Bayesian compressed vector autoregressions

Gary Koop, Dimitris Korobilis, Davide Pettenuzzo

comments by

Sylvia Kaufmann

Study Center Gerzensee

9th ECB Workshop on Forecasting Techniques:

Forecast Uncertainty and Macroeconomic Indicators

Frankfurt a.M., June 3-4, 2016

Content/Structure of the paper

- Data-driven forecasts using a Bayesian compressed VAR, suggestion taken from Bayesian compressed regression analysis
  Note: Compression is not Bayesian

$$Y_t = B^c(\Phi Y_{t-1}) + \epsilon_t,\ \epsilon_t \sim N(0, \Omega)$$

  Parameter reduction achieved by using the structural form, $\tilde{A}$ is lower diagonal with zeros on the diagonal, Σ diagonal

$$\begin{aligned}
\left(\overbrace{I_n + \tilde{A}}^{=A}\right) Y_t &= \Gamma Y_{t-1} + \Sigma E_t,\ E_t \sim N(0, I_n),\ A\Omega A' = \Sigma\Sigma \\
Y_t &= \Gamma Y_{t-1} + \tilde{A}(-Y_t) + \Sigma E_t,\ \Gamma = AB \\
Y_{it} &= \Theta_i^c(\Phi Z_{it}) + \sigma_i E_{it} \\
& Z'_{it} = [Y'_{t-1} Y'_{-it}],\ Y'_{-it} = (Y_{1t}, \ldots, Y_{i-1,t})
\end{aligned}$$

Random simulation of compression matrix Φ_i, $m \times pn + (i-1)$

$$Pr\left(\Phi_{i,jk} = 1/\sqrt{\varphi}\right) = \varphi^2$$
$$Pr\left(\Phi_{i,jk} = 0\right) = 2(1-\varphi)\varphi$$
$$Pr\left(\Phi_{i,jk} = -1/\sqrt{\varphi}\right) = (1-\varphi)^2$$

$m \sim U[1, 5\ln(pn + (i-1))]$, $\varphi \sim U[0.1, 0.8]$,
Note: recommendation is $m \sim U[2\ln(pn + (i-1)), \min(T, pn + (i-1))]$
$\rightarrow$ do m and φ of retained $BCVAR$s cluster around certain values?

- Extension to time-varying parameters (Koop and Korobilis, 2013):

$$Y_{it} = \Theta_{it}^{c}(\Phi Z_{it}) + \sigma_{it} E_{it}$$

- Forecasting exercise, Jan 60-Jun 87...Dec 14 (396 re-estimations):
  Construct medium, large and huge VAR
  (needs BMA for $BCVAR$ and $BCVAR_c$)
  Forecast horizon: $h = 1, 2, 3, 6, 9, 12$, for 7 variables
  Alternative models:
  $AR(1)$, DFM, $FAVAR$, $BVAR$, $BCVAR$, $BCVAR_c$, $BCVAR_{tvp}$
  (model choice according to BIC)
  Evaluation criteria: $MSFE$, multivariate $MSWFE$, $ALPL$, multivariate $ALPL$, significance against $AR(1)$ evaluated according to EPA

- Forecasting exercise, Jan 60-Jun 87...Dec 14 (396 re-estimations):
  Construct medium, large and huge VAR
  (needs BMA for $BCVAR$ and $BCVAR_c$)
  Forecast horizon: $h = 1, 2, 3, 6, 9, 12$, for 7 variables
  Alternative models:
  $AR(1)$, DFM, $FAVAR$, $BVAR$, $BCVAR$, $BCVAR_c$, $BCVAR_{tvp}$
  (model choice according to BIC)
  Evaluation criteria: $MSFE$, multivariate $MSWFE$, $ALPL$, multivariate $ALPL$, significance against $AR(1)$ evaluated according to EPA

$\longrightarrow$ **FAT** paper

Conceptual comments/questions

- Unsupervised versus supervised data compression.
  Supervised because of BMA targeted towards 1 variable?
- Columns of Φ sum to 1 $\rightarrow$ dimension reduction, but not in the sense of variable selection (zero columns in Φ)?
- Conjecture that theorems from Bayesian compressed regressions apply to $BCVAR \rightarrow Z_{t-1}$ are stochastic regressors?
- Is the compressed structural system a VAR? Can one write something like

$$\begin{aligned} Y_{it} &= \Theta_i^c \left(\Phi_i Z_{it}\right) + \sigma_i E_{it} \\ \tilde{Z}_{it} &= \tilde{\Theta}_i^c \tilde{Z}_{i,t-1} + \tilde{E}_{it} \end{aligned}$$

  and is there an interpretation for it?

Conceptual comments/questions continued

- What is the relation to factor analysis (non-stochastic factors)?

$$\begin{aligned}\Sigma_Y &= \mathbf{F}^c\mathbf{F}^{c\prime} + \Sigma^2 \\ &= \underbrace{\mathbf{F}_1^c\mathbf{F}_1^{c\prime}}_{\text{explained}} + \underbrace{\mathbf{F}_2^c\mathbf{F}_2^{c\prime}}_{\text{unexplained}} + \Sigma^2\end{aligned}$$

Minimum rank factor analysis minimizes

$$\phi(\Sigma^2) = \sum_{j=m+1}^{n} \lambda_j \left(\Sigma_Y - \Sigma^2\right)$$

Write Koop et al. as (a very special factor model)

$$Y_t = \begin{bmatrix} \Theta_1^c & & & 0 \\ & \Theta_2^c & & \\ & & \ddots & \\ 0 & & & \Theta_k^c \end{bmatrix} \begin{bmatrix} \Phi_1 Z_{1t} & & & 0 \\ & \Phi_2 Z_{2t} & & \\ & & \ddots & \\ 0 & & & \Phi_k Z_{kt} \end{bmatrix} + \Sigma^2$$

$\rightarrow$ Are $\Phi_i Z_{it}$ stochastic or non-stochastic factors?

$\rightarrow$ BMA procedure (based on BIC) minimizes $\phi(\Sigma^2)$

Comments: Forecasting and forecasting results

- Reduced VAR is re-build to obtain forecasts.
  $\rightarrow$ note: forecasts can be obtained by recursive forecasting

- BMA is optimized conditional on the structural form (imposed sequencing)
  $\rightarrow$ depends on variable order or on ordering of groups of variables?
  $\rightarrow$ I would rank variables of interest last!

- Estimate models and evaluate forecasts for variables separately
  $\rightarrow$ what is left of the VAR part?
  $\rightarrow$ only focus on multivariate forecast evaluations for a smaller set of variables (apply BMA for a small number of variables jointly)?
  $\rightarrow$ alternative model: include all lagged and all current (other) variables? (back to regression compression)

- Significance of forecast improvement is evaluated against $AR(1)$
  $\rightarrow$ significance against other multivariate models difficult to judge

- Results for CPI and IP (point forecasts): Large VAR (IP short term), huge VAR (short term)
  Otherwise, it is difficult to extract regularities

- It would be interesting to evaluate the model's performance around 'turning points': Does it yield forecasts that are way out of other models' forecasts? $\rightarrow$ this is additional information!

Conclusion

- It is a nice paper! Tough work ... also for the computers!
- Bayesian compressed VAR analysis is a nice extension to the model suite for forecasting
- It would be nice to have more theoretical results on the nature of the process
  $\rightarrow$ avenue for future research?