

Mumtaz, Haroon

Working Paper

A generalised stochastic volatility in mean VAR: An updated algorithm

Working Paper, No. 908

Provided in Cooperation with:

School of Economics and Finance, Queen Mary University of London

Suggested Citation: Mumtaz, Haroon (2020) : A generalised stochastic volatility in mean VAR: An updated algorithm, Working Paper, No. 908, Queen Mary University of London, School of Economics and Finance, London

This Version is available at:

<https://hdl.handle.net/10419/247177>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

A Generalised Stochastic Volatility in Mean VAR. An Updated Algorithm

Haroon Mumtaz

Working Paper No. 908

July 2020

ISSN 1473-0278

Ù&@ [|Á ÁÒ&[} [{ æ• Áæ åÁÖä æ &^



A generalised stochastic volatility in mean VAR. An updated algorithm

Haroon Mumtaz*

July 5, 2020

Abstract

In this note we present an updated algorithm to estimate the VAR with stochastic volatility proposed in Mumtaz (2018). The model is re-written so that some of the Metropolis Hastings steps are avoided. The code can be downloaded at [this link](#).

Key words: VAR, Stochastic volatility in mean, error covariance

JEL codes: C3, C11, E3

1 Introduction

In this note we present an updated algorithm to estimate the VAR with stochastic volatility proposed in Mumtaz (2018). The model is re-written so that some of the Metropolis Hastings steps are avoided. This results in two potential improvements: (1) the algorithm requires less tuning and (2) the algorithm is likely to converge faster.

The note is organised as follows: Section 2 describes the VAR model and explains the changes relative to the formulation in Mumtaz (2018). The MCMC algorithm is described in section 3. Simulation evidence is presented in section 3.3 while section 4 considers an empirical application using a VAR with real and financial variables.

*Queen Mary College. Email: h.mumtaz@qmul.ac.uk

2 The VAR with stochastic volatility

Consider the VAR model

$$\tilde{h}_t = \alpha + \theta \tilde{h}_{t-1} + \sum_{j=1}^Q d_j Z_{t-j} + \eta_t \quad (1)$$

$$Z_t = c + \sum_{j=1}^P \beta_j Z_{t-j} + \sum_{k=1}^K b_k \tilde{h}_{t-k} + H_t^{1/2} e_t \quad (2)$$

where Z_t is a matrix of endogenous variables, $\tilde{h}_t = [h_{1t}, h_{2t}, \dots, h_{N,t}]$, $H_t = \text{diag}(\exp(\tilde{h}_t))$. The $M = 2N$ disturbances $\varepsilon_t = \begin{pmatrix} \eta_t \\ e_t \end{pmatrix}$ are distributed normally $N(0, \Sigma)$, $\Sigma = \begin{pmatrix} \Sigma_\eta & \Sigma'_{\eta_e} \\ \Sigma_{\eta_e} & \Sigma_e \end{pmatrix}$ where the diagonal

elements of Σ_e are restricted to equal 1. For example for $M = 4, N = 2$: $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{14} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} \\ \Sigma_{31} & \Sigma_{32} & 1 & \Sigma_{34} \\ \Sigma_{41} & \Sigma_{42} & \Sigma_{43} & 1 \end{pmatrix}$.

Note that this formulation is slightly different from that considered in Mumtaz (2018). In Mumtaz (2018), $\text{diag}(\Sigma_\eta) = 1$ and the variances of η_t are treated as extra parameters. These are drawn by Mumtaz (2018) using a Metropolis algorithm. In this note we consider an alternative approach that avoids this step.

3 MCMC Algorithm

The approach is based on re-parameterising the model in equations 1 and 2. The transition equation can be written as:

$$\tilde{h}_t = \alpha + \theta \tilde{h}_{t-1} + \sum_{j=1}^Q d_j Z_{t-j} + b^{*'} e_t + \eta_t^* \quad (3)$$

$$\text{var}(\eta_t^*) = \Sigma_\eta^*$$

$$b^* = \Sigma_e^{-1} \Sigma_{\eta_e} \quad (4)$$

$$\Sigma_\eta^* = \Sigma_\eta - \Sigma_{\eta_e}' \Sigma_e^{-1} \Sigma_{\eta_e} \quad (5)$$

The observation equation of the system is still given by:

$$Z_t = c + \sum_{j=1}^P \beta_j Z_{t-j} + \sum_{k=1}^K b_k \tilde{h}_{t-k} + H_t^{1/2} e_t$$

$$\text{var}(e_t) = \Sigma_e$$

Note that, in this re-paramerised model:

$$\text{cov}(\eta_t^*, e_t) = 0$$

Moreover, the coefficients b^* and equations 4 and 5 can be used to reconstruct $\text{cov}(\eta_t, e_t) = \Sigma_{\eta_e}$ and the variance Σ_η .

$$\Sigma_{\eta_e} = \Sigma_e b^* \quad (6)$$

$$\Sigma_\eta = \Sigma_\eta^* + \Sigma_{\eta_e}' \Sigma_e^{-1} \Sigma_{\eta_e} \quad (7)$$

Thus equations 6 and 7 can be used to back out Σ_{η_e} and Σ_η instead of attempting to draw these elements directly.

We now describe the steps of the MCMC algorithm that can be used to approximate the posterior distribution in the re-parameterised model. The unknown parameters to be estimated in the transformed

model are: (1) $B^h = \text{vec} \begin{pmatrix} \alpha \\ \theta \\ d_j \\ b^* \end{pmatrix}$, (2) Σ_η^* , (3) $B = \text{vec} \begin{pmatrix} c \\ \beta_j \\ b_k \end{pmatrix}$, (4) Σ_e and (5) \tilde{h}_t .

3.1 Priors

1. B^h, Σ_η^* . Following Banbura *et al.* (2007), we employ a Normal prior for B^h and an inverse Wishart prior for Σ_η^* implemented via dummy observations.

2. B . The prior is set as in 1.

3. Σ_e . Chan and Jeliazkov (2009) decompose Σ as $\Sigma = L^{-1}DL^{-1'}$ with the diagonal elements of D denoted by λ_k and a_{kj} denoting the lower triangular elements of L^{-1} . The prior for a_{kj} is assumed to be $N(0, 1)$ while the prior on D is implicit via the restriction that Σ have diagonal elements that equal 1.
4. \tilde{h}_t . Following Cogley and Sargent (2005) we use a training sample (of 20 pre-sample observations) to set the prior for the elements of the transition equation of the model. Let \hat{v}^{ols} denote the OLS estimate of the VAR covariance matrix estimated on the pre-sample data. The initial value \tilde{h}_0 is set to μ_0 where μ_0 denotes the diagonal elements of the Cholesky decomposition of \hat{v}^{ols} .

3.2 Simulating the posterior distributions

The MCMC algorithm samples from the following conditional posterior distributions:

1. $H\left(B^h | \Sigma_\eta^*, B, \Sigma_e, \tilde{h}_t\right)$. Given \tilde{h}_t and the residual e_t , equation 3 is VAR model. Denote the dummy observations implementing the prior by y_D and x_D and let $y^* = \begin{pmatrix} y_t \\ y_D \end{pmatrix}$, $x^* = \begin{pmatrix} x_t \\ x_D \end{pmatrix}$ where $y_t = \tilde{h}_t$ and x_t denotes all the regressors. The conditional posterior for the coefficients of VAR is normal with mean and variance given by:

$$m = (x^{*'}x^*)^{-1} (x^{*'}y^*)$$

$$v = \Sigma_\eta^* \otimes (x^{*'}x^*)^{-1}$$

2. $H\left(\Sigma_\eta^* | B^h, B, \Sigma_e, \tilde{h}_t\right)$. The conditional posterior is inverse wishart with scale matrix $(y^* - x^* \tilde{B}^h)' (y^* - x^* \tilde{B}^h)$ and degrees of freedom $T + T_d - K$ where \tilde{B}^h denotes the vectorised coefficients B^h reshaped to be conformable with x^* , T_d is the number of dummy observations and k is the number of regressors in each equation.

3. $H\left(B|B^h, \Sigma_\eta^*, \Sigma_e, \tilde{h}_t\right)$. Conditional on \tilde{h}_t , $\begin{pmatrix} \tilde{h}_t \\ Z_t \end{pmatrix}$ has a multi-variate normal distribution

$$\begin{pmatrix} \tilde{h}_t \\ Z_t \end{pmatrix} \sim N \left(\begin{pmatrix} \alpha + \theta \tilde{h}_{t-1} + \sum_{j=1}^Q d_j Z_{t-j} \\ c + \sum_{j=1}^P \beta_j Z_{t-j} + \sum_{k=1}^K b_k \tilde{h}_{t-k} \end{pmatrix}, \underbrace{\begin{pmatrix} \Sigma_{\eta^*} + b^* \Sigma_e b^{*'} & b^* \Sigma_e H_t^{1/2'} \\ H_t^{1/2} \Sigma_e b^{*'} & H_t^{1/2} \Sigma_e H_t^{1/2'} \end{pmatrix}}_{\Sigma} \right)$$

Defining $\bar{h}_t = \tilde{h}_t - \left(\alpha + \theta \tilde{h}_{t-1} + \sum_{j=1}^Q d_j Z_{t-j} \right)$ and $\begin{pmatrix} s_{11,t} & s'_{21,t} \\ s_{21,t} & s_{22,t} \end{pmatrix} = \begin{pmatrix} \Sigma_{\eta^*} + b^* \Sigma_e b^{*'} & b^* \Sigma_e H_t^{1/2'} \\ H_t^{1/2} \Sigma_e b^{*'} & H_t^{1/2} \Sigma_e H_t^{1/2'} \end{pmatrix}$
the conditional distribution $Z_t | \tilde{h}_t$ is also normal with mean and variance:

$$\begin{aligned} m_t &= \bar{h}_t s_{11,t}^{-1} s'_{21,t} \\ v_t &= s_{22,t} - s'_{21,t} s_{11,t}^{-1} s_{21,t} \end{aligned}$$

The VAR model in equation 2 can be written as

$$Z_t^* = c + \sum_{j=1}^P \beta_j Z_{t-j} + \sum_{k=1}^K b_k \tilde{h}_{t-k} + e_t^*$$

where $Z_t^* = Z_t - m_t$ and $\text{var}(e_t^*) = v_t$. This is a VAR with heteroscedastic disturbances. The conditional posterior for the coefficients of this heteroscedastic VAR is normal with mean and variance given by:

$$\bar{m} = \bar{v} \left(\text{vec} \left(\sum_{t=1}^T \left(x_t (Z_t^*)' (v_t)^{-1} \right) \right) + \left(S_0^j \right)^{-1} B_0^{j'} \right)$$

$$\bar{v} = \left(\sum_{t=1}^T \left((v_t)^{-1} \otimes x_t x_t' \right) + \left(S_0^j \right)^{-1} \right)^{-1}$$

where the prior mean and variance is denoted by B_0 and S_0 , respectively and x_t denotes the regressors on the RHS of each equation of the VAR. See Clark (2011) for details on BVARs with stochastic volatility.

4. $H\left(\Sigma_e | B, B^h, \Sigma_\eta^*, \tilde{h}_t\right)$. Chan and Jeliazkov (2009) describe how to sample covariance matrices with

restrictions on some of the elements and we follow their method in implementing the draw from this conditional posterior. Chan and Jeliazkov (2009) decompose Σ_e as $\Sigma_e = L^{-1}DL^{-1'}$. They show that when the diagonal elements of Σ are restricted to equal 1, then the diagonal elements of D (denoted by λ_k) satisfy:

$$\begin{aligned}\lambda_1 &= 1 \\ \lambda_k &= 1 - \sum_{j=1}^{k-1} (a^{kj})^2 \lambda_j, k = 2, 3, \dots, N\end{aligned}\tag{8}$$

where a^{kj} are lower diagonal elements of L^{-1} . They propose an independence Metropolis step to sample a^{kj} with a proposal density of the form:

$$f(a^{kj}|e_t) = N(\mu_k, \tau V_k)$$

where $V = (A_0^{-1} + \sum_{t=1}^T U_t \hat{D}^{-1} U_t)$ and $\mu = V (A_0^{-1} a_0 + \sum_{t=1}^T U_t \hat{D}^{-1} e_t)$. Here U_t is defined as the matrix:

$$U_t = - \begin{pmatrix} 0 & . & . & . & & & & 0 \\ e_{t,1} & 0 & . & . & . & & & \\ 0 & e_{t,1} & e_{t,2} & . & . & . & & \\ 0 & . & . & e_{t,1} & e_{t,2} & e_{t,3} & 0 & . \\ . & . & . & & & & & . \\ . & . & . & & & & & . \\ . & . & . & & & & & 0 \\ 0 & . & . & . & 0 & 0 & e_{t,1} & . & e_{t,N} \end{pmatrix}$$

and the diagonal elements of \hat{D} can be obtained by iterating between the equation for μ and equation

8. The draw is accepted with probability:

$$\alpha = \frac{g(e_t|\Sigma_{new}) f(a_{old}|e_t)}{g(e_t|\Sigma_{old}) f(a_{new}|e_t)}$$

with λ_k restricted to be greater than zero to ensure that Σ_e is positive definite. The expression for the likelihood function used to construct the posterior $g(e_t|\Sigma)$ is given in equation 2.7 in Chan and Jeliazkov (2009).

5. $H\left(\tilde{h}_t|\Sigma_e, B, B^h, \Sigma_\eta^*\right)$ Conditional on the VAR coefficients and the parameters of the transition equation, the model has a multivariate non-linear state-space representation. It is convenient to express the state-space as:

$$\begin{aligned} F_t &= C + \Psi F_{t-1} + N_t \\ Z_t - H_t^{1/2} \mu_{e_t|\eta_t} &= c + \sum_{j=1}^P \beta_j Z_{t-j} + \sum_{k=1}^K b_k \tilde{h}_{t-k} + \tilde{e}_t \\ \text{var}(\tilde{e}_t) &= \Omega_t = H_t^{1/2} \Sigma_{e_t|\eta_t} H_t^{1/2'} \end{aligned} \tag{9}$$

where:

$$F_t = \begin{pmatrix} \eta_{t+1} \\ \eta_t \\ \tilde{h}_t \\ \cdot \\ \tilde{h}_{t-k} \end{pmatrix} \tag{10}$$

$$C = \begin{pmatrix} 0 \\ 0 \\ \alpha + \sum_{j=1}^Q d_j Z_{t-j} \\ \cdot \\ 0 \end{pmatrix} \tag{11}$$

$$\Psi = \begin{pmatrix} 0 & 0 & . & . & 0 \\ 1 & 0 & . & . & . \\ 1 & 0 & \theta & . & 0 \\ 0 & 1 & . & . & 0 \\ . & . & 1 & . & 0 \end{pmatrix} \quad (12)$$

$$N_t = \begin{pmatrix} \eta_{t+1} \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$\mu_{e_t|\eta_t}$ denotes the conditional mean of e_t while $\Sigma_{e_t|\eta_t}$ is the conditional variance. These can be easily calculated using results for multi-variate normal distributions. Partitioning Σ as:

$$\Sigma = \begin{pmatrix} \Sigma_\eta & \Sigma'_{\eta_e} \\ \Sigma_{\eta_e} & \Sigma_e \end{pmatrix}$$

the conditional mean and variance are given by:

$$\begin{aligned} \mu_{e_t|\eta_t} &= \eta_t \Sigma_\eta^{-1} \Sigma'_{\eta_e} \\ \Sigma_{e_t|\eta_t} &= \Sigma_e - \Sigma_{\eta_e} \Sigma_\eta^{-1} \Sigma'_{\eta_e} \end{aligned}$$

Moreover:

$$var(N_t) = \tilde{Q} = \begin{pmatrix} \Sigma_\eta & . & . & . & 0 \\ 0 & . & . & . & . \\ 0 & . & . & . & . \\ 0 & . & . & . & . \\ 0 & . & 0 & 0 & 0 \end{pmatrix}$$

Following recent developments in the seminal paper by Andrieu *et al.* (2010), we employ a particle Gibbs step to sample from the conditional posterior of F_t . Andrieu *et al.* (2010) show how a version of the particle filter, conditioned on a fixed trajectory for one of the particles can be used to produce draws that result in a Markov Kernel with a target distribution that is invariant. However, the usual problem of path degeneracy in the particle filter can result in poor mixing in the original version of particle Gibbs. Recent development, however, suggest that small modifications of this algorithm can largely alleviate this problem. In particular, Lindsten *et al.* (2014) propose the addition of a step that involves sampling the ‘ancestors’ or indices associated with the particle that is being conditioned on. They show that this results in a substantial improvement in the mixing of the algorithm even with a few particles.¹ As explained in Lindsten *et al.* (2014), ancestor sampling breaks the reference path into pieces and this causes the particle system to collapse towards something different than the reference path. In the absence of this step, the particle system tends to collapse to the conditioning path. We employ particle Gibbs with ancestor sampling in this step.

Let $F_t^{(i-1)}$ denote the fixed the fixed trajectory, for $t = 1, 2, \dots, T$ obtained in the previous draw of the Gibbs algorithm. We denote all the parameters of the model by Ξ , and $j = 1, 2, \dots, \tilde{M}$ indexes the particles. The conditional particle filter with ancestor sampling proceeds in the following steps:

1. For $t = 1$

- (a) Draw $F_1^{(j)} \setminus F_0^{(j)}, \Xi$ for $j = 1, 2, \dots, \tilde{M} - 1$. Fix $F_1^{(\tilde{M})} = F_1^{(i-1)}$
- (b) Compute the normalised weights $p_1^{(j)} = \frac{w_1^{(j)}}{\sum_{j=1}^{\tilde{M}} w_1^{(j)}}$ where $w_1^{(j)}$ denotes the conditional likelihood: $\left| \Omega_1^{(j)} \right|^{-0.5} \exp \left(\tilde{e}_1 \left(\Omega_1^{(j)} \right)^{-1} \tilde{e}_1' \right)$ where $\Omega_1^{(j)} = H_1^{(j)} \Sigma_{e_t | \eta_t} H_1^{(j)'} with $H_1^{(j)} = \text{diag} \left(\exp \left(\tilde{h}_{1,[0]}^{(j)} \right) \right)$ and $\tilde{e}_1 = Z_1 - (H_1^{(j)})^{1/2} \mu_{e_t | \eta_t} + c + \sum_{j=1}^P \beta_j Z_{t-j} + \sum_{k=1}^K b_k \tilde{h}_{1,[-k]}^{(j)}$. The subscript $[0]$ denotes the contemporaneous value in the state vector while $[-k]$ denote the k lagged states.$

2. For $t = 2$ to T

- (a) Resample $F_{t-1}^{(j)}$ for $j = 1, 2, \dots, \tilde{M} - 1$ using indices $a_t^{(j)}$ with $\Pr \left(a_t^{(j)} = j \right) \propto p_{t-1}^{(j)}$

¹ See Nonejad (2015) for a recent application of this algorithm.

(b) Draw $F_t^{(j)} \setminus F_{t-1}^{(a_t^{(j)})}, \Xi$ for $j = 1, 2, \dots, \tilde{M} - 1$ using the transition equation of the model. Note that $F_{t-1}^{(a_t^{(j)})}$ denotes the resampled particles in step (a) above.

(c) Fix $F_t^{(\tilde{M})} = F_t^{(i-1)}$

(d) Sample $a_t^{(\tilde{M})}$ with $\Pr(a_t^{(\tilde{M})} = j) \propto p_{t-1}^{(j)} \Pr(F_t^{(i-1)} \setminus F_{t-1}^{(j)}, C, \Psi, \tilde{Q})$ where the density $\Pr(F_t^{(i-1)} \setminus F_{t-1}^{(j)}, C, \Psi, \tilde{Q})$ is computed as $|\tilde{Q}|^{-0.5} \exp\left(N_t^{(j)} (\tilde{Q})^{-1} N_t^{(j)}\right)$. This constitutes the ancestor sampling step.

If $a_t^{(\tilde{M})} = \tilde{M}$ then the algorithm collapses to the simple particle Gibbs.

(e) Update the weights $p_t^{(j)} = \frac{w_t^{(j)}}{\sum_{j=1}^{\tilde{M}} w_t^{(j)}}$ where $w_1^{(j)}$ denotes the conditional likelihood: $|\Omega_t^{(j)}|^{-0.5} - 0.5 \exp\left(\tilde{e}_t (\Omega_t^{(j)})^{-1} \tilde{e}_t'\right)$

3. End

4. Sample $F_t^{(i)}$ with $\Pr(F_t^{(i)} = F_t^{(j)}) \propto p_T^{(j)}$ to obtain a draw from the conditional posterior distribution

We use $\tilde{M} = 20$ particles in our application. The initial values μ_0 defined above are used to initialise step 1 of the filter.

3.3 A Monte-Carlo experiment

We conduct a small Monte-Carlo experiment to evaluate the performance of the algorithm. We generate data from the following DGP

$$\begin{pmatrix} \ln h_{1t} \\ \ln h_{2t} \end{pmatrix} = \begin{pmatrix} 0.85 & -0.1 \\ 0.1 & 0.85 \end{pmatrix} \begin{pmatrix} \ln h_{1t-1} \\ \ln h_{2t-1} \end{pmatrix} + \begin{pmatrix} -0.05 & 0.01 \\ -0.05 & 0.01 \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix}$$

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} 0.3 \\ -0.3 \end{pmatrix} + \begin{pmatrix} 0.5 & -0.1 \\ 0.1 & 0.5 \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} -0.05 & 0.01 \\ -0.05 & 0.01 \end{pmatrix} \begin{pmatrix} \ln h_{1t-1} \\ \ln h_{2t-1} \end{pmatrix} + \begin{pmatrix} h_{1t}^{1/2} e_{1t} \\ h_{2t}^{1/2} e_{2t} \end{pmatrix}$$

$$\begin{pmatrix} \eta_{1t} \\ \eta_{2t} \\ e_{1t} \\ e_{2t} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.2 & 0.3 & -0.4 \\ 0.2 & 1 & 0.6 & 0.2 \\ 0.3 & 0.6 & 1 & -0.2 \\ -0.4 & 0.2 & -0.2 & 1 \end{pmatrix} \right]$$

We generate 600 observations and discard the first 100 to remove the effect of initial conditions. The experiment is repeated 100 times. At each iteration, we estimate the model using an MCMC run of 10000 iterations, with a burn-in of 5000 iterations. The particle Gibbs step employs 20 particles.

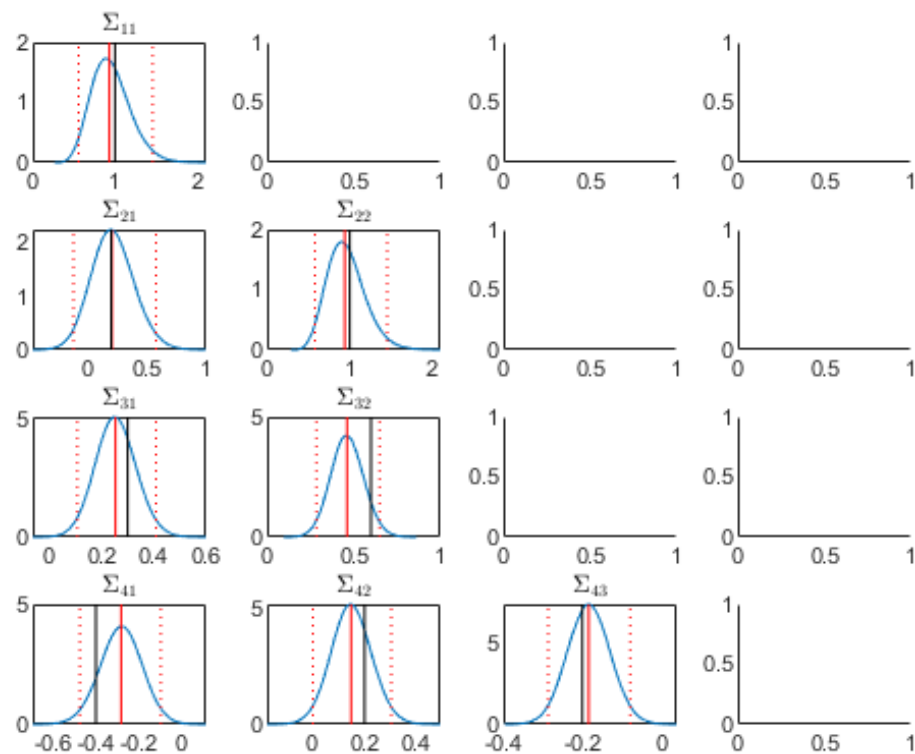


Figure 1: Monte-Carlo estimates of the elements of the covariance matrix. The posterior distribution of estimates is shown by the blue line, the median and 95% error band by the solid and dotted red lines, respectively. The black line is the true value

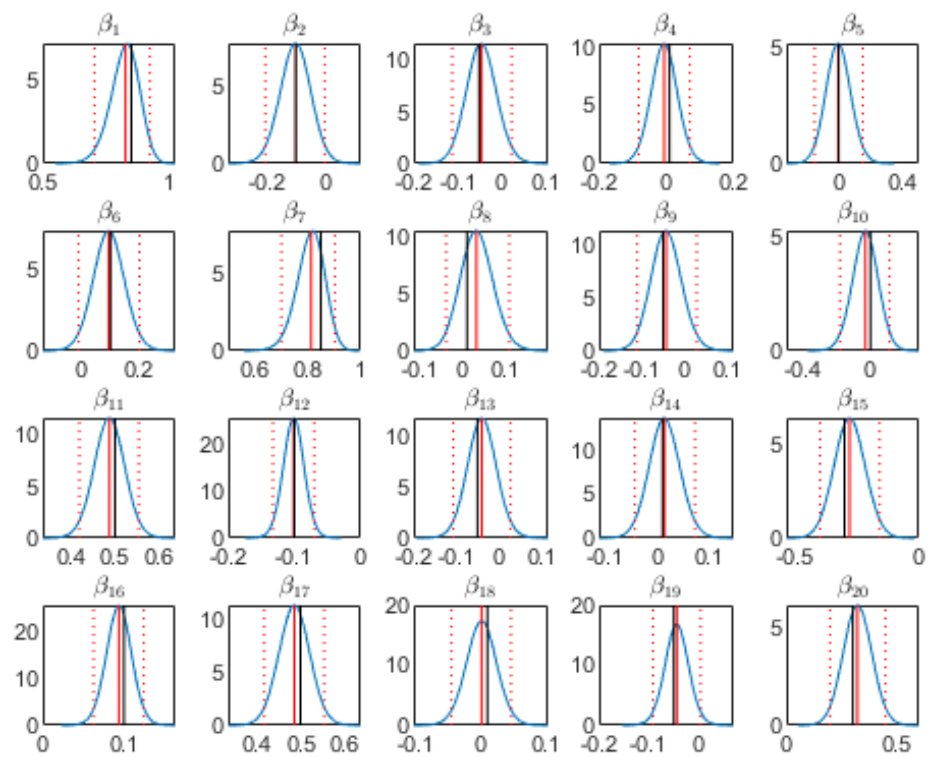


Figure 2: Monte-Carlo estimates of the VAR coefficients. See notes to Figure 1.

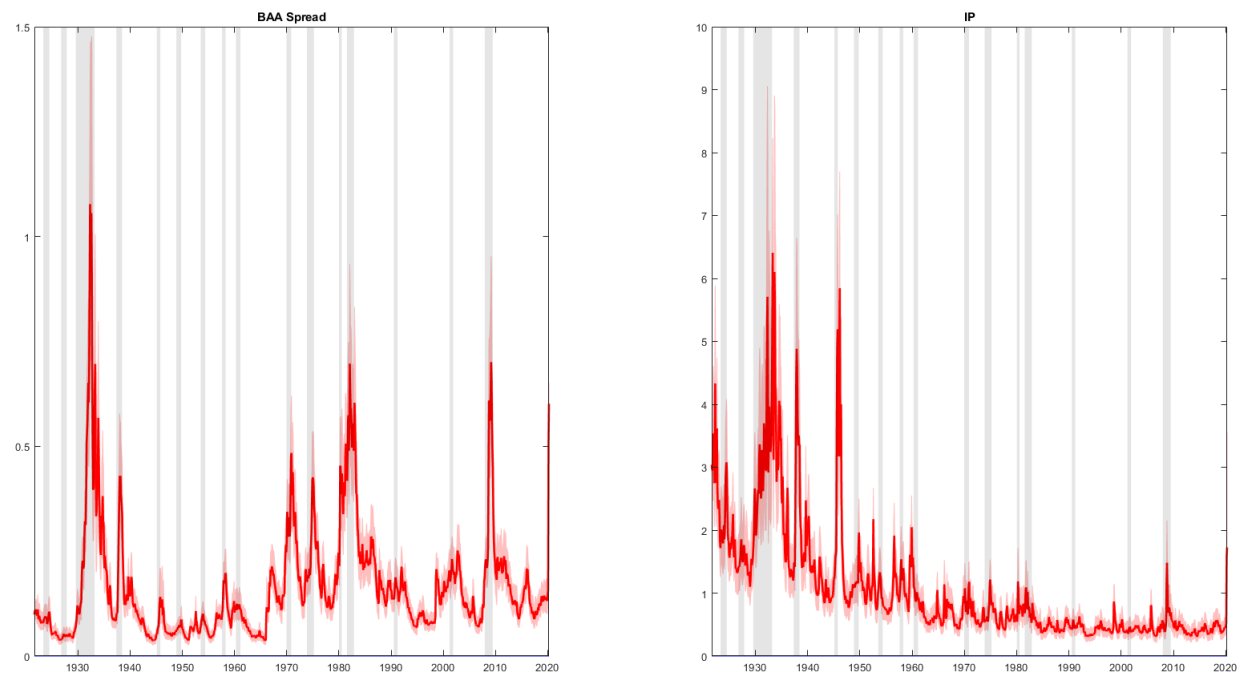


Figure 3: Stochastic volatility

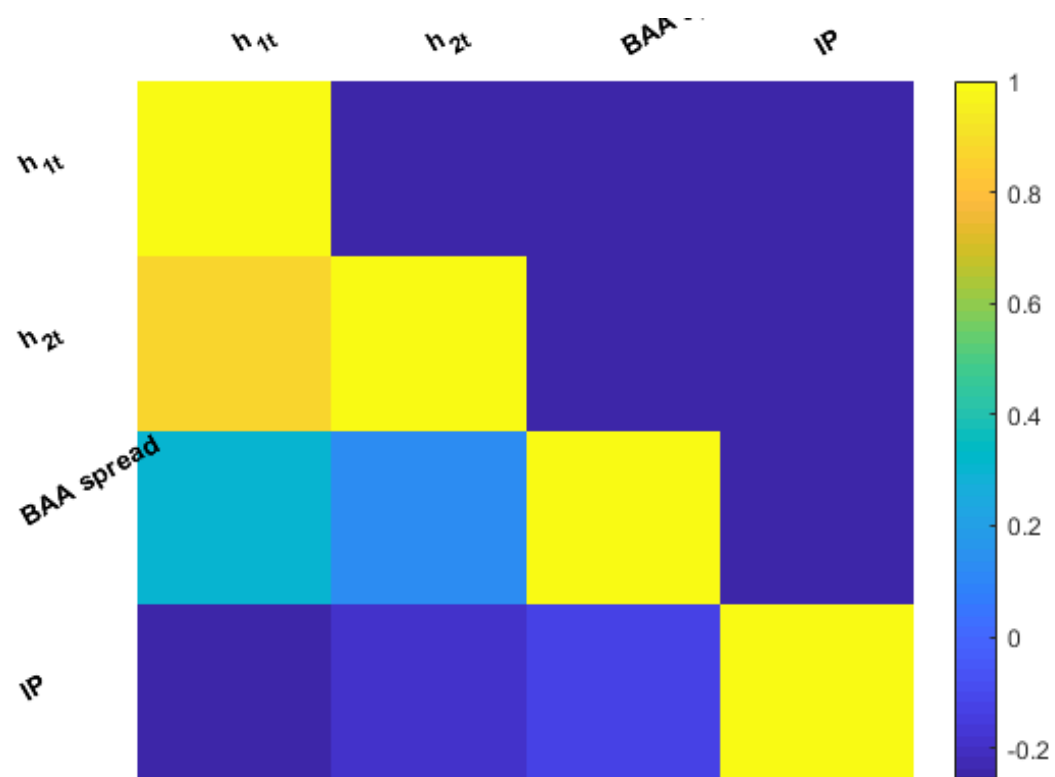


Figure 4: Correlation of the VAR residuals

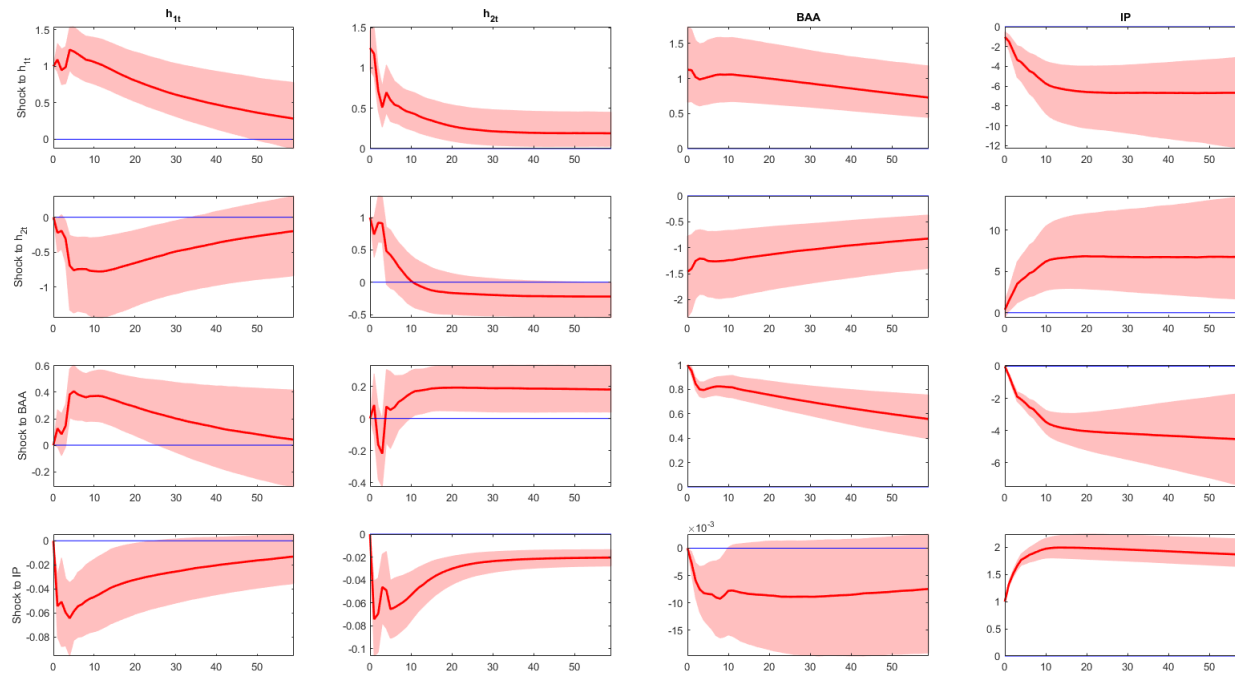


Figure 5: Impulse responses

Figure 1 plots the posterior distribution of Σ averaged over the Monte Carlo draws along with the true values. The posterior estimates are reasonably close to the true values for most elements, with the distribution covering the true values. Figure 2 shows the estimates of the VAR coefficients (i.e. the coefficients of the transition equation and the observation equation including intercepts). The estimates of these parameters are close to the true values.

3.4 A simpler specification of the model

The Metropolis steps in the algorithm can be eliminated completely if certain simplifying assumptions are made. Consider a model where it assumed that the stochastic volatility is common across disturbances (see Carriero *et al.* (2012)). This version of the model is defined as:

$$\lambda_t = \alpha + \theta\lambda_{t-1} + \sum_{j=1}^Q d_j Z_{t-j} + \eta_t \quad (13)$$

$$Z_t = c + \sum_{j=1}^P \beta_j Z_{t-j} + \sum_{k=1}^K b_k \tilde{h}_{t-k} + H_t^{1/2} e_t \quad (14)$$

where Z_t is a matrix of endogenous variables, $H_t = \text{diag}(\exp(\lambda_t) \cdot \mathbf{1}_{1 \times N})$, where λ_t is a scalar process.

The $M = 1 + N$ disturbances $\varepsilon_t = \begin{pmatrix} \eta_t \\ e_t \end{pmatrix}$ are distributed normally $N(0, \Sigma)$, $\Sigma = \begin{pmatrix} \Sigma_\eta & \Sigma'_{\eta_e} \\ \Sigma_{\eta_e} & \Sigma_e \end{pmatrix}$ where

the first diagonal element of Σ_e is restricted to equal 1 for normalisation. For example for $M = 3, N = 2$: $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & 1 & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix}$. The reduced form disturbances $H_t^{1/2} e_t$ have a variance covariance ma-

trix given by $H_t^{1/2} \Sigma_e H_t^{1/2'}$. In the $N = 2$ example this is: $\begin{pmatrix} \exp(\lambda_t)^{0.5} & 0 \\ 0 & \exp(\lambda_t)^{0.5} \end{pmatrix} \begin{pmatrix} 1 & \Sigma_{23} \\ \Sigma_{32} & \Sigma_{33} \end{pmatrix}$

$\begin{pmatrix} \exp(\lambda_t)^{0.5} & 0 \\ 0 & \exp(\lambda_t)^{0.5} \end{pmatrix}$. The main advantage of this formulation is the fact that step 4 of the algo-

rithm, the draw from $H(\Sigma_e | B, B^h, \Sigma_\eta^*, \tilde{h}_t)$ can be carried out directly without resorting to a Metropolis

step. This is because Σ_e can be written as $A^{-1} D A^{-1'}$ where A is lower triangular with ones on the main diagonal and D is a diagonal matrix with the first diagonal element fixed to 1. As discussed in Chan and

Jeliazkov (2009), the conditional posterior is available in closed form for the elements of A and D .

Related to this specification is a version of the benchmark model where only one shock has stochastic volatility. This model is defined as:

$$h_{1t} = \alpha + \theta h_{1t-1} + \sum_{j=1}^Q d_j Z_{t-j} + \eta_t \quad (15)$$

$$Z_t = c + \sum_{j=1}^P \beta_j Z_{t-j} + \sum_{k=1}^K b_k h_{1t-k} + H_t^{1/2} e_t \quad (16)$$

where $H_t = \text{diag}([\exp(h_{1t}), 1_{1 \times N-1}])$ and $\Sigma = \begin{pmatrix} \Sigma_\eta & \Sigma'_{\eta_e} \\ \Sigma_{\eta_e} & \Sigma_e \end{pmatrix}$. As in the example above, only the first diagonal element of Σ_e is restricted to 1 enabling the use of a Gibbs step to draw Σ_e .

4 Empirical application

We estimate a monthly VAR model for the US using the spread of BAA corporate bond yield over the 10 year rate (BAA) and log of industrial production (IP). The sample is monthly and runs from 1919M1 to 2020M5. The lag lengths P, K, Q are set to 12, 6, 6 respectively.

Figure 3 shows the posterior estimates of the stochastic volatilities. While both volatilities rise during NBER recessions, the volatility of the shocks to IP has become smaller over time. It is interesting to note that volatility has risen sharply during the last month of the sample, reflecting uncertainty associated with the pandemic. Figure 4 plots the posterior median of the correlation matrix of the residuals. The shock to the volatility of BAA is positively correlated with the shock to BAA and negatively correlated with the shock to IP. A similar pattern can be seen in the correlations between the shock to the volatility of IP and the level disturbances. Figure 5 shows the estimated (linear) impulse responses based on a Cholesky decomposition. Using this identification the shock to the volatility of BAA is recessionary while a shock to the volatility of IP acts in the opposite manner as it pushes down BAA volatility and the BAA spread. The responses to the level financial shock are similar to the responses to BAA volatility shock. Finally a positive shock to IP reduces both volatilities.

5 Conclusions

This note presents an alternative algorithm to estimate the VAR with stochastic volatility proposed in Muntaz (2018). The main advantage of the new approach is that a Metropolis step is no longer required to sample the variance of the shocks to the stochastic volatilities.

References

- Andrieu, Christophe, Arnaud Doucet and Roman Holenstein, 2010, Particle Markov chain Monte Carlo methods, *Journal of the Royal Statistical Society Series B* **72**(3), 269–342.
- Banbura, Marta, Domenico Giannone and Lucrezia Reichlin, 2007, Bayesian VARs with Large Panels, *CEPR Discussion Papers* 6326, C.E.P.R. Discussion Papers.
- Carriero, Andrea, Todd Clark and Massimiliano Marcellino, 2012, Common Drifting Volatility in Large Bayesian VARs, *CEPR Discussion Papers* 8894, C.E.P.R. Discussion Papers.
- Chan, Joshua Chi-Chun and Ivan Jeliazkov, 2009, MCMC Estimation of Restricted Covariance Matrices, *Journal of Computational and Graphical Statistics* **18**(2), 457–480.
- Clark, Todd E., 2011, Real-Time Density Forecasts From Bayesian Vector Autoregressions With Stochastic Volatility, *Journal of Business & Economic Statistics* **29**(3), 327–341.
- Cogley, T. and T. J. Sargent, 2005, Drifts and Volatilities: monetary policies and outcomes in the Post WWII U.S., *Review of Economic Dynamics* **8**, 262–302.
- Lindsten, Fredrik, Michael I. Jordan and Thomas B. Schön, 2014, Particle Gibbs with Ancestor Sampling, *Journal of Machine Learning Research* **15**, 2145–2184.
- Muntaz, Haroon, 2018, A generalised stochastic volatility in mean VAR, *Economics Letters* **173**, 10 – 14.
- Nonejad, Nima, 2015, REPLICATING THE RESULTS IN ŠA NEW MODEL OF TREND INFLATIONŠ USING PARTICLE MARKOV CHAIN MONTE CARLO, *Journal of Applied Econometrics* pp. n/a–n/a.

School of Economics and Finance



This working paper has been produced by
the School of Economics and Finance at
Queen Mary University of London

Copyright © 2020 Haroon Mumtaz

All rights reserved

School of Economics and Finance Queen
Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: [www.econ.qmul.ac.uk/research/
workingpapers/](http://www.econ.qmul.ac.uk/research/workingpapers/)