ECON 815

Vector Auto Regressions

Winter 2020

Policy Analysis with Macro Data

Consider the following example

$$y_t = \alpha g_t + \gamma x_t + \epsilon_{yt}$$
$$g_t = \beta y_t + \epsilon_{gt}$$

where ϵ_{it} and x_t are uncorrelated.

The first equation describes a model that links output growth y_t to government expenditure growth g_t and some control variable x_t.

The second equation captures a reaction function for fiscal policy.

Goal:

We would like to estimate the impact of fiscal stimulus (ϵ_{gt}) taking into account a fiscal multiplier.

Problem:

Can we interpret correlations as causal relationships?

Can we say something about the size of the fiscal multiplier?

The Identification Problem

Solving the model, we get

$$g_t = \frac{\beta\gamma}{1-\alpha\beta}x_t + \frac{\beta}{1-\alpha\beta}\epsilon_{yt} + \frac{1}{1-\alpha\beta}\epsilon_{gt}$$
$$y_t = \frac{\gamma}{1-\alpha\beta}x_t + \frac{1}{1-\alpha\beta}\epsilon_{yt} + \frac{\alpha}{1-\alpha\beta}\epsilon_{gt}$$

1) Estimating the model would allow us to pin down β (i.e., the reaction function), but not the fiscal multiplier $\frac{\alpha}{1-\alpha\beta}$. We need some other (identifying) assumption to find α.

2) A positive correlation between y_t and g_t can mean anything:

- Economist A ($\alpha = 0$, $\beta > 0$) – growth increases in gov't spending
- Economist B ($\alpha > 0$, $\beta > 0$) – gov't spending spurs growth

VAR Analysis

Consider the model

$$\mathbf{y}_t = \mu + \mathbf{\Gamma}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{\Gamma}_p \mathbf{y}_{t-p} + \mathbf{u}_t$$

which is an autoregression with several lags involving vectors $\mathbf{y}$.

For theoretical exposition, we can always stack vectors of longer lags

$$\begin{bmatrix} \mathbf{y}_t \\ \mathbf{y}_{t-1} \end{bmatrix} = \begin{bmatrix} \mu \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{\Gamma}_1 & \mathbf{\Gamma}_2 \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{y}_{t-1} \\ \mathbf{y}_{t-2} \end{bmatrix} + \begin{bmatrix} \nu_t \\ \mathbf{0} \end{bmatrix}$$

to obtain a first-order VAR with redefined variables

$$\tilde{\mathbf{y}}_t = \tilde{\mu} + \mathbf{\Gamma} \tilde{\mathbf{y}}_{t-1} + \tilde{\mathbf{u}}_t$$

These are **reduced form VARs**, since today's variables only depend on predetermined variables and today's disturbances.

We can use them for forecasting.

Forecasting

For that purpose, we transform the VAR into its MA representation (presuming that $\boldsymbol{\Gamma}$ is stable)

$$\begin{aligned}\mathbf{y}_t &= \mu + \boldsymbol{\Gamma}\mathbf{y}_{t-1} + \mathbf{u}_t \\ &= \mu + \boldsymbol{\Gamma}L\mathbf{y}_t + \mathbf{u}_t \\ &= [\mathbf{I} - \boldsymbol{\Gamma}L]^{-1}(\mu + \mathbf{u}_t) \\ &= \boldsymbol{\Gamma}(L)(\mu + \mathbf{u}_t) \\ &= \bar{\mathbf{y}} + \sum_{i=0}^{\infty} \boldsymbol{\Gamma}^i \mathbf{u}_{t-i}\end{aligned}$$

We can use OLS to estimate $\boldsymbol{\Gamma}$ and use bootstrapping to obtain standard errors.

Interpretation:

The matrices $\boldsymbol{\Gamma}^i$ determine IRFs with respect to innovations/shocks $\mathbf{u}_t$. An element is given by $\gamma_{ml}(i)$ which gives the deviation of $y_{m,t}$ from its mean due to a one-time "shock" in $u_{l,t-i}$.

The Identification Problem Revisited

1) Macro variables are correlated contemporaneously with each other.

2) We think of "shocks" as related to specific variables rather than purely random innovations.

A macroeconomic model has the structure

$$\mathbf{B}_0\mathbf{y}_t = \mathbf{B}_1\mathbf{y}_{t-1} + \mathbf{w}_t$$
$$\mathbf{y}_t = [\mathbf{B}_0 - \mathbf{B}_1 L]^{-1}\mathbf{w}_t$$

where all **structural** shocks $\mathbf{w}_t$

- have mean zero
- are serially uncorrelated
- and $E[w_t, w_t'] = \mathbf{\Sigma}_w = \mathbf{I}$ is a diagonal matrix which we normalize to have unit variances.

$\mathbf{B}_0$ are contemporaneous correlations among variables.

In order to estimate the parameters of this model, we map the model to a reduced form VAR.

$$\begin{aligned} \mathbf{B}_0\mathbf{y}_t &= \mathbf{B}_1\mathbf{y}_{t-1} + \mathbf{w}_t \\ \mathbf{y}_t &= \mathbf{B}_0^{-1}\mathbf{B}_1\mathbf{y}_{t-1} + \mathbf{B}_0^{-1}\mathbf{w}_t \\ \mathbf{y}_t &= \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{u_t} \end{aligned}$$

Hence, we can estimate

$$\mathbf{A}_1 = \mathbf{B}_0^{-1}\mathbf{B}_1$$

from a reduced form VAR and find the covariance matrix

$$\mathbf{\Sigma}_u = E[\mathbf{u_t}\mathbf{u_t}'] = \mathbf{B}_0^{-1}\mathbf{\Sigma}_w\mathbf{B}_0^{-1\prime} = \mathbf{B}_0^{-1}\mathbf{B}_0^{-1\prime}$$

Conclusion:

Identifying a VAR is equivalent of **choosing a unique** matrix of $\mathbf{B}_0$ (or $\mathbf{B}_0^{-1}$ see below) that determines the contemporaneous interactions among variables.

SVAR – Restrictions

Since we have normalized the variances of structural shocks, we have

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

There are n^2 parameters. The covariance matrix is symmetric so that we have only $n(n+1)/2$ independent equations and, thus, we need $n(n-1)/2$ additional restrictions on $\mathbf{B}_0^{-1}$ (from economic reasoning).

Note that different restrictions will yield different IRFs with a unit innovation in the structural shock being interpreted as a one standard deviation shock.

How do we find $n(n-1)/2$ more restrictions?

- recursive structure and ordering (Cholesky decomposition)
- "ad-hoc" restrictons on contemporaneous variables (from theory)
- long-run neutrality restrictions (Blanchard and Quah)

Recursive Identification

Order the variables such that variables higher in the order are determined before variables lower in the order.

Shocks to higher variables contemporaneously influence lower variables, but not the other way around.

This is equivalent to assuming that $\mathbf{B}_0^{-1}$ is lower triangular.

This gives us $n(n-1)/2$ restrictions.

This is the Cholesky Decomposition of the covariance matrix $\mathbf{\Sigma}_u$.

We can estimate $\mathbf{A}_1$ from a reduced form VAR and run the IRF

$$\mathbf{B}_0^{-1}\mathbf{w_t} = \begin{bmatrix} b_{11} & 0 & \dots & \dots & 0 \\ b_{21} & b_{22} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ b_{n1} & b_{n2} & \dots & \dots & b_{nn} \end{bmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

Short-run Restrictions

Theory can tell us about the structure of the matrix $\mathbf{B}_0$ directly.

Consider again the covariance matrix of shocks $E[\mathbf{w_t}\mathbf{w_t'}] = \mathbf{\Sigma}_w$ where we do not (!) normalize the variances.

Using $w_t = \mathbf{B}_0 u_t$, we then have

$$\mathbf{\Sigma}_w = \mathbf{B}_0 \mathbf{\Sigma}_u \mathbf{B}_0'$$

We now impose that the diagonal entries are $\mathbf{1}$ for $\mathbf{B}_0$ which leaves us again with $n(n-1)/2$ restrictions to be chosen.

Note that these now correspond directly to the contemporaneous relationships between the variables.

One needs to rescale the IRFs by the standard deviation of the structural residual w_t.

Long-run Restrictions

We can also use long-run neutrality restrictions from economic theory.

Define $\mathbf{B}(1) \equiv \mathbf{B}_0 - \mathbf{B}_1 = \mathbf{B}_0 \left(\mathbf{I} - \mathbf{A}_1\right) \equiv \mathbf{B}_0 \mathbf{A}(1)$.

The long-run impact matrix is given by

$$\left(\mathbf{I} + \mathbf{B}_0^{-1}\mathbf{B}_1 + \ldots\right)\mathbf{B}_0^{-1} = \left(\mathbf{B}_0(\mathbf{I} - \mathbf{B}_0^{-1}\mathbf{B}_1)\right)^{-1} = \mathbf{B}(1)^{-1} = \boldsymbol{\Theta}(1)$$

on which we will place enough restrictions (e.g., Colesky) for identificaton.

1) We can estimate the long-run matrix from $\mathbf{A}_1$ and $\boldsymbol{\Sigma}_u$ via

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

where we impose sufficient restrictions on $\boldsymbol{\Theta}(1)$.

2) Since $\boldsymbol{\Sigma}_u = \mathbf{B}_0^{-1}\mathbf{B}_0^{-1\prime}$, we can recover short-run restrictions from

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

Application – Productivity and Hours for CAN

Data for Canada (1/1981-3/2013):

- corr(GDP, Hours) = 0.696
- corr(GDP, Prod) = 0.491
- corr(Prod, Hours) = -0.285

Hours move countercyclical relative to (labour) productivity.

In the RBC model, we need a very small Frisch elasticity of the labour supply (η large), so that the income effect dominates the substitution effect.

Or we need other shocks (such as gov't policy or preference shocks) that give rise to a sufficiently strong negative comovement between measured labour productivity and hours.

Empirical Analysis using VARs

Model – VAR in hours n_t and (labour) productivity z_t

We log first-difference productivity and hours to get stationarity

- $\Delta z_t = \log z_t - \log z_{t-1}$
- $\Delta n_t = \log n_t - \log n_{t-1}$

Reduced-form VAR:

$$\begin{pmatrix} \Delta z_t \\ \Delta n_t \end{pmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{pmatrix} \Delta z_{t-1} \\ \Delta n_{t-1} \end{pmatrix} + \begin{pmatrix} \nu_{1t} \\ \nu_{2t} \end{pmatrix}$$

We can estimate this VAR and use it for forecasting.

To investigate the empirical relationship between hours and productivity, we need to make identification assumptions.

Recursive Identification

Reduced form VAR with one lag

- coefficient matrix

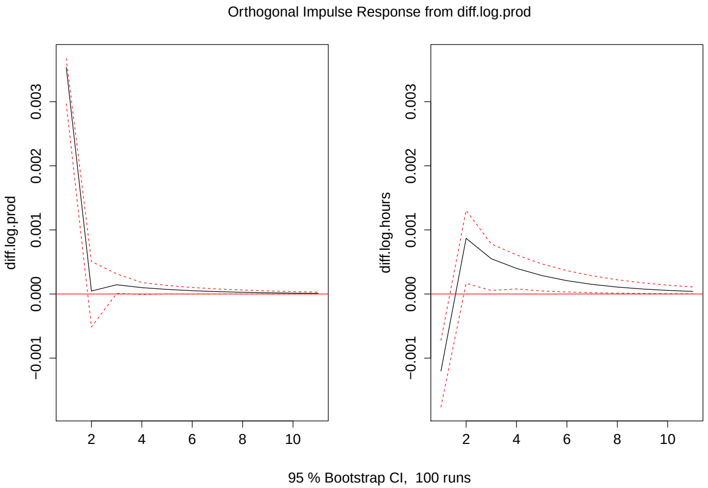
$$\boldsymbol{\Gamma} = \begin{bmatrix} 0.0686 & 0.16273 \\ 0.4527 & 0.6087 \end{bmatrix}$$

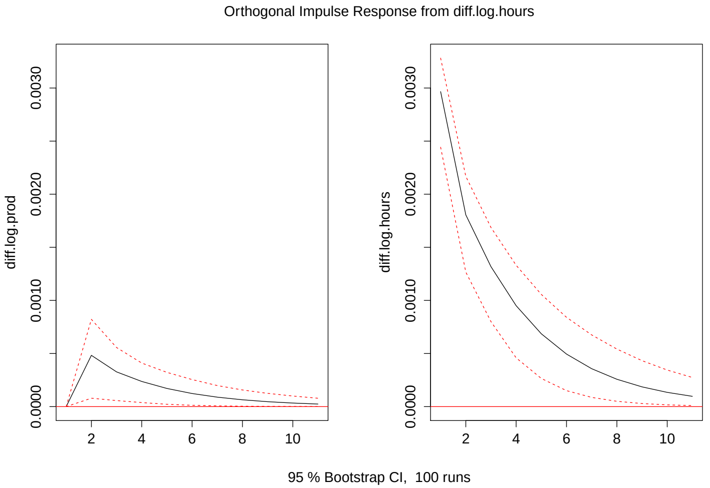
- Identifying restriction

$$\mathbf{B}_0^{-1} = \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix}$$

We interpret ϵ_{1t} as a technology shock and ϵ_{2t} as everything else (what?) that has no direct contemporaneous impact whatsoever on productivity.

IRFs – Recursive VAR




Orthogonal Impulse Response from diff.log.hours
diff.log.prod
diff.log.hours
0.0000
0.0010
0.0020
0.0030
2
4
6
8
10
95 % Bootstrap CI, 100 runs

Long-run Restrictions

We call some shocks *supply or technology* shocks and others *demand* shocks. We allow variables to affect each other contemporaneously and obtain correlations *conditional* on these shocks.

Identification Assumptions:

1) Shocks are orthogonal and normalized to 1 in variance.

2) Productivity is influenced in the long-run only by technology shocks (long-run restriction).

This means that we have the following MA representation

$$\begin{pmatrix} \Delta z_t \\ \Delta n_t \end{pmatrix} = \begin{bmatrix} \Theta_{11}(L) & \Theta_{12}(L) \\ \Theta_{21}(L) & \Theta_{22}(L) \end{bmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} = \boldsymbol{\Theta}(L)\epsilon_t$$

with restriction $\Theta_{12}(1) = 0$ or $\boldsymbol{\Theta}(1)$ being lower triangular.

Evidence from Canadian Data

Step 1 – Reduced form VAR from above to obtain $\mathbf{A}(1)$

Step 2 – Structural VAR:

- restrict long-run impact matrix $\mathbf{\Theta}(1)$ to be lower triangular

$$\mathbf{\Theta}(1) = \begin{bmatrix} 0.0044 & 0 \\ 0.0051 & 0.0082 \end{bmatrix}$$

- calculate the contemporaneous impact matrix

$$\mathbf{B}_0^{-1} = \begin{bmatrix} 0.0033 & -0.0013 \\ 0.0000 & 0.0032 \end{bmatrix}$$

- calculate cond. correlations
  - $corr(\Delta z_t, \Delta n_t|1) = 0.128$
  - $corr(\Delta z_t, \Delta n_t|2) = -0.579$