Manual to accompany MATLAB package for Bayesian VAR models

Gary Koop
University of Strathclyde

Dimitris Korobilis
University of Strathclyde

Glasgow, September 2009

Contents

1 Introduction

This notes manual accompanies the monograph on empirical VAR models and the associated MATLAB code. The ultimate purpose is to introduce academics, students and applied economists to the world of Bayesian time series modelling combining theory with easily digestable computer code. For that reason, we present code in a format that follows the theoretical equations as close as possible, so that users can make the connection easily and understand the models they are estimating. This means that in some cases the code might not be as computationally efficient as it should be in practice, if there is the danger to sacrifice clarity. We try to avoid structure arrays which can be confusing, that is we only represent our variables as vectors or matrices (the SSVS model is the only exception, where we use the MATLAB cell array capabilities).

The directories in the file BAYES_VARS.zip, are:

$BVAR\_Analytical$	VAR models using analytical results
$BVAR\_GIBBS$	VAR models using the Gibbs sampler
$BVAR\_FULL$	Programs to replicate Empirical Illustration 1. This code uses simulation to get the posterior parameters with a flexible choice of 6 different priors.
$SSVS$	VAR with SSVS mixture prior as in George, Sun and Ni (2008)
$VAR\_Selection$	Variable selection in VARs as in Korobilis (2009b)
TVP_VAR_CK	TVP-VAR model using the Carter and Kohn (1994) smoother as in Primiceri(2005)
TVP_VAR_DK	TVP-VAR model using the Durbin and Koopman (2002) smoother
TVP_VAR_GCK	Mixture innovations TVP-VAR as in Koop, Leon-Gonzales and Strachan (2009)
$HierarchicalTVP\_VAR$	Hierarchical TVP-VAR as in Chib and Greenberg (1995)
$Factor\_Models$	Estimation of static and dynamic factor models
$FAVAR$	FAVAR as in Bernanke, Boivin and Eliasz (2005)
TVP_FAVAR	TVP-FAVAR as in Korobilis (2009a)
$GAUSS2MATLAB$	Some usefull GAUSS routines, transcribed for MATLAB

2 VAR models

2.1 Analytical results for VAR models

The simple, reduced-form VAR model can be writen as

$$Y_t = X_t A + \varepsilon_t, \text{ with } \varepsilon_t \sim N(0, \Sigma) \tag{1}$$

As we have shown in the previous subsection, this model can be written in the form

$$y_t = \ (I_M \otimes X_t)\, \alpha + \varepsilon_t \tag{2}$$

or compactly

$$y_t = \ Z_t \alpha + \varepsilon_t \tag{3}$$

where $a = vec(A)$.

In the computations presented henceforth, we will need the OLS estimates of a, A and Σ. Subsequently, using the notation $X = (X_1, ..., X_T)'$, we define

$$\widehat{\alpha} = \left(\sum Z_t' Z_t\right)^{-1} \left(\sum Z_t' y_t\right) \tag{4}$$

the OLS estimate of a,

$$\widehat{A} = (X'X)^{-1} (X'Y) \tag{5}$$

the OLS estimate of A,

$$\widehat{S} = \left(Y - X\widehat{A}\right)' \left(Y - X\widehat{A}\right) \tag{6}$$

the sum of squared errors of the VAR, and

$$\widehat{\Sigma} = \widehat{S} / (T - K) \tag{7}$$

the OLS estimate of Σ.

Code **BVAR_ANALYT.m** (found in folder BVAR_Analytical) gives posterior means and variances of parameters & predictives, using the analytical formulas.

Code **BVAR_FULL.m** (found in the folder BVAR_FULL) estimates the BVAR model combining all the priors discussed below, and provides predictions and impulse responses (check Empirical Application 1 in the monograph)

2.1.1 The Diffuse Prior

The diffuse (or Jeffreys') prior for a and Σ takes the form

$$p(\alpha, \Sigma) \propto |\Sigma|^{-(M+1)/2}$$

The conditional posteriors are easily derived, and it is proven that they are of the form

$$\alpha|\Sigma, y \sim N(\widehat{a}, \Sigma), \quad \Sigma|y \sim IW\left(\widehat{S}, T-K\right)$$

2.1.2 The Natural Conjugate Prior

The natural conjugate prior has the form

$$\alpha|\Sigma \sim N(\underline{\alpha}, \Sigma \otimes \underline{V})$$

and

$$\Sigma^{-1} \sim W\left(\underline{v}, \underline{S}^{-1}\right)$$

The posterior for a is

$$\alpha|\Sigma, y \sim N\left(\overline{\alpha}, \Sigma \otimes \overline{V}\right)$$

where $\overline{V} = \left(\underline{V}^{-1} + X'X\right)^{-1}$ and $\overline{\alpha} = vec\left(\overline{A}\right)$ with $\overline{A} = \overline{V}\left(\underline{V}^{-1}\underline{A} + X'X\widehat{A}\right)$.

The posterior for Σ is

$$\Sigma^{-1}|y \sim W\left(\overline{\nu}, \overline{S}^{-1}\right)$$

where $\overline{\nu} = T + \underline{\nu}$ and $\overline{S} = S + \underline{S} + \widehat{A}'X'X\widehat{A} + \underline{A}'\underline{V}^{-1}\underline{A} - \overline{A}'\left(\underline{V}^{-1} + X'X\right)\overline{A}$.

2.1.3 The Minnesota Prior

The Minnesota Prior refers mainly to restricting the hyperparameters of a. The data-based restrictions are the ones presented in the monograph. The prior for a is still normal and the posteriors are the similar to the Natural conjugate prior case. Σ is assumed known in this case (for example equal to $\widehat{\Sigma}$).

2.2 Estimation of VARs using the Gibbs sampler

2.2.1 The Independent Normal-Wishart Prior-Posterior algorithm

We write the VAR as:

$$y_t = Z_t \beta + \varepsilon_t$$

where $Z_t = I_M \otimes X_t$ and ε_t is $N(0, \Sigma)$.

It can be seen that the restricted VAR can be written as a Normal linear regression model with an error covariance matrix of a particular form. A very general prior for this model (which does not involve the restrictions inherent in the natural conjugate prior) is the independent Normal-Wishart prior:

$$p\left(\beta, \Sigma^{-1}\right) = p(\beta)\, p\left(\Sigma^{-1}\right)$$

where

$$\beta \sim N\left(\underline{\beta}, \underline{V}_\beta\right) \tag{8}$$

and

$$\Sigma^{-1} \sim W\left(\underline{\nu}, \underline{S}^{-1}\right). \tag{9}$$

Note that this prior allows for the prior covariance matrix, $\underline{V}_\beta$, to be anything the researcher chooses, rather than the restrictive $\Sigma \otimes \underline{V}$ form of the natural conjugate prior. For instance, the researcher could choose a prior similar in spirit to the Minnesota prior, but allow for different forms of shrinkage in different equations. A noninformative prior can be obtained by setting $\underline{\nu} = \underline{S} = \underline{V}_\beta^{-1} = 0$.

The *conditional posteriors* are:

Posterior on $\beta = vec(B)$

$$\beta | y, \Sigma^{-1} \sim N\left(\overline{\beta}, \overline{V}_\beta\right), \tag{10}$$

where $\overline{\beta} = \overline{V}_\beta \left(\underline{V}_\beta^{-1} \underline{\beta} + \sum_{i=1}^{T} Z_t' \Sigma^{-1} y_t\right)$ and $\overline{V}_\beta = \left(\underline{V}_\beta^{-1} + \sum_{t=1}^{T} Z_t' \Sigma^{-1} Z_t\right)^{-1}$

Posterior on Σ

$$\Sigma^{-1}|y,\beta \sim W\left(\overline{\nu},\overline{S}^{-1}\right) \tag{11}$$

where $\overline{\nu} = T + \underline{\nu}$ and $\overline{S} = \underline{S} + \sum_{t=1}^{T}(y_t - Z_t\beta)(y_t - Z_t\beta)'$.

The one-step ahead predictive density, conditional on the parameters of the model is:

$$y_t|Z_t,\beta,\Sigma \sim N(Z_t\beta,\Sigma)$$

As we note in the monograph, in order to calculate reasonable predictions, Z_t should contain lags of the dependent variables, and exogenous variables which are observed at time $t-h$, where h is the desired forecast horizon. This result, along with a Gibbs sampler producing draws $\beta^{(r)},\Sigma^{(r)}$ for $r = 1,..,R$ allows for predictive inference.[1] For instance, the predictive mean (a popular point forecast) could be obtained as:

$$E(y_\tau|Z_\tau) = \frac{\sum_{r=1}^{R} Z_t\beta^{(r)}}{R}$$

and other predictive moments can be calculated in a similar fashion. Alternatively, predictive simulation can be done at each Gibbs sampler draw, but this can be computationally demanding. For forecast horizons greater than one, the direct method can be used. This strategy for doing predictive analysis can be used with any of the models discussed below.

Code **BVAR_GIBBS.m** (found in the folder BVAR_Gibbs) estimates this model, but also allows the prior mean and covariance of β (i.e. the hyperparameters $\left(\underline{\beta},\underline{V}_\beta\right)$) to be set as in the Minnesota case.

2.2.2 Stochastic Search Variable Selection in VAR models

In the VAR model

$$Y_t = X_tA + \varepsilon_t \tag{12}$$

we can introduce the SSVS prior (George and McCullogh, 1993) which is a hierarchical prior of the form

$$a|\gamma \sim N(0,D) \tag{13}$$

[1]Typically, some initial draws are discarded as the "burn in". Accordingly, $r = 1,..,R$ should be the post-burn in draws.

where $\alpha = vec(A) = (\alpha_1, ..., \alpha_{KM})'$ and D is a diagonal matrix. If we write its $j-th$ diagonal element as $D_{j,j}$, this prior implies that there is dependence on a hyperparameter $\gamma = (\gamma_1, ..., \gamma_{KM})'$ of the following form

$$D_{j,j} = \begin{cases} \kappa_{0j}^2 \text{ if } \gamma_j = 0 \\ \kappa_{1j}^2 \text{ if } \gamma_j = 1 \end{cases} \tag{14}$$

where we a-priori set the hyperparameters $\kappa_{0j}^2 \to 0$ and $\kappa_{1j}^2 \to \infty$. This prior implies that when $\gamma_j = 0$ the prior variance of the $j-th$ element of a, call it a_j, will be equal to κ_{0j}^2, which is very low since $\kappa_{0j}^2 \to 0$. Subsequently, the posterior of the $j-th$ parameter will be restricted in this case to shrink towards the prior mean, which is 0. In the alternative case, $\gamma_j = 1$, the parameter will remain unrestricted and the posterior will be determined mainly by the likelihood. The SSVS prior in (13) can be written in a mixture of Normals form, which is more illuminating about the effect of each γ_j on the prior of a_j:

$$\alpha_j|\gamma_j \sim \left(1-\gamma_j\right) N\left(0, \kappa_{0j}^2\right) + \gamma_j N\left(0, \kappa_{1j}^2\right)$$

The way in which it is determined whether γ_j is 0 or 1 (and hence whether a_j is restricted or not) is not chosen by the researcher, as in the case of the Minnesota prior which favors only own lags and the constant parameters (and restricts the other R.H.S. variables in a semi-data-based way). The value of γ_j should is determined fully in a data-based fashion, and hence a prior is assigned to γ. In a Bayesian context, a prior on a binomial variable which results in easy computations is the Bernoulli density. Note also that it helps calculations if we assume that the elements of γ are independent of each other and sample each γ_j individually. Subsequently, the prior for γ is of the form

$$\gamma_j|\gamma_{\setminus -j} \sim Bernoulli\left(1, \underline{q}_j\right)$$

This prior can also be written in the form: $\Pr\left(\gamma_j = 1\right) = \underline{q}_j$ and $\Pr\left(\gamma_j = 0\right) = 1 - \underline{q}_j$. A typical "noninformative" value of the hyperparameter $\underline{q}_j$ is 0.5, although the reader might want to consult Chipman et al. (2001) and George and McCullogh (1997) on this issue.

Finally for Σ we assume the standard Wishart prior

$$\Sigma^{-1}, \sim W\left(\underline{v}, \underline{S}^{-1}\right)$$

George, Sun and Ni (2008) provide details on how to implement the restriction search (SSVS prior) on elements of Σ. The MATLAB code implements this approach, but it is not discussed here. The reader is referred to the article by George, Sun and Ni.

The *conditional posteriors* are

1. Sample α from the density

$$\alpha|y,\gamma,\Sigma \sim N(\overline{\alpha}_\alpha, \overline{V}_\alpha),$$

where $\overline{V}_\alpha = [\Sigma^{-1} \otimes (X'X) + (DD)^{-1}]^{-1}$ and $\overline{\alpha}_\alpha = \overline{V}_\alpha[(\Psi\Psi') \otimes (X'X)\hat{\alpha}]$ where $\hat{\alpha}$ is the OLS estimate of α.

2. Sample γ_j from the density

$$\gamma_j|\gamma_{\setminus -j}, b, y, Z \sim Bernoulli\left(1, \overline{q}_j\right) \tag{15}$$

where

$$\overline{q}_j = \frac{\frac{1}{\kappa_{1j}}\exp\left(-\frac{\alpha_j^2}{2\kappa_{1j}^2}\right)\underline{q}_j}{\frac{1}{\kappa_{1j}}\exp\left(-\frac{\alpha_j^2}{2\kappa_{1j}^2}\right)\underline{q}_j + \frac{1}{\kappa_{0j}}\exp\left(-\frac{\alpha_j^2}{2\kappa_{0j}^2}\right)\left(1-\underline{q}_j\right)}$$

3. Sample Σ^{-1} from the density

$$\Sigma^{-1} \sim Wishart\left(\overline{v}, \overline{S}\right)$$

where $\overline{v} = T + \underline{v}$ and $\overline{S} = \left(\underline{S}^{-1} + \sum_{t=1}^{T}(Y_t - Z_t\theta)'(Y_t - Z_t\theta)\right)^{-1}$.

Code **SSVS_VAR.m** and **SSVS_VAR_CONST.m** (found in the folder SSVS_VAR) estimate this model. The first model assumes that all parameters are subject to restriction search. The second code allows the intercepts to be unrestricted, as in the example of George, Sun and Ni (2008).

2.2.3 Flexible Variable Selection in VAR models

Another way to incorporate variable selection in the VAR model is to explicitly restrict the parameter to be zero, when the indicator variable is zero. As we explain in the monograph, the VAR model

$$y_t = Z_t \beta + \varepsilon_t$$

can be written now as

$$y_t = Z_t \theta + \varepsilon_t$$

where $\theta = \Gamma\beta$ and $\Gamma = diag\left(\gamma\right) = diag\left(\gamma_1, ..., \gamma_{KM}\right)$. If we denote by γ_j the j-th element of the vector γ (which is also the j-th diagonal element of the matrix Γ), and by $\gamma_{/-j}$ the vector γ where the j-th element is removed, a Gibbs sampler for this model takes the following form:

Priors:

$$\begin{aligned} \beta &\sim N_{MK}\left(\underline{\beta}, \underline{V}\right) && (16) \\ \gamma_j|\gamma_{\backslash -j} &\sim Bernoulli\left(1, \underline{\pi}\right) && (17) \\ \Sigma^{-1} &\sim Wishart\left(\underline{v}, \underline{S}\right) && (18) \end{aligned}$$

Conditional posteriors:

1. Sample α from the density

$$\beta|\gamma, H, y, Z \sim N_{MK}\left(\overline{\beta}, \overline{V}\right) \tag{19}$$

where $\overline{V} = \left(\underline{V}^{-1} + \sum_{t=1}^{T} Z_t^{*\prime} \Sigma^{-1} Z_t^{*}\right)^{-1}$ and $\overline{\beta} = \overline{V}\left(\underline{V}^{-1}\underline{\beta} + \sum_{t=1}^{T} Z_t^{*\prime} \Sigma^{-1} Y_t\right)$, and $Z_t^{*} = Z_t \Gamma$.

2. Sample γ_j from the density

$$\gamma_j|\gamma_{\backslash -j}, b, y, Z \sim Bernoulli\left(1, \overline{\pi}_j\right) \tag{20}$$

preferably in random order j, where $\overline{\pi}_j = \frac{l_{0j}}{l_{0j} + l_{1j}}$, and

$$\begin{aligned} l_{0j} &= \exp\left(-\frac{1}{2}\, tr\left(\sum_{t=1}^{T}\left(Y_t - Z_t\theta^{*}\right)' \Sigma^{-1}\left(Y_t - Z_t\theta^{*}\right)\right)\right)\pi_{0j} \\ l_{1j} &= \exp\left(-\frac{1}{2}\, tr\left(\sum_{t=1}^{T}\left(Y_t - Z_t\theta^{**}\right)' \Sigma^{-1}\left(Y_t - Z_t\theta^{**}\right)\right)\right)\left(1 - \pi_{0j}\right). \end{aligned}$$

Here we define θ^* to be equal to θ but with the $j-th$ element $\theta_j = \beta_j$ (i.e. when $\gamma_j = 1$). Similarly, we define θ^{**} to be equal to θ but with the $j-th$ element $\theta_j = 0$ (i.e. when $\gamma_j = 0$).

3. Sample Σ^{-1} from the density

$$\Sigma^{-1} \sim Wishart\left(\overline{v}, \overline{S}\right)$$

where $\overline{v} = T + \underline{v}$and $\overline{S} = \left(\underline{S}^{-1} + \sum_{t=1}^{T} \left(Y_t - Z_t\theta\right)' \left(Y_t - Z_t\theta\right)\right)^{-1}$.

Code **VAR_SELECTION.m** (found in the folder VAR_Selection) estimates this model.

3 Time-Varying parameters VAR models

3.1 Homoskedastic TVP-VAR

The basic TVP-VAR can be written as:

$$y_t = Z_t \beta_t + \varepsilon_t, \tag{21}$$

and

$$\beta_{t+1} = \beta_t + u_t, \tag{22}$$

where ε_t is i.i.d. $N(0, \Sigma)$ and u_t is i.i.d. $N(0, Q)$. ε_t and u_s are independent of one another for all s and t.

In this model, using priors of the form

$$\begin{aligned}
\beta_0 &\sim N\left(\underline{\beta}, \underline{V}\right) \\
\Sigma^{-1} &\sim W\left(\underline{\nu}, \underline{S}^{-1}\right) \\
Q &\sim W\left(\underline{\nu}_Q, \underline{S}_Q^{-1}\right)
\end{aligned}$$

we sample β_t (conditional on the values of Σ and Q) using the Kalman filter and a smoother (see the monograph for more information), and Σ from the usual Wishart density as

$$\Sigma^{-1} | \beta_t, Q, y \sim W\left(\overline{\nu}, \overline{S}^{-1}\right) \tag{23}$$

where $\overline{\nu} = T + \underline{\nu}$ and $\overline{S} = \underline{S} + \sum_{t=1}^{T} \left(y_t - Z_t \beta_t\right) \left(y_t - Z_t \beta_t\right)'$.

Finally we sample Q from the Wishart density

$$Q^{-1} | \beta_t, \Sigma, y \sim W\left(\overline{\nu}_Q, \overline{S}_Q^{-1}\right) \tag{24}$$

where $\overline{\nu}_Q = T + \underline{\nu}_Q$ and $\overline{S}_Q = \underline{S}_Q + \sum_{t=1}^{T} \left(\beta_t - \beta_{t-1}\right) \left(\beta_t - \beta_{t-1}\right)'$.

There are two different versions of this model. The one is **Homo_TVP_VAR.m** (found in the folder TVP_VAR_CK) which estimates this model plus impulse responses, using the Carter and Kohn (1994) algorithm. Thes second code is **Homo_TVP_VAR_DK.m** (found in the folder TVP_VAR_DK) which estimates this model plus impulse responses, using the Durbin and Koopman (2002) algorithm.

3.1.1 Variable Selection in the Homoskedastic TVP-VAR

Variable selection is defined by rewritting the TVP-VAR model as

$$\begin{aligned} y_t &= Z_t \theta_t + \varepsilon_t, \\ \beta_{t+1} &= \beta_t + u_t, \end{aligned}$$

where now $\theta_t = \Gamma \beta_t$ and Γ is a diagonal matrix (see also the variable selection in the simple VAR case, when β_t is constant).

The time-varying parameters β_t and the covariance Σ are generated exactly as in the homoskedastic TVP-VAR, described in the previous subsection, but conditional on the RHS variables being $Z_t^* = Z_t \Gamma$. The extra step added to the standard Gibbs sampler for TVP-VAR models, is sampling of the indicators γ_j. These are generated - preferably in random order j - as in (20). The only modification required in this case is that in equations (**??**) and (**??**) the densities $p\left(y|\theta_j, \gamma_{\backslash -j}, \gamma_j = 1\right)$ and $p\left(y|\theta_j, \gamma_{\backslash -j}, \gamma_j = 0\right)$ are derived from the full likelihood of the TVP-VAR model, and hence l_{0j} and l_{1j} are written as

$$\begin{aligned} l_{0j} &= \pi_{0j} \exp\left(-\frac{1}{2} \sum_{t=1}^{T} \left(Y_t - Z_t \theta_t^*\right)' \Sigma_t^{-1} \left(Y_t - Z_t \theta_t^*\right) \right) \\ l_{1j} &= \left(1 - \pi_{0j}\right) \exp\left(-\frac{1}{2} \sum_{t=1}^{T} \left(Y_t - Z_t \theta_t^{**}\right)' \Sigma_t^{-1} \left(Y_t - Z_t \theta_t^{**}\right) \right) \end{aligned}$$

Code **TVP_VAR_SELECTION.m** (found in the folder VAR_Selection) estimates this model.

3.2 Hierarchical TVP-VAR

The hierarchical TVP-VAR, based on the model of Chib and Greenberg (1995) is

$$\begin{aligned} y_t &= Z_t \beta_t + \varepsilon_t \\ \beta_{t+1} &= A_0 \theta_{t+1} + u_t, \\ \theta_{t+1} &= \theta_t + \eta_t. \end{aligned} \tag{25}$$

where

$$\begin{bmatrix} \varepsilon_t \\ u_t \\ \eta_t \end{bmatrix} \overset{iid}{\sim} N\left(0, \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & Q & 0 \\ 0 & 0 & R \end{bmatrix}\right).$$

The *priors* for this model are

- Prior on A_0

$$A_0 \sim N\left(\underline{A}, \underline{V}_A\right)$$

- Prior on θ_t

$$\theta_0 \sim N\left(\underline{\theta}_0, \underline{V}_{\theta_0}\right)$$

- Prior on Σ

$$\Sigma^{-1} \sim W\left(\underline{v}_\Sigma, \underline{S}_\Sigma^{-1}\right)$$

- Prior on Q

$$Q^{-1} \sim W\left(\underline{v}_Q, \underline{S}_Q^{-1}\right)$$

- Prior on R

$$R^{-1} \sim W\left(\underline{v}_R, \underline{S}_R^{-1}\right)$$

By defining priors on these parameters, we also implicitly specify a prior for β_t of the form

$$\beta_t | A_0, \theta_0, Q \sim N\left(A_0\theta_t, Q\right), \text{ for } t = 0, .., T.$$

The *conditional posteriors* are

1. Sample β_t from

$$\beta_t | \Sigma, A_0, \theta_t, Q, y_t \sim N\left(\overline{\beta}_t, \overline{V}\right)$$

where $\overline{\beta}_t = \overline{V}\left(Q^{-1}\left(A_0\theta_t\right) + Z_t\Sigma^{-1}y_t\right)$ and $\overline{V} = \left(Q^{-1} + Z_t'\Sigma^{-1}Z_t\right)^{-1}$.

2. Sample A_0 from

$$A_0 | \beta_t, \theta_t, Q \sim N\left(\overline{A}, \overline{V}_A\right)$$

where $\overline{A} = \overline{V}_A\left(\underline{V}_A\underline{A} + \theta' Q^{-1}\beta\right)$ and $\overline{V}_A = \left(\underline{V}_A + \theta' Q^{-1}\theta\right)^{-1}$.

3. Sample Σ from
$$\Sigma^{-1}|A_0,\beta_t,\theta_t,Q,y_t \sim W\left(\overline{v}_\Sigma,\overline{S}_\Sigma^{-1}\right)$$
where $\overline{v}_\Sigma = T+\underline{v}_\Sigma$ and $\overline{S}_\Sigma = \left(\underline{S}_\Sigma + \sum_{t=1}^{T}\left(y_t - Z_t\beta_t\right)'\left(y_t - Z_t\beta_t\right)\right)$.
4. Sample Q from
$$Q^{-1}|A_0,\beta_t,\theta_t \sim W\left(\overline{v}_Q,\overline{S}_Q^{-1}\right)$$
where $\overline{v}_Q = T+\underline{v}_Q$ and $\overline{S}_Q = \left(\underline{S}_Q + \sum_{t=1}^{T}\left(\beta_t - A_0\theta_t\right)'\left(\beta_t - A_0\theta_t\right)\right)$.
5. Sample R from
$$R^{-1}|\theta_t \sim W\left(\overline{v}_R,\overline{S}_R^{-1}\right)$$
where $\overline{v}_R = T+\underline{v}_R$ and $\overline{S}_R = \left(\underline{S}_R + \sum_{t=1}^{T}\left(\theta_t - \theta_{t-1}\right)'\left(\theta_t - \theta_{t-1}\right)\right)$.
6. Sample θ_t using Carter and Kohn (1994)

Code **HierarchicalTVP_VAR.m** (found in the folder HierarchicalTVP_VAR) estimates the parameters of this model.

3.3 Heteroskedastic TVP-VAR

The Heteroskedastic TVP-VAR takes the form

$$y_t = Z_t\beta_t + \varepsilon_t$$

where $\varepsilon_t \sim N\left(0,\Sigma_t\right)$, and $\Sigma_t = L^{-1}D_tD_tL^{-1\prime}$ where D_t is a diagonal matrix with diagonal elements $d_{it} = \exp\left(\frac{1}{2}h_{it}\right)$ being the error time-varying standard deviations, and L is a lower triangular matrix of time-varying covariances, with ones on the diagonal. For instance, in the $M=3$ case we have

$$L = \begin{bmatrix} 1 & 0 & 0 \\ L_{21} & 1 & 0 \\ L_{31} & L_{32} & 1 \end{bmatrix}$$

If we first stack the unrestricted elements of L by rows into a $\frac{M(M-1)}{2}$ vector as $l_t = \left(L_{21,t},L_{31,t},L_{32,t},..,L_{p(p-1),t}\right)'$ and $h_t = \left(h_{1t},...,h_{Mt}\right)'$, then β_t, l_t and h_t follow independent random walks

$$
\begin{aligned}
\beta_{t+1} &= \beta_t + u_t \\
l_{t+1} &= l_t + \zeta_t \\
h_{t+1} &= h_t + \eta_t
\end{aligned}
$$

The errors in the three state equations are

$$
\begin{bmatrix} u_t \\ \zeta_t \\ \eta_t \end{bmatrix} \overset{iid}{\sim} N\left(0, \begin{bmatrix} Q & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & W \end{bmatrix}\right)
$$

The state-space methods used to estimate β_t can also be used to estimate l_t and h_t. We remind that Z_t is of dimensions $M \times KM$, so that the number of elements of each column vector β_t is $n_\beta = KM$ (for each t). Similarly, the number of elements in each column l_t is $n_l = \frac{M(M-1)}{2}$, and the number of elements in each column vector h_t is $n_h = M$. The priors (initial condition at time $t = 0$) on the time-varying parameters are:

$$
\begin{aligned}
\beta_0 &\sim N\left(0, 4I_{n_\beta}\right) \\
l_0 &\sim N\left(0, 4I_{n_l}\right) \\
h_0 &\sim N\left(0, 4I_{n_h}\right)
\end{aligned} \tag{26}
$$

and the priors on their error covariances, are

$$
\begin{aligned}
Q^{-1} &\sim W\left(1 + n_\beta, \left((k_Q)^2 \cdot (1 + n_\beta) \cdot I_{n_\beta}\right)^{-1}\right) \\
S^{-1} &\sim W\left(1 + n_l, \left((k_S)^2 \cdot (1 + n_l) \cdot I_{n_l}\right)^{-1}\right) \\
W^{-1} &\sim W\left(1 + n_h, \left((k_W)^2 \cdot (1 + n_h) \cdot I_{n_h}\right)^{-1}\right)
\end{aligned} \tag{27}
$$

where the hyperparameters are set to $k_Q = 0.01$, $k_S = 0.1$ and $k_W = 0.01$, and I_m is the identity matrix of dimensions $m \times m$. The user can also specify a prior based on the OLS estimates of a constant parameters VAR on a training sample (see Primiceri (2005) and the code for this approach).

One can inform the priors using a training sample. In particular, assume that

$\widehat{\theta}_{OLS}$ and $V\left(\widehat{\theta}_{OLS}\right)$ are the mean and variance respectively of the OLS estimate (or a Bayesian estimate using noninformative priors) of $\theta = \{\beta, l, h\}$ based on a VAR with constant parameters using an initial, training sample. Then the priors can be rewritten as

$$\begin{aligned}
\beta_0 &\sim N\left(\beta_{OLS}, 4V\left(\beta_{OLS}\right)\right) \\
l_0 &\sim N\left(l_{OLS}, 4V\left(l_{OLS}\right)\right) \\
h_0 &\sim N\left(h_{OLS}, 4V\left(h_{OLS}\right)\right)
\end{aligned} \tag{28}$$

$$\begin{aligned}
Q^{-1} &\sim W\left(1+n_\beta, \left((k_Q)^2 \cdot (1+n_\beta) \cdot V\left(\beta_{OLS}\right)\right)^{-1}\right) \\
S^{-1} &\sim W\left(1+n_l, \left((k_S)^2 \cdot (1+n_l) \cdot V\left(l_{OLS}\right)\right)^{-1}\right) \\
W^{-1} &\sim W\left(1+n_h, \left((k_W)^2 \cdot (1+n_h) \cdot V\left(h\dot{:}_{OLS}\right)\right)^{-1}\right)
\end{aligned} \tag{29}$$

The posterior of β_t is easily obtained, as in the case of the Homoskedastic VAR. The only difference is that now, we draw β_t conditional on the VAR covariance matrix being Σ_t. Draws of l_t and h_t will provide us with draws of L_t and D_t respectively, and then we can recover Σ_t using $\Sigma_t = L_t^{-1} D_t D_t L_t^{-1\prime}$. For detailed info see the monograph, and the appendix Primiceri (2005).

Code **Hetero_TVP_VAR.m** (found in the folder TVP_VAR_CK) estimates the parameters and impulse responses from this model.

4 Factor models

4.1 Static factor model

The static factor model is (ignoring the intercept)

$$y_t = \lambda f_t + \varepsilon_t$$

where y_t is an $M \times 1$ vector of observed time series variables, f_t is a $q \times 1$ vector of unobserved factors with $f_t \sim N(0, I_q)$, λ is an $M \times q$ matrix of coefficients (factor loadings), and $\varepsilon_t \sim N(0, \Sigma)$ with $\Sigma = diag(\sigma_1^2, ..., \sigma_M^2)$. A popular way to identify this model (see Lopes and West, 2004, and Geweke and Zhou, 1996) is to impose λ to be block lower triangular with diagonal elements strictly positive, i.e. $\lambda_{jj} > 0$ and $\lambda_{jk} = 0$ for $j > k$, $j = 1, .., q$. Since the covariance matrix Σ is diagonal, we can treat this model as M indepedent regressions (conditional on knowing f_t). Subsequently we set proper priors of the Normal - inverse-Gamma form, and λ is sampled with the restriction that its diagonal elements come from a truncated normal density, while the upper diagonal elements are zero.

Code **BFM.m** (found in folder Factor_Models) replicates this model, following Lopes and West (2004).

4.2 Dynamic factor model (DFM)

The dynamic factor model assumes that the factors follow a VAR. A simple form of this model is

$$\begin{aligned} y_{it} &= \lambda_{0i} + \lambda_i f_t + \varepsilon_{it} \\ f_t &= \Phi_1 f_{t-1} + .. + \Phi_p f_{t-p} + \varepsilon_t^f \end{aligned}$$

This model needs only a small modification in order to write it as a linear state-space model, with f_t the state variable. This modification is to write the p-lag state equation as a first order Markov system (i.e. transform the VAR(p) equation $f_t = \Phi_1 f_{t-1} + .. + \Phi_p f_{t-p} + \varepsilon_t^f$, into a VAR(1) model; we have seen how to do this in the simple VAR models when we wanted to compute the impulse responses). Conditional on this transformation, a Gibbs sampler is used to draw f_t using the Carter and Kohn (1994) algorithm, while conditional on the draw of f_t the parameters Φ can be estimated using any of the VAR priors of section 2 (see also the monograph). In the first (the measurement) equation, conditional on f_t, we sample each λ_i using the arguments for simple regression models, i.e. a prior of the Normal-Gamma form (see Koop, 2003).

Code **BAYES_DFM.m** (found in folder Factor_Models) estimates the above model.

4.3 Factor-augmented VAR (FAVAR)

The factor augmented VAR builds on the dynamic factor model structure and allows to identify monetary policy shocks. We use the simple formulation of Bernanke, Boivin and Eliasz (2005) which is

$$y_{it} = \lambda_i f_t + \gamma_i r_t + \varepsilon_{it}, \tag{30}$$

$$\begin{pmatrix} f_t \\ r_t \end{pmatrix} = \widetilde{\Phi}_1 \begin{pmatrix} f_{t-1} \\ r_{t-1} \end{pmatrix} + .. + \widetilde{\Phi}_p \begin{pmatrix} f_{t-p} \\ r_{t-p} \end{pmatrix} + \widetilde{\varepsilon}_t^f \tag{31}$$

where $\widetilde{\varepsilon}_t^f$ is i.i.d. $N\left(0, \widetilde{\Sigma}^f\right)$ and r_t is a $k_r \times 1$ vector of observed variables. For instance, Bernanke, Boivin and Eliasz (2005) set r_t to be the Fed Funds rate (a monetary policy instrument) and, thus, $k_r = 1$. All other assumptions about the measurement equation are the same as for the DFM.

Note that conditional on the parameters, the factors can be sampled using state-space methods (see the previous section) where f_t is the unobserved state variable. This is easily implemented if we convert equation (31) from a VAR(p) model into a VAR(1) model (so that f_t is Markov, which is a necessary assumption in order to use the Kalman filter). Subsequently, the parameter matrices $\widetilde{\Phi}$ and $\widetilde{\Sigma}^f$ have to be augmented with zeros in order to conform with the VAR(1) transformation, but we sample the non-zero elements the usual way.

A different option is to use Principal Components to approximate the factors f_t. Bayesian estimation provides us with dynamic factors, with covariance $\widetilde{\Sigma}^f$. Principal Components provide us only with static factors with normalized covariance I (since Principal Components provide a solution only to the factor equation (30) without taking into account the dynamics of the factors in equation (31)). PC estimates is a computationally tractable method regardless of the dimension of the data or the factors we want to extract but are subject to sampling error. MCMC estimation can be cumbersome in very large problems, but having the full posterior of the factors eliminates any sampling error uncertainty. MCMC estimation (and in general likelihood-based estimation) of dynamic factors using the Kalman filter requires strong identification restrictions which may lead to factors with poor economic content. Subsequently the practitioner of this model should be very carefull when choosing a specific method to sample latent factors. Our empirical application proceeds with Principal Components, due to their computational simplicity. No matter the chosen method, the parameters are sampled conditional on the current draw (for MCMC) or final estimate (if using PC) of the factors, i.e. just as if the factors were observed data.

In (30) we have M independent equations, so we can sample the parameters λ_i and σ_i equation-by-equation. Subsequently, we have M univariate regression models and a standard conjugate prior that can be used for the parameters is the Normal-Gamma (see Koop, 2003). Equation (31) is a VAR model on $\begin{pmatrix} f_t \\ r_t \end{pmatrix}$ and the reader is free to use any of the priors discussed in the respective Section about VAR models. For the purpose of the empirical illustration, we use the Noninformative prior. To summarize, we use priors of the form:

$$\begin{aligned} \lambda_i &\sim N(0, cI) \\ \left(\sigma_i^2\right)^{-1} &\sim G(a, \beta) \\ \widetilde{\Phi}, \widetilde{\Sigma}^f &\propto \left|\widetilde{\Sigma}^f\right|^{-(M+k_r+1)/2} \end{aligned}$$

where, in the absense of prior information, c, a and β can be used in a data-based fashion, or set to uninformative values, like 100, 0.01 and 0.01, respectively.

Code **FAVAR.m** (found in folder FAVAR) estimates this model using Principal Components and gives impulse responses for 115 + 3 variables in total. There is the option to use MCMC estimation of the factors, but this is not done automatically. There are directions to the user in order to comment some parts of the code, and uncomment others, in order to do that.

4.4 Time-varying parameters Factor-augmented VAR (FAVAR)

Extending the FAVAR model into a model with time-varying parameters is as "easy" as extending the VAR model into the TVP-VAR model we examined in the previous section. Given that one can still use a principal components approximation of the factors in the TVP-FAVAR model, questions of interest are which parameters should be allowed to vary over time[2]? For the purpose of our empirical illustration, we extend the FAVAR model of the previous subsection, equations (30) - (31) by allowing $\widetilde{\Phi} = \left[\widetilde{\Phi}_1, ..., \widetilde{\Phi}_p\right]$ and $\widetilde{\Sigma}^f$ to be time-varying in a form which is exactly the same as the Heteroskedastic VAR model already discussed (i.e. random walk evolution on each parameter). Exact details are given in the monograph, and section 3.3. We only have to note here that in contrast with the empirical illustration of the Homoskedastic and Heteroskedastic TVP-VARs, we do not use a training sample for the TVP-FAVAR application (though the reader can define a training sample in the same fashion as in section 3.3). Subsequently, the priors on the parameters $\widetilde{\Phi}_t$ and $\widetilde{\Sigma}^f_t$ are given by equations (26) and (27). The priors for the constant parameters λ_i and σ_i are the ones used in the FAVAR model above.

One can allow the loadings matrix to be time varying, as well as the log-variances of the errors in the equation (30). However the loadings matrix λ contains many parameters, so the reader should be carefull to avoid overparametrization when relaxing the assumption of constant loadings.

Code **TVP_FAVAR_FULL.m** (found in folder TVP-FAVAR) estimates the TVP-FAVAR model and gives impulse responses for 115 + 3 variables in total.

[2]This issue is addressed in Korobilis (2009a) and the reader is referred to this paper for more information.

4.5 Data used for factor model applications

All series were downloaded from St. Louis' FRED database and cover the quarters Q1:1959 to Q3:2006. The series HHSNTN, PMNO, PMDEL, PMNV, MOCMQ, MSONDQ (series numbered 152 - 157 in the following table) were kindly provided by Mark Watson and come from the Global Insights Basic Economics Database. All series were seasonally adjusted: either taken adjusted from FRED or by applying to the non-seasonally adjusted series a quarterly X11 filter based on an AR(4) model (after testing for seasonality). Some series in the database were observed only on a monthly basis and quarterly values were computed by averaging the monthly values over the quarter. Following [**?**], the fast moving variables are interest rates, stock returns, exchange rates and commodity prices. The rest of the variables in the dataset are the slow moving variables (output, employment/unemployment etc). All variables are transformed to be approximate stationary. In particular, if $z_{i,t}$ is the original untransformed series, the transformation codes are (column Tcode below): 1 - no transformation (levels), $x_{i,t} = z_{i,t}$; 2 - first difference, $x_{i,t} = z_{i,t} - z_{i,t-1}$; 4 - logarithm, $x_{i,t} = \log z_{i,t}$; 5 - first difference of logarithm, $x_{i,t} = \log z_{i,t} - \log z_{i,t-1}$.

#	Mnemonic	Tcode	Description
1	CBI	1	Change in Private Inventories
2	GDPC96	5	Real Gross Domestic Product, 3 Decimal
3	FINSLC96	5	Real Final Sales of Domestic Product, 3 Decimal
4	CIVA	1	Corporate Inventory Valuation Adjustment
5	CP	5	Corporate Profits After Tax
6	CNCF	5	Corporate Net Cash Flow
7	GDPCTPI	5	Gross Domestic Product: Chain-type Price Index
8	FPI	5	Fixed Private Investment
9	GSAVE	5	Gross Saving

10	PRFI	5	Private Residential Fixed Investment
11	CMDEBT	5	Household Sector: Liabilites: Household Credit Market Debt Outstanding
12	INDPRO	5	Industrial Production Index
13	NAPM	1	ISM Manufacturing: PMI Composite Index
14	HCOMPBS	5	Business Sector: Compensation Per Hour
15	HOABS	5	Business Sector: Hours of All Persons
16	RCPHBS	5	Business Sector: Real Compensation Per Hour
17	ULCBS	5	Business Sector: Unit Labor Cost
18	COMPNFB	5	Nonfarm Business Sector: Compensation Per Hour
19	HOANBS	5	Nonfarm Business Sector: Hours of All Persons
20	COMPRNFB	5	Nonfarm Business Sector: Real Compensation Per Hour
21	ULCNFB	5	Nonfarm Business Sector: Unit Labor Cost
22	UEMPLT5	5	Civilians Unemployed - Less Than 5 Weeks
23	UEMP5TO14	5	Civilian Unemployed for 5-14 Weeks
24	UEMP15OV	5	Civilians Unemployed - 15 Weeks & Over
25	UEMP15T26	5	Civilians Unemployed for 15-26 Weeks
26	UEMP27OV	5	Civilians Unemployed for 27 Weeks & Over
27	NDMANEMP	5	All Employees: Nondurable Goods Manufacturing
28	MANEMP	5	Employees on Nonfarm Payrolls: Manufacturing
29	SRVPRD	5	All Employees: Service-Providing Industries
30	USTPU	5	All Employees: Trade, Transportation & Utilities
31	USWTRADE	5	All Employees: Wholesale Trade
32	USTRADE	5	All Employees: Retail Trade
33	USFIRE	5	All Employees: Financial Activities
34	USEHS	5	All Employees: Education & Health Services
35	USPBS	5	All Employees: Professional & Business Services

36	USINFO	5	All Employees: Information Services
37	USSERV	5	All Employees: Other Services
38	USPRIV	5	All Employees: Total Private Industries
39	USGOVT	5	All Employees: Government
40	USLAH	5	All Employees: Leisure & Hospitality
41	AHECONS	5	Average Hourly Earnings: Construction
42	AHEMAN	5	Average Hourly Earnings: Manufacturing
43	AHETPI	5	Average Hourly Earnings: Total Private Industries
44	AWOTMAN	1	Average Weekly Hours: Overtime: Manufacturing
45	AWHMAN	1	Average Weekly Hours: Manufacturing
46	HOUST	4	Housing Starts: Total: New Privately Owned Housing Units Started
47	HOUSTNE	4	Housing Starts in Northeast Census Region
48	HOUSTMW	4	Housing Starts in Midwest Census Region
49	HOUSTS	4	Housing Starts in South Census Region
50	HOUSTW	4	Housing Starts in West Census Region
51	HOUST1F	4	Privately Owned Housing Starts: 1-Unit Structures
52	PERMIT	4	New Private Housing Units Authorized by Building Permit
53	NONREVSL	5	Total Nonrevolving Credit Outstanding, SA, Billions of Dollars
54	USGSEC	5	U.S. Government Securities at All Commercial Banks
55	OTHSEC	5	Other Securities at All Commercial Banks
56	TOTALSL	5	Total Consumer Credit Outstanding
57	BUSLOANS	5	Commercial and Industrial Loans at All Commercial Banks
58	CONSUMER	5	Consumer (Individual) Loans at All Commercial Banks
59	LOANS	5	Total Loans and Leases at Commercial Banks
60	LOANINV	5	Total Loans and Investments at All Commercial Banks

61	INVEST	5	Total Investments at All Commercial Banks
62	REALLN	5	Real Estate Loans at All Commercial Banks
63	BOGAMBSL	5	Board of Governors Monetary Base, Adjusted for Changes in Reserve Requirements
64	TRARR	5	Board of Governors Total Reserves, Adjusted for Changes in Reserve Requirements
65	BOGNONBR	5	Non-Borrowed Reserves of Depository Institutions
66	NFORBRES	1	Net Free or Borrowed Reserves of Depository Institutions
67	M1SL	5	M1 Money Stock
68	CURRSL	5	Currency Component of M1
69	CURRDD	5	Currency Component of M1 Plus Demand Deposits
70	DEMDEPSL	5	Demand Deposits at Commercial Banks
71	TCDSL	5	Total Checkable Deposits
72	TB3MS	1	3-Month Treasury Bill: Secondary Market Rate
73	TB6MS	1	6-Month Treasury Bill: Secondary Market Rate
74	GS1	1	1-Year Treasury Constant Maturity Rate
75	GS3	1	3-Year Treasury Constant Maturity Rate
76	GS5	1	5-Year Treasury Constant Maturity Rate
77	GS10	1	10-Year Treasury Constant Maturity Rate
78	MPRIME	1	Bank Prime Loan Rate
79	AAA	1	Moody’s Seasoned Aaa Corporate Bond Yield
80	BAA	1	Moody’s Seasoned Baa Corporate Bond Yield
81	sTB3MS	1	TB3MS - FEDFUNDS
82	sTB6MS	1	TB6MS - FEDFUNDS
83	sGS1	1	GS1 - FEDFUNDS
84	sGS3	1	GS3 - FEDFUNDS
85	sGS5	1	GS5 - FEDFUNDS

86	sGS10	1	GS10 - FEDFUNDS
87	sMPRIME	1	MPRIME - FEDFUNDS
88	sAAA	1	AAA - FEDFUNDS
89	sBAA	1	BBB - FEDFUNDS
90	EXSZUS	5	Switzerland / U.S. Foreign Exchange Rate
91	EXJPUS	5	Japan / U.S. Foreign Exchange Rate
92	PPIACO	5	Producer Price Index: All Commodities
93	PPICRM	5	Producer Price Index: Crude Materials for Further Processing
94	PPIFCF	5	Producer Price Index: Finished Consumer Foods
95	PPIFCG	5	Producer Price Index: Finished Consumer Goods
96	PFCGEF	5	Producer Price Index: Finished Consumer Goods Excluding Foods
97	PPIFGS	5	Producer Price Index: Finished Goods
98	PPICPE	5	Producer Price Index Finished Goods: Capital Equipment
99	PPIENG	5	Producer Price Index: Fuels & Related Products & Power
100	PPIIDC	5	Producer Price Index: Industrial Commodities
101	PPIITM	5	Producer Price Index: Intermediate Materials: Supplies & Components
102	CPIAUCSL	5	Consumer Price Index For All Urban Consumers: All Items
103	CPIUFDSL	5	Consumer Price Index for All Urban Consumers: Food
104	CPIENGSL	5	Consumer Price Index for All Urban Consumers: Energy
105	CPILEGSL	5	Consumer Price Index for All Urban Consumers: All Items Less Energy
106	CPIULFSL	5	Consumer Price Index for All Urban Consumers: All Items Less Food
107	CPILFESL	5	Consumer Price Index for All Urban Consumers: All Items Less Food & Energy

108	OILPRICE	5	Spot Oil Price: West Texas Intermediate
109	HHSNTN	1	Uni. of Mich. Index of Consumer Expectations (BCD-83)
110	PMI	1	Purchasing Managers' Index
111	PMNO	1	NAPM New Orders Index
112	PMDEL	1	NAPM Vendor Deliveries Index
113	PMNV	1	NAPM Inventories Index
114	MOCMQ	5	New Orders (NET) - Consumer Goods & Materials, 1996 Dollars (BCI)
115	MSONDQ	5	New Orders - Non-defence Capital Goods, 1996 Dollars (BCI)