Shocks, Sign Restrictions and Identification

Harald Uhlig[1]

[1]University of Chicago
Department of Economics
huhlig@uchicago.edu

Aug 17th, 2015

Outline

1. Supply and Demand
2. Shock identification in BVARs

Key questions:

- What happens after a supply shock?
- What happens after a monetary policy shock?

Goal:

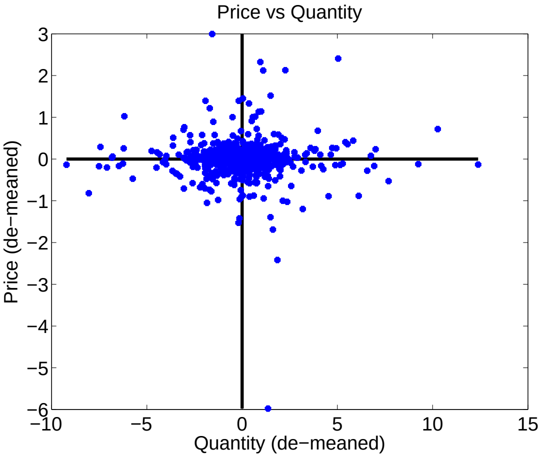
- Provide some generic, econometric perspective, ...
- ... emphasizing the recent debate on sign restrictions.

Sign restrictions: the beginning

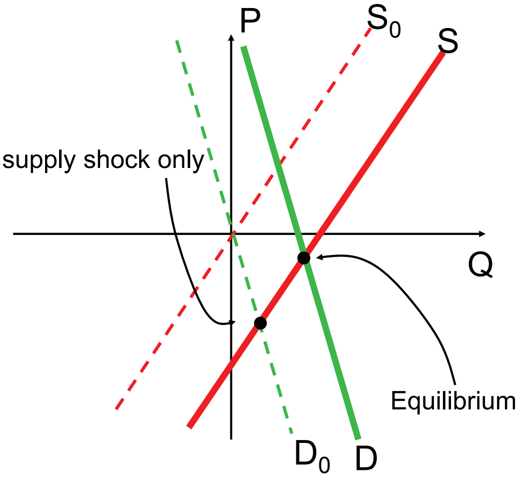
- Faust, 1998.
- Canova and De Nicoló, 2002.
- Uhlig, 2005.
- also: Uhlig, 1998, discussing Faust.
- The latter paper showed, that imposing sign restrictions on the dynamics of the impulse response adds considerable bite, in that particular application. It already contains the 2005 procedure.

Outline

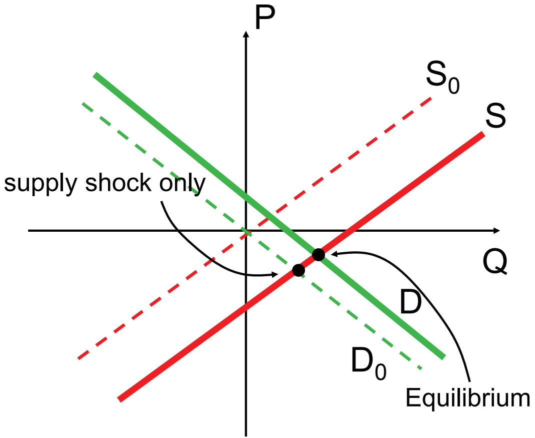
Supply and Demand: the iid sample



Supply and Demand Shocks



Supply and Demand Shocks: another possib.



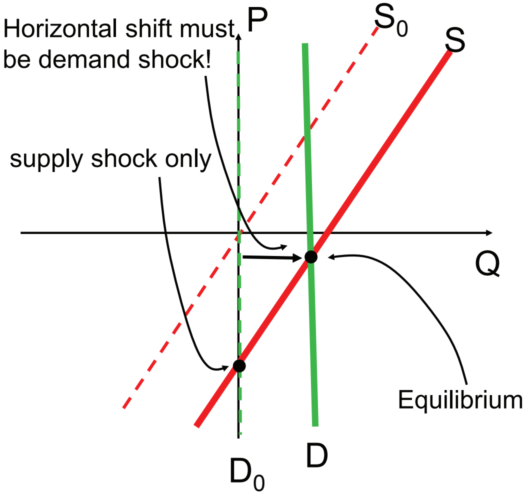
The problem

- Disentangle equilibrium observations (P, Q) into supply shocks and demand shocks.
- The raw data alone is insufficient.
- Additional identifying restrictions are needed.
- Key issue: which restrictions should one impose?

Principle 1: if you know it, impose it!

- D. McCloskey: Economics is Rhetoric.
- “Knowing”: you can convince your skeptical audience.

Example:vertical demand



Principle 2: if you do not know it, do not impose it!

- The convenience of an identification assumption can be a treacherous flytrap.

So, what do we know?

- Supply slopes up.
- Demand slopes down.
- (Do we? Other things?)

Let’s spell out some structure ...

The supply curve

The line

$$\begin{bmatrix} P \\ Q \end{bmatrix} = z_S \lambda_S + x_S \sigma_S \epsilon_S, \; \lambda_S \in \mathbb{R} \tag{1}$$

- z_S: direction of the supply curve
- x_S: direction of supply curve shift.
- $\sigma_S \epsilon_S$: the supply shock, where ϵ_S has unit variance.

The supply curve

The supply **curve** is the **line**

$$\begin{bmatrix} P \\ Q \end{bmatrix} = z_S \lambda_S + x_S \sigma_S \epsilon_S, \; \lambda_S \in \mathbb{R} \tag{2}$$

- Normalization: z_S of unit length, x_S orthogonal to z_S.
- Orthogonality of x_S? Algebraically convenient.
- Baumeister-Hamilton (2014): imposing $x_S(1) = 1$, say, can lead to Cauchy priors ... so don't.
- $$z_S = \begin{bmatrix} \cos(\nu_S) \\ \sin(\nu_S) \end{bmatrix}, \; x_S = \begin{bmatrix} -\sin(\nu_S) \\ \cos(\nu_S) \end{bmatrix} \tag{3}$$

  Ass. 1: Supply slopes upward: $\nu_S \in [0, \pi/2]$

  with add. normalization: supply shift is towards higher Q.
- Remark: the normalization can trip up even the best of us!
- **Distributional assumption:** $\epsilon_S \sim \mathcal{N}(0, 1)$. Dangerous?

The demand curve

The demand **curve** is the **line**

$$\begin{bmatrix} P \\ Q \end{bmatrix} = z_D \lambda_D + x_D \sigma_D \epsilon_D, \ \lambda_D \in \mathbb{R} \tag{4}$$

where $\epsilon_D \sim \mathcal{N}(0, 1)$ and

$$z_D = \begin{bmatrix} \cos(\nu_D) \\ \sin(\nu_D) \end{bmatrix}, x_D = \begin{bmatrix} -\sin(\nu_D) \\ \cos(\nu_D) \end{bmatrix} \tag{5}$$

Ass. 2: Demand slopes downward: $\nu_D \in [\pi/2, \pi]$

with add. normalization: demand shift is towards higher Q.

$+/-$ notation

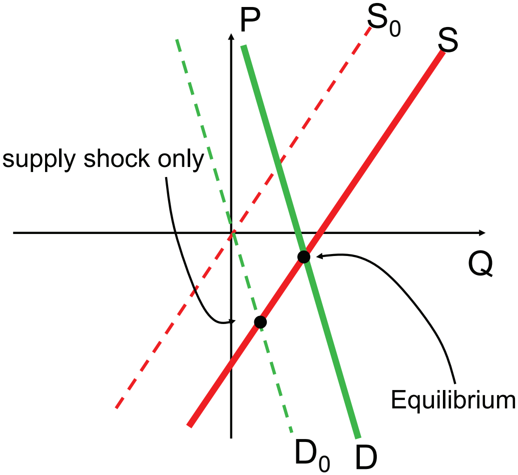
- Supply shocks only:

$$\begin{bmatrix} P \\ Q \end{bmatrix} = \begin{matrix} \epsilon_S & \epsilon_D \end{matrix} \begin{bmatrix} - & ? \\ + & ? \end{bmatrix}$$

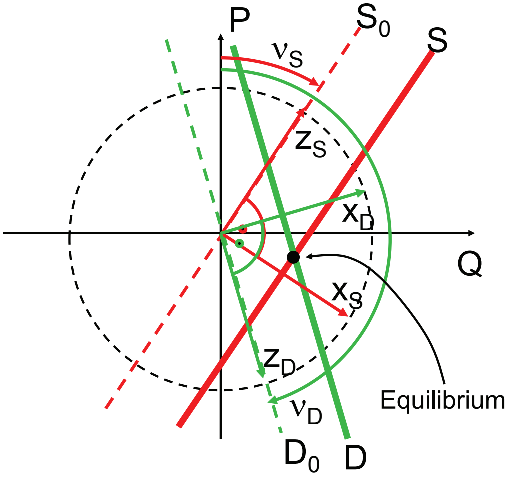
- Supply and demand shocks:

$$\begin{bmatrix} P \\ Q \end{bmatrix} = \begin{matrix} \epsilon_S & \epsilon_D \end{matrix} \begin{bmatrix} - & + \\ + & + \end{bmatrix}$$

Redo the first graph ...



... with the additional pieces



Independence of shocks

Ass. 3: ϵ_S and ϵ_D are mutually independent.

- Why?
- Without independence: if the supply shock hits, does the demand shock move too? Or the other way around?
- Chicken-and-egg problem.
- Perhaps, a third source of randomness is moving both. Find it, model it!

Principle 3: shocks are independent.

The structural parameters

$$\theta = (\nu_S, \nu_D, \sigma_S, \sigma_D) \tag{6}$$

Remark: θ is four-dimensional.

Ass. 4: $\sigma_S > 0$ and $\sigma_D > 0$.

- Simple RBC models: few shocks.
- More data series than shocks creates problems.
- Rest is measurement noise? Better: make them part of the model.

Principle 4: Have at least as many shocks as the length of the data vector.

Convenient definitions

$$
\begin{aligned}
Z &= \begin{bmatrix} z_D & z_S \end{bmatrix} = \begin{bmatrix} \cos(\nu_D) & \cos(\nu_S) \\ \sin(\nu_D) & \sin(\nu_S) \end{bmatrix} \\
X &= \begin{bmatrix} x_S' \\ x_D' \end{bmatrix} = \begin{bmatrix} -\sin(\nu_S) & \cos(\nu_S) \\ \sin(\nu_D) & -\cos(\nu_D) \end{bmatrix} \\
\Omega &= \begin{bmatrix} \sigma_S & 0 \\ 0 & \sigma_D \end{bmatrix}
\end{aligned}
$$

Equilibrium

$$\begin{bmatrix} P \\ Q \end{bmatrix} = X^{-1}\Omega\epsilon \tag{7}$$

Equilibrium distribution: given θ,

$$\begin{bmatrix} P \\ Q \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma(\theta)\right)$$

where

$$\Sigma(\theta) = A(\theta)A(\theta)'$$

with

$$A(\theta) = X^{-1}\Omega = \frac{1}{x_D' z_S}\begin{bmatrix} \sigma_S z_D & \sigma_D z_S \end{bmatrix} =: [a_S, a_D]$$

- First column: $[P, Q]$, after a one-std. dev. pos. supply shock.
- Second column: $[P, Q]$, after a one-std. dev. pos. demand shock.

The identification problem

- θ is four-dimensional.
- $\Sigma(\theta)$ is three-dimensional.
- There is one degree of freedom, in decomposing

$$\Sigma(\theta) = A(\theta)A(\theta)'$$

- If only we knew one of the parameters in θ! Example: vertical demand.
- But: perhaps, we only have the sign restrictions.
- Note: per our normalizations, *sgn* $(\det(A(\theta))) = -1$. Normalization can trip up the best of us.
- Note: $\theta \to A(\theta)$ is invertible.

Sign restrictions: sets of $\theta's$

- Restricting only demand to be downward sloping, i.e. restricting supply shocks to move P, Q in opposite direction:

$$\Theta_D = \{\theta = (\nu_S, \nu_D, \sigma_S, \sigma_D) \mid \nu_D \in [\pi/2, \pi], \nu_S \in (\nu_D - \pi, \nu_D), \sigma_S > 0, \sigma_D > 0\}$$

- Restricting demand and supply:

$$\Theta_{DS} = \{\theta = (\nu_S, \nu_D, \sigma_S, \sigma_D) \mid \nu_S \in [0, \pi/2], \nu_D \in [\pi/2, \pi], (\nu_S, \nu_D) \neq (0, \pi), \sigma_S > 0, \sigma_D > 0\}$$

Sign restrictions only:

Principle 5:
Without an additional exact restriction,
the sign restrictions deliver a one-dimensional
set Θ of $\theta' s$, which all could have
generated the data.

- Even asymptotically, i.e. if Σ is known exactly.
- One could parameterize this set by, say, $\nu_s \in N = [\underline{\nu}_S, \bar{\nu}_S]$.

Sign restrictions only

- How useful are sign restrictions?
- Let’s examine some special cases.
- Sign restriction only on supply versus sign restrictions on both supply and demand.

From Σ to A's

- Cholesky decomposition: $\Sigma = LL'$
- Then,

$$LR = A = [a_S, a_D] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

  for some orthogonal matrix

$$R = \begin{bmatrix} \cos(\mu) & \sin(\mu) \\ \sin(\mu) & -\cos(\mu) \end{bmatrix}$$

  is a candidate decomposition, with $det(A) < 0$.
- For any $A(\theta)$, one can find $R = R(\theta)$ or $\mu = \mu(\theta)$.
- A satisfies the sign restrictions on demand and supply, iff

$$A_{11} \leq 0, A_{21} \geq 0, A_{12} \geq 0, A_{22} \geq 0$$

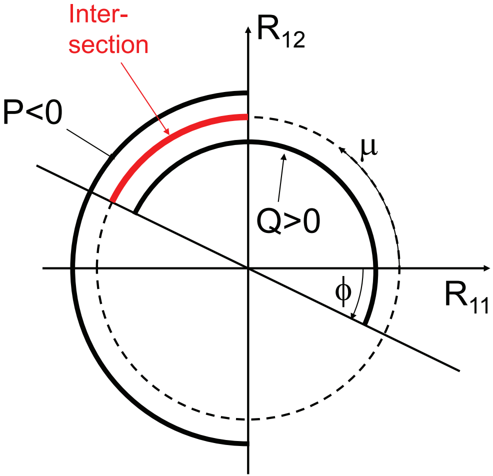
- A satisfies the sign restrictions on supply shocks, iff

$$A_{11} \leq 0, A_{21} > 0$$

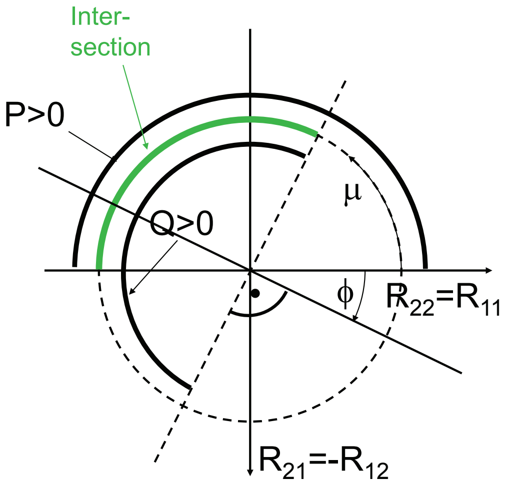
Re-parameterizing θ

- Write
$$\Sigma = \begin{bmatrix} \sigma_P^2 & \sin(\phi)\sigma_P\sigma_Q \\ \sin(\phi)\sigma_P\sigma_Q & \sigma_Q^2 \end{bmatrix}$$
for some $\phi \in (-\pi, \pi)$.
- $\sin(\phi)$ is the correlation between price and quantity.
- $\phi < 0$ iff corr$P, Q < 0$.
- Instead of $\theta = (\nu_S, \nu_D, \sigma_S, \sigma_D)$, use $\psi = (\phi, \mu, \sigma_P, \sigma_Q)$.
- Calculate
$$A = LR = \begin{bmatrix} a_S(\psi) & a_D(\psi) \end{bmatrix}$$
where
$$a_S(\psi) = \begin{bmatrix} \cos(\mu)\sigma_P \\ \sin(\phi+\mu)\sigma_Q \end{bmatrix}, \; a_D(\psi) = \begin{bmatrix} \sin(\mu)\sigma_P \\ -\cos(\phi+\mu)\sigma_Q \end{bmatrix}$$

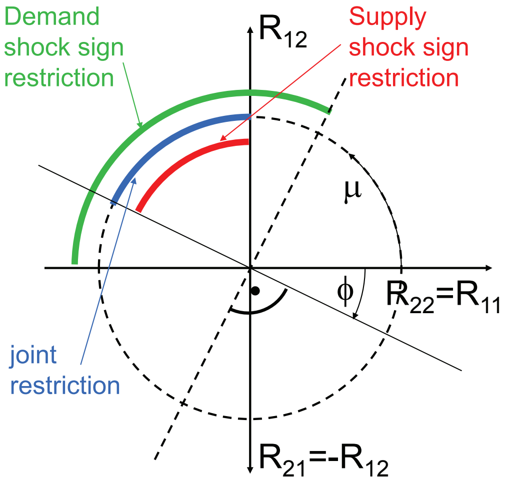
Sign restrictions: supply shock



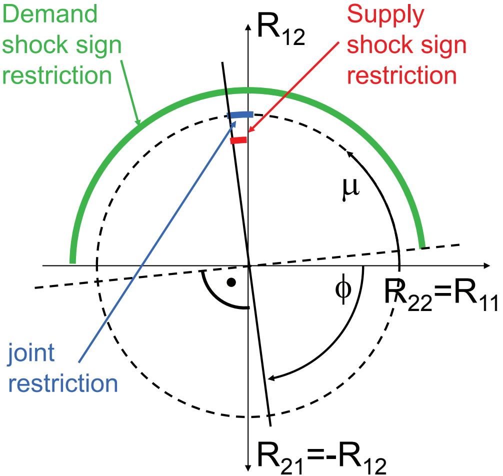
Sign restrictions: demand shock



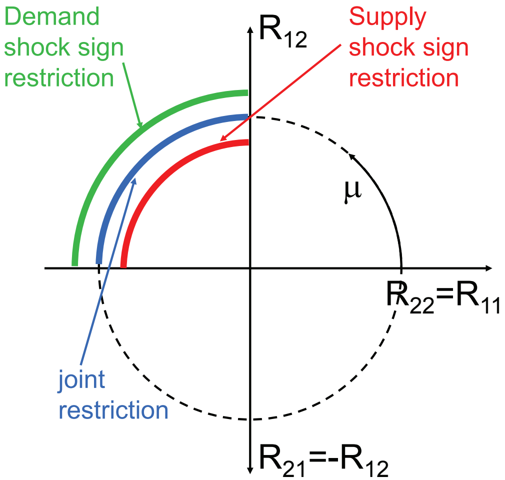
Sign restr. on both, when $\text{corr}(P, Q) > 0$.



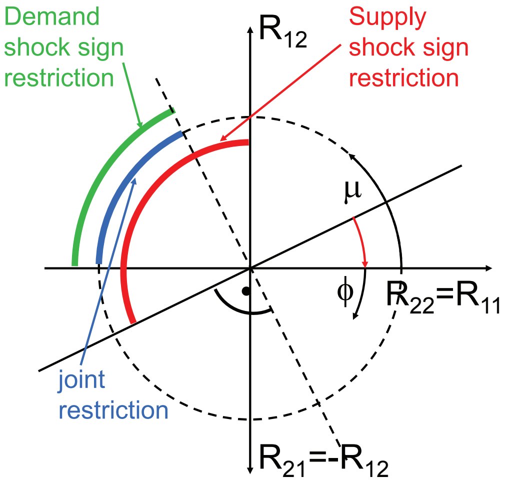
If corr$(P, Q) \to 1$, supply shock is identified.



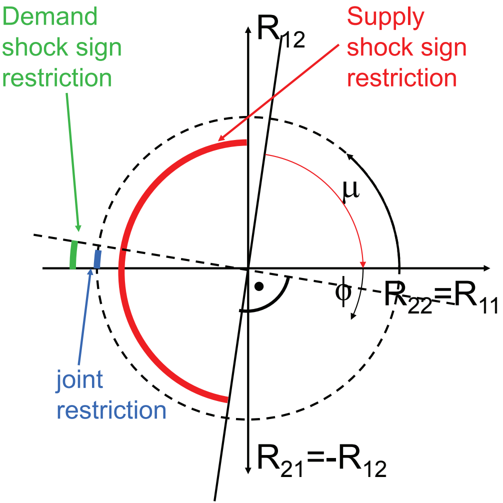
Sign restrictions, if corr$(P, Q) = 0$.



Sign restrictions, if $\text{corr}(P, Q) < 0$.



If corr$(P, Q) \to -1$, demand shock is identified.



Some observations

Principle 7: sign-restricting both can help, sometimes a lot, sometimes not at all.

Draw-and-reject procedure:

- Draw μ from prior on $[0, 2\pi]$.
- Calculate the resulting $A = LR(\mu)$.
- Check sign restrictions:
  - If satisfied, keep draw.
  - If not satisfied, reject draw.

Principle 8: when you reject a lot of draws, it means you have sharp identification! Good!

Set identification or point identification?

Principle 6: different questions usually have different answers.

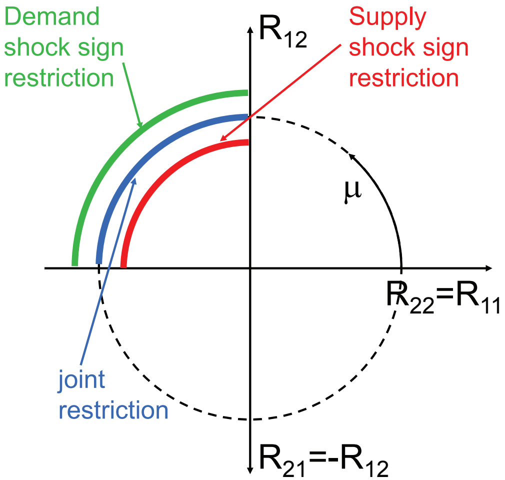
- treat the **set** N or Θ as the object of interest: $\hat{N}, \hat{\Theta}$ with finite data.
  - If you seek to cover the true set N with 95% probability, $\hat{N}$ will typically be **larger** than the true N.
- Uhlig (2005): treat θ as the object of interest.
  - Bayesian prior over $\theta' s$.
  - Posterior = prior asymptotically on Θ. Obvious! (Baumeister-Hamilton,...)
  - Set $\hat{N}$, which contains 95% of all draws will typically be **smaller** than the true N.
- These two $\hat{N}$ are different because they answer different questions.

Classical or Bayesian Perspective?

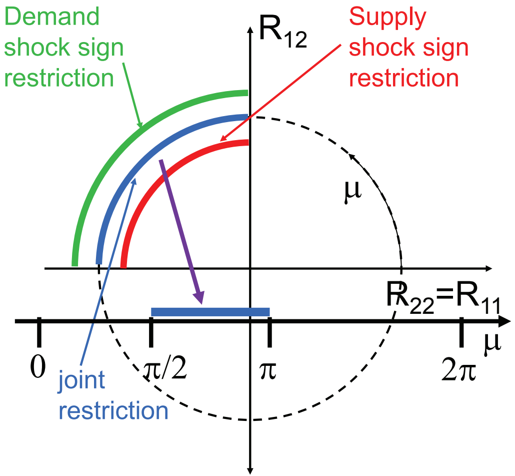
Principle 6: different questions usually have different answers.

- Classical approach: flat likelihood on set Θ.
- Bayesian approach: priors over θ. Posterior = prior on Θ.
- Fry-Pagan, p.948: *"... referring to this range as if it is a confidence interval ... is quite false ... it should not be imbued with probabilistic language."*
  - True from a classical perspective.
  - Wrong for a Bayesian.
- The Bayesian samples from a well-defined probability distribution. There is no "multiple shocks" or "multiple models" issue.
- Moon-Hyungsik-Schorfheide, Ecmta 2012: there is a difference between classical and Bayesian inference, even asymptotically:
  - A Bayesian 95% HPD set is "strictly inside".
  - A classical 95% conf. set is "strictly outside."

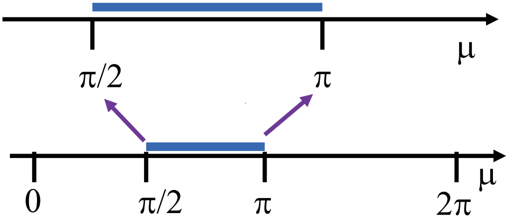
Sign restrictions, if $\text{corr}(P, Q) = 0$, ...



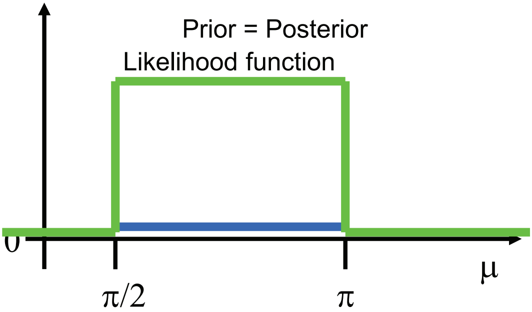
... and examining the interval for μ.



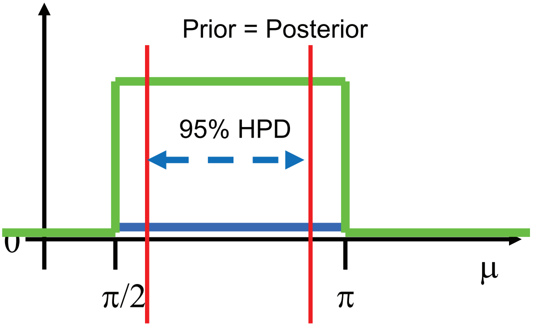
Let's magnify.



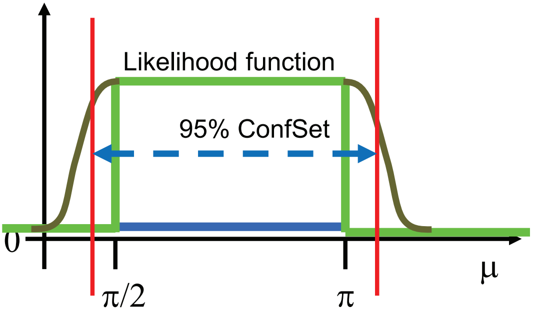
Asymptotically



Bayesian inference



Classical inference



Outline

1. Supply and Demand
2. Shock identification in BVARs

BVARs

- Dynamics: lags matter, observations are rarely iid.
- More than two dimensions.
- Thus: let's use (B)VARs

BVARs

Reduced form: u_t are the one-step ahead prediction errors.

$$
\begin{aligned}
Y_t &= c + \sum_{j=1}^{k} B_j Y_{t-j} + u_t \\
\Sigma &= E[u_t u_t']
\end{aligned}
$$

Structural: ϵ_t are the structural shocks:

$$
\begin{aligned}
\Sigma &= AA' \\
u_t &= A\epsilon_t
\end{aligned}
$$

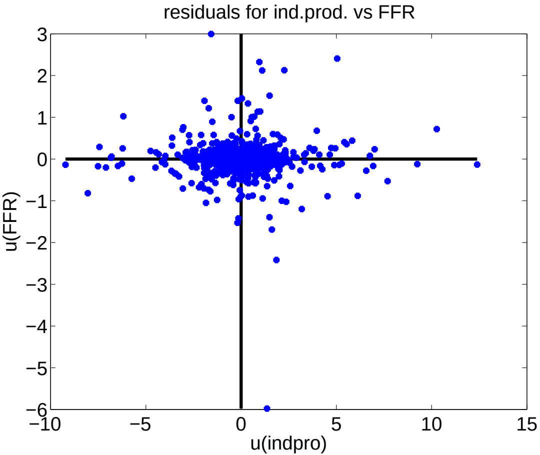
A benchmark macro example.

Focus question:
the effect of a monetary policy shock.

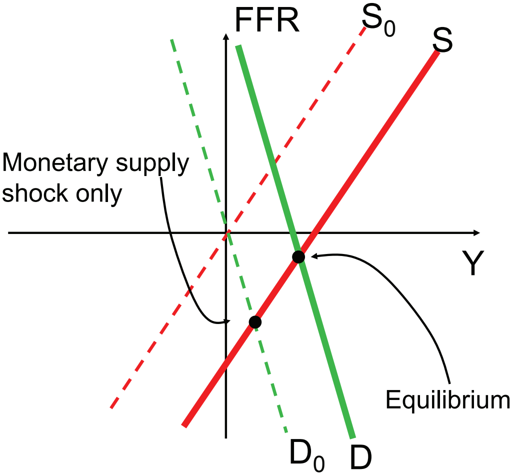
Example: a 4-variable VAR

- Data: monthly real GDP, CPI, FFR, M1, 1959-2015.
- monthly real GDP: a rescaled version of industrial production.

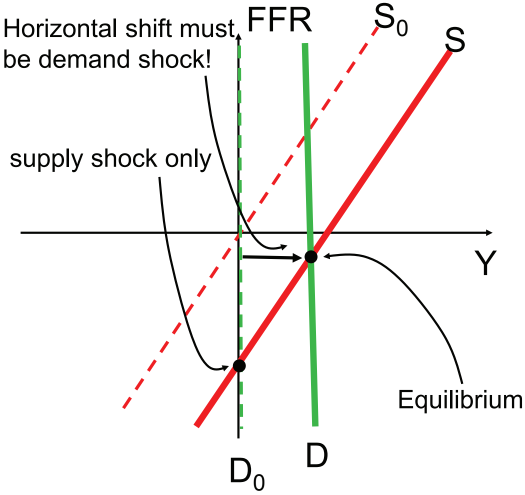
Supply and Demand: the iid sample



Supply and Demand Shocks



Supply and Demand Shocks: Cholesky decomposition.



The problem

- Disentangle equilibrium observations (realY,CPI,FFR,M1) into monetary policy shocks and other shocks.
- The raw data alone is insufficient.
- Additional identifying restrictions are needed.
- Key issue: which restrictions should one impose?

Principle 1: if you know it, impose it!

Principle 2: if you do not know it, do not impose it!

Cholesky decomposition

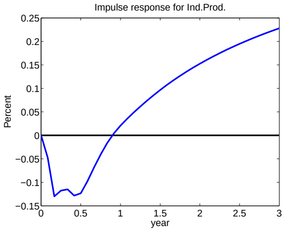
- “Slow-moving” variables, “Fast-moving” variables (Bernanke-Mihov,...)
- But note: there is a non-zero prediction error for all variables!
- So, even the “slow” movers move in response to **some** shock.
- If they move “slowly”, that should mean that their prediction errors are small ...
- ... but it does not imply that they cannot respond to monetary policy shocks just as much as to any other shocks.

Principle 9: the Cholesky decomposition is convenient, clear. However, the “slow-fast” defense makes little sense.

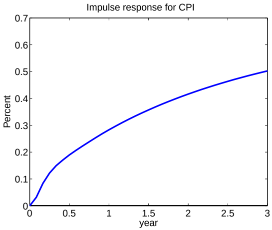
Cholesky decomposition results: price puzzle

Response to third shock, as mon.pol. shock:

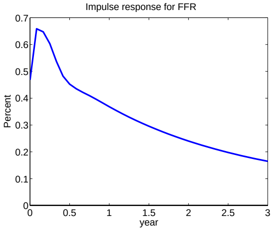
GNP



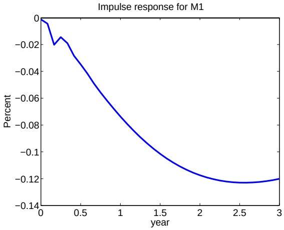
CPI



FFR



M1



Cholesky decomposition: the empirics

Principle 10: if you are worried by the "price puzzle" and the like, do not rely on the Cholesky decomposition. Use it, if you are willing to "live or die" by its implications.

- Results change over time.
- "Price puzzle" seems to have gotten worse.
- Most researchers would then abandon the Cholesky decomp.
- Only use it with much precaution.

Long-run restrictions

- Blanchard-Quah: only supply shocks have long-run effects on output.
- Really? Endogenous growth theory will tell you: everything does.
- And: can we be sure that we have unit roots?

Principle 1: if you know it, impose it!

Principle 2: if you do not know it, do not impose it!

Sign restrictions.

- Sign-restrictions on monetary policy shock (Uhlig, 2005).
- Should one sign-restrict other shocks as well?
- In (P, Q) example: it may help a lot.
- Should one impose additional zero or long-run restrictions?
- Again, it may help a lot.

Sign restrictions on monetary pol. shocks only?

Principle 1: if you know it, impose it!

Principle 2: if you do not know it, do not impose it!

Monetary policy shocks, other shocks:

$$\begin{bmatrix} realY \\ CPI \\ FFR \\ M1 \end{bmatrix} = \begin{array}{c} \begin{matrix} \epsilon_m & \epsilon_1 & \epsilon_2 & \epsilon_3 \end{matrix} \\ \begin{bmatrix} ? & ? & ? & ? \\ - & ? & ? & ? \\ + & ? & ? & ? \\ - & ? & ? & ? \end{bmatrix} \end{array}$$

- Other shocks?
- **Linear combinations of other shocks**?
- Successful application: Uhlig, 2005 (using six variables).
- Result: not much can be said about Y, just with this.

Restrictions on A or A^{-1}?

- $\Sigma = AA'$.
- A: instantaneous reaction of variables to shocks.
- A^{-1}: "policy rule"?
- Arias-Caldara-Rubio-Ramirez (2014): on A^{-1}. Within the period,
  - FFR reacts positively to output and prices, ...
  - .. and not to total reserves or nonborrowed reserves.
- But why?
  - A typical shock will move everything.
  - So policy maker will do inverse mapping from equilibrium outcome back to the shocks, such that these restrictions hold? How are they supposed to do that?
  - Taylor-rule: is in terms of output gap.
  - A positive increase of output can make FFR go up or down in standard NK models: depends on shock!
- Identifying assumptions of A-C-RR seem highly dubious to me.
- Baumeister-Hamilton (2014): economic intuition is about restricting A. I disagree.

Restrictions on A or A^{-1}?

Principle 11: argue about shocks and their propagation. That means imposing restrictions on A, not A^{-1}.

This is violated in A-C-RR (2014). From my perspective, they are violating Principles 1 and 2.

Priors and posteriors

$$Y_t = c + \sum_{j=1}^{k} B_j Y_{t-j} + u_t, \;\; \Sigma = E[u_t u_t']$$

The (conditional) likelihood function for $(c, B_1, \ldots, B_k, \Sigma)$ is proportional to a Normal-Wishart density, which specifies that

- Σ^{-1} is Wishart,
- and, conditionally on Σ, the vector of coefficients $\mathbf{b} = \text{vec}(B_1, \ldots, B_k)$ follows a Normal distribution $\mathcal{N}(\bar{\mathbf{b}}, \Sigma \otimes \mathbf{N}^{-1})$, for some $\mathbf{N}$

Principle 12: the likelihood function is proportional to a Normal-Wishart. This considerably restricts the space to where the posterior can be taken.

Taking draws

- Posterior in $\mathbf{b}, \Sigma$ is prop. to prior times Normal Wishart.
- Take many draws $\mathbf{b}_i, \Sigma_i$.
- Cholesky-decompose $\Sigma_i = L_i L_i'$.
- Any other decomposition $\Sigma_i = A_i A_i'$ satisfies

$$A_i = L_i R_i, \text{ where } R_i R_i' = I$$

- Draw R_i. Check whether A_i satisfies sign restrictions.
- If you draw a single vector: draw r_i with $\| r_i \| = 1$ (“first column of R_i”) and calculate $a_i = L_i r_i$ (“first column of A”). Check sign restrictions.
- **Question:** how to draw R_i or r_i?

In defense of the Haar measure

- **Question:** how to draw R_i or r_i?
- Recall: any other decomposition $\Sigma_i = A_i A_i'$ satisfies

$$A_i = L_i R_i, \text{ where } R_i R_i' = I$$

- ... but the particular choice of the decomposition $\Sigma_i = L_i L_i'$ shouldn't matter.
- That means, it should be equally likely to draw R_i or $\tilde{R}_i = QR_i$, where $QQ' = I$: **Haar measure**, Rubio-Ramirez-Waggoner-Zha.
- With one vector: it should be equally likely to draw r_i or $\tilde{r}_i = Qq_i$, where $QQ' = I$: uniform on sphere.
- Baumeister-Hamilton: transformations of uniform priors look informative. Of course they do.

Principle 13: use the Haar measure to draw R_i or r_i.

Interaction of sign restrictions and reduced-form VAR

- The sampling procedure described samples from a joint prior $(\mathbf{b}, \Sigma, R)$.
- Consider two different $(\mathbf{b}, \Sigma)$. If sign restrictions are easy to satisfy for the first, but difficult to satisfy for the second, the first will be sampled more often, i.e., be given higher marginal probability relative to the no-sign-restriction case.
- But “difficult to satisfy” means the sign restrictions are better at tightly identifying the shock!
- So, the procedure puts more weight on $(\mathbf{b}, \Sigma)$, which offer less tight identification.
- It may be better to use a conditional prior instead:
  - Some standard prior for $(\mathbf{b}, \Sigma$, so take draw.
  - Check whether there is any R then satisfying sign restrictions.
  - If not, discard.
  - If yes, draw a fixed number that satisfy sign restrictions.

Sign restrictions on impact only?

- Should one impose sign restriction on horizons beyond the impact? And for how long?
- Canova-Paustian: in a model without capital, doubts about persistence of signs of shocks.
- Note: in iid (P, Q) example, **no** persistence of shocks.
- It is easy to construct non-persistent theories. Is that what we believe?
- If you believe in your DSGE model, impose it! No reason to use sign restrictions.
- Usually, we only trust certain implications of these models, though.
- Typically, it is economically reasonable that sign restrictions persist for, say, up to a year.

Principle 1: if you know it, impose it!

Principle 2: if you do not know it, do not impose it!

A small example

- A bivariate VAR with one lag.
- Suppose: for some reason, the coefficient matrix is always diagonalizable, and has two identical roots.
- Case 1: the two roots are real. In that case, signs persist forever.
- Case 2: the two roots are complex conjugate. In that case, we get oscillations (“damped sin waves”).
- Sign restrictions up to some horizon K:
  - do not further inform case 1.
  - are incompatible with complex conjugate roots, where the frequency of the oscillations is “too high”.
- So: dynamic sign restrictions can have considerable bite. They do not have to.

How to read Impulse response ranges

- Bayesian procedure: these are proper probability statements about posteriors.
- They provide graphic information about marginal distributions, at each horizon and for each variable plotted.
- Connecting the medians from these marginals does not represent some particular draw or even “median” draw. They simply show how the medians and these marginal distributions evolve.
- This has always been true for Bayesian impulse response error bands.
- Fry-Pagan (2011): seem to think this is a big deal (p. 949).

Some final recommendations

- Sign restrictions may be weak. But what else is there?
- Cholesky-decomposition, long-run restrictions: questionable.
- Variance-maximization procedures, regional data, panel data, high-frequency identification approaches, regime shifts ...: there is lots more.
- Just be careful that general lessons learned from VARs on aggregate time series are not getting lost.

Principle 1: if you know it, impose it!

Principle 2: if you do not know it, do not impose it!