

Vector Autoregressions (VARs)

Gabor Pinter

CCBS, Bank of England

3 February 2015

- 1 Literature
- 2 What are VARs used for?
- 3 Description of basic VAR models
- 4 Impulse response functions
- 5 Forecasting
- 6 Forecast error variance decomposition (FEVD)
- 7 Historical decomposition (HD)

- Seminal paper: Sims [1980] → Nobel-prize, 2011
- Estimation of VARs: Canova [2007]
- Structural VARs: Stock and Watson [2001] and Rubio-Ramirez et al. [2010]
- Bayesian VARs: Blake and Mumtaz [2012]

What are VARs good for?

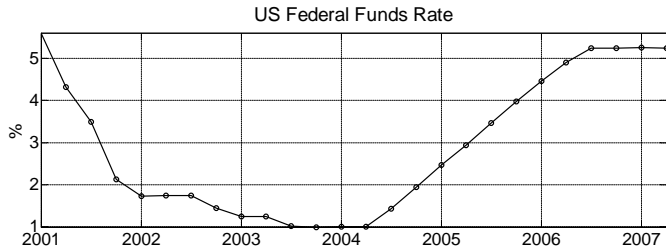
- One of the main tools for central bankers
- VAR models are designed to answer questions like:
 - 1 Impact analysis: What is the impact (size and duration) of a 1% (surprise) change in the policy rate on inflation and output?
 - News versus surprises: What is the impact of surprise changes in technology compared to the impact of news about future expected changes?
 - 2 Narrative: Was the 2008-2009 dip in GDP growth caused by credit supply or credit demand shocks?
 - What fraction of variation in inflation and output in the past 40 years has been due to monetary policy?
 - 3 Forecast: Is the recovery sustainable or is there some probability of an other economic contraction?
 - 4 Data description: does unemployment help predict inflation better than the interest rate?

Motivating example

US Monetary Policy during 2001-2007

- Has US monetary policy been an important driver of the pre-crisis build-up of financial sector leverage?
- A popular narrative is that the low interest rate environment in the US post 9/11 contributed to the balance sheet expansion of the financial sector (housing bubble)

Figure : Pre-crisis US Monetary Policy



Motivating example

US Monetary Policy during 2001-2007

- First, it is unclear whether the low interest rates were a natural consequence of the Fed behaving optimally by responding to deteriorating business cycle conditions via cutting rates.
 - Could low rates be explained by the systematic part of policy, or were low rates in fact the result of non-systematic policy shocks?
 - Identifying monetary policy shocks and separating them from the systematic parts of policy is needed to understand historical interest rate dynamics.
- Second, it is unclear what the partial impact of monetary policy shocks has been on balance sheet expansions, given other factors (e.g. structural shocks related to increased housing demand and commodity price booms)
 - One must disentangle monetary policy shocks from other structural shocks in order to assess the relative importance of monetary policy

What is a VAR

- Suppose we have data on C (consumption) and Y (output). The VAR model of order 1 (VAR(1)) is written as:

$$\begin{aligned}C_t &= \lambda + \beta_1 C_{t-1} + \beta_2 Y_{t-1} + e_t \\Y_t &= \alpha + \rho_1 C_{t-1} + \rho_2 Y_{t-1} + \nu_t \\ \Sigma &= E(e_t, \nu_t)\end{aligned}\tag{1.1}$$

- Each variable depends on the lags of the two \rightarrow each variable is endogenous

Estimating a VAR model

- Just N linear regression
- Estimation of the coefficients is done by OLS:

$$\hat{B} = (X'X)^{-1} X'Y \quad (1.2)$$

$$Y = \begin{pmatrix} C_2 & Y_2 \\ C_3 & Y_3 \\ \vdots & \vdots \\ C_T & Y_T \end{pmatrix}, \quad X = \begin{pmatrix} 1 & C_1 & Y_1 \\ 1 & C_2 & Y_2 \\ 1 & \vdots & \vdots \\ 1 & C_{T-1} & Y_{T-1} \end{pmatrix}$$

- The residuals are used to estimate the variance-covariance matrix:

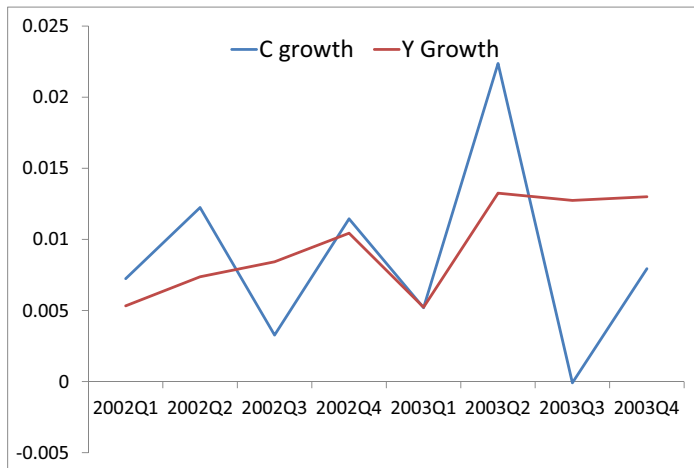
$$\Omega = Y - X\hat{B} \quad (1.3)$$

$$\Sigma = \frac{\Omega'\Omega}{T - K}$$

Estimating a VAR model

Illustration

Figure : Growth of real consumption and output in the UK, 2002Q1-2003Q4



Estimating a VAR model

Illustration

- The data:

$$UK \text{ Data} = \begin{bmatrix} \text{Period} & C \text{ growth} & Y \text{ growth} \\ 2002Q1 & 0.0072 & 0.0053 \\ 2002Q2 & 0.0122 & 0.0073 \\ 2002Q3 & 0.0032 & 0.0084 \\ 2002Q4 & 0.0114 & 0.0104 \\ 2003Q1 & 0.0051 & 0.0052 \\ 2003Q2 & 0.0223 & 0.0132 \\ 2003Q3 & -8.8E - 05 & 0.0127 \\ 2003Q4 & 0.0079 & 0.0129 \end{bmatrix} \quad (1.4)$$

- Y and X matrices:

$$Y = \begin{bmatrix} 0.0122 & 0.0073 \\ 0.0032 & 0.0084 \\ 0.0114 & 0.0104 \\ 0.0051 & 0.0052 \\ 0.0223 & 0.0132 \\ -8.8E - 05 & 0.0127 \\ 0.0079 & 0.0129 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 0.0072 & 0.0053 \\ 1 & 0.0122 & 0.0073 \\ 1 & 0.0032 & 0.0084 \\ 1 & 0.0114 & 0.0104 \\ 1 & 0.0051 & 0.0052 \\ 1 & 0.0223 & 0.0132 \\ 1 & -8.8E - 05 & 0.0127 \end{bmatrix} \quad (1.5)$$

Estimating a VAR model

Illustration

- The OLS estimate:

$$\hat{B} = (X'X)^{-1} X'Y \quad (1.6)$$

- Using the data matrices covering 7 quarters:

$$\hat{B} = \begin{bmatrix} 0.024 & 0.008 \\ -0.46 & -0.09 \\ 1.254 & 0.309 \end{bmatrix} \quad (1.7)$$

- Plug the estimates back into our VAR model 1.1:

$$\begin{aligned} C_t &= 0.024 - 0.46C_{t-1} + 1.254Y_{t-1} + e_t \\ Y_t &= 0.008 - 0.09C_{t-1} + 0.309Y_{t-1} + \nu_t \end{aligned} \quad (1.8)$$

Estimating a VAR model

Lag length

- Choosing the appropriate lag length is very important
- Model Selection Criteria (AIC, SIC etc) and LR tests
- Degrees of freedom \rightarrow a lot of coefficients can lead to imprecise estimates

Lag length in a VAR

- Total number of parameters, K , in a VAR with
 - N number of endogenous variables, and
 - P number of lags

$$K = ((N \times P) + 1) \times N \quad (1.9)$$

- For example, with $N = 4$ and $P = 4$, the number of parameters is $K = 68!$
- If time series is short, then the choice of P is crucial:

$$\begin{aligned} AIC &= \ln \det(\Sigma) + \frac{2N^2P}{T} \\ SIC &= \ln \det(\Sigma) + \frac{N^2P \ln T}{T} \end{aligned} \quad (1.10)$$

- Choose smallest AIC/SIC after estimating from $P = 1 \dots p$

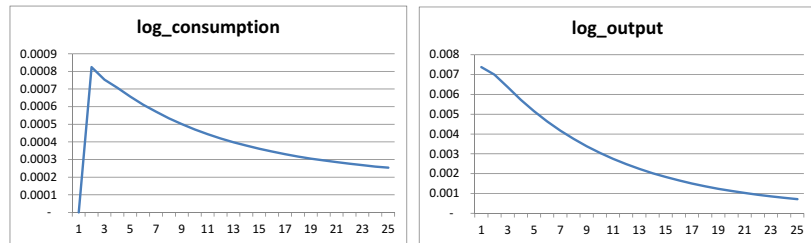
Inference VARs: Impulse Response Functions

- Consider our VAR(1) model:

$$\begin{aligned}C_t &= \lambda + \beta_1 C_{t-1} + \beta_2 Y_{t-1} + e_t \\ Y_t &= \alpha + \rho_1 C_{t-1} + \rho_2 Y_{t-1} + \nu_t\end{aligned}\tag{2.1}$$

- Assume the coefficients are known: What is the impact on C if there is a shock to Y ?
- The impact of this is referred to as the impulse response function (IRF)

Figure : Example for IRFs



Impulse Response Functions

Experiment

- Tracing out the reaction of the variables in the VAR to a shock or innovation in one of the variables
- Consider a VAR(1):

$$\begin{aligned}y_t &= c_1 + b_{11}y_{t-1} + b_{12}x_{t-1} + \nu_{1,t} \\x_t &= c_2 + b_{21}y_{t-1} + b_{22}x_{t-1} + \nu_{2,t}\end{aligned}\tag{2.2}$$

- Consider the following experiment:
 - Time 0: both y and x are zero and $c = 0$
 - Time 1: one unit increase in x

Impulse Response Functions

Tracing out the impact

- Initial Shock:

$$\begin{pmatrix} \nu_{1,t} \\ \nu_{2,t} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (2.3)$$

- Period 1:

$$\begin{pmatrix} y_1 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (2.4)$$

- Period 2:

$$\begin{pmatrix} y_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ x_1 \end{pmatrix} = \begin{pmatrix} b_{11}y_1 + b_{12}x_1 \\ b_{21}y_1 + b_{22}x_1 \end{pmatrix} = \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix} \quad (2.5)$$

Impulse Response Functions

- Period 3:

$$\begin{pmatrix} y_3 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} y_2 \\ x_2 \end{pmatrix} \quad (2.6)$$

$$\begin{pmatrix} y_3 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}^2 \begin{pmatrix} y_1 \\ x_1 \end{pmatrix} \quad (2.7)$$

$$\begin{pmatrix} y_3 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_{11}^2 + b_{12}b_{21} & b_{11}b_{12} + b_{12}b_{22} \\ b_{21}b_{11} + b_{22}b_{21} & b_{21}b_{12} + b_{22}^2 \end{pmatrix} \begin{pmatrix} y_1 \\ x_1 \end{pmatrix} \quad (2.8)$$

$$\begin{pmatrix} y_3 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_{11}b_{12} + b_{12}b_{22} \\ b_{21}b_{12} + b_{22}^2 \end{pmatrix} \quad (2.9)$$

- The impulse response for period i is:

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}^{i-1} \quad (2.10)$$

- VARs are useful for forecasting as all variables are endogenous:

$$\begin{aligned} \begin{pmatrix} \hat{Y}_{T+1} \\ \hat{C}_{T+1} \end{pmatrix} &= \begin{pmatrix} \hat{b}_{11} & \hat{b}_{12} \\ \hat{b}_{21} & \hat{b}_{22} \end{pmatrix} \begin{pmatrix} Y_T \\ C_T \end{pmatrix} \\ \begin{pmatrix} \hat{Y}_{T+2} \\ \hat{C}_{T+2} \end{pmatrix} &= \begin{pmatrix} \hat{b}_{11} & \hat{b}_{12} \\ \hat{b}_{21} & \hat{b}_{22} \end{pmatrix} \begin{pmatrix} \hat{Y}_{T+1} \\ \hat{C}_{T+1} \end{pmatrix} \\ &\vdots \end{aligned} \tag{3.1}$$

- Useful as a short term forecasting tool \rightarrow but little theoretical underpinning

- Conditional forecasting is a useful application
- Assume some of the variables in the VAR are exogenous in the future (e.g. pegged interest rate)
- These may be variables for which data is more readily available or a future path of these variables is available

Conditional Forecasting

- Consider a simple VAR:

$$Y_t = c + BY_{t-1} + A_0\varepsilon_t \quad (3.2)$$

where A_0 is the square root of the error covariance matrix.

- Consider the one step-ahead forecast:

$$Y_{t+1} = c + B(c + BY_{t-1} + A_0\varepsilon_t) + A_0\varepsilon_{t+1} \quad (3.3)$$

$$Y_{t+1} = [c + Bc + B^2Y_{t-1}] + \underbrace{[BA_0\varepsilon_t + A_0\varepsilon_{t+1}]}_{\text{Contribution of future shocks}} \quad (3.4)$$

- In conditional forecasts we calculate future shocks that give us the constrained forecast (e.g. Compass - BoE forecasting platform)

Forecast Error Variance Decomposition (FEVD)

- A concept closely related to impulse responses
- Imagine you make a forecast for the variables in the VAR
- Because of shocks, there will be an error in forecasting
- FEVD refers to working out how much each shock contributed to the forecast error variance

Forecast Error Variance Decomposition (FEVD)

- Consider variable Y one period into the future
- The variables are going to be hit by the two shocks in our model
- The IRF and the scale of the shocks will determine how much each variable moves

$$\begin{aligned}C_t &= \lambda + \beta_1 C_{t-1} + \beta_2 Y_{t-1} + e_t \\ Y_t &= \alpha + \rho_1 C_{t-1} + \rho_2 Y_{t-1} + \nu_t\end{aligned}\tag{4.1}$$

Table : Shock contributions

| Shock | e_t | ν_t |
|--------------------------|---------------------------|-------------------------------|
| Movement in period 1 | $\sigma_e \Psi_{e,1}$ | $\sigma_\nu \Psi_{\nu,1}$ |
| Contribution to variance | $(\sigma_e \Psi_{e,1})^2$ | $(\sigma_\nu \Psi_{\nu,1})^2$ |

Forecast Error Variance Decomposition (FEVD)

- 2 periods in the future – impact of the shocks in period 1 and period 2 and so on

Table : Shock contributions

| Shock | e_t | ν_t |
|--------------------------|---|---|
| Movement in period 2 | $\sigma_e \Psi_{e,1} + \sigma_e \Psi_{e,2}$ | $\sigma_\nu \Psi_{\nu,1} + \sigma_\nu \Psi_{\nu,2}$ |
| Contribution to variance | $(\sigma_e \Psi_{e,1})^2 + (\sigma_e \Psi_{e,2})^2$ | $(\sigma_\nu \Psi_{\nu,1})^2 + (\sigma_\nu \Psi_{\nu,2})^2$ |

Historical Decomposition

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \nu_{1,t} \\ \nu_{2,t} \end{pmatrix} \quad (5.1)$$

$$\begin{aligned} \begin{pmatrix} y_t \\ x_t \end{pmatrix} &= \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \left(\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} y_{t-2} \\ x_{t-2} \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \nu_{1,t-1} \\ \nu_{2,t-1} \end{pmatrix} \right) \\ &+ \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \nu_{1,t} \\ \nu_{2,t} \end{pmatrix} \end{aligned} \quad (5.2)$$

$$\begin{aligned} \begin{pmatrix} y_t \\ x_t \end{pmatrix} &= \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}^2 \begin{pmatrix} y_{t-2} \\ x_{t-2} \end{pmatrix} \\ &+ \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\ &+ \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} \nu_{1,t-1} \\ \nu_{2,t-1} \end{pmatrix} + \begin{pmatrix} \nu_{1,t} \\ \nu_{2,t} \end{pmatrix} \end{aligned} \quad (5.3)$$

Historical Decomposition

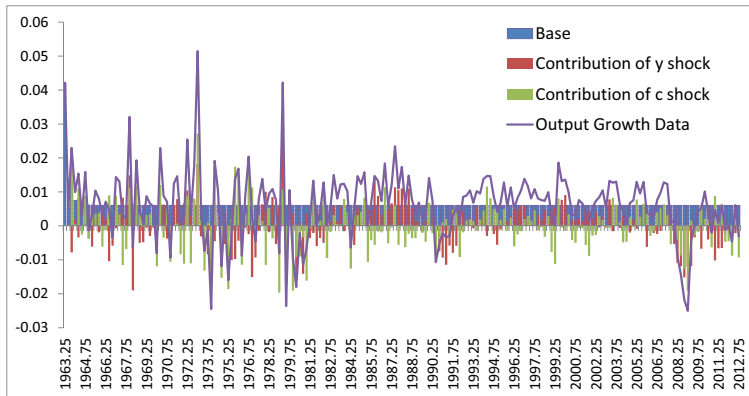
- Historical decomposition measures the contribution of each shock to the level of each series:

$$\begin{aligned} \begin{pmatrix} y_t \\ x_t \end{pmatrix} &= \overbrace{\sum_{i=0}^{T-1} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}^i \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}}^{\text{Base or trend}} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}^T \begin{pmatrix} y_0 \\ x_0 \end{pmatrix} \\ &+ \underbrace{\sum_{i=0}^{T-1} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}^i \begin{pmatrix} \nu_{1,t-i} \\ \nu_{2,t-i} \end{pmatrix}}_{\text{Contributions of the 2 shocks}} \end{aligned} \quad (5.4)$$

Historical Decomposition

- Historical decomposition presented as bar chart where the bars represent contribution by the shocks and the trend

Figure : Example for Historical Decomposition



Plan for next session: Identification

- All the above tools make sense only if the shocks are uncorrelated/structural
- Identification is needed to “make structural shocks out of” reduced-form residuals
- Structural VARs make various identification assumptions to achieve meaningful decompositions

- Andrew P Blake and Haroon Mumtaz. *Applied Bayesian econometrics for central bankers*. Number 4 in Technical Books. Centre for Central Banking Studies, Bank of England, 2012. URL <http://ideas.repec.org/b/ccb/tbooks/4.html>.
- Fabio Canova. *Methods for Applied Macroeconomic Research*. Princeton University Press, 2007.
- Juan F. Rubio-Ramirez, Daniel F. Waggoner, and Tao Zha. Structural vector autoregressions: Theory of identification and algorithms for inference. *Review of Economic Studies*, 77(2):665–696, 2010. URL <http://ideas.repec.org/a/oup/restud/v77y2010i2p665-696.html>.
- Christopher A Sims. Macroeconomics and reality. *Econometrica*, 48(1):1–48, January 1980. URL <http://ideas.repec.org/a/ecm/emetrp/v48y1980i1p1-48.html>.
- James H. Stock and Mark W. Watson. Vector Autoregressions. *Journal of Economic Perspectives*, 15(4):101–115, Fall 2001. URL <http://ideas.repec.org/a/aea/jecper/v15y2001i4p101-115.html>.