Improving inference and forecasting in VAR models using cross-sectional information

Boris Blagov and Jan Prüser

EEA-ESEM Congress 23

31.08.2022

Introduction

- VAR models are frequently used by applied macroeconomists for forecasting and structural analysis
  - They have rich dynamics but also many parameters - every variable is regressed on past values of itself and past values of all the other variables
  - New indicators are also developed while macro time series remain rather short
  - This leads to imprecise estimates and poor out-of-sample performance
- Bayesian estimation has long and successful history to address this problem
  - Introduces additional information **prior** to taking model to the data which helps the estimation
  - Many priors out there that perform well (e.g., Doan, Litterman, and Sims, 1984, Litterman, 1986, Sims and Zha, 1998, Bańbura, Giannone, and Reichlin, 2010, Giannone, Lenza, and Primiceri, 2015, Koop, 2013, Carriero, Clark, and Marcellino, 2016, Korobilis and Pettenuzzo, 2019, Huber and Feldkircher, 2019, Chan, 2020 and Cross, Hou, and Poon, 2020)

Minnesota prior

Example: **one** equation from a VAR with 10 variables and 4 lags

$$y_t = b_{y1}y_{t-1} + \cdots + b_{y4}y_{1,t-4} + b_{x1}x_{t-1} + \cdots + b_{x4}x_{t-4} + \cdots + b_{z1}z_{t-1} + \cdots + b_{z4}z_{t-4} + \epsilon_t.$$

The equation has 40 parameters: $\beta = [b_{y1}, b_{y2}, \ldots, b_{z3}, b_{z4}]'$

- Assume a-priori most of them are **equal to zero** with certainty $\underline{V}$, namely that $p(\beta) \sim N(\underline{\beta}, \underline{V})$
$$\underline{\beta} = [1, 0, \ldots, 0, 0]'$$
- Intuition: akin to a t-test, if the data is informative enough (standard errors are small), the estimated coefficient is pulled away from zero, else the regressor has no impact.

This is called "**shrinking towards zero**" and is extremely effective for both forecasting and structural analysis!

Other shrinkage

- Shrinking towards zero works well for single country VARs but far from the only option!
- If there is data for N countries $\rightarrow$ Panel VAR (PVAR)

$$\beta^1 = [b^1_{y1}, b^1_{y2}, \ldots, b^1_{z3}, b^1_{z4}]'$$
$$\ldots$$
$$\beta^N = [b^N_{y1}, b^N_{y2}, \ldots, b^N_{z3}, b^N_{z4}]'$$

- Assuming identical within country relationships one can use for example Pesaran, Shin and Smith (1999) or relaxing the assumption a bit Jarocinski, 2010
- In Bayesian terms this would be **pooling towards a common mean**, e.g. $\beta = [\overline{b}_{y1}, \overline{b}_{y2}, \ldots, \overline{b}_{z3}, \overline{b}_{z4}]'$

Motivation

- Countries could be homogeneous or heterogeneous
  - Heterogeneous countries: individual BVARs with zero shrinkage
  - Homogenous countries: PVAR with common mean

The euro area:

- Common monetary policy, integrated financial markets
- Disaggregated labour markets, different fiscal rules and legal systems
- Free movement of goods and services
- Some countries share similar characteristics, others do not - clustering
- Clustering does not have to be across all variables for all countries

Contribution

- We propose a flexible pooling prior which combines shrinking towards zero with parameter pooling in a cross-sectional dimension
- It allows clustering of countries in arbitrary number of groups
- The prior allows pooling across specific variables only, i.e. only some country characteristics may be similar
- The degree of similarity across variables and/or countries is endogenously estimated
- Applications:
  - Forecasting of the euro area member states and the G-7
  - Structural analysis
  - Clustering analysis

Literature

- Zellner and Hong, 1989, Jarocinski, 2010, Canova and Ciccarelli, 2009 and Pesaran, Shin, and Smith, 1999 assume common dynamics
- Koop and Korobilis, 2015, Korobilis, 2016 cluster different group of countries
- Our prior is more flexible by allowing for homogeneous and heterogeneous dynamics simultaneously and for both pooling and shrinkage at the same time.

VAR

Let $\boldsymbol{y}_{nt}$ be a $G \times 1$ vector of endogenous variables for country n at time t. Each country VAR can be written as

$$\boldsymbol{y}_{nt} = \sum_{p=1}^{P} \gamma_{np} \boldsymbol{y}_{n,t-p} + \epsilon_{nt}, \qquad \epsilon_{nt} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_n), \tag{1}$$

After some linear algebra we can write all country VARs as one large VAR

$$\tilde{\boldsymbol{y}} = (\boldsymbol{I}_N \otimes \tilde{\boldsymbol{x}})\beta + \tilde{\boldsymbol{u}}, \qquad \tilde{\boldsymbol{u}} \sim N(\boldsymbol{0}, \tilde{\boldsymbol{\Sigma}} \otimes \boldsymbol{I}_T). \tag{2}$$

- β is a large vector where the blocks of country parameters are on top of each other
- $\tilde{\Sigma}$ has a block structure $\leftrightarrow$ **still estimating single country VARs**!

Our Prior

The prior is normal $\beta \sim N(\mathbf{0}, \boldsymbol{V}_\beta)$, provides shrinkage towards zero and pooling towards the other country pairs' VAR coefficients

$$\log p(\beta) \propto \sum_{i=1}^{N}\sum_{j=1}^{K}\frac{\overline{\gamma}_{nj}^{2}}{V_{nj}^{\text{Min}}} + \sum_{j=1}^{K}\sum_{i=1}^{N-1}\sum_{m=i+1}^{N}\frac{(\beta_{j+K(i-1)}-\beta_{j+K(m-1)})^2}{\tau_{i,m}^2\lambda_{i,m,j}^2}. \quad (3)$$

Our Prior

The prior is normal $\beta \sim N(\mathbf{0}, \boldsymbol{V}_\beta)$, provides shrinkage towards zero and pooling towards the other country pairs' VAR coefficients

$$\log p(\beta) \propto \sum_{i=1}^{N} \sum_{j=1}^{K} \frac{\overline{\gamma}_{nj}^2}{V_{nj}^{\text{Min}}} + \sum_{j=1}^{K} \sum_{i=1}^{N-1} \sum_{m=i+1}^{N} \frac{(\beta_{j+K(i-1)} - \beta_{j+K(m-1)})^2}{\tau_{i,m}^2 \lambda_{i,m,j}^2}. \quad (3)$$

Our Prior

The prior is normal $\beta \sim N(\mathbf{0}, \boldsymbol{V}_\beta)$, provides shrinkage towards zero and pooling towards the other country pairs' VAR coefficients

$$\log p(\beta) \propto \sum_{i=1}^{N} \sum_{j=1}^{K} \frac{\overline{\gamma}_{nj}^2}{V_{nj}^{\text{Min}}} + \sum_{j=1}^{K} \sum_{i=1}^{N-1} \sum_{m=i+1}^{N} \frac{(\beta_{j+K(i-1)} - \beta_{j+K(m-1)})^2}{\tau_{i,m}^2 \lambda_{i,m,j}^2}. \quad (3)$$

The pooling part

Abstracting from the sums for a bit

$$\frac{(\beta_{j+K(i-1)} - \beta_{j+K(m-1)})^2}{\tau_{i,m}^2 \lambda_{i,m,j}^2} \tag{4}$$

- $\beta_{j+K(i-1)} - \beta_{j+K(m-1)}$: coefficient j across two countries (i, m) is the same
- $\tau_{i,m}^2$ gives how certain we are in this prior assumption
  - $\tau_{i,m}^2 \to 0$: prior becomes really informative (penalty is strong)
  - $\tau_{i,m}^2 \to \infty$: prior plays no role (penalty is weak)
  - is country-pair specific, i.e. $\tau_{i,m}^2$ is the same for all j and pools coefficients of country i and country m together
- $\lambda_{i,m,j}^2$ intuition is the same, except it is coefficient j-specific and can thus either *prevent* or *allow* exceptions

The pooling part

Abstracting from the sums for a bit

$$\frac{(\beta_{j+K(i-1)} - \beta_{j+K(m-1)})^2}{\tau_{i,m}^2 \lambda_{i,m,j}^2} \quad (4)$$

- $\beta_{j+K(i-1)} - \beta_{j+K(m-1)}$: coefficient j across two countries (i, m) is the same
- $\tau_{i,m}^2$ gives how certain we are in this prior assumption
  - $\tau_{i,m}^2 \to 0$: prior becomes really informative (penalty is strong)
  - $\tau_{i,m}^2 \to \infty$: prior plays no role (penalty is weak)
  - is country-pair specific, i.e. $\tau_{i,m}^2$ is the same for all j and pools coefficients of country i and country m together
- $\lambda_{i,m,j}^2$ intuition is the same, except it is coefficient j-specific and can thus either *prevent* or *allow* exceptions

The pooling part

Abstracting from the sums for a bit

$$\frac{(\beta_{j+K(i-1)} - \beta_{j+K(m-1)})^2}{\tau_{i,m}^2 \lambda_{i,m,j}^2} \quad (4)$$

- $\beta_{j+K(i-1)} - \beta_{j+K(m-1)}$: coefficient j across two countries (i, m) is the same
- $\tau_{i,m}^2$ gives how certain we are in this prior assumption
  - $\tau_{i,m}^2 \to 0$: prior becomes really informative (penalty is strong)
  - $\tau_{i,m}^2 \to \infty$: prior plays no role (penalty is weak)
  - is country-pair specific, i.e. $\tau_{i,m}^2$ is the same for all j and pools coefficients of country i and country m together
- $\lambda_{i,m,j}^2$ intuition is the same, except it is coefficient j-specific and can thus either *prevent* or *allow* exceptions

The pooling part

Abstracting from the sums for a bit

$$\frac{(\beta_{j+K(i-1)} - \beta_{j+K(m-1)})^2}{\tau^2_{i,m}\lambda^2_{i,m,j}} \tag{4}$$

- $\beta_{j+K(i-1)} - \beta_{j+K(m-1)}$: coefficient j across two countries (i, m) is the same
- $\tau^2_{i,m}$ gives how certain we are in this prior assumption
  - $\tau^2_{i,m} \rightarrow 0$: prior becomes really informative (penalty is strong)
  - $\tau^2_{i,m} \rightarrow \infty$: prior plays no role (penalty is weak)
  - is country-pair specific, i.e. $\tau^2_{i,m}$ is the same for all j and pools coefficients of country i and country m together
- $\lambda^2_{i,m,j}$ intuition is the same, except it is coefficient j-specific and can thus either *prevent* or *allow* exceptions

Estimation of the hyperparameters

- τ and λ (and $\bar{\gamma}$ from the Minn prior as well) are called **hyperparameters** - they determine the shape of the prior on the VAR parameters
- How to choose the hyperparameters? Problems with hierarchical estimation:
  - Pairwise pooling introduces overshrinkage towards zero and skews the prior with higher N
  - Relationship between two countries should not be affected by a third country
- $\Rightarrow$ **Empirical Bayes** using a two-step procedure

The estimated hyperparameters **are a measure of similarity** across the countries

Simulation Setup

- Countries: $i \in \{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$
- A stable VAR with $Y = [y^i, \pi^i, r^i]$ with 2 lags is the data generating process (DGP):

$$Y_t^i = B_1^i Y_{t-1}^i + B_2^i Y_{t-2}^i + e_t^i, \qquad e_t^i \sim N(0, \Sigma^i) \tag{5}$$

Consider the following cases

1. The three countries are heterogeneous
2. Two of the three countries, $\mathcal{B}$ and $\mathcal{C}$, have identical dynamics
3. A specific variable shares identical dynamics across countries and other parameters are pairwise identical

		DGP 1			DGP 2			DGP 3		
		$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$	$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$	$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$
y_t equation	y_{t-1}	0.50	0.80	0.20	0.50	0.80	0.80	0.20	0.70	0.90
	π_{t-1}	0.10	-0.20	0.20	0.10	-0.20	-0.20	0.10	-0.20	0.20
	r_{t-1}	-0.40	-0.10	0.00	-0.40	-0.10	-0.10	0.00	0.00	0.00
	y_{t-2}	-0.25	0.00	-0.30	-0.25	0.00	0.00	-0.05	-0.10	-0.30
	π_{t-2}	-0.20	0.20	0.00	-0.20	0.00	0.00	-0.20	0.00	0.00
	r_{t-2}	-0.40	0.10	0.40	-0.40	0.10	0.10	0.00	0.00	0.00
π_t equation	y_{t-1}	-0.20	0.00	0.30	-0.20	0.00	0.00	-0.20	0.10	0.00
	π_{t-1}	0.90	0.40	0.00	0.90	0.40	0.40	0.30	0.00	0.30
	r_{t-1}	-0.10	0.10	0.30	0.10	0.10	0.10	0.00	0.00	0.00
	y_{t-2}	0.00	-0.40	0.30	0.00	-0.40	-0.40	0.00	0.20	0.00
	π_{t-2}	0.00	-0.30	0.50	0.00	-0.30	-0.30	-0.40	-0.30	0.10
	r_{t-2}	-0.20	0.50	-0.50	-0.20	0.50	0.50	0.00	0.00	0.00
r_t equation	y_{t-1}	0.20	0.00	0.50	0.20	0.00	0.00	0.00	0.00	0.00
	π_{t-1}	0.20	0.00	0.50	0.10	0.00	0.00	0.00	0.00	0.00
	r_{t-1}	0.75	0.25	-0.20	0.75	0.25	0.25	0.95	0.95	0.95
	y_{t-2}	0.00	0.30	-0.20	0.00	0.30	0.30	0.00	0.00	0.00
	π_{t-2}	-0.30	0.20	0.00	-0.30	0.20	0.20	0.00	0.00	0.00
	r_{t-2}	-0.25	-0.10	0.20	-0.24	-0.10	-0.10	-0.20	-0.20	-0.20

Table 1: Shading reflects identical parameters across pairs

Simulation Results: Hyperparameters

- Does the prior work? Do the estimated hyperparameters capture the degree of similarity?
- Remember, interplay between two hyperparameters
  1. $\lambda_{i,m,j}$ country-pair and parameter specific
  2. $\tau_{i,m}$ country-pair specific

  Let

$$\Lambda_{i,m,j} = \lambda_{i,m,j}\tau_{i,m}$$

  and

$$\Lambda^{i,m} = [\Lambda_{i,m,1}, \ldots, \Lambda_{i,m,j}, \ldots, \Lambda_{i,m,18}]'$$

- Low values of $\Lambda_{i,m,j}$ suggest similarity between the respective VAR parameters across i and m
- High values of $\Lambda_{i,m,j}$ point to the contrary

DGP2: Identical country pair

		DGP 2			Shrinkage pairs		
		$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$	$\Lambda^{(\mathcal{A},\mathcal{B})}$	$\Lambda^{(\mathcal{A},\mathcal{C})}$	$\Lambda^{(\mathcal{B},\mathcal{C})}$
y_t equation	y_{t-1}	0.50	0.80	0.80	0.14	0.14	0.02
	π_{t-1}	0.10	-0.20	-0.20	0.12	0.12	0.01
	r_{t-1}	-0.40	-0.10	-0.10	0.13	0.13	0.01
	y_{t-2}	-0.25	0.00	0.00	0.12	0.12	0.02
	π_{t-2}	-0.20	0.00	0.00	0.09	0.10	0.01
	r_{t-2}	-0.40	0.10	0.10	0.21	0.21	0.01
π_t equation	y_{t-1}	-0.20	0.00	0.00	0.07	0.07	0.03
	π_{t-1}	0.90	0.40	0.40	0.29	0.29	0.02
	r_{t-1}	0.10	0.10	0.10	0.02	0.02	0.02
	y_{t-2}	0.00	-0.40	-0.40	0.28	0.29	0.03
	π_{t-2}	0.00	-0.30	-0.30	0.16	0.15	0.02
	r_{t-2}	-0.20	0.50	0.50	0.33	0.34	0.02
r_t equation	y_{t-1}	0.20	0.00	0.00	0.11	0.10	0.03
	π_{t-1}	0.10	0.00	0.00	0.05	0.05	0.02
	r_{t-1}	0.75	0.25	0.25	0.30	0.30	0.02
	y_{t-2}	0.00	0.30	0.30	0.26	0.25	0.03
	π_{t-2}	-0.30	0.20	0.20	0.41	0.41	0.02
	r_{t-2}	-0.24	-0.10	-0.10	0.06	0.06	0.02

DGP1: Heterogeneous countries

		DGP 1			Shrinkage pairs		
		$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$	$\Lambda^{(\mathcal{A},\mathcal{B})}$	$\Lambda^{(\mathcal{A},\mathcal{C})}$	$\Lambda^{(\mathcal{B},\mathcal{C})}$
y_t equation	y_{t-1}	0.50	0.80	0.20	0.15	0.16	0.41
	π_{t-1}	0.10	-0.20	0.20	0.15	0.15	0.38
	r_{t-1}	-0.40	-0.10	0.00	0.12	0.16	0.04
	y_{t-2}	-0.25	0.00	-0.30	0.12	0.03	0.16
	π_{t-2}	-0.20	0.20	0.00	0.25	0.15	0.07
	r_{t-2}	-0.40	0.10	0.40	0.21	0.65	0.26
π_t equation	y_{t-1}	-0.20	0.00	0.30	0.06	0.15	0.08
	π_{t-1}	0.90	0.40	0.00	0.30	0.76	0.24
	r_{t-1}	-0.10	0.10	0.30	0.06	0.13	0.06
	y_{t-2}	0.00	-0.40	0.30	0.26	0.07	0.41
	π_{t-2}	0.00	-0.30	0.50	0.16	0.29	0.62
	r_{t-2}	-0.20	0.50	-0.50	0.31	0.11	0.55
r_t equation	y_{t-1}	0.20	0.00	0.50	0.12	0.06	0.20
	π_{t-1}	0.20	0.00	0.50	0.16	0.26	0.56
	r_{t-1}	0.75	0.25	-0.20	0.31	0.83	0.27
	y_{t-2}	0.00	0.30	-0.20	0.23	0.06	0.36
	π_{t-2}	-0.30	0.20	0.00	0.53	0.29	0.12
	r_{t-2}	-0.25	-0.10	0.20	0.07	0.26	0.15

DGP3: Identical variable across countries

		DGP 3			Shrinkage pairs		
		$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$	$\Lambda^{(\mathcal{A},\mathcal{B})}$	$\Lambda^{(\mathcal{A},\mathcal{C})}$	$\Lambda^{(\mathcal{B},\mathcal{C})}$
y_t equation	y_{t-1}	0.20	0.70	0.90	0.26	0.45	0.07
	π_{t-1}	0.10	-0.20	0.20	0.09	0.02	0.10
	r_{t-1}	0.00	0.00	0.00	0.02	0.02	0.02
	y_{t-2}	-0.05	-0.10	-0.30	0.03	0.09	0.07
	π_{t-2}	-0.20	0.00	0.00	0.05	0.05	0.01
	r_{t-2}	0.00	0.00	0.00	0.02	0.01	0.01
π_t equation	y_{t-1}	-0.20	0.10	0.00	0.14	0.06	0.05
	π_{t-1}	0.30	0.00	0.30	0.11	0.02	0.11
	r_{t-1}	0.00	0.00	0.00	0.02	0.02	0.02
	y_{t-2}	0.00	0.20	0.00	0.06	0.01	0.05
	π_{t-2}	-0.40	-0.30	0.10	0.04	0.25	0.17
	r_{t-2}	0.00	0.00	0.00	0.02	0.01	0.02
r_t equation	y_{t-1}	0.00	0.00	0.00	0.01	0.01	0.01
	π_{t-1}	0.00	0.00	0.00	0.01	0.01	0.01
	r_{t-1}	0.95	0.95	0.95	0.02	0.02	0.02
	y_{t-2}	0.00	0.00	0.00	0.01	0.01	0.01
	π_{t-2}	0.00	0.00	0.00	0.01	0.01	0.01
	r_{t-2}	-0.20	-0.20	-0.20	0.02	0.02	0.02

Data

- 10 countries: DE, FR, ES, IT, NL, BE, AT, PT, FI, GR
- 7 variables: real GDP (RGDP), the harmonised index of consumer prices (HICP), EURIBOR, Industrial production (IP), Unemployment (UNEMP), the Economic Sentiment Index (ESI) and the price of crude brent (OIL)

Note: Interest rates and oil prices are identical across countries

- Data are seasonally adjusted and converted to growth rates
- Data are quarterly and span from 2000Q1 to 2019Q4 (80 observations)

Forecasting setup

- Pseudo-out-of sample forecasting with expanding window for $t = 1, \dots, T, \dots, 79$
  1. Each model initially estimated until T starting with $T = 39 \equiv 2009Q3$
  2. Iterative forecasts over 8 forecast horizons ($T + 1 \equiv$ 2009Q4, ..., $T + 8 \equiv$2011Q3)
  3. T incremented by 1, repeat above steps $\forall T$
- Metrics for comparison are the Root Mean Squared Forecast Errors (RMSFEs)

$$RMSFE_{T+1} = \frac{1}{40}\sqrt{\sum_{T=39}^{79}(y^{data}_{T+1} - y^{fore}_{T+1})^2}$$

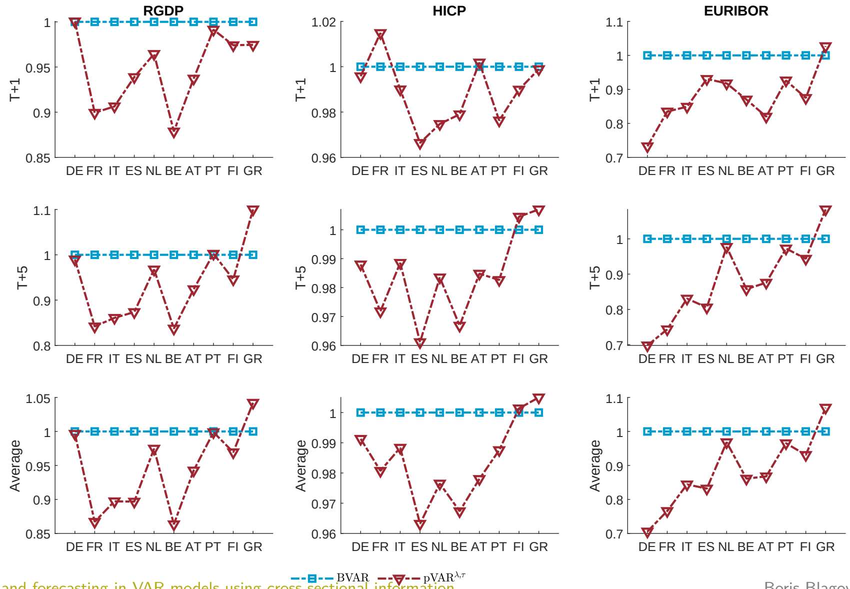
$$RMSFE_{T+5} = \frac{1}{36}\sqrt{\sum_{T=39}^{75}(y^{data}_{T+5} - y^{fore}_{T+5})^2}$$

$$RMSFE_{avg} = \sum_{i=1}^{8} RMSFE_{T+i}$$

- To compare models we take Minnesota BVAR as baseline, i.e.

$$RMSFE^{pVAR}_{T+1} / RMSFE^{Minn}_{T+1} <> 1$$

Forecasting Results I



Competing models

BVAR	Classic Bayesian VAR with Minnesota prior
BVARf	Bayesian VAR with flat prior
BVAR^{+}	BVAR as in Chan, Koop, and Yu (2021)
pVAR$^{\tau}$	Pooling across country pairs with Minnesota
pVAR$^{\lambda,\tau}$	Pooling across country and variable pairs with Minnesota
BPVAR	Classic Bayesian Panel VAR
PVARj	Bayesian Panel VAR as in Jarocinski (2010)

		DE	FR	IT	ES	NL	BE	AT	PT	FI	GR
RGDP	**BVAR (abs)**	0.60	0.35	0.37	0.34	0.42	0.32	0.54	0.62	0.78	1.34
	BVAR	**1.00**	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	BVAR$^+$	1.06	1.07	1.08	1.09	1.02	1.00	0.99	1.03	1.01	1.02
	BVARf	1.24	1.27	1.49	1.47	1.41	1.24	1.20	1.09	1.36	1.33
	pVAR$^{\lambda,\tau}$	1.00	0.90	**0.91**	**0.94**	0.96	0.88	0.94	0.99	**0.97**	0.97
	pVAR$^{\tau}$	1.01	**0.89**	0.91	1.07	**0.96**	**0.84**	**0.93**	**0.97**	0.98	1.00
	BPVAR	1.01	1.07	1.10	1.14	1.09	1.06	0.96	1.00	1.01	1.01
	PVARj	1.08	1.15	1.03	1.17	1.13	1.35	0.96	1.01	1.10	**0.97**

Table 2: **N=10, G=3 setup:** One step ahead root mean squared forecast errors for different models, relative to the BVAR forecast. The first row is the absolute error of the BVAR benchmark. Values in bold denote the lowest relative RMSFE for that country.

HICP		DE	FR	IT	ES	NL	BE	AT	PT	FI	GR
	BVAR (abs)	0.33	0.27	0.30	0.49	0.38	0.36	0.29	0.46	0.28	0.53
	BVAR	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	BVAR^{+}	**0.99**	**1.00**	1.02	1.02	1.02	1.01	1.00	1.01	1.08	1.00
	BVARf	1.14	1.06	1.13	1.19	1.15	1.16	1.14	1.15	1.32	1.08
	pVAR$^{\lambda,\tau}$	1.00	1.01	0.99	0.97	0.97	0.98	1.00	0.98	0.99	1.00
	pVAR$^{\tau}$	0.99	1.01	1.00	0.96	**0.97**	**0.97**	1.00	**0.97**	1.00	1.01
	BPVAR	1.04	1.03	**0.96**	**0.92**	0.97	1.01	**0.96**	0.98	**0.95**	**0.94**
	PVARj	1.08	1.10	1.07	1.00	0.99	1.03	1.02	1.03	1.07	1.05

Table 3: **N=10, G=3 setup:** One step ahead root mean squared forecast errors for different models, relative to the BVAR forecast. The first row is the absolute error of the BVAR benchmark. Values in bold denote the lowest relative RMSFE for that country.

EURIBOR

	DE	FR	IT	ES	NL	BE	AT	PT	FI	GR
BVAR (abs)	0.19	0.18	0.20	0.17	0.20	0.19	0.21	0.23	0.21	0.17
BVAR	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**1.00**
BVAR^{+}	1.01	0.97	0.93	1.04	0.98	0.94	0.94	0.94	0.97	1.16
BVARf	1.46	1.26	1.61	1.57	1.77	1.06	1.19	1.33	1.31	1.67
pVAR$^{\lambda,\tau}$	**0.73**	0.83	0.85	**0.93**	0.92	0.87	0.82	0.93	0.87	1.03
pVAR$^{\tau}$	0.75	**0.83**	0.86	0.98	0.92	**0.85**	0.80	0.91	**0.87**	1.07
BPVAR	0.96	0.93	**0.76**	0.97	**0.88**	1.04	0.86	**0.76**	0.89	1.14
PVARj	1.04	0.99	0.94	1.20	0.92	1.09	**0.73**	1.00	1.00	2.05

Table 4: **N=10, G=3 setup:** One step ahead root mean squared forecast errors for different models, relative to the BVAR forecast. The first row is the absolute error of the BVAR benchmark. Values in bold denote the lowest relative RMSFE for that country.

IRFs I

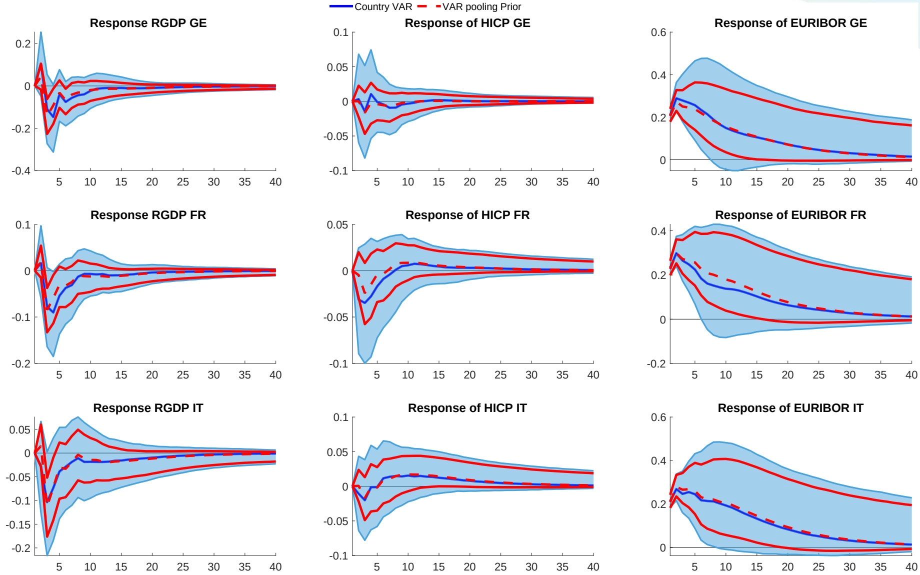


Figure 2: Impulse responses following a one standard deviation shock to the interest rates. Shaded areas represent one standard deviation probability intervals.

IRFs II

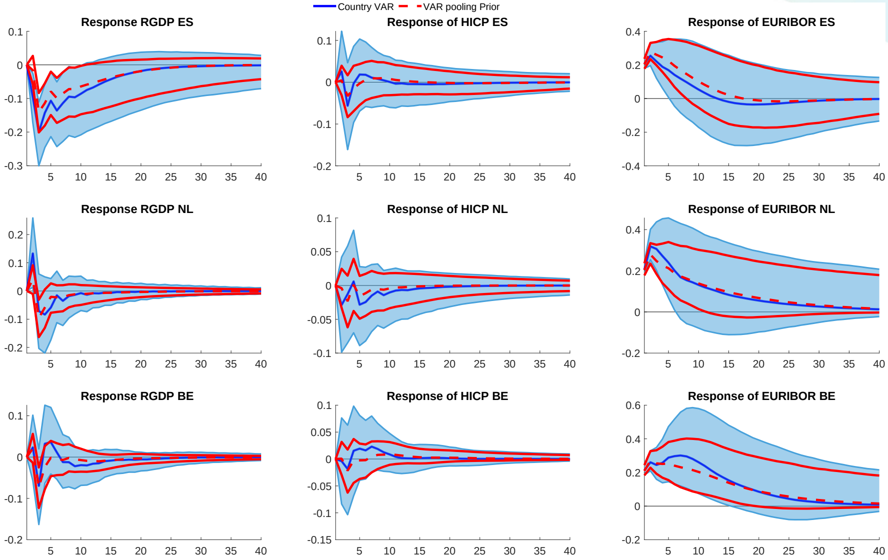


Figure 3: Impulse responses following a one standard deviation shock to the interest rates. Shaded areas represent one standard deviation probability intervals.

IRFs III

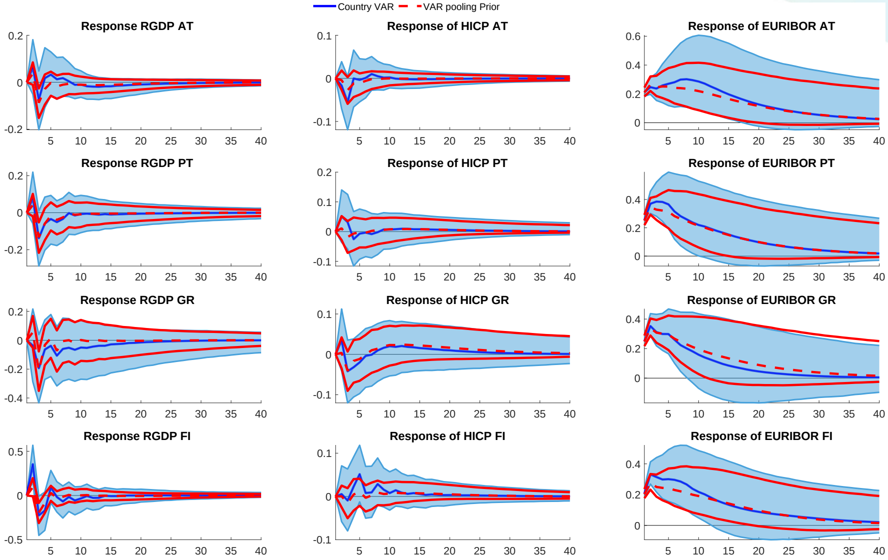


Figure 4: Impulse responses following a one standard deviation shock to the interest rates. Shaded areas represent one standard deviation probability intervals.

Clustering analysis

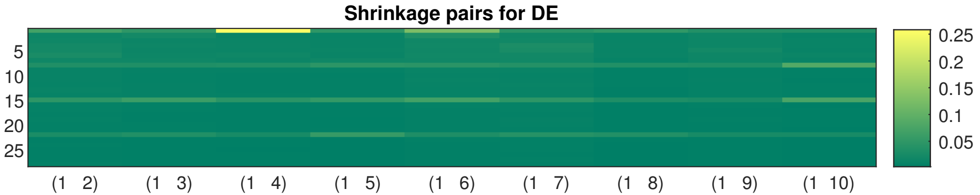


Figure 5: Estimated hyperparameters for the all countries w.r.t. Germany for GDP

$$y_t^1 = b_{y1}^1 y_{t-1}^1 + b_{x1}^1 x_{t-1}^1 + \ldots + b_{y2}^1 y_{t-2}^1 + b_{x2}^1 x_{t-2}^1 \ldots + b_{y4}^1 x_{t-4}^1 + \ldots u_{yt}^1 \quad (6)$$

$$y_t^2 = b_{y1}^2 y_{t-1}^2 + b_{x1}^2 x_{t-1}^2 + \ldots + b_{y2}^2 y_{t-2}^2 + b_{x2}^2 x_{t-2}^2 \ldots + b_{y4}^2 x_{t-4}^2 + \ldots u_{yt}^2 \quad (7)$$

Conclusion

- We propose a flexible VAR framework that endogenously identifies similarities across countries
- It is particularly suitable for the euro area countries
  - Tightly integrated across some dimensions (e.g. financial markets)
  - Widely different in others (e.g. labour markets)
- Euro area application as well as G-7
  - Cross-section information helps parameter identification
  - Better forecasts
  - Narrow error bands with structural identification

References I

- Bańbura, M., D. Giannone, and L. Reichlin (2010). “Large Bayesian Vector Auto Regressions”. In: *Journal of Applied Econometrics* 25.1, pp. 71–92.
- Canova, F. and M. Ciccarelli (2009). “ESTIMATING MULTICOUNTRY VAR MODELS”. In: *International Economic Review* 50.3, pp. 929–959.
- Carriero, A., T. Clark, and M. Marcellino (2016). “Common Drifting Volatility in Large Bayesian VARs”. In: *Journal of Business and Economic Statistics* 34.3, pp. 375–390.
- Chan, J. (2020). “Large Bayesian VARs: A Flexible Kronecker Error Covariance Structure”. In: *Journal of Business and Economic Statistics* 38.1, pp. 68–79.
- Chan, J., G. Koop, and X. Yu (2021). “Large Order-Invariant Bayesian VARs with Stochastic Volatility”. In: *Working Paper.*
- Cross, J., C. Hou, and A. Poon (2020). “Macroeconomic forecasting with large Bayesian VARs: Global-local priors and the illusion of sparsity”. In: *International Journal of Forecasting* 36.3, pp. 899–915.

References II

- Doan, T., R. Litterman, and C. Sims (1984). “Forecasting and Conditional Projection using Realistic Prior Distributions”. In: *Econometric Reviews* 3, pp. 1–100.
- Giannone, D., M. Lenza, and G. Primiceri (2015). “Prior selection for Vector Autoregressions”. In: *Review of Economics and Statistics* 97.2, pp. 436–451.
- Huber, F. and M. Feldkircher (2019). “Adaptive Shrinkage in Bayesian Vector Autoregressive Models”. In: *Journal of Business and Economic Statistics* 37.1, pp. 27–39.
- Jarocinski, M. (2010). “Responses to monetary policy shocks in the east and the west of Europe: a comparison.”. In: *Journal of Applied Econometrics* 25.5, pp. 833–868.
- Koop, G. (2013). “Forecasting with medium and large Bayesian VARs”. In: *Journal of Applied Econometrics* 28.2, pp. 177–203.
- Koop, G. and D. Korobilis (2015). “Model Uncertainty in Panel Vector Autoregressions”. In: *European Economic Review* 81, pp. 115–131.

References III

- Korobilis, D. (2016). “Prior Selection for Panel Vector Autoregressions”. In: *Computational Statistics and Data Analysis* 101, pp. 110–120.
- Korobilis, D. and D. Pettenuzzo (2019). “Adaptive hierarchical Priors for high-dimensional Vector Autoregressions”. In: *Journal of Econometrics* 212.1, pp. 241–271.
- Litterman, R. (1986). “Forecasting With Bayesian vector autoregressions: five years of experience”. In: *Journal of Business and Economic Statistics* 4.1, pp. 25–38.
- Pesaran, H., Y. Shin, and R. Smith (1999). “Pooled Mean Group Estimation of Dynamic Heterogeneous Panels”. In: *Journal of the American Statistical Association* 94.446, pp. 621–634.
- Sims, CA. and T. Zha (1998). “Bayesian methods for dynamic multivariate models”. In: *International Economic Review* 39.4, pp. 949–968.
- Zellner, A. and C. Hong (1989). “Forecasting international growth rates using bayesian shrinkage and other procedures”. In: *Journal of Econometrics* 40.1, pp. 183–202.

Condition posterior distributions of $\tau_{i,m}$ and $\lambda_{i,m,j}$

$$\tau^2_{i,m} \sim IG\left(\frac{K+1}{2}, \frac{1}{v_{\tau^2_{i,m}}} + 0.5\sum_{j=1}^{K}\frac{(\beta_{j+K(i-1)} - \beta_{j+Km})^2}{\lambda^2_{i,m,j}}\right), \tag{8}$$

$$\lambda^2_{i,m,j} \sim IG\left(1, \frac{1}{v_{\lambda^2_{i,m,j}}} + 0.5\frac{(\beta_{j+K(i-1)} - \beta_{j+Km})^2}{\tau^2_{i,m}}\right), \tag{9}$$

$$v_{\tau^2_{i,m}} \sim IG\left(1, 1 + \frac{1}{\tau^2_{i,m}}\right), \tag{10}$$

$$v_{\lambda^2_{i,m,j}} \sim IG\left(1, 1 + \frac{1}{\lambda^2_{i,m,j}}\right). \tag{11}$$

$$\tag{12}$$

Convergence and overshrinkage

	DGP 1			DGP 2			DGP 3		
	T=50	T=120	T=500	T=50	T=120	T=500	T=50	T=120	T=500
pVAR	0.126	0.090	0.064	0.114	0.075	0.046	0.067	0.048	0.029
BVAR	0.117	0.085	0.064	0.113	0.075	0.048	0.073	0.051	0.030
relative	1.054	1.064	1.017	0.985	0.993	0.958	0.749	0.832	0.935

Table 5: Mean absolute errors (MAE), average over all coefficients. Estimated on 200 datasets with varying sample length T. BVAR: Bayesian VAR with Minnesota prior. pVAR: pooling VAR with Minnesota shrinkage.

Forecasting Results for the large model

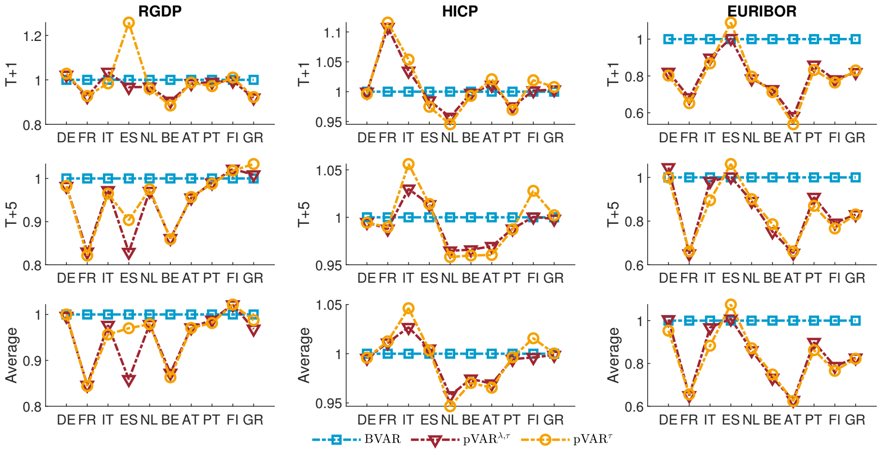


Figure 6: **N=10, G=7 setup:** RMSFEs relative to the BVAR baseline (y-axis) versus countries on the (x-axis). Less is better.