that makes Bayesian inference tractable.
**Proposition A.1.** The TVSS-VAR model can be written as:

\[ y_t^B = Z_t \text{vec}(\mu_t) + \varepsilon_t, \]

(A.15)

\[ \text{vec}(\mu_t) = \text{vec}(\mu_{t-1}) + \eta_t, \]

(A.16)

where \( y_t^B = y_t - B_1 y_{t-1} - \ldots - B_p y_{t-p} \), \( Z_t \equiv I_m \otimes D_t U \), \( D_t = [d_t', d_{t-1}', \ldots, d_{t-p}'] \) and

\[
U = \begin{bmatrix}
I_{(mq)} \\
I_q \otimes B_1 \\
\vdots \\
I_q \otimes B_p
\end{bmatrix}.
\]

**Proof.** Rearranging the terms in Eq. (A.1) we get

\[ (y_t - B_1 y_{t-1} - \ldots - B_p y_{t-p}) = \mu_t d_t - B_1 \mu_{t-1} d_{t-1} - \ldots - B_p \mu_{t-p} d_{t-p} + \varepsilon_t \]

(A.17)

Let \( y_t^B = y_t - B_1 y_{t-1} - \ldots - B_p y_{t-p} \), \( \Theta_t = [\mu_t, B_1 \mu_{t-1}, \ldots, B_p \mu_{t-p}] \) be a \( m \times k \) matrix that collects the coefficients with \( k = (p+1)q \), and \( D_t = [d_t', d_{t-1}', \ldots, d_{t-p}'] \) be a \( k \times 1 \) vector that collects deterministic variables \( d_t \). The model can now be written \( y_t^B = \Theta_t D_t + \varepsilon_t \), which is a standard homoscedastic TVP-VAR model. Then, it is straightforward to write the model in the form of Eq. (4) of Primiceri (2005, p. 824):

\[ y_t^B = I_m \otimes D_t \Theta_t + \varepsilon_t \]

(A.18)

where \( \Theta_t = \text{vec}(\Theta_t) = U \text{vec}(\mu_t) \) with \( U \) being the following \( (mk) \times (mq) \) matrix

\[
U = \begin{bmatrix}
I_{(mq)} \\
I_q \otimes B_1 \\
\vdots \\
I_q \otimes B_p
\end{bmatrix} \] (Villani, 2009). ∗

The model in Eqs. (A.9)-(A.10) is a standard homoscedastic TVP-VAR model as defined in Eqs. (4)-(5) of Primiceri (2005, p. 824) and, thus, the Carter and Kho (1994) algorithm can be implemented. For notational simplicity I assume that only a constant is included in the VAR model, i.e. \( d_t = 1 \ \forall t \), implying that \( \text{vec}(\mu_t) = \mu_t \).
Given the assumption in Eq. (A.3) and defining \( \mu_{\cdot|T} = E(\mu_{\cdot}|\mu_{\cdot}, \beta, Q_{\cdot}, \Sigma) \) and \( P_{\cdot|T} = \text{Var}(\mu_{\cdot}|\mu_{\cdot}, \beta, Q_{\cdot}, \Sigma) \), we can obtain \( \mu_{T|T} \) and \( P_{T|T} \) by applying the following Kalman filter recursions:

\[
\begin{align*}
\mu_{t|t-1} &= \mu_{t-1|t-1} \\
P_{t|t-1} &= P_{t-1|t-1} + Q_{\mu} \\
K_t &= P_{t|t-1}Z_t'(Z_tP_{t|t-1}Z_t' + \Sigma)^{-1} \\
\mu_{t|T} &= \mu_{t|t-1} + K_t(y_t^\beta - Z_t\mu_{t|t-1}) \\
P_{t|T} &= P_{t|t-1} - K_tZ_tP_{t|t-1}
\end{align*}
\]

Given the initial conditions, i.e. \( \mu_{0|0} = \mu \) and \( P_{0|0} = Q \) and running the above recursions from \( t = 1, \ldots, T \), we get \( \mu_{T|T} \) and \( P_{T|T} \) and we use them to draw from the posterior of \( \mu_t \) from \( N(\mu_{T|T} , P_{T|T}) \). Then, we can use \( \mu_{T|T} \) and \( P_{T|T} \) as initial values and apply the Carter and Kohn (1994) recursions backwards for \( t = T - 1, \ldots, 1 \) to smooth the initial Kalman filter estimates with subsequent information and obtain \( \mu_{T|t+1} \) and \( P_{T|t+1} \) as follows:

\[
\begin{align*}
\mu_{t+1|T} &= \mu_{t|T} + P_{t|T}(P_{t|T} + Q_{\mu})^{-1}(\mu_{t+1|T} - \mu_{t|T}) \\
P_{t+1|T} &= P_{t|T} - P_{t|T}(P_{t|T} + Q_{\mu})^{-1}P_{t|T}
\end{align*}
\]

Finally, we draw from the posterior of \( \mu_t \), \( t = T - 1, \ldots, 1 \) using a Normal distribution with mean \( \mu_{t|t+1} \) and variance \( P_{t|t+1} \).

4. Sample \( Q_{\mu} \) conditional on other parameters and data from

\[
Q_{\mu}|\mu^T, \Sigma, \beta, D \sim \text{IW}(\overline{F}, \bar{\zeta})
\]

where \( \overline{F} = \sum_{t=1}^{T}(\mu_t - \mu_{t-1})' (\mu_t - \mu_{t-1}) \) and \( \bar{\zeta} = T + \bar{\zeta} \).
References


