II. Vector autoregressions

D. Identification using inequality constraints

1. Traditional approach to identification

Dynamic structural model:

$$\underset{(n\times n)}{\mathbf{A}}\ \underset{(n\times 1)}{\mathbf{y}_t} = \underset{(n\times 1)}{\boldsymbol{\lambda}} + \underset{(n\times n)}{\mathbf{B}_1}\ \underset{(n\times 1)}{\mathbf{y}_{t-1}} +$$

$$\cdots + \mathbf{B}_m \mathbf{y}_{t-m} + \underset{(n\times n)}{\mathbf{D}^{1/2}}\ \underset{(n\times 1)}{\mathbf{v}_t}$$

$$= \underset{(n\times nm+1)}{\mathbf{B}}\ \underset{(nm+1\times 1)}{\mathbf{x}_{t-1}} + \underset{(n\times n)}{\mathbf{D}^{1/2}}\ \underset{(n\times 1)}{\mathbf{v}_t}$$

$$\mathbf{x}'_{t-1} = (1, \mathbf{y}'_{t-1}, \mathbf{y}'_{t-2}, \ldots, \mathbf{y}'_{t-m})'$$

$$\mathbf{v}_t \sim \text{i.i.d. } N(\mathbf{0}, \mathbf{I}_n)$$

$$\mathbf{D} = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix}$$

$$\mathbf{D}^{1/2} = \begin{bmatrix} \sqrt{d_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{d_{22}} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sqrt{d_{nn}} \end{bmatrix}$$

Example: demand and supply

$$
\begin{aligned}
q_t = \; & k^d + \beta^d p_t + b_{11}^d p_{t-1} + b_{12}^d q_{t-1} + b_{21}^d p_{t-2} \\
& + b_{22}^d q_{t-2} + \cdots + b_{m1}^d p_{t-m} + b_{m2}^d q_{t-m} + \sigma_d v_t^d
\end{aligned}
$$

$$
\begin{aligned}
q_t = \; & k^s + \alpha^s p_t + b_{11}^s p_{t-1} + b_{12}^s q_{t-1} + b_{21}^s p_{t-2} \\
& + b_{22}^s q_{t-2} + \cdots + b_{m1}^s p_{t-m} + b_{m2}^s q_{t-m} + \sigma_s v_t^s
\end{aligned}
$$

$$
\mathbf{y}_t = \begin{bmatrix} q_t \\ p_t \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} 1 & -\beta^d \\ 1 & -\alpha^s \end{bmatrix}
$$

Reduced-form (forecasting equations):

$$\mathbf{y}_t = \mathbf{\Phi}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t$$

$$\mathbf{x}'_{t-1} = (1, \mathbf{y}'_{t-1}, \mathbf{y}'_{t-2}, \ldots, \mathbf{y}'_{t-m})'$$

$\boldsymbol{\varepsilon}_t \sim$ i.i.d. $N(\mathbf{0}, \mathbf{\Omega})$

$$\hat{\mathbf{\Phi}} = \left(\sum_{t=1}^{T} \mathbf{y}_t \mathbf{x}'_{t-1}\right)\left(\sum_{t=1}^{T} \mathbf{x}_{t-1}\mathbf{x}'_{t-1}\right)^{-1}$$

$$\hat{\boldsymbol{\varepsilon}}_t = \mathbf{y}_t - \hat{\mathbf{\Phi}}\mathbf{x}_{t-1}$$

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

Nonorthogonalized impulse-response:

$$\frac{\partial \mathbf{y}_{t+s}}{\partial \boldsymbol{\varepsilon}'_t} = \underset{(n\times n)}{\mathbf{\Psi}_s}$$

Structural model:

$$\mathbf{A}\mathbf{y}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{D}^{1/2}\mathbf{v}_t \quad \mathbf{v}_t \sim N(\mathbf{0}, \mathbf{I}_n)$$

Structural impulse response:

$$\frac{\partial \mathbf{y}_{t+s}}{\partial \mathbf{v}_t'} = \frac{\partial \mathbf{y}_{t+s}}{\partial \boldsymbol{\varepsilon}_t'} \frac{\partial \boldsymbol{\varepsilon}_t}{\partial \mathbf{v}_t'} = \boldsymbol{\Psi}_s \mathbf{H}$$

$$\mathbf{H} = \frac{\partial \boldsymbol{\varepsilon}_t}{\partial \mathbf{v}_t'} = \mathbf{A}^{-1}\mathbf{D}^{1/2}$$

Reduced form:

$$\mathbf{y}_t = \boldsymbol{\Phi}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t \quad \boldsymbol{\varepsilon}_t \sim \text{i.i.d. } N(\mathbf{0}, \boldsymbol{\Omega})$$

$$\boldsymbol{\Phi} = \mathbf{A}^{-1}\mathbf{B}$$

$$\boldsymbol{\varepsilon}_t = \mathbf{A}^{-1}\mathbf{D}^{1/2}\mathbf{v}_t = \mathbf{H}\mathbf{v}_t$$

$$E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') = \boldsymbol{\Omega} = \mathbf{A}^{-1}\mathbf{D}(\mathbf{A}^{-1})'$$

$$\boldsymbol{\Omega} = \mathbf{A}^{-1}\mathbf{D}(\mathbf{A}^{-1})'$$

Can estimate the 3 parameters in $\boldsymbol{\Omega}$ by OLS.

But there are 4 unknown elements in $\mathbf{A}$ and **D** $(\beta^d, \alpha^s, \sigma_d, \sigma_s)$

Traditional approach to identification:
Put enough restrictions on $\mathbf{A}$ and $\mathbf{D}$ so that for any $\boldsymbol{\Omega}$ there is a unique $\mathbf{A}, \mathbf{D}$ for which $\boldsymbol{\Omega} = \mathbf{A}^{-1}\mathbf{D}(\mathbf{A}^{-1})'$

Example: assume short-run demand elasticity $\beta^d = 0$ (meaning $\mathbf{A}$ and $\mathbf{A}^{-1}$ are lower triangular):

$$\mathbf{A} = \begin{bmatrix} 1 & -\beta^d \\ 1 & -\alpha^s \end{bmatrix}$$

If $\mathbf{D}$ is diagonal, then $\mathbf{H}$ also lower triangular:

$$\mathbf{H} = \frac{\partial \mathbf{y}_t}{\partial \mathbf{v}_t'} = \mathbf{A}^{-1}\mathbf{D}^{1/2}$$

If normalize shocks so that v_t^d raises q_t and v_t^s raises $p_t,$ there is unique lower-triangular matrix $\mathbf{H}$ with positive diagonal elements such that $\mathbf{H}\mathbf{H}' = \mathbf{\Omega}$ namely, $\mathbf{H} =$ Cholesky factor of $\mathbf{\Omega}$

Estimate following from reduced form:

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

$\hat{\mathbf{P}}\hat{\mathbf{P}}' = \hat{\boldsymbol{\Omega}}$ (Cholesky factor)

$$\hat{\boldsymbol{\Psi}}_s = \frac{\partial \mathbf{y}_{t+s}}{\partial \boldsymbol{\varepsilon}_t'}$$

$$\hat{\mathbf{H}} = \hat{\mathbf{P}}$$

Then we infer dynamic structural responses:

$$\frac{\partial \mathbf{y}_{t+s}}{\partial \mathbf{v}_t'} = \frac{\partial \mathbf{y}_{t+s}}{\partial \boldsymbol{\varepsilon}_t'} \frac{\partial \boldsymbol{\varepsilon}_t}{\partial \mathbf{v}_t'} = \hat{\boldsymbol{\Psi}}_s \hat{\mathbf{P}}$$

II. Vector autoregressions

D. Identification using inequality constraints

1. Traditional approach to identification

2. Arias, Rubio-Ramirez, and Waggoner approach to partial identification

Obviously assumption that short-run demand elasticity = 0 is very strong.

Can we make inference using weaker assumptions?

We may have confidence in signs:

$$\mathbf{H} = \begin{bmatrix} \partial q_t/\partial v_t^d & \partial q_t/\partial v_t^s \\ \partial p_t/\partial v_t^d & \partial p_t/\partial v_t^s \end{bmatrix} = \begin{bmatrix} + & - \\ + & + \end{bmatrix}$$

Can we use a prior for $\mathbf{H}|\mathbf{\Omega}$ that implies these signs but is otherwise uninformative?

Proposal: $\mathbf{H} = \mathbf{PQ}$ where $\mathbf{\Omega} = \mathbf{PP}'$ and $\mathbf{Q}$ is drawn from a Haar-uniform distribution from the set of all orthogonal matrices ($\mathbf{QQ}' = \mathbf{I}_n$)

How generate a draw for $\mathbf{Q}$?

(1) Generate $(n \times n)$ $\mathbf{X} = [x_{ij}]$ of $N(0, 1)$.

(2) Find $\mathbf{X} = \mathbf{QR}$ for $\mathbf{Q}$ orthogonal and $\mathbf{R}$ upper triangular.

First column of $\mathbf{Q}$ = first column of $\mathbf{X}$ normalized to have unit length:

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

E.g., if $n = 2$, $q_{11} = \cos\theta$ for θ the angle between (x_{11}, x_{21}) and $(1, 0)$ while $q_{21} = \sin\theta$.

$$\mathbf{Q} = \begin{cases} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} & \text{with prob 1/2} \\ \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} & \text{with prob 1/2} \end{cases}$$

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

Algorithm to generate draw subject to sign restrictions:

(1) Generate $\boldsymbol{\Omega}^{-1} \sim W(T, T\hat{\boldsymbol{\Omega}})$

(2) Calculate $\mathbf{P}\mathbf{P}' = \boldsymbol{\Omega}$

(3) Generate orthogonal $\mathbf{Q}$

(4) Calculate $\mathbf{H} = \mathbf{P}\mathbf{Q}$

(5) Keep if satisfies restrictions, otherwise throw out

Issue 1: A prior that is uninformative about a parameter (in this case, the angle of rotation θ) is in general informative about nonlinear transformations of θ.

Baumeister and Hamilton (2014) calculate implicit priors for other objects.

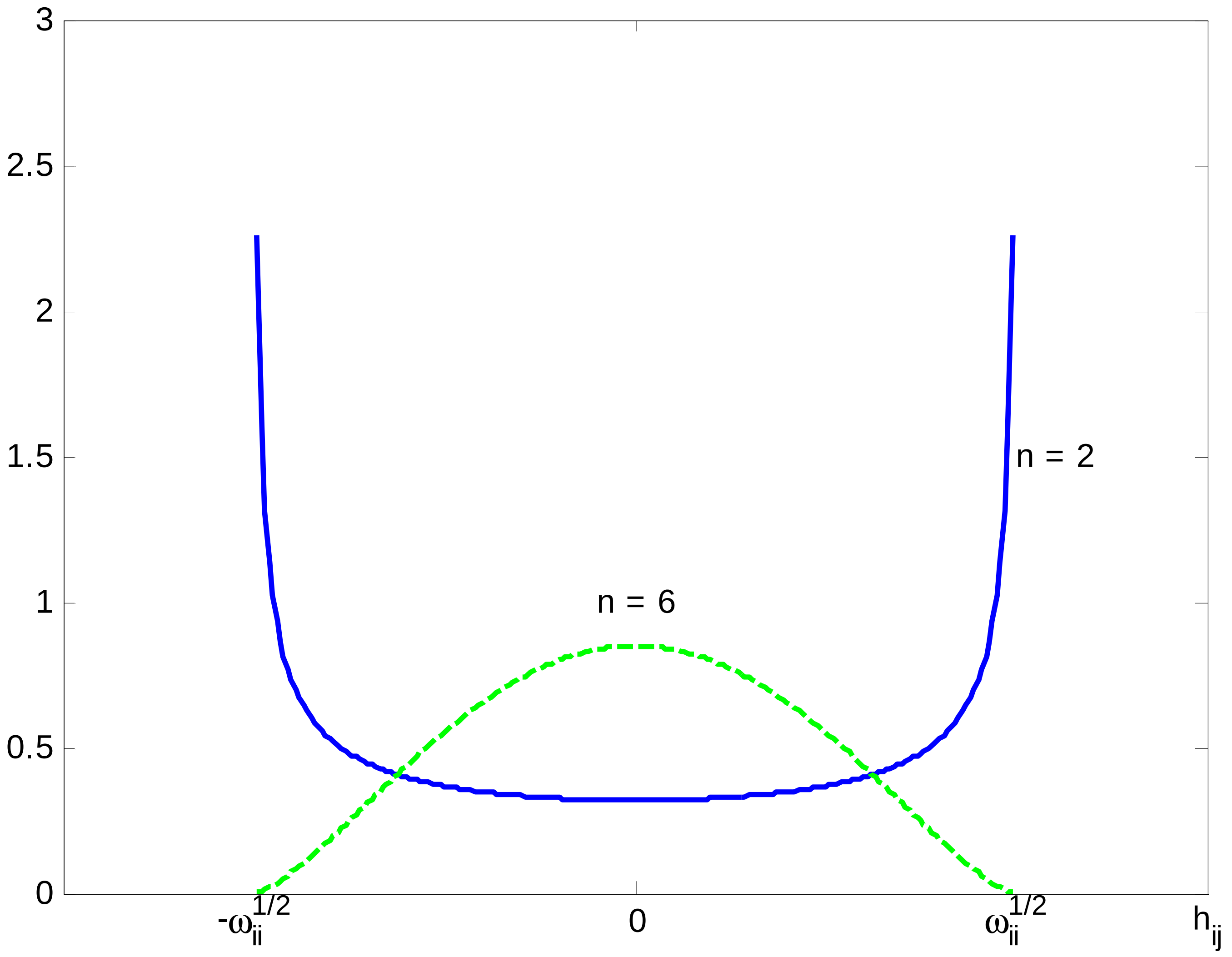
$$q_{i1} = x_{i1} / \sqrt{x_{11}^2 + \cdots + x_{n1}^2}$$

$$\Rightarrow q_{i1}^2 \sim \text{Beta}(1/2, (n-1)/2)$$

$$p(q_{i1}) = \begin{cases} \frac{\Gamma(n/2)}{\Gamma(1/2)\Gamma((n-1)/2)}(1 - q_{i1}^2)^{(n-3)/2} & \text{if } q_{i1} \in [-1, 1] \\ 0 & \text{otherwise} \end{cases}$$

$$h_{11} = p_{11} q_{11} = \sqrt{\omega_{11}}\, q_{11}$$

Effect of one-standard deviation shock on variable i



Alternatively, we might want to normalize shock 1 as something that raises variable 1 by 1 unit:

$$h_{21}^{*} = \frac{h_{21}}{h_{11}} = \frac{p_{21}q_{11}+p_{22}q_{21}}{p_{11}q_{11}} = \frac{p_{21}}{p_{11}} + \frac{p_{22}}{p_{11}}\frac{x_{21}}{x_{11}}$$

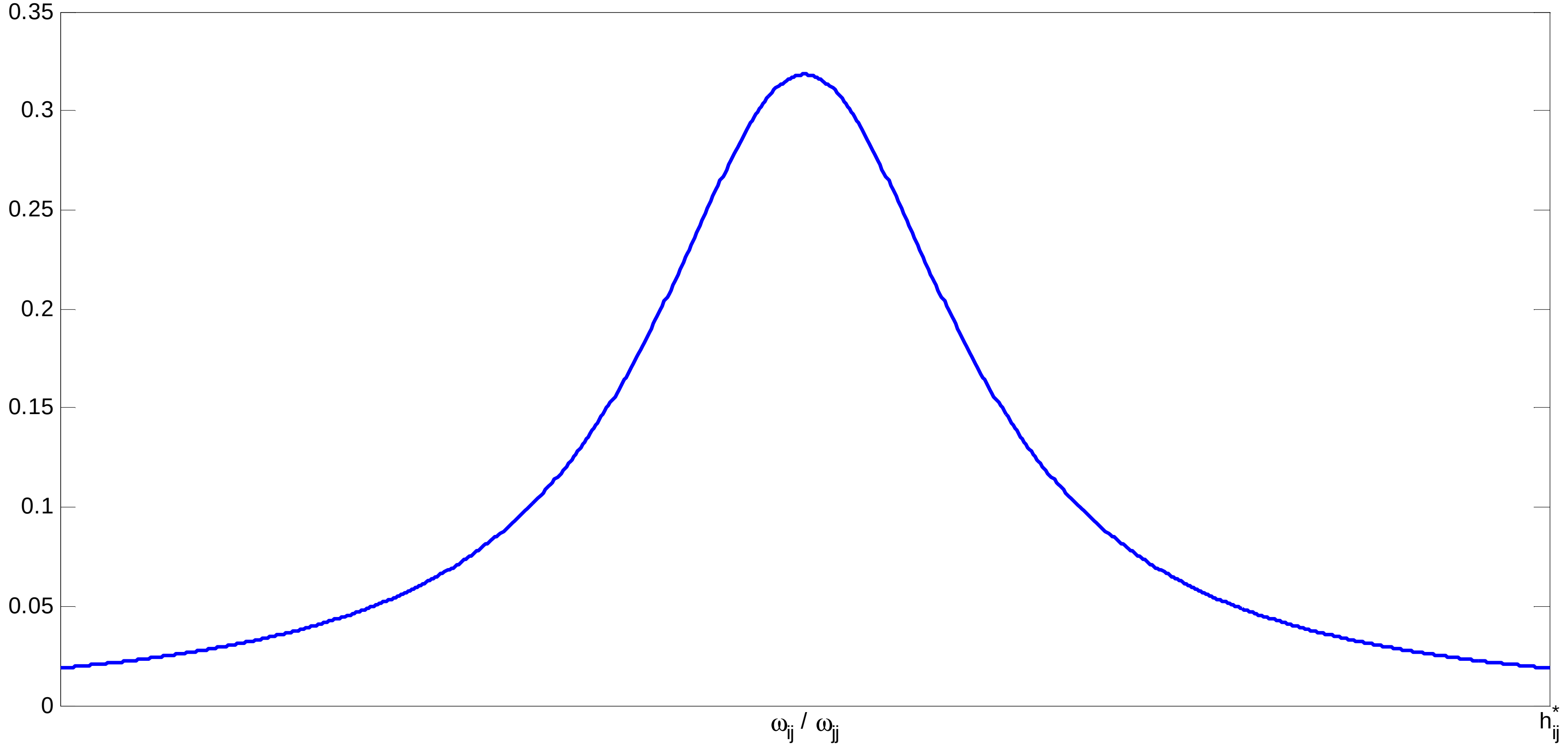
x_{21}/x_{11} ~ Cauchy(0,1)

$\Rightarrow h_{ij}^{*}|\mathbf{\Omega}$ ~ Cauchy($c_{ij}^{*}, \sigma_{ij}^{*}$)

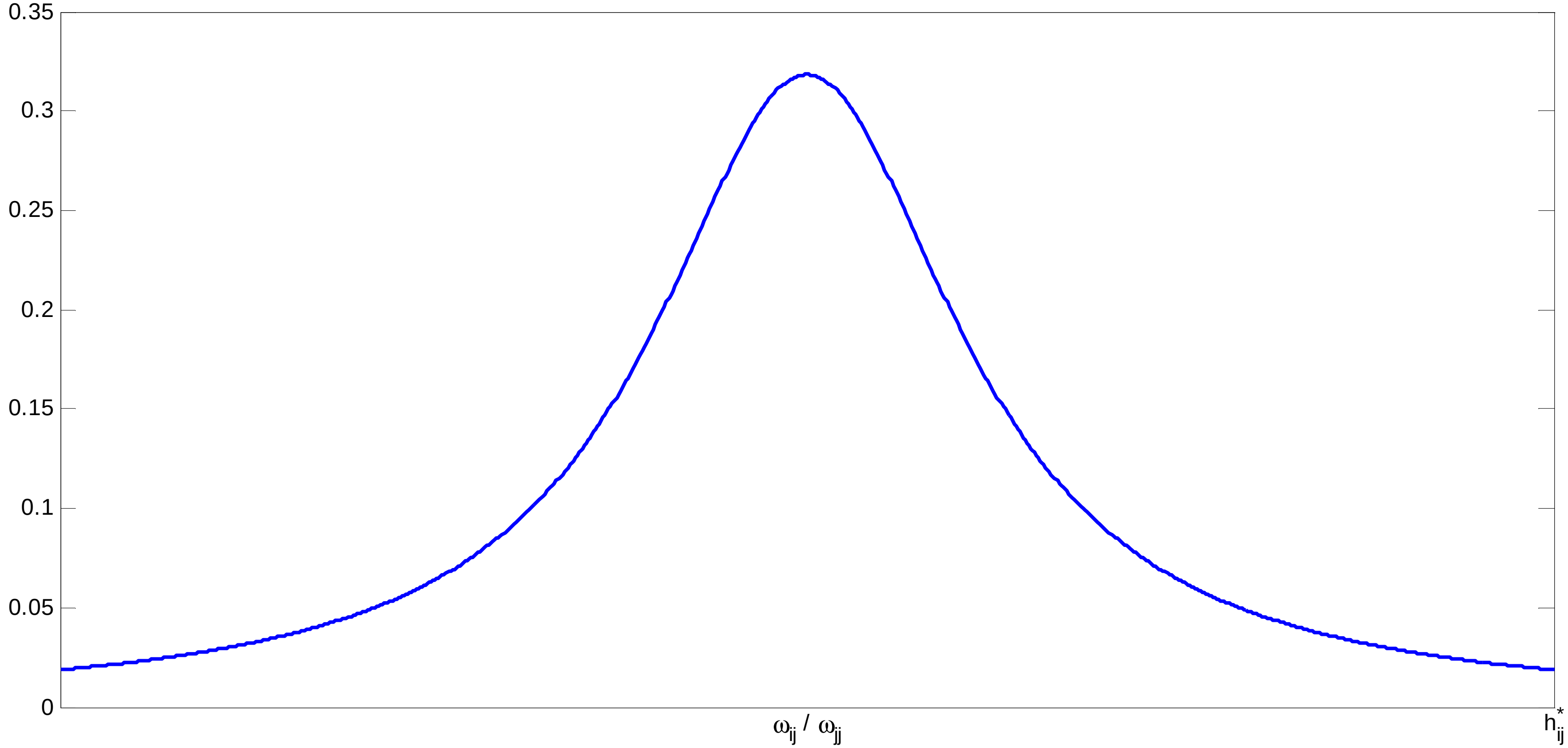
$$c_{ij}^{*} = \omega_{ij}/\omega_{jj}$$

$$\sigma_{ij}^{*} = \sqrt{\frac{\omega_{ii}-\omega_{ij}^{2}/\omega_{jj}}{\omega_{jj}}}$$

Effect on variable i of shock that increases j by one unit



Effect on variable i of shock that increases j by one unit



Sign restrictions confine these distributions to particular regions but do not change their basic features.

Issue 2: the sign restrictions may end implying zero or trivial restrictions on the feasible set.

$$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} p_{11}\cos\theta & p_{11}\sin\theta \\ (p_{21}\cos\theta + p_{22}\sin\theta) & (p_{21}\sin\theta - p_{22}\cos\theta) \end{bmatrix}$$

variable 1 = price, variable 2 = quantity

shock 1 = demand, 2 = supply

$$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} + & + \\ + & - \end{bmatrix}$$

$$h_{11}, h_{12} \geq 0 \Rightarrow \theta \in [0, \pi/2]$$

$$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} p_{11}\cos\theta & p_{11}\sin\theta \\ (p_{21}\cos\theta + p_{22}\sin\theta) & (p_{21}\sin\theta - p_{22}\cos\theta) \end{bmatrix}$$

If $p_{21} > 0$, then $h_{21} \geq 0$ for all $\theta \in [0, \pi/2]$

But $h_{22} \leq 0 \Rightarrow \theta \in [0, \tilde{\theta}]$ for $\cot\tilde{\theta} = p_{21}/p_{22}$

And $\theta \in [0, \tilde{\theta}] \Rightarrow$

$$h_{22}^{*} = \frac{p_{21}}{p_{11}} - \frac{p_{22}}{p_{11}}\cot\theta \in (-\infty, 0]$$

$$h_{21}^{*} = \frac{p_{21}}{p_{11}} + \frac{p_{22}}{p_{11}}\tan\theta \in [\omega_{21}/\omega_{11}, \omega_{22}/\omega_{21}]$$

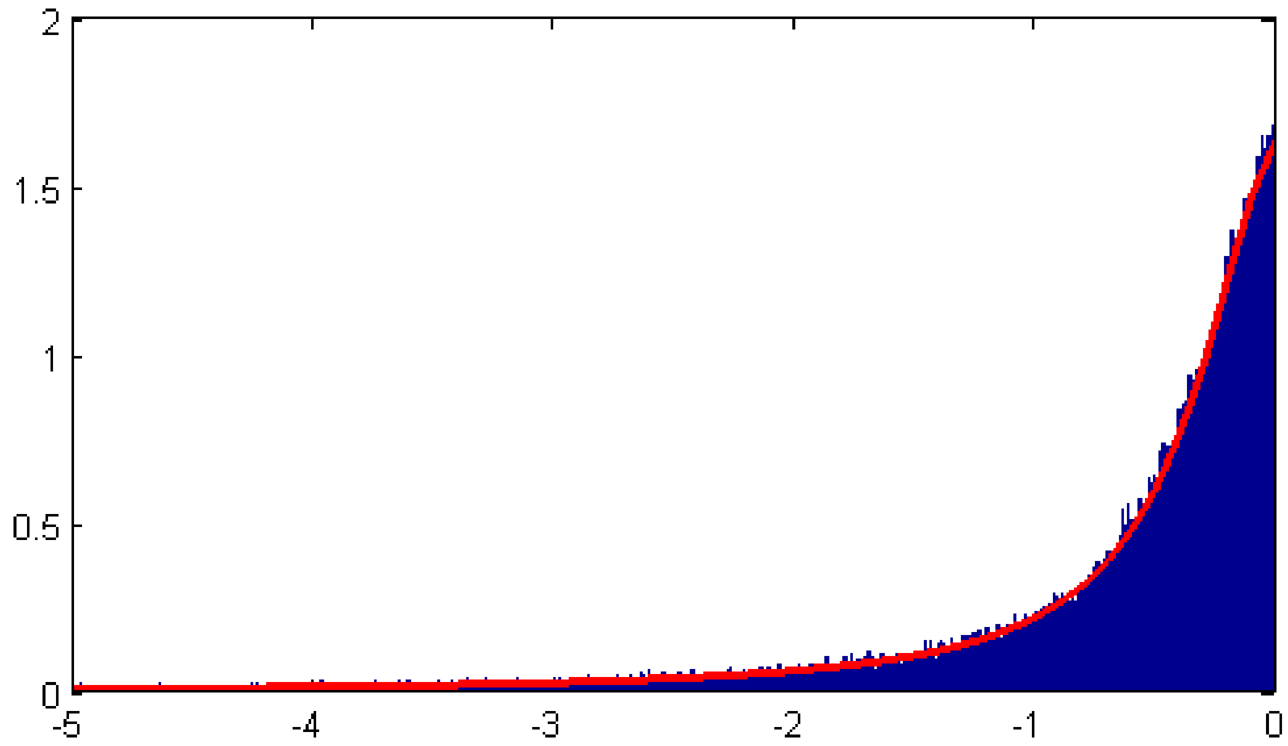
Example: 8-lag VAR fit to growth rates of U.S. real compensation per worker and employment, 1970:Q1-2014:Q2.

$$\hat{\mathbf{\Omega}} = \begin{bmatrix} 0.5920 & 0.0250 \\ 0.0250 & 0.1014 \end{bmatrix}$$

short-run demand elasticity unrestricted

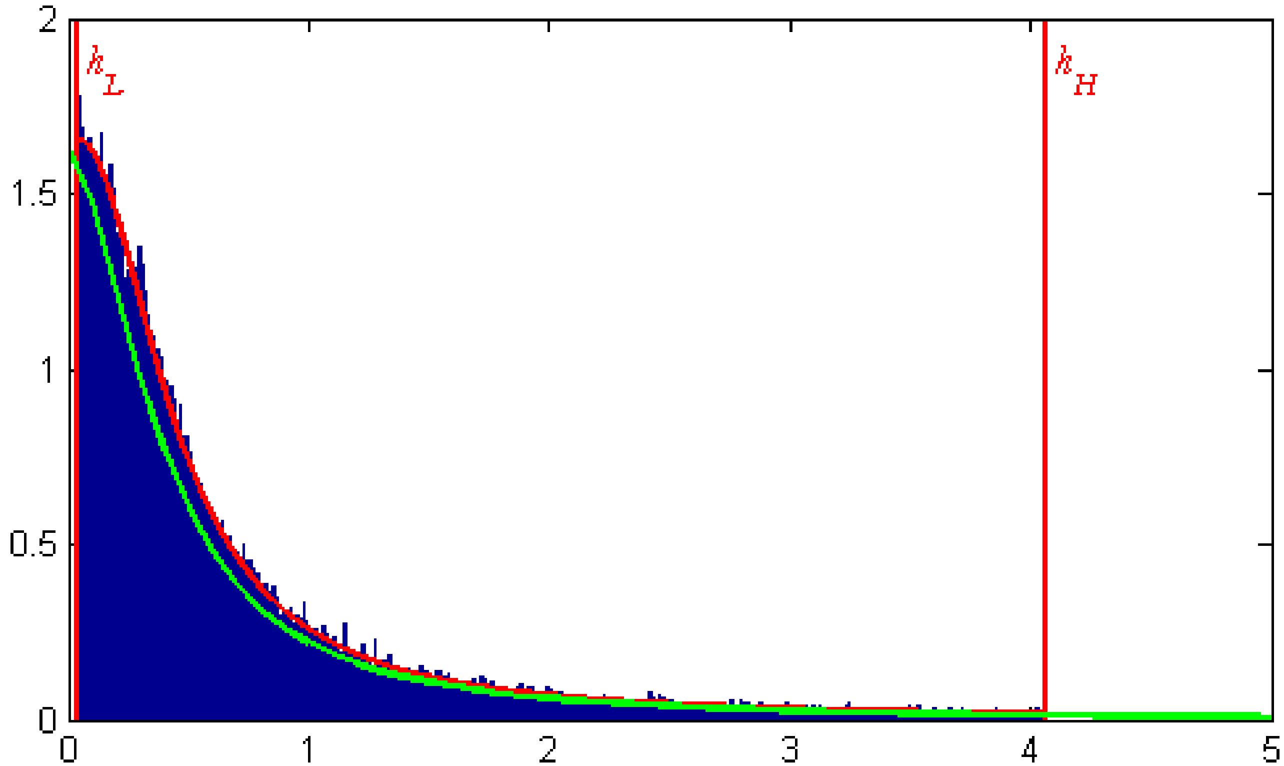
short-run supply elasticity $\in [0.0421, 4.0626]$.

Implied elasticity of labor demand (= h22*)



Red = truncated Cauchy, blue = output of traditional algorithm

Implied elasticity of labor supply (= h21*)



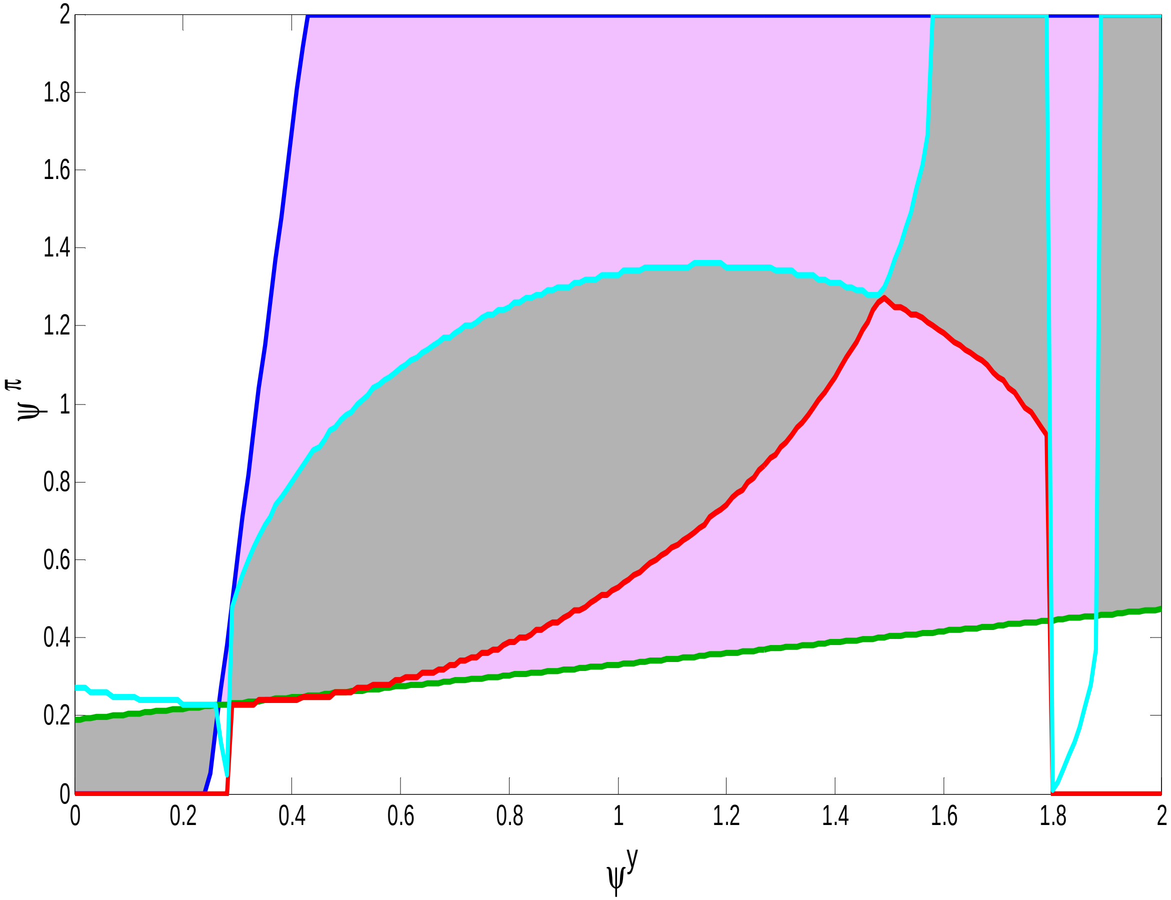
Red = truncated Cauchy, blue = output of traditional algorithm

Issue 3: the sign restrictions may end implying a bizarre topology for set of parameter values of interest.

Example: VAR to U.S. output gap, inflation rate and fed funds rate, 1986:Q1 - 2008:Q4.

- aggregate supply slopes up
- Taylor Rule coefficients positive
- inflation raises aggregate demand

Gray: $\alpha > 0$ and $\beta < 0$; Purple: $\alpha > 0$ and $\beta > 0$



II. Vector autoregressions

D. Identification using inequality constraints

1. Traditional approach to identification

2. Arias, Rubio-Ramirez, and Waggoner approach to partial identification

3. Baumeister and Hamilton approach to proper Bayesian inference with partially identified models

Structural model:

$$\mathbf{A}\mathbf{y}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{u}_t \qquad \mathbf{u}_t \sim N(\mathbf{0}, \mathbf{D})$$

Reduced form:

$$\mathbf{y}_t = \mathbf{\Phi}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t \qquad \boldsymbol{\varepsilon}_t \sim \text{i.i.d. } N(\mathbf{0}, \mathbf{\Omega})$$

$$\boldsymbol{\varepsilon}_t = \mathbf{A}^{-1}\mathbf{u}_t$$

$$E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') = \mathbf{\Omega} \qquad \mathbf{A}\mathbf{\Omega}\mathbf{A}' = \mathbf{D}$$

Structural model:

$$\mathbf{A}\mathbf{y}_t = \boldsymbol{\lambda} + \mathbf{B}_1\mathbf{y}_{t-1} + \cdots + \mathbf{B}_m\mathbf{y}_{t-m} + \mathbf{u}_t$$

$\mathbf{u}_t \sim$ i.i.d. $N(\mathbf{0}, \mathbf{D})$ $\quad$ $\mathbf{D}$ diagonal

Intuition for results that follow:

If we knew row i of $\mathbf{A}$ (denoted $\mathbf{a}_i'$),

then we could estimate coefficients for

ith structural equation ($\mathbf{b}_i$) by

$$\hat{\mathbf{b}}_i = \left(\sum_{t=1}^{T} \mathbf{x}_{t-1}\mathbf{x}_{t-1}'\right)^{-1}\left(\sum_{t=1}^{T} \mathbf{x}_{t-1}\mathbf{y}_t'\mathbf{a}_i\right) = \hat{\boldsymbol{\Phi}}_T'\mathbf{a}_i$$

$$\hat{d}_{ii} = T^{-1}\sum_{t=1}^{T} \hat{u}_t^2 = \mathbf{a}_i'\hat{\boldsymbol{\Omega}}_T\mathbf{a}_i \qquad \hat{\mathbf{D}} = \text{diag}(\mathbf{A}\hat{\boldsymbol{\Omega}}_T\mathbf{A}')$$

Consider Bayesian approach where we begin with arbitrary prior $p(\mathbf{A})$

E.g., prior beliefs about supply and demand elasticities in the form of joint density $p(\alpha^s, \beta^d)$

$$\mathbf{A} = \begin{bmatrix} -\beta^d & 1 \\ -\alpha^s & 1 \end{bmatrix}$$

$p(\mathbf{A})$ could also impose sign restrictions, zeros, or assign small but nonzero probabilities to violations of these constraints.

Will use natural conjugate priors

for other parameters:

$$p(\mathbf{D}|\mathbf{A}) = \prod_{i=1}^{n} p(d_{ii}|\mathbf{A})$$

$$d_{ii}^{-1}|\mathbf{A} \sim \Gamma(\kappa_i, \tau_i)$$

$$E(d_{ii}^{-1}|\mathbf{A}) = \kappa_i/\tau_i$$

$$Var(d_{ii}^{-1}|\mathbf{A}) = \kappa_i/\tau_i^2$$

uninformative priors: $\kappa_i, \tau_i \to 0$

$$\mathbf{B} = \begin{bmatrix} \boldsymbol{\lambda} & \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_m \end{bmatrix}$$

$$p(\mathbf{B}|\mathbf{D},\mathbf{A}) = \prod_{i=1}^{n} p(\mathbf{b}_i|\mathbf{D},\mathbf{A})$$

$$\mathbf{b}_i|\mathbf{A},\mathbf{D} \sim N(\mathbf{m}_i, d_{ii}\mathbf{M}_i)$$

uninformative priors: $\mathbf{M}_i^{-1} \rightarrow \mathbf{0}$

Likelihood:

$$p(\mathbf{Y}_T|\mathbf{A},\mathbf{D},\mathbf{B}) = (2\pi)^{-Tn/2}|\det(\mathbf{A})|^T|\mathbf{D}|^{-T/2} \times$$
$$\exp\left[-(1/2)\sum_{t=1}^{T}(\mathbf{A}\mathbf{y}_t - \mathbf{B}\mathbf{x}_{t-1})'\mathbf{D}^{-1}(\mathbf{A}\mathbf{y}_t - \mathbf{B}\mathbf{x}_{t-1})\right]$$

prior:

$$p(\mathbf{A},\mathbf{D},\mathbf{B}) = p(\mathbf{A})p(\mathbf{D}|\mathbf{A})p(\mathbf{B}|\mathbf{A},\mathbf{D})$$

posterior:

$$p(\mathbf{A},\mathbf{D},\mathbf{B}|\mathbf{Y}_T) = \frac{p(\mathbf{Y}_T|\mathbf{A},\mathbf{D},\mathbf{B})p(\mathbf{A},\mathbf{D},\mathbf{B})}{\int p(\mathbf{Y}_T|\mathbf{A},\mathbf{D},\mathbf{B})p(\mathbf{A},\mathbf{D},\mathbf{B})d\mathbf{A}d\mathbf{D}d\mathbf{B}}$$

$$= p(\mathbf{A}|\mathbf{Y}_T)p(\mathbf{D}|\mathbf{A},\mathbf{Y}_T)p(\mathbf{B}|\mathbf{A},\mathbf{D},\mathbf{Y}_T)$$

Exact Bayesian posterior distribution (all T):

$$\mathbf{b}_i | \mathbf{A}, \mathbf{D}, \mathbf{Y}_T \sim N(\mathbf{m}_i^*, d_{ii}\mathbf{M}_i^*)$$

$$\underset{[1\times(T+k)]}{\tilde{\mathbf{Y}}_i'} = (\mathbf{a}_i'\mathbf{y}_1, \ldots, \mathbf{a}_i'\mathbf{y}_T, \mathbf{m}_i'\mathbf{P}_i)$$

$$\underset{[k\times(T+k)]}{\tilde{\mathbf{X}}_i'} = \begin{bmatrix} \mathbf{x}_0 & \cdots & \mathbf{x}_{T-1} & \mathbf{P}_i \end{bmatrix}$$

$$\mathbf{m}_i^* = \left(\tilde{\mathbf{X}}_i'\tilde{\mathbf{X}}_i\right)^{-1}\left(\tilde{\mathbf{X}}_i'\tilde{\mathbf{y}}_i\right)$$

$$\mathbf{M}_i^* = \left(\tilde{\mathbf{X}}_i'\tilde{\mathbf{X}}_i\right)^{-1} \qquad \mathbf{P}_i\mathbf{P}_i' = \mathbf{M}_i^{-1}$$

If uninformative prior ($\mathbf{M}_i^{-1} = \mathbf{0}$)

then $\mathbf{m}_i^{*\prime} = \mathbf{a}_i'\hat{\mathbf{\Phi}}_T$

Frequentist interpretation of Bayesian posterior distribution as $T \to \infty$:

If prior on $\mathbf{B}$ is not dogmatic (that is, if $\mathbf{M}_i^{-1}$ is finite), then

$$\mathbf{m}_i^* \xrightarrow{p} [E(\mathbf{x}_{t-1}\mathbf{x}_{t-1}')]^{-1} E(\mathbf{x}_{t-1}\mathbf{y}_t')\mathbf{a}_i = \mathbf{\Phi}_0'\mathbf{a}_i$$

$$\mathbf{M}_i^* \xrightarrow{p} \mathbf{0}$$

$$\mathbf{b}_i | \mathbf{A}, \mathbf{D}, \mathbf{Y}_T \xrightarrow{p} \mathbf{\Phi}_0'\mathbf{a}_i$$

Posterior distribution for **D|A**

$$d_{ii}^{-1}|\mathbf{A},\mathbf{Y}_T \sim \Gamma(\kappa_i + (T/2), \tau_i + (\zeta_i^*/2))$$

$$\zeta_i^* = \left(\tilde{\mathbf{Y}}_i'\tilde{\mathbf{Y}}_i\right) - \left(\tilde{\mathbf{Y}}_i'\tilde{\mathbf{X}}_i\right)\left(\tilde{\mathbf{X}}_i'\tilde{\mathbf{X}}_i\right)^{-1}\left(\tilde{\mathbf{X}}_i'\tilde{\mathbf{Y}}_i\right)$$

If $\mathbf{M}_i^{-1} = \mathbf{0}$, $\zeta_i^* = T\mathbf{a}_i'\hat{\mathbf{\Omega}}_T\mathbf{a}_i$

$$\hat{\mathbf{\Omega}}_T = T^{-1}\sum_{t=1}^{T}\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t, \quad \hat{\boldsymbol{\varepsilon}}_t = \mathbf{y}_t - \hat{\mathbf{\Phi}}\mathbf{x}_{t-1}$$

($\hat{\boldsymbol{\varepsilon}}_t$ are unrestricted OLS residuals)

If priors on $\mathbf{B}$ and $\mathbf{D}$ are not dogmatic

(that is, if $\mathbf{M}_i^{-1}, \kappa_i, \tau_i$ are all finite) then

$$\zeta_i^*/T \xrightarrow{p} \mathbf{a}_i'\mathbf{\Omega}_0\mathbf{a}_i$$

$$\mathbf{\Omega}_0 = E(\mathbf{y}_t\mathbf{x}_{t-1}') - E(\mathbf{y}_t\mathbf{x}_{t-1}')\{E(\mathbf{x}_t\mathbf{x}_t')\}^{-1}E(\mathbf{x}_{t-1}\mathbf{y}_t')$$

$$d_{ii}|\mathbf{A}, \mathbf{Y}_T \xrightarrow{p} \mathbf{a}_i'\mathbf{\Omega}_0\mathbf{a}_i$$

Posterior distribution for **A**

$$p(\mathbf{A}|\mathbf{Y}_T) = \frac{k_T p(\mathbf{A})[\det(\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}')]^{T/2}}{\prod_{i=1}^{n}[(2\tau_i/T)+(\zeta_i^*/T)]^{\kappa_i+T/2}}$$

$k_T =$ constant that makes this integrate to 1

$p(\mathbf{A}) =$ prior

If $\mathbf{M}_i^{-1} = \mathbf{0}$, and $\tau_i = \kappa_i = 0$,

$$p(\mathbf{A}|\mathbf{Y}_T) = \frac{k_T p(\mathbf{A})|\det(\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}')|^{T/2}}{\left\{\det\left[\text{diag}(\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}')\right]\right\}^{T/2}}$$

$$p(\mathbf{A}|\mathbf{Y}_T) = \frac{k_T p(\mathbf{A})|\det(\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}')|^{T/2}}{\left\{\det\left[\text{diag}(\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}')\right]\right\}^{T/2}}$$

If evaluated at $\mathbf{A}$ for which

$\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}' = \text{diag}(\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}')$,

$$p(\mathbf{A}|\mathbf{Y}_T) = k_T p(\mathbf{A})$$

$$p(\mathbf{A}|\mathbf{Y}_T) = \frac{k_T p(\mathbf{A})|\det(\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}')|^{T/2}}{\left\{\det\left[\text{diag}(\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}')\right]\right\}^{T/2}}$$

Hadamard's Inequality:

If evaluated at $\mathbf{A}$ for which

$$\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}' \neq \text{diag}(\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}'),$$

$$\det\left[\text{diag}(\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}')\right] > \det(\mathbf{A}\hat{\mathbf{\Omega}}_T\mathbf{A}')$$

$$p(\mathbf{A}|\mathbf{Y}_T) \to 0$$

$$p(\mathbf{A}|\mathbf{Y}_T) \rightarrow \begin{cases} kp(\mathbf{A}) & \text{if } \mathbf{A} \in S(\mathbf{\Omega}_0) \\ 0 & \text{otherwise} \end{cases}$$

$$S(\mathbf{\Omega}_0) = \left\{\mathbf{A}\text{: } \mathbf{A}\mathbf{\Omega}_0\mathbf{A}' \text{ diagonal}\right\}$$

$$\mathbf{\Omega}_0 = E(\mathbf{y}_t\mathbf{x}'_{t-1}) - E(\mathbf{y}_t\mathbf{x}'_{t-1})\{E(\mathbf{x}_t\mathbf{x}'_t)\}^{-1}E(\mathbf{x}_{t-1}\mathbf{y}'_t)$$

Special case: if model is point-identified (so that $S(\mathbf{\Omega})$ consists of a single point), then posterior distribution converges to a point mass at true $\mathbf{A}$

Measure distance $q(\mathbf{A}, \mathbf{\Omega})$ between $\mathbf{A}$ and $S(\mathbf{\Omega})$ by sum of squares of off-diagonal elements of Cholesky factor of $\mathbf{A\Omega A}'$:

$$q(\mathbf{A}, \mathbf{\Omega}) = \sum_{i=2}^{n} \sum_{j=1}^{i-1} p_{ij}^2(\mathbf{A}, \mathbf{\Omega})$$

$$\mathbf{P}(\mathbf{A}, \mathbf{\Omega})[\mathbf{P}(\mathbf{A}, \mathbf{\Omega})]' = \mathbf{A\Omega A}'$$

$q(\mathbf{A}, \mathbf{\Omega}) = 0$ if and only if $\mathbf{A} \in S(\mathbf{\Omega})$

$$H_\delta(\mathbf{\Omega}) = \{\mathbf{A} : q(\mathbf{A}, \mathbf{\Omega}) \leq \delta\}$$

If $p(\mathbf{A})$ bounded and

$\int_{\mathbf{A} \in H_\delta(\mathbf{\Omega}_0)} p(\mathbf{A}) d\mathbf{A} > 0$ for all $\delta > 0$,

then $\text{Prob}[\mathbf{A} \in H_\delta(\mathbf{\Omega}_0) | \mathbf{Y}_T] \to 1$ for all $\delta > 0$.

Application: Labor market dynamics

demand:

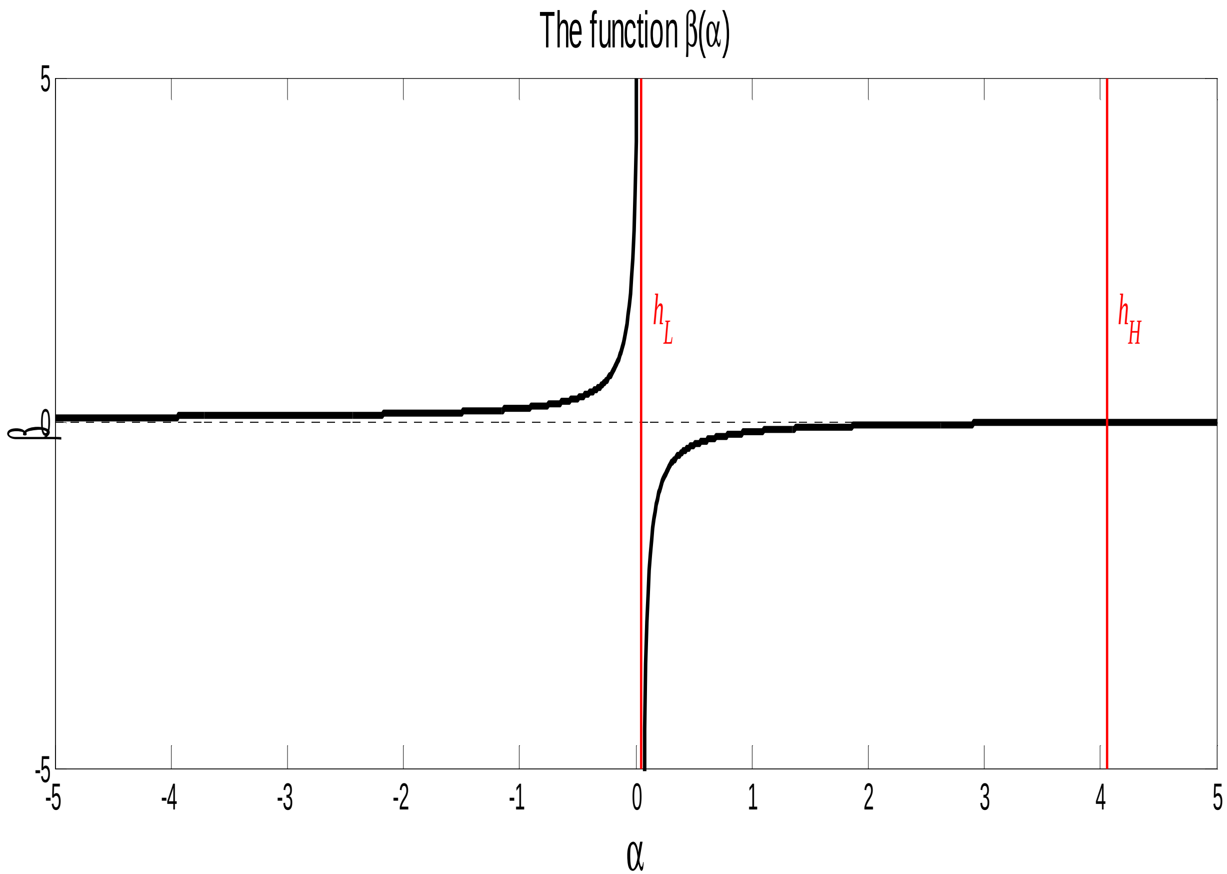
$$\Delta n_t = k^d + \beta^d \Delta w_t + b_{11}^d \Delta w_{t-1} + b_{12}^d \Delta n_{t-1} + b_{21}^d \Delta w_{t-2}$$
$$+ b_{22}^d \Delta n_{t-2} + \cdots + b_{m1}^d \Delta w_{t-m} + b_{m2}^d \Delta n_{t-m} + u_t^d$$

supply:

$$\Delta n_t = k^s + \alpha^s \Delta w_t + b_{11}^s \Delta w_{t-1} + b_{12}^s \Delta n_{t-1} + b_{21}^s \Delta w_{t-2}$$
$$+ b_{22}^s \Delta n_{t-2} + \cdots + b_{m1}^s \Delta w_{t-m} + b_{m2}^s \Delta n_{t-m} + u_t^s$$

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:

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



The function β(α)
5
0
-5
β
-5
-4
-3
-2
-1
0
1
2
3
4
5
α
h_L
h_H

If restrict $\alpha > 0, \beta < 0,$ and if $\hat{\omega}_{12} > 0,$

then $\hat{\alpha}_{MLE} > h_L = \hat{\omega}_{12}/\hat{\omega}_{11}$

Intuition: h_L 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

If restrict $\alpha > 0, \beta < 0,$ and if $\hat{\omega}_{12} > 0,$

then $\hat{\alpha}_{MLE} < h_H = \hat{\omega}_{22}/\hat{\omega}_{12}$

Intuition: h_H^{-1} is coefficient

from OLS regression of $\hat{\varepsilon}_{1t}$ on $\hat{\varepsilon}_{2t}$

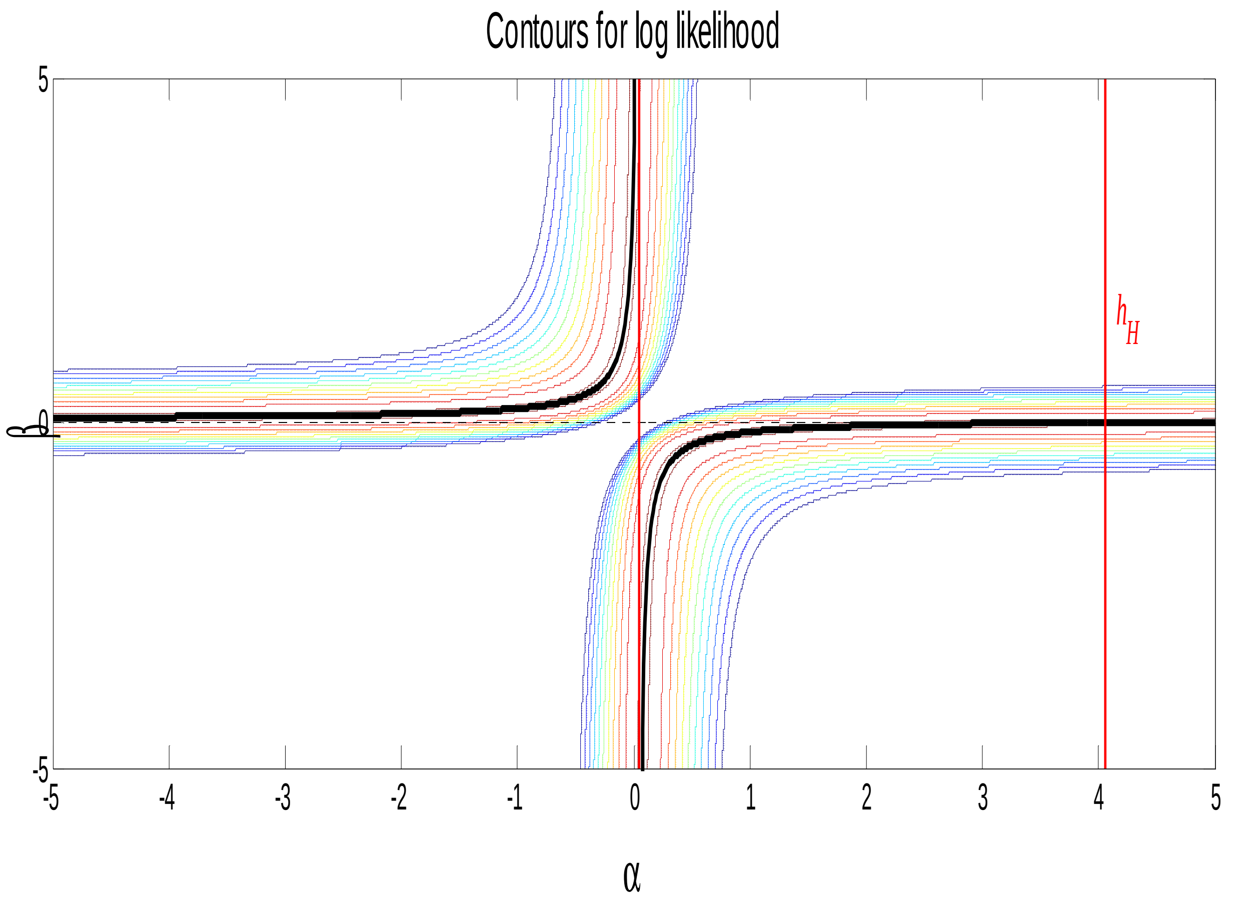
$=$ convex combination of α^{-1} and β^{-1}

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

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

$\Rightarrow h_L < \alpha < h_H$

$\beta \in (-\infty, 0]$



Contours for log likelihood
5
0
-5
β
-5
-4
-3
-2
-1
0
1
2
3
4
5
α
h_H

What do we know from other sources about short-run wage elasticity of labor demand?

- Hamermesh (1996) survey of microeconometric studies: 0.1 to 0.75
- Lichter, et. al. (2014) meta-analysis of 942 estimates: lower end of Hamermesh range
- Theoretical macro models can imply value above 2.5 (Akerlof and Dickens, 2007; Gali, et. al. 2012)

Prior for β: Student t with

location c_β, scale σ_β, d.f. ν_β,

truncated by $\beta \leq 0$

$c_\beta = -0.6, \sigma_\beta = 0.6, \nu_\beta = 3$

$\Rightarrow$ Prob$(\beta < -2.2) = 0.05$

Prob$(\beta > -0.1) = 0.05$

What do we know from other sources about wage elasticity of labor supply?

- Long run: often assumed to be zero because income and substitution effects cancel (e.g., Kydland and Prescott, 1982)
- Short run: often interpreted as Frisch elasticity
- Reichling and Whalen survey of microeconometric studies: 0.27-0.53
- Chetty, et. al. (2013) review of 15 quasi-experimental studies: < 0.5
- Macro models often assume value greater than 2 (Kydland and Prescott, 1982, Cho and Cooley, 1994, Smets and Wouters, 2007)

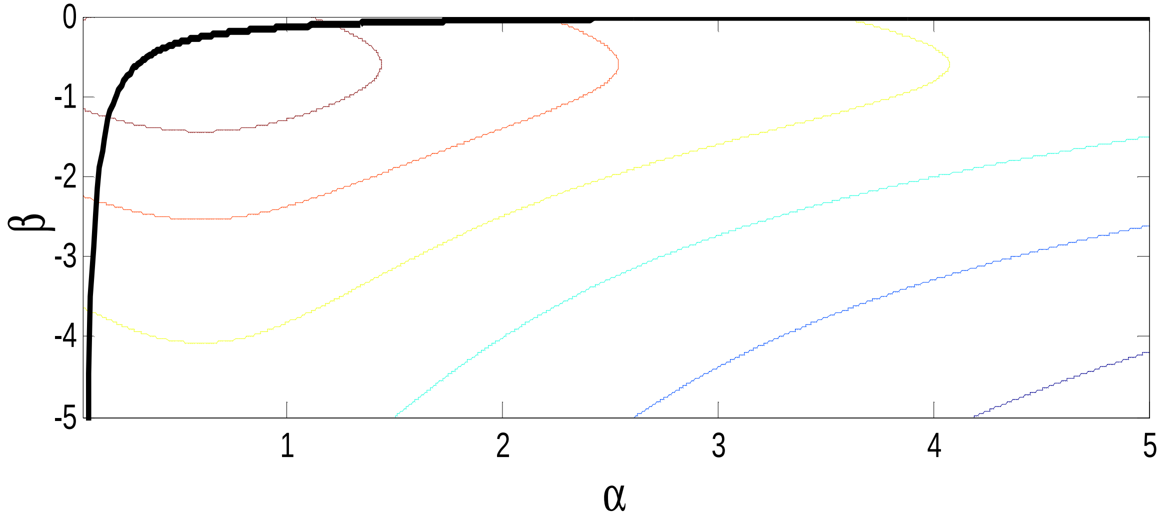
Prior for α: Student t with location c_α, scale σ_α, d.f. ν_a, truncated by $\alpha \geq 0$

$c_\alpha = 0.6, \sigma_\alpha = 0.6, \nu_\alpha = 3$

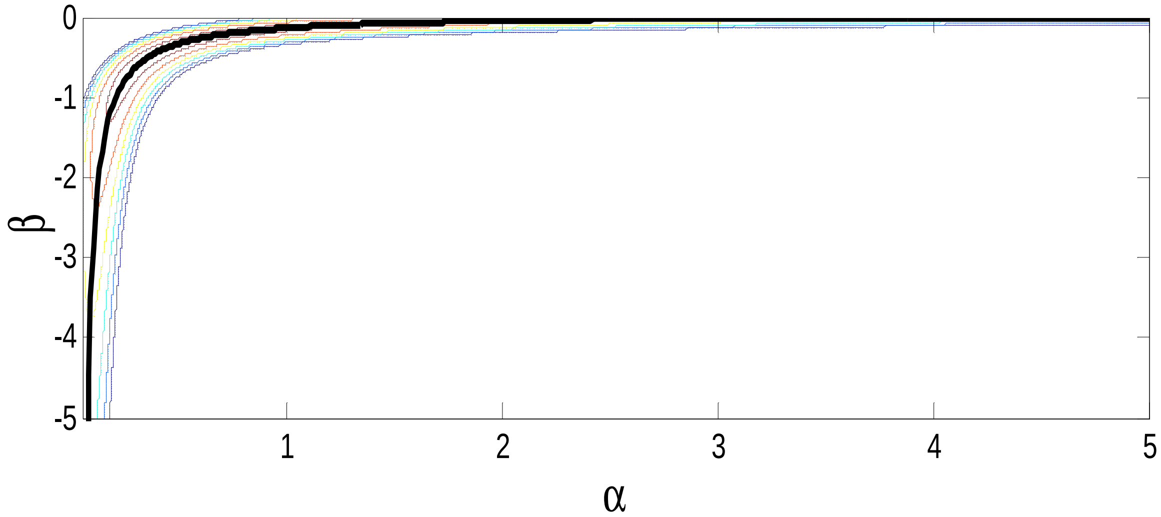
$\Rightarrow$ Prob$(\alpha < 0.1) = 0.05$

Prob$(\alpha > 2.2) = 0.05$

Contours for log of prior



Contours for log of posterior



Could we also use information about long-run labor supply elasticity?

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

(data used for $\mathbf{y}_t$ in VAR as estimated)

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

(data in levels)

$$\mathbf{u}_t = (u_t^d, u_t^s)$$

(vector of structural shocks)

$$\frac{\partial \Delta \tilde{\mathbf{y}}_{t+s}}{\partial \mathbf{u}_t'} = \frac{\partial \boldsymbol{\Delta} \tilde{\mathbf{y}}_{t+s}}{\partial \boldsymbol{\varepsilon}_t'} \frac{\partial \boldsymbol{\varepsilon}_t}{\partial \mathbf{u}_t'} = \boldsymbol{\Psi}_s \mathbf{A}^{-1}$$

$$(\boldsymbol{\Psi}_0 + \boldsymbol{\Psi}_1 L + \boldsymbol{\Psi}_2 L^2 + \cdots)$$

$$= (\mathbf{I}_n - \boldsymbol{\Phi}_1 L - \boldsymbol{\Phi}_2 L^2 - \cdots - \boldsymbol{\Phi}_m L^m)^{-1}$$

$$\frac{\partial \tilde{\mathbf{y}}_{t+s}}{\partial \mathbf{u}_t'} = \frac{\partial \Delta \tilde{\mathbf{y}}_{t+s}}{\partial \mathbf{u}_t'} + \frac{\partial \Delta \tilde{\mathbf{y}}_{t+s-1}}{\partial \mathbf{u}_t'} + \cdots + \frac{\partial \Delta \tilde{\mathbf{y}}_t}{\partial \mathbf{u}_t'}$$

$$= \boldsymbol{\Psi}_s \mathbf{A}^{-1} + \boldsymbol{\Psi}_{s-1} \mathbf{A}^{-1} + \cdots + \boldsymbol{\Psi}_0 \mathbf{A}^{-1}$$

$$\frac{\partial \tilde{\mathbf{y}}_{t+s}}{\partial \mathbf{u}_t'} = \boldsymbol{\Psi}_s \mathbf{A}^{-1} + \boldsymbol{\Psi}_{s-1} \mathbf{A}^{-1} + \cdots + \boldsymbol{\Psi}_0 \mathbf{A}^{-1}$$

$$\lim_{s\to\infty} \frac{\partial \tilde{\mathbf{y}}_{t+s}}{\partial \mathbf{u}_t'} = (\boldsymbol{\Psi}_0 + \boldsymbol{\Psi}_1 + \boldsymbol{\Psi}_2 + \cdots)\mathbf{A}^{-1}$$

$$= (\mathbf{I}_n - \boldsymbol{\Phi}_1 - \boldsymbol{\Phi}_2 - \cdots - \boldsymbol{\Phi}_m)^{-1} \mathbf{A}^{-1}$$

$$= [\mathbf{A}(\mathbf{I}_n - \boldsymbol{\Phi}_1 - \boldsymbol{\Phi}_2 - \cdots - \boldsymbol{\Phi}_m)]^{-1}$$

$$= [\mathbf{A} - \mathbf{B}_1 - \mathbf{B}_2 - \cdots - \mathbf{B}_m]^{-1}$$

$$\lim_{s\to\infty} \frac{\partial \tilde{\mathbf{y}}_{t+s}}{\partial \mathbf{u}_t'} = [\mathbf{A} - \mathbf{B}_1 - \mathbf{B}_2 - \cdots - \mathbf{B}_m]^{-1}$$

Labor demand shock (shock #1)

has zero long run effect on employment

(second element of $\tilde{\mathbf{y}}_{t+s}$) if and only if

$(2, 1)$ element is zero:

$$0 = -\alpha^s - b_{11}^s - b_{21}^s - \cdots - b_{m1}^s$$

$$0 = -\alpha^s - b_{11}^s - b_{21}^s - \cdots - b_{m1}^s$$

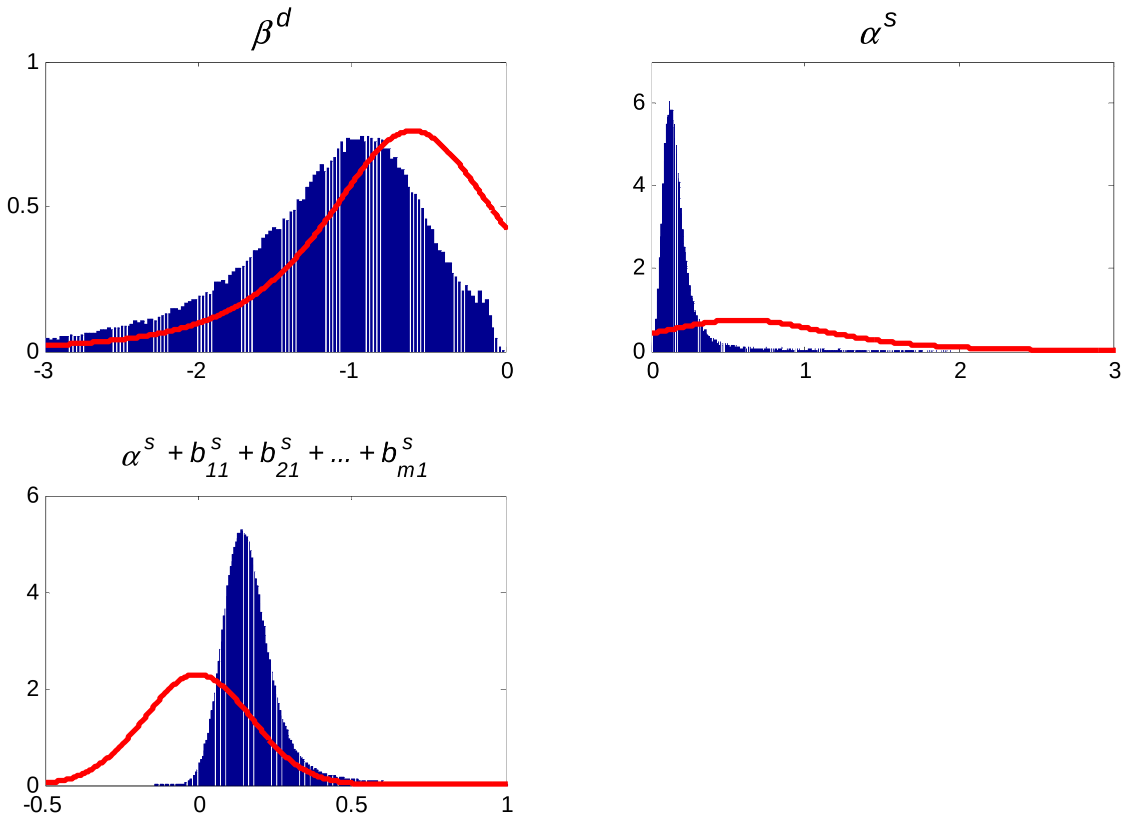
Usual approach: impose this condition as untestable identifying assumption

Our suggestion: instead represent as prior belief,

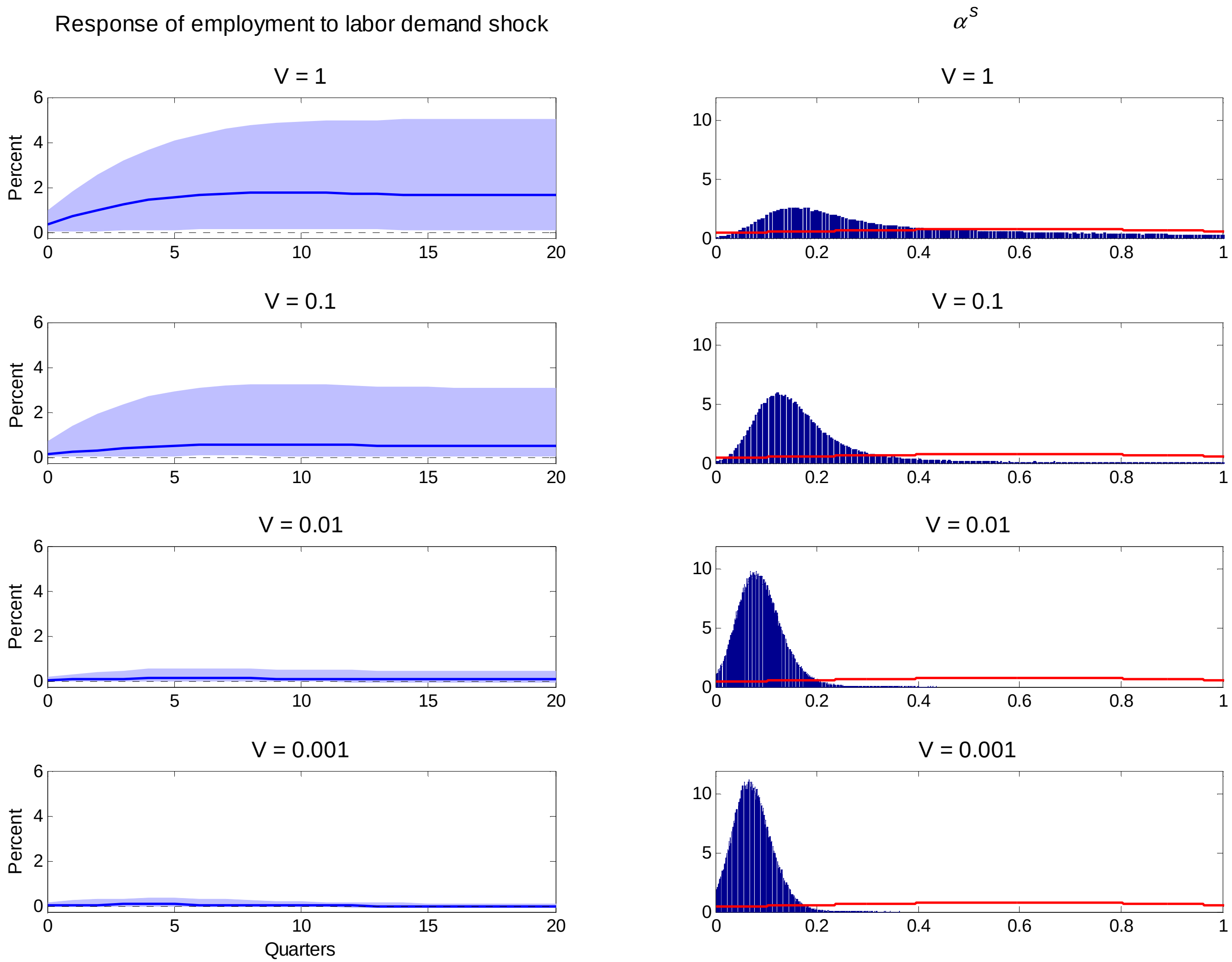
$$(b_{11}^s + b_{21}^s + \cdots + b_{m1}^s)|\mathbf{A},\mathbf{D} \sim N(-\alpha^s, d_{22}V)$$

$V = 0.1 \Rightarrow$ prior given same weight as 10 observations on $\mathbf{y}_t$

Prior and posterior distributions for short-run elasticities and long-run impact



Posterior medians and 95% credibility regions for structural impulse-response functions



Response of employment to labor demand shock
α^s
V = 1
V = 0.1
V = 0.01
V = 0.001
Percent
Quarters

Application 2: Shocks to oil supply and demand

q = quantity of oil produced

y = measure of economic activity

p = real price of oil

oil supply:

$$q_t = \alpha_{qy} y_t + \alpha_{qp} p_t + \mathbf{b}_1' \mathbf{x}_{t-1} + u_{1t}$$

economic activity:

$$y_t = \alpha_{yq} q_t + \alpha_{yp} p_t + \mathbf{b}_2' \mathbf{x}_{t-1} + u_{2t}$$

inverse of oil demand curve:

$$p_t = \alpha_{pq} q_t + \alpha_{py} y_t + \mathbf{b}_3' \mathbf{x}_{t-1} + u_{3t}$$

Note: α_{pq} = inverse of short-run price-elasticity of oil demand

$$\mathbf{A}\mathbf{y}_t = \boldsymbol{\lambda} + \mathbf{B}_1\mathbf{y}_{t-1} + \cdots + \mathbf{B}_m\mathbf{y}_{t-m} + \mathbf{u}_t$$

$$\mathbf{A} = \begin{bmatrix} 1 & -\alpha_{qy} & -\alpha_{qp} \\ -\alpha_{yq} & 1 & -\alpha_{yp} \\ -\alpha_{pq} & -\alpha_{py} & 1 \end{bmatrix}$$

A Bayesian interpretation of traditional identification

Kilian (2009): Cholesky identification

$\alpha_{qy} = \alpha_{qp} = \alpha_{yp} = 0$

oil supply:

$q_t = \alpha_{qy} y_t + \alpha_{qp} p_t + \mathbf{b}_1' \mathbf{x}_{t-1} + u_{1t}$

economic activity:

$y_t = \alpha_{yq} q_t + \alpha_{yp} p_t + \mathbf{b}_2' \mathbf{x}_{t-1} + u_{2t}$

inverse of oil demand curve:

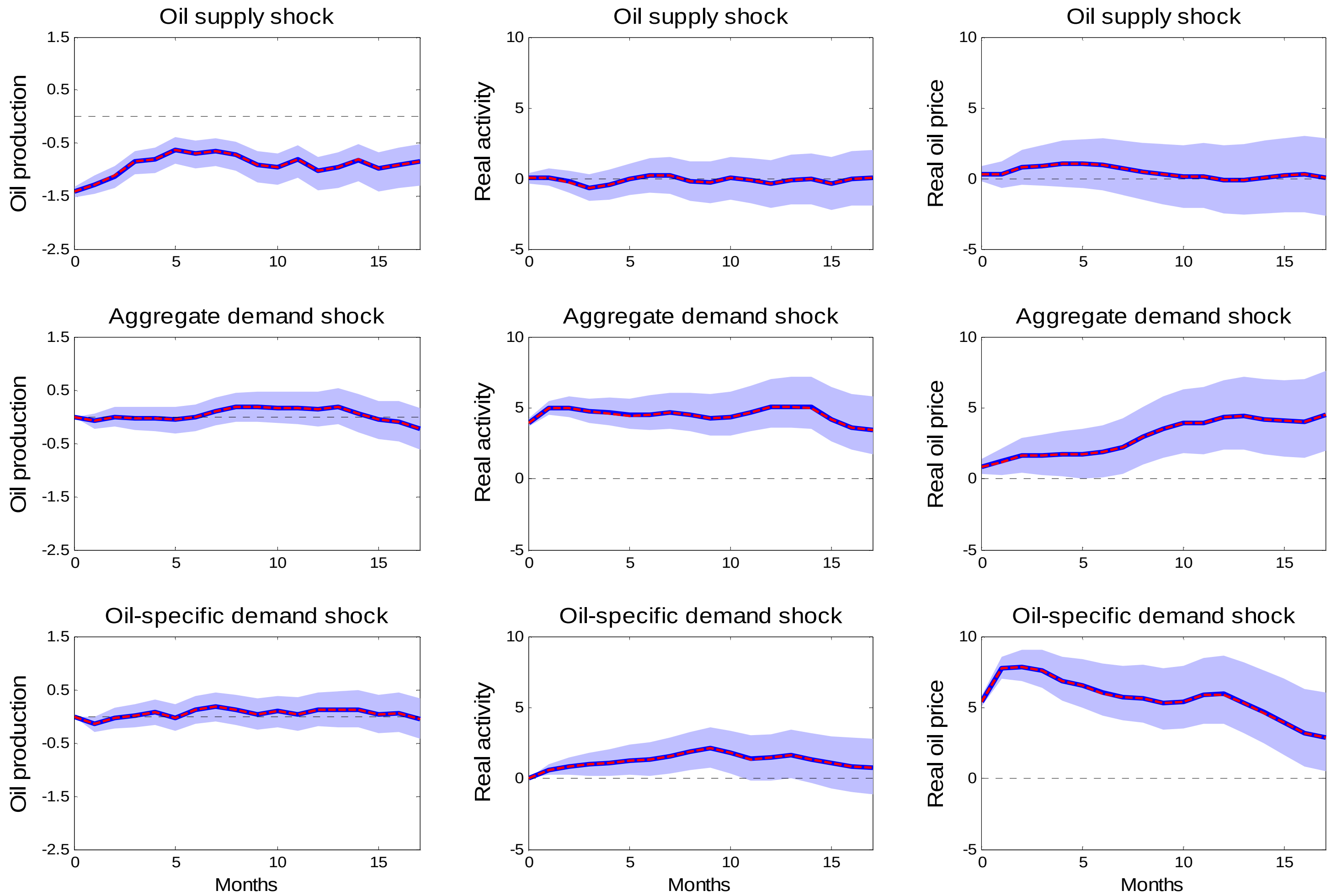
$p_t = \alpha_{pq} q_t + \alpha_{py} y_t + \mathbf{b}_3' \mathbf{x}_{t-1} + u_{3t}$

Bayesian translation: I put absolutely zero possibility on any $\mathbf{A}$ unless the (1,2), (1,3), and (2,3) elements are all zero.

I have no information at all about the (2,1), (3,1), and (3,2) elements.

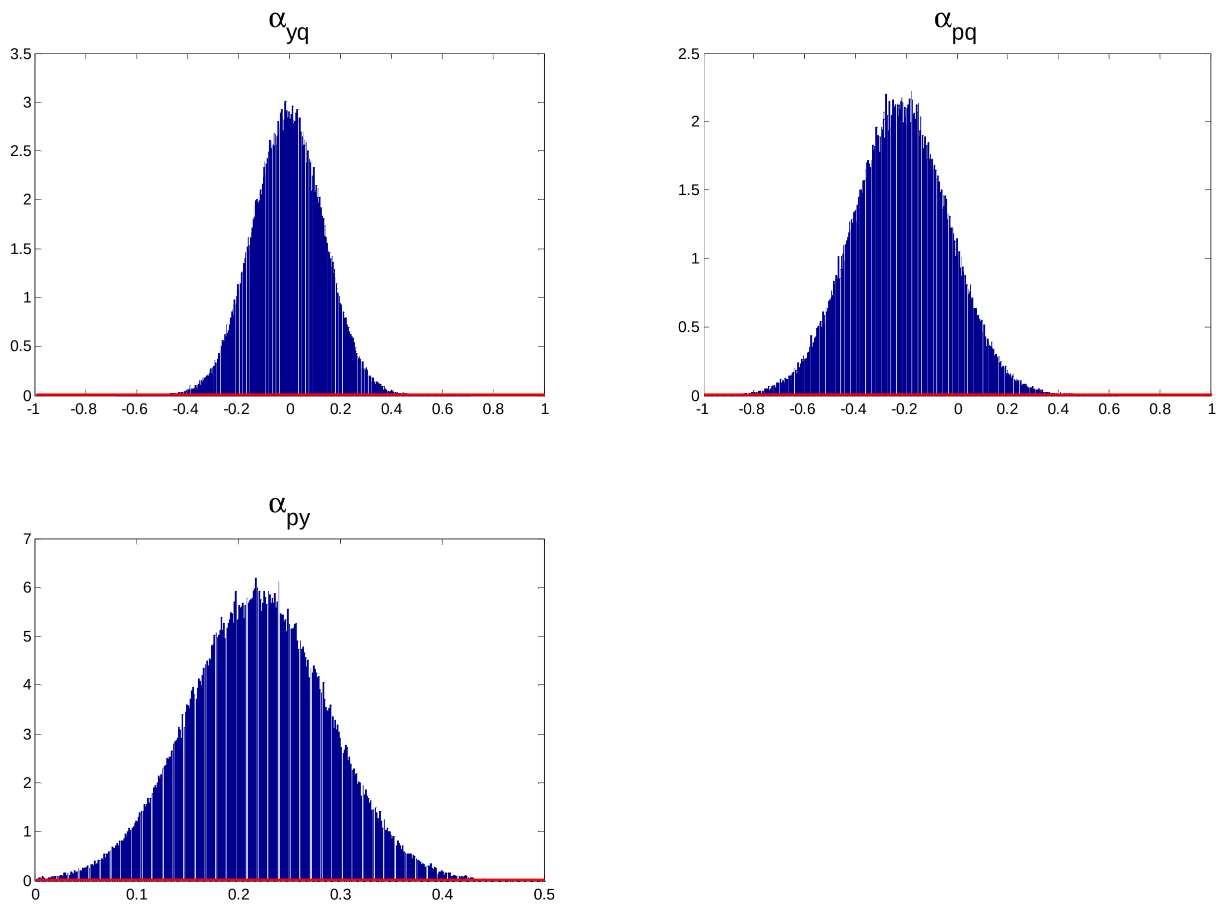
(2,1): $p(\alpha_{yq}) \sim$ Student t with location 0, scale 100, d.f. $=$ 3

Same for $p(\alpha_{pq})$ and $p(\alpha_{py})$



Blue: posterior median IRF as calculated using Baumeister-Hamilton algorithm for above prior.

Red: IRF calculated using Kilian's method for original data set.



Prior (red) and posterior (blue) distributions for unknown elements of **A**