Forecasting With High Dimensional Panel VARs

Gary Koop[1] Dimitris Korobilis[2]

[1]University of Strathclyde [2]University of Glasgow

Overview

- This paper is about Bayesian forecasting using panel VARs (PVARs) with time-varying parameters (TVP-PVARs)

Overview

- This paper is about Bayesian forecasting using panel VARs (PVARs) with time-varying parameters (TVP-PVARs)
- Application: forecasting inflation rates for 19 eurozone countries

Overview

- This paper is about Bayesian forecasting using panel VARs (PVARs) with time-varying parameters (TVP-PVARs)
- Application: forecasting inflation rates for 19 eurozone countries
- Challenges to be overcome:

Overview

- This paper is about Bayesian forecasting using panel VARs (PVARs) with time-varying parameters (TVP-PVARs)
- Application: forecasting inflation rates for 19 eurozone countries
- Challenges to be overcome:
- 1. Over-parameterization: 133 dependent variables

Overview

- This paper is about Bayesian forecasting using panel VARs (PVARs) with time-varying parameters (TVP-PVARs)
- Application: forecasting inflation rates for 19 eurozone countries
- Challenges to be overcome:
- 1. Over-parameterization: 133 dependent variables
- Shrinkage priors designed for multi-country panels (spillovers and linkages)

Overview

- This paper is about Bayesian forecasting using panel VARs (PVARs) with time-varying parameters (TVP-PVARs)
- Application: forecasting inflation rates for 19 eurozone countries
- Challenges to be overcome:
- 1. Over-parameterization: 133 dependent variables
- Shrinkage priors designed for multi-country panels (spillovers and linkages)
- Dimension switching

Overview

- This paper is about Bayesian forecasting using panel VARs (PVARs) with time-varying parameters (TVP-PVARs)
- Application: forecasting inflation rates for 19 eurozone countries
- Challenges to be overcome:
- 1. Over-parameterization: 133 dependent variables
- Shrinkage priors designed for multi-country panels (spillovers and linkages)
- Dimension switching
- Dynamic Model Averaging (DMA)

Overview

- This paper is about Bayesian forecasting using panel VARs (PVARs) with time-varying parameters (TVP-PVARs)
- Application: forecasting inflation rates for 19 eurozone countries
- Challenges to be overcome:
- 1. Over-parameterization: 133 dependent variables
- Shrinkage priors designed for multi-country panels (spillovers and linkages)
- Dimension switching
- Dynamic Model Averaging (DMA)
- 2. Computation: MCMC infeasible

- This paper is about Bayesian forecasting using panel VARs (PVARs) with time-varying parameters (TVP-PVARs)
- Application: forecasting inflation rates for 19 eurozone countries
- Challenges to be overcome:
- 1. Over-parameterization: 133 dependent variables
- Shrinkage priors designed for multi-country panels (spillovers and linkages)
- Dimension switching
- Dynamic Model Averaging (DMA)
- 2. Computation: MCMC infeasible
- Forgetting factor Methods

An Introduction to Panel VARs

- y_{it} contains G dependent variables for country i ($i = 1, .., N$) at time t ($t = 1, .., T$)

An Introduction to Panel VARs

- y_{it} contains G dependent variables for country i ($i = 1, .., N$) at time t ($t = 1, .., T$)
- $Y_t = (y'_{1t}, .., y'_{Nt})'$ all variables for all countries

An Introduction to Panel VARs

- y_{it} contains G dependent variables for country i ($i = 1, .., N$) at time t ($t = 1, .., T$)
- $Y_t = (y'_{1t}, .., y'_{Nt})'$ all variables for all countries
- VAR for country i:

$$y_{it} = A_{1,i} Y_{t-1} + ... + A_{P,i} Y_{t-P} + u_{it}$$

An Introduction to Panel VARs

- y_{it} contains G dependent variables for country i ($i = 1, .., N$) at time t ($t = 1, .., T$)
- $Y_t = (y'_{1t}, .., y'_{Nt})'$ all variables for all countries
- VAR for country i:

$$y_{it} = A_{1,i} Y_{t-1} + ... + A_{P,i} Y_{t-P} + u_{it}$$

- Lots of parameters: $A_{p,i}$ are $G \times NG$ matrices for each lag $p = 1, ..., P$

An Introduction to Panel VARs

- y_{it} contains G dependent variables for country i $(i = 1, .., N)$ at time t $(t = 1, .., T)$
- $Y_t = (y'_{1t}, .., y'_{Nt})'$ all variables for all countries
- VAR for country i:

$$y_{it} = A_{1,i}Y_{t-1} + ... + A_{P,i}Y_{t-P} + u_{it}$$

- Lots of parameters: $A_{p,i}$ are $G \times NG$ matrices for each lag $p = 1, ..., P$
- ε_{it} have Σ_{ii} covariance matrices of dimension $G \times G$.

An Introduction to Panel VARs

- y_{it} contains G dependent variables for country i ($i = 1, .., N$) at time t ($t = 1, .., T$)
- $Y_t = \left(y'_{1t}, .., y'_{Nt}\right)'$ all variables for all countries
- VAR for country i:

$$y_{it} = A_{1,i}Y_{t-1} + ... + A_{P,i}Y_{t-P} + u_{it}$$

- Lots of parameters: $A_{p,i}$ are $G \times NG$ matrices for each lag $p = 1, ..., P$
- ε_{it} have Σ_{ii} covariance matrices of dimension $G \times G$.
- May have correlation between countries: $cov\left(\varepsilon_{it}, \varepsilon_{jt}\right) = \Sigma_{ij}$

An Introduction to Panel VARs

- y_{it} contains G dependent variables for country i ($i = 1, .., N$) at time t ($t = 1, .., T$)
- $Y_t = (y'_{1t}, .., y'_{Nt})'$ all variables for all countries
- VAR for country i:

$$y_{it} = A_{1,i} Y_{t-1} + ... + A_{P,i} Y_{t-P} + u_{it}$$

- Lots of parameters: $A_{p,i}$ are $G \times NG$ matrices for each lag $p = 1, ..., P$
- ε_{it} have Σ_{ii} covariance matrices of dimension $G \times G$.
- May have correlation between countries: $cov\left(\varepsilon_{it}, \varepsilon_{jt}\right) = \Sigma_{ij}$
- This is the unrestricted PVAR

An Introduction to Panel VARs

- y_{it} contains G dependent variables for country i $(i = 1, .., N)$ at time t $(t = 1, .., T)$
- $Y_t = (y'_{1t}, .., y'_{Nt})'$ all variables for all countries
- VAR for country i:

$$y_{it} = A_{1,i}Y_{t-1} + ... + A_{P,i}Y_{t-P} + u_{it}$$

- Lots of parameters: $A_{p,i}$ are $G \times NG$ matrices for each lag $p = 1, ..., P$
- ε_{it} have Σ_{ii} covariance matrices of dimension $G \times G$.
- May have correlation between countries: $cov(\varepsilon_{it}, \varepsilon_{jt}) = \Sigma_{ij}$
- This is the unrestricted PVAR
- Write compactly as:

$$Y_t = X'_t\alpha + \varepsilon_t$$

An Introduction to Panel VARs

- y_{it} contains G dependent variables for country i ($i = 1, .., N$) at time t ($t = 1, .., T$)
- $Y_t = (y'_{1t}, .., y'_{Nt})'$ all variables for all countries
- VAR for country i:

$$y_{it} = A_{1,i} Y_{t-1} + ... + A_{P,i} Y_{t-P} + u_{it}$$

- Lots of parameters: $A_{p,i}$ are $G \times NG$ matrices for each lag $p = 1, ..., P$
- ε_{it} have Σ_{ii} covariance matrices of dimension $G \times G$.
- May have correlation between countries: $cov(\varepsilon_{it}, \varepsilon_{jt}) = \Sigma_{ij}$
- This is the unrestricted PVAR
- Write compactly as:

$$Y_t = X'_t \alpha + \varepsilon_t$$

- $K = p \times (N \times G)^2$ VAR parameters and $\frac{N \times G \times (N \times G + 1)}{2}$ error covariance terms.

Preview of our Data Set

- $G = 19$ eurozone countries and $N = 7$ variables

Preview of our Data Set

- $G = 19$ eurozone countries and $N = 7$ variables
-

Variables	**Explanation**	Trans
HICP	Indices of Consumer Prices	$\Delta \ln$
IP	Industrial production index	$\Delta \ln$
UN	Harmonised unemployment rates (%)	*lev*
REER	Real Effective Exchange Rate	$\Delta \ln$
SURVEY1	Financial situation over the next 12 months	*lev*
SURVEY2	Economic situation over next 12 months	*lev*
SURVEY3	Price trends over the next 12 months	*lev*

Preview of our Data Set

- $G = 19$ eurozone countries and $N = 7$ variables
-

Variables	**Explanation**	Trans
HICP	Indices of Consumer Prices	$\Delta \ln$
IP	Industrial production index	$\Delta \ln$
UN	Harmonised unemployment rates (%)	*lev*
REER	Real Effective Exchange Rate	$\Delta \ln$
SURVEY1	Financial situation over the next 12 months	*lev*
SURVEY2	Economic situation over next 12 months	*lev*
SURVEY3	Price trends over the next 12 months	*lev*

- 133 variables leads to very large VAR

Preview of our Data Set

- $G = 19$ eurozone countries and $N = 7$ variables
-

Variables	Explanation	Trans
HICP	Indices of Consumer Prices	$\Delta \ln$
IP	Industrial production index	$\Delta \ln$
UN	Harmonised unemployment rates (%)	*lev*
REER	Real Effective Exchange Rate	$\Delta \ln$
SURVEY1	Financial situation over the next 12 months	*lev*
SURVEY2	Economic situation over next 12 months	*lev*
SURVEY3	Price trends over the next 12 months	*lev*

- 133 variables leads to very large VAR
- Data from 1999M1-2014M12 means fairly small sample

Achieving Parsimony Through Factor Structure for VAR Coefficients

- Canova and Ciccarelli (IER, 2009, CC09) suggest :

$$\begin{aligned} \alpha &= \Xi_1\theta_1 + \Xi_2\theta_2 + .. + \Xi_q\theta_q + e \\ &= \Xi\theta + e \end{aligned}$$

Achieving Parsimony Through Factor Structure for VAR Coefficients

- Canova and Ciccarelli (IER, 2009, CC09) suggest :

$$\begin{aligned} \alpha &= \Xi_1 \theta_1 + \Xi_2 \theta_2 + .. + \Xi_q \theta_q + e \\ &= \Xi \theta + e \end{aligned}$$

- $\Xi = (\Xi_1, .., \Xi_q)$ are known matrices and $\theta = \left(\theta_1', .., \theta_q'\right)'$ is an $R \times 1$ vector of unknown parameters with $R < K$

Achieving Parsimony Through Factor Structure for VAR Coefficients

- Canova and Ciccarelli (IER, 2009, CC09) suggest :

$$\begin{aligned} \alpha &= \Xi_1\theta_1 + \Xi_2\theta_2 + .. + \Xi_q\theta_q + e \\ &= \Xi\theta + e \end{aligned}$$

- $\Xi = (\Xi_1, .., \Xi_q)$ are known matrices and $\theta = \left(\theta_1', .., \theta_q'\right)'$ is an $R \times 1$ vector of unknown parameters with $R < K$
- e is $N\left(0, \Sigma \otimes V\right)$ where $V = \sigma^2 I$.

Achieving Parsimony Through Factor Structure for VAR Coefficients

- Canova and Ciccarelli (IER, 2009, CC09) suggest :

$$\begin{aligned} \alpha &= \Xi_1\theta_1 + \Xi_2\theta_2 + .. + \Xi_q\theta_q + e \\ &= \Xi\theta + e \end{aligned}$$

- $\Xi = (\Xi_1, .., \Xi_q)$ are known matrices and $\theta = \left(\theta_1', .., \theta_q'\right)'$ is an $R \times 1$ vector of unknown parameters with $R < K$
- e is $N(0, \Sigma \otimes V)$ where $V = \sigma^2 I$.
- Much more parsimonious

Achieving Parsimony Through Factor Structure for VAR Coefficients

- Canova and Ciccarelli (IER, 2009, CC09) suggest :

$$\begin{aligned} \alpha &= \Xi_1\theta_1 + \Xi_2\theta_2 + .. + \Xi_q\theta_q + e \\ &= \Xi\theta + e \end{aligned}$$

- $\Xi = (\Xi_1, .., \Xi_q)$ are known matrices and $\theta = \left(\theta_1', .., \theta_q'\right)'$ is an $R \times 1$ vector of unknown parameters with $R < K$
- e is $N(0, \Sigma \otimes V)$ where $V = \sigma^2 I$.
- Much more parsimonious
- VAR composed of common factor, a factor specific to each country and a factor specific to each variable

Achieving Parsimony Through Factor Structure for VAR Coefficients

- Canova and Ciccarelli (IER, 2009, CC09) suggest :

$$\begin{aligned} \alpha &= \Xi_1\theta_1 + \Xi_2\theta_2 + .. + \Xi_q\theta_q + e \\ &= \Xi\theta + e \end{aligned}$$

- $\Xi = (\Xi_1, .., \Xi_q)$ are known matrices and $\theta = \left(\theta_1', .., \theta_q'\right)'$ is an $R \times 1$ vector of unknown parameters with $R < K$
- e is $N(0, \Sigma \otimes V)$ where $V = \sigma^2 I$.
- Much more parsimonious
- VAR composed of common factor, a factor specific to each country and a factor specific to each variable
- Ξ_1 is $K \times 1$ vector of ones, θ_1 a scalar.

Achieving Parsimony Through Factor Structure for VAR Coefficients

- Canova and Ciccarelli (IER, 2009, CC09) suggest :

$$\begin{aligned} \alpha &= \Xi_1\theta_1 + \Xi_2\theta_2 + .. + \Xi_q\theta_q + e \\ &= \Xi\theta + e \end{aligned}$$

- $\Xi = (\Xi_1, .., \Xi_q)$ are known matrices and $\theta = \left(\theta_1', .., \theta_q'\right)'$ is an $R \times 1$ vector of unknown parameters with $R < K$
- e is $N\left(0, \Sigma \otimes V\right)$ where $V = \sigma^2 I$.
- Much more parsimonious
- VAR composed of common factor, a factor specific to each country and a factor specific to each variable
- Ξ_1 is $K \times 1$ vector of ones, θ_1 a scalar.
- Ξ_2 is $K \times N$ matrix containing zeros and ones defined so as to pick out coefficients for each country and θ_2 is an $N \times 1$ vector.

Achieving Parsimony Through Factor Structure for VAR Coefficients

- Canova and Ciccarelli (IER, 2009, CC09) suggest :

$$\begin{aligned}\alpha &= \Xi_1\theta_1 + \Xi_2\theta_2 + .. + \Xi_q\theta_q + e \\ &= \Xi\theta + e\end{aligned}$$

- $\Xi = (\Xi_1, .., \Xi_q)$ are known matrices and $\theta = \left(\theta_1', .., \theta_q'\right)'$ is an $R \times 1$ vector of unknown parameters with $R < K$
- e is $N(0, \Sigma \otimes V)$ where $V = \sigma^2 I$.
- Much more parsimonious
- VAR composed of common factor, a factor specific to each country and a factor specific to each variable
- Ξ_1 is $K \times 1$ vector of ones, θ_1 a scalar.
- Ξ_2 is $K \times N$ matrix containing zeros and ones defined so as to pick out coefficients for each country and θ_2 is an $N \times 1$ vector.
- Ξ_3 is $K \times G$ matrix containing zeros and ones defined so as to pick out coefficients for each variable and θ_3 is an $G \times 1$ vector.

Achieving Parsimony Through Factor Structure for VAR Coefficients

- Canova and Ciccarelli (IER, 2009, CC09) suggest :

$$\begin{aligned} \alpha &= \Xi_1\theta_1 + \Xi_2\theta_2 + .. + \Xi_q\theta_q + e \\ &= \Xi\theta + e \end{aligned}$$

- $\Xi = (\Xi_1, .., \Xi_q)$ are known matrices and $\theta = \left(\theta_1', .., \theta_q'\right)'$ is an $R \times 1$ vector of unknown parameters with $R < K$
- e is $N\left(0, \Sigma \otimes V\right)$ where $V = \sigma^2 I$.
- Much more parsimonious
- VAR composed of common factor, a factor specific to each country and a factor specific to each variable
- Ξ_1 is $K \times 1$ vector of ones, θ_1 a scalar.
- Ξ_2 is $K \times N$ matrix containing zeros and ones defined so as to pick out coefficients for each country and θ_2 is an $N \times 1$ vector.
- Ξ_3 is $K \times G$ matrix containing zeros and ones defined so as to pick out coefficients for each variable and θ_3 is an $G \times 1$ vector.
- e picks up remaining heterogeneity in coefficients.

Achieving Parsimony Through Factor Structure for VAR Coefficients

- We use CC09 but many other structures are possible

Achieving Parsimony Through Factor Structure for VAR Coefficients

- We use CC09 but many other structures are possible
- We also use country-specific VAR factor structure

Achieving Parsimony Through Factor Structure for VAR Coefficients

- We use CC09 but many other structures are possible
- We also use country-specific VAR factor structure
- For country i, Ξ defined so, that θ loads only on the G^2 coefficients that are on lags of country i variables, not country j

Achieving Parsimony Through Factor Structure for VAR Coefficients

- We use CC09 but many other structures are possible
- We also use country-specific VAR factor structure
- For country i, Ξ defined so, that θ loads only on the G^2 coefficients that are on lags of country i variables, not country j
- I.e. $e = 0$, the coefficients on country j variables are zero and PVAR breaks down into N individual VARs, one for each country

Achieving Parsimony Through Factor Structure for VAR Coefficients

- We use CC09 but many other structures are possible
- We also use country-specific VAR factor structure
- For country i, Ξ defined so, that θ loads only on the G^2 coefficients that are on lags of country i variables, not country j
- I.e. $e = 0$, the coefficients on country j variables are zero and PVAR breaks down into N individual VARs, one for each country
- Intuition: working with VARs one country at a time close to being adequate, but occasional inter-linkages captured through e.

Achieving Parsimony Through Factor Structure for VAR Coefficients

- We use CC09 but many other structures are possible
- We also use country-specific VAR factor structure
- For country i, Ξ defined so, that θ loads only on the G^2 coefficients that are on lags of country i variables, not country j
- I.e. $e = 0$, the coefficients on country j variables are zero and PVAR breaks down into N individual VARs, one for each country
- Intuition: working with VARs one country at a time close to being adequate, but occasional inter-linkages captured through e.
- Use DMA/DMS to choose between two factor structures in dynamic fashion

Achieving Parsimony Through Dimension Reduction

- Forecasting inflation is hard (often hard to beat univariate models)

Achieving Parsimony Through Dimension Reduction

- Forecasting inflation is hard (often hard to beat univariate models)
- Probably a PVAR with $G = 7$ is over-parameterized

Achieving Parsimony Through Dimension Reduction

- Forecasting inflation is hard (often hard to beat univariate models)
- Probably a PVAR with $G = 7$ is over-parameterized
- We want to let data decide if this is so

Achieving Parsimony Through Dimension Reduction

- Forecasting inflation is hard (often hard to beat univariate models)
- Probably a PVAR with $G = 7$ is over-parameterized
- We want to let data decide if this is so
- Do DMA or DMS over PVARs of different dimensions

Achieving Parsimony Through Dimension Reduction

- Forecasting inflation is hard (often hard to beat univariate models)
- Probably a PVAR with $G = 7$ is over-parameterized
- We want to let data decide if this is so
- Do DMA or DMS over PVARs of different dimensions
- G_C is set of core variables of interest (inflation, unemployment rate and industrial production)

Achieving Parsimony Through Dimension Reduction

- Forecasting inflation is hard (often hard to beat univariate models)
- Probably a PVAR with $G = 7$ is over-parameterized
- We want to let data decide if this is so
- Do DMA or DMS over PVARs of different dimensions
- G_C is set of core variables of interest (inflation, unemployment rate and industrial production)
- Work with PVARs of dimension G_C or larger

Achieving Parsimony Through Dimension Reduction

- Forecasting inflation is hard (often hard to beat univariate models)
- Probably a PVAR with $G = 7$ is over-parameterized
- We want to let data decide if this is so
- Do DMA or DMS over PVARs of different dimensions
- G_C is set of core variables of interest (inflation, unemployment rate and industrial production)
- Work with PVARs of dimension G_C or larger
- $G = 3, 4, .., 7$

Achieving Parsimony Through Dimension Reduction

- Forecasting inflation is hard (often hard to beat univariate models)
- Probably a PVAR with $G = 7$ is over-parameterized
- We want to let data decide if this is so
- Do DMA or DMS over PVARs of different dimensions
- G_C is set of core variables of interest (inflation, unemployment rate and industrial production)
- Work with PVARs of dimension G_C or larger
- $G = 3, 4, .., 7$
- PVARs of different dimensions receive different weight at different points in time

Moving from the PVAR to the TVP-PVAR

- TVP-PVAR is
$$Y_t = X_t' \alpha_t + u_t,$$
where $X_t = I \otimes \left(Y_{t-1}', ..., Y_{t-p}'\right)'$, and $u_t \sim N\left(0, \Sigma_t\right)$.

Moving from the PVAR to the TVP-PVAR

- TVP-PVAR is
$$Y_t = X_t' \alpha_t + u_t,$$
where $X_t = I \otimes \left(Y_{t-1}', ..., Y_{t-p}'\right)'$, and $u_t \sim N\left(0, \Sigma_t\right)$.
- Note: α_t and Σ_t time varying

Moving from the PVAR to the TVP-PVAR

- TVP-PVAR is
$$Y_t = X_t' \alpha_t + u_t,$$
where $X_t = I \otimes \left(Y_{t-1}', ..., Y_{t-p}'\right)'$, and $u_t \sim N\left(0, \Sigma_t\right)$.
- Note: α_t and Σ_t time varying
- TVP-VAR literature uses random walk evolution of α_t and multivariate stochastic volatility for Σ_t

Moving from the PVAR to the TVP-PVAR

- TVP-PVAR is
$$Y_t = X_t' \alpha_t + u_t,$$
where $X_t = I \otimes \left(Y_{t-1}', ..., Y_{t-p}'\right)'$, and $u_t \sim N(0, \Sigma_t)$.
- Note: α_t and Σ_t time varying
- TVP-VAR literature uses random walk evolution of α_t and multivariate stochastic volatility for Σ_t
- In our case, over-parameterized and does not take panel structure into account

A Factor Structure for TVP-PVAR Coefficients

- Use the hierarchical prior:

$$\begin{aligned}\alpha_t &= \Xi\theta_t + e_t \\ \theta_t &= \theta_{t-1} + w_t,\end{aligned}$$

A Factor Structure for TVP-PVAR Coefficients

- Use the hierarchical prior:

$$\begin{aligned} \alpha_t &= \Xi\theta_t + e_t \\ \theta_t &= \theta_{t-1} + w_t, \end{aligned}$$

- θ_t is $R \times 1$ with $R \lll K$

A Factor Structure for TVP-PVAR Coefficients

- Use the hierarchical prior:

$$\begin{aligned} \alpha_t &= \Xi\theta_t + e_t \\ \theta_t &= \theta_{t-1} + w_t, \end{aligned}$$

- θ_t is $R \times 1$ with $R \lll K$
- Ξ is defined as above

A Factor Structure for TVP-PVAR Coefficients

- Use the hierarchical prior:

$$\begin{aligned} \alpha_t &= \Xi\theta_t + e_t \\ \theta_t &= \theta_{t-1} + w_t, \end{aligned}$$

- θ_t is $R \times 1$ with $R \lll K$
- Ξ is defined as above
- $w_t \sim N(0, W_t)$

A Factor Structure for TVP-PVAR Coefficients

- Use the hierarchical prior:

$$\begin{aligned} \alpha_t &= \Xi\theta_t + e_t \\ \theta_t &= \theta_{t-1} + w_t, \end{aligned}$$

- θ_t is $R \times 1$ with $R \lll K$
- Ξ is defined as above
- $w_t \sim N(0, W_t)$
- CC09 used this model a single Ξ and $\Sigma_t = \Sigma$ (homoskedasticity) using MCMC methods (and set $\sigma^2 = 0$)

A Factor Structure for TVP-PVAR Coefficients

- Use the hierarchical prior:

$$\begin{aligned} \alpha_t &= \Xi\theta_t + e_t \\ \theta_t &= \theta_{t-1} + w_t, \end{aligned}$$

- θ_t is $R \times 1$ with $R \lll K$
- Ξ is defined as above
- $w_t \sim N(0, W_t)$
- CC09 used this model a single Ξ and $\Sigma_t = \Sigma$ (homoskedasticity) using MCMC methods (and set $\sigma^2 = 0$)
- Too computationally burdensome for forecasting exercise unless G and N are both small

A Hierarchical Prior for the Error Covariance Matrix

- But error covariance matrix is also very parameter rich. Can we shrink it as well using a Ξ?

A Hierarchical Prior for the Error Covariance Matrix

- But error covariance matrix is also very parameter rich. Can we shrink it as well using a Ξ?
- We use ideas from Primiceri (ReStud, 2005) to do so.

A Hierarchical Prior for the Error Covariance Matrix

- But error covariance matrix is also very parameter rich. Can we shrink it as well using a Ξ?
- We use ideas from Primiceri (ReStud, 2005) to do so.
- Decompose as $\Sigma_t = B_t^{-1} H_t \left(H_t B_t^{-1} \right)'$

A Hierarchical Prior for the Error Covariance Matrix

- But error covariance matrix is also very parameter rich. Can we shrink it as well using a Ξ?
- We use ideas from Primiceri (ReStud, 2005) to do so.
- Decompose as $\Sigma_t = B_t^{-1} H_t \left(H_t B_t^{-1} \right)'$
- B_t is lower triangular matrix with ones on the diagonal,

A Hierarchical Prior for the Error Covariance Matrix

- But error covariance matrix is also very parameter rich. Can we shrink it as well using a Ξ?
- We use ideas from Primiceri (ReStud, 2005) to do so.
- Decompose as $\Sigma_t = B_t^{-1} H_t \left(H_t B_t^{-1} \right)'$
- B_t is lower triangular matrix with ones on the diagonal,
- H_t is diagonal matrix

A Hierarchical Prior for the Error Covariance Matrix

- But error covariance matrix is also very parameter rich. Can we shrink it as well using a Ξ?
- We use ideas from Primiceri (ReStud, 2005) to do so.
- Decompose as $\Sigma_t = B_t^{-1} H_t \left(H_t B_t^{-1} \right)'$
- B_t is lower triangular matrix with ones on the diagonal,
- H_t is diagonal matrix
- Can write PVAR as

$$\begin{aligned} Y_t &= X_t' \alpha_t + B_t^{-1} H_t \varepsilon_t \\ Y_t &= X_t' \gamma_t + Z_t' \beta_t + H_t \varepsilon_t, \end{aligned}$$

A Hierarchical Prior for the Error Covariance Matrix

- Use same factor structures as before, but for both γ_t and β_t:

$$\begin{aligned} \begin{bmatrix} \gamma_t \\ \beta_t \end{bmatrix} &= \begin{bmatrix} \Xi^{\gamma} & 0 \\ 0 & \Xi^{\beta} \end{bmatrix} \theta_t + u_t \\ \theta_t &= \theta_{t-1} + v_t. \end{aligned}$$

A Hierarchical Prior for the Error Covariance Matrix

- Use same factor structures as before, but for both γ_t and β_t:

$$\begin{aligned} \begin{bmatrix} \gamma_t \\ \beta_t \end{bmatrix} &= \begin{bmatrix} \Xi^{\gamma} & 0 \\ 0 & \Xi^{\beta} \end{bmatrix} \theta_t + u_t \\ \theta_t &= \theta_{t-1} + v_t. \end{aligned}$$

- $u_t \sim N\left(0, H_t \otimes \left(\sigma^2 I\right)\right)$

A Hierarchical Prior for the Error Covariance Matrix

- Use same factor structures as before, but for both γ_t and β_t:

$$\begin{aligned} \begin{bmatrix} \gamma_t \\ \beta_t \end{bmatrix} &= \begin{bmatrix} \Xi^{\gamma} & 0 \\ 0 & \Xi^{\beta} \end{bmatrix} \theta_t + u_t \\ \theta_t &= \theta_{t-1} + v_t. \end{aligned}$$

- $u_t \sim N\left(0, H_t \otimes \left(\sigma^2 I\right)\right)$
- Note lower triangularity of B_t and block diagonality of Ξ means standard state methods can be used (see Primiceri, 2005)

A Hierarchical Prior for the Error Covariance Matrix

- Use same factor structures as before, but for both γ_t and β_t:

$$\begin{aligned}
\begin{bmatrix} \gamma_t \\ \beta_t \end{bmatrix} &= \begin{bmatrix} \Xi^{\gamma} & 0 \\ 0 & \Xi^{\beta} \end{bmatrix} \theta_t + u_t \\
\theta_t &= \theta_{t-1} + v_t.
\end{aligned}$$

- $u_t \sim N\left(0, H_t \otimes \left(\sigma^2 I\right)\right)$
- Note lower triangularity of B_t and block diagonality of Ξ means standard state methods can be used (see Primiceri, 2005)
- Let 1 subscripts mean CC09, and 2 subscripts mean country-specific VAR factor structure

A Hierarchical Prior for the Error Covariance Matrix

- Use same factor structures as before, but for both γ_t and β_t:

$$\begin{bmatrix} \gamma_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} \Xi^{\gamma} & 0 \\ 0 & \Xi^{\beta} \end{bmatrix} \theta_t + u_t$$
$$\theta_t = \theta_{t-1} + v_t.$$

- $u_t \sim N\left(0, H_t \otimes \left(\sigma^2 I\right)\right)$
- Note lower triangularity of B_t and block diagonality of Ξ means standard state methods can be used (see Primiceri, 2005)
- Let 1 subscripts mean CC09, and 2 subscripts mean country-specific VAR factor structure
- We consider four models which have: i) Ξ_1^{γ} and Ξ_1^{β}, ii) Ξ_1^{γ} and Ξ_2^{β}, iii) Ξ_2^{γ} and Ξ_1^{β} and iv) Ξ_2^{γ} and Ξ_2^{β}.

Estimating the TVP-PVAR

- Details in paper: here the ideas

Estimating the TVP-PVAR

- Details in paper: here the ideas
- If we knew Σ_t, W_t and σ^2 have Normal linear state space model

Estimating the TVP-PVAR

- Details in paper: here the ideas
- If we knew Σ_t, W_t and σ^2 have Normal linear state space model
- Standard state space methods (involving Kalman filter) provide predictive densities

Estimating the TVP-PVAR

- Details in paper: here the ideas
- If we knew Σ_t, W_t and σ^2 have Normal linear state space model
- Standard state space methods (involving Kalman filter) provide predictive densities
- Replace W_t with estimate

$$\widehat{W}_t = \left(\frac{1}{\lambda} - 1 \right) var\left(\theta_t | \mathcal{D}_{t-1} \right)$$

Estimating the TVP-PVAR

- Details in paper: here the ideas
- If we knew Σ_t, W_t and σ^2 have Normal linear state space model
- Standard state space methods (involving Kalman filter) provide predictive densities
- Replace W_t with estimate

$$\widehat{W}_t = \left(\frac{1}{\lambda} - 1\right) var\left(\theta_t | \mathcal{D}_{t-1}\right)$$

- $\mathcal{D}_{t-1}$ denotes data available through period $t-1$

Estimating the TVP-PVAR

- Details in paper: here the ideas
- If we knew Σ_t, W_t and σ^2 have Normal linear state space model
- Standard state space methods (involving Kalman filter) provide predictive densities
- Replace W_t with estimate

$$\widehat{W}_t = \left(\frac{1}{\lambda} - 1 \right) var \left(\theta_t | \mathcal{D}_{t-1} \right)$$

- $\mathcal{D}_{t-1}$ denotes data available through period $t-1$
- $0 < \lambda \leq 1$ is forgetting factor (more motivation below)

Estimating the TVP-PVAR

- Details in paper: here the ideas
- If we knew Σ_t, W_t and σ^2 have Normal linear state space model
- Standard state space methods (involving Kalman filter) provide predictive densities
- Replace W_t with estimate

$$\widehat{W}_t = \left(\frac{1}{\lambda} - 1 \right) var\left(\theta_t | \mathcal{D}_{t-1}\right)$$

- $\mathcal{D}_{t-1}$ denotes data available through period $t - 1$
- $0 < \lambda \leq 1$ is forgetting factor (more motivation below)
- $var\left(\theta_t | \mathcal{D}_{t-1}\right)$ available from Kalman filter iteration at time $t - 1$

Estimating the TVP-PVAR

- Details in paper: here the ideas
- If we knew Σ_t, W_t and σ^2 have Normal linear state space model
- Standard state space methods (involving Kalman filter) provide predictive densities
- Replace W_t with estimate

$$\widehat{W}_t = \left(\frac{1}{\lambda} - 1\right) var\left(\theta_t | \mathcal{D}_{t-1}\right)$$

- $\mathcal{D}_{t-1}$ denotes data available through period $t-1$
- $0 < \lambda \leq 1$ is forgetting factor (more motivation below)
- $var\left(\theta_t | \mathcal{D}_{t-1}\right)$ available from Kalman filter iteration at time $t-1$
- Σ_t estimates using exponentially weighted moving average (EWMA) with decay factor κ

Estimating the TVP-PVAR

- Details in paper: here the ideas
- If we knew Σ_t, W_t and σ^2 have Normal linear state space model
- Standard state space methods (involving Kalman filter) provide predictive densities
- Replace W_t with estimate

$$\widehat{W}_t = \left(\frac{1}{\lambda} - 1\right) var\left(\theta_t | \mathcal{D}_{t-1}\right)$$

- $\mathcal{D}_{t-1}$ denotes data available through period $t-1$
- $0 < \lambda \leq 1$ is forgetting factor (more motivation below)
- $var\left(\theta_t | \mathcal{D}_{t-1}\right)$ available from Kalman filter iteration at time $t-1$
- Σ_t estimates using exponentially weighted moving average (EWMA) with decay factor κ
- σ^2 consider grid of values

Dynamic Treatment of Model Uncertainty

- We have now defined many different models and methods for estimating/forecasting any one of them

Dynamic Treatment of Model Uncertainty

- We have now defined many different models and methods for estimating/forecasting any one of them
- Models differ in choice of:

Dynamic Treatment of Model Uncertainty

- We have now defined many different models and methods for estimating/forecasting any one of them
- Models differ in choice of:
- Dimension of the TVP-PVAR: $G = 3, 4, 5, 6, 7$

Dynamic Treatment of Model Uncertainty

- We have now defined many different models and methods for estimating/forecasting any one of them
- Models differ in choice of:
- Dimension of the TVP-PVAR: $G = 3, 4, 5, 6, 7$
- Structure for Ξ: (i) Ξ_1^γ and Ξ_1^β, ii) Ξ_1^γ and Ξ_2^β, iii) Ξ_2^γ and Ξ_1^β and iv) Ξ_2^γ and Ξ_2^β

Dynamic Treatment of Model Uncertainty

- We have now defined many different models and methods for estimating/forecasting any one of them
- Models differ in choice of:
- Dimension of the TVP-PVAR: $G = 3, 4, 5, 6, 7$
- Structure for Ξ: (i) Ξ_1^γ and Ξ_1^β, ii) Ξ_1^γ and Ξ_2^β, iii) Ξ_2^γ and Ξ_1^β and iv) Ξ_2^γ and Ξ_2^β
- Forgetting factor, λ (gradual change in coefficients: $\lambda = 0.99$ and no change: $\lambda = 1$ thus including PVAR)

Dynamic Treatment of Model Uncertainty

- We have now defined many different models and methods for estimating/forecasting any one of them
- Models differ in choice of:
- Dimension of the TVP-PVAR: $G = 3, 4, 5, 6, 7$
- Structure for Ξ: (i) Ξ_1^γ and Ξ_1^β, ii) Ξ_1^γ and Ξ_2^β, iii) Ξ_2^γ and Ξ_1^β and iv) Ξ_2^γ and Ξ_2^β
- Forgetting factor, λ (gradual change in coefficients: $\lambda = 0.99$ and no change: $\lambda = 1$ thus including PVAR)
- Decay factor, $\kappa = \{0.94, 0.96, 1\}$ with $\kappa = 1$ meaning homoskestacity

Dynamic Treatment of Model Uncertainty

- We have now defined many different models and methods for estimating/forecasting any one of them
- Models differ in choice of:
- Dimension of the TVP-PVAR: $G = 3, 4, 5, 6, 7$
- Structure for Ξ: (i) Ξ_1^γ and Ξ_1^β, ii) Ξ_1^γ and Ξ_2^β, iii) Ξ_2^γ and Ξ_1^β and iv) Ξ_2^γ and Ξ_2^β
- Forgetting factor, λ (gradual change in coefficients: $\lambda = 0.99$ and no change: $\lambda = 1$ thus including PVAR)
- Decay factor, $\kappa = \{0.94, 0.96, 1\}$ with $\kappa = 1$ meaning homoskestacity
- Grid for σ^2 (20 values over a wide interval)

Dynamic Treatment of Model Uncertainty

- We have now defined many different models and methods for estimating/forecasting any one of them
- Models differ in choice of:
- Dimension of the TVP-PVAR: $G = 3, 4, 5, 6, 7$
- Structure for Ξ: (i) Ξ_1^γ and Ξ_1^β, ii) Ξ_1^γ and Ξ_2^β, iii) Ξ_2^γ and Ξ_1^β and iv) Ξ_2^γ and Ξ_2^β
- Forgetting factor, λ (gradual change in coefficients: $\lambda = 0.99$ and no change: $\lambda = 1$ thus including PVAR)
- Decay factor, $\kappa = \{0.94, 0.96, 1\}$ with $\kappa = 1$ meaning homoskestacity
- Grid for σ^2 (20 values over a wide interval)
- 2400 models.

Brief Introduction to DMA and DMS

- Let $M^{(i)}$ for $i = 1, ..., J$ be the set of models under consideration

Brief Introduction to DMA and DMS

- Let $M^{(i)}$ for $i = 1, ..., J$ be the set of models under consideration
- Attach probability to each model using data available at time $t-1$: $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$

Brief Introduction to DMA and DMS

- Let $M^{(i)}$ for $i = 1, ..., J$ be the set of models under consideration
- Attach probability to each model using data available at time $t - 1$: $p\left(M^{(i)} | \mathcal{D}_{t-1}\right)$
- DMA takes forecasts from all models and averages using these probabilities

Brief Introduction to DMA and DMS

- Let $M^{(i)}$ for $i = 1, ..., J$ be the set of models under consideration
- Attach probability to each model using data available at time $t-1$: $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- DMA takes forecasts from all models and averages using these probabilities
- DMS forecasts at $t-1$ using model with highest $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$

Brief Introduction to DMA and DMS

- Let $M^{(i)}$ for $i = 1, ..., J$ be the set of models under consideration
- Attach probability to each model using data available at time $t-1$: $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- DMA takes forecasts from all models and averages using these probabilities
- DMS forecasts at $t-1$ using model with highest $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- Computation is problem: order of 2^{TJ} possible combinations (models and time periods)

Brief Introduction to DMA and DMS

- Let $M^{(i)}$ for $i = 1, \ldots, J$ be the set of models under consideration
- Attach probability to each model using data available at time $t-1$: $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- DMA takes forecasts from all models and averages using these probabilities
- DMS forecasts at $t-1$ using model with highest $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- Computation is problem: order of 2^{TJ} possible combinations (models and time periods)
- Raftery, Karny and Ettler (Tech, 2010) pioneered field through forgetting factor methods

Brief Introduction to DMA and DMS

- Let $M^{(i)}$ for $i = 1, ..., J$ be the set of models under consideration
- Attach probability to each model using data available at time $t-1$: $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- DMA takes forecasts from all models and averages using these probabilities
- DMS forecasts at $t-1$ using model with highest $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- Computation is problem: order of 2^{TJ} possible combinations (models and time periods)
- Raftery, Karny and Ettler (Tech, 2010) pioneered field through forgetting factor methods
- $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$ obtained in a fast, recursive manner, in the spirit of Kalman filtering

Brief Introduction to DMA and DMS

- Let $M^{(i)}$ for $i = 1, ..., J$ be the set of models under consideration
- Attach probability to each model using data available at time $t-1$: $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- DMA takes forecasts from all models and averages using these probabilities
- DMS forecasts at $t-1$ using model with highest $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- Computation is problem: order of 2^{TJ} possible combinations (models and time periods)
- Raftery, Karny and Ettler (Tech, 2010) pioneered field through forgetting factor methods
- $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$ obtained in a fast, recursive manner, in the spirit of Kalman filtering
- Selects models which have forecast well in the recent past

$$p\left(M^{(i)}|\mathcal{D}_{t-1}\right) \propto \prod_{i=1}^{t-1}\left[p\left(Y_t^C|M^{(i)}, \mathcal{D}_{t-1}\right)\right]^{\mu^i}$$

Brief Introduction to DMA and DMS

- Let $M^{(i)}$ for $i = 1, ..., J$ be the set of models under consideration
- Attach probability to each model using data available at time $t-1$: $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- DMA takes forecasts from all models and averages using these probabilities
- DMS forecasts at $t-1$ using model with highest $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- Computation is problem: order of 2^{TJ} possible combinations (models and time periods)
- Raftery, Karny and Ettler (Tech, 2010) pioneered field through forgetting factor methods
- $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$ obtained in a fast, recursive manner, in the spirit of Kalman filtering
- Selects models which have forecast well in the recent past

$$p\left(M^{(i)}|\mathcal{D}_{t-1}\right) \propto \prod_{i=1}^{t-1}\left[p\left(Y_t^C|M^{(i)},\mathcal{D}_{t-1}\right)\right]^{\mu^i}$$

- μ is forgetting factor with similar interpretation to λ

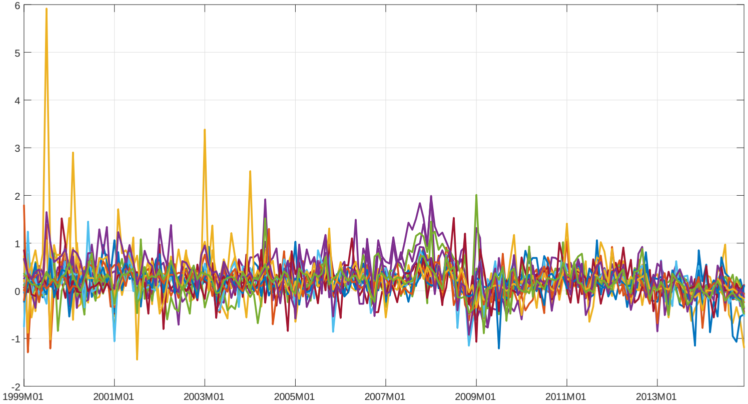
Brief Introduction to DMA and DMS

- Let $M^{(i)}$ for $i = 1, ..., J$ be the set of models under consideration
- Attach probability to each model using data available at time $t-1$: $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- DMA takes forecasts from all models and averages using these probabilities
- DMS forecasts at $t-1$ using model with highest $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$
- Computation is problem: order of 2^{TJ} possible combinations (models and time periods)
- Raftery, Karny and Ettler (Tech, 2010) pioneered field through forgetting factor methods
- $p\left(M^{(i)}|\mathcal{D}_{t-1}\right)$ obtained in a fast, recursive manner, in the spirit of Kalman filtering
- Selects models which have forecast well in the recent past

$$p\left(M^{(i)}|\mathcal{D}_{t-1}\right) \propto \prod_{i=1}^{t-1}\left[p\left(Y_t^C|M^{(i)},\mathcal{D}_{t-1}\right)\right]^{\mu^i}$$

- μ is forgetting factor with similar interpretation to λ
- Y_t^C are common to all models (inflation, the unemployment rate and industrial production)

Euro Area Inflation



Inflation in Eurozone Countries

Estimation Using TVP-PVAR-DMA

- We do DMA over the key specification choices (Ξ and G) and DMS so as to "estimate" factors which are similar to parameters (λ, κ and σ^2)

Estimation Using TVP-PVAR-DMA

- We do DMA over the key specification choices (Ξ and G) and DMS so as to "estimate" factors which are similar to parameters (λ, κ and σ^2)
- Call our approach TVP-PVAR-DMA

- We do DMA over the key specification choices (Ξ and G) and DMS so as to "estimate" factors which are similar to parameters (λ, κ and σ^2)
- Call our approach TVP-PVAR-DMA
- Following figures show:

Estimation Using TVP-PVAR-DMA

- We do DMA over the key specification choices (Ξ and G) and DMS so as to "estimate" factors which are similar to parameters (λ, κ and σ^2)
- Call our approach TVP-PVAR-DMA
- Following figures show:
- Which choices for Ξ are supported by the data (CC09 or our country-specific VAR factor structure)

Estimation Using TVP-PVAR-DMA

- We do DMA over the key specification choices (Ξ and G) and DMS so as to “estimate” factors which are similar to parameters (λ, κ and σ^2)
- Call our approach TVP-PVAR-DMA
- Following figures show:
- Which choices for Ξ are supported by the data (CC09 or our country-specific VAR factor structure)
- Which dimension is chosen

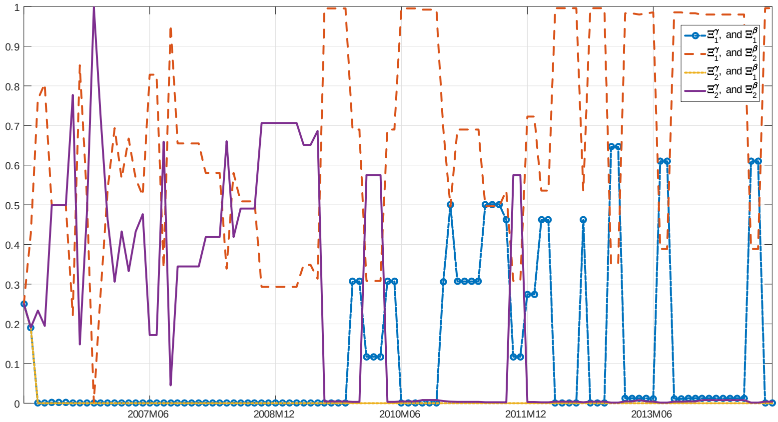
- We do DMA over the key specification choices (Ξ and G) and DMS so as to “estimate” factors which are similar to parameters (λ, κ and σ^2)
- Call our approach TVP-PVAR-DMA
- Following figures show:
- Which choices for Ξ are supported by the data (CC09 or our country-specific VAR factor structure)
- Which dimension is chosen
- Much time variation in each.

Estimation Using TVP-PVAR-DMA

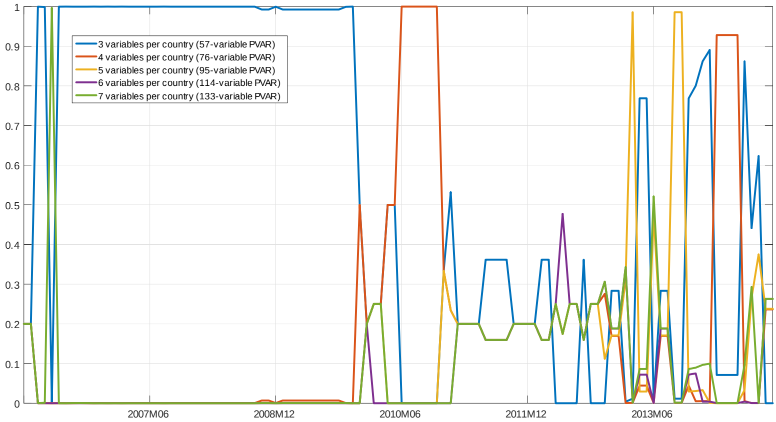
- We do DMA over the key specification choices (Ξ and G) and DMS so as to "estimate" factors which are similar to parameters (λ, κ and σ^2)
- Call our approach TVP-PVAR-DMA
- Following figures show:
- Which choices for Ξ are supported by the data (CC09 or our country-specific VAR factor structure)
- Which dimension is chosen
- Much time variation in each.
- CC09 which imposes interlinkages between countries supported post-eurozone crisis, but not as strongly before.

Estimation Using TVP-PVAR-DMA

- We do DMA over the key specification choices (Ξ and G) and DMS so as to "estimate" factors which are similar to parameters (λ, κ and σ^2)
- Call our approach TVP-PVAR-DMA
- Following figures show:
- Which choices for Ξ are supported by the data (CC09 or our country-specific VAR factor structure)
- Which dimension is chosen
- Much time variation in each.
- CC09 which imposes interlinkages between countries supported post-eurozone crisis, but not as strongly before.
- Co-movements increased with crisis?



DMA Probabilities Attached to Different Combinations for Ξ



DMA Probabilities Attached to Different PVAR Dimensions

Forecasting using TVP-PVAR-DMA

- Compare TVP-PVAR-DMA to:

Forecasting using TVP-PVAR-DMA

- Compare TVP-PVAR-DMA to:
- Individual country TVP-PVARs (19-TVP-VAR)

Forecasting using TVP-PVAR-DMA

- Compare TVP-PVAR-DMA to:
- Individual country TVP-PVARs (19-TVP-VAR)
- A Large TVP-VAR (LTVP-VAR) using the methods of Koop and Korobilis (2013)

Forecasting using TVP-PVAR-DMA

- Compare TVP-PVAR-DMA to:
- Individual country TVP-PVARs (19-TVP-VAR)
- A Large TVP-VAR (LTVP-VAR) using the methods of Koop and Korobilis (2013)
- A dynamic factor model (DFM) with time variation in parameters

Forecasting using TVP-PVAR-DMA

- Compare TVP-PVAR-DMA to:
- Individual country TVP-PVARs (19-TVP-VAR)
- A Large TVP-VAR (LTVP-VAR) using the methods of Koop and Korobilis (2013)
- A dynamic factor model (DFM) with time variation in parameters
- The TVP-PVAR without DMA (TVP-PVAR)

Forecasting using TVP-PVAR-DMA

- Compare TVP-PVAR-DMA to:
- Individual country TVP-PVARs (19-TVP-VAR)
- A Large TVP-VAR (LTVP-VAR) using the methods of Koop and Korobilis (2013)
- A dynamic factor model (DFM) with time variation in parameters
- The TVP-PVAR without DMA (TVP-PVAR)
- The TVP-PVAR extended to place a factor structure on error covariances (TVP-PVAR-X)

Forecasting using TVP-PVAR-DMA

- Compare TVP-PVAR-DMA to:
- Individual country TVP-PVARs (19-TVP-VAR)
- A Large TVP-VAR (LTVP-VAR) using the methods of Koop and Korobilis (2013)
- A dynamic factor model (DFM) with time variation in parameters
- The TVP-PVAR without DMA (TVP-PVAR)
- The TVP-PVAR extended to place a factor structure on error covariances (TVP-PVAR-X)
- Exact details in paper but made as comparable as possible

Forecasting using TVP-PVAR-DMA

- Forecast evaluation for $h = 1, 3, 6, 12$ for period 2006M1-2014M12

Forecasting using TVP-PVAR-DMA

- Forecast evaluation for $h = 1, 3, 6, 12$ for period 2006M1-2014M12
- Large tables with MSFEs and averages log predictive likelihoods for 19 countries are in paper

Forecasting using TVP-PVAR-DMA

- Forecast evaluation for $h = 1, 3, 6, 12$ for period 2006M1-2014M12
- Large tables with MSFEs and averages log predictive likelihoods for 19 countries are in paper
- Key points in tables:

Forecasting using TVP-PVAR-DMA

- Forecast evaluation for $h = 1, 3, 6, 12$ for period 2006M1-2014M12
- Large tables with MSFEs and averages log predictive likelihoods for 19 countries are in paper
- Key points in tables:
- TVP-PVAR-DMA forecasts best most often (e.g. 29 of 76 comparisons has lowest MSFE)

Forecasting using TVP-PVAR-DMA

- Forecast evaluation for $h = 1, 3, 6, 12$ for period 2006M1-2014M12
- Large tables with MSFEs and averages log predictive likelihoods for 19 countries are in paper
- Key points in tables:
- TVP-PVAR-DMA forecasts best most often (e.g. 29 of 76 comparisons has lowest MSFE)
- But even where it is not best, it is never far from the best-forecasting model

Forecasting using TVP-PVAR-DMA

- Forecast evaluation for $h = 1, 3, 6, 12$ for period 2006M1-2014M12
- Large tables with MSFEs and averages log predictive likelihoods for 19 countries are in paper
- Key points in tables:
- TVP-PVAR-DMA forecasts best most often (e.g. 29 of 76 comparisons has lowest MSFE)
- But even where it is not best, it is never far from the best-forecasting model
- 19-TVP-PVARs is second best, but occasionally forecasts very poorly (Ireland)

Forecasting using TVP-PVAR-DMA

- Forecast evaluation for $h = 1, 3, 6, 12$ for period 2006M1-2014M12
- Large tables with MSFEs and averages log predictive likelihoods for 19 countries are in paper
- Key points in tables:
- TVP-PVAR-DMA forecasts best most often (e.g. 29 of 76 comparisons has lowest MSFE)
- But even where it is not best, it is never far from the best-forecasting model
- 19-TVP-PVARs is second best, but occasionally forecasts very poorly (Ireland)
- DFM often forecast well, but predictive likelihoods for Greece and Latvia much worse than TVP-PVAR-DMA

Forecasting using TVP-PVAR-DMA

- Forecast evaluation for $h = 1, 3, 6, 12$ for period 2006M1-2014M12
- Large tables with MSFEs and averages log predictive likelihoods for 19 countries are in paper
- Key points in tables:
- TVP-PVAR-DMA forecasts best most often (e.g. 29 of 76 comparisons has lowest MSFE)
- But even where it is not best, it is never far from the best-forecasting model
- 19-TVP-PVARs is second best, but occasionally forecasts very poorly (Ireland)
- DFM often forecast well, but predictive likelihoods for Greece and Latvia much worse than TVP-PVAR-DMA
- The Large TVP-PVAR (ignoring panel structure of problem) forecasts relatively poorly

Forecasting using TVP-PVAR-DMA

- Forecast evaluation for $h = 1, 3, 6, 12$ for period 2006M1-2014M12
- Large tables with MSFEs and averages log predictive likelihoods for 19 countries are in paper
- Key points in tables:
- TVP-PVAR-DMA forecasts best most often (e.g. 29 of 76 comparisons has lowest MSFE)
- But even where it is not best, it is never far from the best-forecasting model
- 19-TVP-PVARs is second best, but occasionally forecasts very poorly (Ireland)
- DFM often forecast well, but predictive likelihoods for Greece and Latvia much worse than TVP-PVAR-DMA
- The Large TVP-PVAR (ignoring panel structure of problem) forecasts relatively poorly
- TVP-PVAR and TVP-PVAR-X are okay, but addition of DMA leads to clear improvements

When Are Forecast Improvements of TVP-PVAR-DMA Achieved?

- Previous results are for average forecast performance, one country at a time

When Are Forecast Improvements of TVP-PVAR-DMA Achieved?

- Previous results are for average forecast performance, one country at a time
- Next figure plots cumulative sums of log predictive likelihoods

When Are Forecast Improvements of TVP-PVAR-DMA Achieved?

- Previous results are for average forecast performance, one country at a time
- Next figure plots cumulative sums of log predictive likelihoods
- Sum also taken across countries

When Are Forecast Improvements of TVP-PVAR-DMA Achieved?

- Previous results are for average forecast performance, one country at a time
- Next figure plots cumulative sums of log predictive likelihoods
- Sum also taken across countries
- An overall measure of forecast performance

When Are Forecast Improvements of TVP-PVAR-DMA Achieved?

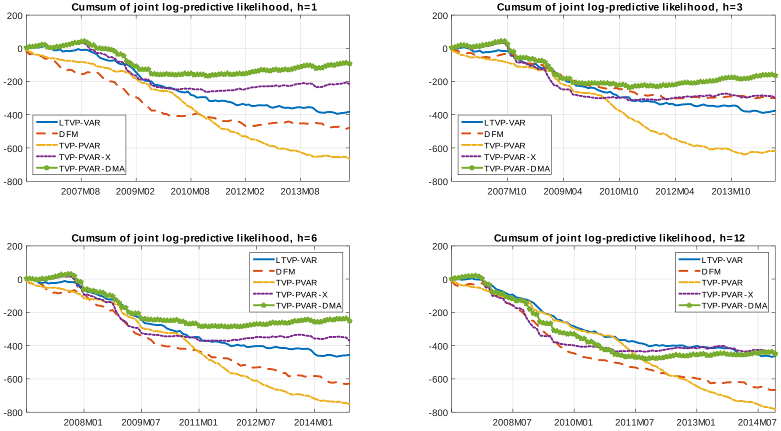
- Previous results are for average forecast performance, one country at a time
- Next figure plots cumulative sums of log predictive likelihoods
- Sum also taken across countries
- An overall measure of forecast performance
- Eurozone crisis usually dated from end of 2009

When Are Forecast Improvements of TVP-PVAR-DMA Achieved?

- Previous results are for average forecast performance, one country at a time
- Next figure plots cumulative sums of log predictive likelihoods
- Sum also taken across countries
- An overall measure of forecast performance
- Eurozone crisis usually dated from end of 2009
- This is time when TVP-PVAR-DMA become apparent

When Are Forecast Improvements of TVP-PVAR-DMA Achieved?

- Previous results are for average forecast performance, one country at a time
- Next figure plots cumulative sums of log predictive likelihoods
- Sum also taken across countries
- An overall measure of forecast performance
- Eurozone crisis usually dated from end of 2009
- This is time when TVP-PVAR-DMA become apparent
- In stable times, all approaches forecast roughly the same, but in unstable times DMA does better



Cumulative Sums of Log Predictive Likelihoods (across countries)

Summary

- Bayesian methods for forecasting with large TVP-PVARs developed which:

Summary

- Bayesian methods for forecasting with large TVP-PVARs developed which:
- Are computationally feasible (forgetting factors)

Summary

- Bayesian methods for forecasting with large TVP-PVARs developed which:
- Are computationally feasible (forgetting factors)
- Incorporate hierarchical priors for working with multi-country data

Summary

- Bayesian methods for forecasting with large TVP-PVARs developed which:
- Are computationally feasible (forgetting factors)
- Incorporate hierarchical priors for working with multi-country data
- Allow for dynamic model averaging or selection over more parsimonious specifications

Summary

- Bayesian methods for forecasting with large TVP-PVARs developed which:
- Are computationally feasible (forgetting factors)
- Incorporate hierarchical priors for working with multi-country data
- Allow for dynamic model averaging or selection over more parsimonious specifications
- Thus: begin with a parameter rich model with a range of prior/specification/forgetting factor choices and let data decide which ones to attach more weight to

Summary

- Bayesian methods for forecasting with large TVP-PVARs developed which:
- Are computationally feasible (forgetting factors)
- Incorporate hierarchical priors for working with multi-country data
- Allow for dynamic model averaging or selection over more parsimonious specifications
- Thus: begin with a parameter rich model with a range of prior/specification/forgetting factor choices and let data decide which ones to attach more weight to
- Euro area inflation forecasting exercise shows the benefits of our approach.