A Primer on Vector Autoregressions

Ambrogio Cesa-Bianchi

[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 describe the job of macroeconometricians as consisting of the following
  - Describe and summarize macroeconomic time series
  - Make forecasts
  - Recover the structure of the macroeconomy from the data Main focus of these notes
  - Advise macroeconomic policy-makers
- Vector autoregressive models (VARs) are a statistical tool to perform these tasks

What can we do with VARs?

- Consider a bivariate VAR with the following variables: real GDP growth (y_t) and the policy rate (r_t)
- A VAR can help us answering the following questions
  - [1] What is the dynamic behavior of these variables? How do these variables interact?
  - [2] What is the most likely behavior 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?

VAR Basics

What is a Vector Autoregression (VAR)?

- Consider a (2×1) vector of zero-mean time series x_t, composed of t observations and an initial condition x_0

$$x_t = \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1t} \\ x_{21} & x_{22} & \dots & x_{2t} \end{bmatrix} \quad \text{and} \quad x_0 = \begin{bmatrix} x_{10} \\ x_{20} \end{bmatrix}$$

- Assume that the two time series in x_t are covariance stationary, which means (for $i = 1, 2$)
  * Constant mean $\mathbb{E}[x_{it}] = \mu_i$
  * Constant variance $\mathbb{V}[x_{it}] = \sigma_i$
  * Constant autocovariance $COV[x_{it}, x_{it+\tau}] = \gamma_i(\tau)$

- A **structural VAR** of order 1 is given by

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

  where
  * Φ and B are (2×2) matrices of coefficients
  * ε_t is an (2×1) vector of unobservable zero-mean white noise processes

Three different ways of writing the same thing

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

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

- For example, we can write it in matrix form

$$\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

- Or as a system of linear equations

$$\begin{cases} x_{1t} = \phi_{11} x_{1,t-1} + \phi_{12} x_{1,t-1} + b_{11}\varepsilon_{1t} + b_{12}\varepsilon_{2t} \\ x_{2t} = \phi_{21} x_{2,t-1} + \phi_{22} x_{2,t-1} + b_{21}\varepsilon_{1t} + b_{22}\varepsilon_{2t} \end{cases}$$

The structural shocks

- We defined ε_t as a *vector of unobservable zero mean white noise processes*. **What does it mean?**

- The elements of ε_t are serially uncorrelated and independent of each other

- In other words we assumed

$$\varepsilon_t = (\varepsilon'_{1t}, \varepsilon'_{2t})' \sim \mathcal{N}(0, I_2)$$

where

$$\mathbb{V}(\varepsilon_t) = \sigma_\varepsilon = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad CORR(\varepsilon_t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Why is it called 'structural' VAR?

- Go back to our bivariate structural VAR(1)

$$\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

- The structural VAR can be thought of as a description of the true structure of the economy
  - E.g.: an approximation of the structure of a DSGE model
- The structural shocks are shocks with a well-defined economic interpretation
  - E.g.: TFP shocks or monetary policy shocks
  - As $\varepsilon_t \sim \mathcal{N}(0, I_2)$ 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). To make a concrete example, assume that
  - x_{1t} and x_{2t} are output growth (y_t) and the policy rate (r_t), both demeaned
  - ε_{1t} and ε_{2t} are a demand shock (ε_t^{Demand}) and a monetary policy shock (ε_t^{MonPol})
  - B is known (we'll get back to this in a second)

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

- 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. How can we estimate B?

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

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

$$u_t = B\varepsilon_t \Rightarrow \begin{bmatrix} u_{1t} \\ u_{2t} = \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix} \Rightarrow \begin{cases} u_{yt} = b_{11}\varepsilon_t^{Demand} + b_{12}\varepsilon_t^{MonPol} \\ u_{rt} = b_{21}\varepsilon_t^{Demand} + b_{22}\varepsilon_t^{MonPol} \end{cases}$$

- Why is this useful? The VAR becomes

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

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

The reduced-form VAR

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

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

- In matrix form

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

- Or as a system of linear equations

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

The reduced-form covariance matrix

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

$$\sigma_u = \begin{bmatrix} \sigma_y^2 & \sigma_{yr}^2 \\ \sigma_{yr}^2 & \sigma_r^2 \end{bmatrix}$$

- Differently from the structural shocks (which are orthogonal), the reduced-form residuals are correlated among each other

- This is because the elements of u_t inherit all the contemporaneous relations among the endogenous variables x_t
  - To see that, remember how the reduced form residuals are defined

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

The reduced-form covariance matrix

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

$$\sigma_u = \begin{bmatrix} \sigma_y^2 & \sigma_{yr}^2 \\ \sigma_{yr}^2 & \sigma_r^2 \end{bmatrix}$$

- Differently from the structural shocks (which are orthogonal), the reduced-form residuals are correlated among each other

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

- To make causal statements (e.g. the effects on y_t of a shock to ε_t^{MonPol}) we need to find a way to recover B

- This is the essence of **identification** in VARs

The Wold representation

- Before turning to identification, let's introduce another representation of the VAR that will be useful later
- Start from the structural VAR representation

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

- The **Wold representation** can be obtained by substituting recursively the elements on the right hand side of the equal sign

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

The Wold representation (cont'd)

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

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

  - The sum of current and past structural shocks
  - An initial condition

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

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

- But: we assumed that x_t is covariance stationary. How do these infinite sums relate to that assumption?
  - Aren't the increasing powers of Φ exploding?

Stability of the VAR

- A VAR is stable if the effect of shocks progressively dissipate over time. For that to happen we need Φ^j to converge to zero

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

- Why does this matter? If shocks have permanent effects
  - The mean and the variance of x_t will depend on the history of shocks
  - Violates covariance stationary assumption $\rightarrow$ VAR displays unstable dynamics

- **Definition** 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 (i.e. its unconditional mean)

The unconditional mean of the VAR

- First note that if the eigenvalues of Φ are less than 1 in modulus we have

$$\Phi^{\infty} = 0 \quad \text{and} \quad \sum_{j=0}^{\infty} \Phi^{j} = \boxed{(I_2 - \Phi)^{-1}}$$

Geometric series

- Because of white noise assumption of the ε_t, the unconditional mean is simply given by

$$\mathbb{E}[x_t] = \Phi^{\infty} x_{t-\infty} + \sum_{j=0}^{\infty} \Phi^{j} B \mathbb{E}\left[\varepsilon_{t-j}\right] = 0$$

- Note that if the VAR had a constant (α) an additional term would show up in the Wold representation

$$\mathbb{E}[x_t] = \Phi^{\infty} x_{t-\infty} + \sum_{j=0}^{\infty} \Phi^{j} \alpha + \sum_{j=0}^{\infty} \Phi^{j} B \mathbb{E}\left[\varepsilon_{t-j}\right] = 0$$

- The unconditional mean in this case would be

$$\mathbb{E}[x_t] = (I_2 - \Phi)^{-1} \alpha$$

The general form of the stationary structural VAR(p) model

- The basic bivariate VAR(1) model used so far may be too parsimonious to sufficiently summarize the dynamic relations of the data

- Model can be enriched with along the following dimensions
  - Increase the number of endogenous variables (k)
  - Increase the number of lags (p)
  - Add deterministic terms (e.g. time trend or seasonal dummy variables)
  - Add exogenous variables (e.g. price of oil from the point of view of a small country)

- The general form of the VAR(p) model with deterministic terms (Z_t) and exogenous variables (W_t) is given by

$$x_t = \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \ldots + \Phi_p x_{t-p} + \Lambda Z_t + \Psi W_t + B\varepsilon_t$$

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 $\hat{u}_t$:

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

- **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, assume that the 'true' (and unobserved) model of the economy is given by

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

- It is obvious that the reduced form innovations are a linear combination of the two 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!

Reduced-form VARs do not tell us anything about causality

- To see that, assume that the 'true' (and unobserved) model of the economy is given by

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

- We have already seen that the reduced form innovations are a linear combination of the two structural shocks

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

- So, an increase in u_{rt} could be due to
  - [1] A positive demand shock that increases both output growth and the policy rate ($b_{21} > 0$)
  - [2] Or a monetary policy shock that decreases output growth and increases the policy rate ($b_{22} > 0$)

- How to know whether 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

$$u_t \Rightarrow B\varepsilon_t$$

- To do that, we can exploit the relation between reduced form and structural innovations ($u_t = B\varepsilon_t$) to write

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

- 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 combinations of B that give you the same σ_u

The identification problem (cont'd)

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

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

- Can be rewritten as

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

- **Problem** Because of the symmetry of the σ_u matrix, the second and the third equation are identical

- We are left with 4 unknowns (the elements of B) but only 3 equations!

Identification Schemes

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**
  - 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 (recursive) contemporaneous restrictions
- Zero (recursive) long-run restrictions
- Sign restrictions
- External instruments
- Combining sign restrictions and external instruments
- Other (narrative sign restrictions, maximization of forecast error variance,...)

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, Eichenbaum, Evans (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

- We have seen above how to assume that monetary policy has no contemporaneous effects on output growth

By assumption

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

- **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:

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

Zero contemporaneous restrictions

Impact effects

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

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

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

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

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

$$\begin{cases} y_t = \sigma_y^2 \\ r_t = \sigma_{yr}/\sigma_y^2 \end{cases}$$

Zero contemporaneous restrictions

Aka Cholesky identification

- 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

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

Cholesky decomposition of a matrix [Back to basics]

- 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 a lower triangular matrix (and therefore P' is upper triangular)

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

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

Zero contemporaneous restrictions

Aka Cholesky identification

- 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 $P = B$

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 is neutral in the long-run and has no cumulative effect on the level of output
- 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?

- Re-write the VAR as

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

- If a shock ε_t hits in t, its cumulative (long run) impact on x_t would be

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

- 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 $C \equiv (I - \Phi)^{-1}$ is the cumulative effect that ε_t has on output growth from time t to ∞, i.e. the effect on output level

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

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

- Take the first equation: $y_{t,t+\infty} = c_{11}\varepsilon_t^{Demand} + c_{12}\varepsilon_t^{MonPol}$
  - The coefficient c_{12} represents the impact of a monetary policy shock (hitting in t) on the **level of GDP** in the long-run
  - If you believe in the long-run neutrality of monetary policy 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 (so, it admits a unique Cholesky decomposition)
  $$\Omega = PP'$$
  3. Because of our assumption that C is lower triangular, it follows that $P = C$
- We achieved identification: $B = (I - \Phi) P$

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

$$B = (I_2 - \Phi)P = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{2,2} \end{bmatrix} \right) \begin{bmatrix} \sigma_y^2 & 0 \\ \sigma_{yr}/\sigma_y^2 & \sqrt{\sigma_r^2 - \frac{\sigma_{yr}^2}{\sigma_y^2}} \end{bmatrix}$$

- Note that the B matrix **is not triangular** (differently from what we had in the zero contemporaneous restrictions identification)

- The impact effects are left unrestricted, the restrictions is on the C matrix
  * We'll check later that the restrictions is satisfied in a simple example with true data

Common Identification Schemes

Sign restrictions

Sign restrictions

- **Intuition** Exploits prior beliefs (typically informed by theoretical models) about the sign that certain shocks should have on certain endogenous variables

- **Intuition** Faust (1998), Canova and De 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 for 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
  $$\sigma_u = PP'$$
  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

Orthonormal matrix [Back to basics]

- An orthonormal matrix Q is a real square matrix whose columns and rows are orthogonal unit vectors

- 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 orthonormal 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 Q 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 $i = 1, 2, ..., N$) represent admissible solutions to the identification problem
- In this sense, sign restricted VARs are only **set identified**

Common Identification Schemes

External Instruments (or Proxy SVARs)

External instruments

- **Intuition** Exploits 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 policy to some development in the literature)
- 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:

$$\begin{aligned} \mathbb{E}\left[\varepsilon_t^{Demand} z_t'\right] &= 0, \\ \mathbb{E}[\varepsilon_t^{MonPol} z_t'] &= c, \end{aligned}$$

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

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

External instruments

- **How does it work?** Recall that the reduced form residuals are a linear combination of the two structural shocks

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

- First stage regression

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

- The OLS estimate of β identifies b_{22} up to a scaling factor
- The OLS estimate of ξ_t collects everything else that is uncorrelated with ε_t^{MonPol}

External instruments

- How to get the remaining impact coefficient b_{12}?
- First stage regression

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

- Construct the fitted values

$$\hat{u}_{rt} = \hat{\beta} z_t$$

- Second stage regression to get a consistent estimate of the ratio b_{12}/b_{11}:

$$u_{yt} = \underbrace{\gamma}_{b_{12}/b_{22}} \hat{u}_{rt} + \zeta_t,$$

- If we normalize the effect of ε_t^{MonPol} on r_t to 1 (that is, we fix $b_{22} = 1$) we can easily recover b_{21} from the OLS estimates of γ

Structural Dynamic Analysis

Structural Dynamic Analysis

- We have seen how to ‘identify’ structural shocks by imposing some restrictions on the data

- But what can we do with that?
  - 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**

- We'll turn to this structural dynamic analysis next

Structural 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 to single out the effect of a shock (e.g. its impact an persistence) 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

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

- Define a 2×1 vector of impulse selection (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

$$\begin{cases} IR_t = Bs, & \text{for } t = 0, \\ IR_t = \Phi \cdot IR_{t-1} & \text{for } t = 2, ..., h. \end{cases}$$

- 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

$$\begin{bmatrix} IR_h^y \\ IR_h^r \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} IR_h^y \\ IR_h^r \end{bmatrix}$$

The companion matrix [Back to basics]

- So far, we considered simple VAR(1) specifications. But what to do if the VAR has $p > 1$?
- Every VAR(p) can be written as a VAR(1) using the **companion representation**
  - For example, take a VAR(2)

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

  - Re-write the VAR(2) as

Companion matrix

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

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

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

Structural 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

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

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

Basic properties of the variance [Back to basics]

- If X is a random variable x and a is a constant
  - $\mathbb{V}(x+a) = \mathbb{V}(x)$
  - $\mathbb{V}(ax) = a^2\mathbb{V}(x)$
- If Y is a random variable and b is a constant
  - $\mathbb{V}(aX+bY) = a^2\mathbb{V}(x) + b^2\mathbb{V}(Y) + 2abCOV(X,Y)$
- 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

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

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

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

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

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

Structural 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 the endogenous variables away from their equilibrium?

- *HD* allow 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 GDP growth in 1973?

How to compute historical decompositions

- As an example, let's compute the *HD* of of the endogenous variables when $t = 2$ in our simple bivariate VAR
- Historical decompositions can be easily understood from the Wold representation of the VAR

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

- Using the Wold representation, we can write x_2 as a function of present and past structural shocks (ε^{Demand} and ε^{MonPol}) plus the initial condition (x_0)

$$x_2 = \underbrace{\Phi^2 x_0}_{init} + \underbrace{\Phi B}_{\boldsymbol{\Theta_1}} \varepsilon_1 + \underbrace{B}_{\boldsymbol{\Theta}_0} \varepsilon_2$$

- Re-write x_2 in matrix form

$$\begin{bmatrix} y_2 \\ r_2 \end{bmatrix} = \begin{bmatrix} init_y \\ init_r \end{bmatrix} + \begin{bmatrix} \theta^1_{11} & \theta^1_{12} \\ \theta^1_{21} & \theta^1_{22} \end{bmatrix} \begin{bmatrix} \varepsilon^{Demand}_2 \\ \varepsilon^{MonPol}_2 \end{bmatrix} + \begin{bmatrix} \theta^0_{11} & \theta^0_{12} \\ \theta^0_{21} & \theta^0_{22} \end{bmatrix} \begin{bmatrix} \varepsilon^{Demand}_2 \\ \varepsilon^{MonPol}_2 \end{bmatrix}$$

How to compute historical decompositions (cont'd)

- Then x_2 can be expressed as

$$\begin{cases} y_2 = init_y + \theta^1_{11}\varepsilon^{Demand}_1 + \theta^1_{12}\varepsilon^{MonPol}_1 + \theta^0_{11}\varepsilon^{Demand}_2 + \theta^0_{12}\varepsilon^{MonPol}_2 \\ r_2 = init_r + \theta^1_{21}\varepsilon^{Demand}_1 + \theta^1_{22}\varepsilon^{MonPol}_1 + \theta^0_{21}\varepsilon^{Demand}_2 + \theta^0_{22}\varepsilon^{MonPol}_2 \end{cases}$$

- The historical decomposition is given by

$$\underbrace{\begin{cases} HD^{\varepsilon^{Demand}}_{y_2} = \theta^1_{11}\varepsilon^{Demand}_1 + \theta^2_{11}\varepsilon^{Demand}_2 \\ HD^{\varepsilon^{MonPol}}_{y_2} = \theta^1_{12}\varepsilon^{MonPol}_1 + \theta^2_{12}\varepsilon^{MonPol}_2 \\ HD^{init}_{y_2} = init_y \end{cases}}_{\text{This sums up to } y_2} \quad \underbrace{\begin{cases} HD^{\varepsilon^{Demand}}_{r_2} = \theta^1_{21}\varepsilon^{Demand}_1 + \theta^0_{21}\varepsilon^{Demand}_2 \\ HD^{\varepsilon^{MonPol}}_{r_2} = \theta^1_{22}\varepsilon^{MonPol}_1 + \theta^0_{22}\varepsilon^{MonPol}_2 \\ HD^{init}_{r_2} = init_r \end{cases}}_{\text{This sums up to } r_2}$$

Practical Examples

The VAR Toolbox

- We'll see in practice how VARs work through a set of examples using the **VAR Toolbox 3.0**
- The VAR Toolbox is a collection of Matlab routines to perform VAR analysis
  - Codes are available at `https://github.com/ambropo/VAR-Toolbox`
  - No installation is required. Simply clone the folder from Github and add the folder (with subfolders) to your Matlab path
  - To save figures in high quality format, you need to download an install Ghostscript (freely available at `www.ghostscript.com`).
    - The first time you'll be saving a figure using the Toolbox, you'll be asked to locate the Ghostscript on your local drive
- We'll start with a very simple example and then replicate the results from a few well-known papers

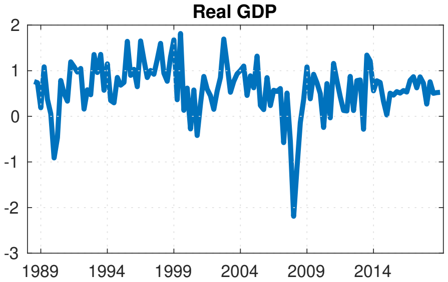
Adding the VAR Toolbox path to Matlab

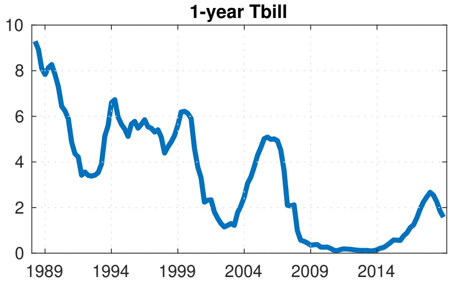
- To avoid clashes with functions from other toolboxes, it is recommendable to add and remove the Toolbox at beginning and end of your scripts
- If you download the toolbox to `/User/VAR-Toolbox/`, you can simply add the following lines at the beginning and end of your script

```
addpath(genpath('/User/VAR-Toolbox/v3dot0/'))
...
rmpath(genpath('/User/VAR-Toolbox/v3dot0'))
```

A simple bivariate VAR model

- Our first example will be a simple bivariate VAR as the one considered above
- US quarterly data from 1989q1to 2019q4 on output growth (y_t) and the 1-year T-bill (r_t)





A simple bivariate VAR model

- As both GDP growth an the 1-year rate are non-zero means, we fit the data with a VAR(1) with a constant

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} \alpha_y \\ \alpha_\pi \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ \pi_{t-1} \end{bmatrix} + \begin{bmatrix} u_{yt} \\ u_t^\pi \end{bmatrix}$$

- This means we will estimate the following parameters
  * $2+4$ coefficients, namely the elements of α and Φ
  * 2 variances of the reduced-form residuals, namely σ_y^2 and σ_π^2
  * 1 covariance of the reduced-form residuals, namely $\sigma_{y\pi}$

- We will store these coefficients in two Matlab matrices

$$\texttt{F} = \begin{bmatrix} \alpha_1 & \phi_{11} & \phi_{12} \\ \alpha_2 & \phi_{21} & \phi_{22} \end{bmatrix} \qquad \texttt{sigma} = \begin{bmatrix} \sigma_y^2 & \sigma_{yr}^2 \\ \sigma_{yr}^2 & \sigma_r^2 \end{bmatrix}$$

A simple bivariate VAR model

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

$$\texttt{X} = \begin{bmatrix} y_1 & r_1 \\ y_2 & r_2 \\ \dots & \dots \\ y_T & r_T \end{bmatrix} = (y_t', r_t') = x_t'$$

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

```
% Set the deterministic variable in the VAR (1=constant, 2=trend)
det = 1;
% Set number of nlags
nlags = 12;
% Estimate VAR by OLS
[VAR, VARopt] = VARmodel(ENDO,nlags,det);
```

A simple bivariate VAR model: VAR output

- The code estimates the VAR equation by equation with OLS
- Results are stored in the `VAR` and `VARopt` structures
- The six estimated parameters (i.e. α and Φ) can be printed at screen by simply typing `disp(VAR.Ft)` to get

```
>> disp(VAR.F)
    0.3630    0.3788    0.0041
   -0.0729    0.2607    0.9541
```

- For the three elements of σ_u type `disp(VAR.sigma)` to get

```
>> disp(VAR.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)
Real GDP (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 inflation are contemporaneously positively correlated
  - It means that, on average, when GDP growth increases inflation increases, too
- Does it mean that a shock to output always increase inflation?
  - No! Recall that reduced from 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

Stability and equilibrium

- We've already seen that a VAR is stable when $|eig(\Phi)| < 1$
  - If this condition is not met, the infinite sums in the Wold representation do not converge

- You can check the maximum value of Φ's eigenvalues in the `VAR` structure, by typing `disp(VAR.maxEig)` to get

```
>> disp(VAR.maxEig)
    0.9559
```

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

- In practice:

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

Stability and equilibrium (cont'd)

- As our VAR is stable, its Wold representation will converge to the (finite) unconditional mean of the data

- To see that, first consider the Wold representation in the presence of a constant

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

- For t large enough and taking expectations we get

$$\mathbb{E}\left[x_t\right] = \sum_{j=0}^{t-1} \Phi^t \alpha = \left(I_2 - \Phi\right)^{-1} \alpha$$

- In absence of shocks, the VAR's variable will converge to its equilibrium $\left(I_2 - \Phi\right)^{-1} \alpha$ at a rate that depends on Φ

Examples of different identification schemes

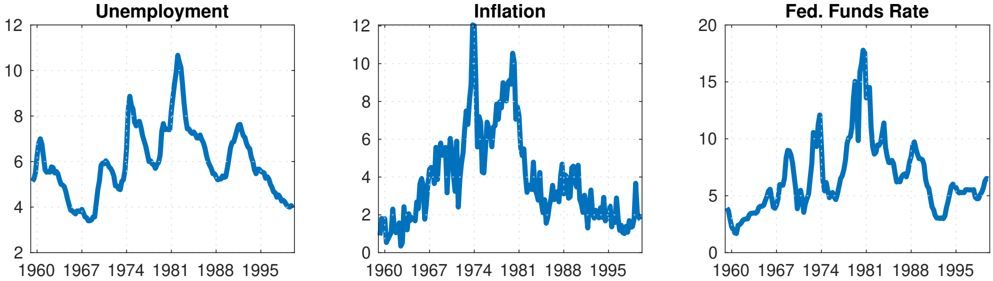
- Zero short-run restrictions
  - Stock and Watson (2001). "Vector Autoregressions," *Journal of Economic Perspectives*
- Zero long-run restrictions
  - Blanchard and Quah (1989). "The Dynamic Effects of Aggregate Demand and Supply Disturbances", *American Economic Review*
- Sign Restrictions
  - Uhlig (2005). "What are the effects of monetary policy on output? Results from an agnostic identification procedure," *Journal of Monetary Economics*
- External instruments
  - Gertler and Karadi (2015). "What are the effects of monetary policy on output? Results from an agnostic identification procedure," *American Economic Journal: Macroeconomics*
- External instruments & Sign restrictions
  - Cesa-Bianchi and Sokol (2020). "Financial Shocks, Credit Spreads, and the International Credit Channel," *Unpublished manuscript*

Practical Examples

Stock and Watson (2001, JEP)

Stock and Watson (2001): Zero short-run restrictions

- Stock and Watson (2001). “Vector Autoregressions,” *Journal of Economic Perspectives*
- US quarterly data from 1960QI to 2000Q4



Monetary policy shocks, inflation and unemployment

- **Objective** Infer the causal influence of monetary policy on unemployment and inflation
- Assume a VAR with $p = 4$ with inflation (π_t), unemployment (u_t), and the fed funds rate (r_t)
- **Key identifying assumptions**
  * MP (r_t) reacts contemporaneously to movements in inflation and in unemployment
  * MP shocks (ε_t^{MonPol}) do not affect inflation and unemployment within the quarter of the shock

$$\begin{bmatrix} \pi_t \\ u_t \\ r_t \end{bmatrix} = \sum_{p=1}^{4} \Phi_p x_{t-p} + \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \\ \varepsilon_t^{MonPol} \end{bmatrix}$$

Replicating Stock and Watson (2001) with the VAR Toolbox

- In Matlab, set lag length to 4 and estimate a VAR with a constant

```
% Set up and estimate VAR
det = 1;
nlags = 4;
[VAR, VARopt] = VARmodel(X,nlags,det);
```

- Then set the option for recursive identification `VARopt.ident ='short'` and compute the *IR* with the `VARir` function. Note that the **ordering of the variables matter!**

```
% For zero contemporaneous restrictions set:
VARopt.ident = 'short';
% Compute IR
[IR, VAR] = VARir(VAR,VARopt);
```

- Finally note that the second output of the `VARir` function is `VAR` again
  - This is because the `VAR` structure is updated with the *B* matrix corresponding to the identification scheme chosen

Replicating Stock and Watson (2001) with the VAR Toolbox (cont'd)

- The `VARirband` function allows to compute confidence intervals

```
% Compute IR
[IR, VAR] = VARir(VAR,VARopt);
```

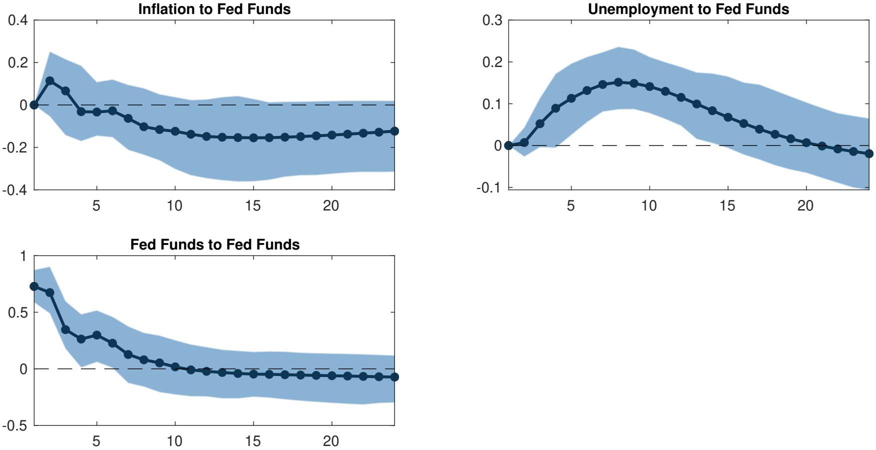
- You can control the options of the bootstrapping procedure by modifying the `VARopt` structure (before running `VARir`)

- For example

```
% Some options for the bootstrap
VARopt.ndraws = 1000; % Number of draws
VARopt.pctg = 95; % Level for confidence intervals
VARopt.method = 'bs'; % 'bs' sampling with replacement; 'wild' wild bootstrap
```

The effect of a monetary policy shock

- Monetary policy shock raises inflation in the short run (price puzzle) and increases unemployment

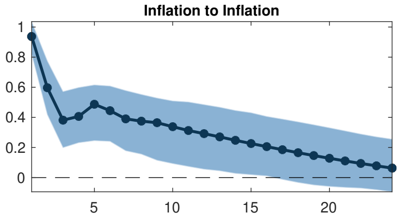


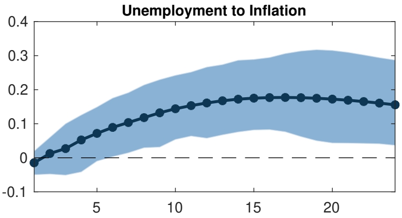
The other two shocks are identified by definition... but how can we interpret them?

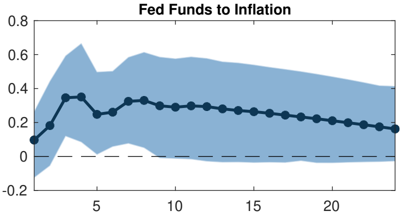
- How about ε_t^1 and ε_t^2?
  - The shock ε_t^1 affects all variables contemporaneously
  - The shock ε_t^2 affects r_t contemporaneously but not π_t
- Can we interpret these shocks? Are the assumptions consistent with any theoretical mechanism?
- Some shocks may be better identified than others

The other two shocks are identified by definition... but how can we interpret them?

- Shock to ε_t^1 behaves as a negative aggregate supply shock

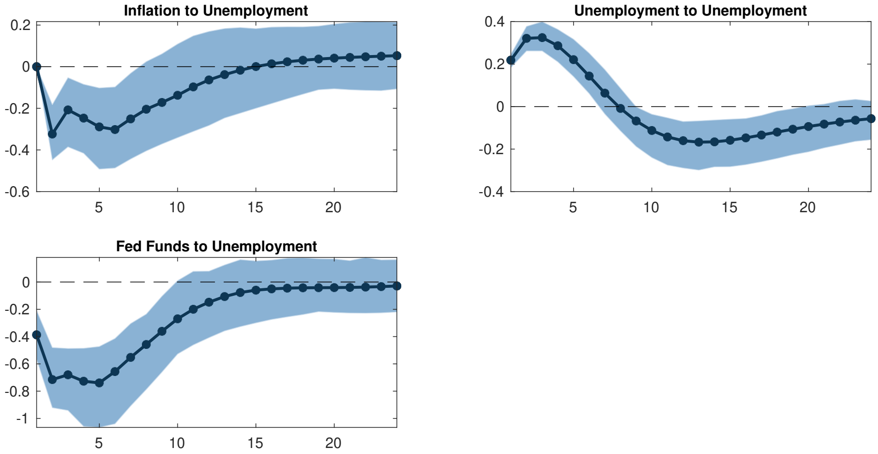






The other two shocks are identified by definition... but how can we interpret them?

- Shock to ε_t^2 behaves as a negative aggregate demand shock



Forecast error variance & Historical decompositions

- In Matlab, set compute the *VD* with the `VARvd` function
  - The matrix `VD` is a simple *H* horizon, *k* shocks, *k* variables matrix

```
% Compute VD
[VD, VAR] = VARvd(VAR,VARopt);
```

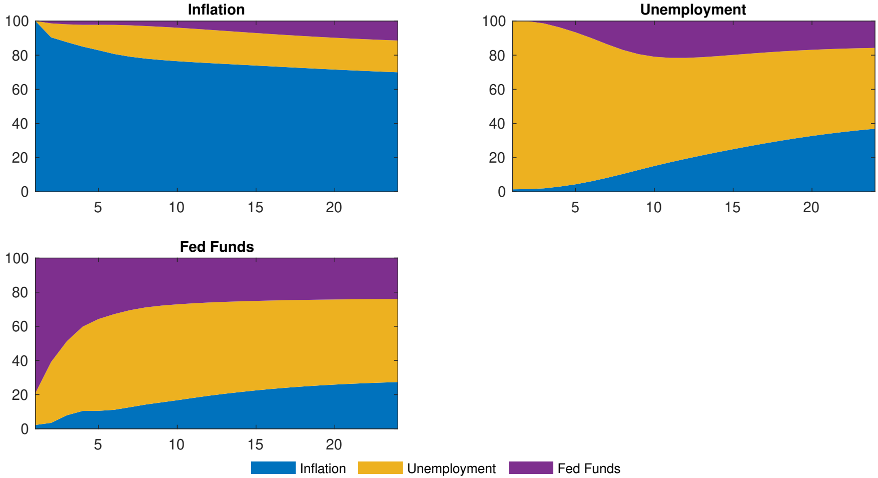
- Similarly, the *HD* can be computed with the `VARhd` function

```
% Compute HD
[HD, VAR] = VARhd(VAR,VARopt);
```

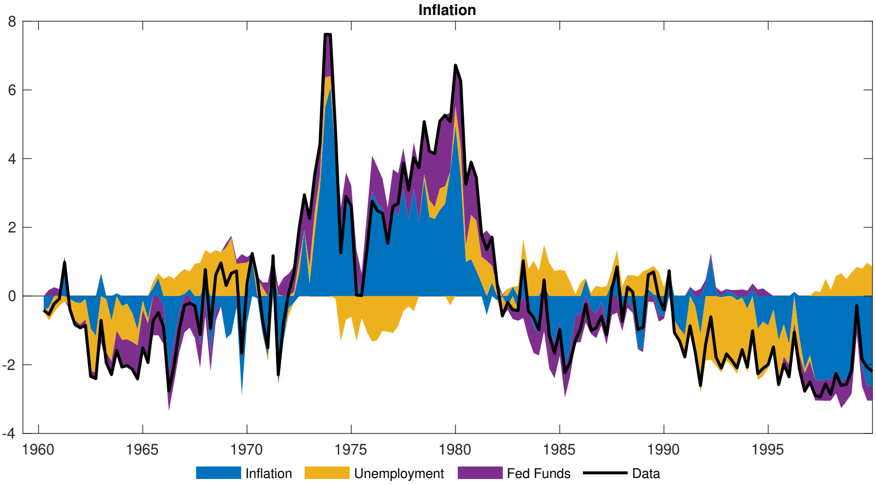
- Differently from `VD`, the output of `VARhd` is a structure `HD`

```
>> disp(HD)
     shock: [164x3x3 double]
      init: [164x3 double]
     const: [164x3 double]
     trend: [164x3 double]
    trend2: [164x3 double]
       exo: [164x3x0 double]
      endo: [164x3 double]
```

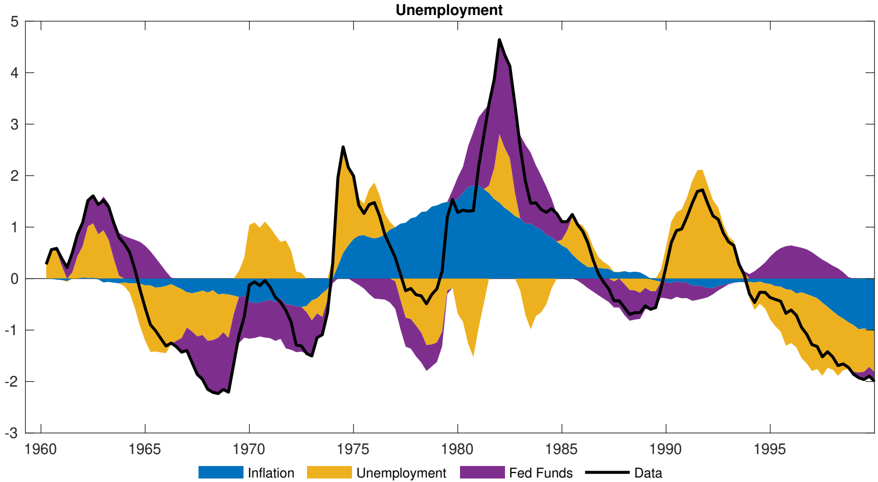
Forecast error variance decomposition



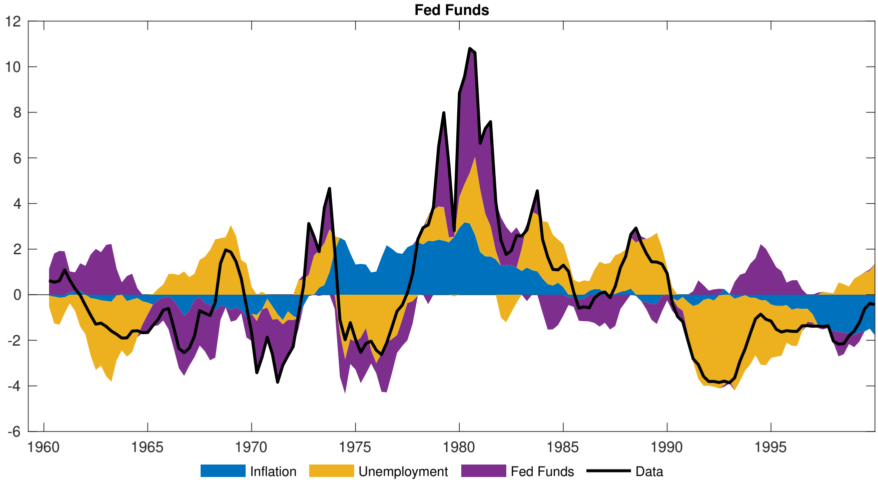
Historical decomposition



Historical decomposition



Historical decomposition



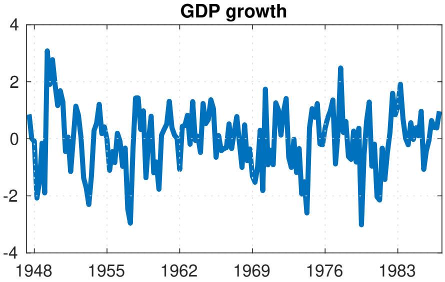
Practical Examples

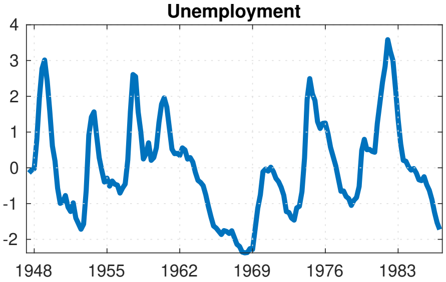
Blanchard and Quah (1989, AER)

Blanchard and Quah (1989): Zero long-run restrictions

- Blanchard and Quah (1989). “The Dynamic Effects of Aggregate Demand and Supply Disturbances”, *American Economic Review*

- US quarterly data from 1948Q1 to 1987Q4





What is the effect of demand and supply shocks?

- **Objective** Identify the effects of demand and supply shocks on output and unemployment
- Assume a bivariate VAR with $p = 8$ with output growth (y_t) and unemployment (u_t)
- **Key identifying assumption** Demand-side shocks have no long-run effect on the level of output, while supply-side shocks do
  - Blanchard and Quah impose zero long-run restrictions on the cumulative effect of demand shocks on output growth (i.e. on output level) to identify the shocks

$$\begin{bmatrix} y_{t,t+\infty} \\ u_{t,t+\infty} \end{bmatrix} = \begin{bmatrix} c_{11} & 0 \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Supply} \\ \varepsilon_t^{Demand} \end{bmatrix}$$

Monetary policy shocks, inflation and unemployment

- In Matlab, set lag length to 8 and estimate a VAR with a constant

```
% Set up and estimate VAR
det = 1;
nlags = 8;
[VAR, VARopt] = VARmodel(X,nlags,det);
```

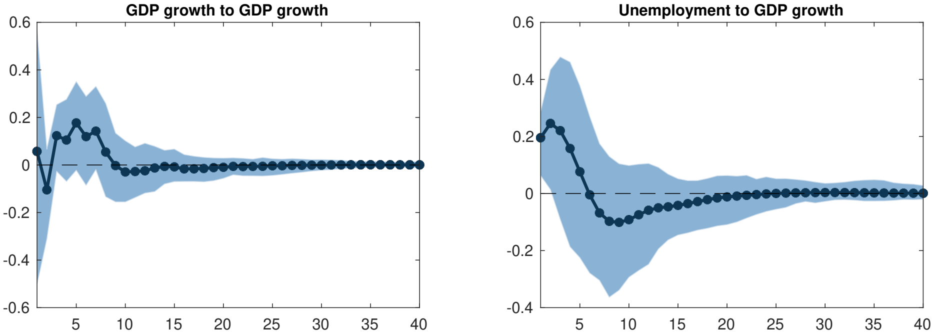
- Then set the option for zero long-run restrictions `VARopt.ident ='long'` and compute the *IR* with the `VARir` function. Note that the **ordering of the variables matter!**

```
% For zero contemporaneous restrictions set:
VARopt.ident = 'long';
% Compute IR
[IR, VAR] = VARir(VAR,VARopt);
```

- The *B* matrix implied by the zero long-run restrictions is stored in `VAR.B`

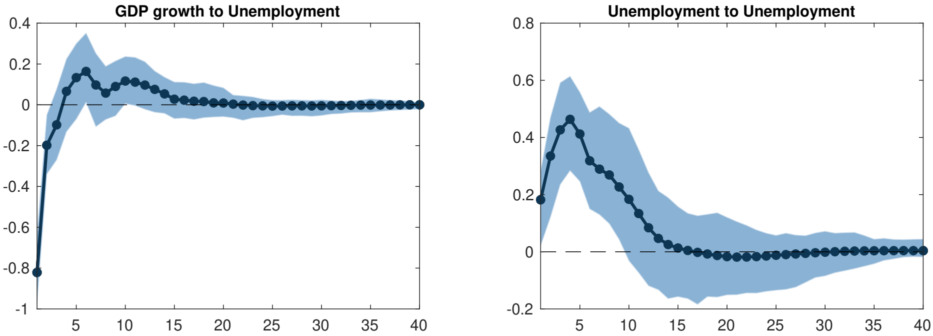
Aggregate supply shock

- Aggregate supply shock initially increases unemployment (puzzle of hours to productivity shocks)



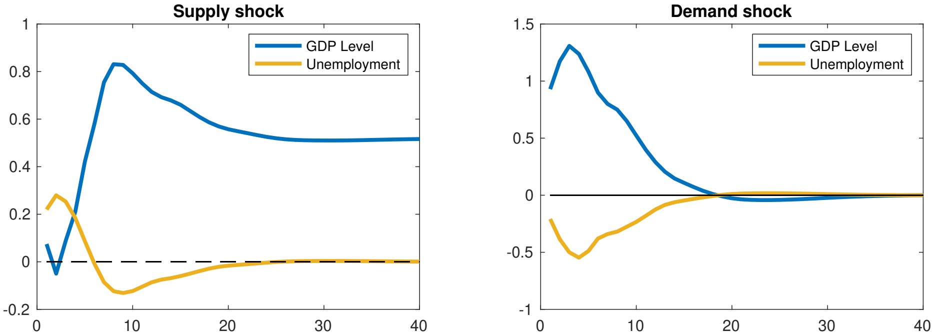
Aggregate demand shock

- Aggregate demand shocks have a hump-shaped effect on output and unemployment



What is the long run effect of demand and supply shocks on output level?

- Blanchard and Quah report (Figure 1) the cumulative sum of the impulse responses of output growth (i.e. the response of output level)
- By assumption, it should be zero for demand shocks ✔

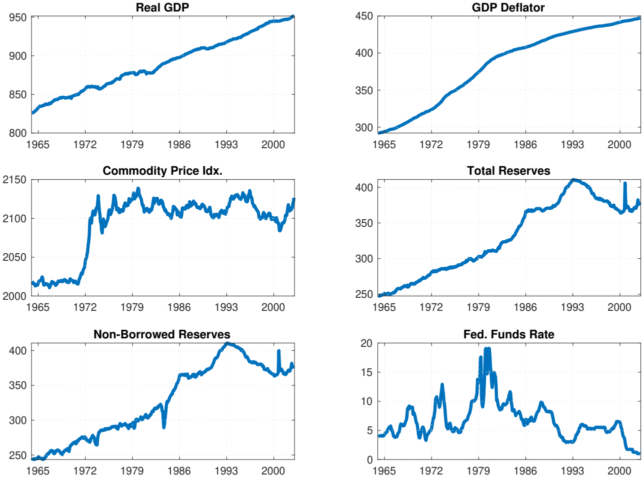


Practical Examples

Uhlig (2005, JME)

Uhlig (2005, JME): Sign restrictions

- Uhlig (2005). “What are the effects of monetary policy on output? Results from an agnostic identification procedure,” *Journal of Monetary Economics*
- US monthly data from 1965M1 to 2003M12



What are the effects of monetary policy on output?

- **Objective** Infer the causal influence of monetary policy on real GDP
- Assume a VAR with $p = 12$ with real GDP, real GDP deflator, a commodity price index, total reserves, non-borrowed reserves, and the fed. funds rate
- **Key identifying assumptions** According to conventional wisdom, monetary contractions should
  - Raise the federal funds rate
  - Lower prices
  - Decrease non-borrowed reserves
- Real GDP is left unrestricted

Monetary policy shock: The sign restrictions

- Uhlig imposes the following sign restrictions on the impulse responses of the VAR

	Monetary Policy Shock
Real GDP	?
Real GDP deflator	< 0
Commodity price index	?
Total reserves	?
Non-borrowed reserves	< 0
Fed. Funds Rate	> 0

- Restrictions are imposed for 6 periods

Monetary policy shock: The sign restrictions

- In Matlab, the sign restrictions can be set as follows

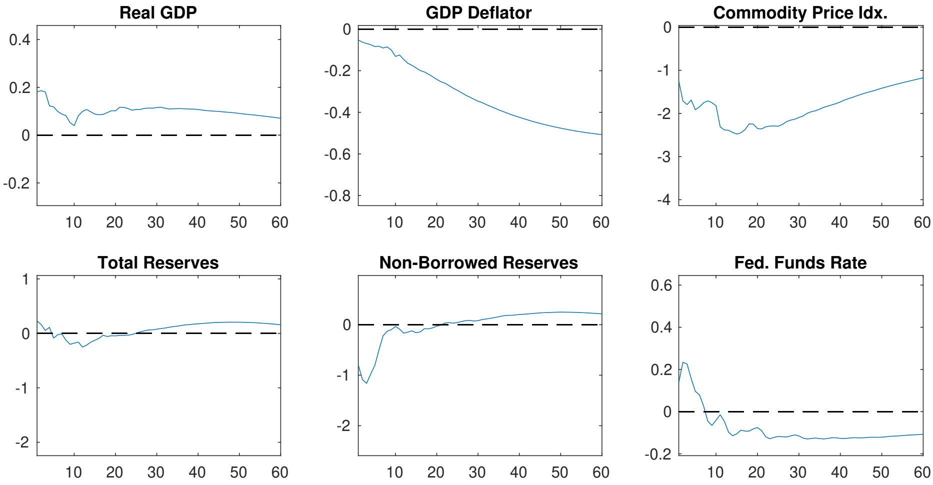
```
% Define the shock names
VARopt.snames = {'Mon. Policy Shock'};
% Define sign restrictions : positive 1, negative -1, unrestricted 0
SIGN = [ 0,0,0,0,0,0;  % Real GDP
        -1,0,0,0,0,0;  % Deflator
        -1,0,0,0,0,0;  % Commodity Price
         0,0,0,0,0,0;  % Total Reserves
        -1,0,0,0,0,0;  % NonBorr. Reserves
         1,0,0,0,0,0]; % Fed Funds
% Define the number of steps the restrictions are imposed for:
VARopt.sr_hor = 6;
```

- The routine is then implemented with the `SR` function

```
% Function SR performs the sign restrictions identification and computes
% IRs, VDs, and HDs. All the results are stored in SRout
SRout = SR(VAR,SIGN,VARopt);
```

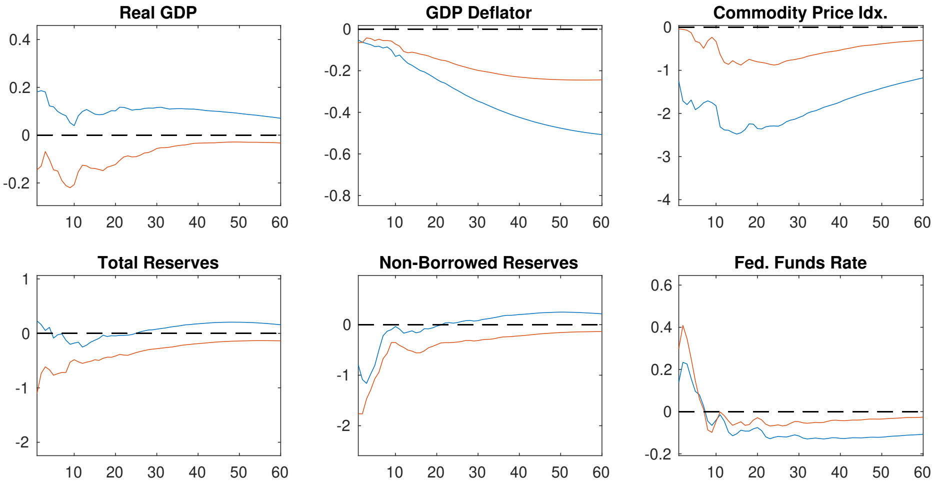
What happens when you do sign restrictions

- Start drawing orthonormal matrices Q until you find one that satisfies the restrictions...



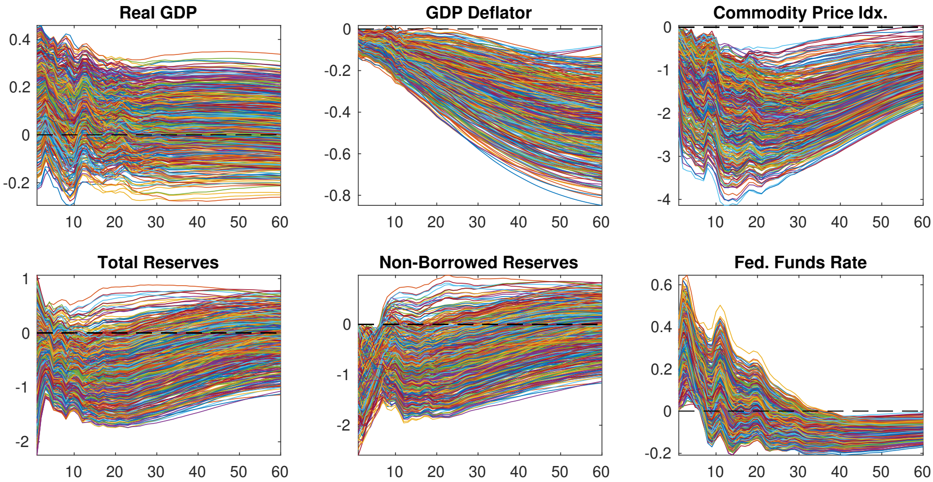
What happens when you do sign restrictions

- Drawing new Qs again until you find another one...



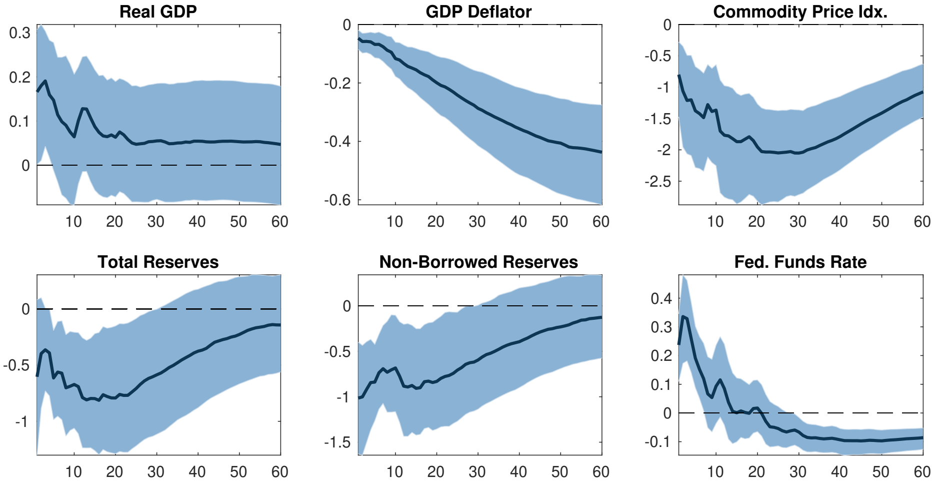
What happens when you do sign restrictions

- After a while...



What are the effects of monetary policy on output?

- Ambiguous effect on real GDP $\Longrightarrow$ Long-run monetary neutrality



Practical Examples

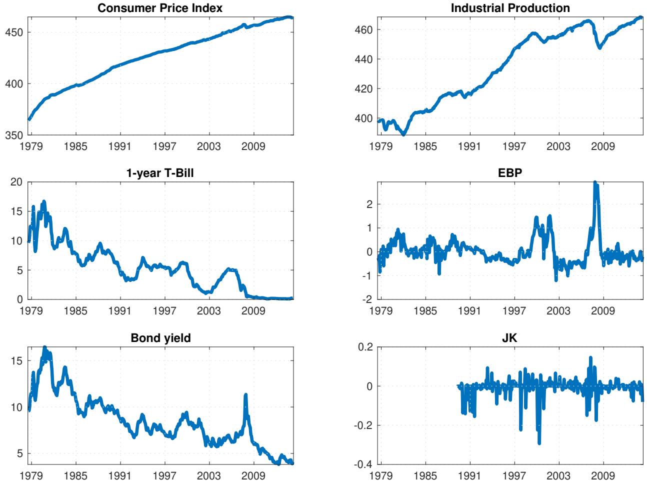
Gertler and Karadi (2015, AEJ:M)

Practical Examples

Cesa-Bianchi and Sokol (2020)

Cesa-Bianchi and Sokol (2020): Combining sign restrictions & external instruments

- Cesa-Bianchi and Sokol (2020). “Financial Shocks, Credit Spreads, and the International Credit Channel procedure,” *BoE Working Paper No. 693*
- Gertler and Karadi (2015)’s data set augmented with corporate bond yields
- Instrument for monetary policy shocks from Jarocinski and Karadi (2020, AEJ:M)



What are the effects of financial shocks?

- **Objective** Infer the causal influence of financial shocks on the economy and compare it to the financial transmission of monetary policy shocks

- **Challenge**
  - Financial shocks look like demand or supply shocks
  - Decrease output, increase credit spreads, reduce the policy rate (ambiguous about inflation)

- **Intuition**
  - First identify a monetary policy shocks with external instruments
  - Then exploit an ambiguity in the response of bond yields to impose sign restrictions

$$r_t^B = r_t(\downarrow) + cs_t(\uparrow)$$

- **Key identifying assumption** The increase in credit spreads dominates over the fall in the policy rate for financial shocks
  - When a financial shock hits r_t^B goes up, while it falls for demand/supply shocks

Monetary policy shock: The sign restrictions

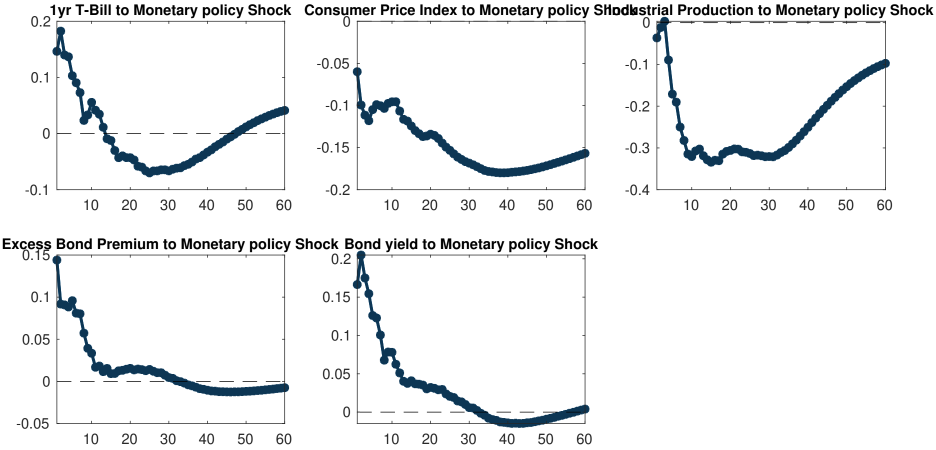
- Imposes the following sign restrictions on the impulse responses of the VAR

	Monetary
	Monetary Policy Shock
Real GDP)	?
Real GDP deflator)	< 0
Commodity price index	?
Total reserves	?
Non-borrowed reserves	< 0
Fed. Funds Rate	> 0

- Restrictions are imposed for 6 periods

The effect of monetary policy shocks

- Impulse responses are in line with Gertler and Karadi (2015) and Jarocinski and Karadi (2020)
  - Note: bootstrapping procedure still to be added to the toolbox



What are the effects of financial shocks on the economy?

- Persistent effects on output and inflation neutrality