Structural vector autoregressive models

Kevin Kotzé

Contents

Introduction

- SVAR models allow for:
  - contemporaneous variables that may be treated as explanatory variables
  - specific restrictions on the parameters in the coefficient and residual covariance matrices

- Allowing for contemporaneous variables is important in many economic studies, where we often deal with quarterly data

- Allows for the identification of specific independent shocks that are not affected by covariance terms

Introduction

- With the VAR model, errors must have positive definite covariance matrix
- This leads to difficulties when trying to evaluate the effect of an independent shock
- SVAR models become an indispensable tool for studying relationships and the effects of shocks in macroeconomics

Incorporating contemporaneous variables

- Start off by assuming that each variable is symmetrical

- For the two variable case let,
  - $y_{1,t}$ be affected by current and past realizations of $y_{2,t}$
  - $y_{2,t}$ be affected by current and past realizations of $y_{1,t}$

$$
\begin{aligned}
y_{1,t} &= b_{10} - b_{12} y_{2,t} + \gamma_{11} y_{1,t-1} + \gamma_{12} y_{2,t-1} + \varepsilon_{1,t} \\
y_{2,t} &= b_{20} - b_{21} y_{1,t} + \gamma_{21} y_{1,t-1} + \gamma_{22} y_{2,t-1} + \varepsilon_{2,t}
\end{aligned}
$$

  - where both $y_{1,t}$ and $y_{2,t}$ are stationary
  - $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are white noise with σ_1 and σ_2 std
  - $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are uncorrelated, since we want to identify the effect of each independent shock
  - Hence covariance elements in Σ_ε are set to zero

- Note: b_{12} describes the contemporaneous effect of a change in $y_{2,t}$ on $y_{1,t}$ and vice versa for b_{21}

Incorporating contemporaneous variables

- Given the model:

$$
\begin{aligned}
y_{1,t} &= b_{10} - b_{12}y_{2,t} + \gamma_{11}y_{1,t-1} + \gamma_{12}y_{2,t-1} + \varepsilon_{1,t} \\
y_{2,t} &= b_{20} - b_{21}y_{1,t} + \gamma_{21}y_{1,t-1} + \gamma_{22}y_{2,t-1} + \varepsilon_{2,t}
\end{aligned}
$$

  - There will be an indirect contemporaneous effect of $\varepsilon_{1,t}$ on $y_{2,t}$ if $b_{21} \neq 0$
  - Similarly, $\varepsilon_{2,t}$ affects $y_{1,t}$ if $b_{12} \neq 0$

- Much richer characterisation of dynamics than in previous lecture
  - In previous model, $\varepsilon_{2,t}$ could only affect $y_{1,t-1}$, and v.v.

- However, the inclusion of contemporaneous parameters does present some challenges with parameter estimation

Standard VAR: Structural Form

- To express the above *structural-form* of the model as a *reduced-form* expression:

$$B\boldsymbol{y}_t = \Gamma_0 + \Gamma_1 \boldsymbol{y}_{t-1} + \varepsilon_t$$

- where

$$B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}, \quad \boldsymbol{y}_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix}, \quad \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}$$

$$\Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \quad \text{and} \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

Standard VAR: Reduced-Form

- Premultiplication by B^{-1} gives us the VAR in *reduced-form*:

$$\boldsymbol{y}_t = A_0 + A_1 \boldsymbol{y}_{t-1} + \boldsymbol{u}_t$$

- where $A_0 = B^{-1}\Gamma_0$, $A_1 = B^{-1}\Gamma_1$ and $\boldsymbol{u}_t = B^{-1}\varepsilon_t$
- Now where:
  - a_{i0} is the i element in A_0
  - a_{ij} is row i column j of matrix A_1
  - $\boldsymbol{u}_t$ has elements $u_{1,t}$ and $u_{2,t}$

$$\begin{aligned} y_{1,t} &= a_{10} + a_{11} y_{1,t-1} + a_{12} y_{2,t-1} + u_{1,t} \\ y_{2,t} &= a_{20} + a_{21} y_{1,t-1} + a_{22} y_{2,t-1} + u_{2,t} \end{aligned}$$

Standard VAR: Reduced-Form

- By using the relationship $\boldsymbol{u}_t = B^{-1}\varepsilon_t$, or:

$$\begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_{y,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

- We can show that,

$$u_{1,t} = (\varepsilon_{1,t} - b_{12}\varepsilon_{2,t})/(1 - b_{12}b_{21})$$
$$u_{2,t} = (\varepsilon_{2,t} - b_{21}\varepsilon_{1,t})/(1 - b_{12}b_{21})$$

Standard VAR: Variance/covariance

- Since $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are white noise processes
  - The residuals $u_{1,t}$ and $u_{2,t}$ have zero means, constant variances, and have little autocorrelation
  - However, as $\boldsymbol{u}_t$ is dependent upon both $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$, there may be some evidence of covariation

- The covariance of the two terms is:

$$\begin{aligned}\textsf{cov}\left[u_{1,t}, u_{2,t}\right] &= \mathbb{E}\left[(\varepsilon_{1,t} - b_{12}\varepsilon_{2,t})(\varepsilon_{2,t} - b_{21}\varepsilon_{1,t})\right] / (1 - b_{12}b_{21})^2 \\ &= -\left[(b_{21}\sigma_1^2 + b_{12}\sigma_2^2)\right] / (1 - b_{12}b_{21})^2\end{aligned}$$

- Since they are all time invariant, the variance/covariance matrix will be,

$$\Sigma_{\boldsymbol{u}} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

- where $\textsf{var}\left[u_{i,t}\right] = \sigma_{ii}$ and $\sigma_{12} = \sigma_{21} = \textsf{cov}\left[u_{1,t}, u_{2,t}\right]$

Estimation

- Note that in the *Reduced-Form*:
    - RHS contains only predetermined variables
    - Error terms are serially uncorrelated with constant variance
- Hence we can use OLS - consistent and asymptotically efficient

Identification

- The structural equations can't be estimated directly (due to feedback effects from contemporaneous variables)
  - However, we can estimate the *reduced-form* of the VAR model
  - This would allow for us to obtain the residuals $u_{1,t}$ and $u_{2,t}$ and the coefficients in the A_0 and A_1 matrices
  - Could we use these to recover the *structural-form* parameter estimates given the relationships between the structural and reduced forms?

Identification

- Unfortunately not, since the *structural-form* contains 10 parameters:
  - $b_{10}, b_{20}, \gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}, b_{12}, b_{21}, \sigma_1, \sigma_2$
- while the *reduced-form* contains 9 parameters:
  - $a_{10}, a_{20}, a_{11}, a_{12}, a_{21}, a_{22}$, var$[u_{1,t}]$, var$[u_{2,t}]$, cov$[u_{1,t}, u_{2,t}]$
- And there is no mapping that enables us to obtain the *structural-form* parameters from the *reduced-form* parameters

Identification

- However, it may be possible to show that:
  - If one variable in the *structural-form* is restricted to a calibrated value then the structural system could be exactly identified?????

Recursive estimation

- Consider the method of recursive estimation (Sims, 1980)
  - Suppose that you are willing to assume that $b_{21} = 0$ in the structural system:

$$
\begin{aligned}
y_{1,t} &= b_{10} - b_{12} y_{2,t} + \gamma_{11} y_{1,t-1} + \gamma_{12} y_{2,t-1} + \varepsilon_{1,t} \\
y_{2,t} &= b_{20} \qquad\qquad\; + \gamma_{21} y_{1,t-1} + \gamma_{22} y_{2,t-1} + \varepsilon_{2,t}
\end{aligned}
$$

$$
\text{such that} \;\; B^{-1} = \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix}
$$

- Premultiplying by B^{-1} yields

$$
\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} b_{10} - b_{12} b_{20} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} - b_{12}\gamma_{21} & \gamma_{12} - b_{12}\gamma_{22} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} .
$$

$$
\begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} - b_{12}\varepsilon_{2,t} \\ \varepsilon_{2,t} \end{bmatrix}
$$

Recursive estimation

- Take note of the previous expression:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \cdots + \begin{bmatrix} \varepsilon_{1,t} - b_{12}\varepsilon_{2,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

- Hence, by setting $b_{21} = 0$, the shocks from $\varepsilon_{1,t}$ do not effect contemporaneous values of $y_{2,t}$
- However both $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ affect $y_{1,t}$
- Note also that $\varepsilon_{1,t-1}$ could still influence $y_{2,t}$ through its effect on $y_{1,t-1}$
- Furthermore, by returning to the relationship $\boldsymbol{u}_t = B^{-1}\varepsilon_t$,

$$\begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

- We have $\varepsilon_{2,t} = u_{1,t}$, and using $b_{12} = -\text{cov}[u_{1,t}, u_{2,t}]/\sigma_2^2$, which allows us to get $\varepsilon_{1,t} = b_{12}\varepsilon_{2,t} + u_{1,t}$

Mapping the reduced to structural form

- From the reduced form (where all the coefficient matrices are premultiplied by B^{-1});

$$
\begin{aligned}
y_{1,t} &= a_{10} + a_{11} y_{1,t-1} + a_{12} y_{2,t-1} + u_{1,t} \\
y_{2,t} &= a_{20} + a_{21} y_{1,t-1} + a_{22} y_{2,t-1} + u_{2,t}
\end{aligned}
$$

$$
\begin{aligned}
a_{10} &= b_{10} - b_{12} b_{20} & a_{11} &= \gamma_{11} - b_{12}\gamma_{21} \\
a_{12} &= \gamma_{12} - b_{12}\gamma_{22} & a_{20} &= b_{20} \\
a_{21} &= \gamma_{21} & a_{22} &= \gamma_{22}
\end{aligned}
$$

$$
\begin{aligned}
\text{var}[u_1] &= \sigma_1^2 + b_{12}^2 \sigma_2^2 \\
\text{var}[u_2] &= \sigma_2^2 \\
\text{cov}[u_1, u_2] &= -b_{12}\sigma_2^2
\end{aligned}
$$

Cholesky decomposition

- In the above example, we were able to recover the $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ sequences use the relationship $u_{1,t} = \varepsilon_{1,t} - b_{12}\varepsilon_{2,t}$ and $u_{2,t} = \varepsilon_{2,t}$
  - When $b_{21} = 0$, $y_{1,t}$ does not have a contemporaneous effect on $y_{2,t}$ and $\varepsilon_{1,t}$ does not affect $y_{2,t}$
  - Observed values of $u_{2,t}$ are attributed to pure shocks in $y_{2,t}$
  - This procedure of setting the the lower triangle of the B coefficient matrix equal to zero is termed applying the Cholesky decomposition
  - It turns out that the number of restrictions that we need to impose is equivalent to the number of terms in the lower (or upper) triangle of the B matrix, which is $[(K^2 - K)/2]$
  - The alternative ordering of the Cholesky decomposition is to let $b_{12} = 0$ (i.e. the upper triangle)

IRF: MA representation

- In many cases it is useful to express a $AR(p)$ process as a $MA(q)$ process
  - For example, the stationary univariate $AR(1)$ model:

$$y_t = \phi y_{t-1} + \varepsilon_t$$

  - has the $MA(\infty)$ representation,

$$y_t = \sum_{i=0}^{\infty} \theta_i \varepsilon_{t-i}$$

- This representation is particularly useful for calculating impact multipliers and impulse response functions

VMA representation

- Just as every stable $AR(p)$ has a $MA(q)$ representation; every $VAR(p)$ has a $VMA(q)$ representation

- From;

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

- Where μ_1 and μ_2 are mean values for $y_{1,t}$ and $y_{2,t}$;

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \cdot \begin{bmatrix} u_{1,t-i} \\ u_{2,t-i} \end{bmatrix}$$

VMA representation

- Now since, $\boldsymbol{u}_t = B^{-1}\varepsilon_t$, and where,

$$B^{-1} = \frac{1}{\det}\begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} = \frac{1}{1 - b_{12}b_{21}}\begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$$

- We have:

$$\begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} = \frac{1}{1 - b_{12}b_{21}} \sum_{i=0}^{\infty} \cdot \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

- such that the SVAR model can be written as,

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \frac{1}{1 - b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \cdot \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-i} \\ \varepsilon_{2,t-i} \end{bmatrix}$$

- This expression may be used to describe the effect of a shock in ε_t on the endogenous variables

VMA representation

- The impact multipliers, which describe the effect of shocks on the endogenous variables, are summarised in matrix Θ_i

$$\Theta_i = \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}_i = \frac{a_1^i}{1 - b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$$

- where $\mu = [\mu_1 \; \mu_2]'$ and $\boldsymbol{y}_t = [y_{1,t} \; y_{2,t}]'$ we are left with,

$$\boldsymbol{y}_t = \mu + \sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i}$$

- This is a particularly useful expression, as the Θ_i matrix describes the effects of the shocks, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ on the entire paths of $y_{1,t}$ and $y_{2,t}$

VMA representation

- For example, where the numbers in brackets refer to the lags of $\theta_{jk}(i)$:
  - $\theta_{12}(0)$ is the instant impact of 1 unit change in $\varepsilon_{2,t}$ on $y_{1,t}$
  - $\theta_{11}(1)$ is the instant impact of 1 unit change in $\varepsilon_{1,t-1}$ on $y_{1,t}$
  - $\theta_{12}(1)$ is the instant impact of 1 unit change in $\varepsilon_{2,t-1}$ on $y_{1,t}$

Impulse response functions

- The impact multipliers $\theta_{11}(i), \theta_{12}(i), \theta_{21}(i)$ and $\theta_{22}(i)$ are used to generate the impulse response functions for different values of i
  - Visually represent the behaviour of $y_{1,t}$ and $y_{2,t}$ in response to various shocks, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$
- To avoid the problem of an under-identified system we use the Cholesky decomposition;

$$\begin{aligned} u_{1,t} &= \varepsilon_{1,t} - b_{12}\varepsilon_{2,t} \\ u_{2,t} &= \varepsilon_{2,t} \end{aligned}$$

  - Note that all the errors from $u_{2,t}$ are attributed to $\varepsilon_{2,t}$
  - We can then find $\varepsilon_{1,t}$ using b_{12}, $u_{1,t}$ and $\varepsilon_{1,t}$
- Although the Cholesky decomposition constrains the system such that $\varepsilon_{1,t}$ has no direct effect on $y_{2,t}$, you should note that lagged values of $y_{1,t}$ affect the contemporaneous value of $y_{2,t}$

Ordering of Cholesky decomposition

- The ordering of the Cholesky decomposition (i.e. whether to set b_{12} or b_{21} to 0) depends on the magnitude of the correlation between $u_{1,t}$ and $u_{2,t}$

- When $\rho_{12} = \sigma_{12} / \left(\sqrt{\sigma_{11}} \sqrt{\sigma_{22}} \right)$;
    - If the correlation is zero then ordering is immaterial
    - If the correlation is unity then it is inappropriate to attribute the shock to a single source
    - If the correlation is between 0 and 1 then you usually need to consider both ordering - if the results are different then you need to investigate further

- Try where possible to relate ordering to theoretical consideration. (i.e. shock to the US exchange rate may affect SA exchange rate immediately, but not the other way around)

Impulse response functions

- Note that with zero off-diagonal elements in the variance-covariance matrix we could consider the effect of independent shocks
- Or alternatively we could order the variables from most exogenous to most endogenous when using a Cholenski decomposition

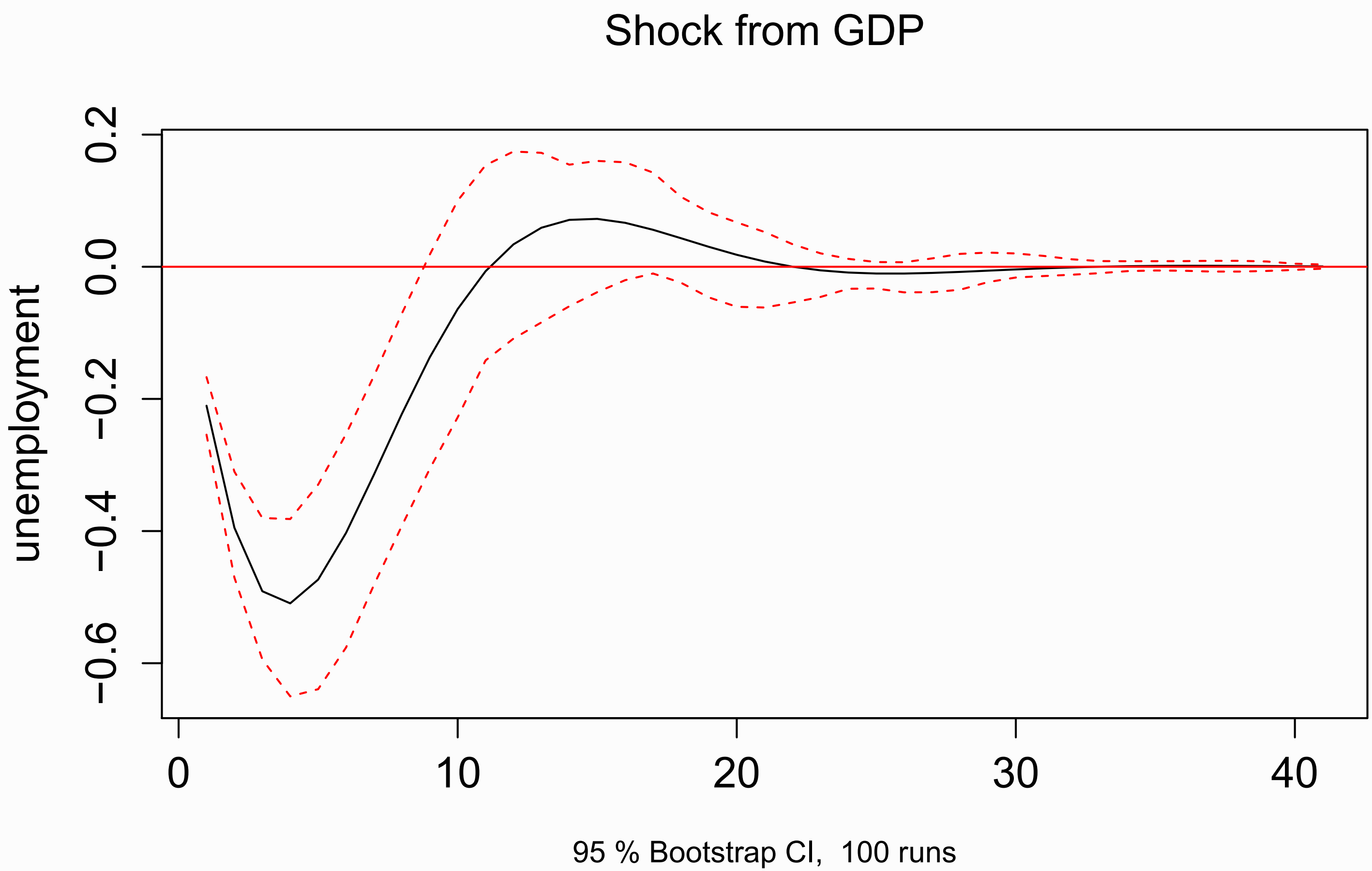

Figure : IRF - unemployment shock on output

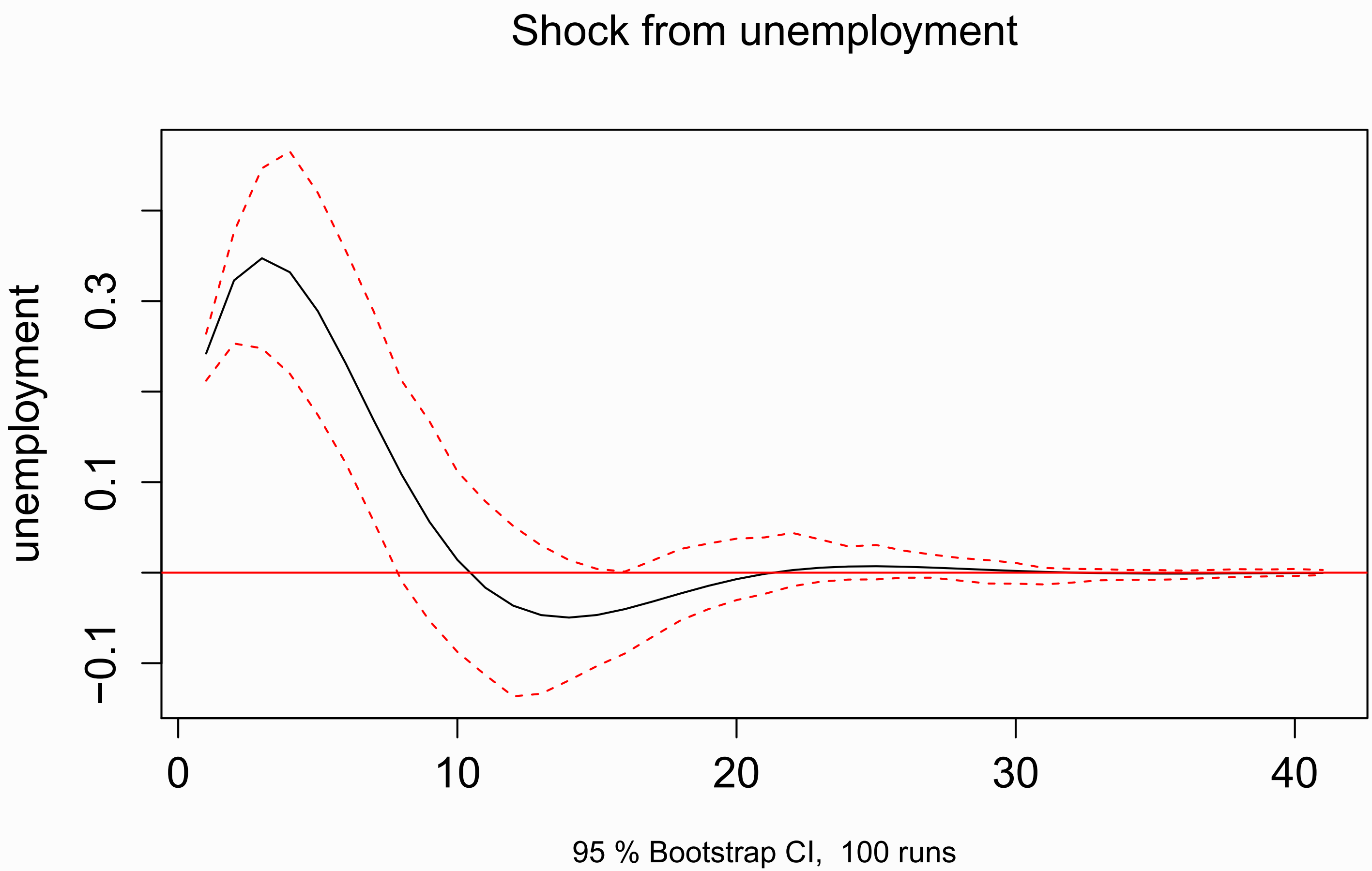

Figure : IRF - unemployment shock on unemployment

Variance Decompositions

- If you knew the coefficients of A_0 and A_1 and wanted to forecast values of $\boldsymbol{y}_{t+h}$ conditional on $\boldsymbol{y}_t$
  - The conditional expectation of $\boldsymbol{y}_{t+1}$ is

$$\mathbb{E}_t[\boldsymbol{y}_{t+1}] = A_0 + A_1\boldsymbol{y}_t$$

- and the conditional expectation of $\boldsymbol{y}_{t+2}$ is

$$\mathbb{E}_t[\boldsymbol{y}_{t+2}] = [I + A_1]A_0 + A_1^2\boldsymbol{y}_t$$

- such that the conditional expectation of $\boldsymbol{y}_{t+H}$ is

$$\mathbb{E}_t[\boldsymbol{y}_{t+H}] = [I + A_1 + A_1^2 + \ldots + A_1^{H-1}]A_0 + A_1^H\boldsymbol{y}_t$$

Variance Decompositions: Forecast errors

- One-step ahead forecast error is $\left(\boldsymbol{y}_{t+1} - \mathbb{E}_t[\boldsymbol{y}_{t+1}]\right)$
- This equals $\boldsymbol{u}_{t+1}$, since $\mathbb{E}_t[\mathbf{y}_{t+1}] = A_0 + A_1\boldsymbol{y}_t$ and $\boldsymbol{y}_{t+1} = A_0 + A_1\boldsymbol{y}_t + \boldsymbol{u}_{t+1}$
- Two-step ahead forecast error is $\left(\boldsymbol{u}_{t+2} + A_1\boldsymbol{u}_{t+1}\right)$
- H-step ahead forecast error is $\left(\boldsymbol{u}_{t+H} + A_1\boldsymbol{u}_{t+H-1} + A_1^2\boldsymbol{u}_{t+H-2} + \ldots + A_1^{H-1}\boldsymbol{u}_{t+1}\right)$
- Of course it is possible to write the forecast errors in terms of the *structural-form* errors, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$
- The forecast error variance decomposition tells us the proportion of the expected variance in a variable that is due to each of the shocks in the model
  - If $\varepsilon_{2,t}$ explains none of the forecast error variance of $y_{1,t}$; then $y_{1,t}$ is exogenous as it evolves independent of $\varepsilon_{2,t}$ and $y_{2,t}$
  - If $\varepsilon_{2,t}$ explains all the forecast error variance of $y_{1,t}$; then $y_{1,t}$ is entirely endogenous

Variance Decomposition

- Variance decomposition also has identification problems (as per above)
  - Cholesky decomposition necessitates that all one period forecast error of $y_{2,t}$ is due to $\varepsilon_{2,t}$
  - Similarly for alternate ordering

- It is often useful to examine the variance decompositions at different horizons
  - as H increases the decompositions should converge

- Analysis of impulse responses and variance decompositions may be termed innovation accounting

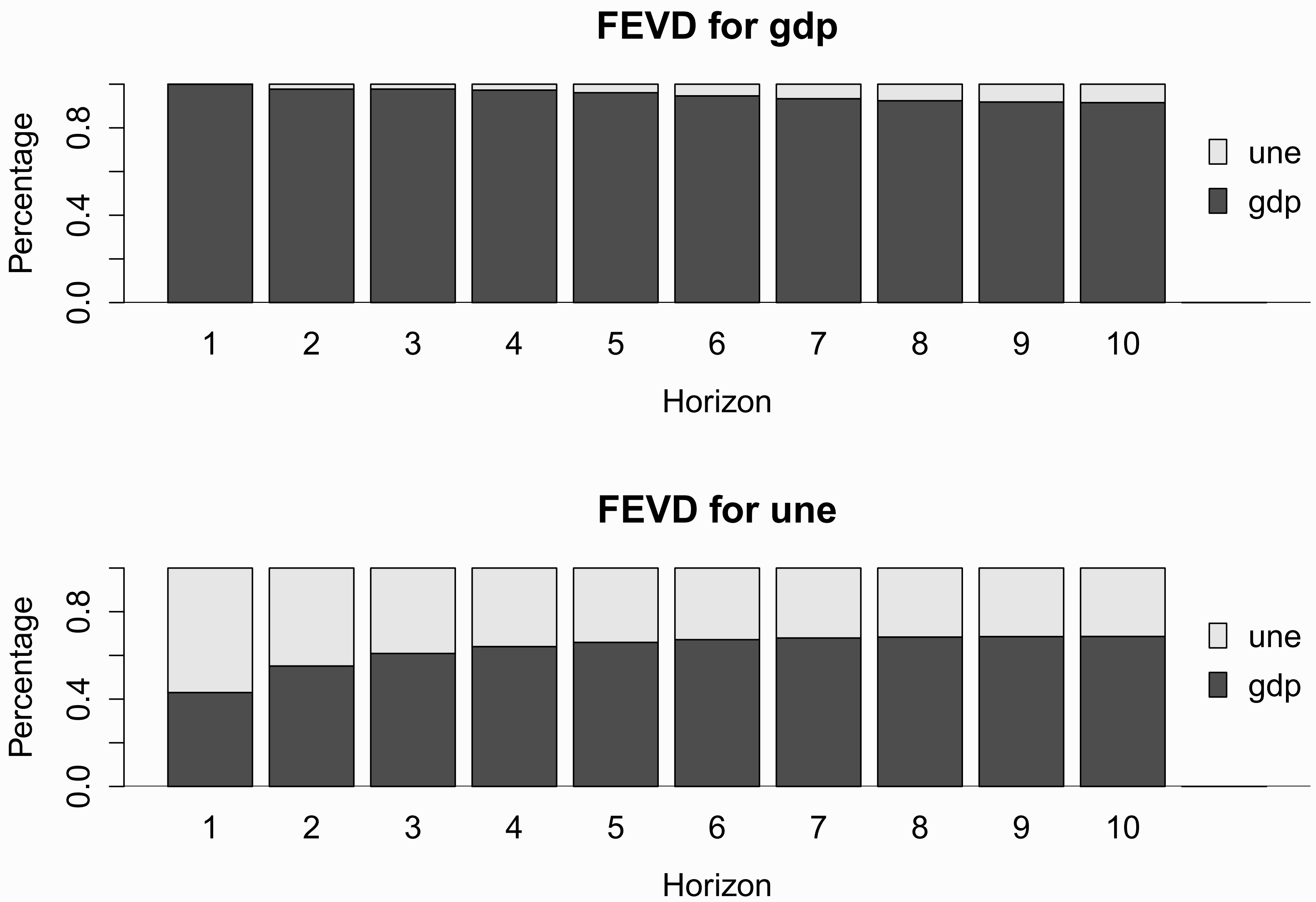

Figure : Variance Decomposition

Structural Decomposition

- In a three variable model, where $C = B^{-1}$ the Cholesky decomposition would suggest,

$$\begin{aligned} u_{1,t} &= \varepsilon_{1,t} \\ u_{2,t} &= c_{21}\varepsilon_{1,t} + \varepsilon_{2,t} \\ u_{3,t} &= c_{31}\varepsilon_{1,t} + c_{32}\varepsilon_{2,t} + \varepsilon_{3,t} \end{aligned}$$

- Sims (1986) and Bernanke (1986) provide examples of theoretical restrictions that may differ from the upper or lower triangle
  - Involves estimating the relationships among the structural shocks using an economic model
  - For example, they would consider the decomposition,

$$\begin{aligned} u_{1t} &= \varepsilon_{1t} + c_{13}\varepsilon_{3t} \\ u_{2t} &= c_{21}\varepsilon_{1t} + \varepsilon_{2t} \\ u_{3t} &= c_{31}\varepsilon_{2t} + \varepsilon_{3t} \end{aligned}$$

Structural Decomposition

- Note that with this structural decomposition:
  - We have lost the triangular structure
  - where each variable is affected by its own structural innovation and the structural innovation in one other variable
  - The condition for $(K^2 - K)/2$ restrictions is satisfied, so the conditions for exact identification are maintained

Example of identifying restrictions

- Suppose that we have a 2 variable model with a sample size of 5
- This gives us 5 residuals for $u_{1,t}$ and $u_{2,t}$

	1	2	3	4	5
$u_{1,t}$	1.0	-0.5	0.0	-1.0	0.5
$u_{2,t}$	0.5	-1.0	0.0	-0.5	1.0

- Note that both $u_{1,t}$ and $u_{2,t}$ sum to zero
- $\sigma_1 = 0.5, \sigma_{12} = \sigma_{21} = 0.4, \text{ and } \sigma_2 = 0.5$, which gives a variance/covariance

$$\Sigma_{\boldsymbol{u}} = \begin{bmatrix} 0.5 & 0.4 \\ 0.4 & 0.5 \end{bmatrix}$$

Example of identifying restrictions

- Since we premultiplied ε_t by B^{-1} to get $\boldsymbol{u}_t$
- We can derive values for Σ_ε from $\Sigma_{\boldsymbol{u}}$ as

$$\Sigma_\varepsilon = B\Sigma_{\boldsymbol{u}}B'$$

- Hence,

$$\begin{bmatrix} \mathsf{var}(\varepsilon_1) & 0 \\ 0 & \mathsf{var}(\varepsilon_2) \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} 0.5 & 0.4 \\ 0.4 & 0.5 \end{bmatrix} \begin{bmatrix} 1 & b_{21} \\ b_{12} & 1 \end{bmatrix}$$

Example of identifying restrictions

- This leaves us with,

$$
\begin{aligned}
\mathsf{var}(\varepsilon_1) &= 0.5 + 0.8b_{12} + 0.5b_{12}^2 \\
0 &= 0.5b_{21} + 0.4b_{21}b_{12} + 0.4 + 0.5b_{12} \\
0 &= 0.5b_{21} + 0.4b_{21}b_{12} + 0.4 + 0.5b_{12} \\
\mathsf{var}(\varepsilon_2) &= 0.5b_{21}^2 + 0.8b_{21} + 0.5
\end{aligned}
$$

- Since the middle lines are identical we have 3 independent equations to solve for 4 unknowns

Identification: Cholesky decomposition

- When $b_{12} = 0$ we have,

$$\begin{aligned} \mathsf{var}(\varepsilon_1) &= 0.5 \\ 0 &= 0.5b_{21} + 0.4 \text{ s.t. } \; b_{21} = -0.8 \\ 0 &= 0.5b_{21} + 0.4 \text{ s.t. } \; b_{21} = -0.8 \\ \mathsf{var}(\varepsilon_2) &= 0.5b_{21}^2 + 0.8b_{21} + 0.5 = 0.18 \end{aligned}$$

- Since $\varepsilon_{1,t} = u_{1,t}$ and $\varepsilon_{2,t} = -0.8u_{1,t} + u_{2,t}$

	1	2	3	4	5
$\varepsilon_{1,t}$	1.0	-0.5	0.0	-1.0	0.5
$\varepsilon_{2,t}$	-0.3	-0.6	0.0	0.3	0.6

Alternative identification restrictions

- If one shock, $\varepsilon_{2,t}$ has a one-for-one affect on $y_{1,t}$ s.t. $b_{12} = 1$

$$\mathsf{var}(\varepsilon_1) = 0.5 + 0.8 b_{12} + 0.5 b_{12}^2 = 1.8$$
$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

- From which we could derive ε_t

Alternative identification restrictions

- Although there is little theory that informs us on the variance of shocks
- If it is given that **var**$(\varepsilon_1) = 1.8$ we could work out values for b_{12}

$$\textbf{var}(\varepsilon_1) = 1.8 = 0.5 + 0.8b_{12} + 0.5b_{12}^2$$
$$\vdots \quad \vdots \quad \vdots$$

- From which we could derive ε_t

Alternative identification restrictions

- If we assume that $b_{12} = b_{21}$
- Then replacing b_{21} with b_{12} in the following

$$
\begin{aligned}
0 &= 0.5b_{21} + 0.4b_{21}b_{12} + 0.4 + 0.5b_{12} \\
&\vdots \qquad\qquad\qquad\qquad \vdots
\end{aligned}
$$

- Allows us to derive values for b_{12} and we can then solve for the rest

Long-run restrictions

- Suggested that economic theory does not always provide enough meaningful contemporaneous restrictions
- As an alternative we could impose restrictions on the long-run properties of shocks, allowing for the neutrality of the effects of certain shocks over time
- Blanchard & Quah (1989) consider the use of such restriction on a model for output (demand) and unemployment (supply)
- This bivariate VAR would need a single restriction
- Suggested that output growth and unemployment were driven by two orthogonal structural shocks
- Demand side shocks have a temporary effect on real GNP
- Supply side productivity shocks have a permanent effect on real GNP
- Rate of unemployment is considered stationary, so no shock could change unemployment permanently

Decomposition using Blanchard-Quah

- If the logarithm of output, $y_{1,t}$, is $I(1)$ then output growth, $\Delta y_{1,t}$, is $I(0)$
- Assume rate of unemployment, $y_{2,t}$, is affected by the same variables and is $I(0)$
- The bivariate moving average representation, where $\boldsymbol{y}_t$ is a vector of both variables is

$$\boldsymbol{y}_t = \sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i}$$

Decomposition using Blanchard-Quah

- Which may be expanded as

$$\begin{bmatrix} \Delta y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \theta_{11}(0) & \theta_{12}(0) \\ \theta_{21}(0) & \theta_{22}(0) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} + \ldots$$
$$\begin{bmatrix} \theta_{11}(1) & \theta_{12}(1) \\ \theta_{21}(1) & \theta_{22}(1) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix} + \ldots$$

- where the effect of $\varepsilon_{1,t-1}$ on $\Delta y_{1,t}$ is summarized by $\theta_{11}(1)$

Long-run restrictions

- Now, if $\varepsilon_{1,t}$ has no long-run cumulative impact on $\Delta y_{1,t}$ we could impose the restriction

$$\sum_{i=0}^{\infty} \theta_{11}(i) = 0$$

- which may be included in the coefficient matrix for the moving average representation,

$$\sum_{i=0}^{\infty} \Theta_i = \begin{bmatrix} 0 & \sum_{i=0}^{\infty} \theta_{12,i} \\ \sum_{i=0}^{\infty} \theta_{21,i} & \sum_{i=0}^{\infty} \theta_{22,i} \end{bmatrix} = \sum_{i=0}^{\infty} \begin{bmatrix} 0 & \theta_{12}(i) \\ \theta_{21}(i) & \theta_{22,}(i) \end{bmatrix}$$

Restrictions

- Hence, we can impose restrictions on either the short-run contemporaneous parameters, or the long-run moving average components
- Alternatively we could use a combination of the two
- The only condition is that the number of restrictions must equal $[(K^2 - K)/2]$

Limitations of the VAR approach

- A major limitation of the traditional VAR approach is that it is highly parametrised
- In addition all of the effects of omitted variables will be contained in the residuals
- This may lead to major distortions in the impulse responses, making them of little use for structural interpretations
- Measurement errors or mis-specifications of the model make interpretation of the impulse responses difficult
- We can't make use of an infinite number of MA coefficients, since the dataset is finite (this may lead to a bias in the parameter estimates)

Summary

- Sims (1980) introduced SVAR models as an alternative to the large-scale macroeconometric models that were used during that time
- The SVAR methodology has gained widespread use in applied time series research
- Allows for the incorporation of contemporaneous variables and an investigation into the impact of individual shocks
- To identify the structural VAR model, we need to impose restrictions
- Widely-used identification methods rely on short-run or long-run restrictions
- The short-run restrictions were originally suggested by Sims (1986)
- Blanchard & Quah (1989) introduced long-run restrictions

Summary

- A system of K variables would require that we impose $(K^2 - K)/2$ identifying restrictions for exact identification
- The use of the Cholesky decomposition would ensure that the identified shocks from the VAR model will be orthogonal (uncorrelated) and unique
- However, the choice of the this method for imposing restrictions could affect the results of the model
- An impulse response function describes how a given (structural) shock affects a variable over time
- The forecast error variance decomposition attributes the forecast error variance to specific structural shocks at different horizons