Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information*

James D. Hamilton, UCSD

Aarhus University CREATES Lecture

November 10, 2015

*Based on joint research with Christiane Baumeister, University of Notre Dame

Can we give structural interpretation to VARs using only sign restrictions?

- Parameters only set identified: data cannot distinguish different models within set
- Frequentist methods
  - Awkward and computationally demanding [Moon, Schorfheide, and Granziera, 2013]
- Bayesian methods
  - Numerically simple [Rubio-Ramírez, Waggoner, and Zha (2010)]
  - For some questions, estimate reflects only the prior [Poirier (1998); Moon and Schorfheide (2012)]

Today’s lecture

- Calculate small-sample and asymptotic Bayesian posterior distributions for partially identified structural VAR
- Characterize regions of parameter space about which data are uninformative
- Explicate the prior that is implicit in traditional sign-restricted structural VAR algorithms
- Propose that researchers use informative priors and report difference between prior and posterior distributions
- Illustrate with simple model of labor market
- Code available at http://econweb.ucsd.edu/~jhamilton/BHcode.zip

Outline

1. Bayesian inference for partially identified structural VARs
2. Implicit priors in traditional approach
3. Empirical application: shocks to labor supply and demand

1. Bayesian inference for partially identified structural vector autoregressions

Structural model of interest:

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

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

$\mathbf{D}$ diagonal

Example: demand and supply

$$q_t = k^d + \beta^d p_t + b_{11}^d p_{t-1} + b_{12}^d q_{t-1} + b_{21}^d p_{t-2} + b_{22}^d q_{t-2} + \cdots + b_{m1}^d p_{t-m} + b_{m2}^d q_{t-m} + u_t^d$$

$$q_t = k^s + \alpha^s p_t + b_{11}^s p_{t-1} + b_{12}^s q_{t-1} + b_{21}^s p_{t-2} + b_{22}^s q_{t-2} + \cdots + b_{m1}^s p_{t-m} + b_{m2}^s q_{t-m} + u_t^s$$

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

Reduced-form (can easily estimate):

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{\Phi}_m \mathbf{y}_{t-m} + \boldsymbol{\varepsilon}_t$$

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

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

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

$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{c} & \mathbf{\Phi}_1 & \mathbf{\Phi}_2 & \cdots & \mathbf{\Phi}_m \end{bmatrix}$$

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

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

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

Reduced form:

$$\mathbf{y}_t = \mathbf{c} + \boldsymbol{\Phi}_1\mathbf{y}_{t-1} + \cdots + \boldsymbol{\Phi}_m\mathbf{y}_{t-m} + \boldsymbol{\varepsilon}_t$$

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

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

$\mathbf{A}\boldsymbol{\Omega}\mathbf{A}' = \mathbf{D}$ (diagonal)

Problem: there are more unknown elements in $\mathbf{D}$ and $\mathbf{A}$ than in $\boldsymbol{\Omega}$.

Supply and demand example:

4 structural parameters in $\mathbf{A}, \mathbf{D}$

$(\alpha^s, \beta^d, d_{11}, d_{22})$

only 3 parameters known from $\mathbf{\Omega}$

$(\omega_{11}, \omega_{12}, \omega_{22})$

We can achieve partial identification from

$\alpha^s \geq 0,\ \beta^d \leq 0$

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

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

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

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: $\kappa_i, \tau_i \rightarrow 0$

$$\mathbf{B} = \left[\begin{array}{ccccc} \boldsymbol{\lambda} & \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_m \end{array} \right]$$

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

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

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

Recommended default priors (Minnesota prior)

Doan, Litterman, Sims (1984)

Sims and Zha (1998)

◦ elements of $\mathbf{m}_i$ corresponding to lag 1 given by $\mathbf{a}_i$

◦ all other elements of $\mathbf{m}_i$ are zero

◦ $\mathbf{M}_i$ diagonal with smaller values on bigger lags

$\Rightarrow$ prior belief that each element of $\mathbf{y}_t$ behaves like a random walk

τ_i function of $\mathbf{A}$ (or prior mode of $p(\mathbf{A})$) and scale of data

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

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

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

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

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

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

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

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

Hadamard's Inequality:

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

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

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

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

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

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

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

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

2. Prior beliefs that are implicit in the traditional approach

Alternatively could specify priors in terms of impact matrix:

$$\mathbf{y}_t = \mathbf{\Phi}\mathbf{x}_{t-1} + \mathbf{H}\mathbf{u}_t$$

$$\mathbf{H} = \frac{\partial \mathbf{y}_t}{\partial \mathbf{u}_t'} = \mathbf{A}^{-1}$$

We found solution for all priors on $\mathbf{A}$ and joint for $p(\mathbf{A}, \mathbf{D})$ when $\mathbf{D}|\mathbf{A}$ is natural conjugate.

Traditional approach best understood as $p(\mathbf{H}|\mathbf{\Omega})$.

(1) Calculate Cholesky factor $\mathbf{\Omega} = \mathbf{P}\mathbf{P}'$.

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

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

(4) Generate candidate $\mathbf{H} = \mathbf{PQ}$ and keep if it satisfies sign restrictions.

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 $q_{21} = \sin\theta$.

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

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

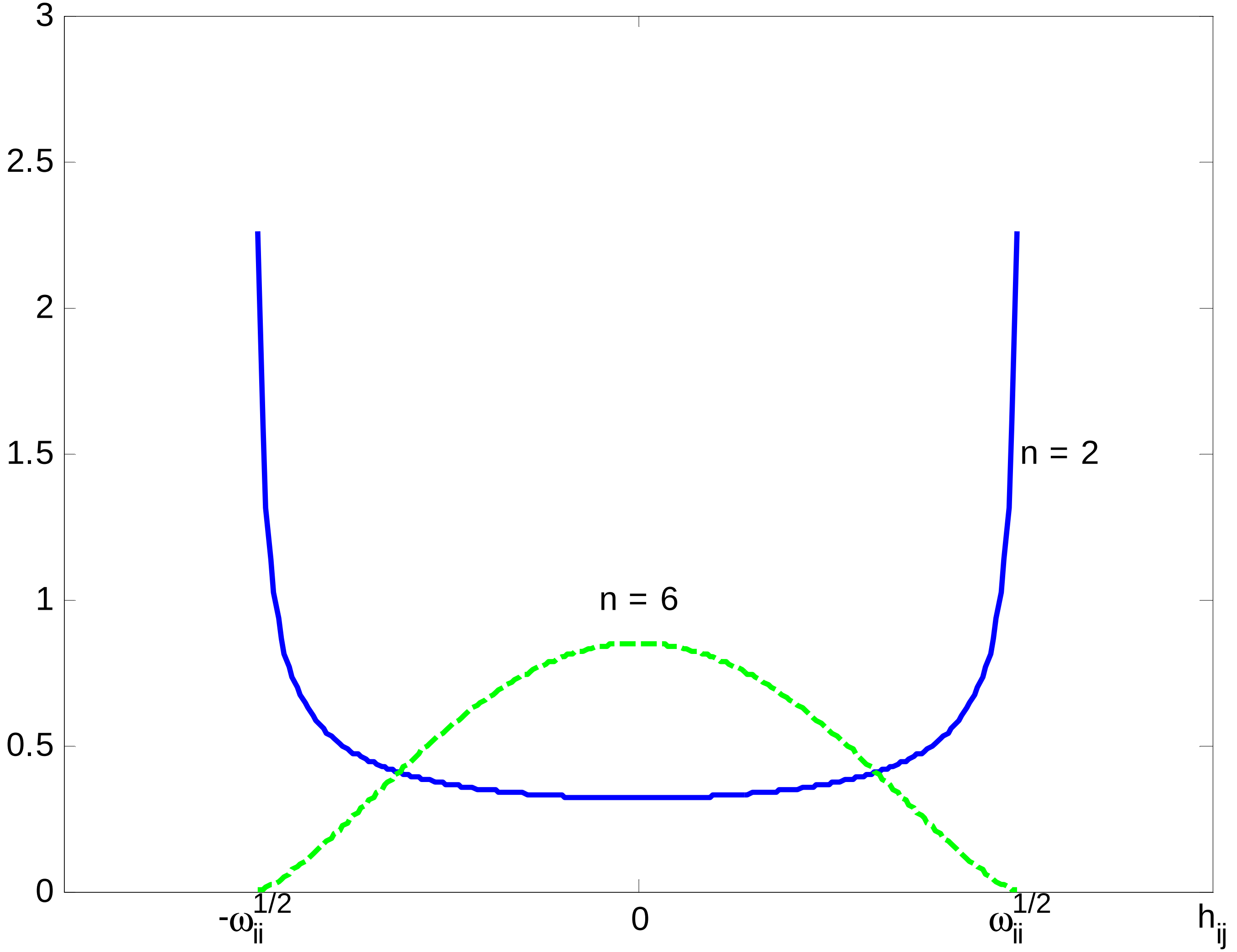
$$q_{i1} = x_{i1} / \sqrt{x_{11}^2 + \cdots + x_{n1}^2}$$

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

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

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

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:

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

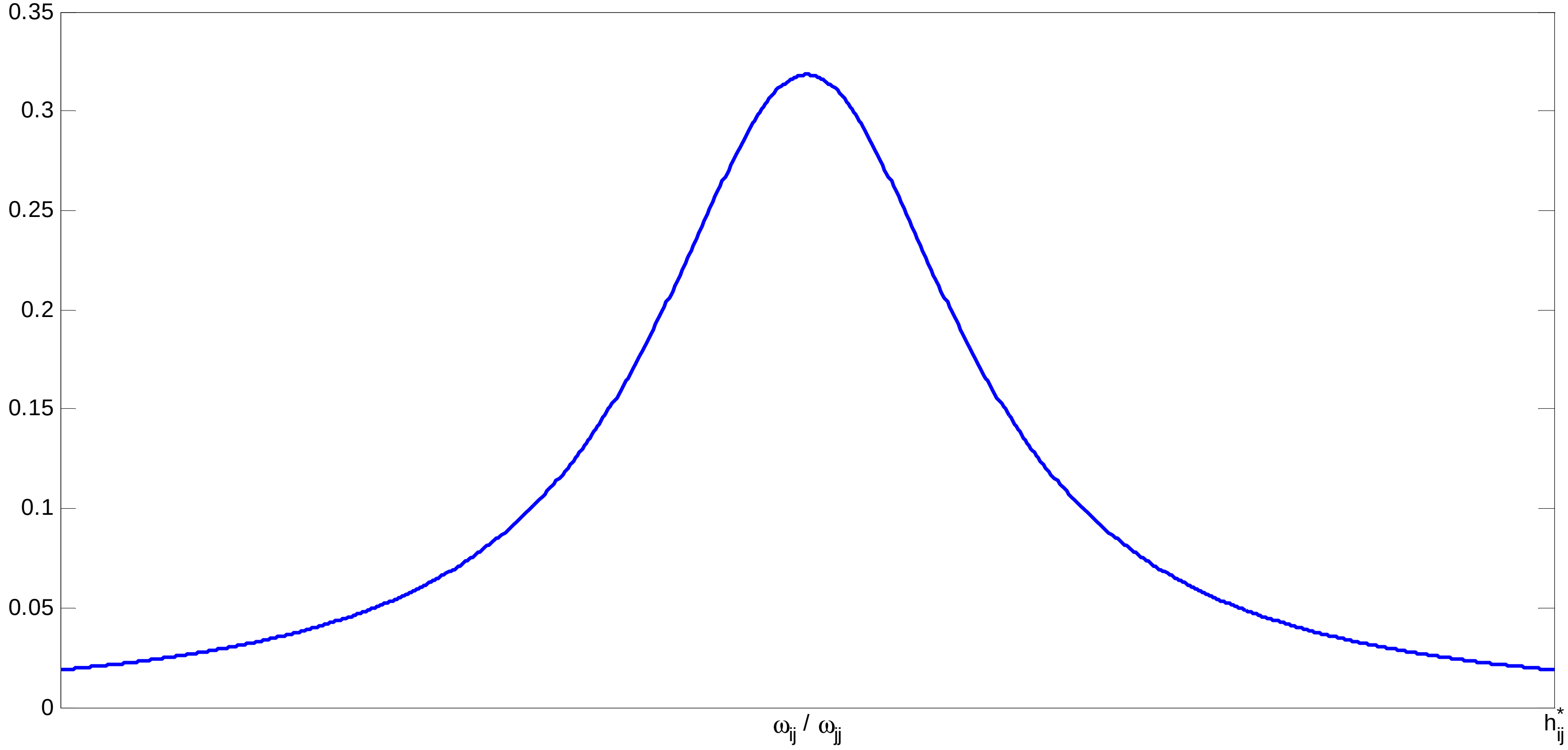
e.g., response of quantity to demand shock that raises price by 1% is the short-run elasticity of supply

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

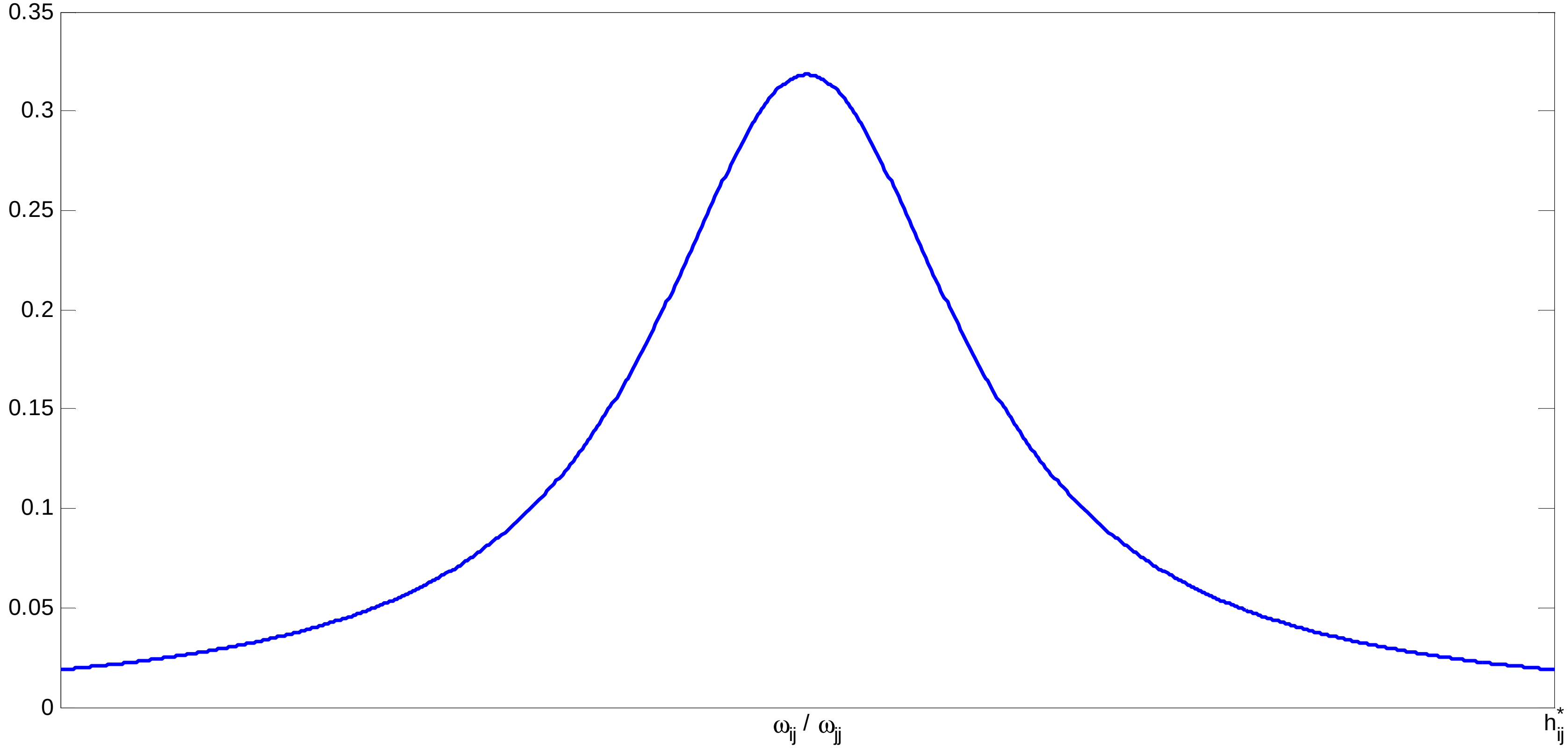
$\Rightarrow h_{ij}^{*}|\boldsymbol{\Omega}$ ~ Cauchy($c_{ij}^{*}, \sigma_{ij}^{*}$)

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

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.

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

variable 1 $=$ price, variable 2 $=$ quantity

shock 1 $=$ demand, 2 $=$ supply

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

Can show if $p_{21} > 0,$ sign restrictions require

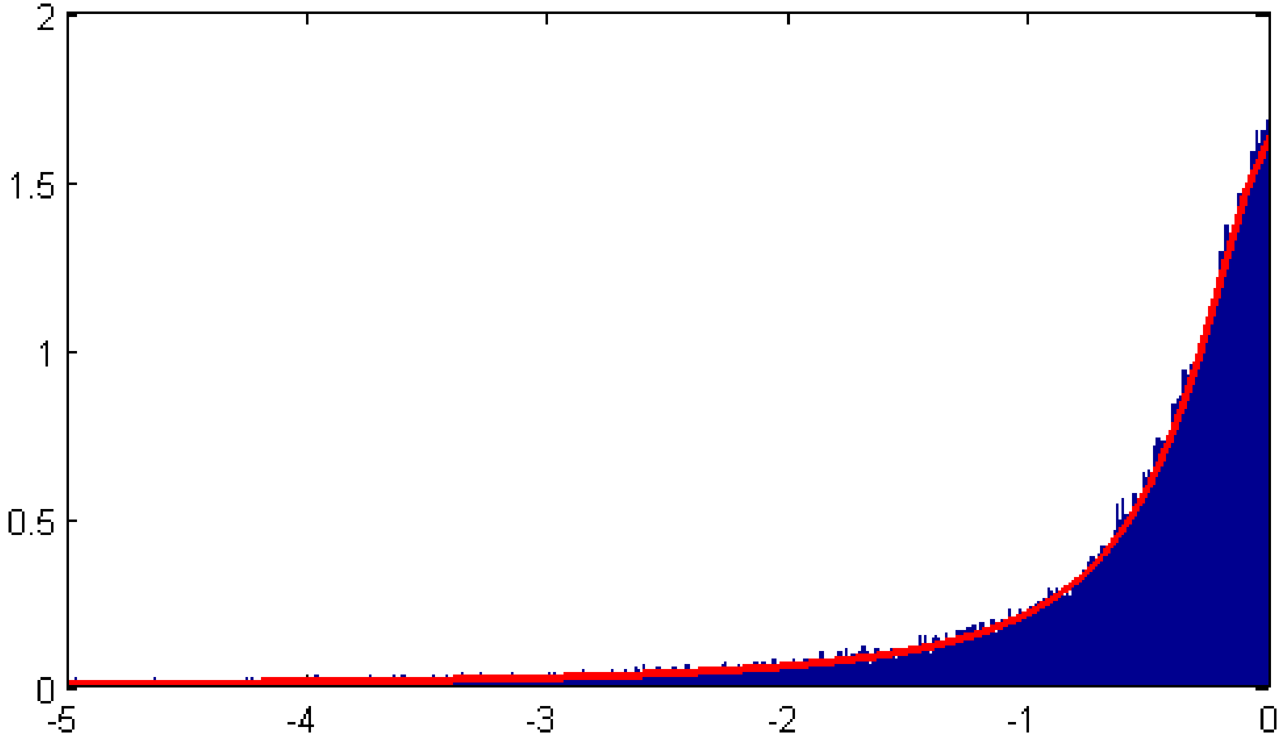
$\theta \in [0, \tilde{\theta}]$ for $\cot\tilde{\theta} = p_{21}/p_{22}$

$\Rightarrow h_{22}^{*} \in (-\infty, 0]$ (demand elasticity unrestricted)

$h_{21}^{*} \in [\omega_{21}/\omega_{11}, \omega_{22}/\omega_{21}]$ (supply elasticity in certain range)

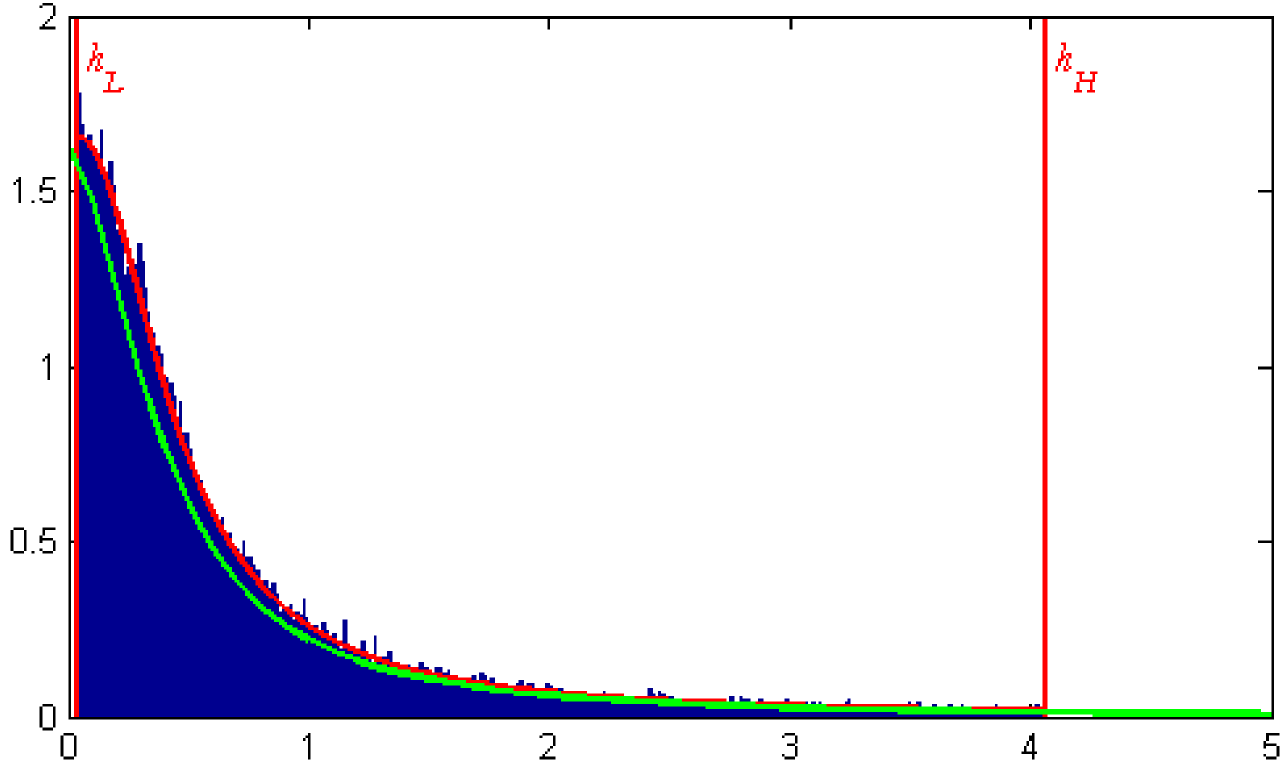
Apply traditional algorithm to 8-lag VAR fit to growth rates of U.S. real compensation per worker and U.S. employment, 1970:Q1-2014:Q2.

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

3. 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:

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

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 $\alpha \geq 0$

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

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

Prob($\alpha > 2.2) = 0.05$

We might also use information about long-run labor supply elasticity

Proposition: labor demand shock has zero long run effect on employment iff

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

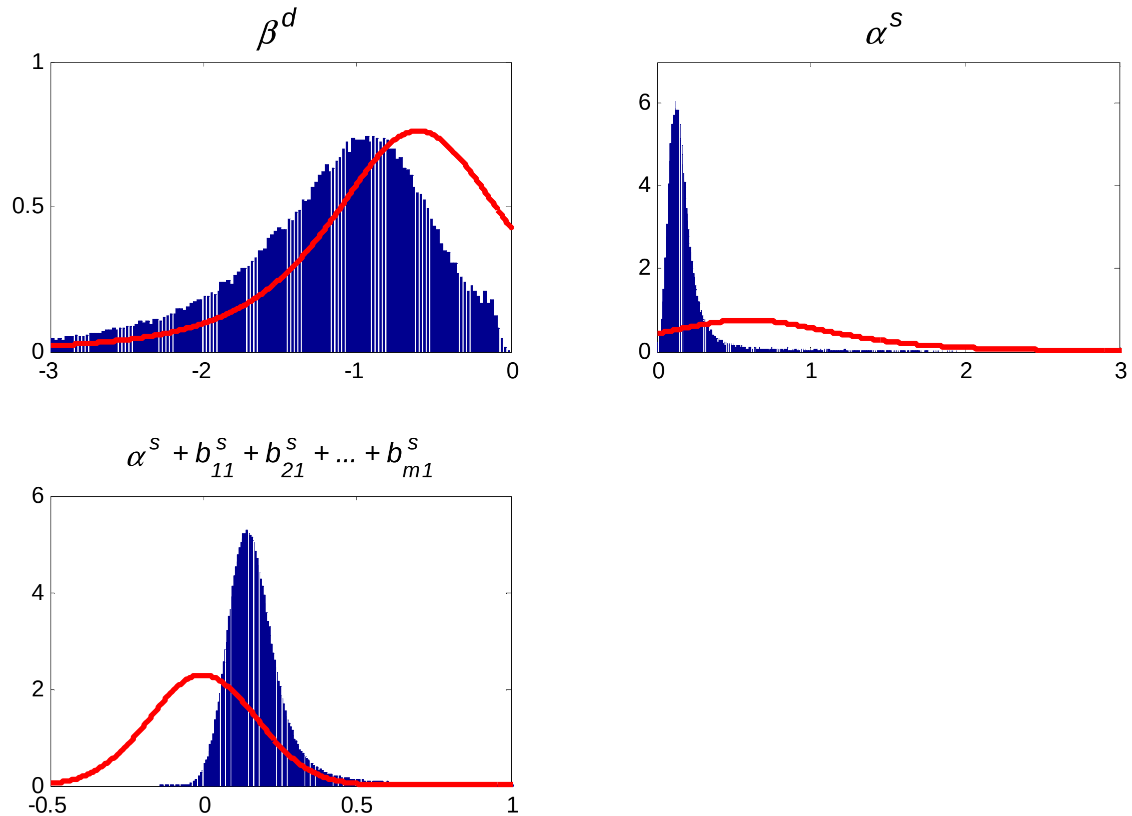
Usual approach: impose this condition as untestable identifying assumption

Our suggestion: instead represent as prior belief,

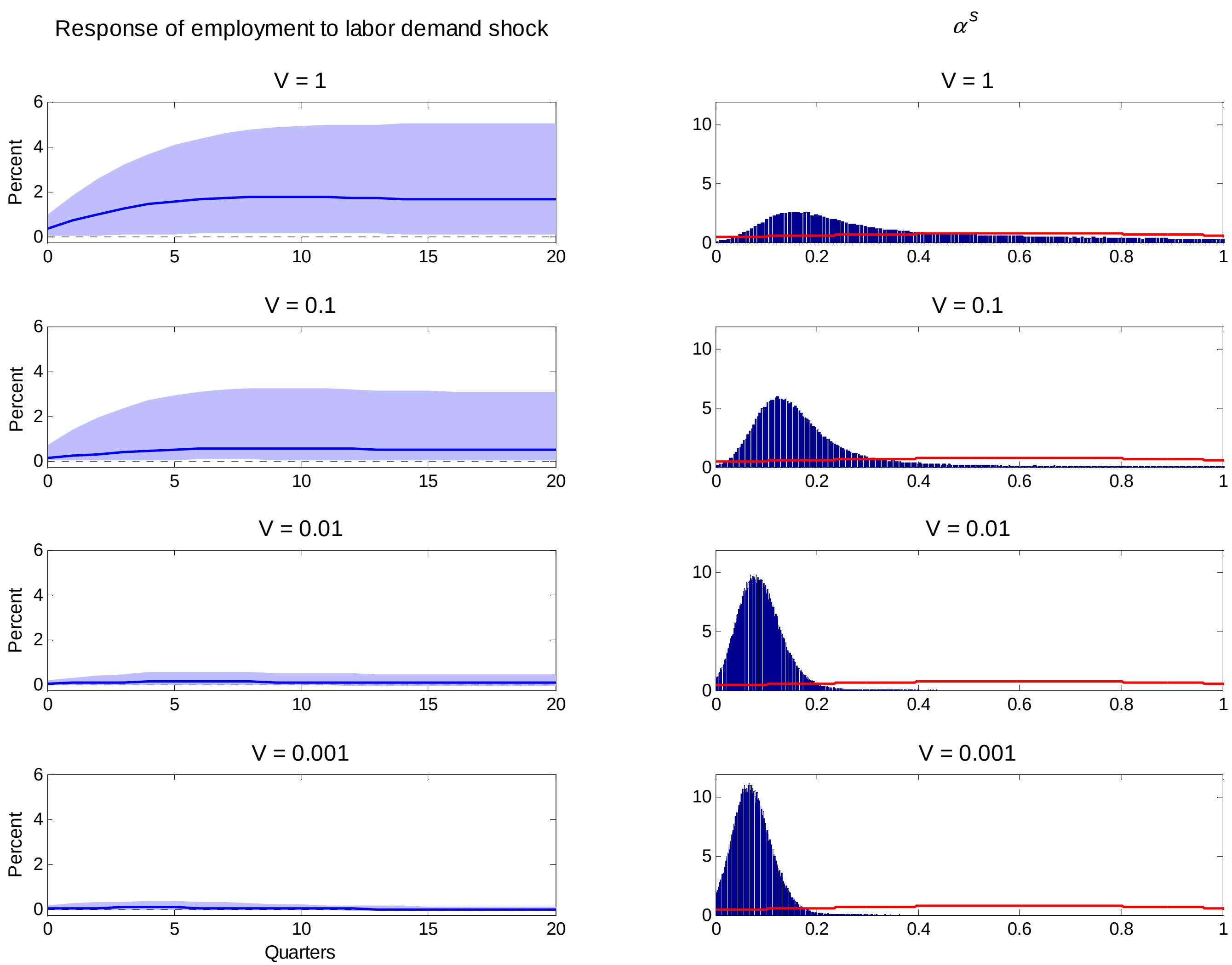
$(b^s_{11} + b^s_{21} + \cdots + b^s_{m1})|\mathbf{A},\mathbf{D} \sim N(-\alpha^s, d_{22}V)$

$V = 0.1 \Rightarrow$ 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