Bayesian Compressed Vector Autoregressions

Gary Koop[a], Dimitris Korobilis[b], and Davide Pettenuzzo[c]

[a]University of Strathclyde

[b]University of Glasgow

[c]Brandeis University

9th ECB Workshop on Forecasting Techniques

Forecast Uncertainty and Macroeconomic Indicators

June 2-3, 2016

Large VARs

- Since Sims (1980), Vector autoregressions (VARs) are an important tool in applied macroeconomics
- Recently, big focus on forecasting in "data-rich" environments, relying on large VARs with dozens of dependent variables
- Typically these large models have many more parameters than observations
  - E.g., with $n = 100$ variables and $p = 12$ lags, there are $120,000$ parameters to estimate (excluding intercepts)
- Solutions involve dynamic factor models, shrinkage methods (LASSO, Elastic net, etc.), Bayesian variable selection
- Recent studies indicate that large Bayesian VARs can be quite competitive in forecasting
  - Banbura, Giannone and Reichlin (2010);
  - Carriero, Kapetanios and Marcellino (2009, 2011, 2012);
  - Koop (2011), Koop and Korobilis (2013); Korobilis (2013)

What we do in this paper

1. We build on ideas from the machine learning literature and apply Bayesian "compressed regression" methods to large VARs

   Main idea:

   - Compress the VAR regressors through random projection
   - Use BMA to average across different random projections

2. We apply Bayesian compressed VARs to forecast a 130-variable VARs with 13 lags (similar to Banbura et al (2010)), with more than $200,000$ parameters to estimate

   - Find good forecasting performance, relative to a host of alternative methods including DFM, FAVAR, and BVAR with Minnesota priors

3. Extend the Bayesian compressed VARs to feature time-varying coefficients and volatilities, and further improve forecasting performance

Bayesian Compressed Regression (BCR)

- Start with the case of a scalar dependent variable y_t, $t = 1, ..., T$, predictor matrix $x_t = (x_{t,1}, ..., x_{t,k})'$, and linear regression model

$$y_t = x_t'\beta + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$$

  When $k \gg T$, estimation is either impossible (e.g. MLE), or computationally very hard (e.g. Bayesian regression with natural conjugate priors)
- Guhaniyogi and Dunson (2015) consider a compressed regression specification

$$y_t = (\Phi x_t)' \beta^c + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$$

  where Φ is an $(m \times k)$ compression matrix with $m \ll k$
- Conditional on Φ, estimating β^c and forecasting y_{t+1} is now very straightforward, and can be carried out using standard (Bayesian) regression methods

Projection matrix

- The elements $\{\Phi_{ij}\}$ can be generated quickly, e.g.

$$\Phi_{ij} \sim \mathcal{N}(0,1)$$

  Alternatively, Achlioptas (2003) use a sparse random projection

$$\Phi_{ij} = \begin{cases} -\sqrt{\varphi} & \text{with probability} & 1/2\varphi \\ 0 & \text{with probability} & 1-1/\varphi \\ \sqrt{\varphi} & \text{with probability} & 1/2\varphi \end{cases}$$

  where $\varphi = 1$ or 3.
- We follow the scheme of Guhaniyogi and Dunson (2015)

$$\Phi_{ij} = \begin{cases} -\frac{1}{\sqrt{\varphi}} & \text{with probability} & \varphi^2 \\ 0 & \text{with probability} & 2\left(1-\varphi\right)\varphi \\ \frac{1}{\sqrt{\varphi}} & \text{with probability} & \left(1-\varphi\right)^2 \end{cases}$$

  where $\varphi \in (0.1, 0.9)$ and is estimated from the data; the rows of Φ are normalized using Gram-Schmidt orthonormalization.

RP vs. PCA

Model Averaging

- Guhaniyogi and Dunson (2015) show that BCR produces a predictive density for y_{t+1} that (under mild conditions) converges to its true predictive density (large k, small T asymptotics)
- To limit sensitivity of results to choice of m and φ, generate R random compressions based on different (m, φ) pairs.
- Use BMA to integrate out (m, φ) from predictive density of y_{t+1}:

$$p\left(y_{t+1}|\mathcal{Y}^t\right) = \sum_{r=1}^{R} p\left(y_{t+1}|M_r, \mathcal{Y}^t\right) p\left(M_r|\mathcal{Y}^t\right)$$

where $p\left(M_r|\mathcal{Y}^t\right)$ denotes model M_r posterior probability (computed using standard BMA formula) and M_r denotes the r-th pair of (m, φ) values, where:

  - $\varphi \sim \mathcal{U}\left(0.1, 0.9\right)$
  - $m \sim \mathcal{U}\left(2\ln\left(k\right), \min\left(T, k\right)\right)$

VAR setup

- VAR(p) for $n \times 1$ vector of dependent variables is :

$$Y_t = a_0 + \sum_{j=1}^{p} A_j Y_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \Omega)$$

  Rewrite this compactly as

$$Y_t = BX_t + \varepsilon_t$$

  where B is an $n \times k$ matrix of coefficients, X_t is $k \times 1$, and $k = np + 1$. Also, note that Ω has $n(n+1)/2$ free parameters

- Potentially, many parameters to estimate. E.g., when $n = 130$ and $p = 13$, B has $220,000+$ parameters to estimate, while Ω has $8,500+$ unconstrained elements

Bayesian Compressed VAR (BCVAR)

- Define the Compressed VAR as

$$Y_t = B^c \left(\Phi X_t \right) + \varepsilon_t$$

  where the projection matrix Φ is $m \times k$, $m \ll k$
- Conditional on a given Φ (its elements randomly drawn as before), estimation and forecasts for the compressed VAR above are trivial and very fast to compute
- Note:
  - h-step ahead forecasts (for $h > 1$) not available analytically. For those, rewrite compressed VAR as

$$Y_t = \left(B^c \Phi \right) X_t + \varepsilon_t$$

    and iterate forward in the usual way
  - The compressed VAR above imposes the same compression (ΦX_t) in all equations; may be too restrictive
  - So far, no compression is applied to the elements of Ω

Compressing the VAR covariance matrix

- Ω has $n\left(n+1\right)/2$ unconstrained elements, so we modify the BCVAR to allow also for their compression
- Following common practice (e.g., Primiceri, 2005 and Eisenstat, Chan and Strachan, 2015) we use a triangular decomposition of Ω

$$A\Omega A' = \Sigma\Sigma,$$

Σ is a diagonal matrix with diagonal elements σ_i
A is a lower triangular matrix with ones on the diagonal

- Define $A = I + \widetilde{A}$, where $\widetilde{A}$ is lower triangular but with zeros on the diagonal, and rewrite uncompressed VAR as

$$\begin{aligned} Y_t &= \Gamma X_t + \widetilde{A}\left(-Y_t\right) + \Sigma E_t \\ &= \Theta Z_t + \Sigma E_t \end{aligned}$$

where $E_t \sim \mathcal{N}\left(0, I_n\right)$, $Z_t = \left[X_t, -Y_t\right]$ and $\Theta = \left[\Gamma, \widetilde{A}\right]'$

Compressing the VAR covariance matrix - contn'd

- Compression can be accomplished as follows:

$$Y_t = \Theta^c \left(\Phi Z_t \right) + \Sigma E_t$$

  where Φ is now an $m \times (k+n)$ random compression matrix
  Note that we would still be relying on the same compression matrix (Φ) for all equations

- Alternatively, we can allow each equation to have its own random compression matrix (of size $m_i \times (k+i-1)$):

$$Y_{i,t} = \Theta_i^c \left(\Phi_i Z_{i,t} \right) + \sigma_i E_{i,t}$$

  Having n compression matrices (each of different dimension and with different randomly drawn elements) allows for the explanatory variables of different equations to be compressed in potentially different ways

Estimation and Predictions

- Estimation is performed equation-by-equation (Zellner, 1971), conditional on a known (generated) Φ_i
- We choose a standard natural conjugate prior:

$$\begin{aligned} \Theta_i^c &\sim \mathcal{N}\left(\underline{\Theta}_i^c, \sigma_i^2 \underline{V}_i\right) \\ \sigma_i^{-2} &\sim \mathcal{G}\left(\underline{s}^{-2}, \underline{\nu}\right) \end{aligned}$$

  where $i = 1, ..., n$.
  Posterior location and scale parameters for $\Theta_i^c, \sigma_i^{-2}$ are available analytically
- 1-step ahead forecasts are also available analytically
- h-step ahead forecasts (for $h > 1$) require some extra work
  Rewrite compressed VAR as

$$Y_{i,t} = \left(\Theta_i^c \Phi_i\right) Z_{i,t} + \sigma_i E_{i,t}$$

  and iterate forward in the usual way, one equation at a time

Model averaging

- We generate many random $\Phi^{(r)}$ (or $\Phi_i^{(r)}$), $r = 1, ..., R$ based on different (m, φ) pairs, then implement BMA as follows
- First, we rely on BIC instead of the marginal likelihood. We compute model M_r BIC as

$$BIC_r = \ln\left(\left|\overline{\Sigma}_r\right|\right) + \frac{\ln(t)}{t}\left(n \times \sum_{i=1}^{n} m_i\right)$$

Posterior model probability is approximated by

$$\Pr\left(M_r|\mathcal{Y}^t\right) \approx \frac{\exp\left(-\frac{1}{2}BIC_r\right)}{\sum_{\varsigma=1}^{R}\exp\left(-\frac{1}{2}BIC_\varsigma\right)}$$

- Next,

$$p\left(Y_{t+h}|\mathcal{Y}^t\right) = \sum_{r=1}^{R} p\left(Y_{t+h}|M_r, \mathcal{Y}^t\right) p\left(M_r|\mathcal{Y}^t\right)$$

where $h = 1, ..., H$

Data

- We use the "FRED-MD" monthly macro data (McCracken and Ng, 2015), 2015-05 vintage
  - 134 series covering: (1) the real economy (output, labor, consumption, orders and inventories), (2) money and prices, (3) financial markets (interest rates, exchange rates, stock market indexes).
- Series are transformed as in Banbura et al (2010) by applying logarithms, excepts when series are already expressed in rates
- Final sample is 1960M3 - 2014M12 (658 obs.)
- We focus on forecasting: Employment (PAYEMS), Inflation (CPIAUCSL), Federal fund rate (FEDFUNDS), Industrial production (INDPRO), Unemployment rate (UNRATE), Producer Price Index (PPIFGS), and 10 year US Treasury Bond yield (GS10).

VAR specifications

- We have three sets of VARs: **Medium**, **Large**, and **Huge**
- All VARs include seven key variables of interest: Employment, Inflation, Fed Fund rate, IPI, Unemployment, PPI, and 10 yr bond yield
- **Medium** VAR has 19 variables - similar to Banbura et al (2010)
- **Large** VAR has 46 variables - similar to Carriero et al (2011)
- **Huge** VAR has 129 variables
- Note: All four VARs produce forecasts for the variables of interest, but imply different information sets

Forecast evaluation

- We forecast $h = 1$ to 12 months ahead
- Initial estimation based on first half of the sample, $t = 1, ..., T_0$; forecast evaluation over the remaining half, $t = T_0 + 1, ..., T - h$ ($T_0 = 1987M7$, $T = 2014M12$)
- Forecasts are computed recursively, using an expanding estimation window.
- We evaluate forecasts relative to an AR(1) benchmark and focus on
  - Mean squared forecast error (MSFE)
  - Cumulative sum of squared forecast errors (Cum SSE)
  - Average (log) predictive likelihoods (ALPLs)
- Competing methods are **DFM** using PCA as in Stock and Watson (2002), **FAVAR** using PCA as in Bernanke et al (2005) with selection of lags and factors using BIC, and **BVAR** with Minnesota prior as in Banbura et al (2010)

Relative MSFE ratios, Large VAR

Variable	*DFM*	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*	*DFM*	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*
			$h = 1$					$h = 2$		
PAYEMS	1.129	0.906	**0.788****	0.888***	0.905**	0.871	0.698***	**0.521*****	0.805***	0.837***
CPIAUCSL	1.149	1.100	1.009	0.998	**0.953**	1.172	1.118	1.110	0.943	**0.909****
FEDFUNDS	2.436	1.671	2.461	**1.093**	1.103	1.979	1.415	2.594	**0.991**	1.150
INDPRO	0.824**	0.890**	**0.783*****	0.838***	0.914***	0.852	0.966	**0.772****	0.934*	0.929*
UNRATE	0.841*	**0.786*****	0.824*	0.798***	0.855**	0.677*	**0.661*****	0.666*	0.722***	0.758**
PPIFGS	1.039	1.013	1.045	0.987	**0.986**	1.149	1.056	1.162	1.012	**1.004**
GS10	1.010	**0.983**	1.106	1.006	0.992	1.020	**0.987**	1.149	1.052	1.044
			$h = 3$					$h = 6$		
PAYEMS	0.799	0.710***	**0.489*****	0.757***	0.748***	0.826	0.832**	**0.647***	0.773**	0.762***
CPIAUCSL	1.129	1.067	1.154	0.948	**0.913****	1.042	0.970	0.999	0.910**	**0.902****
FEDFUNDS	1.695	**1.026**	2.241	1.034	1.108	1.252	**0.944**	1.224	1.098	1.088
INDPRO	0.911	0.943	**0.862**	0.957*	0.952	**0.923**	0.962	0.980	0.959	0.984
UNRATE	**0.603***	0.631***	0.615*	0.677***	0.731**	**0.600**	0.663***	0.617	0.671***	0.712**
PPIFGS	1.151	1.018	1.177	1.021	**1.013**	1.108	1.014	1.095	1.012	**0.998**
GS10	1.041	**1.030**	1.222	1.057	1.059	1.033	**1.018**	1.115	1.043	1.029
			$h = 9$					$h = 12$		
PAYEMS	0.891	0.930	**0.840**	0.865	0.844**	**0.919**	0.966	0.999	0.972	0.940
CPIAUCSL	1.052	0.975	0.932	0.887**	**0.867*****	1.046	0.984	0.904	0.902**	**0.879*****
FEDFUNDS	1.085	**0.999**	1.139	1.060	1.028	1.065	**0.994**	1.259	1.092	1.041
INDPRO	**0.963**	0.972	1.018	1.004	0.999	**0.956**	0.979	1.056	1.002	1.020
UNRATE	**0.663**	0.684***	0.715	0.698**	0.739**	0.715*	**0.710*****	0.831	0.728**	0.756**
PPIFGS	1.063	1.000	1.051	**0.985**	0.992	1.082	1.002	1.039	1.004	**0.972**
GS10	1.005	**1.001**	1.050	1.009	1.019	1.011	**1.001**	1.054	1.023	1.014

▸ Medium VAR ▸ Huge VAR

Average (log) predictive likelihoods, Large VAR

Variable	*DFM*	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*	*DFM*	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*
			$h=1$					$h=2$		
PAYEMS	0.064***	0.107***	**0.259*** **	0.065***	0.063***	0.136***	0.168***	**0.399*** **	0.120***	0.113***
CPIAUCSL	-0.216	-0.239	-0.775	-0.078	**0.104**	-1.142	-0.581	-2.312	-0.222	**-0.220**
FEDFUNDS	0.016	0.060***	**0.149*** **	-0.017	-0.018	-0.001	**0.021** **	-0.023	-0.004	-0.012
INDPRO	-0.057	-0.034	-0.059	-0.011	**0.045*** **	0.054	0.105	**0.240** **	0.102*	0.146
UNRATE	0.114***	**0.138*** **	0.122**	0.096***	0.061***	0.211**	0.144**	**0.235** **	0.169***	0.144**
PPIFGS	-0.047	-0.030	-0.711	-0.023	**0.028**	-0.656	-0.213	-1.207	**0.023**	-0.144
GS10	**0.040***	0.036	-0.012	0.001	0.011	-0.002	**0.003**	-0.010	-0.021	-0.027
			$h=3$					$h=6$		
PAYEMS	0.136***	0.141***	**0.407*** **	0.148***	0.149***	0.092**	0.068***	**0.282*** **	0.114***	0.157***
CPIAUCSL	-0.593	**-0.232**	-1.949	-0.446	-0.269	-0.058	**0.048**	-0.889	-0.189	-0.002
FEDFUNDS	-0.012	0.027***	**0.032**	0.001	-0.001	-0.011	0.007*	**0.158*** **	-0.013	-0.014
INDPRO	0.060*	0.199	0.044	**0.223**	0.074*	**0.023**	-0.053	-0.294	-0.045	-0.081
UNRATE	**0.401**	0.116	0.356	0.357*	0.340*	**0.977**	0.495	0.502	0.958	0.869
PPIFGS	-0.228	**0.004**	-1.109	-0.020	-0.009	-0.147	-0.108	-0.857	-0.116	**-0.060**
GS10	**0.011**	0.002	-0.025	-0.002	-0.029	-0.006	**0.002**	-0.009	-0.021	-0.024
			$h=9$					$h=12$		
PAYEMS	0.063**	0.027*	**0.105***	0.065***	0.099***	**0.052**	0.039	0.016	0.041**	0.024
CPIAUCSL	-0.203	**0.027**	-0.943	-0.081	-0.228	**0.029**	-0.127	-0.784	-0.106	-0.053
FEDFUNDS	-0.004	-0.006	**0.148*** **	-0.024	-0.022	-0.005	-0.005	**0.136*** **	-0.028	-0.016
INDPRO	0.040	**0.122**	-0.168	0.092	-0.002	**0.087**	-0.057	-0.231	-0.109	0.058
UNRATE	1.437	**1.495**	0.180	1.326	1.136	1.097	**1.878**	-0.016	1.367	1.040
PPIFGS	-0.106	0.014	-0.629	0.029	**0.061**	-0.214	-0.185	-0.711	-0.150	**-0.145**
GS10	-0.011	-0.011	**0.024**	-0.022	-0.024	-0.017	-0.008	**0.010**	-0.018	-0.040

Forecast evaluation - contn'd

- We also look at the multivariate mean squared forecast error proposed by Christoffersen and Diebold (1998). Define the weighted forecast error of model i at time $\tau + h$ as

$$we_{i,\tau+h} = \left(e'_{i,\tau+h} \times W \times e_{i,\tau+h}\right)$$

$e_{i,\tau+h} = Y_{\tau+h} - \widehat{Y}_{i,\tau+h}$ is the $(N \times 1)$ vector of forecast errors, and W is an $(N \times N)$ matrix of weights

- We set the matrix W to be a diagonal matrix featuring on the diagonal the inverse of the variances of the series to be forecast
- Next, define

$$WMSFE_{ih} = \frac{\sum_{\tau=\underline{t}}^{\overline{t}-h} we_{i,\tau+h}}{\sum_{\tau=\underline{t}}^{\overline{t}-h} we_{bcmk,\tau+h}}$$

where $\underline{t}$ and $\overline{t}$ denote the start and end of the out-of-sample period

Forecast evaluation - contn'd

- Finally, we consider the multivariate average log predictive likelihood differentials between model i and the benchmark AR(1),

$$MVALPL_{ih} = \frac{1}{\overline{t} - \underline{t} - h + 1} \sum_{\tau=\underline{t}}^{\overline{t}-h} \left(MVLPL_{i,\tau+h} - MVLPL_{bcmk,\tau+h} \right),$$

where:

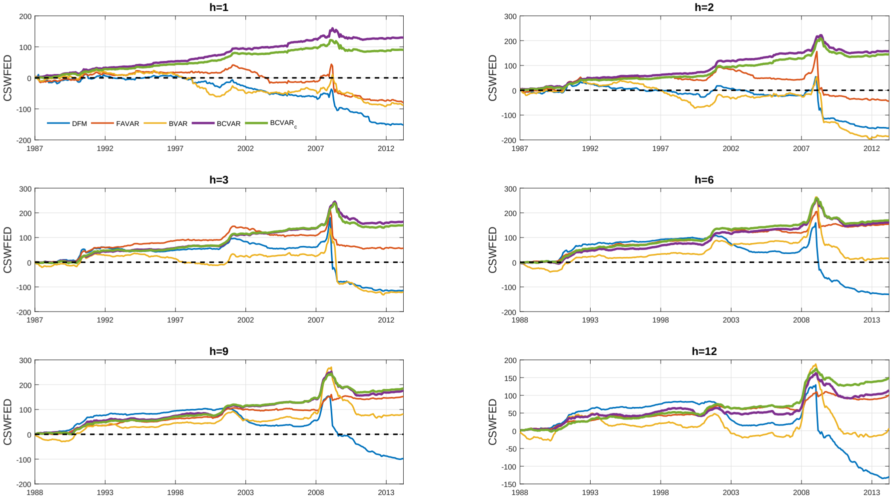
  - $MVLPL_{i,\tau+h}$ denote the multivariate log predictive likelihoods of model i at time $\tau + h$
  - and $MVLPL_{bcmk,\tau+h}$ denote the multivariate log predictive likelihoods of the benchmark model at time $\tau + h$

both computed under the assumption of joint normality.

Multivariate forecast comparisons

Fcst h.										
	Medium VAR									
	WTMSFE					*MVALPL*				
	DFM	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*	*DFM*	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*
h= 1	1.208	1.143	1.194	**0.936*** **	0.938***	0.587***	0.788***	**1.005*** **	0.919***	0.318***
h= 2	1.079	1.125	1.154	0.937*	**0.936** **	0.935***	0.912***	**1.222*** **	1.120***	0.514***
h= 3	1.032	1.051	1.082	0.949	**0.939* **	1.031***	1.053***	**1.362*** **	1.222***	0.575***
h= 6	1.029	0.961	1.005	**0.936**	0.938	0.983***	1.216***	**1.472*** **	1.383***	0.690***
h= 9	1.021	0.934	0.973	**0.926* **	0.930*	0.890**	1.336***	**1.502*** **	1.471***	0.703***
h=12	1.025	**0.941**	1.002	0.954	0.960	0.833*	1.434**	1.425***	**1.465*** **	0.699***
	Large VAR									
	DFM	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*	*DFM*	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*
h= 1	1.217	1.080	1.172	**0.968**	0.975	0.705***	0.830***	**0.913*** **	0.827***	0.257***
h= 2	1.175	1.051	1.224	**0.961**	0.979	0.974***	0.970***	0.937***	**1.071*** **	0.323***
h= 3	1.121	0.969	1.192	**0.962**	0.964	1.091***	1.089***	1.024***	**1.159*** **	0.424***
h= 6	1.019	**0.949** **	0.994	0.954	0.950*	1.176***	1.170***	**1.347*** **	1.280***	0.551***
h= 9	0.996	0.967*	0.988	0.953	**0.944* **	1.270***	1.311***	1.337***	**1.381*** **	0.574***
h=12	0.995	0.971	1.033	0.980	**0.960**	1.222***	**1.390*** **	1.134***	1.362***	0.476***
	Huge VAR									
	DFM	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*	*DFM*	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*
h= 1	1.094	1.050	1.055	**0.920*** **	0.944***	**0.938*** **	0.931***	0.760***	0.921***	0.272***
h= 2	1.081	1.023	1.098	**0.916** **	0.923**	1.124***	1.148***	0.875***	**1.203*** **	0.468***
h= 3	1.058	0.971	1.061	**0.917* **	0.924*	1.230***	**1.282*** **	1.012***	1.276***	0.510***
h= 6	1.062	0.927*	0.993	0.924	**0.920**	1.233***	1.426***	1.018***	**1.483*** **	0.675***
h= 9	1.045	0.930**	0.962	0.919	**0.915* **	1.137***	1.542***	1.027***	**1.555*** **	0.737***
h=12	1.058	0.955*	0.997	0.949	**0.933**	0.999***	**1.631*** **	0.712	1.593***	0.645***

Weighted Cum. Sum SSE diffs (Huge VAR)



Time-variation in Parameters: The Compressed TVP-VAR

- We generalize the compressed VAR to the case of a VAR with time-varying parameters and volatilities (BCVAR-TVP)
- The model becomes

$$Y_{i,t} = \Theta^c_{i,t}\left(\Phi_i Z_{i,t}\right) + \sqrt{\sigma^2_{i,t}} E_{i,t}.$$

- To estimate $\Theta^c_{i,t}$ and $\sigma^2_{i,t}$, we assume that they evolve according to:

$$\begin{aligned}
\Theta^c_{i,t} &= \Theta^c_{i,t-1} + \sqrt{\frac{\left(1-\lambda_{i,t}\right) var\left(\Theta^c_{i,t|t-1}\right)}{\lambda_{i,t}}} u_{i,t}, \\
\sigma^2_{i,t} &= \kappa_{i,t}\sigma^2_{i,t-1} + \left(1-\kappa_{i,t}\right)\widehat{E}^2_{i,t}.
\end{aligned}$$

where $\lambda_{i,t}$ and $\kappa_{i,t}$ are the forgetting and decay factors, typically in the range of $(0.9, 1)$, and control how quickly discounting of past data occurs.

Out-of-sample performance: Compressed TVP-VAR

Variable	MSFE						ALPL					
Medium VAR	$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$	$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$
PAYEMS	**0.691*****	**0.547*****	**0.525*****	**0.605****	**0.704****	**0.805**	**0.346*****	**0.409*****	**0.374*****	0.155	-0.315	-0.593
CPIAUCSL	**0.912*****	**0.871*****	**0.881*****	**0.857*****	**0.814*****	**0.827*****	**0.254**	**0.317***	**0.286***	**0.229****	**0.283***	**0.275**
FEDFUNDS	**0.888**	**0.932**	**0.931**	**0.958**	0.952	1.022	**0.679*****	**0.557***	**0.490**	**0.315**	0.057	**0.360**
INDPRO	0.912***	0.929	0.940	0.973	0.989	0.974	-0.064	0.067	-0.225	-0.220	-0.239	-0.139
UNRATE	0.807***	0.663***	0.599**	0.561**	0.596**	0.636**	0.132***	0.279***	**0.478****	**0.970**	0.904	1.210
PPIFGS	0.977	**0.984**	**0.989**	**1.014**	0.999	1.019	**0.323**	**0.281**	**0.346**	**0.372**	**0.336**	**0.318**
GS10	1.023	1.022	1.048	1.027	1.020	1.039	0.008	**0.018**	-0.073	**0.040****	-0.046	**0.004**
Multivariate	**0.920*****	**0.894*****	**0.887*****	**0.891****	**0.892****	**0.922***	**1.661*****	**1.791*****	**1.786*****	**1.680*****	1.496***	1.240***
Large VAR	$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$	$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$
PAYEMS	**0.692*****	0.560***	0.573***	**0.627*****	**0.715****	**0.816***	**0.330*****	0.373***	0.243*	-0.007	-0.232	-0.924
CPIAUCSL	**0.941**	**0.879*****	**0.878*****	**0.841*****	**0.811*****	**0.792*****	**0.221**	**0.242*****	**0.186**	**0.257**	**0.265****	**0.252**
FEDFUNDS	**0.903***	**0.861***	**0.888**	**0.922**	**0.968**	1.010	**0.641*****	**0.729*****	**0.683*****	**0.437**	**0.164**	**0.210**
INDPRO	0.922***	0.939*	0.963	0.933	0.967	0.985	-0.003	0.056	-0.150	-0.315	-0.378	-0.260
UNRATE	0.819***	0.709***	0.670***	0.664***	0.699***	0.741***	0.120***	0.223***	**0.413****	0.732	0.407	-0.956
PPIFGS	**0.969**	**0.984**	**0.991**	**0.991**	**0.980**	0.972	**0.295***	**0.298**	**0.342**	**0.322**	**0.245**	**0.189**
GS10	1.036	1.038	**1.023**	1.035	1.003	1.012	0.015	-0.053	-0.003	-0.040	-0.028	-0.088
Multivariate	**0.930*****	**0.890*****	**0.887*****	**0.880*****	**0.891*****	**0.912****	**1.580*****	**1.718*****	**1.680*****	**1.435*****	1.167***	0.800
Huge VAR	$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$	$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$
PAYEMS	0.713***	0.604***	0.563***	0.669*	0.790	0.880	**0.337*****	0.378***	0.314***	-0.007	-0.572	-0.701
CPIAUCSL	0.972	**0.848*****	**0.876****	**0.853***	**0.835*****	**0.825*****	**0.149**	**0.353***	**0.352**	**0.294**	**0.199**	**0.276**
FEDFUNDS	**0.869***	**0.912**	**0.941**	0.970	1.008	1.102	**0.642****	**0.586***	**0.492**	**0.204**	0.146	0.186
INDPRO	0.937	0.992	0.960	0.989	1.014	1.006	0.000	-0.078	-0.265	-0.379	-0.340	-0.199
UNRATE	0.839**	0.673**	0.590**	0.551**	0.596**	0.638**	0.110***	0.233***	**0.426****	0.879	0.852	0.750
PPIFGS	0.985	1.021	**1.009**	**1.010**	**0.990**	1.024	**0.291***	**0.428**	**0.377**	**0.351**	**0.168**	**0.400**
GS10	1.033	1.044	**1.024**	1.034	1.022	1.033	-0.005	-0.025	-0.013	**0.052****	-0.002	-0.063
Multivariate	0.938***	**0.913****	**0.895****	**0.906**	0.922	0.950	**1.566*****	**1.698*****	**1.754*****	**1.539*****	1.235***	1.041***

Conclusions

- Apply Bayesian "compressed regression" methods to large VARs
- Method works by:
  - Compressing the VAR regressors through random projection
  - Averaging across different random projections
- BCVAR as an alternative to the existing dimension reduction and shrinkage methods for large VARs
- Apply BCVAR to forecast a 130-variable macro VARs
  - BCVAR forecasts are quite accurate, in many instances improving over BVAR and FAVAR
  - Computationally much faster than BVAR, but slower than FAVAR (based on PCA+OLS)
- Extension to time-varinyg parameters and volatilities is computationally very fast and leads to further improvements in forecast accuracy

Appendix

Random Projection vs. Principal Component Analysis

- Random Projection (RP) is a projection method similar to Principal Component Analysis (PCA)
  - High-dimensional data is projected onto a low-dimensional subspace using a random matrix, whose columns have unit length
  - Unlike PCA, "loadings" are not estimated from data, rather generated randomly ("Data Oblivious" method)
- Inexpensive in terms of time/space. Random projection can be generated without even seeing the data
- Theoretical results show that RP preserves volumes and affine distances, or the structure of data (e.g., clustering)
  - Johnson-Lindenstrauss (1984) lemma: Any n point subset of Euclidean space can be embedded in $k = O\left(\log n/\epsilon^2\right)$ dimensions without distorting the distances between any pair of points by more than a factor of $1 \pm \epsilon$, for any $0 < \epsilon < 1$

◂ Projection matrix

Relative MSFE ratios, Medium VAR

Variable	*DFM*	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*	*DFM*	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR$_c$*
			$h=1$					$h=2$		
PAYEMS	1.076	1.041	0.892	**0.799*** **	0.813***	0.915	0.871	**0.571*** **	0.688***	0.711***
CPIAUCSL	1.139	1.102	**0.925**	0.951	0.935**	1.080	1.115	0.964	0.940	**0.925** **
FEDFUNDS	2.262	2.043	2.727	1.027	**0.952**	1.463	1.691	2.437	**0.999**	1.002
INDPRO	0.866**	0.896*	**0.821** **	0.830***	0.894***	0.924	0.997	**0.835***	0.914*	0.934
UNRATE	0.866	**0.758*** **	0.764**	0.762***	0.798***	0.772	0.626**	**0.582** **	0.626***	0.662***
PPIFGS	0.995	0.996	0.968	**0.966**	0.982	1.050	1.050	1.060	1.030	**1.000**
GS10	1.139	**0.973**	1.095	1.007	0.997	1.040	1.032	1.082	**0.995**	1.002
			$h=3$					$h=6$		
PAYEMS	0.834	0.827	**0.530*** **	0.630***	0.649***	0.960	0.825	**0.687***	0.694**	0.703**
CPIAUCSL	1.089	1.083	0.993	0.969	**0.956**	1.031	0.976	0.983	**0.962**	0.967
FEDFUNDS	1.282	1.372	1.862	1.063	**1.030**	1.195	0.998	1.230	**0.976**	0.987
INDPRO	0.923	0.959	0.924	**0.917***	0.939	**0.959**	0.980	1.022	0.965	0.963
UNRATE	0.769	0.620**	**0.526** **	0.563***	0.603***	0.835	0.637*	**0.512***	0.535**	0.575**
PPIFGS	1.033	1.034	1.070	1.051	**1.026**	1.047	**1.021**	1.092	1.046	1.042
GS10	**1.027**	1.040	1.144	1.058	1.031	**1.014**	1.019	1.123	1.054	1.038
			$h=9$					$h=12$		
PAYEMS	0.996	0.850	0.798	**0.779***	0.798*	1.027	**0.885**	0.901	0.887	0.893
CPIAUCSL	1.004	0.958*	0.968	0.956	**0.943**	1.009	0.972	0.985	**0.966**	0.970
FEDFUNDS	1.160	0.926	1.065	**0.904**	0.925	1.137	**0.925**	1.135	0.960	0.990
INDPRO	0.965	**0.952**	1.022	0.953	0.969	0.970	**0.965**	0.992	0.965	0.975
UNRATE	0.897	0.645	0.569*	**0.562** **	0.601**	0.939	0.661*	0.640*	**0.601** **	0.636**
PPIFGS	1.020	**0.994**	1.075	1.051	1.031	1.029	**1.001**	1.105	1.062	1.046
GS10	1.016	**1.006**	1.047	1.023	1.016	1.010	**1.006**	1.054	1.031	1.019

◂ Results

Relative MSFE ratios, Huge VAR

Variable	*DFM*	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR*$_c$	*DFM*	*FAVAR*	*BVAR*	*BCVAR*	*BCVAR*$_c$
			$h=1$					$h=2$		
PAYEMS	0.796**	0.956	**0.702*** **	0.761***	0.779***	0.716*	0.679**	**0.447*** **	0.641***	0.661***
CPIAUCSL	0.946	0.948	**0.872* **	0.947	0.943	1.006	1.003	0.939	0.890**	**0.879** **
FEDFUNDS	2.027	1.701	1.984	**0.891**	0.927	1.796	1.472	2.187	**0.929**	0.961
INDPRO	0.839**	0.830***	**0.785*** **	0.856***	0.913***	0.852	0.844*	**0.786** **	0.924**	0.939
UNRATE	0.778**	**0.736*** **	0.883	0.781***	0.813***	0.609**	**0.558*** **	0.646**	0.644***	0.703***
PPIFGS	0.952	0.985	**0.947**	0.985	1.004	1.074	1.051	1.071	1.029	**1.016**
GS10	1.098	**0.995**	1.100	0.998	1.019	1.044	1.057	1.150	**1.024**	1.032
			$h=3$					$h=6$		
PAYEMS	0.719	0.616**	**0.412*** **	0.591***	0.591***	0.910	0.742*	**0.516** **	0.645**	0.662**
CPIAUCSL	0.980	0.984	0.978	**0.903**	0.923	0.959	0.910	1.019	0.916	**0.893**
FEDFUNDS	1.581	1.182	1.848	**0.961**	0.985	1.380	**0.957**	1.275	0.995	0.986
INDPRO	0.941	0.917	**0.865**	0.951	0.947	1.033	**0.952**	0.972	0.961	0.970
UNRATE	0.550**	**0.495*** **	0.540**	0.590***	0.643***	0.638	0.479**	**0.435** **	0.562***	0.601***
PPIFGS	1.096	**1.032**	1.089	1.040	1.043	1.111	**1.037**	1.130	1.058	1.055
GS10	1.070	1.104	1.212	1.058	**1.057**	1.070	1.047	1.176	1.053	**1.033**
			$h=9$					$h=12$		
PAYEMS	1.006	0.863	**0.649**	0.744*	0.748*	1.079	0.949	**0.788**	0.854	0.842
CPIAUCSL	0.939	0.883**	0.994	0.876	**0.871** **	0.950	0.904***	0.992	0.880*	**0.862** **
FEDFUNDS	1.293	**0.950**	1.066	0.960	0.973	1.281	1.024	1.141	1.037	**1.007**
INDPRO	1.037	0.992	1.042	**0.973**	0.990	0.988	**0.981**	1.056	0.990	0.985
UNRATE	0.775	0.512***	**0.448** **	0.577***	0.610***	0.872	0.555***	**0.497** **	0.603***	0.640***
PPIFGS	1.064	**1.003**	1.127	1.062	1.019	1.104	**1.023**	1.159	1.069	1.037
GS10	1.022	1.013	1.070	1.025	**1.003**	1.039	1.021	1.084	1.035	**1.016**

◂ Results