No class Wed Oct 18

Set identification using sign restrictions

Could we still draw structural conclusions using much weaker identifying assumptions, e.g., supply curve slopes up and demand curve slopes down?

$\boldsymbol{\varepsilon}_t$ = vector of VAR forecast errors

$$\boldsymbol{\Omega} = E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$$

$$\hat{\boldsymbol{\varepsilon}}_t = \mathbf{y}_t - \hat{\mathbf{c}} - \hat{\boldsymbol{\Phi}}_1 \mathbf{y}_{t-1} - \cdots - \hat{\boldsymbol{\Phi}}_p \mathbf{y}_{t-p}$$

$$\hat{\boldsymbol{\Omega}} = T^{-1} \sum_{t=1}^{T} \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t'$$

$\mathbf{v}_t$ = vector of structural shocks

$$E(\mathbf{v}_t \mathbf{v}_t') = \mathbf{I}_n$$

$$\boldsymbol{\varepsilon}_t = \mathbf{H} \mathbf{v}_t$$

$$E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \boldsymbol{\Omega} = \mathbf{H}\mathbf{H}'$$

$E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \boldsymbol{\Omega} = \mathbf{H}\mathbf{H}'$

One example of an $\mathbf{H}$ we could consider is Cholesky factor $\boldsymbol{\Omega} = \mathbf{P}\mathbf{P}'$ for $\mathbf{P}$ lower triangular. The set of all possible $\mathbf{H}$ can be characterized as $\mathbf{H} = \mathbf{P}\mathbf{Q}$ for $\mathbf{Q} \in O_n,$ the set of all orthonormal $(n \times n)$ matrices

$$O_n = \{\mathbf{Q} : \mathbf{Q}\mathbf{Q}' = \mathbf{I}_n\}$$

Proof:

(1) If $\mathbf{H} = \mathbf{PQ}$ then $\mathbf{HH}' = \mathbf{PQQ}'\mathbf{P}' = \mathbf{\Omega}$

(2) If $\mathbf{HH}' = \mathbf{\Omega}$ then $\mathbf{HH}' = \mathbf{PP}'$ and $\mathbf{P}^{-1}\mathbf{HH}'(\mathbf{P}')^{-1} = \mathbf{I}_n$ so $\mathbf{P}^{-1}\mathbf{H} = \mathbf{Q}$ must be an orthonormal matrix (that is, $\mathbf{H}$ must be of form $\mathbf{H} = \mathbf{PQ}$)

What does O_n look like for $n = 2$?

$$\mathbf{Q} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ or } \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

for $\theta \in [-\pi, \pi]$.

If we generated $\theta \sim U[-\pi, \pi]$ and then selected one of the above matrices with prob 1/2, this is described as a distribution over O_2 that is Haar-uniform.

Rubio-Ramírez, Waggoner and Zha (2010) algorithm for generating a Haar-uniform draw from O_n.

(1) Generate an $(n \times n)$ matrix $\mathbf{X}$ of independent $N(0, 1)$ variables.

(2) Calculate the QR decomposition $\mathbf{X} = \mathbf{QR}$ where $\mathbf{Q}$ is orthonormal and $\mathbf{R}$ is upper triangular

Matlab: [Q,R] = qr(X)

How the QR decomposition works: first column of $\mathbf{Q}$ is simply first column of $\mathbf{X}$ normalized to have unit length:

$$\begin{bmatrix} q_{11} \\ q_{21} \\ \vdots \\ q_{n1} \end{bmatrix} = \begin{bmatrix} x_{11}/\sqrt{x_{11}^2 + \cdots + x_{n1}^2} \\ x_{21}/\sqrt{x_{11}^2 + \cdots + x_{n1}^2} \\ \vdots \\ x_{n1}/\sqrt{x_{11}^2 + \cdots + x_{n1}^2} \end{bmatrix}$$

$$\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} = \begin{bmatrix} x_{11}/\sqrt{x_{11}^2 + x_{21}^2} \\ x_{21}/\sqrt{x_{11}^2 + x_{21}^2} \end{bmatrix}$$

For $n = 2$, q_{11} is cosine of angle formed by x_{11}, x_{21} and q_{21} is the sine.

$$\mathbf{Q} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ or } \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

$\theta \sim U(-\pi, \pi)$

Algorithm for generating possible draws for $\mathbf{H}$.

(1) Either fix $\boldsymbol{\Omega}$ and $\boldsymbol{\Gamma}$ at MLEs $\hat{\boldsymbol{\Omega}} = T^{-1}\sum_{t=1}^{T}\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t'$ and $\hat{\boldsymbol{\Gamma}}' = \left(\sum_{t=1}^{T}\mathbf{y}_t\mathbf{x}_t'\right)\left(\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right)^{-1} \Rightarrow \hat{\boldsymbol{\Psi}}_s$ or draw $\boldsymbol{\Omega}^{-1}$ from Wishart with $T - p$ degrees of freedom and scale matrix $T\hat{\boldsymbol{\Omega}}$ and use this to draw vec$(\boldsymbol{\Gamma}) \sim N\left(\text{vec}\left[\hat{\boldsymbol{\Gamma}}\right], \boldsymbol{\Omega} \otimes \left[\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]^{-1}\right)$.

(2) Find Cholesky factor $\mathbf{\Omega} = \mathbf{P}\mathbf{P}'$, draw $\mathbf{Q}$ from Haar distribution, and calculate candidate $\mathbf{H} = \mathbf{P}\mathbf{Q}$.

(3) Calculate signs of chosen magnitudes in $\mathbf{\Psi}_s\mathbf{H}$ and keep draw if these satisfy theory, otherwise throw out.

(e.g., monetary contraction raises interest rate, lowers output and inflation on impact ($s = 0$)

(4) Researchers typically report median accepted draw for element i,j of $\mathbf{\Psi}_s\mathbf{H}$ as if it is estimate of effect of structural shock j on variable i and 68% of draws around the median as if they were "error bands" (this is problematic!)

Example:

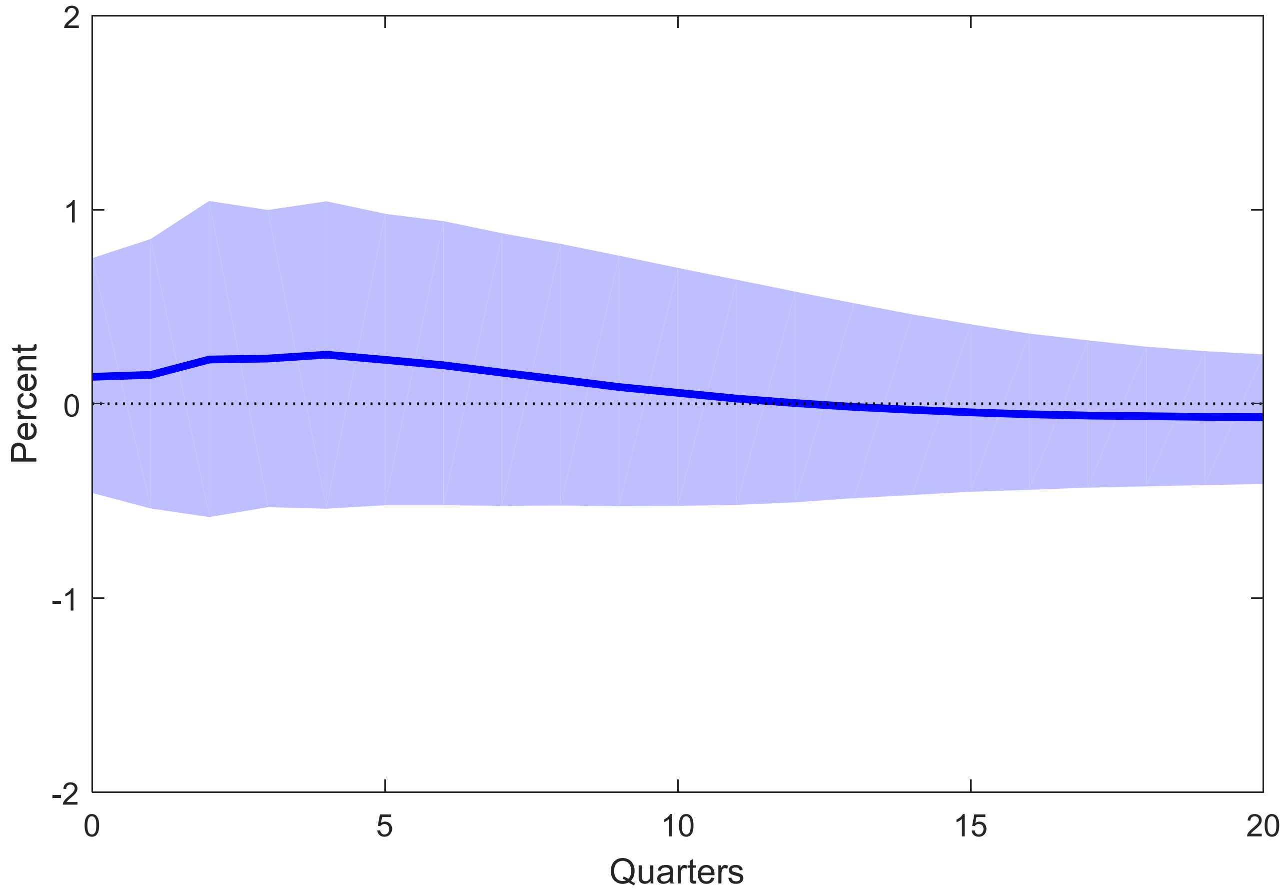
y_{1t} = fed funds rate

y_{2t} = log output gap

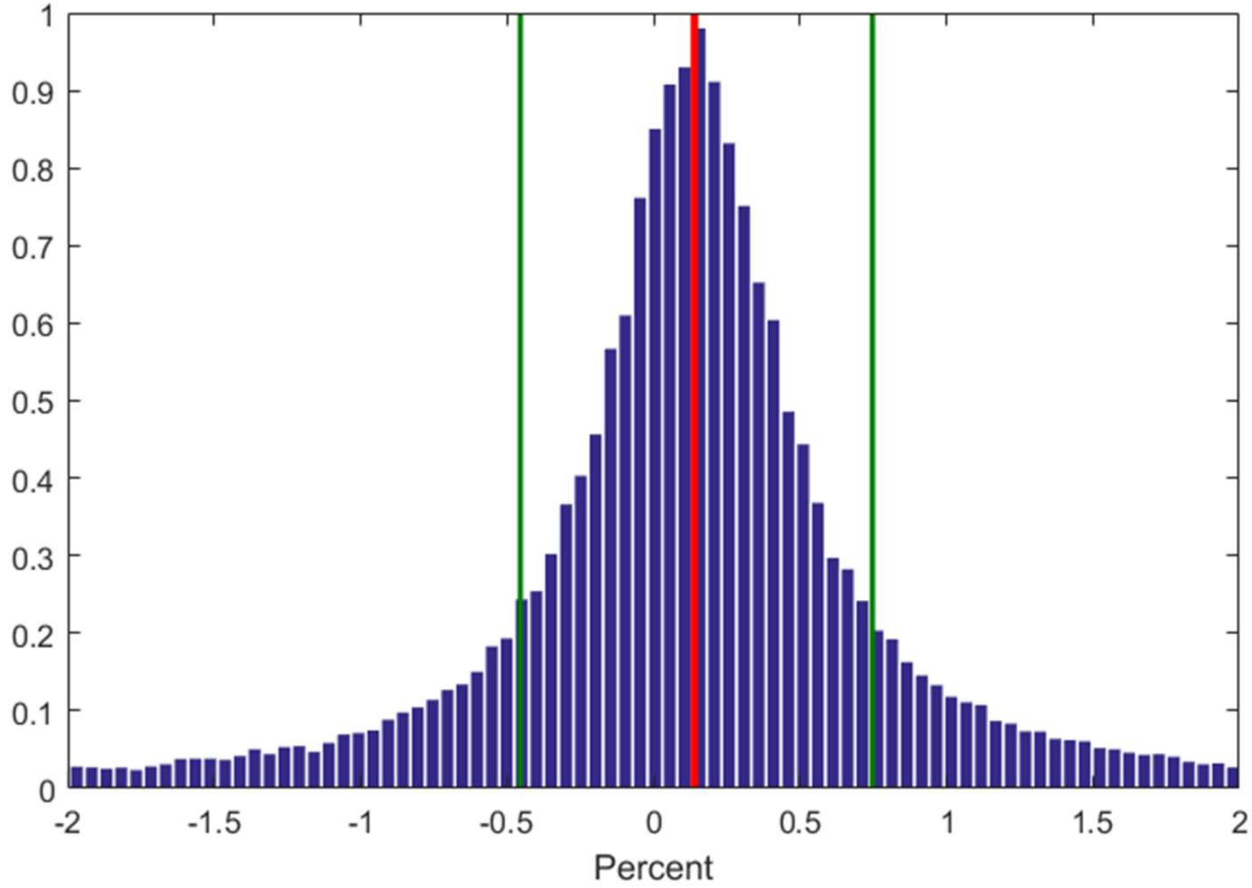
y_{3t} = inflation

Let's run the algorithm to find the effect on output of a monetary policy shock that raises fed funds rate by 0.25%, with one change– we forget to throw any of the draws out!

Supposed effect on output gap of 0.25% monetary contraction without making any assumptions (68% “error bands”)



Supposed effect at horizon zero on output gap of 0.25% monetary contraction



This magnitude is 0.25 times the ratio of the (2,1) element of $\mathbf{H}$ to the (1,1) element $=$ 0.25 times ratio of effect of shock 1 (monetary policy?) on output to its effect on fed funds rate

$\mathbf{h}_1 = \mathbf{P}\mathbf{q}_1$

$$= \begin{bmatrix} p_{11} & 0 & 0 \\ p_{21} & p_{22} & 0 \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \begin{bmatrix} x_{11}/\sqrt{x_{11}^2 + x_{21}^2 + x_{31}^2} \\ x_{21}/\sqrt{x_{11}^2 + x_{21}^2 + x_{31}^2} \\ x_{31}/\sqrt{x_{11}^2 + x_{21}^2 + x_{31}^2} \end{bmatrix}$$

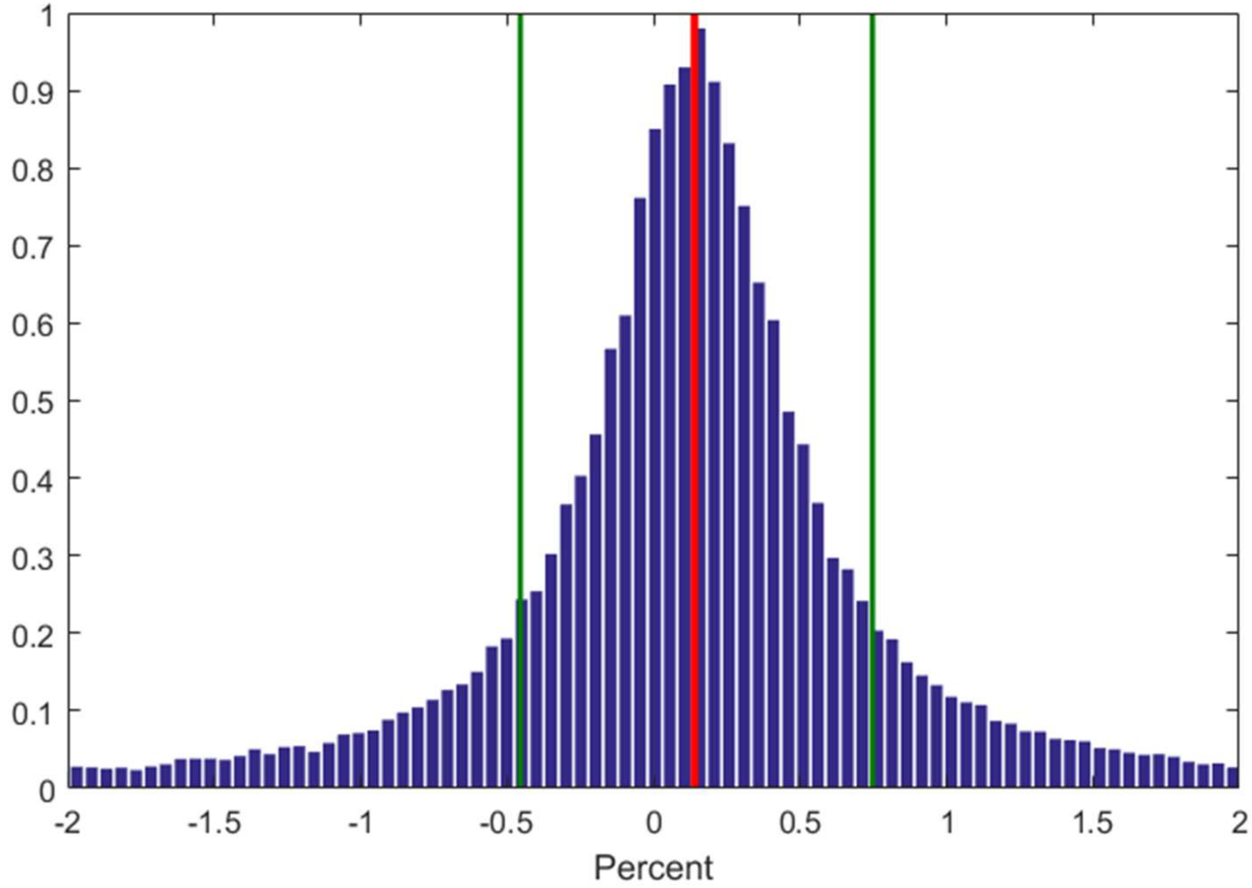
$$h_{21}/h_{11} = \frac{p_{21}x_{11} + p_{22}x_{21}}{p_{11}x_{11}} = \frac{p_{21}}{p_{11}} + \frac{p_{22}}{p_{11}}\frac{x_{21}}{x_{11}}$$

$$x_{ij} \sim N(0,1)$$

x_{21}/x_{11} ~ Cauchy(0,1)

h_{21}/h_{11} ~ Cauchy($p_{21}/p_{11}, p_{22}/p_{11}$)

Supposed effect at horizon zero on output gap of 0.25% monetary contraction



If we reported all the draws instead of 68% "error bands," answer would just be the real line.

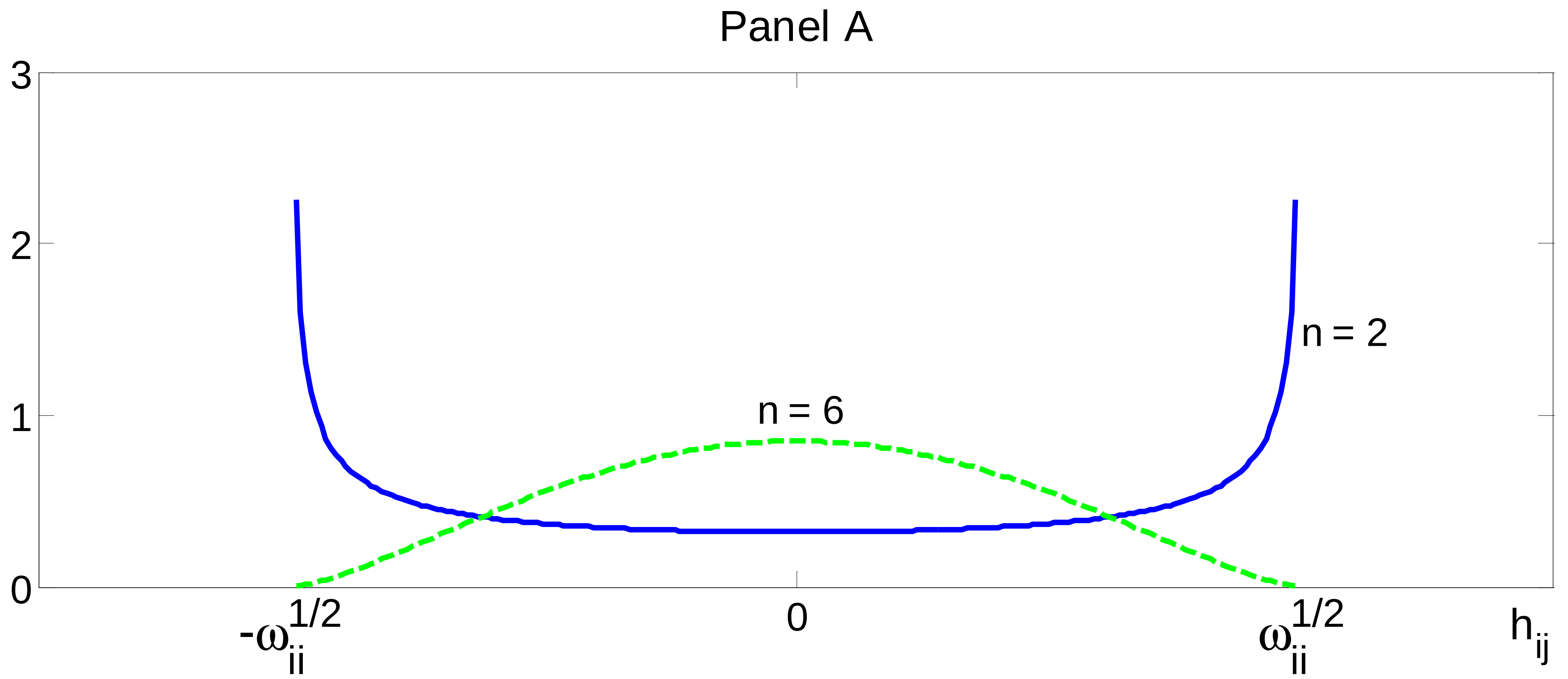
Implicit distribution has made it appear we learned more than we did.

What about distribution of individual elements h_{ij}?

$$h_{11} = p_{11} x_{11} / \sqrt{x_{11}^2 + x_{21}^2 + \cdots + x_{n1}^2}$$

$$p_{11} = \sqrt{\omega_{11}}$$

Analytic distribution of h_{ij}



Although the procedure implies a uniform distribution for the angle of rotation θ associated with the matrix $\mathbf{Q}$, we are not interested in inference about θ.

The algorithm implies a nonuniform distribution for structural impulse-response coefficients and this is what we are looking at with median and "error bands".

How do sign restrictions change any of this?

Δw_t = growth rate of real labor compensation

Δn_t = growth rate of total employment

$\mathbf{y}_t = (\Delta w_t, \Delta n_t)'$

demand: $\Delta n_t = k^d + \beta^d \Delta w_t + b_{11}^d \Delta w_{t-1} + b_{12}^d \Delta n_{t-1} + b_{21}^d \Delta w_{t-2}$

$$+ b_{22}^d \Delta n_{t-2} + \cdots + b_{m1}^d \Delta w_{t-m} + b_{m2}^d \Delta n_{t-m} + u_t^d$$

supply:

$$\Delta n_t = k^s + \alpha^s \Delta w_t + b_{11}^s \Delta w_{t-1} + b_{12}^s \Delta n_{t-1} + b_{21}^s \Delta w_{t-2}$$

$$+ b_{22}^s \Delta n_{t-2} + \cdots + b_{m1}^s \Delta w_{t-m} + b_{m2}^s \Delta n_{t-m} + u_t^s$$

sign restrictions: $\beta^d \leq 0,\ \alpha^s \geq 0$.

For fixed α^s, MLE of β^d can be found by an IV regression of $\hat{\varepsilon}_{2t}$ on $\hat{\varepsilon}_{1t}$ using $\hat{\varepsilon}_{2t} - \alpha\hat{\varepsilon}_{1t}$ as instrument:

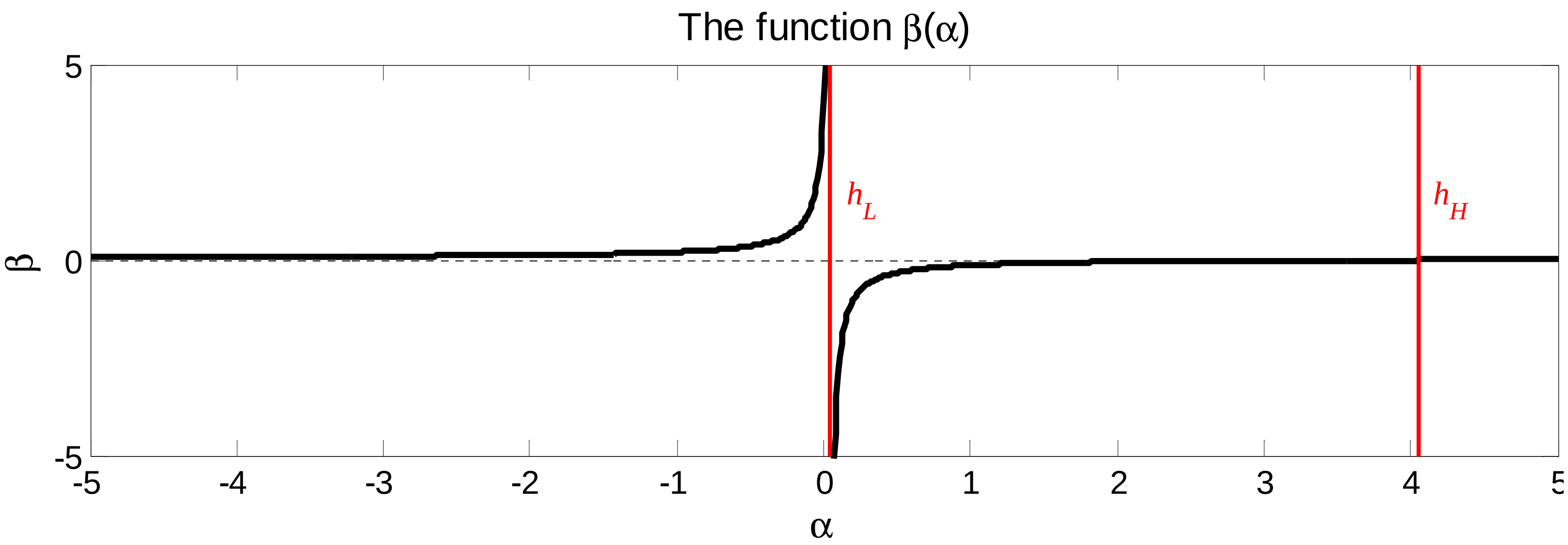
$$\hat{\beta}(\alpha) = \frac{\sum_{t=1}^{T}(\hat{\varepsilon}_{2t}-\alpha\hat{\varepsilon}_{1t})\hat{\varepsilon}_{2t}}{\sum_{t=1}^{T}(\hat{\varepsilon}_{2t}-\alpha\hat{\varepsilon}_{1t})\hat{\varepsilon}_{1t}} = \frac{(\hat{\omega}_{22}-\alpha\hat{\omega}_{12})}{(\hat{\omega}_{12}-\alpha\hat{\omega}_{11})}$$

$$\hat{\beta}(\alpha) = \frac{\sum_{t=1}^{T}(\hat{\varepsilon}_{2t}-\alpha\hat{\varepsilon}_{1t})\hat{\varepsilon}_{2t}}{\sum_{t=1}^{T}(\hat{\varepsilon}_{2t}-\alpha\hat{\varepsilon}_{1t})\hat{\varepsilon}_{1t}} = \frac{(\hat{\omega}_{22}-\alpha\hat{\omega}_{12})}{(\hat{\omega}_{12}-\alpha\hat{\omega}_{11})}$$

In the data, $\hat{\omega}_{12} > 0$.

At $\alpha = h_H = \hat{\omega}_{22}/\hat{\omega}_{12}$, numerator switches from positive to negative.

At $\alpha = h_L = \hat{\omega}_{12}/\hat{\omega}_{11}$, denominator switches from positive to negative.



$\alpha > 0$ and $\beta < 0$ restricts $h_L < \alpha < h_H$

but allows any $\beta < 0$.

Intuition: $h_L = \hat{\omega}_{12}/\hat{\omega}_{11}$ is coeff from OLS regression of $\hat{\varepsilon}_{2t}$ on $\hat{\varepsilon}_{1t}$

$=$ convex combination of α and β

$\Rightarrow \beta < h_L, \alpha > h_L$

since $h_L > 0,$ this restricts $\alpha,$ not β

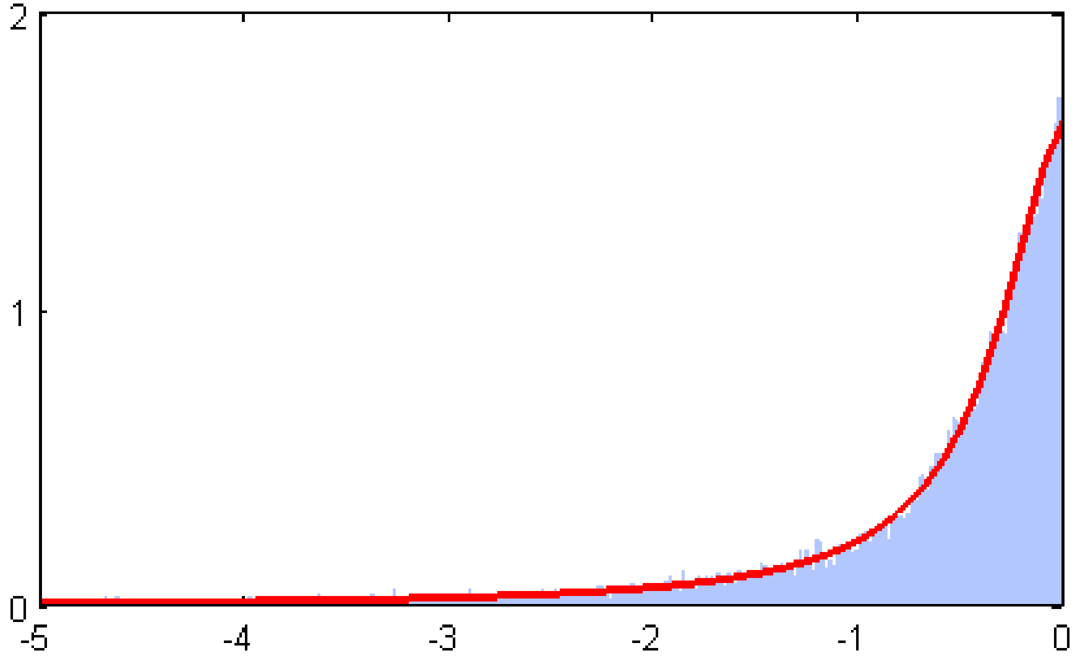
Intuition: $h_H^{-1} = \hat{\omega}_{12}/\hat{\omega}_{22}$ is coefficient from OLS regression of $\hat{\varepsilon}_{1t}$ on $\hat{\varepsilon}_{2t}$

$=$ convex combination of α^{-1} and β^{-1}

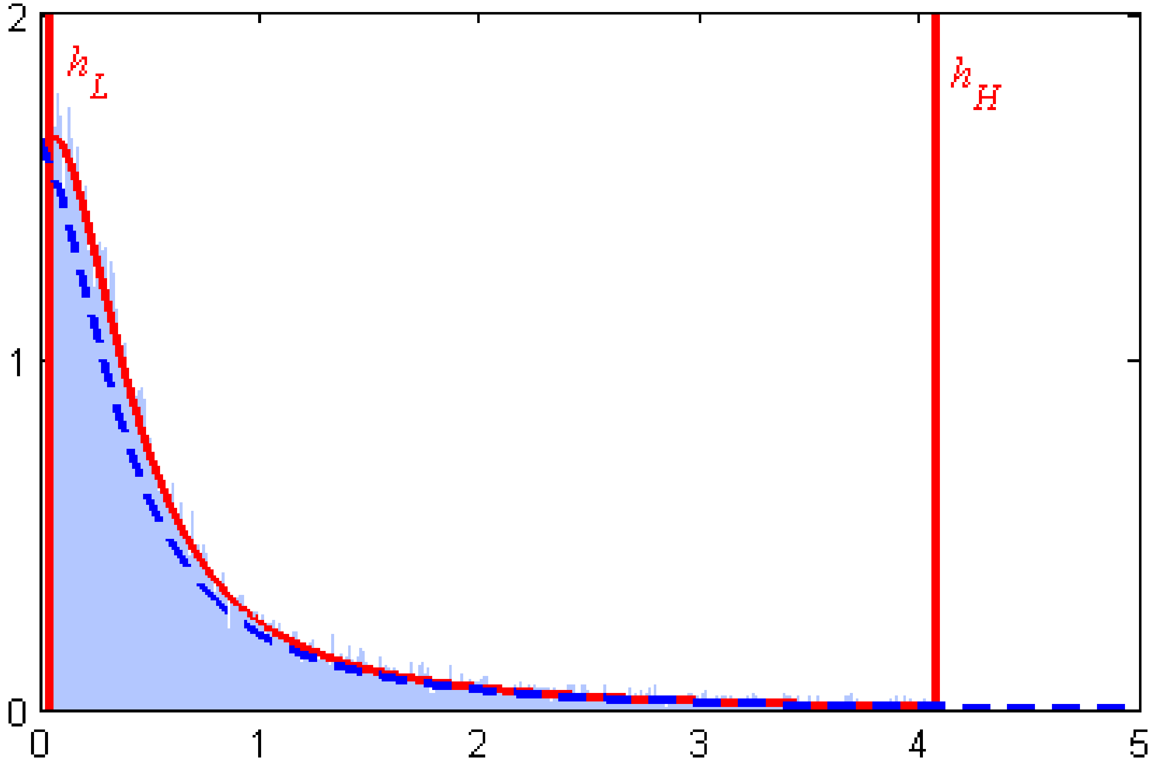
$\Rightarrow \beta^{-1} < h_H^{-1}, \alpha^{-1} > h_H^{-1}$

since $h_H > 0,$ this restricts $\alpha,$ not β

$\Rightarrow h_L < \alpha < h_H$



Distribution for draws of β when sign restrictions are imposed is Cauchy truncated to be negative.



Distribution for draws of α when sign restrictions are imposed is Cauchy truncated to be between h_L and h_H.