A Bayesian Interpretation of Traditional Approaches to Identification

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

Identification Strategy

- Traditional approach:

  assume perfect knowledge of structure to achieve identification

- Bayesian approach:

  represent imperfect information about elements in **A** in the form of a Bayesian prior distribution $p(\mathbf{A})$

  ⇨ generalization of traditional methods

How does this relate to non-Bayesian approaches?

Traditional point-identified structural VARs are a special case of a Bayesian prior that is dogmatic.

$\Rightarrow$ Application 5: Recursive identification

Example: Kilian AER (2009)

q_t = world oil production

y_t = real global economic activity

p_t = real price of oil

Structural Model of the World Oil Market

- Dynamic oil supply equation:

$$q_t = \gamma y_t + \alpha p_t + \mathbf{b}_s' \mathbf{x}_{t-1} + u_t^s$$

- Determinants of world economic activity:

$$y_t = \xi q_t + \psi p_t + \mathbf{b}_y' \mathbf{x}_{t-1} + u_t^y$$

- Dynamic oil demand equation:

$$q_t = \delta y_t + \beta p_t + \mathbf{b}_d' \mathbf{x}_{t-1} + u_t^d$$

Put in Canonical Form

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

$$\mathbf{A} = \begin{bmatrix} 1 & -\gamma & -\alpha \\ -\xi & 1 & -\psi \\ 1 & -\delta & -\beta \end{bmatrix}$$

What Does Cholesky Identification Imply?

$$\alpha = \gamma = \psi = 0$$

supply: $q_t = \gamma y_t + \alpha p_t + \mathbf{b}_s' \mathbf{x}_{t-1} + u_t^s$

econ activity: $y_t = \xi q_t + \psi p_t + \mathbf{b}_y' \mathbf{x}_{t-1} + u_t^y$

inverse demand: $p_t = \widetilde{\delta} y_t + \widetilde{\beta} q_t + \mathbf{b}_d' \mathbf{x}_{t-1} + u_t^d$

Amounts to the claim:

Oil producers take longer than one month to respond to price or income ($\gamma = \alpha = 0$)

Oil prices take longer than one month to affect economic activity ($\psi = 0$)

Bayesian Translation of Cholesky Identification

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

  $\Rightarrow$ dogmatic prior: degenerate distribution

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

  $\Rightarrow$ totally uninformative prior

Special Case of Bayesian Prior Beliefs

$$\widetilde{\mathbf{A}} = \begin{bmatrix} 1 & 0 & 0 \\ -\xi & 1 & 0 \\ -\widetilde{\beta} & -\widetilde{\delta} & 1 \end{bmatrix}$$

$$p(\widetilde{\mathbf{A}}) = p(\xi)p(\widetilde{\beta})p(\widetilde{\delta})$$

How to Represent "No Information"?

- Requirement:

  Prior has to be a proper density

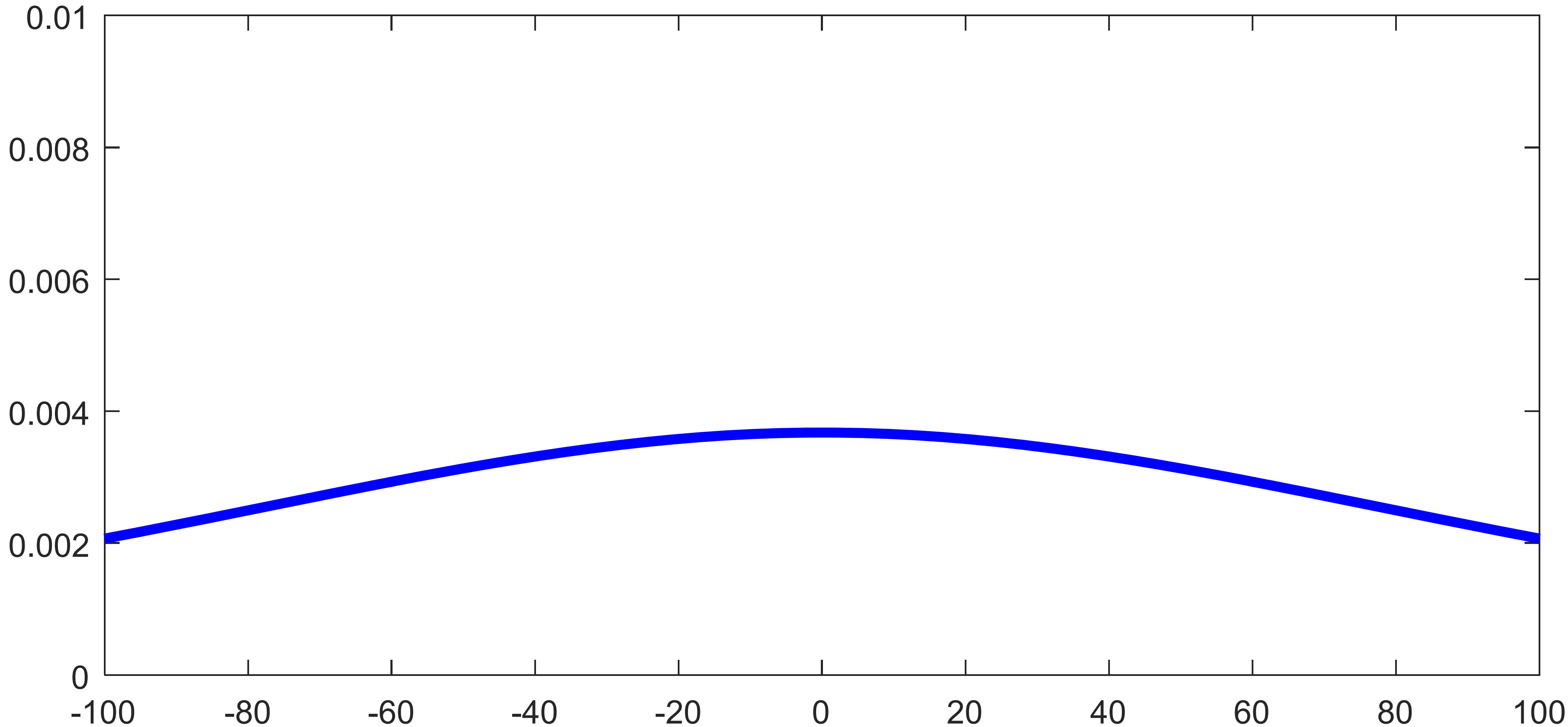
- Proposal:

  (2,1) element: $p(\xi) \sim$ Student t with location 0,

  scale 100, d.f. $= 3$

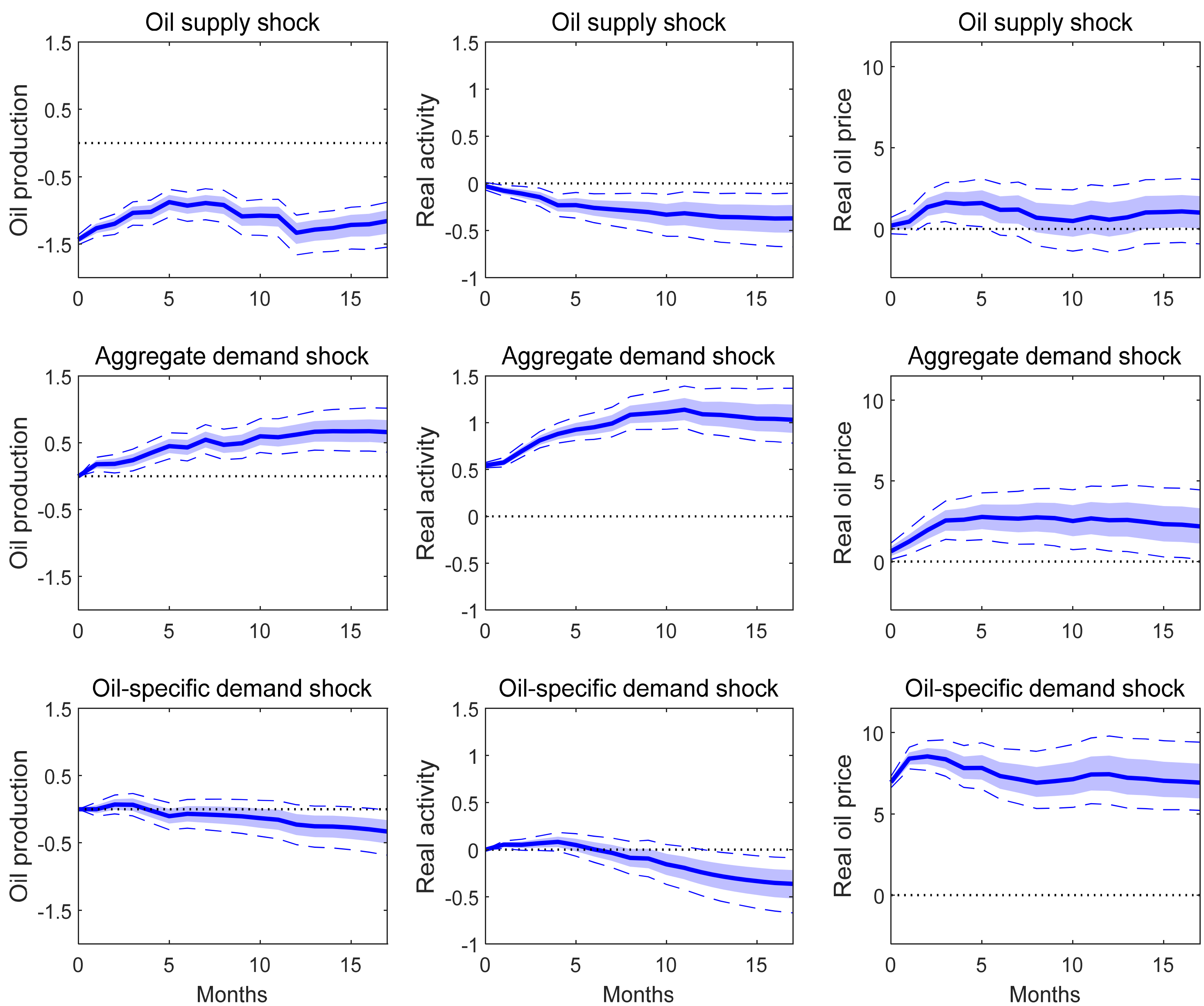
Same for $p(\widetilde{\beta})$ and $p(\widetilde{\delta})$

Prior for Lower-Triangular Elements of $\tilde{\mathbf{A}}$

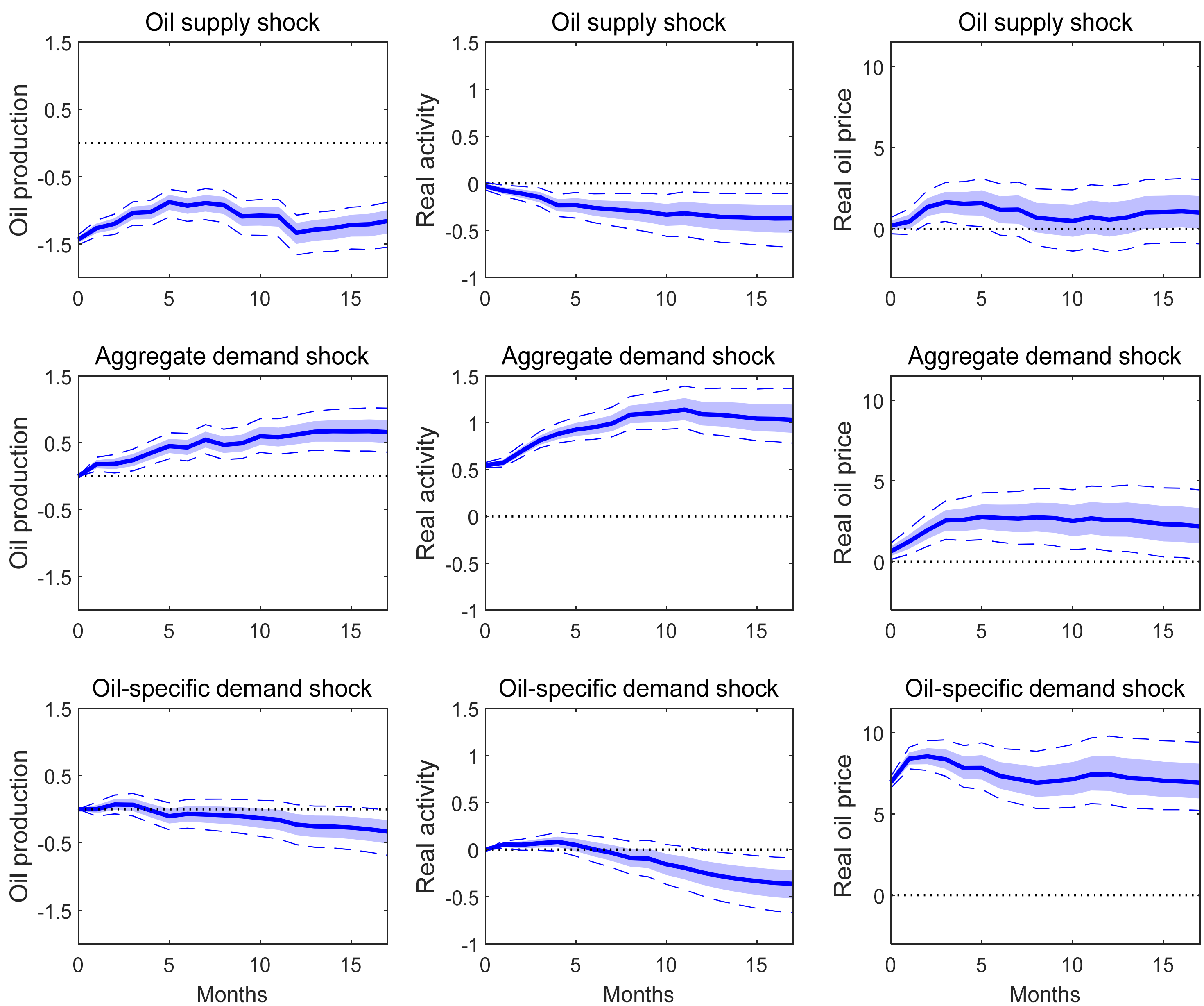


$p(\xi)$: If oil production goes up by 1% this month, maybe economic activity goes up by 100%, maybe it goes down by 100%.

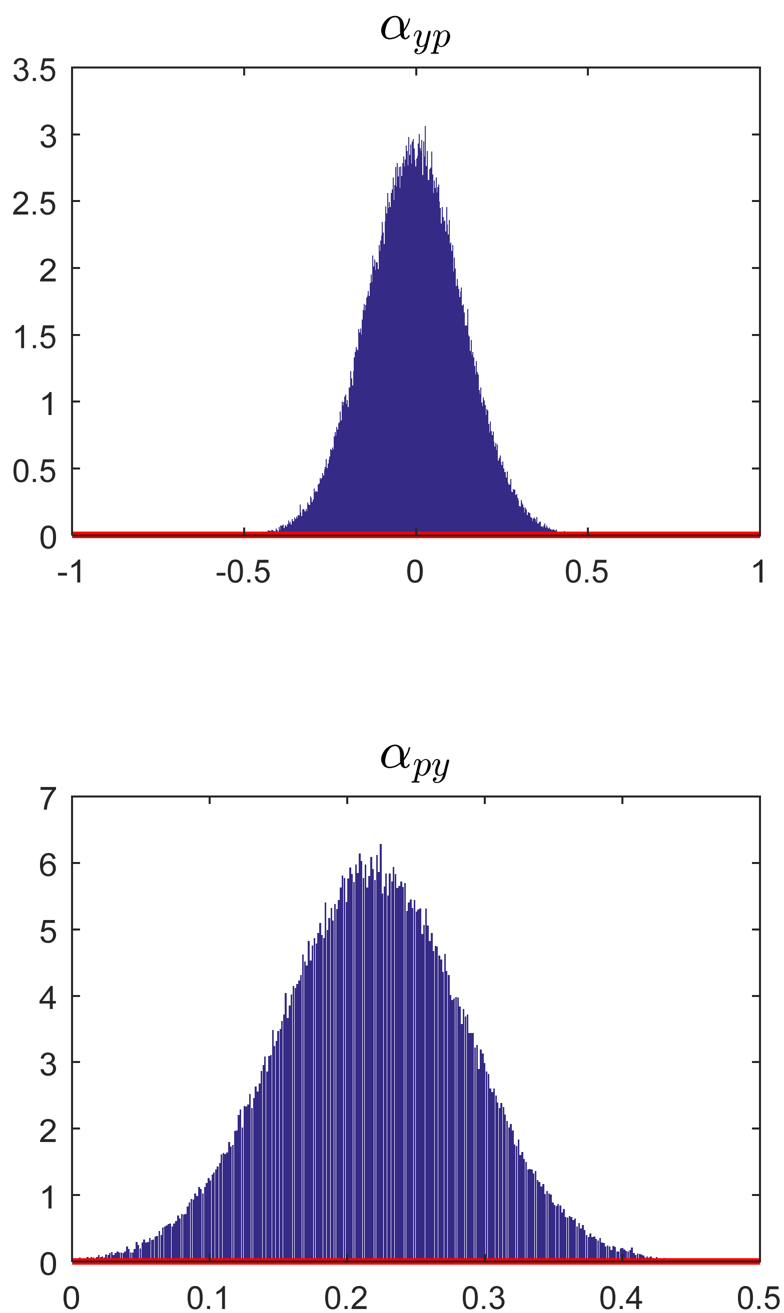
Structural IRF of Bayesian who was dogmatic about some and uninformed about other parameters



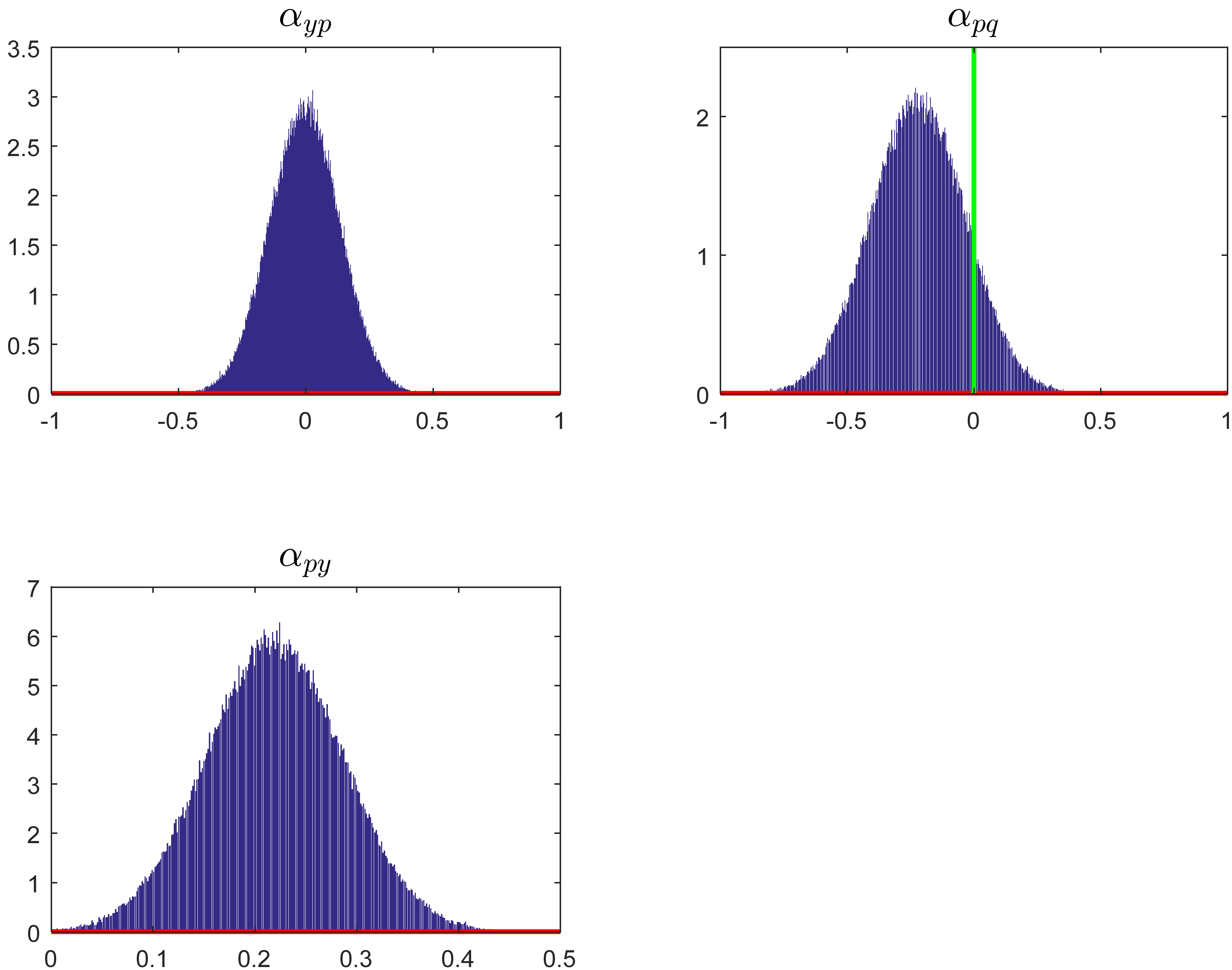
Structural IRF of frequentist who used Cholesky identifying assumptions



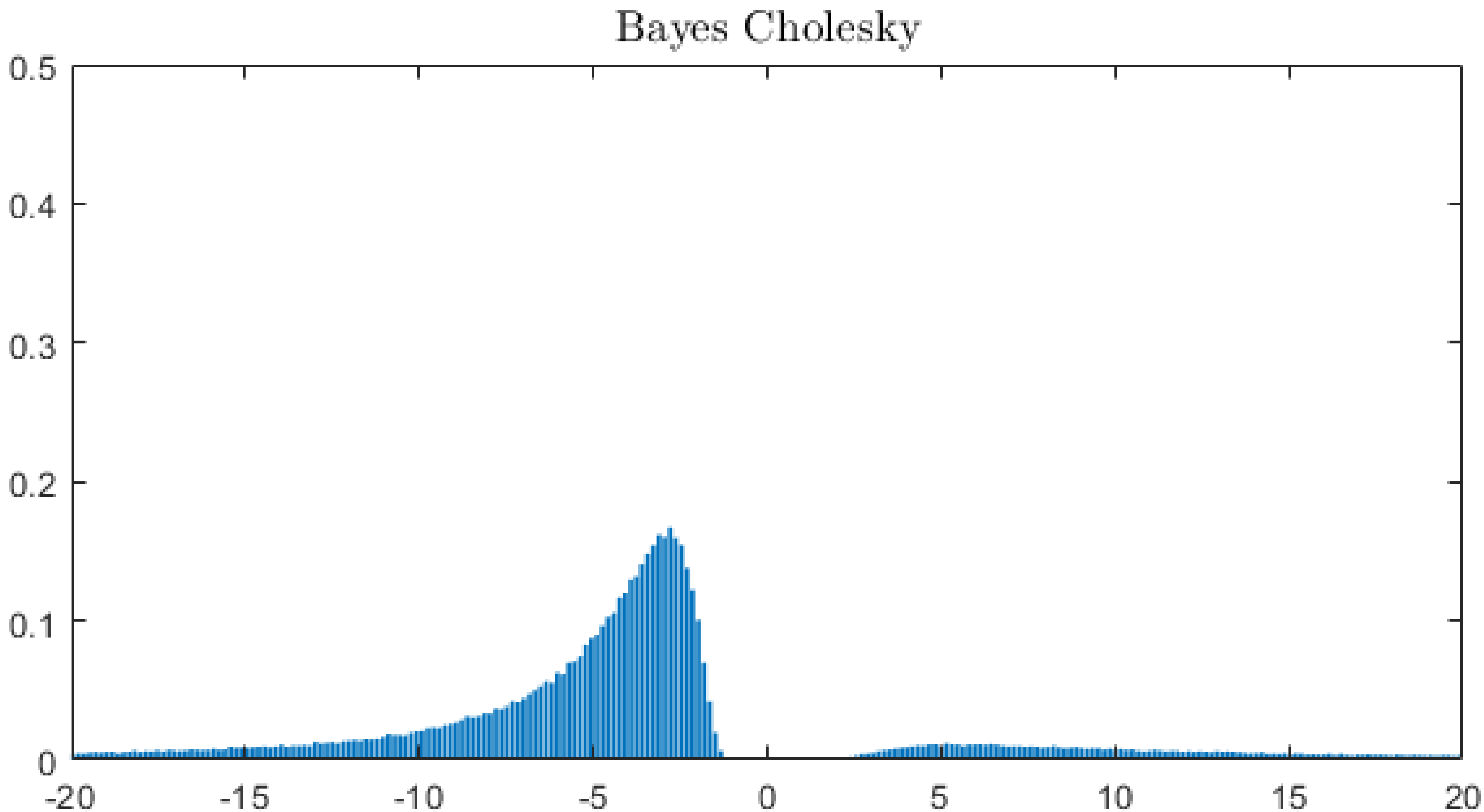
Prior (red) and posterior (blue) distributions for unknown elements of $\widetilde{\mathbf{A}}$



Prior (red) and posterior (blue) distributions for unknown elements of $\widetilde{\mathbf{A}}$



Bayesian posterior distribution of short-run demand elasticity β



17.5% probability that $\beta > 0$

97% probability that $|\beta| > 2$

Why is the demand elasticity so large?

- Cholesky point estimate: $\beta = \frac{1}{\tilde{\beta}} = -5.9562$
- A 10% increase in the price of oil with no change in income would result in a 60% drop in consumption within the month.
- $\alpha_{qp} = 0$ plays a key role in this conclusion.
- Suppose that $\alpha > 0$, positive correlation between demand shock and q_t would bias OLS estimate upward (closer to zero)

$\Rightarrow$ implies bias of estimated demand toward larger absolute value (Baumeister and Hamilton, ET 2024)

⇨ casts doubt on the Cholesky assumption of $\alpha = 0$

A Bayesian generalization of traditional identification

Consider next a Bayesian who is extremely confident (but not absolutely certain) that supply elasticity is very small:

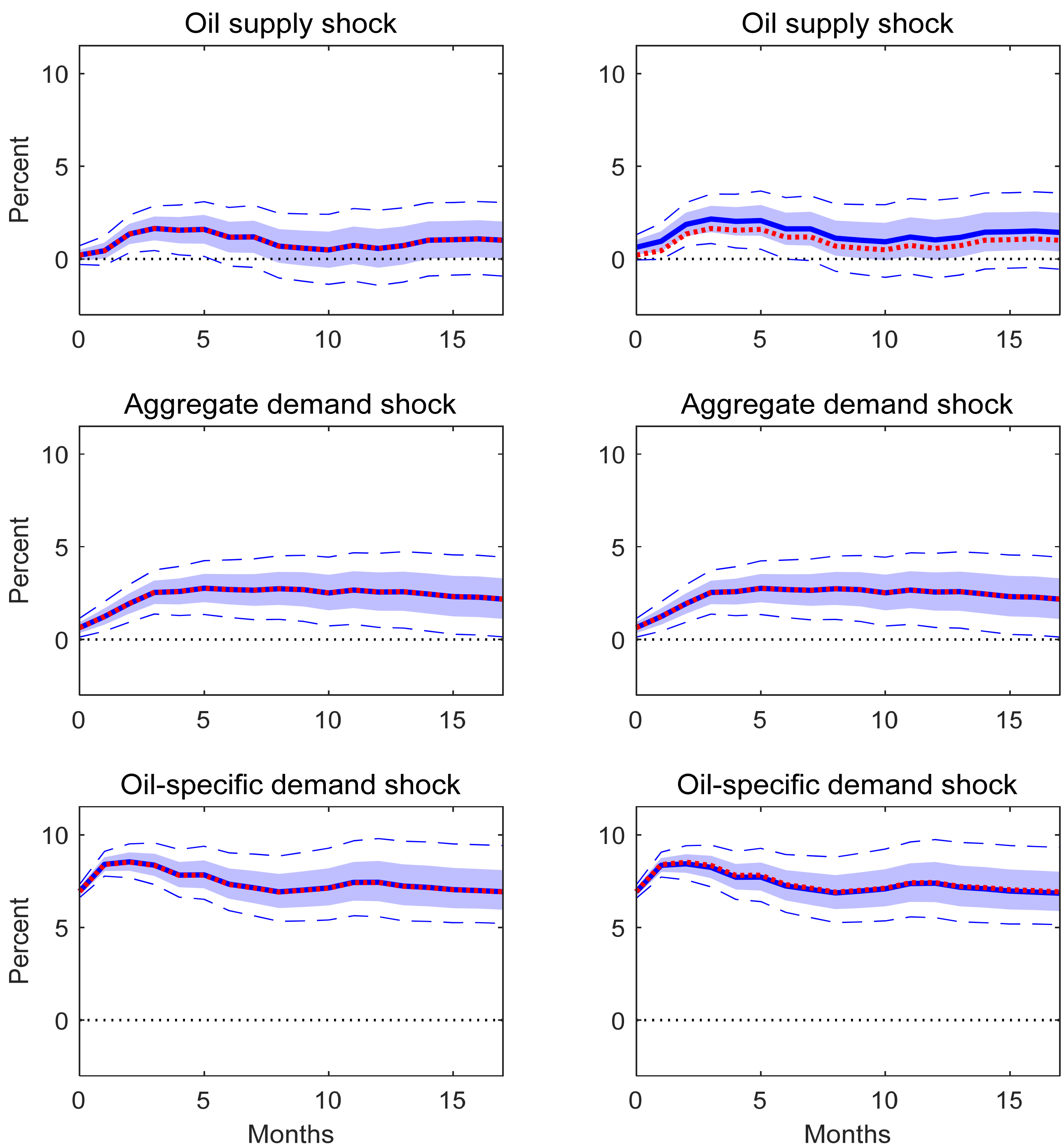
$p(\alpha) \sim U(0, 0.025)$

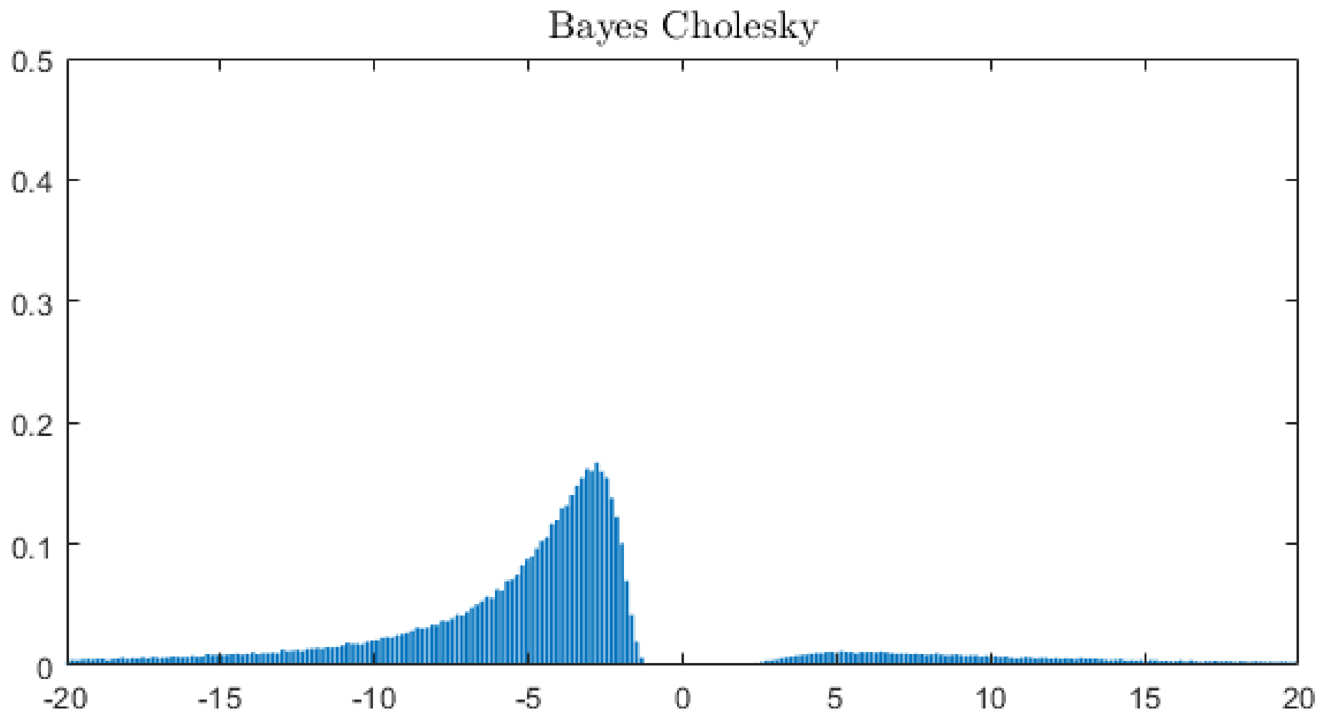
$$p(\mathbf{A}) \propto \left[1 + \frac{1}{\nu}\left(\frac{\xi}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}} \left[1 + \frac{1}{\nu}\left(\frac{\delta}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}} \times$$

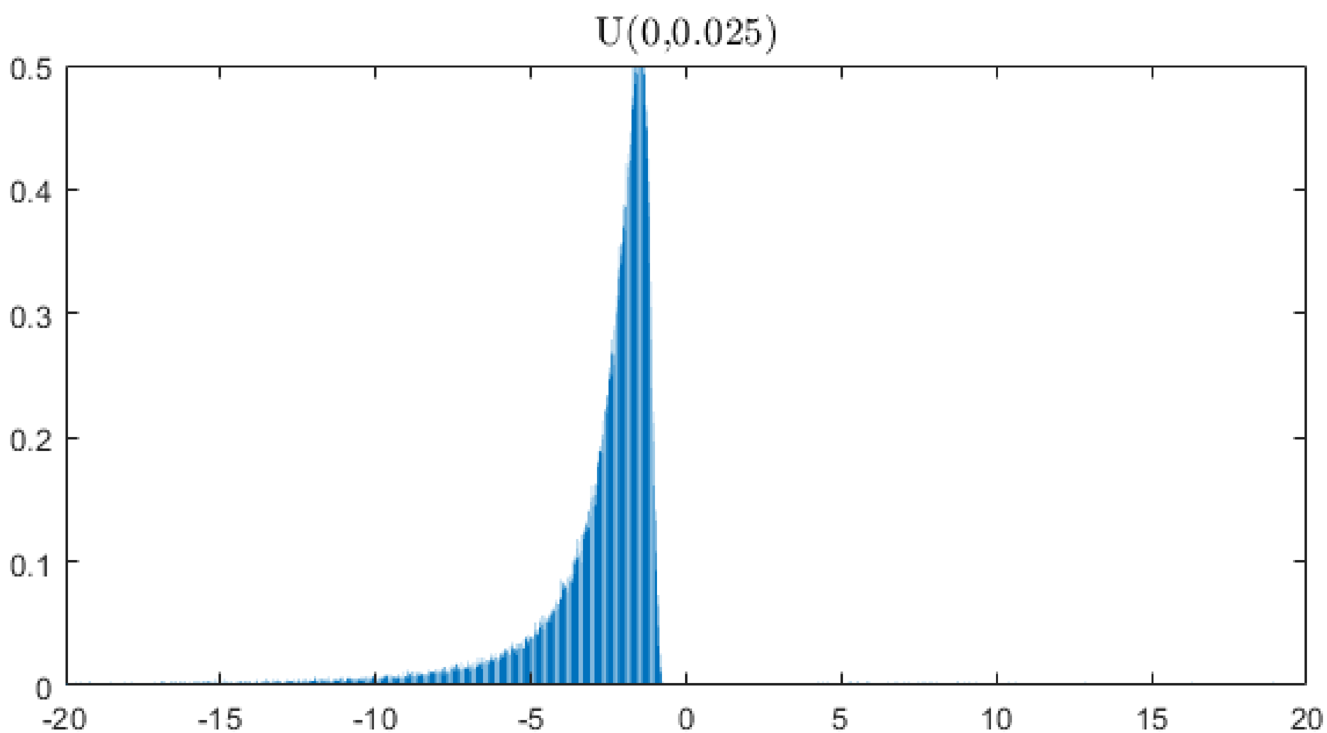
$$\left[1 + \frac{1}{\nu}\left(\frac{\beta}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}} \text{ if } \alpha \in [0, 0.025]$$

$p(\mathbf{A}) = 0$ otherwise

$\alpha = 0$ $\qquad$ $\alpha \sim U(0,0.025)$

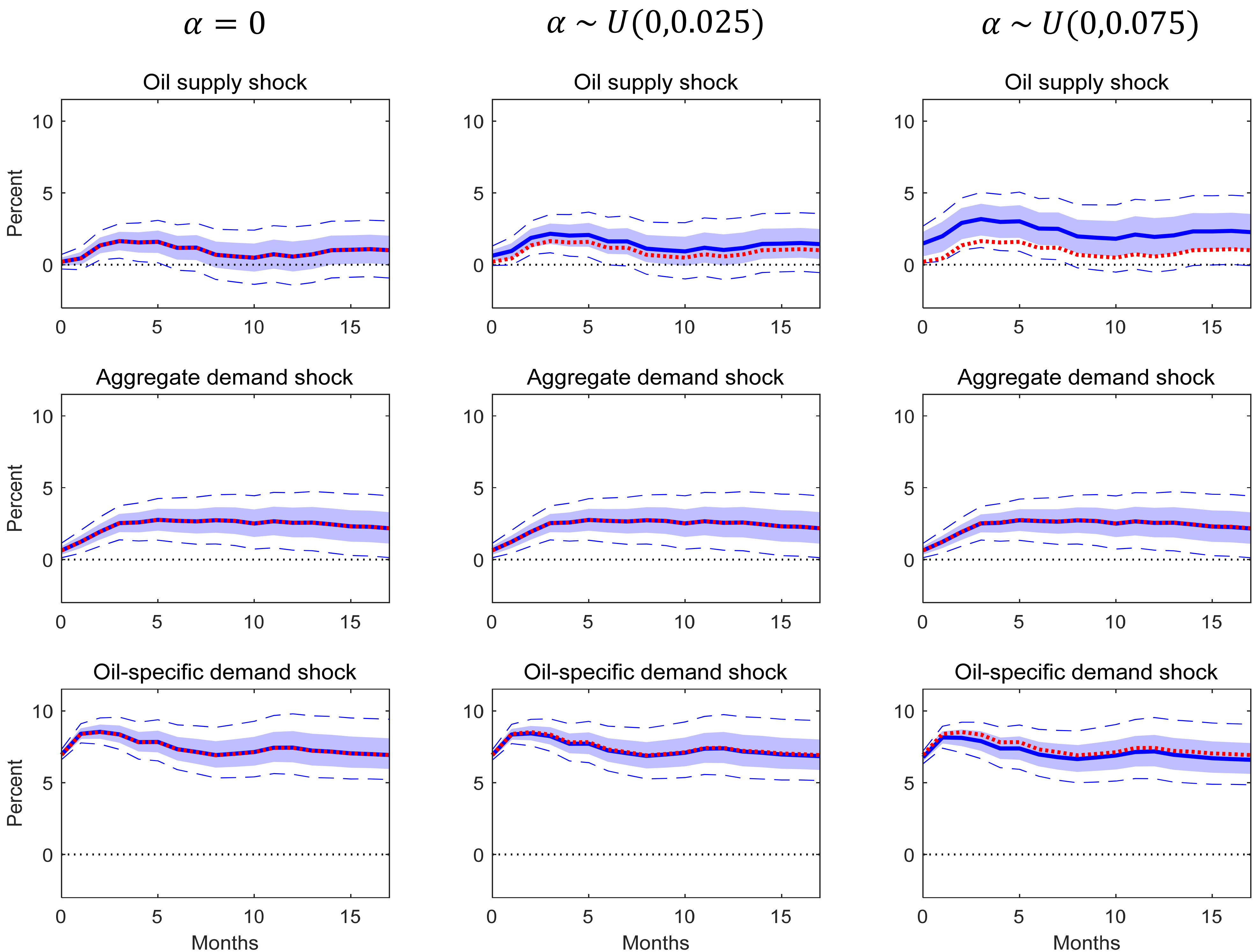



Bayes Cholesky
0.5
0.4
0.3
0.2
0.1
0
-20
-15
-10
-5
0
5
10
15
20


U(0,0.025)
0.5
0.4
0.3
0.2
0.1
0
-20
-15
-10
-5
0
5
10
15
20

Suppose we relax further:

$$p(\alpha) \sim U(0, 0.075)$$



$\alpha = 0$
$\alpha \sim U(0,0.025)$
$\alpha \sim U(0,0.075)$
Oil supply shock
Aggregate demand shock
Oil-specific demand shock
Percent
Months

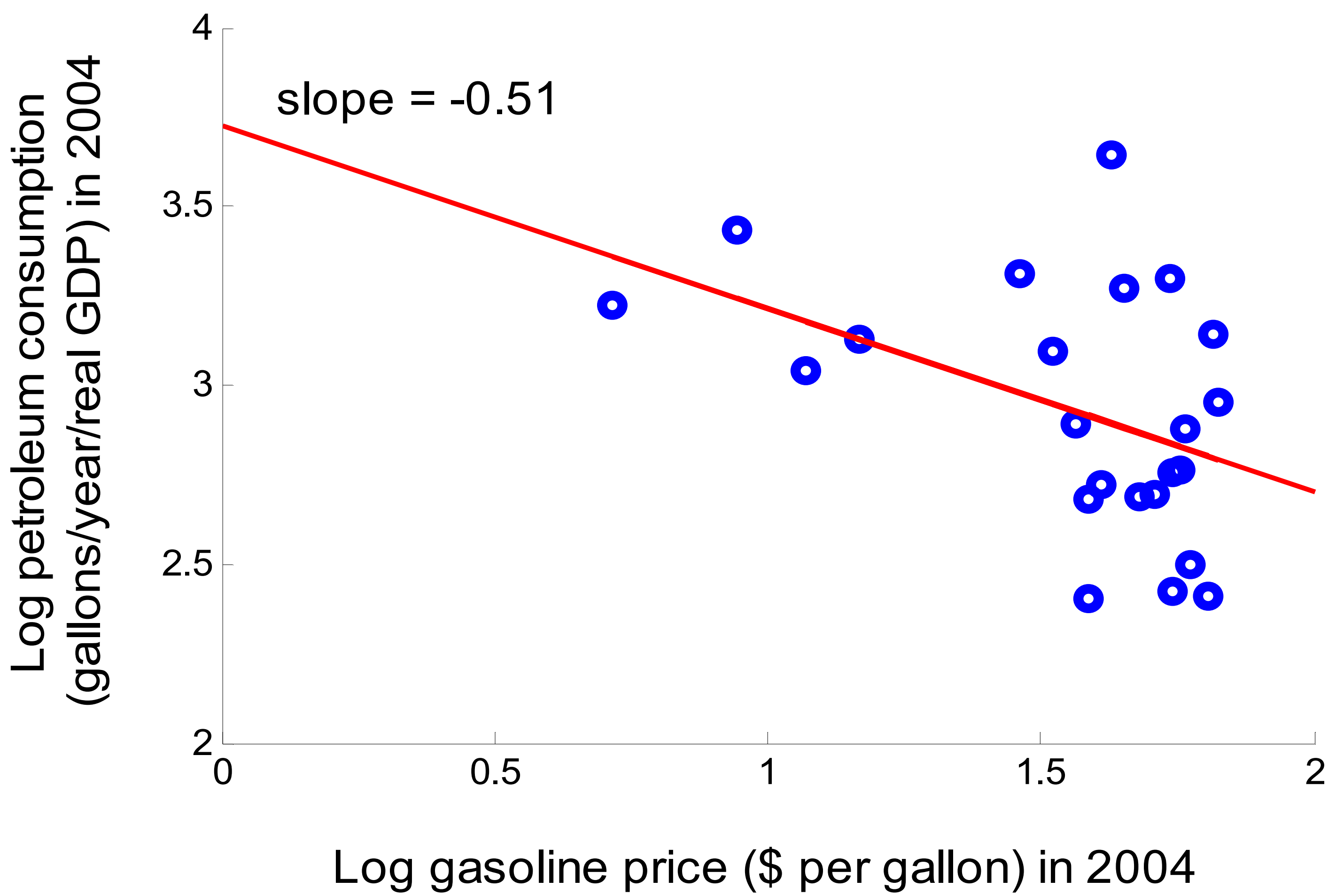
- As we become less certain, posterior credible sets widen substantially and we lose confidence in structural conclusions.

- Could compensate in part by also using prior information about demand elasticity β (and other parameters).

⇨ Bring in inexact information from multiple sources rather than claiming to have exact prior knowledge about a few parameters

What do we know about the price elasticity of demand?

- Cross-country regression of log of petroleum use per dollar of GDP on real price of gasoline for 23 OECD countries in 2004



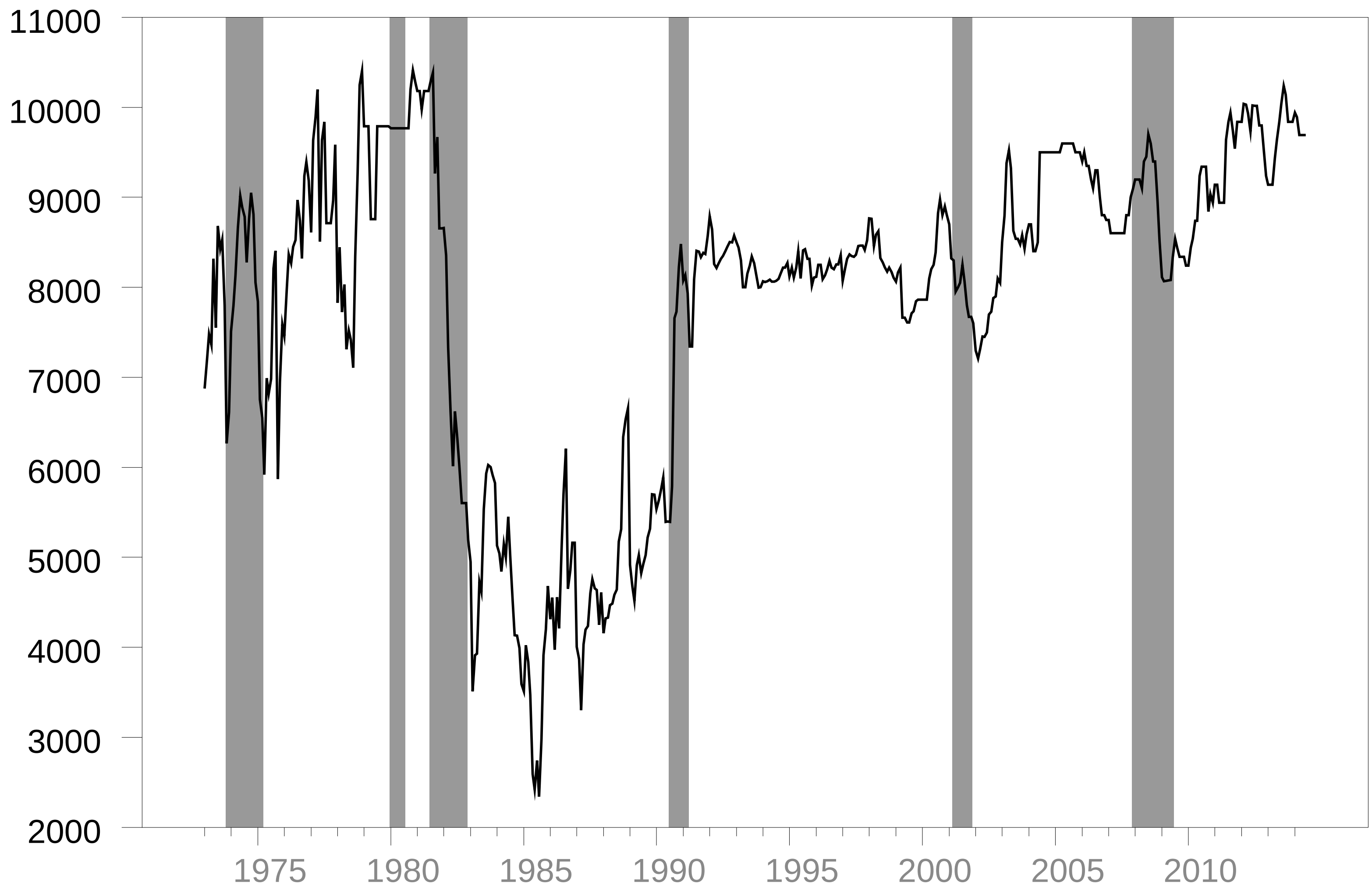
What do we know about the price elasticity of demand?

- Cross-sectional evidence based on household surveys
  - Newey and Hausman (1995): -0.81
  - Yatchew and No (2001): -0.9

- Estimates from previous literature surveys:
  - Dahl and Sterner (1991): -0.86
  - Graham and Glaister (2004): -0.77
  - Brons et al. (2008): -0.84

⇒ short-run elasticity < long-run elasticity

What about the price elasticity of supply?



Saudi Arabian oil production can respond quickly to changing economic conditions:
short-run supply elasticity is not zero

What else do we know about the price elasticity of oil supply?

- Estimates from multiple historical episodes
  - Caldara, Cavallo, and Iacoviello (JME 2019): 0.081
- Estimates from individual oil wells in North Dakota
  - Bjørnland, Nordvik, and Rohrer (JAE 2021): 0.3-0.9
- Estimates from micro data of shale producers
  - Aastveit, Bjørnland, and Gundersen (2021): 0.62

Objects of Structural Interest

BH recommendation: use literature and other data sets as source of prior information about structural parameters $\mathbf{A}$

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

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

Most applications of structural VARs focus instead on impacts of structural shocks $\mathbf{H}$

$$\mathbf{H} = \mathbf{A}^{-1}\mathbf{D}^{1/2}$$

- Uhlig (Adv Econometrics 2017): **H** is better focus because we care about equilibrium effects of policy interventions.

- However, the question of identification comes down to what we know from prior evidence or theory.

Sources of Prior Information

- Rows of **A** correspond to behavior of individual agents (e.g. consumers, producers, govt policy)

- Prior information in the form of:
  - Elasticities (Baumeister and Hamilton 2019; Aastveit et al., 2020; Brinca et al., 2021)
  - Policy rules (Baumeister and Hamilton 2018; Nguyen 2019; Belongia and Ireland, 2021)
  - Behavioral equations from economic theory (Aruoba et al. 2022; Lukmanova and Rabitsch 2021)

- Rows of **H** correspond to the general equilibrium consequences of changes in those agents' behavior
  ⇨ include via composite prior

Take-Away #5

- Better use non-dogmatic priors to reflect doubts that we entertain about identifying assumptions
  ⇒ replace zero restrictions with prior density

- Use all available information to form prior beliefs

- Bayesian posterior distribution incorporates both
  - uncertainty from a finite data set *and*
  - incomplete confidence in identifying assumptions

A Common Error in Estimating Elasticities

- Some studies have tried to calculate behavioral elasticities using the ratios of the elements of a single column of **H** (instead of correct formula that would use inverse of **H**).
- Example: Calculate price elasticity of demand as ratio of change in oil consumption to change in price that results from a shock to supply.
- Studies that do this:
  - Kilian and Murphy (2012, 2014), Güntner (2014), Riggi and Venditti (2015), Kilian and Lütkepohl (2017), Ludvigson et al. (2017), Antolín-Díaz and Rubio-Ramírez (2018), Basher et al. (2018), Herrera and Rangaraju (2020), Zhou (2020) … … …

Why is this Wrong?

For the 3-equation oil market model, $\mathbf{H} = \mathbf{A}^{-1}\mathbf{D}^{1/2}$ with

$$\mathbf{A}^{-1} = |\mathbf{A}|^{-1}\begin{bmatrix} -\beta - \delta\psi & \alpha\delta - \beta\gamma & \alpha + \gamma\psi \\ -\psi - \beta\xi & \alpha - \beta & \psi + \alpha\xi \\ \delta\xi - 1 & \delta - \gamma & 1 - \gamma\xi \end{bmatrix}$$

Why is this Wrong?

For the 3-equation oil market model, $\mathbf{H} = \mathbf{A}^{-1}\mathbf{D}^{1/2}$ with

$$\mathbf{A}^{-1} = |\mathbf{A}|^{-1}\begin{bmatrix} -\beta-\delta\psi & \alpha\delta-\beta\gamma & \alpha+\gamma\psi \\ -\psi-\beta\xi & \alpha-\beta & \psi+\alpha\xi \\ \delta\xi-1 & \delta-\gamma & 1-\gamma\xi \end{bmatrix}$$

Thus, the ratio of impact responses amounts to:

$$\frac{h_{11}}{h_{31}} = \frac{-\beta-\delta\psi}{\delta\xi-1} \neq \beta$$

$$
\begin{aligned}
&\underbrace{\beta}_{\text{response to price}} \underbrace{\sqrt{d_{11}}|\mathbf{A}|^{-1}(\delta\xi - 1)}_{\text{change in price}} \\
&+ \underbrace{\delta}_{\text{response to income}} \underbrace{\sqrt{d_{11}}|\mathbf{A}|^{-1}(-\psi - \beta\xi)}_{\text{change in income}} \\
&= \underbrace{\sqrt{d_{11}}|\mathbf{A}|^{-1}(-\beta - \delta\psi)}_{\text{total change}} .
\end{aligned}
$$

$$
\underbrace{\beta}_{\text{response to price}} \quad \underbrace{\sqrt{d_{11}}|\mathbf{A}|^{-1}(\delta\xi - 1)}_{\text{change in price}}
$$

$$
+ \quad \underbrace{\delta}_{\text{response to income}} \quad \underbrace{\sqrt{d_{11}}|\mathbf{A}|^{-1}(-\psi - \beta\xi)}_{\text{change in income}}
$$

$$
= \underbrace{\sqrt{d_{11}}|\mathbf{A}|^{-1}(-\beta - \delta\psi)}_{\text{total change}} .
$$

- Dividing this by the change in price yields $\frac{-\beta-\delta\psi}{\delta\xi-1}$
- This equals β only in the special case when $\delta = 0$
- When $\delta \neq 0$, this estimate is a combination of sensitivity of demand to price and sensitivity of demand to income.

Take-Away #6

- In a system with $n > 2$ variables, calculating ratios of elements of a single column of **H** could be used to calculate $n - 1$ different "measures" of *each individual elasticity.*

- One could define "price elasticity of demand" to be the ratio of change in consumption to change in price that results from any of the $n - 1$ shocks other than the demand shock.

Example

- Kilian and Murphy (2012, 2014) proposed two different measures of price elasticity of oil supply based on responses of production and price to either of two types of demand shocks.

- If we estimate these two magnitudes using Kilian and Murphy (2014) data and method but without imposing constraint on elasticity, they differ by factor of five.

- Kilian and Murphy force the measures to be close by imposing that both magnitudes have to be smaller than 0.0258.

- The correct frequentist or Bayesian approach results in a unique and optimal estimate of each individual elasticity that is invariant with respect to how the model is parameterized.

⇨ Use inverse of **H**!

What if only a single column of **H** is identified?

- Example: proxy variable/instrument for only one structural shock
- Possible to uncover structural parameters of jth row of **A** by using jth column of **H** together with $\boldsymbol{\Omega}$

  (see Baumeister and Hamilton, ET 2024)
- How?

What if only a single column of **H** is identified?

- Goal: estimate the coefficients of the first structural equation

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

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

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

The first column of $\mathbf{A}^{-1}$ is $\mathbf{h}_1$.

Postmultiply the above expression by $\mathbf{h}_1$

$$\boldsymbol{\Omega}^{-1}\mathbf{h}_1 = \mathbf{A}'\mathbf{D}^{-1}\mathbf{e}_1$$

where $\mathbf{e}_1$ is the first column of $\mathbf{I}_n$.

What if only a single column of **H** is identified?

Given that the first structural shock u_{1t} is uncorrelated with the other structural shocks, $\mathbf{D}^{-1}\mathbf{e}_1 = d_{11}^{-1}\mathbf{e}_1$.

Let $\mathbf{a}_1$ denote the first column of $\mathbf{A}'$, then we have: $\mathbf{\Omega}^{-1}\mathbf{h}_1 = d_{11}^{-1}\mathbf{a}_1$.

- We have estimates for

  $\hat{\mathbf{\Omega}} = T^{-1}\sum_{t=1}^{T}\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t'$ (from OLS VAR residuals)

  $\mathbf{h}_1$ (up to a constant of proportionality g)

Where does $\boldsymbol{h}_1$ come from?

- Example: Have instrument z_t such that

$$E(u_{1t}z_t) = g \neq 0 \text{ (relevance)}$$

$$E(\mathbf{u}_{2t}z_t) = \mathbf{0} \qquad \text{(exogeneity)}$$

$$\boldsymbol{\varepsilon}_t = \mathbf{h}_1 u_{1t} + \mathbf{H}_2 \mathbf{u}_{2t}$$

Multiply by z_t and take expectations

$$E(\boldsymbol{\varepsilon}_t z_t) = \mathbf{h}_1 E(u_{1t}z_t) = \mathbf{h}_1 g$$

$$T^{-1}\sum_{t=1}^{T} \hat{\boldsymbol{\varepsilon}}_t z_t \xrightarrow{p} \mathbf{h}_1 g$$

Estimate of structural parameters

- Estimate $\mathbf{a}_1$ up to an unknown constant:

  Define $\hat{\mathbf{v}}_1 = \hat{\boldsymbol{\Omega}}^{-1}\hat{\mathbf{h}}_1 g = (\sum_{t=1}^{T} \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t')^{-1}(\sum_{t=1}^{T} \hat{\boldsymbol{\varepsilon}}_t z_t)$

  $\Rightarrow$ coefficient from OLS regression of z_t on $\hat{\boldsymbol{\varepsilon}}_t$

- If we normalize one of the coefficients in the first structural equation to be unity, e.g., the coefficient on the fed funds rate in the montary policy rule, then we get an estimate of $\mathbf{a}_1$: $\hat{\mathbf{a}}_1' = \hat{\mathbf{v}}_1' / \hat{\mathbf{v}}_1' \mathbf{e}_1$

Estimate of structural shock

- Can calculate the historical value of the first structural shock at date t as:

$$u_{1t} = \widehat{\mathbf{a}}_1' \hat{\boldsymbol{\varepsilon}}_t$$

Application 6: Taylor Rule Coefficients

- Use Bauer-Swanson (2023) high-frequency instrument for monetary policy shocks z_t^{MP}

- Estimate quarterly VAR(4) consisting of inflation rate (π_t), real GDP growth (y_t) and fed funds/shadow rate (r_t) for period 1955Q4-2019Q4 and compute residuals $\hat{\boldsymbol{\varepsilon}}_t$

- Regress $\hat{\boldsymbol{\varepsilon}}_t$ on z_t^{MP} for 1988Q1-2019Q4 to obtain OLS coefficient $\hat{\boldsymbol{\gamma}}$ (impact effect of instrument)

- Calculate $\hat{\mathbf{v}}_1 = \widehat{\boldsymbol{\Omega}}^{-1}\,\hat{\boldsymbol{\gamma}}$ with $\widehat{\boldsymbol{\Omega}}$ estimated on the full sample and divide $\hat{\mathbf{v}}_1$ by its first element (r_t)