Lecture 6: (S)VAR models: part 2

Geert Mesters (Universitat Pompeu Fabra, Barcelona GSE and VU Amsterdam)

Motivation

Motivation

- Our objective is to learn about the parameters of the structural vector autoregressive model
- We know that we need extra information in order to fully identify the model
- To this extend this set of slides discusses several prominent identification methods

Summary of reduced and structural forms

Reduced form

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \ldots + \mathbf{A}_p\mathbf{y}_{t-p} + \boldsymbol{u}_t \qquad \boldsymbol{u}_t \sim WN(0, \boldsymbol{\Sigma}_u)$$

$$\mathbf{y}_t = \sum_{i=0}^{\infty} \boldsymbol{\Phi}_i \boldsymbol{u}_{t-i}$$

Structural form

$$\mathbf{B}_0\mathbf{y}_t = \mathbf{B}_1\mathbf{y}_{t-1} + \ldots + \mathbf{B}_p\mathbf{y}_{t-p} + \boldsymbol{w}_t \qquad \boldsymbol{w}_t \sim WN(0, \mathbf{I}_K)$$

$$\mathbf{y}_t = \sum_{i=0}^{\infty} \boldsymbol{\Theta}_i \boldsymbol{w}_{t-i}$$

Relationship

$$\boldsymbol{u}_t = \mathbf{B}_0^{-1}\boldsymbol{w}_t \qquad \mathbf{A}_i = \mathbf{B}_0^{-1}\mathbf{B}_i \qquad \boldsymbol{\Theta}_i = \boldsymbol{\Phi}_i\mathbf{B}_0^{-1}$$

Objective

- The goal is to identify the matrix $\mathbf{B}_0$ using
    - Short run restrictions
    - Long run restrictions
    - Heteroskedasticity based identification
    - Identification using external instruments

Identification by short run restrictions

Short run restrictions

Given the relationship

$$\boldsymbol{u}_t = \mathbf{B}_0^{-1}\boldsymbol{w}_t$$

We may compute the variance[1]

$$\boldsymbol{\Sigma}_u = \mathbf{B}_0^{-1}\mathbf{B}_0^{-1\prime}$$

- This equality is a system of nonlinear equations, with unknown parameters in $\mathbf{B}_0$

[1] Note that we exploit the identity $\mathbf{E}\boldsymbol{w}_t\boldsymbol{w}_t' = \mathbf{I}_K$

Short run restrictions

The system

$$\mathbf{\Sigma}_u = \mathbf{B}_0^{-1}\mathbf{B}_0^{-1'}$$

cannot be solved in general as

- $\mathbf{\Sigma}_u$ has $K(K+1)/2$ free parameters
- $\mathbf{B}_0^{-1}$ has K^2 free parameters
- This implies that we need to impose $K(K-1)/2$ additional restrictions on $\mathbf{B}_0^{-1}$

Short run restrictions

$$\mathbf{\Sigma}_u = \mathbf{B}_0^{-1}\mathbf{B}_0^{-1'}$$

- Imposing $K(K-1)/2$ restriction on $\mathbf{B}_0$ can be done in different ways; proportionality, zeros, etc
- But, it must be such that $\mathbf{B}_0$ is full rank, see Ramirez, Waggoner and Zha (2010)

Recursively identified models

- A prominent way of imposing restrictions is via a recursive structure

Note that for any positive definite matrix we may decompose

$$\mathbf{\Sigma}_u = \mathbf{P}\mathbf{P}'$$

where **P** is lower triangular. This implies

$$\mathbf{\Sigma}_u = \mathbf{B}_0^{-1}\mathbf{B}_0^{-1'} \rightarrow \mathbf{B}_0^{-1} = \mathbf{P}$$

It is important to keep in mind that the orthogonalization of the reduced form residuals by applying a lower-triangular decomposition is appropriate only if the recursive structure embodied in **P** can be justified on economic grounds

Sources for restrictions

- Information delays
- Physical constraints
- Institutional knowledge
- Assumptions about market structure
- Homogeneity of demand functions
- Extraneous parameter estimates
- High-frequency data

See Kilian and Lutkepohl (2017) for an extensive list of examples

Example crude oil

Let

$$\begin{bmatrix} \Delta prod_t \\ rea_t \\ rpoil_t \end{bmatrix}$$

where

- $\Delta prod_t$ denotes the percent change in world crude oil production
- rea_t is a business cycle index measuring global real economic activity
- $rpoil_t$ is the log of the real price of oil

Example crude oil

Killian (2009) impose that

$$\begin{bmatrix} u_t^{prod} \\ u_t^{rea} \\ u_t^{poil} \end{bmatrix} = \begin{bmatrix} b_0^{11} & 0 & 0 \\ b_0^{21} & b_0^{22} & 0 \\ b_0^{31} & b_0^{32} & b_0^{33} \end{bmatrix} \begin{bmatrix} w_t^{\text{oil supply}} \\ w_t^{\text{aggregate demand}} \\ w_t^{\text{oil-specific demand}} \end{bmatrix}$$

Implies

- vertical short-run oil supply curve and downward-sloping short-run oil demand curve
- delay restriction: oil-specific demand shocks raise the real price of oil, but without affecting global real economic activity within the same month

Monetary policy example

Consider the example of Christiano, Eichenbaum, and Evans (1999)

$$\boldsymbol{Y}_t = \begin{bmatrix} X_t^s \\ r_t \\ X_t^f \end{bmatrix}$$

where X_t^s slow moving, r_t policy rate and X_t^f fast moving.

The benchmark timing assumption is

$$\begin{bmatrix} u_t^s \\ u_t^r \\ u_t^f \end{bmatrix} = \begin{bmatrix} b_0^{11} & 0 & 0 \\ b_0^{21} & b_0^{22} & 0 \\ b_0^{31} & b_0^{32} & b_0^{33} \end{bmatrix} \begin{bmatrix} \epsilon_t^s \\ \epsilon_t^r \\ \epsilon_t^f \end{bmatrix}$$

Implementation short run restrictions

The main benefit of the short run approach is that it is easy to implement

- Estimate the reduced form coefficients $\hat{\mathbf{A}}(L), \hat{\mathbf{\Sigma}}_u$
- Solve $\hat{\mathbf{\Sigma}}_u = \hat{\mathbf{B}}_0^{-1} \hat{\mathbf{B}}_0^{-1'}$
- Construct the impulse responses using $\hat{\mathbf{\Theta}}_i = \hat{\mathbf{\Phi}}_i \hat{\mathbf{B}}_0^{-1}$

Criticism

Some criticism of this approach, see Rudebush (1998) and Killian (2012) for more

- Questionable credibility of lack of in-period response of X_t^s to r_t (would gdp not respond within the quarter to a policy change?)
- The implicit policy reaction function often does not accord with theory or practical experience
- Implementations often ignore changes in policy reaction functions (different periods at the FED implied different policy function)
- VAR information is typically far less than standard information sets
- Estimated monetary policy shocks dont match futures market

Identification by long run restrictions

Long run restrictions

- One alternative idea has been to impose restrictions on the long-run response of variables to shocks
- Focus on long-run properties of models that most economists can more easily agree on
- Example: most economists agree that demand shocks such as monetary policy shocks are neutral in the long-run, whereas productivity shocks are not

Long run restrictions

The long-run cumulative effects of the structural shocks are summarized by the matrix

$$\boldsymbol{\Theta}(1) = \sum_{i=0}^{\infty} \boldsymbol{\Theta}_i$$

The main idea is to restrict certain long run relationships to zero in order to identify $\mathbf{B}_0$

We have that

$$\boldsymbol{\Theta}(1) = \mathbf{B}(1)^{-1} = \mathbf{A}(1)^{-1}\mathbf{B}_0^{-1}$$

which implies

$$\mathbf{B}_0^{-1} = \mathbf{A}(1)\mathbf{B}(1)^{-1}$$

Long run restrictions

Now notice

$$\begin{aligned} \boldsymbol{\Sigma}_u &= \mathbf{B}_0^{-1}\mathbf{B}_0^{-1'} \\ &= \mathbf{A}(1)\mathbf{B}(1)^{-1}\left[\mathbf{A}(1)\mathbf{B}(1)^{-1}\right]' \\ &= \mathbf{A}(1)\mathbf{B}(1)^{-1}\mathbf{B}(1)^{-1'}\mathbf{A}(1)' \end{aligned}$$

Pre-multiply by $\mathbf{A}(1)^{-1}$ and post-multiply by $\mathbf{A}(1)^{-1'}$ to get

$$\begin{aligned} \mathbf{A}(1)^{-1}\boldsymbol{\Sigma}_u\mathbf{A}(1)^{-1'} &= \mathbf{B}(1)^{-1}\mathbf{B}(1)^{-1'} \\ &= \boldsymbol{\Theta}(1)\boldsymbol{\Theta}(1)' \end{aligned}$$

Long run restrictions

We have obtained

$$\mathbf{A}(1)^{-1}\mathbf{\Sigma}_u\mathbf{A}(1)^{-1'} = \mathbf{\Theta}(1)\mathbf{\Theta}(1)'$$

- This equality is a system of nonlinear equations, with unknown parameters in $\mathbf{\Theta}(1)$
- Again we have K^2 unknowns and only $K(K+1)/2$ observables; so we need to impose $K(K-1)/2$ restrictions on $\mathbf{\Theta}(1)$

After imposing restrictions we solve for $\mathbf{B}_0$ using

$$\mathbf{\Theta}(1) = \mathbf{A}(1)^{-1}\mathbf{B}_0^{-1}$$

Example: Blanchard & Quah (1989)

Let

$$\mathbf{y}_t = \begin{bmatrix} \Delta gdp_t \\ ur_t \end{bmatrix} \qquad \boldsymbol{w}_t = \begin{bmatrix} w_t^{\text{supply}} \\ w_t^{\text{demand}} \end{bmatrix}$$

The restriction imposed is that

$$\boldsymbol{\Theta}(1) = \begin{bmatrix} \theta_{11}(1) & 0 \\ \theta_{21}(1) & \theta_{22}(1) \end{bmatrix}$$

- This requires GDP to return to its initial level in response to an aggregate demand shock

Identification by heteroskedasticity

Heteroskedasticity based identification

- Another strand of the literature exploits certain statistical properties of the data for identification.
- In particular, changes in the conditional or unconditional volatility of the VAR errors can be used to assist in the identification of structural shocks, e.g. Rigobon (2003) and Rigobon and Sack (2003)

Heteroskedasticity based identification

Consider

$$\mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{A}_p \mathbf{y}_{t-p} + \boldsymbol{u}_t$$

Idea, exploit

$$\begin{aligned} \mathbb{E}(\boldsymbol{u}_t \boldsymbol{u}_t') &= \boldsymbol{\Sigma}_1 \qquad t = 1, \ldots, T_1 \\ \mathbb{E}(\boldsymbol{u}_t \boldsymbol{u}_t') &= \boldsymbol{\Sigma}_2 \qquad t = T_1, \ldots, T \end{aligned}$$

with $\boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$ and other parameters do not change

There always exists $\mathbf{G}$ and $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_K)$ such that

$$\boldsymbol{\Sigma}_1 = \mathbf{G}\mathbf{G}' \qquad \boldsymbol{\Sigma}_2 = \mathbf{G}\boldsymbol{\Lambda}\mathbf{G}'$$

Heteroskedasticity based identification

We may take $\mathbf{B}_0^{-1} = \mathbf{G}$, which implies

$$\mathbf{\Sigma}_1 = \mathbf{B}_0^{-1}\mathbf{B}_0^{-1\prime} \qquad \mathbf{\Sigma}_2 = \mathbf{B}_0^{-1}\mathbf{\Lambda}\mathbf{B}_0^{-1\prime}$$

and

$$\mathbb{E}(\boldsymbol{w}_t \boldsymbol{w}_t') = \mathbf{I}_K \qquad t = 1, \ldots, T_1$$

$$\mathbb{E}(\boldsymbol{w}_t \boldsymbol{w}_t') = \mathbf{\Lambda} \qquad t = T_1, \ldots, T$$

- Now we have 2 sets of nonlinear equations, with $\mathbf{B}_0^{-1}$ as unobserved variables
- As an example, take $K = 2$, then we have $b_0^{11}, b_0^{12}, b_0^{21}, b_0^{22}, \lambda_1, \lambda_2$ as unknown parameters, but we also have $3 + 3 = 6$ observable parameters in $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$

Some comments

- Note the difference with the previous identification methods: here one relies on essentially statistical assumptions
- The shocks are only identified up to rotation, which implies that one still needs to sort out which shock is which
- Weak identification in this context is studied in Lewis (2018) who proposes several tests for weak identification

Identification using instrumental variables

IV based identification

Recall that the identification of $\mathbf{B}_0$ is akin to a simultaneous equations problem

$$\mathbf{B}_0 \mathbf{y}_t = \mathbf{B}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{B}_p \mathbf{y}_{t-p} + \boldsymbol{w}_t$$

- An external instrument z_t can be used to identify a particular structural shock, and thus the corresponding impulse response
- This is often referred to as SVAR-IV or Proxy-SVAR, see Stock (2008), Mertens & Raven (2012,2018), Olea, Stock & Watson (2018)

Instrument condition

- Without loss of generality assume that the structural shock of interest is the first shock $w_{1,t}$

Key Assumptions

1. $\mathbb{E}(z_t w_{1,t}) = c \neq 0$ (Relevant)
2. $\mathbb{E}(z_t w_{j,t}) = 0$ for all $j \neq 1$ and t (Exogeneity)

Further, because the scales of $\boldsymbol{u}_t = \mathbf{B}_0^{-1} \boldsymbol{w}_t$ is not uniquely pinned down we restrict $b_0^{11} = 1$

Identifying the impulse response

Let $\mathbf{B}_{0,1}^{-1}$ denote the first column of $\mathbf{B}_0^{-1}$

$$\mathbf{\Theta}_{i,1} = \mathbf{\Phi}_i \mathbf{B}_{0,1}^{-1},$$

To get this column note that

$$\mathbf{\Gamma} = \mathbb{E}(z_t \boldsymbol{u}_t) = \mathbb{E}(z_t \mathbf{B}_0^{-1} \boldsymbol{w}_t) = c\mathbf{B}_{0,1}^{-1}$$

which identifies the column up to scale and using $b_0^{11} = 1$ we obtain

$$\mathbf{\Theta}_{i,1} = \mathbf{\Phi}_i \mathbf{\Gamma} / \mathbf{\Gamma}_{11}$$

Estimation

Replace population quantities by sample estimates

$$\hat{\boldsymbol{\Theta}}_{i,1} = \hat{\boldsymbol{\Phi}}_i \hat{\boldsymbol{\Gamma}} / \hat{\boldsymbol{\Gamma}}_1$$

where $\hat{\boldsymbol{\Gamma}} = \frac{1}{T} \sum_{t=1}^{T} z_t \hat{\mathbf{u}}_t$, where $\hat{\mathbf{u}}_t$ is the reduced form residual.

- The strength of the instrument can be assessed using the F-statistic for the first stage regression of z_t on $\hat{\mathbf{u}}_t$
- Weak inference robust inference is discussed in Olea, Stock & Watson (2018)

Example

Let

$$\mathbf{y}_t = \begin{bmatrix} FFR_t \\ ur_t \end{bmatrix} \qquad \boldsymbol{w}_t = \begin{bmatrix} w_t^{\text{mp}} \\ w_t^{\text{other}} \end{bmatrix}$$

- The goal is to identify the impulse response unemployment to the monetary policy shock
- A good instrument is the high frequency monetary policy shock of Kuttner (2001)

Other identification schemes

Other identification schemes

- Sign restrictions: instead of imposing zeros only impose that the sign of $\mathbf{B}_0$ are known, leads to partial identification and typically requires Bayesian estimation
- Non-Gaussian identification: assume that the errors are non-Gaussian. This again identifies the shocks up to rotation (same as for heteroskedasticity). High potential but the distance from the Gaussian distribution needs to be very large for reliable inference.

Confidence bounds for impulse responses

Confidence bounds

Constructing confidence bounds can be done using

- Asymptotic distribution
- Bootstrap

Confidence bounds via asymptotic distribution

Recall that

$$\mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{A}_p \mathbf{y}_{t-p} + \boldsymbol{u}_t \qquad \boldsymbol{u}_t \sim WN(0, \boldsymbol{\Sigma}_u)$$

We can work out

$$\begin{bmatrix} \sqrt{T}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \\ \sqrt{T}(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}) \end{bmatrix} \xrightarrow{d} N\left(0, \begin{bmatrix} \boldsymbol{\Sigma}_\alpha & 0 \\ 0 & \boldsymbol{\Sigma}_\sigma \end{bmatrix}\right)$$

and the impulse responses $\hat{\theta}_{lk,i} = g_{lk,i}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\sigma}})$ are a function of these parameters, the delta method implies

$$\sqrt{T}(\hat{\theta}_{lk,i} - \theta_{lk,i}) \xrightarrow{d} N(0, \sigma^2_{lk,i})$$

where $\sigma^2_{lk,i}$ is a complicated function of $\boldsymbol{\Sigma}_\alpha$ and $\boldsymbol{\Sigma}_\sigma$, see Killian & Lutkepohl (2017, Chapter 12) for details

Confidence bounds via bootstrap

- In practice most researchers prefer to use bootstrap methods to construct confidence bounds
- This is easy to implement and generally works better in finite sample
- We only discuss a simple implementation which is valid only if the errors satisfy $\boldsymbol{u}_t \sim IID(0, \boldsymbol{\Sigma}_u)$

The bootstrap DGP

- From the original data we can obtain estimates $\hat{\mathbf{A}}_1, \dots, \hat{\mathbf{A}}_p, \hat{\boldsymbol{\Sigma}}_u$

This allow to construct Bootstrap GDP

$$\mathbf{y}_t^* = \hat{\mathbf{A}}_1 \mathbf{y}_{t-1}^* + \dots + \hat{\mathbf{A}}_p \mathbf{y}_{t-p}^* + \boldsymbol{u}_t^* \qquad \boldsymbol{u}_t^* \sim \hat{F}_u$$

for the errors there are two options

- Parametric: assume F_u known and take $\hat{F}_u$ as its sample equivalent, for instance if $\boldsymbol{u}_t \sim NID(0, \boldsymbol{\Sigma}_u)$ then $\boldsymbol{u}_t^* \sim NID(0, \hat{\boldsymbol{\Sigma}}_u)$
- Non-parametric: take $\hat{F}_u$ as the empirical distribution function of the residuals $\hat{u}_t$, amounts to random sampling from the residuals with replacement

The bootstrap DGP

Given Bootstrap GDP

$$\mathbf{y}_t^* = \hat{\mathbf{A}}_1 \mathbf{y}_{t-1}^* + \ldots + \hat{\mathbf{A}}_p \mathbf{y}_{t-p}^* + \boldsymbol{u}_t^* \qquad \boldsymbol{u}_t^* \sim \hat{F}_u$$

1. For $r = 1, \ldots, B$, with B large, construct samples $\{u_t^{*,r}\}_{t=1}^T$
2. Based on the new errors construct data samples $\{\mathbf{y}_t^*\}_{t=1}^T$ for $r = 1, \ldots, B$
3. For each new data sample construct the impulse responses of interest: $\hat{\mathbf{\Theta}}_i^{*,r}$ for $r = 1, \ldots .B$
4. We take the desired percentiles of the empirical bootstrap distribution as confidence intervals: e.g. $[\hat{\theta}_{lk,i,\gamma/2}, \hat{\theta}_{lk,i,1-\gamma/2}]$

Some comments

- When the errors are heteroskedastic and serially correlated different bootstrap procedures need to be adopted; see Goncalves and Kilian (2004, 2007), Bruggemann, Jentsch, and Trenkler (2016), Killian & Lutkepohl (2017, Chapter 12)

References & Material

Reference & Material

- References:
  Structural Vector Autoregressive Analysis' by Lutz Kilian and Helmut Lutkepohl, Cambridge University Press, 2017.