Impulse Response Functions

Wouter J. Den Haan

University of Amsterdam

April 28, 2011

General definition IRFs

- The IRF gives the j^{th}-period response when the system is shocked by a one-standard-deviation shock.
- Suppose

$$y_t = \rho y_{t-1} + \varepsilon_t \text{ and } \varepsilon_t \text{ has a variance equal to } \sigma^2$$

- Consider a sequence of shocks $\{\bar{\varepsilon}_t\}_{t=1}^{\infty}$. Let the generated series for y_t be given by $\{\bar{y}_t\}_{t=1}^{\infty}$.
- Consider an alternative series of shocks such that

$$\tilde{\varepsilon}_t = \begin{cases} \bar{\varepsilon}_t + \sigma & \text{if } t = \tau \\ \bar{\varepsilon}_t & \text{o.w.} \end{cases}$$

- The IRF is then defined as

$$IRF(j) = \tilde{y}_{\tau-1+j} - \bar{y}_{\tau-1+j}$$

IRFs for linear processes

- Linear processes: The IRF is independent of the particular draws for $\bar{\varepsilon}_t$
- Thus we can simply start at the steady state (that is when $\bar{\varepsilon}_t$ has been zero for a very long time)
- $$IRF(j) = \sigma\rho^{j-1}$$
- Often you cannot get an analytical formula for the impulse response function, but simple iteration on the law of motion (driving process) gives you the exact same answer
- Note that the IRF is not stochastic

IRFs in theoretical models

- When you have solved for the policy functions then it is trivial to get the IRFs by simply giving the system a one standard deviation shock and iterating on the policy functions.
- Shocks in the model are *structural* shocks, such as
  - productivity shock
  - preference shock
  - monetary policy shock

IRFs in empirical models

What we are going to do?

- Describe an empirical model that has turned out to be very useful (for example for forecasting)
  - Reduced form VAR
- Make clear we do not have to worry about variables being I(1)
- Describe a way to back out structural shocks (this is the hard part)
  - Structural VAR

Reduced Form Vector AutoRegressive models (VARs)

- Let y_t be an $n \times 1$ vector of n variables (typically in logs)

$$y_t = \sum_{j=1}^{J} A_j y_{t-j} + u_t$$

where A_j is an $n \times n$ matrix.

- This system can be estimated by OLS (equation by equation) even if y_t contains $I(1)$ variables
- constants and trend terms are left out to simplify the notation

Spurious regression

- Let z_t and x_t be $I(1)$ variables that have nothing to do with each other
- Consider the regression equation

$$z_t = ax_t + u_t$$

- The least-squares estimator is given by

$$\hat{a}_T = \frac{\sum_{t=1}^{T} x_t z_t}{\sum_{t=1}^{T} x_t^2}$$

- Problem:

$$\lim_{T \longrightarrow \infty} \hat{a}_T \neq 0$$

Source of spurious regressions

- The problem is not that z_t and x_t are I(1)
- The problem is that there is not a single value for a such that u_t is stationary
- If z_t and x_t are cointegrated then there is a value of a such that

$$z_t - a x_t \text{ is stationary}$$

- Then least-squares estimates of a are consistent
- but you have to change formula for standard errors

How to avoid spurious regressions?

Answer: Add enough lags.

- Consider the following regression equation

$$z_t = ax_t + bz_{t-1} + u_t$$

- Now there are values of the regression coefficients so that u_t is stationary, namely

$$a = 0 \text{ and } b = 1$$

- So as long as you have enough lags in the VAR you are fine (but be careful with inferences)

Structural VARs

Consider the reduced-form VAR

$$y_t = \sum_{j=1}^{J} A_j y_{t-j} + u_t$$

- For example suppose that y_t contains
  - the interest rate set by the central bank
  - real GDP
  - residential investment
- What affects
  - the error term in the interest rate equation?
  - the error term in the output equation?
  - the error term in the housing equation?

Structural shocks

- Suppose that the economy is being hit by "structural shocks", that is shocks that are not responses to economic events
- Suppose that there are 10 structural shocks. Thus

$$u_t = Be_t$$

  where B is a 3×10 matrix.
- Without loss of generality we can assume that

$$\mathsf{E}[e_t e_t'] = I$$

Structural shocks

- Can we identify B from the data?

$$\mathsf{E}[u_t u_t'] = B\mathsf{E}[e_t e_t']B' = BB'$$

- We can get an estimate for $\mathsf{E}[u_t u_t']$ using

$$\hat{\Sigma} = \sum_{t=J+1}^{T} \hat{u}_t \hat{u}_t' / (T - J)$$

- But B has 30 and $\hat{\Sigma}$ only 9 elements.

Identification of B

- Can we identify B if there are only three structural shocks?
- B has 9 distinct elements
- $\hat{\Sigma}$ is symmetric
- Not all equations are independent. $\Sigma_{1,2} = \Sigma_{2,1}$. For example

$$\Sigma_{1,2} = b_{11}b_{21} + b_{12}b_{22} + b_{13}b_{23}$$

but also

$$\Sigma_{2,1} = b_{21}b_{11} + b_{22}b_{12} + b_{23}b_{13}$$

- In other words, different B matrices lead to the same Σ matrix

Identification of B

- We need additional identification assumptions

$$\begin{bmatrix} u_t^i \\ u_t^y \\ u_t^r \end{bmatrix} = B \begin{bmatrix} e_t^1 \\ e_t^2 \\ e_t^{\mathsf{mp}} \end{bmatrix}$$

- And suppose we impose

$$B = \begin{bmatrix} & 0 & 0 \\ & & 0 \\ & & \end{bmatrix}$$

- Then I can solve for the remaining elements of B from

$$\hat{B}'\hat{B} = \hat{\Sigma}$$

- In Matlab use B=chol(S)′

Identification of B

- Suppose instead we use

$$\begin{bmatrix} u_t^y \\ u_t^i \\ u_t^r \end{bmatrix} = D \begin{bmatrix} e_t^1 \\ e_t^2 \\ e_t^{\mathsf{mp}} \end{bmatrix}$$

- And that we impose

$$D = \begin{bmatrix} & 0 & 0 \\ & & 0 \\ & & \end{bmatrix}$$

- This corresponds with imposing

$$B = \begin{bmatrix} & & 0 \\ & 0 & 0 \\ & & \end{bmatrix}$$

- This does not affect the IRF of e_t^{mp}. All that matters for the IRF is whether a variable is ordered before or after r_t

First-order VAR

$$y_t = A_1 y_{t-1} + Be_t$$

- Now we can calculate IRFs, variances, and autocovariances analytically
- Mainly because you can easily calculate the MA representation

$$y_t = Be_t + A_1 Be_{t-1} + A_1^2 Be_{t-2} + \cdots$$

State-space notation

Every VAR can be presented as a first-order VAR. For example let

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = A_1 \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + A_2 \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \end{bmatrix} + B \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix}$$

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ I_{2\times 2} & 0_{2\times 2} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{1,t-2} \\ y_{2,t-2} \end{bmatrix} + \begin{bmatrix} B & 0_{2\times 2} \\ 0_{2\times 2} & 0_{2\times 2} \end{bmatrix} \begin{bmatrix} e_{1,} \\ e_{2,} \\ 0 \\ 0 \end{bmatrix}$$

VAR used by Gali

$$z_t = \sum_{j=1}^{J} A_j z_{t-j} + B\varepsilon_t$$

with

$$z_t = \begin{bmatrix} \Delta \ln(y_t/h_t) \\ \Delta \ln(h_t) \end{bmatrix}$$

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{t,\text{technology}} \\ \varepsilon_{t,\text{non-technology}} \end{bmatrix}$$

Identifying assumption (Blanchard-Quah)

- Non-technology shock does not have a long-run impact on productivity
- Long-run impact is zero if
  - Response of the *level* goes to zero
  - Responses of the *differences* sum to zero

Get MA representation

$$
\begin{aligned}
z_t &= A(L)z_t + B\varepsilon_t \\
&= (I - A(L))^{-1}B\varepsilon_t \\
&= D(L)\varepsilon_t \\
&= D_0\varepsilon_t + D_1\varepsilon_{t-1} + \cdots
\end{aligned}
$$

Note that $D_0 = B$

Sum of responses

$$\sum_{j=0}^{\infty} D_j = D(1) = (I - A(1))^{-1} B$$

Blanchard-Quah assumption:

$$\sum_{j=0}^{\infty} D_j = \begin{bmatrix} & 0 \\ & \end{bmatrix}$$

If you ever feel bad about getting too much criticism ....

-

If you ever feel bad about getting too much criticism ....

-
- be glad you are not a structural VAR

Structural VARs & critiques

- From MA to AR
    - Lippi & Reichlin
- From prediction errors to structural shocks
    - Fernández-Villaverde, Rubio-Ramirez, Sargent
- Problems in finite samples
    - Chari, Kehoe, McGratten

From MA to AR

Consider the two following *different* MA(1) processes

$$
\begin{aligned}
y_t &= \varepsilon_t + \frac{1}{2}\varepsilon_{t-1}, \;\; \mathsf{E}_t\left[\varepsilon_t\right] = 0,\, \mathsf{E}_t\left[\varepsilon_t^2\right] = \sigma^2 \\
x_t &= e_t + 2e_{t-1}, \;\; \mathsf{E}_t\left[e_t\right] = 0,\, \mathsf{E}_t\left[e_t^2\right] = \sigma^2/4
\end{aligned}
$$

- Different IRFs
- Same variance and covariance

$$
\mathsf{E}\left[y_t y_{t-j}\right] = \mathsf{E}\left[x_t x_{t-j}\right]
$$

From MA to AR

- AR representation:

$$\begin{aligned} y_t &= (1+\theta L)\,\varepsilon_t \\ \frac{1}{(1+\theta L)} y_t &= \varepsilon_t \\ \frac{1}{(1+\theta L)} &= \sum_{j=0}^{\infty} a_j L^j \end{aligned}$$

- Solve for a_js from

$$1 = a_0 + (a_1 + a_0\theta)\,L + (a_2 + a_1\theta)\,L^2 + \cdots$$

From MA to AR

Solution:

$$\begin{aligned} a_0 &= 1 \\ a_1 &= -a_0\theta \\ a_2 &= -a_1\theta = a_0\theta^2 \\ &\cdots \end{aligned}$$

You need

$$|\theta| < 1$$

Prediction errors and structural shocks

Solution to economic model

$$
\begin{aligned}
x_{t+1} &= Ax_t + B\varepsilon_{t+1} \\
y_{t+1} &= Cx_t + D\varepsilon_{t+1}
\end{aligned}
$$

- x_t: state variables
- y_t: observables (used in VAR)
- ε_t: structural shocks

Prediction errors and structural shocks

- From the VAR you get

$$\begin{aligned} e_{t+1} &= y_{t+1} - \mathsf{E}_t\left[y_{t+1}\right] \\ &= Cx_t + D\varepsilon_{t+1} - \mathsf{E}_t\left[Cx_{t+1}\right] \\ &= C\left(x_t - \mathsf{E}_t\left[x_t\right]\right) + D\varepsilon_{t+1} \end{aligned}$$

- Problem: Not guaranteed that

$$x_t = \mathsf{E}_t\left[x_t\right]$$

Prediction errors and structural shocks

- Suppose: $y_t = x_t$
  - that is, all state variables are observed
- Then

$$x_t = \mathrm{E}_t\left[x_t\right]$$

Prediction errors and structural shocks

- Suppose: $y_t \neq x_t$
- F-V,R-R,S show that $x_t = \mathsf{E}_t\left[x_t\right]$ if

  the eigenvalues of $A - BD^{-1}C$

  must be strictly less than 1 in modulus

Finite sample problems

- Summary of discussion above
  - Life is excellent if you observe all state variables
  - But,
    - we don't observe capital (well)
    - even harder to observe news about future changes
  - If ABCD condition is satisfied, you are still ok *in theory*
- Problem: you may need ∞-order VAR for observables
  - recall that k_t has complex dynamics

Finite sample problems

1. Bias of estimated VAR
   - apparently bigger for VAR estimated in first differences
2. Good VAR may need many lags

Alleviating finite sample problems

Do with model exactly what you do with data:

- NOT: compare data results with model IRF
- YES:
  - generate N samples of length T
  - calculate IRFs as in data
  - compare average across N samples with data analogue

This is how Kydland & Prescott calculated business cycle stats