II. Vector autoregressions

D. Identification using inequality constraints
   1. Traditional approach to identification
Dynamic structural model:

\[
\begin{align*}
\mathbf{A} \mathbf{y}_t &= \lambda + \mathbf{B}_1 \mathbf{y}_{t-1} + \\
&\quad \cdots + \mathbf{B}_m \mathbf{y}_{t-m} + \mathbf{D}^{1/2} \mathbf{v}_t \\
&= \mathbf{B} \mathbf{x}_{t-1} + \mathbf{D}^{1/2} \mathbf{v}_t \\
\mathbf{x}'_{t-1} &= (1, \mathbf{y}'_{t-1}, \mathbf{y}'_{t-2}, \ldots, \mathbf{y}'_{t-m})' \\
\mathbf{v}_t &\sim \text{i.i.d. } \mathcal{N}(0, \mathbf{I}_n)
\end{align*}
\]
\[ D = \begin{bmatrix}
    d_{11} & 0 & \cdots & 0 \\
    0 & d_{22} & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & d_{nn}
\end{bmatrix} \]

\[ D^{1/2} = \begin{bmatrix}
    \sqrt{d_{11}} & 0 & \cdots & 0 \\
    0 & \sqrt{d_{22}} & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & \sqrt{d_{nn}}
\end{bmatrix} \]
Example: demand and supply

\[ q_t = k^d + \beta^d p_t + b^d_{11} p_{t-1} + b^d_{12} q_{t-1} + b^d_{21} p_{t-2} \]
\[ + b^d_{22} q_{t-2} + \cdots + b^d_{m1} p_{t-m} + b^d_{m2} q_{t-m} + \sigma_d \nu^d_t \]

\[ q_t = k^s + \alpha^s p_t + b^s_{11} p_{t-1} + b^s_{12} q_{t-1} + b^s_{21} p_{t-2} \]
\[ + b^s_{22} q_{t-2} + \cdots + b^s_{m1} p_{t-m} + b^s_{m2} q_{t-m} + \sigma_s \nu^s_t \]

\[ y_t = \begin{bmatrix} q_t \\ p_t \end{bmatrix} \quad A = \begin{bmatrix} 1 & -\beta^d \\ 1 & -\alpha^s \end{bmatrix} \]
Reduced-form (forecasting equations):
\[ y_t = \Phi x_{t-1} + \varepsilon_t \]
\[ x'_{t-1} = (1, y'_{t-1}, y'_{t-2}, \ldots, y'_{t-m})' \]
\[ \varepsilon_t \sim \text{i.i.d. } N(0, \Omega) \]
\[ \hat{\Phi} = \left( \sum_{t=1}^{T} y_t x'_{t-1} \right) \left( \sum_{t=1}^{T} x_{t-1} x'_{t-1} \right)^{-1} \]
\[ \hat{\varepsilon}_t = y_t - \hat{\Phi} x_{t-1} \]
\[ \hat{\Omega} = T^{-1} \sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}'_t \]
Nonorthogonalized impulse-response:
\[ \frac{\partial y_{t+s}}{\partial \varepsilon'_t} = \Psi_s \]
\[ (n \times n) \]
Structural model:
\[Ay_t = Bx_{t-1} + D^{1/2}v_t \quad v_t \sim N(0, I_n)\]

Structural impulse response:
\[
\frac{\partial y_{t+s}}{\partial v'_t} = \frac{\partial y_{t+s}}{\partial \varepsilon_t} \frac{\partial \varepsilon_t}{\partial v'_t} = \Psi_s H
\]

\[H = \frac{\partial \varepsilon_t}{\partial v'_t} = A^{-1}D^{1/2}\]

Reduced form:
\[y_t = \Phi x_{t-1} + \varepsilon_t \quad \varepsilon_t \sim \text{i.i.d. } N(0, \Omega)\]
\[\Phi = A^{-1}B\]
\[\varepsilon_t = A^{-1}D^{1/2}v_t = Hv_t\]
\[E(\varepsilon_t\varepsilon'_t) = \Omega = A^{-1}D(A^{-1})'\]
\[ \Omega = A^{-1}D(A^{-1})' \]

Can estimate the 3 parameters in \( \Omega \) by OLS.
But there are 4 unknown elements in \( A \) and \( D \) \( (\beta^d, \alpha^s, \sigma_d, \sigma_s) \)
Traditional approach to identification:
Put enough restrictions on $A$ and $D$
so that for any $\Omega$ there is a unique
$A, D$ for which $\Omega = A^{-1} D (A^{-1})'$
Example: assume short-run demand elasticity $\beta^d = 0$ (meaning $A$ and $A^{-1}$ are lower triangular):

$$A = \begin{bmatrix}
1 & -\beta^d \\
1 & -\alpha^s
\end{bmatrix}$$
If $D$ is diagonal, then $H$ also lower triangular:

$$H = \frac{\partial y_t}{\partial v_t'} = A^{-1}D^{1/2}$$

If normalize shocks so that $v_t^d$ raises $q_t$ and $v_t^s$ raises $p_t$, there is unique lower-triangular matrix $H$ with positive diagonal elements such that $HH' = \Omega$ namely, $H =$ Cholesky factor of $\Omega$
Estimate following from reduced form:
\[
\hat{\Omega} = T^{-1} \sum_{t=1}^{T} \hat{\epsilon}_t \hat{\epsilon}_t'
\]
\[
\hat{P}' \cdot \hat{P}' = \hat{\Omega} \quad \text{(Cholesky factor)}
\]
\[
\hat{\Psi}_s = \frac{\partial y_{t+s}}{\partial \epsilon_t'}
\]
\[
\hat{H} = \hat{P}
\]
Then we infer dynamic structural responses:
\[
\frac{\partial y_{t+s}}{\partial v_t'} = \frac{\partial y_{t+s}}{\partial \epsilon_t'} \cdot \frac{\partial \epsilon_t}{\partial v_t'} = \hat{\Psi}_s \hat{P}
\]
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:

\[
H = \begin{bmatrix}
\frac{\partial q_t}{\partial v_t^d} & \frac{\partial q_t}{\partial v_t^s} \\
\frac{\partial p_t}{\partial v_t^d} & \frac{\partial p_t}{\partial v_t^s}
\end{bmatrix} = \begin{bmatrix} + & - \\ + & + \end{bmatrix}
\]

Can we use a prior for \(H|\Omega\) that implies these signs but is otherwise uninformative?
Proposal: $H = PQ$ where $\Omega = PP'$ and $Q$ is drawn from a Haar-uniform distribution from the set of all orthogonal matrices ($QQ' = I_n$)
How generate a draw for $Q$?

1. Generate $(n \times n) \, X = [x_{ij}]$ of $N(0, 1)$.
2. Find $X = QR$ for $Q$ orthogonal and $R$ upper triangular.
First column of \( Q = \) first column of \( X \) 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 \( \theta \) the angle between \( (x_{11}, x_{21}) \) and \((1, 0)\) while \( q_{21} = \sin \theta \).
\[ Q = \begin{cases} 
\begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta 
\end{bmatrix} & \text{with prob 1/2} \\
\begin{bmatrix}
\cos \theta & \sin \theta \\
\sin \theta & -\cos \theta 
\end{bmatrix} & \text{with prob 1/2}
\end{cases} \]

\[ \theta \sim U(-\pi, \pi) \]
Algorithm to generate draw subject to sign restrictions:

1. Generate $\Omega^{-1} \sim W(T, T\hat{\Omega})$
2. Calculate $PP' = \Omega$
3. Generate orthogonal $Q$
4. Calculate $H = PQ$
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 $\theta$) is in general informative about nonlinear transformations of $\theta$. Baumeister and Hamilton (2014) calculate implicit priors for other objects.
\[ q_{i1} = \frac{x_{i1}}{\sqrt{x_{11}^2 + \cdots + x_{n1}^2}} \]

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

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

\[ h_{11} = p_{11}q_{11} = \sqrt{\omega_{11}} q_{11} \]
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:

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

\[
x_{21}/x_{11} \sim \text{Cauchy}(0,1)
\]

\[
\Rightarrow h_{ij}^*|\Omega \sim \text{Cauchy}(c_{ij}^*, \sigma_{ij}^*)
\]

\[
c_{ij}^* = \omega_{ij}/\omega_{jj}
\]

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

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

\( h_{11}, h_{12} \geq 0 \Rightarrow \theta \in [0, \pi/2] \)
\[
\begin{bmatrix}
  h_{11} & h_{12} \\
  h_{21} & h_{22}
\end{bmatrix} = 
\begin{bmatrix}
  p_{11} \cos \theta \\
  (p_{21} \cos \theta + p_{22} \sin \theta) \\
  (p_{21} \sin \theta - p_{22} \cos \theta)
\end{bmatrix}
\]

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 \( \theta \in [0, \tilde{\theta}] \Rightarrow \)

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

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

\[ \hat{\Omega} = \begin{bmatrix} 0.5920 & 0.0250 \\ 0.0250 & 0.1014 \end{bmatrix} \]

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.
- 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:

\[ Ay_t = Bx_{t-1} + u_t \quad u_t \sim N(0, D) \]

Reduced form:

\[ y_t = \Phi x_{t-1} + \varepsilon_t \quad \varepsilon_t \sim \text{i.i.d. } N(0, \Omega) \]

\[ \varepsilon_t = A^{-1} u_t \]

\[ E(\varepsilon_t \varepsilon_t') = \Omega \quad A\Omega A' = D \]
Structural model:

\[ Ay_t = \lambda + B_1 y_{t-1} + \cdots + B_m y_{t-m} + u_t \]

\[ u_t \sim \text{i.i.d. } N(0, D) \quad D \text{ diagonal} \]

Intuition for results that follow:

If we knew row \( i \) of \( A \) (denoted \( a'_i \)),

then we could estimate coefficients for

\( i \)th structural equation (\( b_i \)) by

\[ \hat{b}_i = \left( \sum_{t=1}^{T} x_{t-1}x'_{t-1} \right)^{-1} \left( \sum_{t=1}^{T} x_{t-1}y'_t a_i \right) = \hat{\Phi}_T a_i \]

\[ \hat{d}_{ii} = T^{-1} \sum_{t=1}^{T} \hat{u}_t^2 = a'_i \hat{\Omega}_T a_i \quad \hat{D} = \text{diag}(A\hat{\Omega}_T A') \]
Consider Bayesian approach where we begin with arbitrary prior \( p(A) \).

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

\[
A = \begin{bmatrix}
-\beta^d & 1 \\
-\alpha^s & 1
\end{bmatrix}
\]
\( p(A) \) 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: \( \kappa_i, \tau_i \to 0 \)
\[
\mathbf{B} = \begin{bmatrix}
\lambda & \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_m
\end{bmatrix}
\]

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

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

uninformative priors: \(\mathbf{M}_i^{-1} \rightarrow 0\)
Likelihood:

\[ p(Y_T|A, D, B) = (2\pi)^{-Tn/2} |\text{det}(A)|^T |D|^{-T/2} \times \]
\[ \exp \left[ -(1/2) \sum_{t=1}^{T} (Ay_t - Bx_{t-1})'D^{-1}(Ay_t - Bx_{t-1}) \right] \]

prior:

\[ p(A, D, B) = p(A)p(D|A)p(B|A, D) \]

posterior:

\[ p(A, D, B|Y_T) = \frac{p(Y_T|A,D,B)p(A,D,B)}{\int p(Y_T|A,D,B)p(A,D,B)dAdDdB} \]
\[ = p(A|Y_T)p(D|A, Y_T)p(B|A, D, Y_T) \]
Exact Bayesian posterior distribution (all $T$):

$$b_i|A, D, Y_T \sim N(m_i^*, d_{ii}M_i^*)$$

$$\tilde{Y}_i' = (a_i'y_1, \ldots, a_i'y_T, m_i'P_i)$$

$$\tilde{X}_i' = \begin{bmatrix} x_0 & \cdots & x_{T-1} & P_i \end{bmatrix}$$

$$m_i^* = (\tilde{X}_i'\tilde{X}_i)^{-1} (\tilde{X}_i'\tilde{y}_i)$$

$$M_i^* = (\tilde{X}_i'\tilde{X}_i)^{-1} P_i P_i' = M_i^{-1}$$

If uninformative prior ($M_i^{-1} = 0$) then $m_i^{*'} = a_i'\hat{\Phi}_T$
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

$\mathbf{m}_i^* \xrightarrow{p} \left[ E(\mathbf{x}_{t-1}\mathbf{x}_{t-1}') \right]^{-1} E(\mathbf{x}_{t-1}\mathbf{y}_t') \mathbf{a}_i = \Phi_0^' \mathbf{a}_i$

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

$\mathbf{b}_i | \mathbf{A}, \mathbf{D}, \mathbf{Y}_T \xrightarrow{p} \Phi_0^' \mathbf{a}_i$
Posterior distribution for $D|A$

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

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

If $M_i^{-1} = 0$, $\zeta_i^* = T a_i'\hat{\Omega} T a_i$

$$\hat{\Omega}_T = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}_t, \quad \hat{\varepsilon}_t = y_t - \hat{\Phi}x_{t-1}$$

($\hat{\varepsilon}_t$ 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

$$\zeta_i^*/T \xrightarrow{p} \mathbf{a}_i \mathbf{\Omega}_0 \mathbf{a}_i$$

$$\mathbf{\Omega}_0 = E(\mathbf{y}_t \mathbf{x}_{t-1}^t) - E(\mathbf{y}_t \mathbf{x}_{t-1}^t) \{E(\mathbf{x}_t \mathbf{x}_t^t)\}^{-1} E(\mathbf{x}_{t-1} \mathbf{y}_t^t)$$

$$d_i | \mathbf{A}, \mathbf{Y}_T \xrightarrow{p} \mathbf{a}_i \mathbf{\Omega}_0 \mathbf{a}_i$$
Posterior distribution for $\mathbf{A}$

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

$k_T$ = constant that makes this integrate to 1

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

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

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

If evaluated at \( A \) for which
\[ A \hat{\Omega}_{T} A' = \text{diag}(A \hat{\Omega}_{T} A'), \]
\[ p(A|Y_T) = k_T p(A) \]
\[ p(A|Y_T) = \frac{k_T p(A) |\det(A\hat{\Omega} TA')|^{T/2}}{\left\{ \det[\text{diag}(A\hat{\Omega} TA')] \right\}^{T/2}} \]

Hadamard’s Inequality:
If evaluated at \( A \) for which

\[ A\hat{\Omega} TA' \neq \text{diag}(A\hat{\Omega} TA'), \]

\[ \det[\text{diag}(A\hat{\Omega} TA')] > \det(A\hat{\Omega} TA') \]

\[ p(A|Y_T) \to 0 \]
\[ p(A|Y_T) \rightarrow \begin{cases} 
kp(A) & \text{if } A \in S(\Omega_0) \\
0 & \text{otherwise} 
\end{cases} \]

\[ S(\Omega_0) = \{ A : A\Omega_0A' \text{ diagonal} \} \]

\[ \Omega_0 = E(y_t x_{t-1}') - E(y_t x_{t-1}') \{ E(x_t x_t') \}^{-1} E(x_{t-1} y_t') \]
Special case: if model is point-identified (so that $S(\Omega)$ consists of a single point), then posterior distribution converges to a point mass at true $\Lambda$.
Measure distance $q(\mathbf{A}, \Omega)$ between $\mathbf{A}$ and $S(\Omega)$ by sum of squares of off-diagonal elements of Cholesky factor of $\mathbf{A}\Omega\mathbf{A}'$:

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

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

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

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

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

\[
\int_{\mathbf{A} \in H_\delta(\Omega_0)} p(\mathbf{A}) d\mathbf{A} > 0 \text{ for all } \delta > 0,
\]

then $\text{Prob}[\mathbf{A} \in H_\delta(\Omega_0)|\mathbf{Y}_T] \to 1$ for all $\delta > 0$. 
Application: Labor market dynamics

demand:
\[ \Delta n_t = k^d + \beta^d \Delta w_t + b^d_{11} \Delta w_{t-1} + b^d_{12} \Delta n_{t-1} + b^d_{21} \Delta w_{t-2} \]
\[ + b^d_{22} \Delta n_{t-2} + \cdots + b^d_{m1} \Delta w_{t-m} + b^d_{m2} \Delta n_{t-m} + u^d_t \]
supply:
\[ \Delta n_t = k^s + \alpha^s \Delta w_t + b^s_{11} \Delta w_{t-1} + b^s_{12} \Delta n_{t-1} + b^s_{21} \Delta w_{t-2} \]
\[ + b^s_{22} \Delta n_{t-2} + \cdots + b^s_{m1} \Delta w_{t-m} + b^s_{m2} \Delta n_{t-m} + u^s_t \]
For fixed $\alpha^s$, MLE of $\beta^d$ 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 $\beta(\alpha)$
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 $\alpha$ and $\beta$

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

since $h_L > 0$, this restricts $\alpha$, not $\beta$
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 $\hat{\varepsilon}_{2t}$

$= \text{convex combination of } \alpha^{-1} \text{ and } \beta^{-1}$

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

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

$\Rightarrow h_L < \alpha < h_H$

$\beta \in (-\infty, 0]$
Contours for log likelihood

$H_{-5}^{-4} -3 -2 -1 0 1 2 3 4 5$

$\alpha$

$h_H$
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 $\beta$: Student $t$ with location $c_\beta$, scale $\sigma_\beta$, d.f. $\nu_\beta$, 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 $\alpha$: Student $t$ with location $c_\alpha$, scale $\sigma_\alpha$, d.f. $\nu_\alpha$, truncated by $\alpha \geq 0$

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

$\Rightarrow \text{Prob}(\alpha < 0.1) = 0.05$

$\text{Prob}(\alpha > 2.2) = 0.05$
Could we also use information about long-run labor supply elasticity?

\[ \Delta \tilde{y}_t = (\Delta w_t, \Delta n_t)' \]

(data used for \( y_t \) in VAR as estimated)

\[ \tilde{y}_t = (w_t, n_t)' \]

(data in levels)

\[ u_t = (u_t^d, u_t^s) \]

(vector of structural shocks)
\[
\frac{\partial \Delta \tilde{y}_{t+s}}{\partial u'_t} = \frac{\partial \Delta \tilde{y}_{t+s}}{\partial \varepsilon'_t} \frac{\partial \varepsilon_t}{\partial u'_t} = \Psi_s A^{-1}
\]

\[
(\Psi_0 + \Psi_1 L + \Psi_2 L^2 + \cdots)
\]

\[
= (I_n - \Phi_1 L - \Phi_2 L^2 - \cdots - \Phi_m L^m)^{-1}
\]

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

\[
= \Psi_s A^{-1} + \Psi_{s-1} A^{-1} + \cdots + \Psi_0 A^{-1}
\]
\[
\frac{\partial \tilde{y}_{t+s}}{\partial u'_t} = \Psi_s A^{-1} + \Psi_{s-1} A^{-1} + \cdots + \Psi_0 A^{-1}
\]

\[
\lim_{s \to \infty} \frac{\partial \tilde{y}_{t+s}}{\partial u'_t} = (\Psi_0 + \Psi_1 + \Psi_2 + \cdots) A^{-1}
\]

\[
= (I_n - \Phi_1 - \Phi_2 - \cdots - \Phi_m)^{-1} A^{-1}
\]

\[
= [A(I_n - \Phi_1 - \Phi_2 - \cdots - \Phi_m)]^{-1}
\]

\[
= [A - B_1 - B_2 - \cdots - B_m]^{-1}
\]
\[
\lim_{s \to \infty} \frac{\partial \tilde{y}_{t+s}}{\partial u_t'} = \left[ A - B_1 - B_2 - \cdots - B_m \right]^{-1}
\]

Labor demand shock (shock #1) has zero long run effect on employment (second element of \( \tilde{y}_{t+s} \)) if and only if (2, 1) element is zero:

\[
0 = -\alpha_s^s - b_{11}^s - b_{21}^s - \cdots - b_{m1}^s
\]
\[0 = -\alpha^s - b_{11}^s - b_{21}^s - \cdots - b_{m1}^s\]

**Usual approach:** impose this condition as untestable identifying assumption

**Our suggestion:** instead represent as prior belief,

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

\[V = 0.1 \Rightarrow \text{prior given same weight as 10 observations on } \mathbf{y}_t\]
Prior and posterior distributions for short-run elasticities and long-run impact

\[ \beta^d \]

\[ \alpha^s + b_{11}^s + b_{21}^s + \ldots + b_{m1}^s \]
Posterior medians and 95% credibility regions for structural impulse-response functions
Response of employment to labor demand shock

\[ V = 1 \]

\[ V = 0.1 \]

\[ V = 0.01 \]

\[ V = 0.001 \]

\[ \alpha^g \]

\[ V = 1 \]

\[ V = 0.1 \]

\[ V = 0.01 \]

\[ V = 0.001 \]
Application 2: Shocks to oil supply and demand

$q = \text{quantity of oil produced}$

$y = \text{measure of economic activity}$

$p = \text{real price of oil}$
oil supply:
\[ q_t = \alpha_{qy}y_t + \alpha_{qp}p_t + b_1'x_{t-1} + u_{1t} \]

economic activity:
\[ y_t = \alpha_{yq}q_t + \alpha_{yp}p_t + b_2'x_{t-1} + u_{2t} \]

inverse of oil demand curve:
\[ p_t = \alpha_{pq}q_t + \alpha_{py}y_t + b_3'x_{t-1} + u_{3t} \]

Note: \( \alpha_{pq} \) = inverse of short-run price-elasticity of oil demand
\[Ay_t = \lambda + B_1 y_{t-1} + \cdots + B_m y_{t-m} + u_t\]

\[A = \begin{bmatrix}
1 & -\alpha_{qy} & -\alpha_{qp} \\
-\alpha_{yq} & 1 & -\alpha_{yp} \\
-\alpha_{pq} & -\alpha_{py} & 1
\end{bmatrix}\]
A Bayesian interpretation of traditional identification

Kilian (2009): Cholesky identification

\[ \alpha_{qy} = \alpha_{qp} = \alpha_{yp} = 0 \]

oil supply:

\[ q_t = \alpha_{qy} y_t + \alpha_{qp} p_t + b_1' x_{t-1} + u_{1t} \]

economic activity:

\[ y_t = \alpha_{yq} q_t + \alpha_{yp} p_t + b_2' x_{t-1} + u_{2t} \]

inverse of oil demand curve:

\[ p_t = \alpha_{pq} q_t + \alpha_{py} y_t + b_3' x_{t-1} + u_{3t} \]
Bayesian translation: I put absolutely zero possibility on any $A$ 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): $p(\alpha_{yq}) \sim \text{Student t with location 0, scale 100, d.f. = 3}$
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 $\mathbf{A}$