VARMA versus VAR for Macroeconomic Forecasting

George Athanasopoulos

Department of Econometrics and Business Statistics

Monash University

Farshid Vahid

School of Economics

Australian National University

Outline

- VAR models dominate
- Why VARMA?
  - More parsimonious representation
  - Closed with respect to linear transformations

- VAR models dominate
- Why VARMA?
  - More parsimonious representation
  - Closed with respect to linear transformations

Difficult to Identify

"If univariate ARIMA modelling is difficult then VARMA modelling is even more difficult - some might say impossible!" - Chatfield

- VAR models dominate
- Why VARMA?
  - More parsimonious representation
  - Closed with respect to linear transformations

Difficult to Identify

"If univariate ARIMA modelling is difficult then VARMA modelling is even more difficult - some might say impossible!" - Chatfield

Identification Problem

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} - \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix}$$

- VAR models dominate
- Why VARMA?
  - More parsimonious representation
  - Closed with respect to linear transformations

Difficult to Identify

"If univariate ARIMA modelling is difficult then VARMA modelling is even more difficult - some might say impossible!" - Chatfield

Identification Problem

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} - \begin{bmatrix} \theta_{11} & \theta_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix}$$

MONASH University

- VAR models dominate
- Why VARMA?
  - More parsimonious representation
  - Closed with respect to linear transformations

Difficult to Identify

"If univariate ARIMA modelling is difficult then VARMA modelling is even more difficult - some might say impossible!" - Chatfield

Identification Problem

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} - \begin{bmatrix} \theta_{11} & \theta_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix}$$

$$y_{2,t} = \varepsilon_{2,t} \Rightarrow y_{2,t-1} = \varepsilon_{2,t-1}$$

- VAR models dominate
- Why VARMA?
  - More parsimonious representation
  - Closed with respect to linear transformations

Difficult to Identify

"If univariate ARIMA modelling is difficult then VARMA modelling is even more difficult - some might say impossible!" - Chatfield

Identification Problem

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} - \begin{bmatrix} \theta_{11} & \theta_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix}$$

$$y_{2,t} = \varepsilon_{2,t} \Rightarrow y_{2,t-1} = \varepsilon_{2,t-1} \Rightarrow (\phi_{12}, \theta_{12})$$

- VAR models dominate
- Why VARMA?
  - More parsimonious representation
  - Closed with respect to linear transformations

Difficult to Identify

"If univariate ARIMA modelling is difficult then VARMA modelling is even more difficult - some might say impossible!" - Chatfield

Identification Problem

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} - \begin{bmatrix} \theta_{11} & \theta_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix}$$

$$y_{2,t} = \varepsilon_{2,t} \Rightarrow y_{2,t-1} = \varepsilon_{2,t-1} \Rightarrow (\phi_{12}, \theta_{12})$$

- SCM framework: Tiao & Tsay (1989) completed by Athanasopoulos & Vahid (2006)
- Echelon form: Hannan & Kavalieris (1984); Poskitt (1992); Lütkepohl & Poskitt (1996), Athanasopoulos, Poskitt & Vahid (2007)

Outline

Definition of a SCM: For a given K-dimensional $VARMA(p,q)$

$$\mathbf{y}_t = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \eta_t - \mathbf{\Theta}_1 \eta_{t-1} - \ldots - \mathbf{\Theta}_q \eta_{t-q} \qquad (1)$$

Definition of a SCM: For a given K-dimensional $VARMA(p,q)$

$$\mathbf{y}_t = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \eta_t - \mathbf{\Theta}_1 \eta_{t-1} - \ldots - \mathbf{\Theta}_q \eta_{t-q} \qquad (1)$$

$$z_{r,t} = {\alpha_r}' \mathbf{y}_t \sim SCM(p_r, q_r)$$

$$\text{if } \alpha_r \text{ satisfies} \quad \begin{array}{l} {\alpha_r}' \mathbf{\Phi}_{p_r} \neq \mathbf{0}^T \text{ where } 0 \leq p_r \leq p \\ {\alpha_r}' \mathbf{\Phi}_l = \mathbf{0}^T \quad \text{for } l = p_r + 1, ..., p \\ {\alpha_r}' \mathbf{\Theta}_{q_r} \neq \mathbf{0}^T \text{ where } 0 \leq q_r \leq q \\ {\alpha_r}' \mathbf{\Theta}_l = \mathbf{0}^T \text{ for } l = q_r + 1, ..., q \end{array}$$

Definition of a SCM: For a given K-dimensional $VARMA(p,q)$

$$\mathbf{y}_t = \mathbf{\Phi}_1\mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p\mathbf{y}_{t-p} + \eta_t - \mathbf{\Theta}_1\eta_{t-1} - \ldots - \mathbf{\Theta}_q\eta_{t-q} \qquad (1)$$

$$z_{r,t} = {\alpha_r}'\mathbf{y}_t \sim SCM(p_r, q_r)$$

if α_r satisfies

$$\begin{aligned} &{\alpha_r}'\mathbf{\Phi}_{p_r} \neq \mathbf{0}^T \text{ where } 0 \leq p_r \leq p \\ &{\alpha_r}'\mathbf{\Phi}_l = \mathbf{0}^T \text{ for } l = p_r + 1, \ldots, p \\ &{\alpha_r}'\mathbf{\Theta}_{q_r} \neq \mathbf{0}^T \text{ where } 0 \leq q_r \leq q \\ &{\alpha_r}'\mathbf{\Theta}_l = \mathbf{0}^T \text{ for } l = q_r + 1, \ldots, q \end{aligned}$$

SCM Methodology: Find K-linearly independent vectors $\mathbf{A} = (\alpha_1, \ldots, \alpha_K)'$ which transform (1) into

$$\mathbf{A}\mathbf{y}_t = \mathbf{\Phi}_1^*\mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p^*\mathbf{y}_{t-p} + \varepsilon_t - \mathbf{\Theta}_1^*\varepsilon_{t-1} - \ldots - \mathbf{\Theta}_q^*\varepsilon_{t-q} \qquad (2)$$

where $\mathbf{\Phi}_i^* = \mathbf{A}\mathbf{\Phi}_i, \varepsilon_t = \mathbf{A}\eta_t$ and $\mathbf{\Theta}_i^* = \mathbf{A}\mathbf{\Theta}_i\mathbf{A}^{-1}$

Definition of a SCM: For a given K-dimensional $VARMA(p,q)$

$$\mathbf{y}_t = \mathbf{\Phi}_1\mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p\mathbf{y}_{t-p} + \eta_t - \mathbf{\Theta}_1\eta_{t-1} - \ldots - \mathbf{\Theta}_q\eta_{t-q} \qquad (1)$$

$$z_{r,t} = {\alpha_r}'\mathbf{y}_t \sim SCM(p_r, q_r)$$

if α_r satisfies

$$\begin{aligned} {\alpha_r}'\mathbf{\Phi}_{p_r} &\neq \mathbf{0}^T \text{ where } 0 \leq p_r \leq p \\ {\alpha_r}'\mathbf{\Phi}_l &= \mathbf{0}^T \text{ for } l = p_r + 1, ..., p \\ {\alpha_r}'\mathbf{\Theta}_{q_r} &\neq \mathbf{0}^T \text{ where } 0 \leq q_r \leq q \\ {\alpha_r}'\mathbf{\Theta}_l &= \mathbf{0}^T \text{ for } l = q_r + 1, ..., q \end{aligned}$$

SCM Methodology: Find K-linearly independent vectors $\mathbf{A} = (\alpha_1, \ldots, \alpha_K)'$ which transform (1) into

$$\mathbf{Ay}_t = \mathbf{\Phi}_1^*\mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p^*\mathbf{y}_{t-p} + \varepsilon_t - \mathbf{\Theta}_1^*\varepsilon_{t-1} - \ldots - \mathbf{\Theta}_q^*\varepsilon_{t-q} \qquad (2)$$

where $\mathbf{\Phi}_i^* = \mathbf{A\Phi}_i, \varepsilon_t = \mathbf{A}\eta_t$ and $\mathbf{\Theta}_i^* = \mathbf{A\Theta}_i\mathbf{A}^{-1}$

Series of C/C tests: $E\begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix}\begin{bmatrix} y_{1,t} & y_{2,t} \end{bmatrix}[\alpha_1] = \mathbf{0}$

$$\alpha_1'\mathbf{y}_t \sim SCM(0,0)$$

Example: $K = 3$

$$\begin{aligned} \alpha_1' \mathbf{y}_t &\sim SCM(1,1) \\ \alpha_2' \mathbf{y}_t &\sim SCM(1,0) \\ \alpha_3' \mathbf{y}_t &\sim SCM(0,0) \end{aligned}$$

Example: $K = 3$

$$\begin{aligned} \alpha_1' \mathbf{y}_t &\sim SCM(1,1) \\ \alpha_2' \mathbf{y}_t &\sim SCM(1,0) \\ \alpha_3' \mathbf{y}_t &\sim SCM(0,0) \end{aligned}$$

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \mathbf{y}_t = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \phi_{13}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \phi_{23}^{(1)} \\ 0 & 0 & 0 \end{bmatrix} \mathbf{y}_{t-1} + \varepsilon_t - \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \varepsilon_{t-1}$$

Example: $K = 3$

$$\begin{aligned} \alpha_1' \mathbf{y}_t &\sim SCM(1,1) \\ \alpha_2' \mathbf{y}_t &\sim SCM(1,0) \\ \alpha_3' \mathbf{y}_t &\sim SCM(0,0) \end{aligned}$$

$$\begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & 1 & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & 1 \end{bmatrix} \mathbf{y}_t = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \phi_{13}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \phi_{23}^{(1)} \\ 0 & 0 & 0 \end{bmatrix} \mathbf{y}_{t-1} + \varepsilon_t - \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \varepsilon_{t-1}$$

- Normalise diagonally (test for improper normalisations)

Example: $K = 3$

$$\begin{aligned} \alpha_1' \mathbf{y}_t &\sim SCM(1,1) \\ \alpha_2' \mathbf{y}_t &\sim SCM(1,0) \\ \alpha_3' \mathbf{y}_t &\sim SCM(0,0) \end{aligned}$$

$$\begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & 1 & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & 1 \end{bmatrix} \mathbf{y}_t = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \phi_{13}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \phi_{23}^{(1)} \\ 0 & 0 & 0 \end{bmatrix} \mathbf{y}_{t-1} + \varepsilon_t - \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \varepsilon_{t-1}$$

- Normalise diagonally (test for improper normalisations)
- Reduce parameters of **A** to produce a Canonical SCM

Example: $K = 3$

$$\begin{aligned} \alpha_1' \mathbf{y}_t &\sim SCM(1,1) \\ \alpha_2' \mathbf{y}_t &\sim SCM(1,0) \\ \alpha_3' \mathbf{y}_t &\sim SCM(0,0) \end{aligned}$$

$$\begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & 1 & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & 1 \end{bmatrix} \mathbf{y}_t = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \phi_{13}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \phi_{23}^{(1)} \\ 0 & 0 & 0 \end{bmatrix} \mathbf{y}_{t-1} + \varepsilon_t - \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \varepsilon_{t-1}$$

- Normalise diagonally (test for improper normalisations)
- Reduce parameters of **A** to produce a Canonical SCM

Example: $K = 3$

$$\begin{aligned} \alpha_1' \mathbf{y}_t &\sim SCM(1,1) \\ \alpha_2' \mathbf{y}_t &\sim SCM(1,0) \\ \alpha_3' \mathbf{y}_t &\sim SCM(0,0) \end{aligned}$$

$$\begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & 1 & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & 1 \end{bmatrix} \mathbf{y}_t = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \phi_{13}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \phi_{23}^{(1)} \\ 0 & 0 & 0 \end{bmatrix} \mathbf{y}_{t-1} + \varepsilon_t - \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \varepsilon_{t-1}$$

- Normalise diagonally (test for improper normalisations)
- Reduce parameters of **A** to produce a Canonical SCM

$$\begin{bmatrix} 1 & 0 & 0 \\ \alpha_{21} & 1 & 0 \\ \alpha_{31} & \alpha_{32} & 1 \end{bmatrix} \mathbf{y}_t = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \phi_{13}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \phi_{23}^{(1)} \\ 0 & 0 & 0 \end{bmatrix} \mathbf{y}_{t-1} + \varepsilon_t - \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \varepsilon_{t-1}$$

Example: $K = 3$

$$\begin{aligned} \alpha_1' \mathbf{y}_t &\sim SCM(1,1) \\ \alpha_2' \mathbf{y}_t &\sim SCM(1,0) \\ \alpha_3' \mathbf{y}_t &\sim SCM(0,0) \end{aligned}$$

$$\begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & 1 & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & 1 \end{bmatrix} \mathbf{y}_t = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \phi_{13}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \phi_{23}^{(1)} \\ 0 & 0 & 0 \end{bmatrix} \mathbf{y}_{t-1} + \varepsilon_t - \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \varepsilon_{t-1}$$

- Normalise diagonally (test for improper normalisations)
- Reduce parameters of **A** to produce a Canonical SCM

$$\begin{bmatrix} 1 & 0 & 0 \\ \alpha_{21} & 1 & 0 \\ \alpha_{31} & \alpha_{32} & 1 \end{bmatrix} \mathbf{y}_t = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \phi_{13}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \phi_{23}^{(1)} \\ 0 & 0 & 0 \end{bmatrix} \mathbf{y}_{t-1} + \varepsilon_t - \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \varepsilon_{t-1}$$

Empirical Results:

Example: $K = 3$

$$\begin{aligned} \alpha_1' \mathbf{y}_t &\sim SCM(1,1) \\ \alpha_2' \mathbf{y}_t &\sim SCM(1,0) \\ \alpha_3' \mathbf{y}_t &\sim SCM(0,0) \end{aligned}$$

$$\begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & 1 & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & 1 \end{bmatrix} \mathbf{y}_t = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \phi_{13}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \phi_{23}^{(1)} \\ 0 & 0 & 0 \end{bmatrix} \mathbf{y}_{t-1} + \varepsilon_t - \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \varepsilon_{t-1}$$

- Normalise diagonally (test for improper normalisations)
- Reduce parameters of **A** to produce a Canonical SCM

$$\begin{bmatrix} 1 & 0 & 0 \\ \alpha_{21} & 1 & 0 \\ \alpha_{31} & \alpha_{32} & 1 \end{bmatrix} \mathbf{y}_t = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \phi_{13}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \phi_{23}^{(1)} \\ 0 & 0 & 0 \end{bmatrix} \mathbf{y}_{t-1} + \varepsilon_t - \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \varepsilon_{t-1}$$

Empirical Results:

1. Average performance across many trivariate systems
2. A four variable example
3. Simulation: Why do VARMA models do better than VARs

Outline

1. Introduction
2. Canonical SCM
3. **Forecast performance**
4. Example
5. Simulation
6. Summary of findings and future research

Forecasting: 40 monthly macroeconomic variables from 8 general categories of economic activity, 1959:1-1998:12 (N=480)

Forecasting: 40 monthly macroeconomic variables from 8 general categories of economic activity, 1959:1-1998:12 (N=480)

- 50 $\times$ 3 variable systems

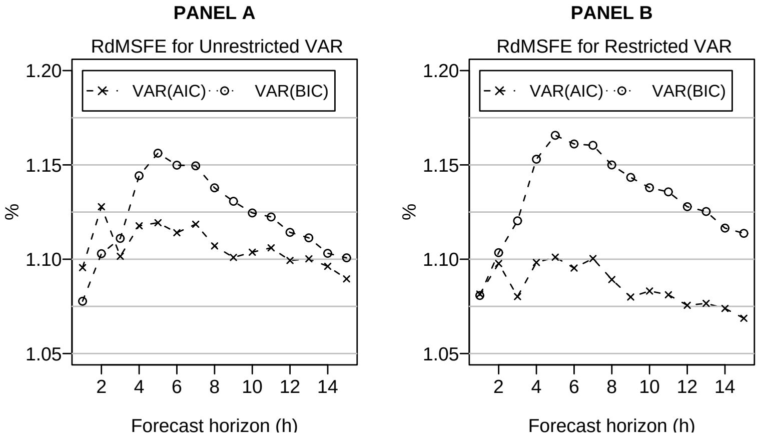
Forecasting: 40 monthly macroeconomic variables from 8 general categories of economic activity, 1959:1-1998:12 (N=480)

- 50×3 variable systems
- Test sample: $N_1 = 300$
  - Estimated canonical SCM VARMA
  - Unrestricted VAR(AIC) and VAR(BIC)
  - Restricted VAR(AIC) and VAR(BIC)

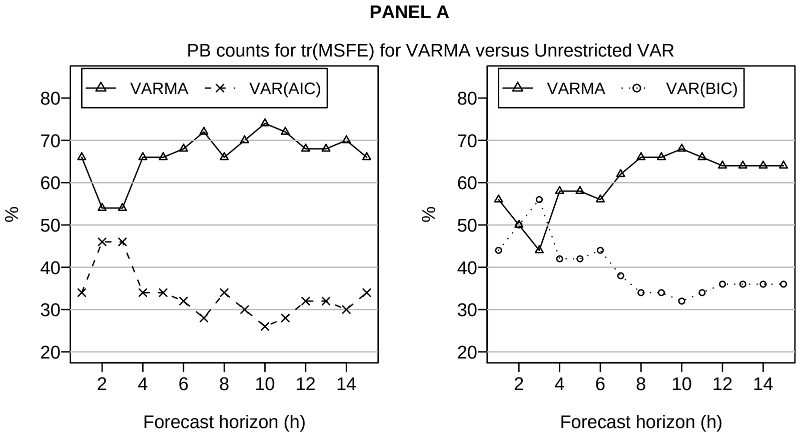
Forecasting: 40 monthly macroeconomic variables from 8 general categories of economic activity, 1959:1-1998:12 (N=480)

- 50×3 variable systems
- Test sample: $N_1 = 300$
  - Estimated canonical SCM VARMA
  - Unrestricted VAR(AIC) and VAR(BIC)
  - Restricted VAR(AIC) and VAR(BIC)
- Hold-out sample: $N_2 = 180$
  - Produced $N_2 - h + 1$ out-of-sample forecasts for each h=1 to 15
  - Forecast error measures: $|MSFE|$ and $tr(MSFE)$
  - Percentage Better: PB_h
  - Relative Ratios:
    - $\overline{RRdMSFE}_h = \frac{1}{50} \sum_{i=1}^{50} \frac{|MSFE_i^{VAR}|}{|MSFE_i^{VARMA}|}$

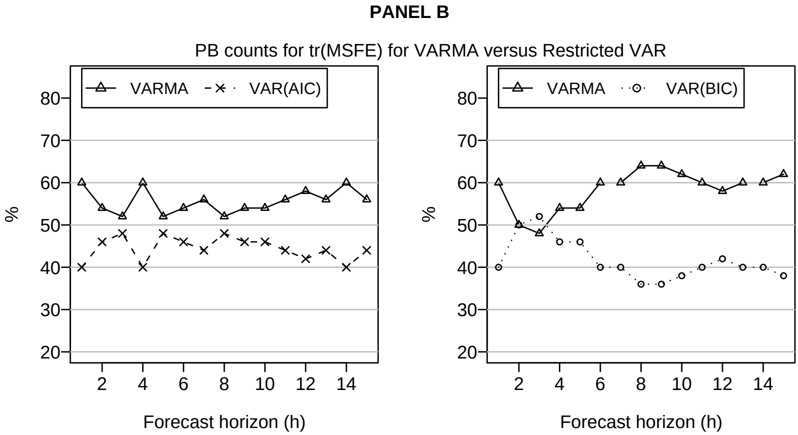
Relative Ratios



Percentage Better: Unrestricted VAR



Percentage Better: Restricted VAR



Diebold-Mariano test for tr(MSFE):

VARMA sign better (5%, 25%, 25%,5%) VARMA sign worst

	Forecast horizon (h)				
	1	4	8	12	15
$VAR(AIC)$ - Unrest	24,46,20,14	16,38,10,0	14,32,6,4	10,22,10,0	10,22,8,0
$VAR(BIC)$ - Unrest	24,50,22,12	10,38,18,10	12,26,22,8	12,18,16,4	12,18,16,2
$VAR(AIC)$ - Rest	34,54,26,12	14,36,16,0	14,26,8,2	12,22,10,0	10,24,10,0
$VAR(BIC)$ - Rest	32,54,22,14	10,38,18,8	12,26,22,6	10,20,20,6	10,16,14,2

Diebold-Mariano test for tr(MSFE):

VARMA sign better (5%, 25%, 25%,5%) VARMA sign worst

	Forecast horizon (h)				
	1	*4*	*8*	*12*	*15*
$VAR(AIC)$ - Unrest	24,46,20,14	16,38,10,0	14,32,6,4	10,22,10,0	10,22,8,0
$VAR(BIC)$ - Unrest	24,50,22,12	10,38,18,10	12,26,22,8	12,18,16,4	12,18,16,2
$VAR(AIC)$ - Rest	34,54,26,12	14,36,16,0	14,26,8,2	12,22,10,0	10,24,10,0
$VAR(BIC)$ - Rest	32,54,22,14	10,38,18,8	12,26,22,6	10,20,20,6	10,16,14,2

Messages:

Diebold-Mariano test for tr(MSFE):

VARMA sign better (5%, 25%, 25%,5%) VARMA sign worst

	Forecast horizon (h)				
	1	*4*	*8*	*12*	*15*
VAR(*AIC*) - Unrest	24,46,20,14	16,38,10,0	14,32,6,4	10,22,10,0	10,22,8,0
VAR(*BIC*) - Unrest	24,50,22,12	10,38,18,10	12,26,22,8	12,18,16,4	12,18,16,2
VAR(*AIC*) - Rest	34,54,26,12	14,36,16,0	14,26,8,2	12,22,10,0	10,24,10,0
VAR(*BIC*) - Rest	32,54,22,14	10,38,18,8	12,26,22,6	10,20,20,6	10,16,14,2

Messages:

1. There are cases where VARMA significantly outperform VAR and vice versa

Diebold-Mariano test for tr(MSFE):

VARMA sign better (5%, 25%, 25%,5%) VARMA sign worst

	Forecast horizon (h)				
	1	*4*	*8*	*12*	*15*
$VAR(AIC)$ - Unrest	24,46,20,14	16,38,10,0	14,32,6,4	10,22,10,0	10,22,8,0
$VAR(BIC)$ - Unrest	24,50,22,12	10,38,18,10	12,26,22,8	12,18,16,4	12,18,16,2
$VAR(AIC)$ - Rest	34,54,26,12	14,36,16,0	14,26,8,2	12,22,10,0	10,24,10,0
$VAR(BIC)$ - Rest	32,54,22,14	10,38,18,8	12,26,22,6	10,20,20,6	10,16,14,2

Messages:

1. There are cases where VARMA significantly outperform VAR and vice versa
2. VARMA models significantly outperform VAR more than the reverse

Diebold-Mariano test for tr(MSFE):

VARMA sign better (5%, 25%, 25%,5%) VARMA sign worst

	Forecast horizon (h)				
	1	*4*	*8*	*12*	*15*
$VAR(AIC)$ - Unrest	24,46,20,14	16,38,10,0	14,32,6,4	10,22,10,0	10,22,8,0
$VAR(BIC)$ - Unrest	24,50,22,12	10,38,18,10	12,26,22,8	12,18,16,4	12,18,16,2
$VAR(AIC)$ - Rest	34,54,26,12	14,36,16,0	14,26,8,2	12,22,10,0	10,24,10,0
$VAR(BIC)$ - Rest	32,54,22,14	10,38,18,8	12,26,22,6	10,20,20,6	10,16,14,2

Messages:

1. There are cases where VARMA significantly outperform VAR and vice versa
2. VARMA models significantly outperform VAR more than the reverse
3. As h increases the number significant differences decreases

Diebold-Mariano test for tr(MSFE):

VARMA sign better (5%, 25%, 25%,5%) VARMA sign worst

	Forecast horizon (h)				
	1	*4*	*8*	*12*	*15*
VAR(*AIC*) - Unrest	24,46,20,14	16,38,10,0	14,32,6,4	10,22,10,0	10,22,8,0
VAR(*BIC*) - Unrest	24,50,22,12	10,38,18,10	12,26,22,8	12,18,16,4	12,18,16,2
VAR(*AIC*) - Rest	34,54,26,12	14,36,16,0	14,26,8,2	12,22,10,0	10,24,10,0
VAR(*BIC*) - Rest	32,54,22,14	10,38,18,8	12,26,22,6	10,20,20,6	10,16,14,2

Messages:

1. There are cases where VARMA significantly outperform VAR and vice versa
2. VARMA models significantly outperform VAR more than the reverse
3. As h increases the number significant differences decreases
4. Restrictions do not improve VAR performance when significant differences

Outline

MONASH University

Example:

Example:

- Four variables (also six variables):
    - GDP growth rate
    - inflation rate
    - spread (10 yr gvt bill yield) − (3-month treasury bill rate)
    - 3-month treasury bill rate

  - in line with term structure literature: Ang, Piazzesi, Wei (2006)
  - variations in New Keynesian DSGE - contributions in Taylor (1999)
- Quarterly data

Example:

- Four variables (also six variables):
  - GDP growth rate
  - inflation rate
  - spread (10 yr gvt bill yield) − (3-month treasury bill rate)
  - 3-month treasury bill rate

  - in line with term structure literature: Ang, Piazzesi, Wei (2006)
  - variations in New Keynesian DSGE - contributions in Taylor (1999)
- Quarterly data

Message: We should start considering VARMA

Outline

Why do VARMA forecast better?

Estimated a VARMA(2,1): - SCM(2,0) - SCM(1,1)
- SCM(1,0) - SCM(0,0)

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ * & * & * & 1 \end{pmatrix} \mathbf{y}_t = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{y}_{t-1} + \begin{pmatrix} * & * & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{y}_{t-2}$$

$$- \begin{pmatrix} 0 & 0 & 0 & 0 \\ * & * & * & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{e}_{t-1} + \mathbf{e}_t$$

Why do VARMA forecast better?

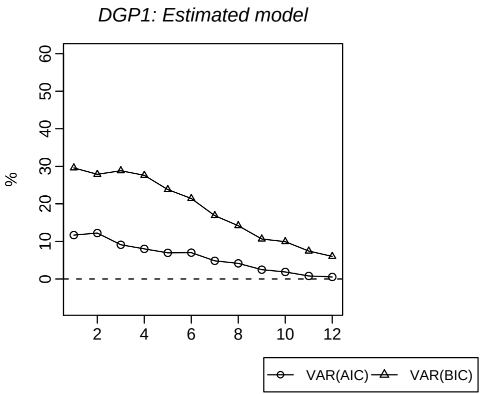
Estimated a VARMA(2,1): - SCM(2,0) - SCM(1,1)
- SCM(1,0) - SCM(0,0)

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ * & * & * & 1 \end{pmatrix} \mathbf{y}_t = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{y}_{t-1} + \begin{pmatrix} * & * & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{y}_{t-2}$$

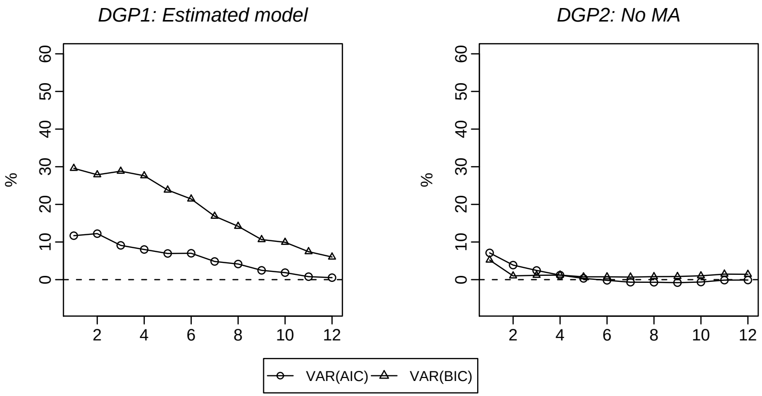
$$- \begin{pmatrix} 0 & 0 & 0 & 0 \\ * & * & * & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{e}_{t-1} + \mathbf{e}_t$$

- Simulate from the benchmark estimated model assuming $\mathbf{e} \sim N(\mathbf{0}, \mathbf{\Sigma})$
- $n = 164 \longrightarrow$ estimate VARMA(2,1), VAR(AIC), VAR(BIC)
- $nout = 42 \longrightarrow$ compute 1 to 12-step ahead forecasts
- $iterations = 100 \longrightarrow$ calculate $\overline{|MSFE|}$ for all models
- Compare the percentage difference using $\overline{|MSFE_{VARMA}|}$ as a base
- Repeat by changing specific features and compare with the benchmark

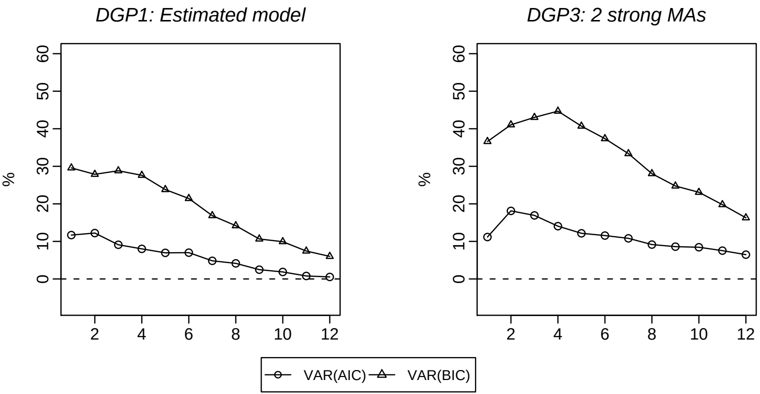
Why do VARMA forecast better:



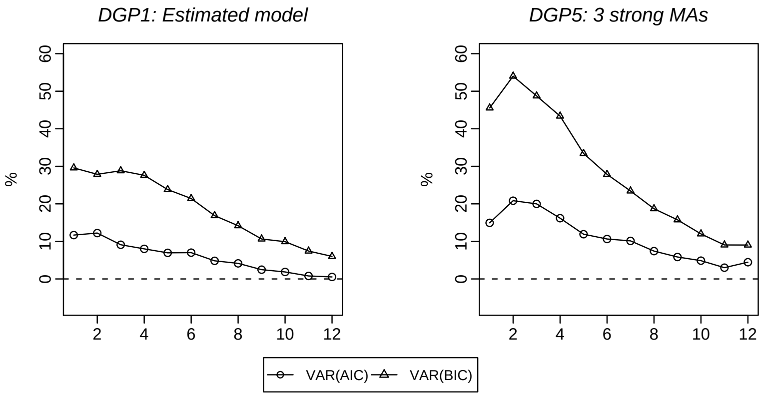
Why do VARMA forecast better:



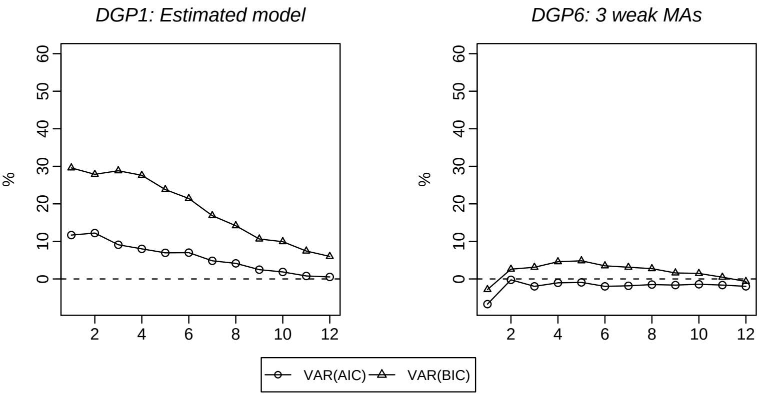
Why do VARMA forecast better:



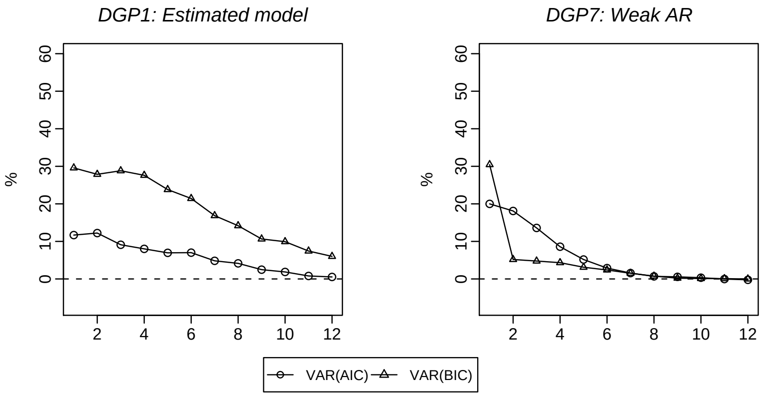
Why do VARMA forecast better:



Why do VARMA forecast better:



Why do VARMA forecast better:



Outline

Summary of findings:

1. We can obtain better forecasts for macroeconomic variables by considering VARMA models

Summary of findings:

1. We can obtain better forecasts for macroeconomic variables by considering VARMA models
2. With the methodological developments and the improvement in computer power there is no compelling reason to restrict the class of models to VARs only

Summary of findings:

1. We can obtain better forecasts for macroeconomic variables by considering VARMA models
2. With the methodological developments and the improvement in computer power there is no compelling reason to restrict the class of models to VARs only
3. The existence of VMA components cannot be well-approximated by finite order VARs

Summary of findings:

1. We can obtain better forecasts for macroeconomic variables by considering VARMA models
2. With the methodological developments and the improvement in computer power there is no compelling reason to restrict the class of models to VARs only
3. The existence of VMA components cannot be well-approximated by finite order VARs
4. Are these favourable results specific the SCM methodology? No! Athanasopoulos, Poskitt and Vahid (2007) show that similar conclusions emerge when one uses the "Echelon" form approach

Summary of findings:

1. We can obtain better forecasts for macroeconomic variables by considering VARMA models
2. With the methodological developments and the improvement in computer power there is no compelling reason to restrict the class of models to VARs only
3. The existence of VMA components cannot be well-approximated by finite order VARs
4. Are these favourable results specific the SCM methodology? No! Athanasopoulos, Poskitt and Vahid (2007) show that similar conclusions emerge when one uses the "Echelon" form approach

Future Research:

1. Developing a fully automated identification process
2. Developing an alternative estimation approach which avoids fitting a long VAR to estimate the lagged innovations
3. Move into the non-stationary world

Thank you!!!!!