A Primer on Vector Autoregressions

VAR Toolbox 4.0

Ambrogio Cesa-Bianchi

June 29, 2026

[DISCLAIMER]

These notes are meant to provide intuition on the basic mechanisms of VARs

As such, most of the material covered here is treated in a very informal way

If you crave a formal treatment of these topics, you should stop here and buy a copy of Hamilton's "Time Series Analysis"

The Matlab codes accompanying the notes are available at:
`https://github.com/ambropo/VAR-Toolbox`

The job of macro-econometricians

- In their 2001 Journal of Economic Perspectives' article "Vector Autoregressions" Stock and Watson (2001) describe the job of macroeconometricians as consisting of the following tasks
  - Describe and summarize macroeconomic time series
  - Make forecasts
  - Recover the structure of the macroeconomy from the data
  - Advise macroeconomic policy-makers

- Vector autoregressive models (VARs) are a statistical tool to perform these tasks

What can we do with VARs?

- Consider an (over-simplistic) bivariate VAR with the following variables:
  - Real GDP growth (y_t)
  - Policy rate (r_t)

- A VAR can help us answer the following questions
  [1] What is the dynamic behavior of these variables? How do these variables interact?
  [2] What is the most likely path of GDP in the next few quarters?
  [3] What is the effect of a monetary policy shock on GDP?
  [4] What has been the historical contribution of monetary policy shocks to GDP fluctuations?

(S)VARs Basics

The Reduced-Form VAR

The reduced-form VAR

- A VAR with k endogenous variables and p lags describes the evolution of the $k \times 1$ vector x_t as a function of its own lags, a constant, and a serially uncorrelated error term

$$x_t = c + \Phi_1 x_{t-1} + \cdots + \Phi_p x_{t-p} + u_t$$

- where
  * c is a $k \times 1$ vector of intercepts
  * $\Phi_1, \ldots, \Phi_p$ are $k \times k$ matrices of autoregressive coefficients
  * u_t is a $k \times 1$ vector of serially uncorrelated reduced-form residuals with zero mean and time-invariant covariance matrix $\mathbb{V}(u_t) \equiv \Sigma_u$ (homoskedasticity)

A bivariate running example

- The running example in this primer is a bivariate VAR(1) in $x_t = (y_t, r_t)'$, where
  - y_t is the log difference of US real GDP
  - r_t is the 1-year US Treasury bill yield

Warning: Do not try this at home!

The bivariate VAR(1) used throughout is a purely pedagogical device. With only two variables and one lag, the algebra is tractable and the toolbox mechanics are easy to follow. But this is not a credible empirical model: no referee, advisor, or editor would accept it as serious applied work. Use it to understand how the toolbox works; do not use it as a template for research.

- The reduced-form representation of this simple VAR(1) is

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} u_{yt} \\ u_{rt} \end{bmatrix}$$

The reduced-form VAR

- This formulation of the VAR is called the **reduced-form VAR** representation

$$x_t = c + \Phi x_{t-1} + u_t$$

- In matrix form

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} u_{yt} \\ u_{rt} \end{bmatrix}$$

- Or as a system of linear equations

$$\begin{cases} y_t = c_y + \phi_{11} y_{t-1} + \phi_{12} r_{t-1} + u_{yt} \\ r_t = c_r + \phi_{21} y_{t-1} + \phi_{22} r_{t-1} + u_{rt} \end{cases}$$

The reduced-form covariance matrix

- A key object of interest in VARs is the covariance matrix of the **reduced-form residuals**

$$\Sigma_u \equiv \mathbb{E}\left[u_t u_t'\right] = \begin{bmatrix} \sigma_y^2 & \sigma_{yr} \\ \sigma_{yr} & \sigma_r^2 \end{bmatrix}$$

- Σ_u is symmetric (same off-diagonal elements σ_{yr}) and positive definite

- Typically, the reduced-form residuals are correlated among each other ($\sigma_{yr} \neq 0$)

- This is because the elements of u_t inherit all the contemporaneous relations among the endogenous variables x_t

- Thus, we can't use u_t to make causal statements about structural shocks hitting the economy
  - e.g. what are the effects of a monetary policy 'shock' on output

(S)VARs Basics

The Structural VAR Representation

The structural VAR representation

- Let the only two structural shocks be a demand and a monetary policy shock
- The reduced-form VAR(1) above can be written in its **structural form** as:

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

- The **structural shocks** ε_t have a well-defined economic interpretation
  - Serially uncorrelated, mutually uncorrelated, with zero mean and unit variance: $\mathbb{E}[\varepsilon_t] = 0$, $\mathbb{E}[\varepsilon_t \varepsilon_t'] = I_2$
- The **structural impact matrix** B maps structural shocks into the variables
  - The (i, j) element of B is the *impact* response of variable i to shock j
- The structural and reduced-form VAR are linked by $u_t = B\varepsilon_t$, so that

$$\Sigma_u = \mathbb{E}\left[u_t u_t'\right] = \mathbb{E}\left[B\varepsilon_t \left(B\varepsilon_t\right)'\right] = B\mathbb{E}(\varepsilon_t \varepsilon_t')B' = B\Sigma_\varepsilon B' = BB'$$

Three different ways of writing the same thing

- There are different ways to write the same structural VAR(1)

$$x_t = c + \Phi x_{t-1} + B\varepsilon_t$$

- For example, we can write it in matrix form

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

- Or as a system of linear equations

$$\begin{cases} y_t = c_y + \phi_{11} y_{t-1} + \phi_{12} r_{t-1} + b_{11}\varepsilon_t^{Demand} + b_{12}\varepsilon_t^{MonPol} \\ r_t = c_r + \phi_{21} y_{t-1} + \phi_{22} r_{t-1} + b_{21}\varepsilon_t^{Demand} + b_{22}\varepsilon_t^{MonPol} \end{cases}$$

Why reduced-form residuals are correlated

- Now that we have introduced the structural shocks ε_t, we can explain why Σ_u typically has non-zero off-diagonal elements
- The reduced-form residuals u_t are **mixtures** of structural shocks: $u_t = B\varepsilon_t$, so

$$\begin{cases} u_{yt} = b_{11}\varepsilon_t^{Demand} + b_{12}\varepsilon_t^{MonPol} \\ u_{rt} = b_{21}\varepsilon_t^{Demand} + b_{22}\varepsilon_t^{MonPol} \end{cases}$$

- Although structural shocks are orthogonal ($\mathbb{E}[\varepsilon_t\varepsilon_t'] = I_2$), the reduced-form residuals are correlated because they share the same underlying shocks

Why is it called 'structural' VAR?

- Go back to our bivariate structural VAR(1)

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

- The structural VAR can be thought of as a description of the true structure of the economy
  * e.g. an approximation of the solution of a DSGE model
- The structural shocks are shocks with a well-defined economic interpretation
  * e.g. TFP shocks or monetary policy shocks
  * Since the structural shocks are mutually uncorrelated, we can move one shock keeping the other shocks fixed
  * That is: we can focus on the causal effect of one shock at the time

Structural VARs can answer many interesting questions...

- Go back to our bivariate structural VAR(1), and assume for the moment that Φ and B are known
- In matrix form:

Dynamic matrix

Impact matrix

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \boxed{\begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \boxed{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

- What is the effect of monetary policy shocks on output?
  - The coefficient b_{12} captures the 'impact effect' of a monetary policy shock on output growth
  - The Φ matrix allows us to trace the 'dynamic effect' of the monetary policy shock over time
  - (We can add additional variables and look at other shocks: aggregate supply, oil price,...)

… but the estimation of structural VARs is tricky

- **Problem** The structural shocks ε_t are unobserved. So, how can we estimate B?

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

- Best we can do is to 'bundle' the ε_t into a single object:

$$u_t = B\varepsilon_t \Rightarrow \begin{bmatrix} u_{yt} \\ u_{rt} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix} \Rightarrow \begin{cases} u_{yt} = b_{11}\varepsilon_t^{Demand} + b_{12}\varepsilon_t^{MonPol} \\ u_{rt} = b_{21}\varepsilon_t^{Demand} + b_{22}\varepsilon_t^{MonPol} \end{cases}$$

- Why is this useful? The VAR becomes

$$x_t = c + \Phi x_{t-1} + u_t$$

- Now we can estimate Φ and u_t with OLS (where u_t will be OLS residuals)

(S)VARs Basics

Wold Representations & Stability

The Wold representation

- Let's introduce another representation of the VAR that will be useful later
- Start from the structural VAR representation

$$x_t = c + \Phi x_{t-1} + B\varepsilon_t$$

- The **Wold representation** can be obtained by substituting the VAR recursively

$$\begin{aligned}
x_t &= c + \Phi x_{t-1} + B\varepsilon_t \\
&= c + \Phi \underbrace{\boxed{\left(c + \Phi x_{t-2} + B\varepsilon_{t-1}\right)}}_{x_{t-1}} + B\varepsilon_t = (I + \Phi)c + \Phi^2 x_{t-2} + \Phi B\varepsilon_{t-1} + B\varepsilon_t \\
&= \ldots \\
&= \Phi^t x_0 + \sum_{j=0}^{t-1} \Phi^j c + \sum_{j=0}^{t-1} \Phi^j B\varepsilon_{t-j}
\end{aligned}$$

The Wold representation

- The Wold representation shows that each observation (x_t) can be re-written as a combination of three terms

$$x_t = \underbrace{\boxed{\Phi^t x_0}}_{\text{Initial condition}} + \underbrace{\boxed{\sum_{j=0}^{t-1} \Phi^j c}}_{\text{Constant contribution}} + \underbrace{\boxed{\sum_{j=0}^{t-1} \Phi^j B \varepsilon_{t-j}}}_{\text{Current \& past shocks}}$$

  * An initial condition
  * The cumulative contribution of the constant
  * The sum of current and past structural shocks

- Now let $t \rightarrow \infty$ to get

$$x_t = \Phi^{\infty} x_{t-\infty} + \sum_{j=0}^{\infty} \Phi^j c + \sum_{j=0}^{\infty} \Phi^j B \varepsilon_{t-j}$$

- Matrix Φ plays a crucial role $\rightarrow$ Determines whether a shock from the infinitely distant past still has an effect on x_t today

VAR stability

- VARs are typically designed to study cyclical variations around trends
  - The effect of shocks progressively dissipate over time
- Stability is the formal expression of this idea. For this to happen, two conditions must hold:
  - The infinite sums $\sum_{j=0}^{\infty} \Phi^j$ must converge to finite values
  - The first term $\Phi^{\infty} x_{t-\infty}$ must vanish
- **Mathematically** A VAR is called stable iff all the eigenvalues of Φ are less than 1 in modulus. More formally:

$$\det\left(\Phi - \lambda I_2\right) = 0 \quad |\lambda| < 1$$

- **Implication** In the absence of shocks, the VAR will converge to its equilibrium

[BOX] Why eigenvalues govern stability?

- **Scalar case**
  - Suppose $x_t = \phi x_{t-1}$. After j periods, $x_t = \phi^j x_0$ — a geometric series
  - Converges to zero if $|\phi| < 1$; diverges if $|\phi| > 1$
  - The scalar ϕ encodes all persistence in a single number
- **VAR case**
  - The matrix Φ has k eigenvalues $\lambda_1, \ldots, \lambda_k$ (solutions to $\det(\Phi - \lambda I) = 0$)
  - Raising to the j-th power: $\Phi^j = P \,\mathrm{diag}(\lambda_1^j, \ldots, \lambda_k^j)\, P^{-1}$ — a geometric series running independently along each eigenvalue direction
  - Converges to zero iff $|\lambda_i| < 1$ for all i
- **Key insight**
  - The largest eigenvalue in modulus governs how slowly the system returns to equilibrium
  - Eigenvalue close to $1 \Rightarrow$ highly persistent responses
  - Eigenvalue close to $0 \Rightarrow$ rapid decay

Equilibrium (the unconditional mean)

- Start from an arbitrary x_0 (say because a shock pushed the system away from equilibrium)
- Where will x_t be in the infinitely distant future? From the h-step-ahead Wold representation

$$\mathbb{E}[x_h \mid x_0] = \Phi^h x_0 + \sum_{j=0}^{h-1} \Phi^j c + \sum_{j=0}^{h-1} \Phi^j B\, \mathbb{E}[\varepsilon_{h-j}]$$

- Under stability, as $h \to \infty$ only the constant terms survive:

Geometric series

$$\lim_{h\to\infty} \mathbb{E}[x_h \mid x_0] = \sum_{j=0}^{\infty} \Phi^j c = \boxed{(I_2 - \Phi)^{-1}} c \equiv \mu$$

- The equilibrium the VAR converges to is the unconditional mean $\mu = (I - \Phi)^{-1} c$

(S)VARs Basics

The Identification Problem

Back to our reduced-form VAR

- We have seen above that with OLS we can only estimate the reduced-form VAR (and not the structural VAR)
- Assume we already have an OLS estimate of $\hat{\Phi}$ and $\hat{u}_t$:

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} u_{yt} \\ u_{rt} \end{bmatrix}$$

- **Question** What is the effect of a monetary policy shock on GDP growth?
- Unfortunately, the reduced-form innovations (u_{yt} or u_{rt}) are not going to help us in answering the question

Reduced-form VARs do not tell us anything about causality

- To see that, recall we assumed that the 'true' (and unobserved) model of the economy is

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

- Also recall that the reduced-form innovations are a linear combination of the structural shocks

$$\begin{aligned} u_{yt} &= b_{11}\varepsilon_t^{Demand} + b_{12}\varepsilon_t^{MonPol} \\ u_{rt} &= b_{21}\varepsilon_t^{Demand} + b_{22}\varepsilon_t^{MonPol} \end{aligned}$$

- An increase in u_{rt} is not a monetary policy shock! could be due to
  - [1] A positive demand shock that increases both output growth and the policy rate
  - [2] Or a monetary policy shock that decreases output growth and increases the policy rate
- How to know whether it is [1] or [2]? This is the very nature of the **identification problem**!

The identification problem

- The identification problem consists in finding a mapping from the reduced-form VAR to its structural counterpart

$$u_t = B\varepsilon_t$$

- To do that, we can exploit the relation between reduced-form and structural innovations

$$\Sigma_u = BB'$$

- The identification problem simply boils down to finding a B matrix that satisfies $\Sigma_u = BB'$
- Unfortunately this is not as easy as it sounds. Why?
  - Hint: There are infinite different Bs that give you the same Σ_u

The identification problem (cont'd)

- Think of $\Sigma_u = BB'$ as a system of equations

$$\begin{bmatrix} \sigma_y^2 & \sigma_{yr} \\ \sigma_{yr} & \sigma_r^2 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix}$$

- Can be rewritten as

$$\begin{cases} \sigma_y^2 = b_{11}^2 + b_{12}^2 \\ \sigma_{yr} = b_{11} b_{21} + b_{12} b_{22} \\ \sigma_{yr} = b_{11} b_{21} + b_{12} b_{22} \\ \sigma_r^2 = b_{21}^2 + b_{22}^2 \end{cases}$$

- **Problem** Because of the symmetry of Σ_u, the second and the third equation are identical!
- We are left with 4 unknowns (the elements of B) but only 3 equations!

[BOX] The rotation representation of the identification problem

- Let B_0 be any matrix with $\Sigma_u = B_0 B_0'$ (e.g. the Cholesky factor). For any orthogonal Q ($QQ' = I_k$), $B_0 Q$ is *also* a valid solution: $(B_0 Q)(B_0 Q)' = B_0\, QQ'\, B_0' = \Sigma_u$

- The set of *all* solutions to $\Sigma_u = BB'$ is the orbit $\{B_0 Q : QQ' = I_k\}$. Every identification scheme is just a rule for selecting one particular Q

- In the bivariate case, $Q(\theta) = \left[\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right]$ is a plane rotation: the angle θ is the single free parameter left unconstrained by the 3 equations of $\Sigma_u = BB'$

- **Geometric intuition** The data pin down the *shape* of the residual covariance ellipse, but not the *directions* of the structural axes within it. Identification amounts to anchoring those directions

Estimation

Adding the VAR Toolbox path to Matlab

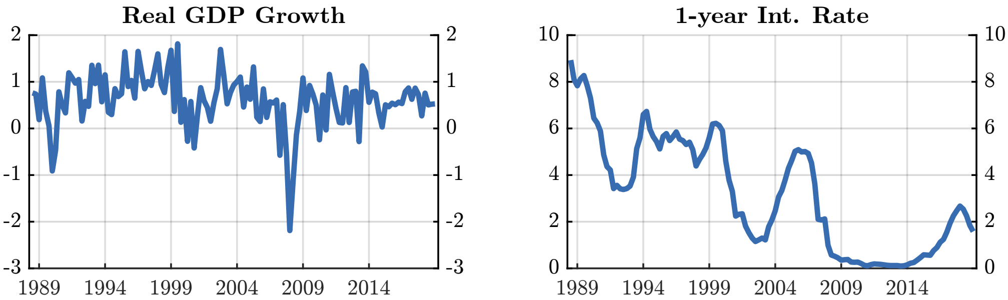
- To avoid clashes with functions from other toolboxes, add and remove the toolbox at the beginning and end of your scripts

- If you clone or download the toolbox to a folder named `VARToolbox`, add each subfolder explicitly to be safer:

```
% Identify the toolbox root from this file's location (one up from Primer)
root = fileparts(fileparts(mfilename('fullpath')));

% Add each relevant subfolder explicitly
addpath(fullfile(root, 'VAR'));
addpath(fullfile(root, 'Figure'));
addpath(fullfile(root, 'Stats'));
addpath(fullfile(root, 'Utils'));
addpath(fullfile(root, 'Primer'));
addpath(fullfile(root, 'Auxiliary'));
disp('VAR Toolbox 4.0 path set.');
.
.
% Remove the toolbox from the path at the end of the script
rmpath(genpath(root))
```

A simple bivariate VAR model

- Our first example will be a simple bivariate VAR as the one considered above
- US quarterly data from 1989:Q1 to 2019:Q4 on output growth (y_t) and the 1-year T-bill (r_t)



A simple bivariate VAR model

- Fit GDP growth and the 1-year rate with a VAR(1) with a constant

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} u_{yt} \\ u_{rt} \end{bmatrix}$$

- This means we will estimate the following parameters
  - $2+4$ coefficients, namely the elements of c and Φ
  - 2 variances of the reduced-form residuals, namely σ_y^2 and σ_r^2
  - 1 covariance of the reduced-form residuals, namely σ_{yr}

- We will store these coefficients in two Matlab matrices

$$\texttt{F} = \begin{bmatrix} c_y & \phi_{11} & \phi_{12} \\ c_r & \phi_{21} & \phi_{22} \end{bmatrix} \qquad \texttt{sigma} = \begin{bmatrix} \sigma_y^2 & \sigma_{yr} \\ \sigma_{yr} & \sigma_r^2 \end{bmatrix}$$

A simple bivariate VAR model

- In Matlab we store the data in the matrix `X`

$$\texttt{X} = \begin{bmatrix} y_1 & r_1 \\ y_2 & r_2 \\ \dots & \dots \\ y_T & r_T \end{bmatrix} \qquad \big(t\text{-th row: } x_t' = (y_t,\, r_t)\big)$$

- The VAR can then be easily estimated with a few lines of code

```
% Align variables to a common sample by removing leading/trailing NaNs
[X, fo, lo] = CommonSample(X);
% Initialise VAR options
VARopt = VARoption;
% Deterministic component: 1=constant, 2=constant+trend
detc = 1;
% Lag order
nlags = 1;
% Estimate VAR by OLS (reduced form only)
VAR_redform = VARmodel(X,nlags,detc,VARopt);
```

The `VARopt` and `VAR_redform` structures

- The `VARopt` structure stores all the options and settings that control estimation and identification:
  - Created by `VARoption` with sensible defaults for every field
  - Includes identification scheme (`ident`), number of horizons (`nsteps`), variable names (`vnames`)
  - Can be reused and modified across multiple calls to `VARmodel` and related functions

- The `VAR_redform` structure stores all the results of the VAR estimation
  - Includes reduced-form coefficients (`F`), residuals (`resid`), covariance matrix (`sigma`), eigenvalues (`maxEig`)
  - Also includes identification results (once computed, not done here yet): B matrix, impulse responses (`IR`), variance decomposition (`VD`), etc.
  - Contains the input data: the endogenous variables (`ENDO`), number of lags (`nlag`), etc.

VAR output

- The six estimated parameters (i.e. c and Φ) can be printed at screen by simply typing `disp(VAR_redform.F)` to get

```
>> disp(VAR_redform.F)

0.3630     0.3788     0.0041
-0.0729     0.2607     0.9541
```

- For the estimated reduced-form covariance matrix Σ_u type `disp(VAR_redform.sigma)`

```
>> disp(VAR_redform.sigma)

0.2891     0.0782
0.0782     0.1473
```

OLS estimation: Typical VAR output (cont'd)

- The off-diagonal elements of Σ capture the **average** contemporaneous relation between the endogenous variables

	GDP growth (u_y)	1-year T-bill (u_r)
GDP growth (u_y)	0.2891	$\boxed{0.0782}$
1-year T-bill (u_r)	0.0782	0.1473

Cov(u_y,u_r)>0

- In our example output growth and interest rates are contemporaneously positively correlated
  * It means that, on average, when GDP growth increases interest rates increase, too
- Does it mean that a shock to interest rates always increases output growth?
  * No! Recall that reduced-form residuals are not informative about structural shocks

Model checking & tuning

- These notes do not cover this aspect in detail but...
- ... before interpreting the VAR results you should check a number of assumptions
- Loosely speaking, you need to check that the reduced-form residuals are
  - Normally distributed
  - Not autocorrelated
  - Not heteroskedastic (i.e. have constant variance)
- ... and that the VAR is stable (next slide)

Stability and equilibrium

- We've already seen that a VAR is stable when all eigenvalues of its companion matrix lie inside the unit circle (for the VAR(1) here, the companion matrix is Φ)
  - If this condition is not met, the infinite sums in the Wold representation do not converge
- You can check the maximum modulus of the companion matrix's eigenvalues in the `VAR` structure, by typing `disp(VAR_redform.maxEig)` to get

```
>> disp(VAR_redform.maxEig)

0.9559
```

- You can also check all of the companion matrix's eigenvalues by executing Matlab's `eig` function on `Fcomp` (which, note, is built excluding the constant c from `F`)
- In practice:

```
>> disp(eig(VAR_redform.Fcomp))
0.3769
0.9559
```

[BOX] VAR estimation: log-levels vs. log-differences

- Standard VAR theory assumes covariance stationarity — but many macro aggregates (GDP, consumption, investment) contain stochastic trends in log-levels
- **Log-differences** Removes the unit root ⇒ a model for growth rates. Shocks are transitory in growth rates, but cumulating the responses leaves a *permanent* level effect. The constant is the average growth rate
- **Log-levels** OLS is still consistent despite unit roots (Sims et al., 1990) (though t/F tests may be non-standard). Preserves low-frequency comovement. Shocks are transitory *in levels*; the constant has no stand-alone interpretation
- This handbook estimates the benchmark VAR in log-differences

[BOX] Choosing the lag order

- Too **few** lags $\longrightarrow$ omitted-variable bias and serially correlated residuals; too **many** $\longrightarrow$ wasted degrees of freedom and inflated parameter uncertainty
- Information criteria: $\text{AIC} = \ln|\hat{\Sigma}_u| + \frac{2}{T}pk^2$ and $\text{BIC} = \ln|\hat{\Sigma}_u| + \frac{\ln T}{T}pk^2$. BIC penalizes more heavily $\Rightarrow$ more parsimonious models
- Rule of thumb: quarterly VARs use 4–8 lags, monthly 12–24; keep parameters per equation well below $T/10$
- Use `[AIC,BIC,logL] = VARlag(X,maxlag,detc)`. Always verify the chosen lag order yields serially uncorrelated residuals

Identification

How to solve the identification problem?

- Identification problem (recap)
  - Identification $\longrightarrow$ Find a B that satisfies $\Sigma_u = BB'$
  - There are infinite of such Bs
- In our simple example, we have to solve a system of 3 equations in 4 unknowns. How can we do it? Add a fourth equation!
- Economic theory can help in providing the 'missing' equation
  - Make an assumption about the structure of the economy based on your beliefs (e.g. long-run monetary neutrality)
  - Try to map this assumption into an equation that involves the VAR parameters
- The additional equation is known as a **restriction**
  - That is: the additional equation restricts the set of infinite B matrices to a single one (or few ones) that are consistent with your assumption

Common identification schemes

- Zero contemporaneous restrictions
- Zero long-run restrictions
- Sign restrictions
- Narrative sign restrictions
- External instruments (proxy SVARs)
- Combining sign restrictions & external instruments
- Identification with exogenous variables

Common Identification Schemes

Zero Short-Run Restrictions

Zero contemporaneous restrictions

- **Intuition** Identification is achieved by assuming that some shocks have zero contemporaneous effect on some of the endogenous variables
- **References** Sims (1980), Christiano et al. (1999)
- For example, assume that monetary policy works with a lag and has no contemporaneous effects on output
- But how can we impose restrictions on the effect of a structural shock?

Zero contemporaneous restrictions

- **Solution** Impose zero restrictions on the impact matrix B

- The b_{12} coefficient captures the contemporaneous effect of monetary policy on output growth

By assumption

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & \boxed{0} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

- **Implication** We now have 3 structural parameters to estimate (instead of 4) and 3 restrictions implied by Σ_u

Zero contemporaneous restrictions

How to achieve identification?

- The system of equations implied by $\Sigma_u = BB'$ now becomes

$$\begin{bmatrix} \sigma_y^2 & \sigma_{yr} \\ - & \sigma_r^2 \end{bmatrix} = \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} \\ 0 & b_{22} \end{bmatrix}$$

- This yields

$$\begin{cases} \sigma_y^2 = b_{11}^2 \\ \sigma_{yr} = b_{11} b_{21} \\ \sigma_r^2 = b_{21}^2 + b_{22}^2 \end{cases}$$

- And can be easily solved to get:

$$\begin{cases} b_{11} = \sigma_y \\ b_{21} = \sigma_{yr}/\sigma_y \\ b_{22} = \sqrt{\sigma_r^2 - \frac{\sigma_{yr}^2}{\sigma_y^2}} \end{cases}$$

Zero contemporaneous restrictions

Impact effects

- We can now derive the impact effects of shocks by simply re-writing the structural VAR as

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \sigma_y & 0 \\ \sigma_{yr}/\sigma_y & \sqrt{\sigma_r^2 - \frac{(\sigma_{yr})^2}{\sigma_y^2}} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

- A one standard deviation shock to monetary policy ($\varepsilon_t^{MonPol} = 1$) in t leads to

$$\begin{cases} y_t = \boxed{0} \xrightarrow{} \text{By assumption} \\ r_t = \sqrt{\sigma_r^2 - \frac{(\sigma_{yr})^2}{\sigma_y^2}} \end{cases}$$

- A one standard deviation shock to aggregate demand ($\varepsilon_t^{Demand} = 1$) in t leads to

$$\begin{cases} y_t = \sigma_y \\ r_t = \sigma_{yr}/\sigma_y \end{cases}$$

Zero contemporaneous restrictions
Cholesky implementation

- This identification scheme is normally implemented via a Cholesky decomposition of Σ_u

- A Cholesky decomposition allows us to decompose Σ_u into the product of a lower triangular matrix P times its transpose

$$\Sigma_u = PP'$$

- In matrix form we have

Lower Cholesky factor

$$\begin{bmatrix} \sigma_y^2 & \sigma_{yr} \\ \sigma_{yr} & \sigma_r^2 \end{bmatrix} = \boxed{\begin{bmatrix} p_{11} & 0 \\ p_{21} & p_{22} \end{bmatrix}} \begin{bmatrix} p_{11} & p_{21} \\ 0 & p_{22} \end{bmatrix}$$

[BOX] Cholesky decomposition of a matrix

- The Cholesky decomposition is (roughly speaking) the square root of a matrix
  - As for a square root, you can't compute a Cholesky decomposition for a non positive-definite matrix
- A symmetric and positive-definite matrix A can be decomposed as:

$$A = PP'$$

  where P is the unique lower triangular matrix with positive diagonal entries (and therefore P' is upper triangular)

- The formula for the decomposition of a 2×2 matrix is

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad P = \begin{bmatrix} \sqrt{a} & 0 \\ \frac{b}{\sqrt{a}} & \sqrt{c - \frac{b^2}{a}} \end{bmatrix}$$

Zero contemporaneous restrictions

Cholesky implementation

- To see why the zero contemporaneous restrictions identification can be implemented with a Cholesky decomposition, first note that Σ_u is a positive semi-definite matrix

- Then we can use the Cholesky decomposition to write

$$\Sigma_u = PP'$$

- But remember that we assumed that B is also lower triangular ($b_{12} = 0$) and that

$$\Sigma_u = BB'$$

- As both P and B are lower triangular, it must follow that $P = B$

Zero short-run restrictions: In Matlab

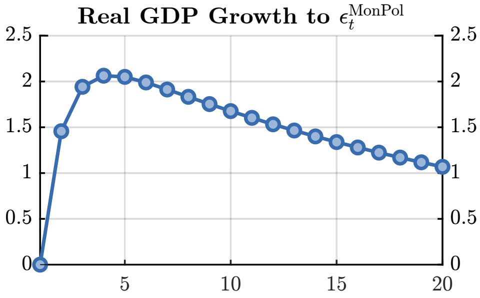
- Set `VARopt.ident = 'short'` and call `VARmodel` again, now saving the results in `VAR_short`:

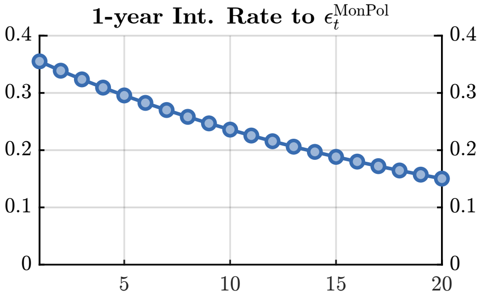
```
% Select zero short-run identification
VARopt.ident = 'short';

% Compute impulse responses and plot the MP shock (pick=2)
VAR_short = VARmodel(X,nlags,detc,VARopt);
VARopt.pick = 2;
VARirplot(VAR_short.IR,VARopt);
```

- `VAR_short.B` contains the impact matrix consistent with zero short-run restrictions
- `VAR_short.IR` contains the impulse responses
- The in-built function `VARirplot` can be used to plot the impulse responses

Zero short-run restrictions: Impulse responses

- Monetary policy shock (shock 2) raises the interest rate and lowers output growth





[BOX] Variable ordering matters!

- The `'short'` option assumes B is lower triangular $\Rightarrow$ the ordering of the variables in `X` matters
- The shock in the first equation affects *all* variables on impact; the shock in the second has *zero* contemporaneous effect on the first variable but affects all later ones; and so on
- The ordering encodes an economic assumption: the variable placed first is treated as *causally prior* to all others at the same point in time
- Choose the ordering to be consistent with economic theory and prior intuition

Common Identification Schemes

Zero Long-Run Restrictions

Zero long-run restrictions

- **Intuition** Identification is achieved by assuming that some shocks have zero cumulative effect on some of the endogenous variables in the long run
- **References** Blanchard and Quah (1989), Gali (1999)
- For example, assume that monetary policy shocks have no long-run effect on the level of output, while supply shocks do
- But how can we impose restrictions on the long-run cumulative effect of a structural shock?

Zero long-run restrictions

How to compute the cumulative long-run effects of shocks?

- Start from our standard structural VAR representation:

$$x_t = \Phi x_{t-1} + B\varepsilon_t$$

- If a shock ε_t hits in t, its **cumulative** impact on x_t in the long run is given by

$$x_{t,t+\infty} = \boxed{B\varepsilon_t} + \boxed{\Phi B\varepsilon_t} + \boxed{\Phi^2 B\varepsilon_t} + \ldots + \Phi^\infty B\varepsilon_t$$

Impact in t

etc...

Impact in t+1

- Note: for output growth, $y_{t,t+\infty}$ is the effect of ε_t on the level of output

Zero long-run restrictions

How to compute the cumulative long-run effects of shocks?

- Re-write the VAR as

$$x_t = \Phi x_{t-1} + B\varepsilon_t$$

- If a shock ε_t hits in t, its **cumulative** impact on x_t in the long run is given by

$$x_{t,t+\infty} = \underbrace{\boxed{B\varepsilon_t}}_{\text{Impact in } t} + \underbrace{\boxed{\Phi B\varepsilon_t}}_{\text{Impact in } t+1} + \underbrace{\boxed{\Phi^2 B\varepsilon_t}}_{\text{etc...}} + \ldots + \Phi^{\infty} B\varepsilon_t$$

- If the VAR is stable, we can rewrite

$$x_{t,t+\infty} = \sum_{j=0}^{\infty} \Phi^j B\varepsilon_t = (I-\Phi)^{-1} B\varepsilon_t = C\varepsilon_t$$

where $C \equiv (I-\Phi)^{-1} B$ captures the cumulative effect of ε_t on x_t from t to ∞

Zero long-run restrictions

How to compute the cumulative long-run effects of shocks?

- What is the intuition for C?

- Go back to our output growth / policy rate example. Now assume that the only two shocks driving the economy are a supply and a monetary policy shock:

$$\begin{bmatrix} y_{t,t+\infty} \\ r_{t,t+\infty} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Supply} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

- Take the first equation: $y_{t,t+\infty} = c_{11}\varepsilon_t^{Supply} + c_{12}\varepsilon_t^{MonPol}$
  * c_{12} captures the impact of a monetary policy shock (hitting in t) on the **level of GDP** in the long run
  * If you believe in long-run monetary neutrality (no long-run effect from monetary policy shocks) you would expect $c_{12} = 0$

Zero long-run restrictions

How to achieve identification?

- Remember that $C \equiv (I - \Phi)^{-1} B$ is **unobserved** as we don't know B. So, how does this help with the identification of B?

- To achieve identification define $\Omega \equiv CC'$ and note that

  1. Ω is known!
  $$\Omega = \left((I - \Phi)^{-1}\right) BB' \left((I - \Phi)^{-1}\right)' = \left((I - \Phi)^{-1}\right) \Sigma_u \left((I - \Phi)^{-1}\right)'$$

  2. Ω is a positive-definite symmetric matrix $\rightarrow$ It admits a unique Cholesky decomposition
  $$\Omega = P_\Omega P_\Omega'$$

  3. Because of our assumption that C is lower triangular, it follows that $C = P_\Omega$

- We achieved identification: $B = (I - \Phi) P_\Omega$

Zero long-run restrictions

How to achieve identification?

- As before, we can rewrite the structural VAR with the B matrix implied by the zero long-run restriction
$$B = (I_2 - \Phi)P_\Omega = (I_2 - \Phi) \times \texttt{chol}\left(\left((I - \Phi)^{-1}\right)\Sigma_u\left((I - \Phi)^{-1}\right)'\right)$$
where `chol` denotes the Cholesky factor

- Note that the B matrix **is not triangular**
  - This is different from what we had in the zero contemporaneous restrictions identification
- The impact effects are left unrestricted, the restrictions are on the C matrix
  - We'll check later that the restrictions are satisfied in a simple example with true data

[BOX] The Cholesky decomposition, again

- The same Cholesky logic as for zero short-run restrictions applies here, with the long-run matrix Ω in place of Σ_u
- Define $\Omega \equiv (I - \Phi)^{-1} \Sigma_u \left((I - \Phi)^{-1}\right)'$ — a known, positive-definite, symmetric matrix
- Ω admits a unique Cholesky factor P_Ω; the lower-triangular structure of $C = (I - \Phi)^{-1} B$ then identifies the system
- As with short-run restrictions, the number of zero restrictions needed grows *faster* than the number of variables — Cholesky is especially useful as the VAR dimension increases

Zero long-run restrictions: In Matlab

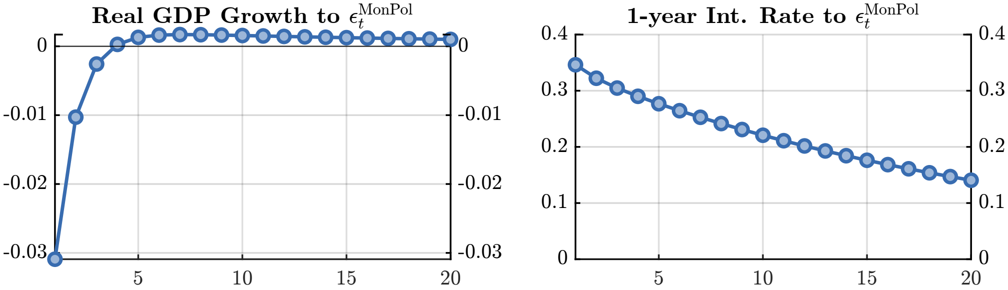
- Set `VARopt.ident = 'long'` and call `VARmodel` again, saving the results in `VAR_long`:

```
% Select zero long-run (Blanchard-Quah) identification
VARopt.ident = 'long';

% Compute impulse responses; plot the MP shock (pick=2)
VAR_long = VARmodel(X,nlags,detc,VARopt);
VARopt.pick = 2;
VARirplot(VAR_long.IR,VARopt);
```

- `VAR_long.B` contains the impact matrix consistent with the zero long-run restrictions
- `VAR_long.IR` contains the impulse responses
- Long-run multiplier matrix $C = (I - \Phi)^{-1}B$ can be verified by typing: `(eye(2)-VAR_long.Fcomp)\VAR_long.B`

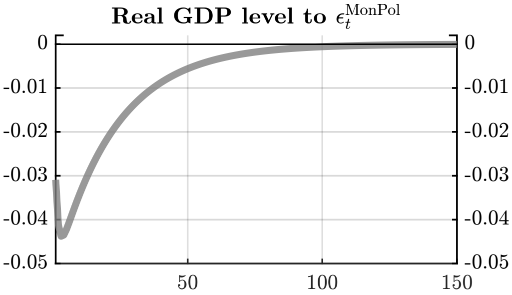
Zero long-run restrictions: Impulse responses

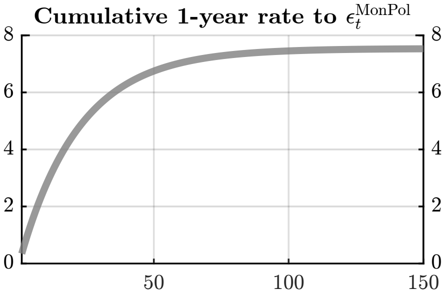
- Monetary policy shock (shock 2) has a zero long-run effect on the output level by construction



Zero long-run restrictions: Impulse responses

- Can also be verified by plotting the cumulative impulse responses over a long-enough horizon





Common Identification Schemes

Sign Restrictions

Sign restrictions

- **Intuition** Exploit prior beliefs (typically informed by theoretical models) about the sign that certain shocks should have on certain endogenous variables
- **Reference** Faust (1998), Canova and Nicolo (2002), Uhlig (2005)
- For example
  - Demand shocks should lead to an increase in output and interest rates
  - Monetary policy shocks should lead to a fall in output and an increase in interest rates

	Demand (ε_t^{Demand})	Monetary Policy (ε_t^{MonPol})
Output growth (y_t)	+	-
Short-rate Int. Rate (r_t)	+	+

- But how can we impose restrictions on the signs of the effect of a structural shock?

Sign restrictions
How to achieve identification?

- The key intuition is based on the following three steps
  1. Consider a **random orthonormal matrix** Q such that
  $$QQ' = I_2$$
  2. Consider the lower triangular B matrix corresponding to the Cholesky factor of Σ_u
  $$\Sigma_u = BB' = PP'$$
  3. The following equality holds
  $$\Sigma_u = PP' = PQQ'P' = \underbrace{(PQ)}_{B}\underbrace{(PQ)'}_{B'}$$
- The matrix $B = PQ$ is a valid 'candidate' impact matrix that solves the identification problem!
  * Differently from P, the matrix PQ is not lower triangular anymore

[BOX] Orthogonal matrix

- An orthogonal matrix Q is a real square matrix whose columns form an orthonormal set
- What does it mean? Take for example two 2×1 vectors q_1 and q_2, then the matrix $Q = (q_1, q_2)$ is orthogonal if
  * The vectors have unit norm: $\| q_i \| = 1$
  * The vectors are mutually orthogonal: $q_1^T q_2 = 0$
- It follows that

$$QQ' = I \quad \text{and} \quad Q' = Q^{-1}$$

- **Note** You can draw infinite matrices that satisfy the above conditions (we'll see how to do it in Matlab below)

Sign restrictions

How to achieve identification?

- But Q is a random matrix... How can we check that $B = PQ$ represents a plausible solution?
- **Solution** Check that the effects of shocks implied by $B = PQ$ satisfy a set of a priori sign restrictions. That is:

[1] Consider the structural representation of our VAR

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

[2] Then check that the elements of B satisfy

	Demand (ε_t^{Demand})	Monetary Policy (ε_t^{MonPol})
Output growth (y_t)	$b_{11} > 0$?	$b_{12} < 0$?
Short-rate Int. Rate (r_t)	$b_{21} > 0$?	$b_{22} > 0$?

Sign restriction in steps

- Perform N replications of the following steps
  - [1] Draw a random orthonormal matrix Q
  - [2] Compute $B = PQ$ where P is the Cholesky decomposition of the reduced-form residuals Σ_u
  - [3] Compute the impact effects of shocks associated with B
  - [4] Are the sign restrictions satisfied?
    - [4.1] Yes. Store B and go back to [1]
    - [4.2] No. Discard B and go back to [1]

- All matrices in the set $B^{(i)}$ (for $i = 1, 2, \ldots, N$) represent admissible solutions to the identification problem

- In this sense, sign-restricted VARs are only **set identified**

[BOX] Set identification vs. point identification

- **Point-identified** (Cholesky, long-run) A *unique* B satisfies the restrictions. The only source of uncertainty is sampling variability
- **Set-identified** (sign restrictions) *Many* B_j are consistent with the data and the restrictions — genuinely different structural models, even in an infinite sample
- Consequence 1: the median IRF across draws need not converge to a unique value as $T \longrightarrow \infty$ — it summarizes the *set*, not a point estimate
- Consequence 2: credible bands reflect *both* sampling uncertainty **and** the width of the identified set; the two cannot be disentangled without further restrictions

Sign restrictions: In Matlab

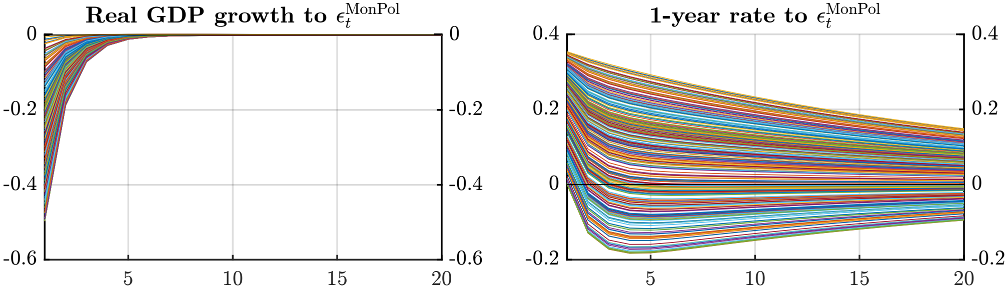
- Encode restrictions in `R` ($+1$ positive, -1 negative, 0 unrestricted)
- Set `VARopt.ident = 'sign'`, and call `VARmodel` again, saving the results in `VAR_sr`

```
% Sign restriction matrix: positive 1, negative -1, unrestricted 0
R = [ 1, -1;   % Real GDP
      1,  1]; % 1-year rate

% Number of random draws and restriction horizon
VARopt.ndraws = 1000;
VARopt.sr_hor = 1;

% Identify with sign restrictions; print acceptance rate as a diagnostic
VARopt.ident = 'sign';
VARopt.R = R;
VAR_sr = VARmodel(X,nlags,detc,VARopt);
```

- All accepted draws of the impact matrix stored in `VAR_sr.Ball` (dimension $2 \times 2 \times 1000$)

Sign restrictions: Impulse responses

- All accepted rotations

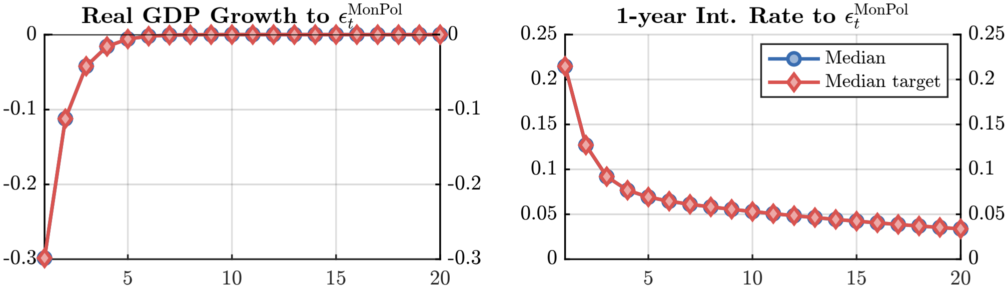


[BOX] Element-wise median vs. Fry–Pagan median-target rotation

- **Element-wise median** (`VAR_sr.IRmed`) Coordinate-by-coordinate median across accepted draws. Convenient, but the implied impact matrix `VAR_sr.Bmed` matches *no* accepted draw and need not satisfy $\Sigma_u = B_{\text{med}} B'_{\text{med}}$ — a 'Frankenstein' object ⇒ VD shares need not sum to one, HD need not reproduce the data
- **Fry–Pagan median-target** (`VAR_sr.IRfp`) The single accepted draw whose responses are closest (least squares) to the element-wise median. A *genuine* structural model: it satisfies the sign restrictions and $\Sigma_u = B_{j^*} B'_{j^*}$ ⇒ VD sums to one, HD reproduces the data; cost: it selects one draw from the set
- **Why they differ** Admissible impact matrices lie on a curved set (orthonormal rotations of a Cholesky factor of Σ_u); the coordinate-wise median generally lies *off* it. Present even with `inference = 0`, and grows with the VAR dimension and the width of the identified set — hence tiny here (next slide)

Sign restrictions: Impulse responses

- Median rotation vs. Fry–Pagan median-target rotation



Common Identification Schemes

Narrative Sign Restrictions

Narrative sign restrictions

- **Intuition** Sign restrictions may leave a wide identified set. *Narrative restrictions* exploit well-documented historical episodes to sharpen identification by screening out rotations that are sign-consistent but historically implausible

- **References** Antolín-Díaz and Rubio-Ramírez (2018)

- Historical records about specific events carry two types of information
  - The **sign** of a shock at a particular date (e.g. the Volcker tightening was a positive MP shock)
  - Which shock was the **dominant driver** of a particular variable at a particular date

Narrative restrictions: Two types

- Applied as additional acceptance criteria *after* the sign restrictions are satisfied

- **Type 1 - sign at a date:** At date t^*, the m-th structural shock must have a specified sign

$$s \cdot \varepsilon_{m,t^*} > 0, \qquad s \in \{+1, -1\}$$

- **Type 2 - dominance at a date:** At date t^*, shock m must account for more of the unexpected movement in variable i than all other shocks combined

$$\left| b_{im}\, \varepsilon_{m,t^*} \right| > \sum_{k \neq m} \left| b_{ik}\, \varepsilon_{k,t^*} \right|$$

where b_{ik} is the (i, k) element of B

Narrative sign restrictions in steps

- Perform N replications of the following steps
  - [1] Draw a random orthonormal matrix Q
  - [2] Compute $B = PQ$ and check sign restrictions
  - [3] Also check that the implied structural shocks $\varepsilon_t = B^{-1}u_t$ satisfy the narrative restrictions at each specified date t^*
    - [3.1] All restrictions satisfied: retain B. Go back to [1]
    - [3.2] Any restriction violated: discard B. Go back to [1]
- The accepted draws form a **smaller, differently located set** than sign restrictions alone
- A large decrease in acceptance rate signals that the narrative evidence is genuinely informative

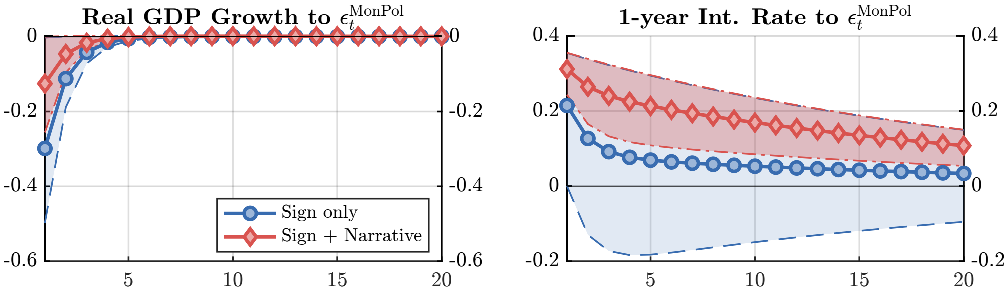
Narrative sign restrictions: In Matlab

- Build the `R` (now a structure) with `R.sign`, `R.narr_sign`, and `R.narr_dom`
- Keep `VARopt.ident='sign'`, call `VARmodel` again, saving the results in `VAR_nsr`

```
% Sign restrictions
R.sign = [ 1, -1;   % Real GDP
           1,  1]; % 1-year rate

% Sign of MP shock for both episodes
R.narr_sign.shock  = [2;        2      ];
R.narr_sign.period = {'1994q1'; '2001q1'};
R.narr_sign.sign   = [1;       -1      ];

% MP shock is the dominant driver of i1yr only in 1994q1
R.narr_dom.shock  = [2;        ];
R.narr_dom.period = {'1994q1'; };
R.narr_dom.var    = [2;        ];

% Identify with sign + narrative restrictions; print acceptance rate as a diagnostic
VARopt.R = R;
VAR_nsr = VARmodel(X, nlags, detc, VARopt);
```

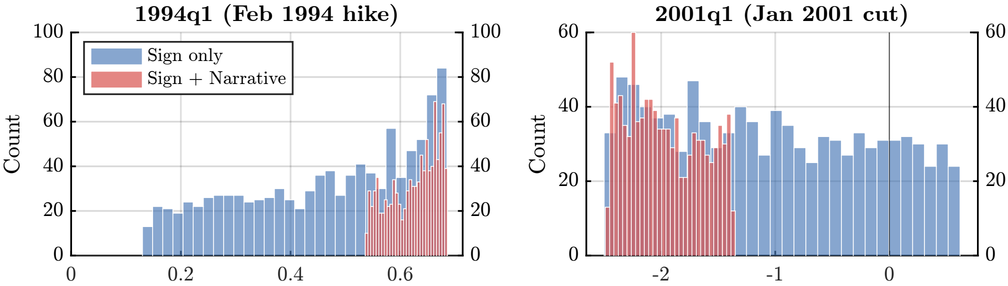
Narrative vs. Sign-only impulse responses

- Narrative restrictions shrink the identified set and shift the median



Narrative vs. Sign-only impulse responses

- The additional constraints screen out sign-consistent but historically implausible rotations



Common Identification Schemes

External Instruments (aka Proxy SVARs)

External instruments

- **Intuition** Exploit information from a variable that is *external* to the VAR, but that is correlated with a particular shock of interest and uncorrelated with other shocks (the *instrument*)
- **References** Stock and Watson (2012), Mertens and Ravn (2013)
- For example, assume that you have some 'narrative' series of policy surprises (i.e. that are not just a response of the central bank to some development in the economy)
- But how can this help in finding the B matrix?

External instruments

- **Key element** Presence of an *instrument* that is correlated with a shock of interest and uncorrelated with all other shocks

- For example, assume that such an instrument exists (z_t) and satisfies the following properties:

$$\begin{aligned} \mathbb{E}[\varepsilon_t^{Other} z_t'] &= 0, \\ \mathbb{E}[\varepsilon_t^{MonPol} z_t'] &= \alpha, \end{aligned}$$

- Then, we can identify one column (in this example, the second one) of the B matrix:

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} - & b_{12} \\ - & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Other} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

External instruments identification: How does it work?

- Recall that the reduced-form residuals are a linear combination of the two structural shocks

$$\begin{cases} u_{yt} = b_{11}\varepsilon_t^{Other} + b_{12}\varepsilon_t^{MonPol} \\ u_{rt} = b_{21}\varepsilon_t^{Other} + b_{22}\varepsilon_t^{MonPol} \end{cases}$$

- The OLS estimate of β in the 'first stage' regression identifies b_{22} up to a scaling factor

$$u_{rt} = \beta z_t + \xi_t$$

- To see that, recall that the OLS β can be written as $\beta = Cov(u_{rt}, z_t)/Var(z_t)$. Focus on the Cov term and plug in the definition of u_{rt} to get

$$Cov(u_{rt}, z_t) = Cov(b_{21}\varepsilon_t^{Other} + b_{22}\varepsilon_t^{MonPol}, z_t) = b_{22}Cov(\varepsilon_t^{MonPol}, z_t) = b_{22}\alpha \rightarrow \quad \beta = \frac{b_{22}\alpha}{Var(z_t)}$$

- As α is an unknown constant, $b_{22} = \beta Var(z_t)/\alpha$ is only identified to a scaling factor

External instruments identification: How does it work?

- The OLS estimate of γ in the following 'second stage' regression identifies the ratio b_{12}/b_{22}

$$u_{yt} = \gamma \hat{u}_{rt} + \zeta_t = \gamma \left(\frac{b_{22}\alpha}{Var(z_t)} \right) z_t + \zeta_t$$

- To see that, and recalling again that $\gamma = Cov(u_{yt}, \hat{u}_{rt})/Var(\hat{u}_{rt})$
  - Focus on the Cov term and plug in the definition of u_{yt} and $\hat{u}_{rt}$ to get

$$Cov(u_{yt}, \hat{u}_{rt}) = Cov\left(b_{11}\varepsilon_t^{Other} + b_{12}\varepsilon_t^{MonPol}, \frac{b_{22}\alpha}{Var(z_t)} z_t \right) = \frac{b_{12}b_{22}\alpha^2}{Var(z_t)}$$

  - Then focus on the Var term to get

$$Var(\hat{u}_{rt}) = Var\left(\frac{b_{22}\alpha}{Var(z_t)} z_t \right) = \frac{b_{22}^2\alpha^2}{Var\left(z_t\right)}$$

  - It follows that $\gamma = \frac{b_{12}}{b_{22}}$

External instruments: Partial identification

- In sum, we can normalize the effect of ε_t^{MonPol} on r_t to 1

$$b_{22} = 1$$

- And quantify the effect of ε_t^{MonPol} on y_t as

$$b_{12} = \gamma$$

- In other words, we have identified the column of the B matrix of the structural VAR representation up to a scaling factor

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} - & \gamma \\ - & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Other} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

  * **Note** It is actually possible to work out the true values of the second column of B (footnote 4 of Gertler and Karadi (2015))

External instruments: In Matlab

- Pass the instrument via `VARopt.IV` and set `VARopt.ident = 'iv'`
- Call `VARmodel` again, saving the results in `VAR_iv`:

```
% Set identification to 'iv', attach instrument, and compute impulse responses
VARopt.ident     = 'iv';
VARopt.IV        = mps;
VARopt.snames    = {bfeps('MonPol'), bfeps('Other')};
VAR_iv = VARmodel(X,nlags,detc,VARopt);
```

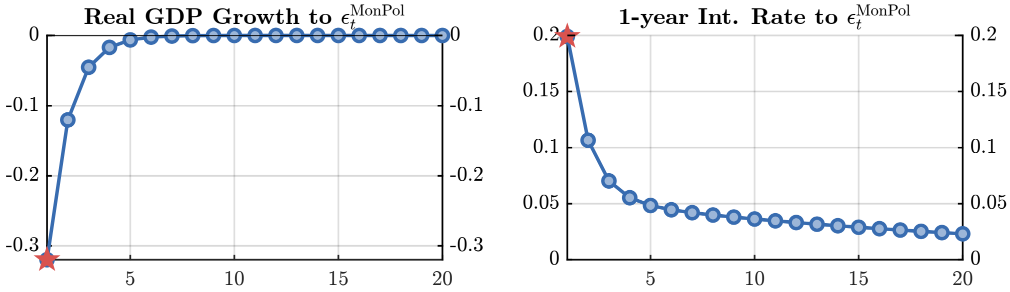
- IV-identified column: `VAR_iv.B(:,1)`; first-stage output in `VAR_iv.FirstStage`

[BOX] Warning: the instrument mps is fictitious

- **Warning** The instrument `mps` used in this example is fictitious and must *never* be used in applied work
- A valid instrument needs *relevance* (correlated with the target shock) **and** *exogeneity* (orthogonal to all other shocks). A canonical example: the high-frequency MP surprises of Gertler and Karadi (2015)
- Here `mps` = sign-restriction MP shock + random noise: it makes no economic sense, but works mechanically
- The bivariate VAR(1) is too parsimonious for any real-world instrument — the example illustrates the *mechanics* of proxy SVARs only

External instruments: Impulse responses

- Impact responses (stars) from the first/second-stage regressions; dynamic responses from the estimated VAR



[BOX] Assessing instrument strength: the first-stage F-statistic

- The first-stage F-statistic tests whether the instrument is sufficiently correlated with the reduced-form residual it is meant to instrument
- A *weak* instrument — low partial correlation with the target — produces biased and imprecise IV estimates, *regardless* of whether exogeneity holds
- Rule of thumb (Stock and Yogo, 2002): the first-stage F should exceed 10 (one instrument, one regressor, 5% size-distortion tolerance)
- In the toolbox: `VAR_iv.FirstStage.F`. Gertler and Karadi (2015) report F well above 10

Common Identification Schemes

Combining Sign Restrictions & External Instruments

Combining sign restrictions & external instruments

- **Intuition** Identifies one (or more) columns of B with external instruments and conditional on that the remaining columns with sign restrictions

- **References** Cesa-Bianchi and Sokol (2021)

- For example, assume that there are two shocks that imply similar signs (so that sign restrictions are not enough to identify the shocks), but you have an instrument for one of the two shocks

- How can we find the B matrix?

Combining sign restrictions & external instruments

- Consider a k-variable version of our simple structural VAR(1)

$$\begin{bmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{kt} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1k} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{k1} & \phi_{k2} & \cdots & \phi_{kk} \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ \vdots \\ x_{kt-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kk} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{kt} \end{bmatrix}$$

- Assume that
  - The first structural shock (ε_{1t}) can be identified with an external instrument
  - The remaining structural shocks ($\varepsilon_{2t}, \ldots, \varepsilon_{kt}$) can be identified with sign restrictions

Combining sign restrictions & external instruments

- Partition the structural matrix B as $[b\ \mathcal{B}]$

$$B = \left[\underbrace{\begin{matrix} b_{11} \\ b_{21} \\ \vdots \\ b_{k1} \end{matrix}}_{b} \quad \underbrace{\begin{matrix} b_{12} & \cdots & b_{1k} \\ b_{22} & \cdots & b_{2k} \\ \vdots & \ddots & \vdots \\ b_{k2} & \cdots & b_{kk} \end{matrix}}_{\mathcal{B}} \right]$$

- Column vector b captures the impact of the first shock, matrix $\mathcal{B}$ captures the impact of the remaining shocks

- We have seen above how to identify b with external instruments

- **Question** Once b is known, how can we find a $\mathcal{B}$ matrix that satisfies a set of sign restrictions?

Combining sign restrictions & external instruments

- Let P be the Cholesky decomposition of Σ_u. Find a normal vector q of dimension $k \times 1$ that rotates the first column of P into the vector b, so that

$$Pq = b$$

- Given q, build a $(k \times (k-1))$ matrix $\mathcal{Q}$ such that $Q = [q\ \mathcal{Q}]$ is orthonormal

$$[q\ \mathcal{Q}][q\ \mathcal{Q}]' = QQ' = I_k$$

- As Q is an orthonormal matrix we have

$$\Sigma_u = PP' = PQQ'P' = (PQ)(PQ)'$$

- So $B = PQ$ is a valid candidate matrix that solves the identification problem as
  - $\Sigma_u = (PQ)(PQ)'$ holds
  - The first column of PQ is b

Combining sign restrictions & external instruments: Steps

[1] Identify b, the first column of $B = [b\ \mathcal{B}]$, with the external instrument

[2] Compute the Cholesky decomposition P of the reduced-form residuals' covariance matrix Σ_u

[3] Find a normal vector q that rotates the first column of P into the vector b, namely $Pq = b$

[3.i] Given q, build the remaining $k-1$ columns of an orthonormal matrix $Q = [q\ \mathcal{Q}]$

[3.ii] The matrix PQ then represents a candidate identification scheme because:

$$(PQ)(PQ)' = \Sigma_u \quad \text{and} \quad P[q\ \mathcal{Q}] = [b\ \mathcal{B}]$$

[3.iii] If $\mathcal{B}$ satisfies the sign restrictions, retain it. Otherwise, go back to [3.i]

[4] Go back to [1] and repeat N times

Combining sign restrictions & external instruments: In Matlab

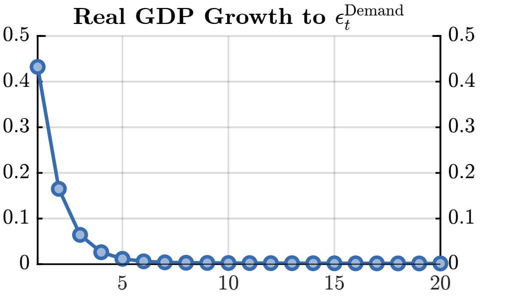
- Single `VARmodel` call with `ident='sign+iv'`: IV stage pins column 1 of B, sign restrictions rotate the orthogonal complement

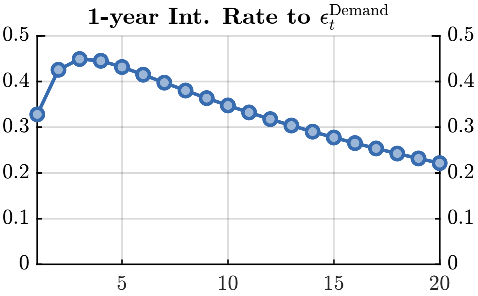
```
% Sign restriction for the demand shock (conditional on the IV-identified MP column)
% Positive 1, Negative -1, Unrestricted 0:
R = [ 1;  % Real GDP: positive
1]; % 1-year rate: positive

% Figure path and shock names
VARopt.figname = 'graphics/iv_sign';
VARopt.snames = {bfeps('MonPol'), bfeps('Demand')};

% Run hybrid identification: ident='sign+iv' handles IV stage and sign rotation
% in a single call. VARopt.IV = mps is already set from section 8.
VARopt.ident = 'sign+iv';
VARopt.R = R;
VAR_ivsr = VARmodel(X, nlags, detc, VARopt);
```

- IV column fixed at `VAR_ivsr.Bmed(:,1)` across all draws; median demand-shock responses in `VAR_ivsr.IRmed`

Combining sign restrictions & external instruments: Impulse responses

- Demand shock responses; MP column is fixed at the IV estimate for every draw





Common Identification Schemes

Identification with Exogenous Variables

Identification with exogenous variables

- **Intuition** Suppose you have at your disposal an exogenous shock s_t
  - It is *contemporaneously exogenous* to the VAR system
  - Rather than using it as an instrument in a second-stage projection ...
  - ...include it *directly* as a contemporaneous regressor in the VAR
- The model becomes

$$x_t = c + \sum_{j=1}^{p} \Phi_j x_{t-j} + \delta s_t + u_t$$

where δ is a $k \times 1$ vector of impact coefficients

- The OLS estimate $\hat{\delta}$, scaled by the sample standard deviation of s_t, gives the first column of B: the contemporaneous response to a one-standard-deviation shock

Exogenous variables vs. Proxy SVAR

- Both approaches exploit external variation in the same observed series to identify one structural shock
- The key distinction
  - **Proxy SVAR**: z_t enters as an instrument in a second-stage IV regression on the reduced-form residuals
  - **Exogenous variable**: s_t enters as a contemporaneous regressor *directly* in each VAR equation
- Under the same identification assumptions and a common lag structure, the two estimators are **asymptotically equivalent** Plagborg-Møller and Wolf (2021)
- Any divergence in finite samples is informative about potential VAR misspecification

Identification with exogenous variables: In Matlab

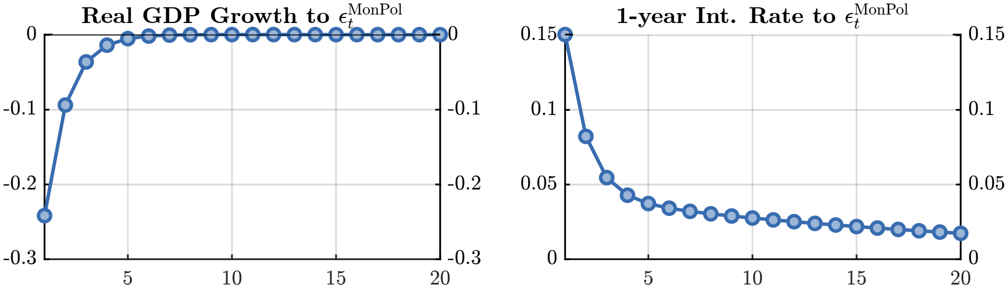
- Pass the exogenous shock via `VARopt.exoshock` and set `VARopt.ident = 'exog'`

```
VARopt_exog = VARoption;
VARopt_exog.vnames     = Xvnames;
VARopt_exog.nsteps     = 20;
VARopt_exog.ident      = 'exog';
VARopt_exog.exoshock   = mps;
VARopt_exog.snames     = {bfeps('MonPol')};
VAR_exog = VARmodel(X, nlags, detc, VARopt_exog);
```

- First column of `VAR_exog.B` equals $\hat{\delta}\hat{\sigma}_s$; remaining columns are zero

Identification with exogenous variables: Impulse responses

- Responses virtually identical to the proxy SVAR at all horizons (asymptotic equivalence)



Inference

Quantifying uncertainty

- Impulse responses, variance decompositions, and historical decompositions are functions of estimated parameters $(\hat{\Phi}, \hat{\Sigma}_u)$ — they inherit sampling uncertainty
- Two distinct situations require different inferential approaches
  - **Point-identified** models (Cholesky, long-run, IV, exogenous): unique B; only source of uncertainty is parameter uncertainty ⇒ **bootstrap**
  - **Set-identified** models (sign restrictions, narrative): B is set-identified; uncertainty has two components: parameter uncertainty + identification uncertainty ⇒ **Bayesian draws**

Bootstrap for point-identified models

- Activate with `VARopt.inference = 1`
- Two variants, selected via `VARopt.method`
  - `'bs'` (residual bootstrap): resample residuals with replacement; preserves autocorrelation structure but assumes homoskedastic residuals
  - `'wild'`: multiply each residual by an independent ± 1 Rademacher draw; robust to conditional heteroskedasticity; default for proxy SVARs

```
VARopt.inference = 1;        % activate bootstrap
VARopt.method    = 'bs';     % 'bs' = residual bootstrap (default)
VARopt.ndraws    = 1000;     % number of bootstrap draws (default)
VARopt.pctg      = 90;       % confidence level in percent
VAR_id = VARmodel(X, nlags, detc, VARopt);
```

Bayesian inference for sign-restricted models

- For sign-restricted models, the accepted draws $\{B_j\}$ carry **both** sources of uncertainty: the rotation draws capture identification uncertainty; an independent draw of VAR parameters from a Normal-inverse-Wishart posterior captures parameter uncertainty
- Control parameter uncertainty via `VARopt.inference`
  - `inference = 1` (default): both sources of uncertainty
  - `inference = 0`: identification uncertainty only (OLS parameters fixed)
- Credible bands stored in `VAR_sr.IRinf` and `VAR_sr.IRsup`; pointwise median in `VAR_sr.IRmed` (element-wise median across draws — not a valid structural model)

[BOX] Controlling parameter uncertainty in sign-restricted VARs

- `VARopt.inference = 1` (default) Draws jointly from the posterior over (Φ, Σ_u) *and* over admissible rotations $\Rightarrow$ both parameter **and** identification uncertainty
- `VARopt.inference = 0` Fixes (Φ, Σ_u) at their OLS estimates and randomises only over rotations $\Rightarrow$ identification uncertainty *alone*
- The second setting is useful for isolating the width of the identified set from sampling variability in the reduced-form estimates

Dynamic Analysis

Structural dynamic analysis

- Assume for the moment that we know how to 'identify' structural shocks $\longrightarrow$ B matrix is known
- What can we do with our structural VAR?
  - Quantify the dynamic effect of a shock over time $\Rightarrow$ **Impulse responses**
  - Quantify how important a shock is in explaining the variation in the endogenous variables (on average) $\Rightarrow$ **Forecast error variance decomposition**
  - Quantify how important a shock was in driving the behavior of the endogenous variables in a specific time period in the past $\Rightarrow$ **Historical decompositions**
- This is what is called structural dynamic analysis

Dynamic Analysis

Impulse Responses

Impulse response functions

- Impulse response functions (*IR*) answer the following question:

> What is the response over time of the VAR's endogenous variables to an innovation in the structural shocks, assuming that the other structural shocks are kept to zero?

- *IR* allow us to single out the effect of a shock (e.g. its impact and persistence on all variables in the system) keeping all else equal

- **Example** What is the impact of a monetary policy shock to GDP?

How to compute impulse response functions

- Consider our simple bivariate VAR

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

- Define a 2×1 impulse selection vector (s) that takes value of one for the structural shock that we want to consider.

- For example, to compute the IR to the demand shock, define s as:

$$s = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- The impulse responses to ε_t^{Demand} can be easily computed with the following equation

$$x_t = \Phi x_{t-1} + Bs$$

How to compute impulse response functions (cont'd)

- The IR can be computed recursively as follows

$$\begin{cases} IR_h = Bs & \text{for } h = 0 \\ IR_h = \Phi IR_{h-1} & \text{for } h = 1, \ldots, H \end{cases}$$

- Note that the impact response is simply given by the elements of the impact matrix B selected by s...

$$\begin{bmatrix} IR_0^y \\ IR_0^r \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix}$$

- ... while the responses at longer horizons are given the transition matrix

$$\begin{bmatrix} IR_t^y \\ IR_t^r \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} IR_{t-1}^y \\ IR_{t-1}^r \end{bmatrix}$$

[BOX] The companion matrix

- Every VAR(p) can be written as a VAR(1) using the **companion representation**
  - For example, take a VAR(2)

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11}^1 & \phi_{12}^1 \\ \phi_{21}^1 & \phi_{22}^1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \phi_{11}^2 & \phi_{12}^2 \\ \phi_{21}^2 & \phi_{22}^2 \end{bmatrix} \begin{bmatrix} y_{t-2} \\ r_{t-2} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

  - Re-write the VAR(2) as

Companion matrix

$$\begin{bmatrix} y_t \\ r_t \\ y_{t-1} \\ r_{t-1} \end{bmatrix} = \boxed{\begin{bmatrix} \phi_{11}^1 & \phi_{12}^1 & \phi_{11}^2 & \phi_{12}^2 \\ \phi_{21}^1 & \phi_{22}^1 & \phi_{21}^2 & \phi_{22}^2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}} \begin{bmatrix} y_{t-1} \\ r_{t-1} \\ y_{t-2} \\ r_{t-2} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \\ 0 \\ 0 \end{bmatrix}$$

  - Which is a VAR(1) where $\tilde{\Phi}$ is the **companion matrix**

$$\tilde{x}_t = \tilde{\Phi}\tilde{x}_{t-1} + \tilde{B}\varepsilon_t$$

Impulse responses: In Matlab

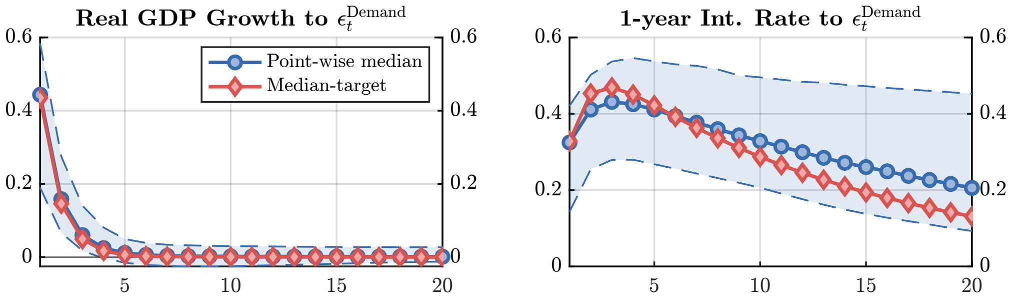
- With identification set, `VARmodel` computes IRFs; `VARirplot` plots them
- Use as an example the sign-restricted VAR from before
- Plot responses to monetary policy shock (shock 2) with confidence bands from the Bayesian posterior

```
VARopt.inference = 1;        % enable Bayesian credible bands
VAR_sr = VARmodel(X,nlags,detc,VARopt);
VARopt.figname = 'graphics/sign_IR';
VARirplot(VAR_sr.IRmed, VARopt, VAR_sr.IRinf, VAR_sr.IRsup);
```

- Growth-rate vs. level IRFs: cumulating output growth responses gives the **output level** response (relevant for long-run neutrality checks)

Impulse responses: Sign-restricted bivariate VAR

- Shaded area = 95% credible bands; solid line = element-wise median



[BOX] Growth-rate IRFs vs. level IRFs

- The toolbox computes IRFs for the variables *as they enter the VAR* — here, the growth rate of GDP. IR_h^y is the deviation of the growth rate from baseline at horizon h
- For the *level* of output, the relevant object is the **cumulative** IRF:

$$CIR_h^y = \sum_{i=0}^{h} IR_i^y$$

- The long-run (monetary neutrality) restriction is $CIR_\infty^y = 0$ — a restriction on the *cumulative* sum, not on any individual growth-rate response
- Growth-rate responses need not return to zero; cumulative (level) responses converge to zero by construction under neutrality

Dynamic Analysis

Forecast Error Variance Decompositions

Forecast error variance decompositions

- Forecast error variance decompositions (*VD*) answer the following question:

  > *What portion of the variance of the VAR's forecast errors (at a given horizon h) is due to each structural shock?*

- *VD* provide information about the relative importance of each structural shock in affecting the forecast error variance of the VAR's endogenous variables

- **Example** What is the (average) importance of demand shocks in driving GDP forecast errors?

How to compute forecast error variance decompositions

- The forecast error of a variable at horizon $t + h$ is the change in the variable that couldn't have been forecast between $t - 1$ and $t + h$ due to the realization of the structural shocks.

- For example, at $h = 0$ we can compute the forecast error as

$$x_t - \mathbb{E}_{t-1}[x_t] = \Phi x_{t-1} + B\varepsilon_t - \Phi x_{t-1} = B\varepsilon_t$$

- At $h = 1$, we have

$$\begin{aligned} x_{t+1} - \mathbb{E}_{t-1}[x_{t+1}] &= \Phi x_t + B\varepsilon_{t+1} - \Phi^2 x_{t-1} = \\ &= \Phi(\Phi x_{t-1} + B\varepsilon_t) + B\varepsilon_{t+1} - \Phi^2 x_{t-1} = \Phi B\varepsilon_t + B\varepsilon_{t+1} \end{aligned}$$

- So, in general we have

$$FE_{t+h} = x_{t+h} - E_{t-1}[x_{t+h}] = \sum_{i=0}^{h} \Phi^{h-i} B\varepsilon_{t+i}$$

- What is the variance of FE_{t+h}?

[BOX] Basic properties of the variance

- If X is a random variable x and a is a constant
  - $Var\left(x+a\right)=Var\left(x\right)$
  - $Var\left(ax\right)=a^2 Var\left(x\right)$
- If Y is a random variable and b is a constant
  - $Var\left(aX+bY\right)=a^2 Var\left(x\right)+b^2 Var\left(Y\right)+2abCov\left(X,Y\right)$
- Since the structural errors are independent, it follows that $COV\left(\varepsilon_{t+1}^{Demand},\varepsilon_{t+1}^{MonPol}\right)=0$

How to compute forecast error variance decompositions (cont'd)

- For simplicity consider $h = 0$, namely

$$Var\left(FE_t\right) = Var\left(x_t - E_{t-1}[x_t]\right) = Var\left(B\varepsilon_t\right)$$

- Recalling that $Var\left(\varepsilon_t\right) = I_2$ and that the structural shocks are orthogonal to each other, the variance of the forecast error can be computed as

$$\begin{aligned} Var\left(y_t - E_{t-1}[y_t]\right) &= b_{11}^2 Var\left(\varepsilon_t^{Demand}\right) + b_{12}^2 Var\left(\varepsilon_t^{MonPol}\right) = b_{11}^2 + b_{12}^2 \\ Var\left(r_t - E_{t-1}[r_t]\right) &= b_{21}^2 Var\left(\varepsilon_t^{Demand}\right) + b_{22}^2 Var\left(\varepsilon_t^{MonPol}\right) = b_{21}^2 + b_{22}^2 \end{aligned}$$

- What portion of the variance of the forecast error at $h = 0$ is due to each structural shock?

$$\underbrace{\left\{ \begin{aligned} VD_{y_0}^{\varepsilon^{Demand}} &= \frac{b_{11}^2}{b_{11}^2 + b_{12}^2} \\ VD_{y_0}^{\varepsilon^{MonPol}} &= \frac{b_{12}^2}{b_{11}^2 + b_{12}^2} \end{aligned} \right.}_{\text{This sums up to 1}} \qquad \underbrace{\left\{ \begin{aligned} VD_{r_0}^{\varepsilon^{Demand}} &= \frac{b_{21}^2}{b_{21}^2 + b_{22}^2} \\ VD_{r_0}^{\varepsilon^{MonPol}} &= \frac{b_{22}^2}{b_{21}^2 + b_{22}^2} \end{aligned} \right.}_{\text{This sums up to 1}}$$

Forecast error variance decomposition: Formula

- The VD of variable i at horizon h attributable to shock j is

$$VD_{i,h}^{\varepsilon^j} = \frac{\sum_{s=0}^{h} (\theta_{ij}^{s})^2}{\sum_{s=0}^{h} \sum_{m=1}^{k} (\theta_{im}^{s})^2}$$

where θ_{ij}^{s} is the (i, j) element of $\Theta_s \equiv \Phi^s B$ (the impulse response matrix at horizon s)

- Three properties
  - **Adding-up:** VD shares sum to 1 across shocks at each horizon
  - **Long-run:** as $h \to \infty$, VD converges to the long-run contribution
  - **Identification:** FEVD inherits whatever is recovered by the identification scheme

FEVD: In Matlab

- VD is computed and stored automatically by `VARmodel`, plotted by `VARvdplot` with two 'modes'
- Stacked area chart

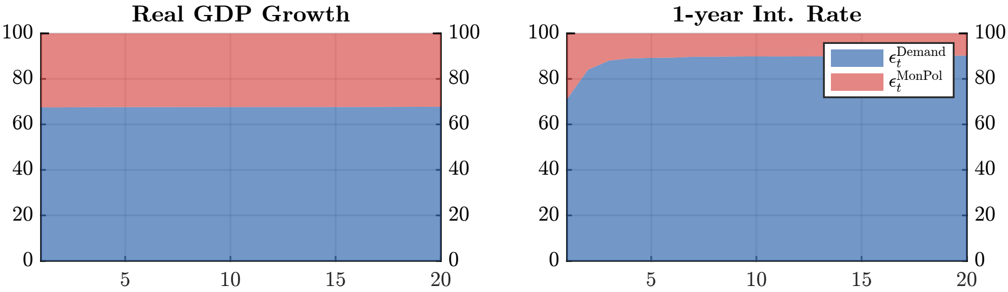
```
VARopt.figname = 'graphics/sign_VD';
VARvdplot(VAR_sr.VDmed, VARopt);
```

- Line with confidence intervals

```
VARopt.figname = 'graphics/sign_VD_bands';
VARopt.pick    = 2;
VARopt.color   = pantone('Tomato');
VARvdplot(VAR_sr.VDmed, VARopt, VAR_sr.VDinf, VAR_sr.VDsup);
VARopt.pick    = 0;
VARopt.color   = [];
```

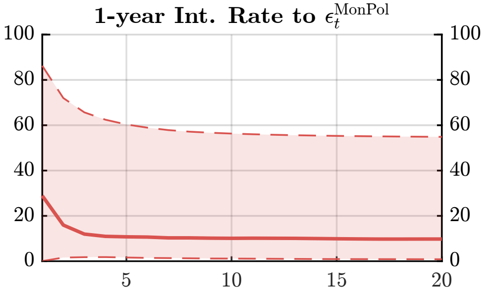
FEVD: Sign-restricted bivariate VAR

- **Stacked area**: contribution of each shock to forecast error variance share at every horizon



FEVD: Uncertainty bands

- **Bands**: MP shock contribution with Bayesian credible bands across accepted sign-restriction draws



[BOX] Three properties of the variance decomposition

- **Adding-up** For any variable and horizon, the shares across all k shocks sum to one: $\sum_{j=1}^{k} VD_{y,h}^{\varepsilon^j} = 1$ (under full identification; under partial identification they need not)
- **Long-run limit** For a covariance-stationary VAR, as $h \to \infty$ the FEVD converges to the unconditional variance shares (depending on both Φ and B)
- **Inherits identification** Shares are computed from $\Theta_i = \Phi^i B$, so they depend on the *full* B matrix. Unlike IRFs — where one column of B suffices — the FEVD share of shock j requires the full B, since the denominator aggregates variance across all k shocks

Dynamic Analysis

Historical Decompositions

Historical decompositions

- Historical decompositions (*HD*) answer the following question:

  > What is the historical contribution of each structural shock in driving deviations of the VAR's endogenous variables away from their equilibrium?

- *HD* allow us to track, at each point in time, the role of structural shocks in driving the VAR's endogenous variables away from their steady state

- **Example** What was the contribution of oil shocks in driving the fall in GDP growth in 1973:Q4?

How to compute historical decompositions

- As an example, let's compute the *HD* of the endogenous variables for $t = 2$ in our simple bivariate VAR

- Historical decompositions can be easily understood from the Wold representation of the VAR

$$x_t = \Phi^t x_0 + \sum_{j=0}^{t-1} \Phi^j c + \sum_{j=0}^{t-1} \Phi^j B \varepsilon_{t-j}$$

- Using the Wold representation, we can write x_2 as a function of present and past structural shocks, the initial condition, and the constant

$$x_2 = \underbrace{\Phi^2 x_0}_{init} + \underbrace{(I + \Phi)c}_{const} + \underbrace{\Phi B}_{\Theta_1} \varepsilon_1 + \underbrace{B}_{\Theta_0} \varepsilon_2$$

- Re-write x_2 in matrix form

$$\begin{bmatrix} y_2 \\ r_2 \end{bmatrix} = \begin{bmatrix} init_y \\ init_r \end{bmatrix} + \begin{bmatrix} const_y \\ const_r \end{bmatrix} + \begin{bmatrix} \theta_{11}^1 & \theta_{12}^1 \\ \theta_{21}^1 & \theta_{22}^1 \end{bmatrix} \begin{bmatrix} \varepsilon_1^{Demand} \\ \varepsilon_1^{MonPol} \end{bmatrix} + \begin{bmatrix} \theta_{11}^0 & \theta_{12}^0 \\ \theta_{21}^0 & \theta_{22}^0 \end{bmatrix} \begin{bmatrix} \varepsilon_2^{Demand} \\ \varepsilon_2^{MonPol} \end{bmatrix}$$

How to compute historical decompositions (cont'd)

- Then x_2 can be expressed as

$$\begin{cases} y_2 = init_y + const_y + \theta^1_{11}\varepsilon_1^{Demand} + \theta^1_{12}\varepsilon_1^{MonPol} + \theta^0_{11}\varepsilon_2^{Demand} + \theta^0_{12}\varepsilon_2^{MonPol} \\ r_2 = init_r + const_r + \theta^1_{21}\varepsilon_1^{Demand} + \theta^1_{22}\varepsilon_1^{MonPol} + \theta^0_{21}\varepsilon_2^{Demand} + \theta^0_{22}\varepsilon_2^{MonPol} \end{cases}$$

- The historical decomposition is given by

$$\underbrace{\begin{cases} HD_{y_2}^{\varepsilon^{Demand}} = \theta^1_{11}\varepsilon_1^{Demand} + \theta^0_{11}\varepsilon_2^{Demand} \\ HD_{y_2}^{\varepsilon^{MonPol}} = \theta^1_{12}\varepsilon_1^{MonPol} + \theta^0_{12}\varepsilon_2^{MonPol} \\ HD_{y_2}^{init} = init_y \\ HD_{y_2}^{const} = const_y \end{cases}}_{\text{This sums up to } y_2} \qquad \underbrace{\begin{cases} HD_{r_2}^{\varepsilon^{Demand}} = \theta^1_{21}\varepsilon_1^{Demand} + \theta^0_{21}\varepsilon_2^{Demand} \\ HD_{r_2}^{\varepsilon^{MonPol}} = \theta^1_{22}\varepsilon_1^{MonPol} + \theta^0_{22}\varepsilon_2^{MonPol} \\ HD_{r_2}^{init} = init_r \\ HD_{r_2}^{const} = const_r \end{cases}}_{\text{This sums up to } r_2}$$

Historical decomposition: In Matlab

- HD is computed and stored automatically by `VARmodel`; plot with `VARhdplot`

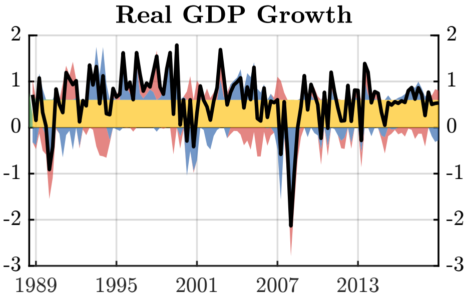
```
VARopt.figname = 'graphics/sign_HD';
VARhdplot(VAR_sr.HDmed, VARopt);
```

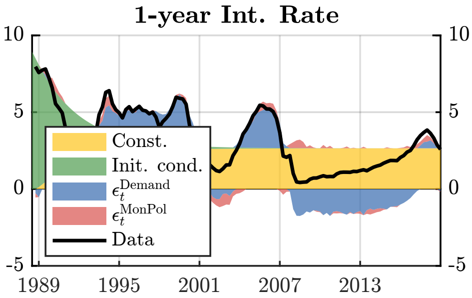
```
VARopt.figname = 'graphics/sign_HD_bands';
VARopt.pick    = 2;
VARopt.color   = pantone('Tomato');
VARhdplot(VAR_sr.HDmed, VARopt, VAR_sr.HDinf, VAR_sr.HDsup);
VARopt.pick    = 0;
VARopt.color   = [];
```

- `VAR_sr.HD` is a struct with fields `shock`, `init`, `const`, `trend`, `endo`

Historical decomposition: Sign-restricted bivariate VAR

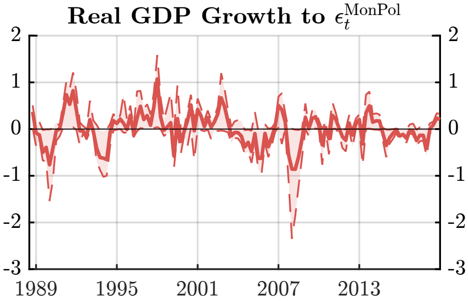
- Stacked-area: contribution of each shock to the deviation of each variable from its unconditional mean

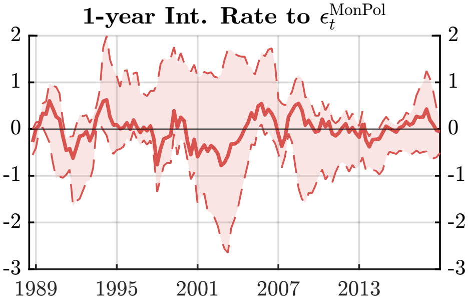




Historical decomposition: Uncertainty bands

- MP shock contribution with Bayesian credible bands; width reflects identification uncertainty across accepted draws





[BOX] HD: `init` and `const` in log-levels vs. log-differences

- The economic content of the `init` and `const` components depends on the data transformation
- **Log-differences VAR** The constant converges to the average growth rate (meaningful, stable across samples); the initial condition is the deviation from the long-run average at the sample start, and decays toward zero as the VAR is stable
- **Log-levels VAR** Neither has a stand-alone interpretation: the constant converges to the sample mean of log GDP (grows with the trend, sample-dependent); the initial condition is the first-period log-level, which for a growing economy lies below the mean

Local Projections

Local projections: Background

- **Intuition** Estimate the IRF at each horizon directly
  - No fitted VAR to iterate forward
  - Regress the h-step-ahead outcome on the shock and lagged controls
- **Reference** Jorda (2005)
- One regression per horizon $h = 0, 1, \dots, H$:

$$x_{i,t+h} = \alpha_h + \beta_h s_t + \gamma_h' w_t + e_{t+h}$$

  - $\{\hat{\beta}_h\}_{h=0}^{H}$ traces the response of $x_{i,t+h}$ to s_t
  - Std. errors: Newey-West HAC, bandwidth $L = h - 1$

LP vs. VAR: Trade-offs

- **Robustness** Robust to VAR misspecification
  - Each horizon estimated independently, no cross-horizon restrictions
- **Efficiency** VAR IRFs more efficient when the VAR is correct
  - LP fits a separate regression at each horizon
- **Equivalence** Same population IRFs as the lag order grows Plagborg-Møller and Wolf (2021)
- **Complements**, not substitutes
  - Divergence at long horizons signals VAR misspecification

LP-OLS: Estimation in the VAR Toolbox

- Pass the full `ENDO` matrix, `LPmodel` loops over its columns (same shock, controls, lags)
- Controls always enter `CTRL` contemporaneously in `LPmodel`
  - `LPmodel` lags them internally by the pre-specified `nlags`, contemporaneous values never used
- `mps` (shock s_t) and `detc` (deterministic component) carry over from the earlier examples

```
% Set up LP-OLS options
LPopt          = LPoption;
LPopt.nsteps   = VARopt.nsteps; % match VAR horizon
LPopt.pctg     = 90;            % 90% confidence bands
LPopt.impact   = 0;             % 1-std-dev shock
LPopt.mnem     = Xmnem;         % mnemonics names
LPopt.vnames   = Xvnames;       % labels for plots
% Run LP-OLS for all endogenous variables in a single matrix call
ENDO = X;   % outcomes x_{i,t+h}
CTRL = X;   % controls w_t (lags constructed internally)
LP = LPmodel(ENDO, mps, CTRL, nlags, detc, LPopt);
```

LP-IV: Instrumented treatment

- When the treatment is endogenous, instrument it via `LPopt.IV`;
  - Treatment `ENDO(:,2)` enters as the second argument in `LPmodel`
  - `LPopt.IV` holds the external instrument
- First-stage F-statistic per horizon in `LP_IV.dlngdp.hN.Fstat_fs`; check > 10 (Stock-Yogo threshold)

```
%  Set up and run LP-IV
LPopt_IV          = LPoption;
LPopt_IV.nsteps   = VARopt.nsteps;
LPopt_IV.pctg     = 90;
LPopt_IV.impact   = 0;
LPopt_IV.IV       = mps;    % external instrument
LPopt_IV.mnem     = Xmnem; % name per-variable sub-structs
LP_IV = LPmodel(ENDO, ENDO(:,2), CTRL, nlags, detc, LPopt_IV);
```

The `LP` struct

- `LPmodel` returns a single struct `LP`
  - Per-variable sub-structs `LP.dlngdp`, `LP.i1yr` (full univariate output)
  - Top-level `IR`, `INF`, `SUP` ($H \times N$): IRFs + Newey-West bands
  - Input echo: `ENDO`, `TREAT`, `CTRL`, `EXOG`, `nlag`, `const`
  - Per-horizon `h1`, …, `hH`; joint-test fields `jtest_chi2/df/pval`

[BOX] Newey-West standard errors in LP

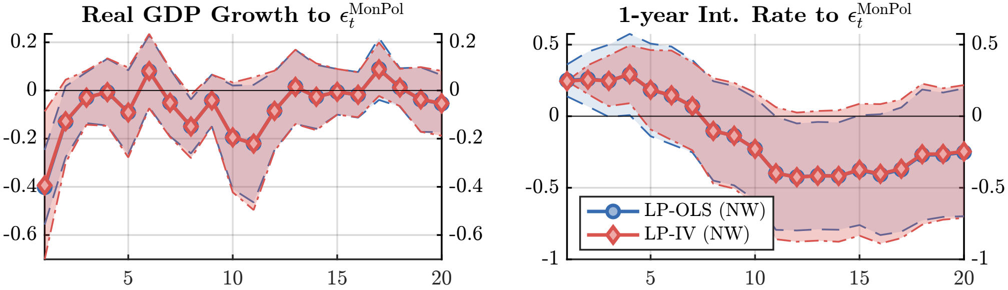
- LP error e_{t+h} = the h-step-ahead forecast error
- Under a correct VAR(p), it is an MA($h-1$):

$$e_{t+h} = \varepsilon_{t+h} + \psi_1 \varepsilon_{t+h-1} + \cdots + \psi_{h-1} \varepsilon_{t+1}$$

  - Serial correlation up to order $h-1$, even with white-noise ε_t
- **Fix** Newey-West HAC, bandwidth $L = h - 1$
  - Bands `INF`, `SUP` use `bstd_NW`

LP-OLS vs. LP-IV: Bivariate example

- Both normalized to a 25 bps rise in the 1-year rate on impact
  - **LP-OLS** reduced-form effect of an `mps` surprise
  - **LP-IV** causal effect of the rate change instrumented by `mps`

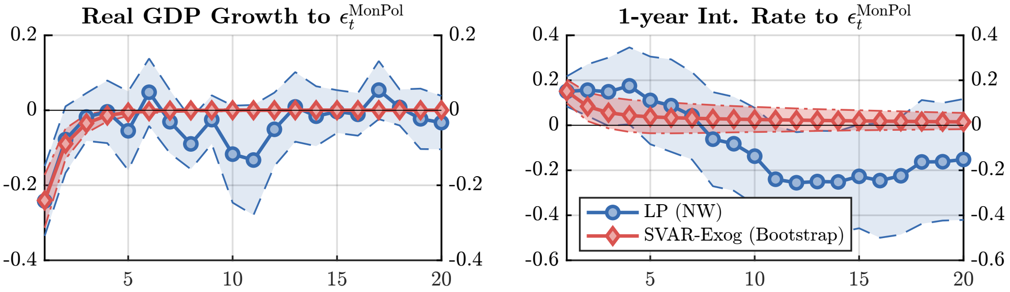


LP vs. Exogenous-variable SVAR: Two routes to the same object

- Recall that LP and the exogenous-variable SVAR are **the same object** on impact
- **Exog-variable SVAR** s_t enters each equation contemporaneously
  - Impact identified with the VAR dynamics, then iterated forward
- **Local projections** project the outcome on s_t at each horizon
  - No dynamic model fitted
- Asymptotically equivalent as the lag order grows Plagborg-Møller and Wolf (2021)
  - Divergence at $h > 0$ signals VAR misspecification

LP vs. Exogenous-variable SVAR: Bivariate example

- Same external instrument, two routes
  - **LP** each horizon estimated directly
  - **Exog-SVAR** IV-identified impact iterated through the VAR dynamics



Appendix

References I

Antolín-Díaz, J. and J. F. Rubio-Ramírez (2018): "Narrative Sign Restrictions for SVARs," *American Economic Review*, 108, 2802–2829.

Blanchard, O. J. and D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79, 655–73.

Canova, F. and G. D. Nicolo (2002): "Monetary disturbances matter for business fluctuations in the G-7," *Journal of Monetary Economics*, 49, 1131–1159.

Cesa-Bianchi, A. and A. Sokol (2021): "Financial shocks, credit spreads, and the international credit channel," *Journal of International Economics*, 1035–43.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999): "Monetary policy shocks: What have we learned and to what end?" in *Handbook of Macroeconomics*, ed. by J. B. Taylor and M. Woodford, Elsevier, vol. 1 of *Handbook of Macroeconomics*, chap. 2, 65–148.

Faust, J. (1998): "The robustness of identified VAR conclusions about money," International Finance Discussion Papers 610, Board of Governors of the Federal Reserve System (U.S.).

Gali, J. (1999): "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review*, 89, 249–271.

Gertler, M. and P. Karadi (2015): "Monetary Policy Surprises, Credit Costs, and Economic Activity," *American Economic Journal: Macroeconomics*, 7, 44–76.

Jorda, O. (2005): "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 95, 161–182.

Mertens, K. and M. O. Ravn (2013): "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States," *American Economic Review*, 103, 1212–47.

Plagborg-Møller, M. and C. K. Wolf (2021): "Local Projections and VARs Estimate the Same Impulse Responses," *Econometrica*, 89, 955–980.

Sims, C. A. (1980): "Macroeconomics and Reality," *Econometrica*, 48, 1–48.

Sims, C. A., J. H. Stock, and M. W. Watson (1990): "Inference in Linear Time Series Models with Some Unit Roots," *Econometrica*, 58, 113–44.

References II

Stock, J. and M. Watson (2012): "Disentangling the Channels of the 2007-2009 Recession," *Brookings Papers on Economic Activity*, Spring, 81–135.

Stock, J. H. and M. W. Watson (2001): "Vector Autoregressions," *Journal of Economic Perspectives*, 15, 101–115.

Stock, J. H. and M. Yogo (2002): "Testing for Weak Instruments in Linear IV Regression," NBER Technical Working Papers 0284, National Bureau of Economic Research, Inc.

Uhlig, H. (2005): "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics*, 52, 381–419.