Recent Developments in Modeling Energy Market Dynamics

Christiane Baumeister, University of Notre Dame
James D. Hamilton, UCSD

Overview

- Key role of elasticities in assessing the relative importance of supply-side vs demand-side shocks, in studying their dynamics, and in conducting policy counterfactuals

  - Pitfalls in structural models of the world oil market (Baumeister and Hamilton, *AER* 2019; Baumeister and Hamilton, *Econometric Theory* 2023)

  - Heterogeneous adjustment dynamics across major oil consumers and producers (Baumeister and Hamilton, 2023)

World oil market

- Dynamic demand equation:

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

$$\mathbf{x}_{t-1} = (1, \mathbf{y}_{t-1}', \mathbf{y}_{t-2}', \ldots, \mathbf{y}_{t-m}')' \text{ with } \mathbf{y}_t = (q_t, y_t, p_t)'$$

World oil market

- Dynamic demand equation:

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

$$\mathbf{x}_{t-1} = (1, \mathbf{y}_{t-1}', \mathbf{y}_{t-2}', \ldots, \mathbf{y}_{t-m}')' \quad \text{with } \mathbf{y}_t = (q_t, y_t, p_t)'$$

- Meaning of demand elasticity β:

  If price were to increase by 1% with income held constant, by how much would quantity demanded change?

World oil market

- Dynamic 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 demand equation:

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

Structural model:

$$\mathbf{A}\mathbf{y}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{u}_t \qquad E(\mathbf{u}_t\mathbf{u}_t') = \mathbf{D} \text{ (diagonal)}$$

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

$$\mathbf{B} = [\mathbf{b}_s' \;\; \mathbf{b}_y' \;\; \mathbf{b}_d']'$$

Structural model:

$$\mathbf{A}\mathbf{y}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{u}_t \qquad E(\mathbf{u}_t\mathbf{u}_t') = \mathbf{D} \text{ (diagonal)}$$

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

$$\mathbf{B} = [\mathbf{b}_s' \ \ \mathbf{b}_y' \ \ \mathbf{b}_d']'$$

Reduced form:

$$\mathbf{y}_t = \mathbf{A}^{-1}\mathbf{B}\mathbf{x}_{t-1} + \mathbf{A}^{-1}\mathbf{u}_t$$

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

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

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

Normalization of shocks to unit variance

- $\mathbf{y}_t = \Pi \mathbf{x}_{t-1} + \mathbf{H}\mathbf{u}_t^* \qquad E(\mathbf{u}_t^* \mathbf{u}_t^{*\prime}) = \mathbf{I}_n$

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

  $\mathbf{H} = \mathbf{A}^{-1}\mathbf{D}^{1/2}$ = impact on $\mathbf{y}_t$ of one-standard-deviation structural shocks

- If model just-identified, MLE of **H**: $\hat{\mathbf{H}}_{MLE}\hat{\mathbf{H}}'_{MLE} = \hat{\boldsymbol{\Omega}}_{MLE}$

- MLE estimate of short-run demand elasticity:

$$\hat{\beta}_{MLE} = -\hat{h}^{33}_{MLE}/\hat{h}^{31}_{MLE}$$

  where h^{ij} denotes the (i, j) element of $\mathbf{H}^{-1}$

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 the *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:

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

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

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

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

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

- Dividing this by the change in price yields $\frac{-\beta-\delta\psi}{\delta\xi-1}$
- This equals β 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.

- 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 (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 0.025.

- 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.

Empirical illustration: Cholesky identification

- Recursive structure implies: $\alpha = \gamma = \psi = 0$

Empirical illustration: Cholesky identification

- Recursive structure implies: $\alpha = \gamma = \psi = 0$
- Under this assumption, MLE can be obtained by dividing demand equation by β and rewrite as:

$$
\begin{aligned}
p_t &= \beta^{-1} q_t - (\delta/\beta) y_t - \beta^{-1} \mathbf{b}_d' \mathbf{x}_{t-1} - \beta^{-1} u_t^d \\
&= \tilde{\beta} q_t + \tilde{\delta} y_t + \tilde{\mathbf{b}}_d' \mathbf{x}_{t-1} + \tilde{u}_t^d
\end{aligned}
$$

- Estimate by OLS since error term is uncorrelated with any explanatory variables

Empirical illustration: Cholesky identification

- Recursive structure implies: $\alpha = \gamma = \psi = 0$
- Under this assumption, MLE can be obtained by dividing demand equation by β and rewrite as:

$$\begin{aligned} p_t &= \beta^{-1} q_t - (\delta/\beta) y_t - \beta^{-1} \mathbf{b}_d' \mathbf{x}_{t-1} - \beta^{-1} u_t^d \\ &= \tilde{\beta} q_t + \tilde{\delta} y_t + \tilde{\mathbf{b}}_d' \mathbf{x}_{t-1} + \tilde{u}_t^d \end{aligned}$$

- Estimate by OLS since error term is uncorrelated with any explanatory variables

- $\tilde{\beta}_{MLE} = -0.167893$

  $\Rightarrow$ elasticity estimate: $\hat{\beta} = 1/\tilde{\beta} = -5.9562$

Why is β so large?

- A 10% increase in the price of oil with no change in income would result in a 60% drop in consumption within the month.

Why is β so large?

- A 10% increase in the price of oil with no change in income would result in a 60% drop in consumption within the month.
- $\alpha = 0$ plays a key role in this conclusion.
- Suppose that $\alpha > 0$, positive correlation between $\tilde{u}_t^d$ and q_t would bias OLS estimate upward (closer to zero) $\Rightarrow$ implies bias of estimated demand elasticity toward larger absolute value

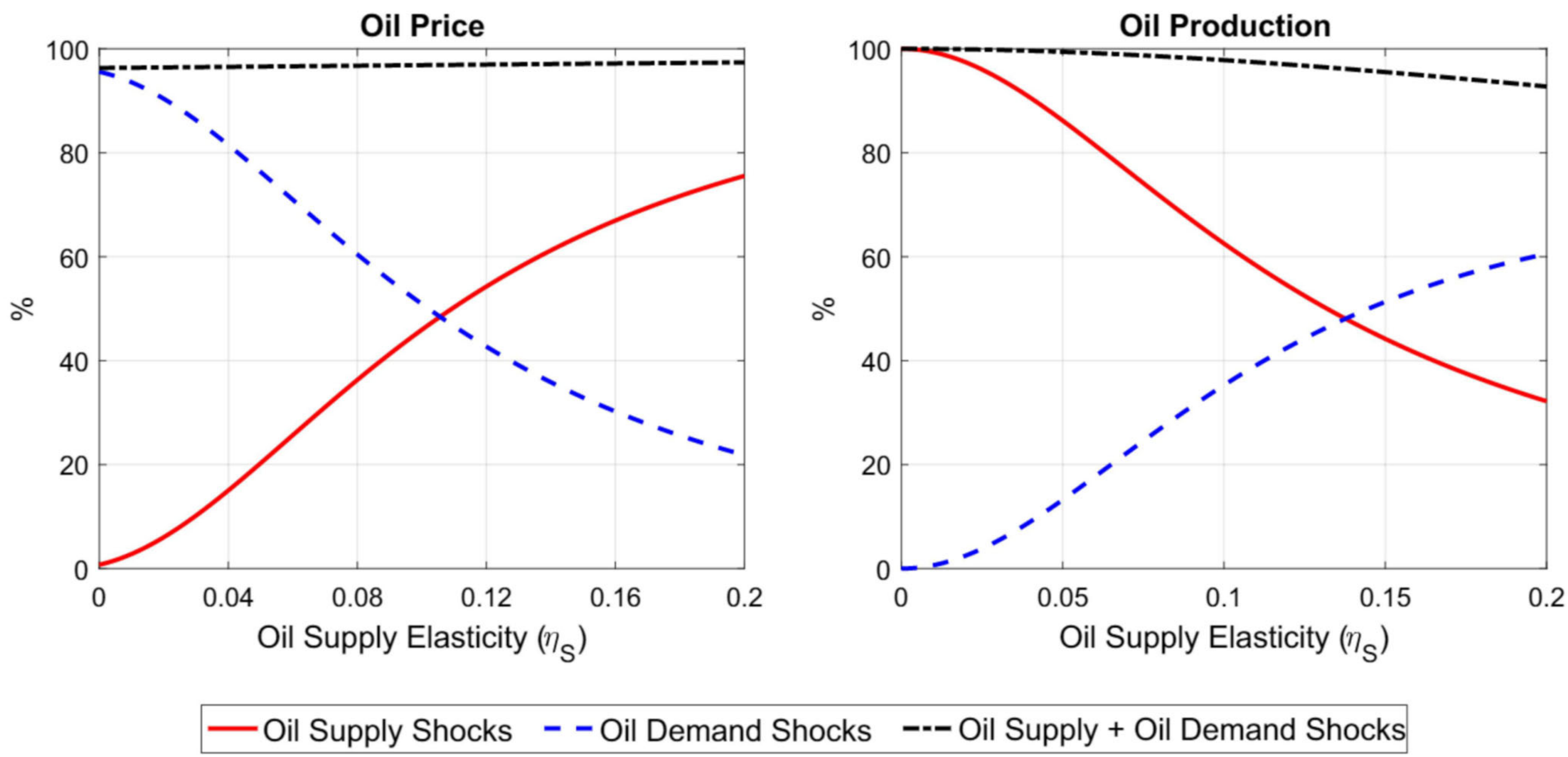
Why is β so large?

- A 10% increase in the price of oil with no change in income would result in a 60% drop in consumption within the month.
- $\alpha = 0$ plays a key role in this conclusion.
- Suppose that $\alpha > 0$, positive correlation between $\tilde{u}_t^d$ and q_t would bias OLS estimate upward (closer to zero) ⇒ implies bias of estimated demand elasticity toward larger absolute value

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

Other implications of a zero supply elasticity

- Caldara, Cavallo, and Iacoviello (JME 2019) show that small changes in the elasticities also have implications for quantifying the determinants of fluctuations in oil prices and production.



Structural inference with inexact identifying information

- Frequentist: exact prior information about some aspects of the model
- Bayesian: represent prior information in the form of a probability density $p(\boldsymbol{\psi})$ for $\boldsymbol{\psi} = (\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{D})$

$\Rightarrow p(\boldsymbol{\psi})$ is higher for values that are more plausible

$\Rightarrow p(\boldsymbol{\psi})$ is lower for values that are less likely

- Posterior density combines prior and data:

$$p(\boldsymbol{\psi}|\mathbf{Y}) = \frac{p(\boldsymbol{\psi})f(\mathbf{Y}|\boldsymbol{\psi})}{\int p(\boldsymbol{\psi})f(\mathbf{Y}|\boldsymbol{\psi})d\boldsymbol{\psi}}$$

Bayesian interpretation of Cholesky identification

Consider Bayesian who was certain

before seeing data that $\gamma = \alpha = \psi = 0$.

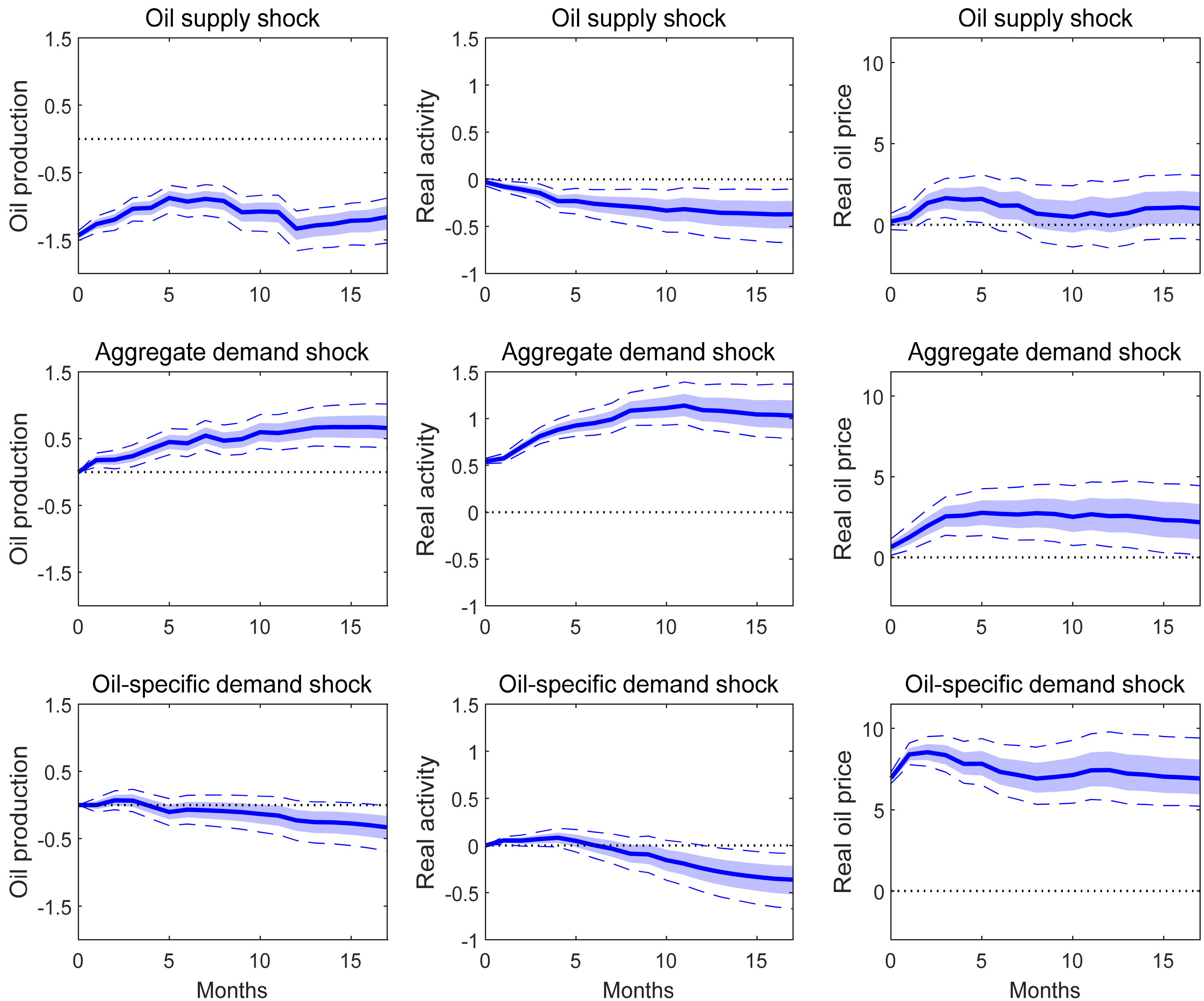
Had no useful information about ξ, δ, β

Student t priors with $c = 0, \sigma = 100, \nu = 3$.

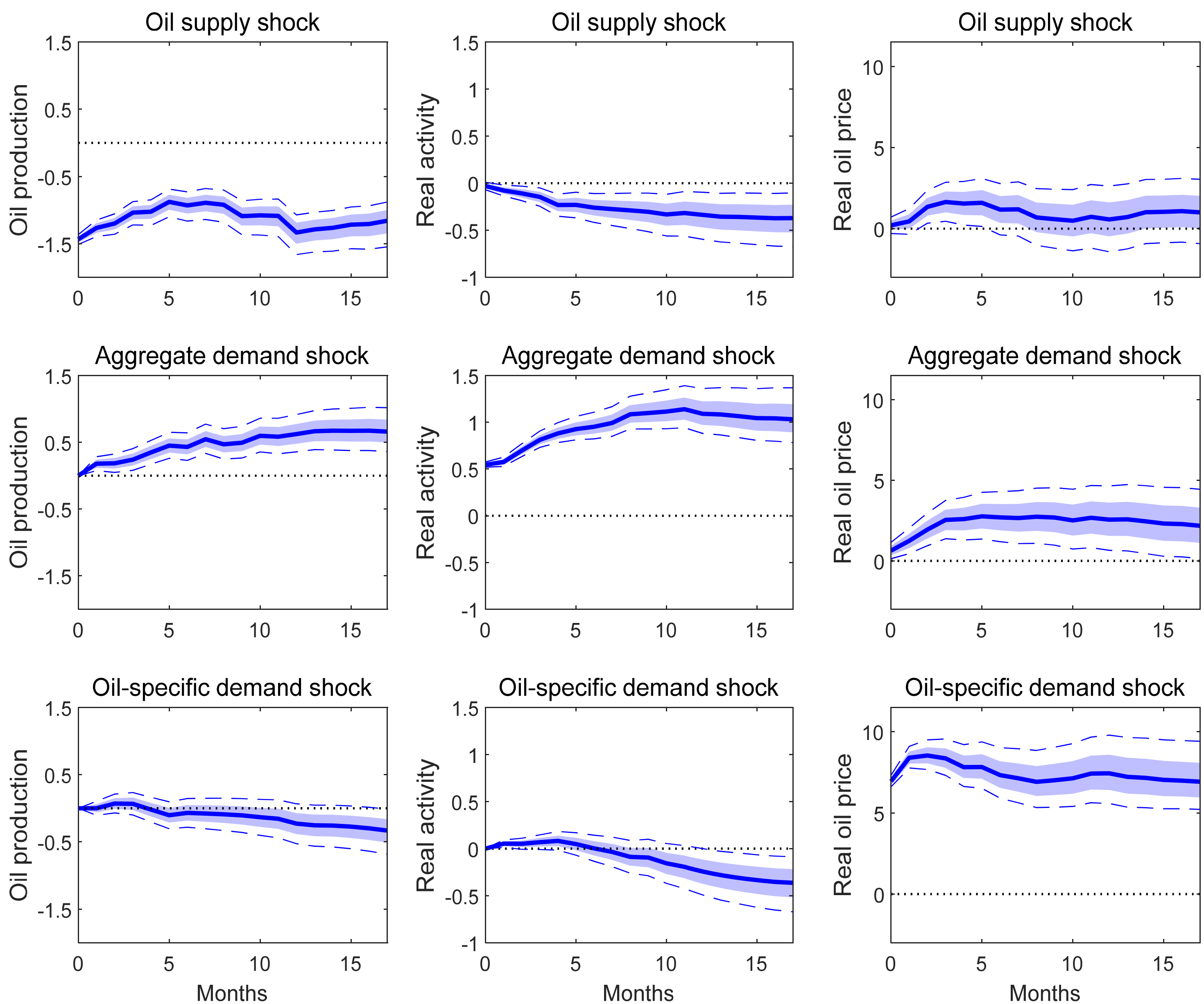
$$p(\mathbf{A}) \propto \left[1 + \tfrac{1}{\nu}\left(\tfrac{\xi}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}} \left[1 + \tfrac{1}{\nu}\left(\tfrac{\delta}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}} \times$$

$$\left[1 + \tfrac{1}{\nu}\left(\tfrac{\beta}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}}$$

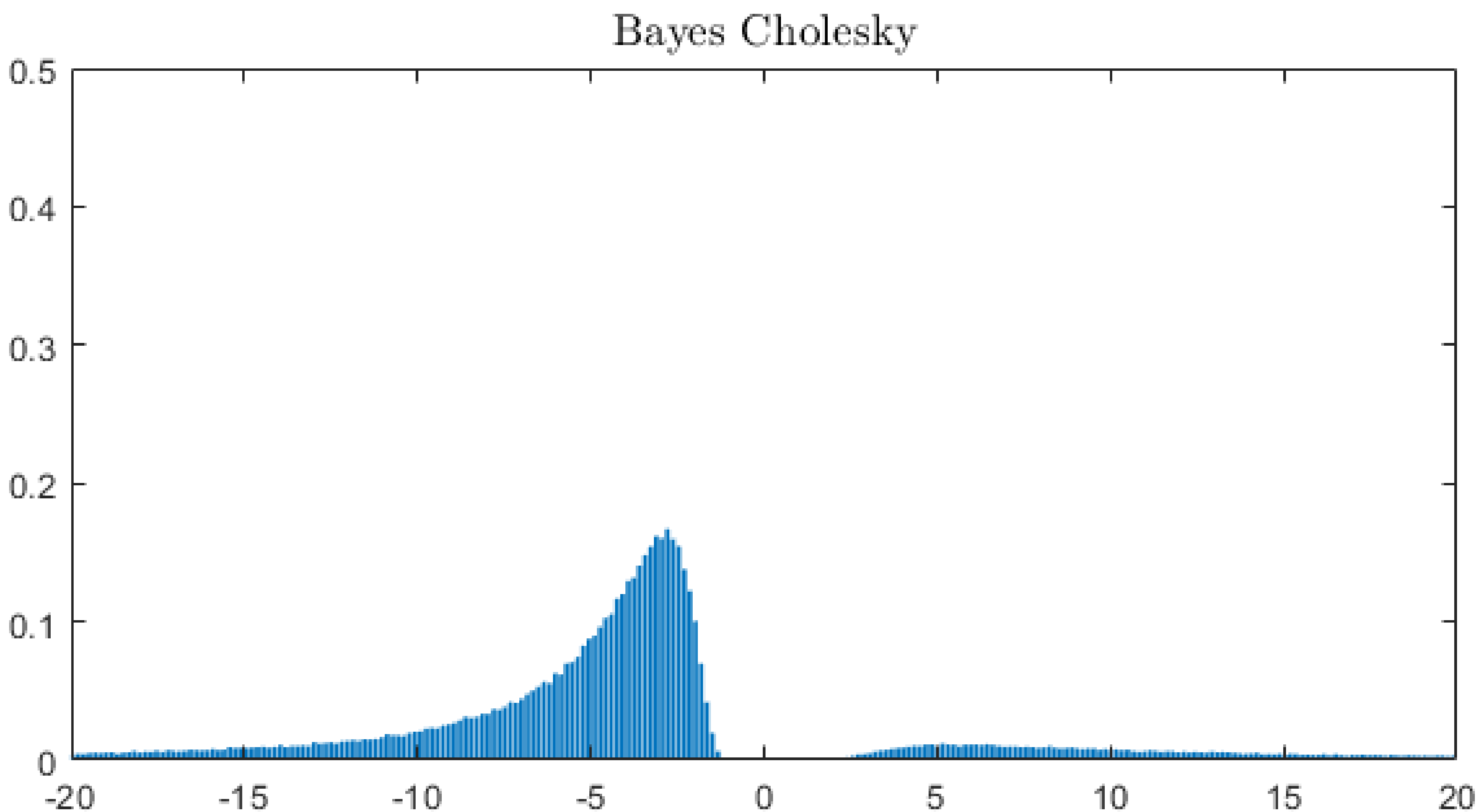
Structural IRF of Bayesian who was dogmatic about some and uninformed about other parameters



Structural IRF of frequentist who used Cholesky identifying assumptions



Bayesian posterior distribution of short-run demand elasticity β



17.5% probability that $\beta > 0$
97% probability that $|\beta| > 2$

A Bayesian generalization of traditional identification

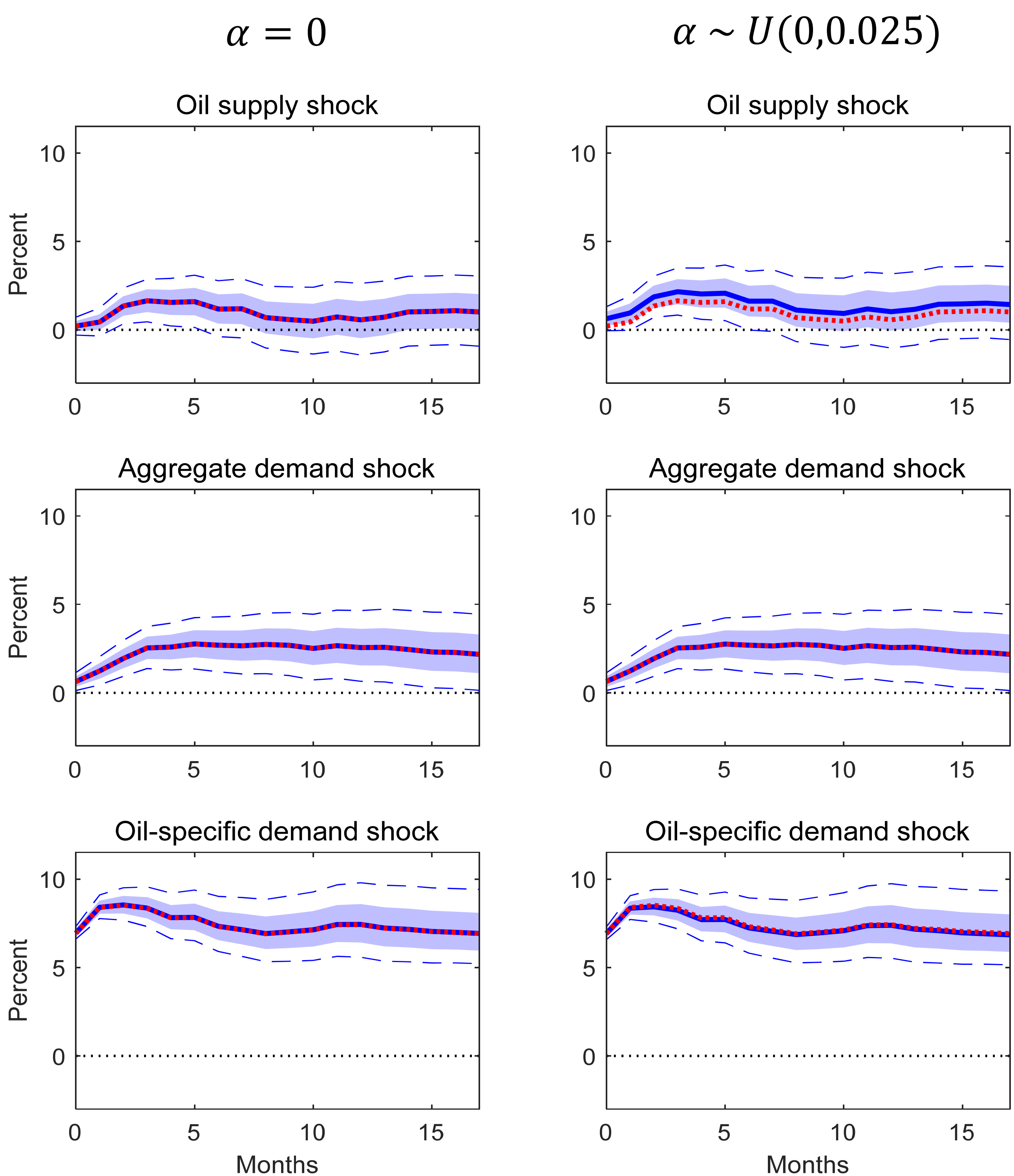
Consider next a Bayesian who is extremely confident (but not absolutely certain) that supply elasticity is very small:

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

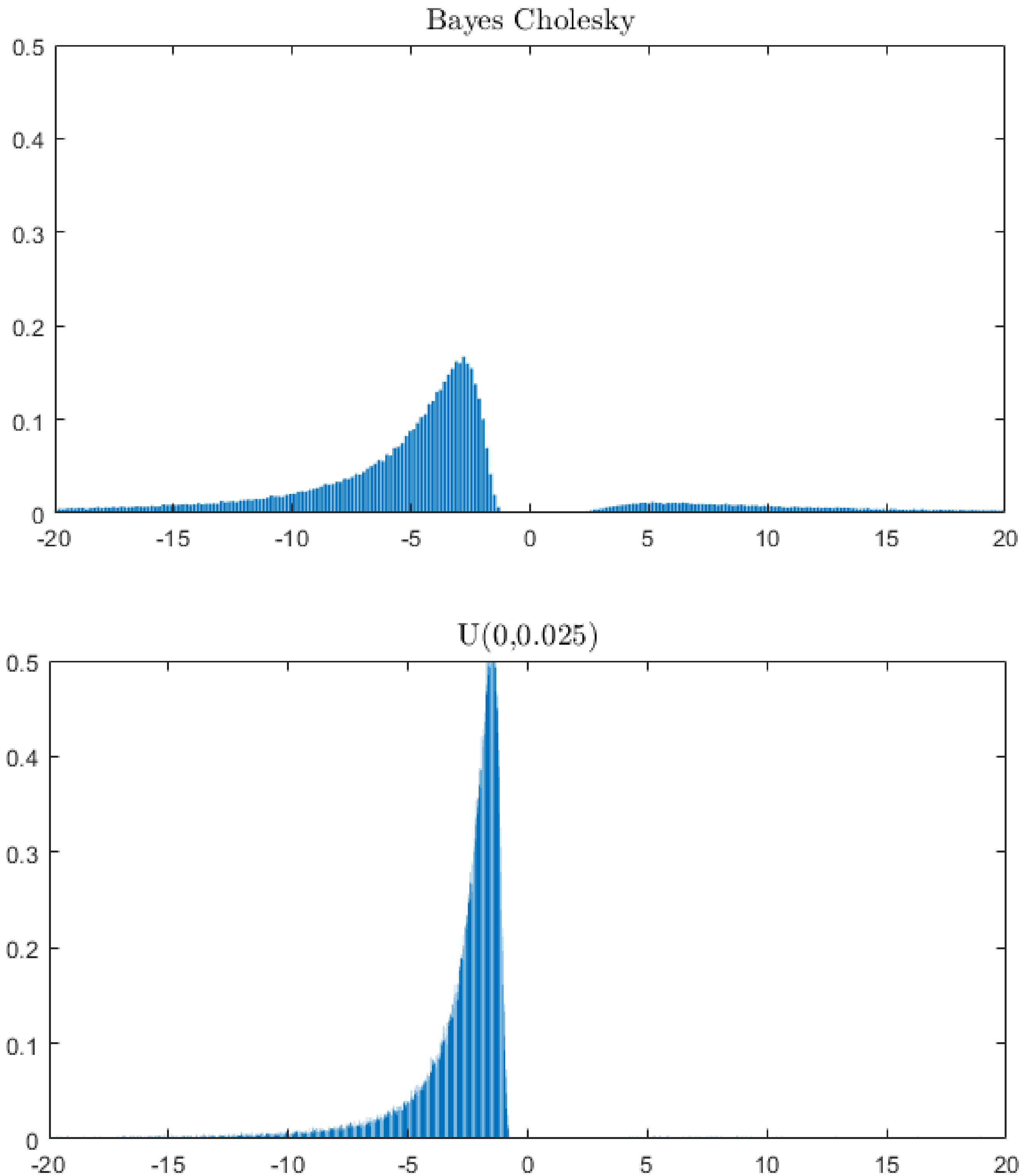
$$p(\mathbf{A}) \propto \left[1 + \tfrac{1}{\nu}\left(\tfrac{\xi}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}} \left[1 + \tfrac{1}{\nu}\left(\tfrac{\delta}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}} \times$$

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

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



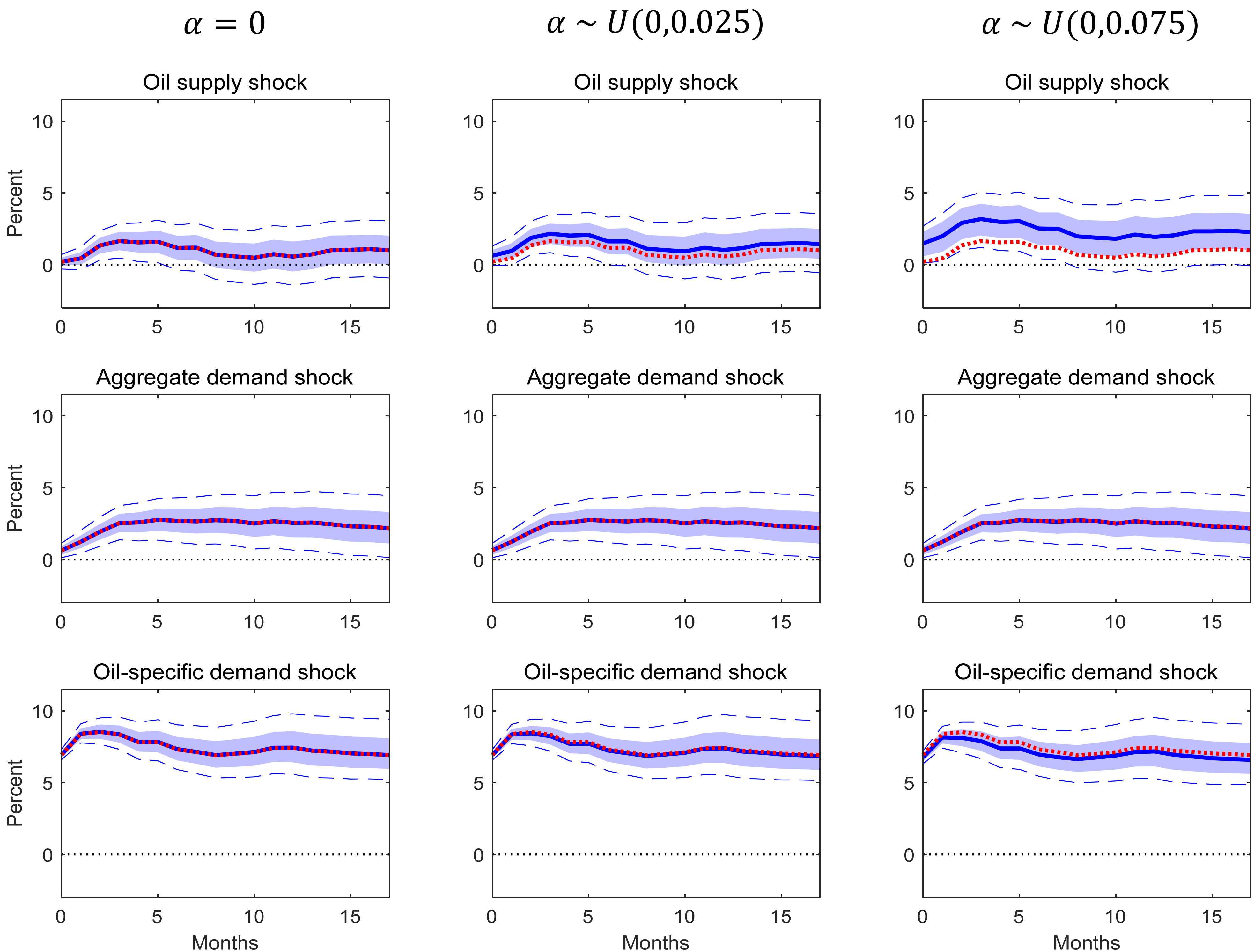
$\alpha = 0$
$\alpha \sim U(0,0.025)$
Oil supply shock
Oil supply shock
Aggregate demand shock
Aggregate demand shock
Oil-specific demand shock
Oil-specific demand shock
Percent
Months
Months



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
Oil supply shock
Oil supply shock
Aggregate demand shock
Aggregate demand shock
Aggregate demand shock
Oil-specific demand shock
Oil-specific demand shock
Oil-specific demand shock
Percent
Months
Months
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

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

Set identification using sign restrictions

- Prior knowledge in the form of signs of effects of structural shocks (see, e.g., Uhlig 2005; Rubio-Ramírez, Waggoner, and Zha 2010)

- For example: oil market model of Kilian and Murphy (2012) with $\mathbf{y}_t = (q_t, y_t, p_t)'$

$$\text{sign}(\mathbf{H}) = \begin{bmatrix} - & + & + \\ - & + & - \\ + & + & + \end{bmatrix}$$

RWZ sign restriction algorithm

Step 1 Take a draw $(\mathbf{\Pi}, \mathbf{\Omega})$ from the posterior

Step 2 Compute the Cholesky factor **P** of $\mathbf{\Omega}$

Step 3 Generate an $(n \times n)$ matrix $\mathbf{X} = [x_{ij}]$ from $N(0,1)$

Step 4 Take the QR decomposition of $\mathbf{X} = \mathbf{QR}$ with $\mathbf{Q}$ an orthonormal matrix $(\mathbf{Q}'\mathbf{Q} = \mathbf{I}_n)$

Step 5 Compute IRFs using $\mathbf{H} = \mathbf{PQ}'$

Step 6 Keep **H** if it satisfies the sign restrictions; otherwise discard it.

- Researchers typically report median and 68% of retained set of values.

- BUT each of these draws is perfectly consistent

  with all the observed data and

  with all the specified restrictions.

A frequentist critique

- Conditional on MLE estimate of reduced-form parameters, calculate upper and lower bounds of the set across all generated draws:

$$(\hat{\underline{\zeta}}_{ijs}, \hat{\bar{\zeta}}_{ijs})$$

- Accounting for estimation uncertainty will further increase set

- Moon and Schorfheide (2012), Watson (2019), and Giacomini and Kitagawa (2021) criticized practice of reporting 68% error bands strictly smaller than

$$(\hat{\underline{\zeta}}_{ijs}, \hat{\bar{\zeta}}_{ijs})$$

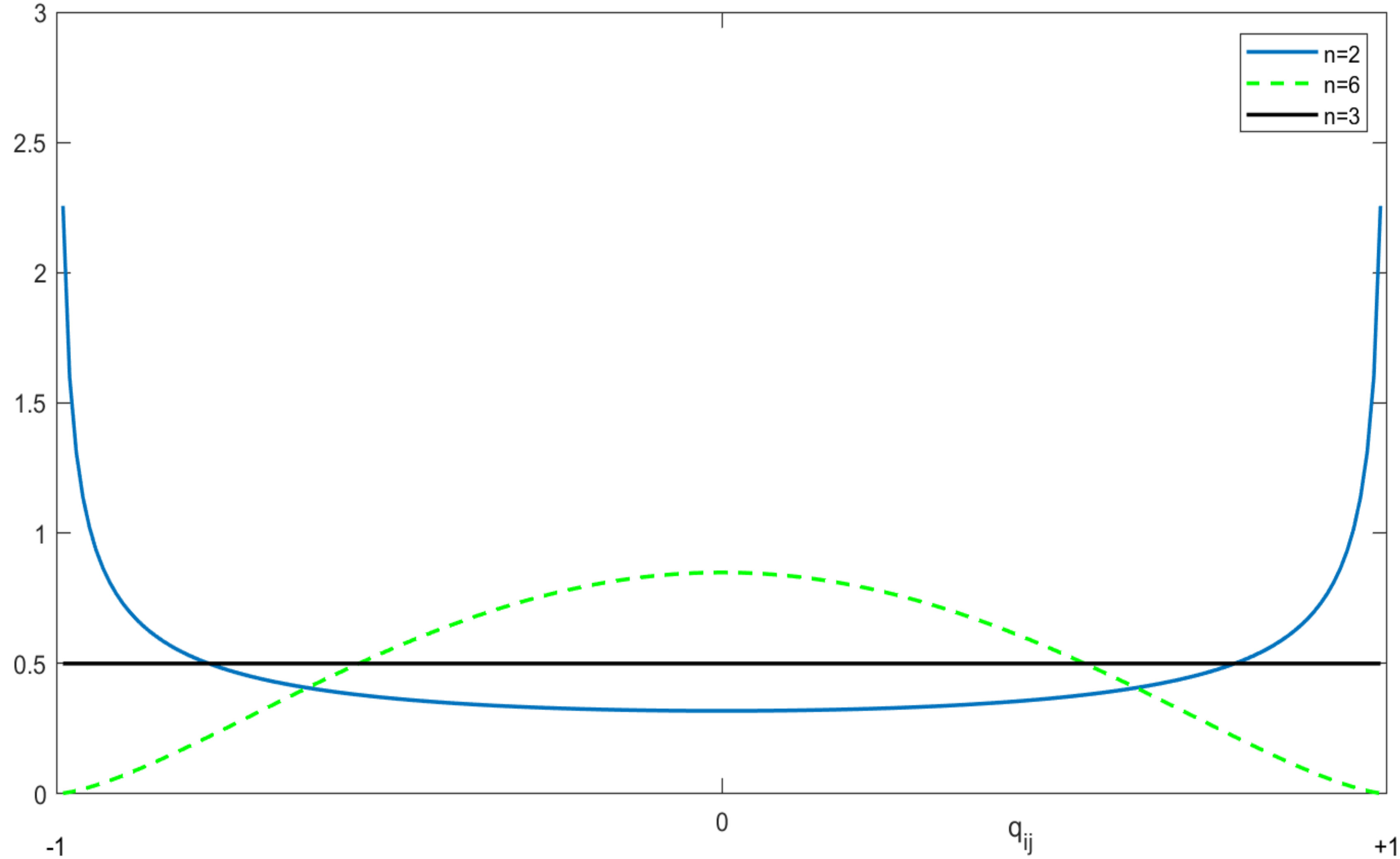
A Bayesian critique

- Baumeister and Hamilton (2015) showed that the (i,j) element of $\mathbf{Q}$ has the following density

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

- Reporting median and 68% credibility sets can only be justified if **this** distribution was acknowledged as prior information

- What prior information does this distribution incorporate?



- If $n = 3$, all values are equally likely.
- If $n = 2$, effect of first shock on first variable is more likely to be large.
- If $n > 3$, values near zero are more likely.

Why is this problematic?

- This prior implies that before seeing the data, we would have the **same** prior knowledge regardless of the dataset, economic content, sample period …

  $\Rightarrow$ reasonable?

- This practice of ruling out values that are perfectly consistent with data and restrictions means that *additional* prior information is used that is not made explicit.

- Rubio-Ramirez and co-authors argue that using this prior information is justified because the prior implies a uniform distribution with respect to a certain measure.

- BUT a frequentist who is unpersuaded by the validity of the prior information would find the reported confidence bands to be much too narrow.

- Users make no effort to persuade others that the prior information they have relied on is convincing.

Why is this such a big deal?

- In set-identified structural VARs, the credibility of the prior is key because the influence of the prior does not vanish asymptotically (Baumeister and Hamilton, 2015; Giacomini and Kitagawa, 2021)

- For some parameters, data are completely uninformative, and you get back the prior.

  $\Rightarrow$ Bayesian with uniform prior would claim to rule out some values for those parameters even though there is no basis in the data for doing so

  $\Rightarrow$ Thus, uniform prior is used in an informative way!

Quantitative relevance of the prior

- Inoue and Kilian (2022) argue that prior for **Q** is negligible in tightly identified VAR models given that they imply a narrow identified set

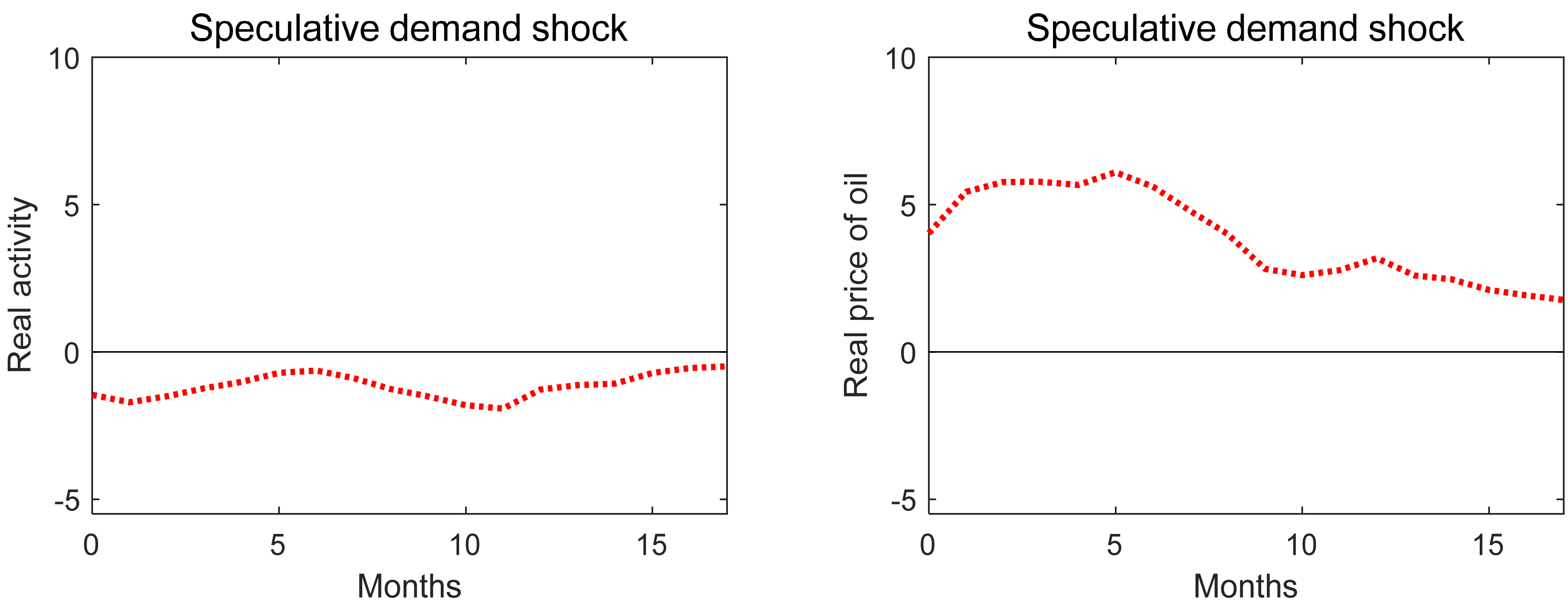
- As we add information in the form of more and tighter constraints, we will end up discarding most of the RWZ draws.

- BUT using only a handful of numbers may give a very inaccurate estimate of the identified set.

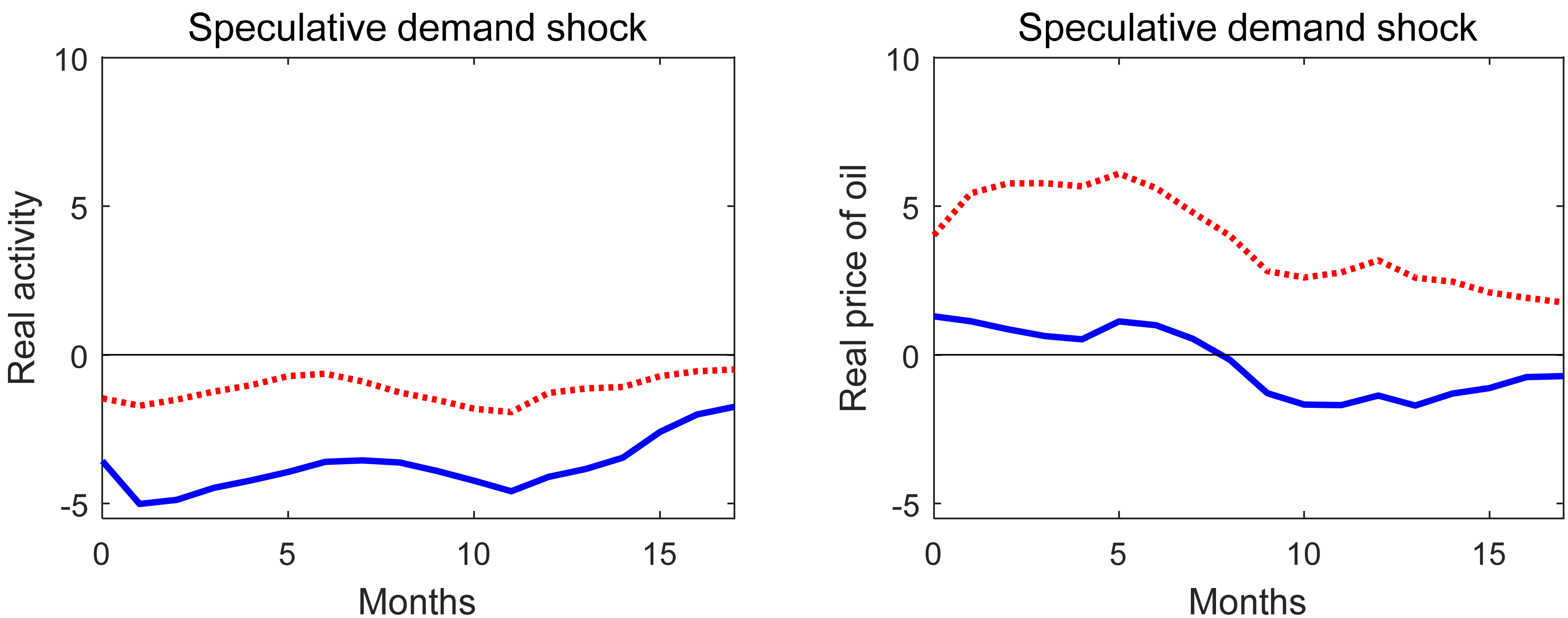
Example

- Kilian and Murphy (2014) generated 5 million draws for the vector of possible parameters.
- Rule out various draws based on a long list of criteria (e.g., supply elasticity < 0.0258)
- The end result of running the code is that only 16 of the original 5 million draws remain.
- Select the draw with “impact price elasticity of oil demand in use closest to the posterior median”

Effect of speculative demand shock on real activity and price as originally reported in KM14



Effect of speculative demand shock calculated using KM14 code with two different random number seeds



Dashed red: seed = 316 (used by KM14). Blue: seed = 613.

What to do going forward?

(1) Robust Bayesian approach (Giacomini and Kitagawa, 2021): inference based on *full set of possible priors*

→ Bayesian robust posterior credible set corresponds to frequentist confidence set asymptotically

(2) Upper and lower bound based on MLE (no prior)

(3) Use *informative prior* about objects of interest derived from economic theory or prior evidence that is stated upfront and summarizes knowledge in the form of a probability distribution (BH 2015, 2018, 2019, 2020, 2022, 2023)

Example: labor market dynamics

demand:

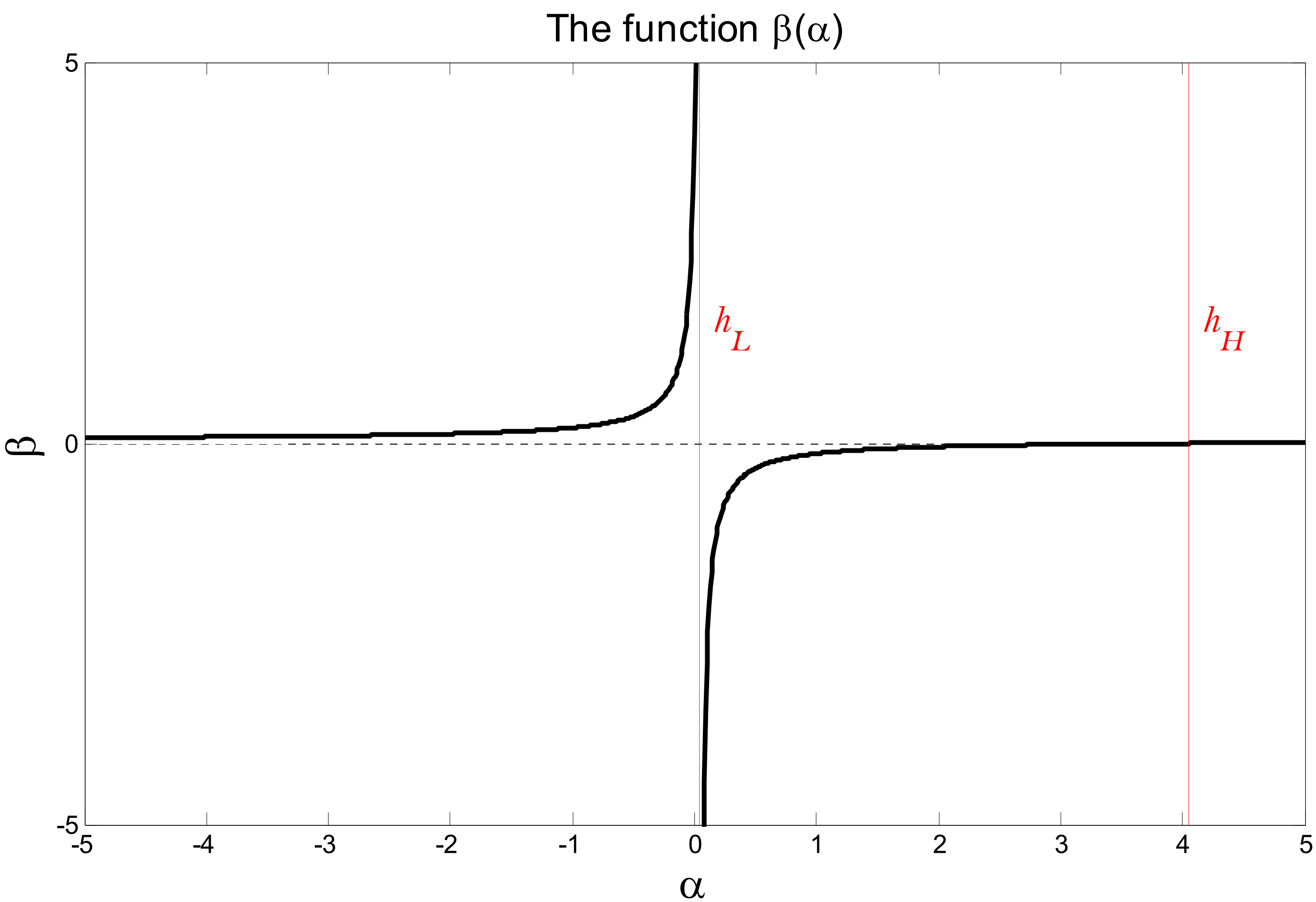
$$\Delta n_t = k^d + \beta^d \Delta w_t + b_{11}^d \Delta w_{t-1} + b_{12}^d \Delta n_{t-1} + b_{21}^d \Delta w_{t-2}$$
$$+ b_{22}^d \Delta n_{t-2} + \cdots + b_{m1}^d \Delta w_{t-m} + b_{m2}^d \Delta n_{t-m} + u_t^d$$

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:

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



$$\hat{\beta}(\alpha) = \frac{(\hat{\omega}_{22} - \alpha\hat{\omega}_{12})}{(\hat{\omega}_{12} - \alpha\hat{\omega}_{11})}$$

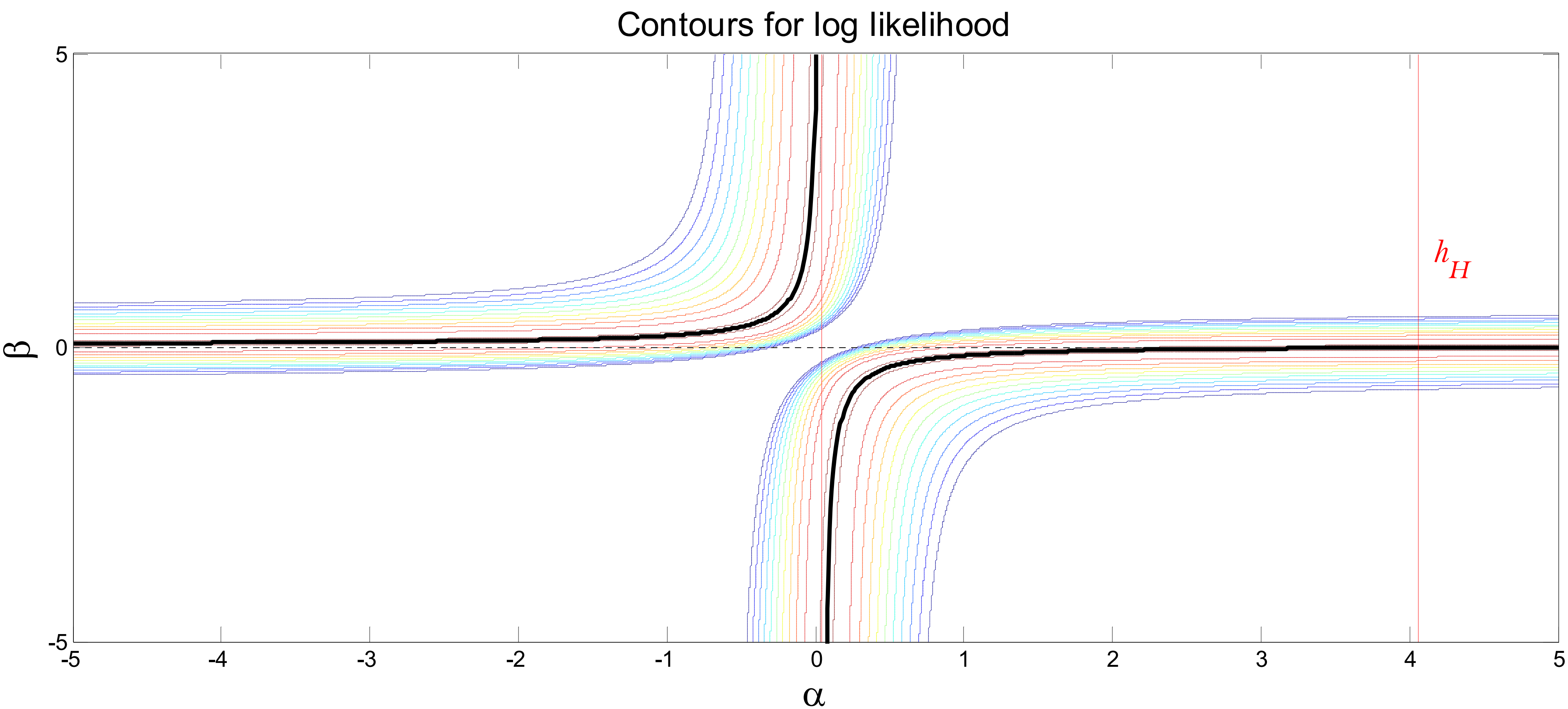
for $\beta = 0$,

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

for $\beta = -\infty$,

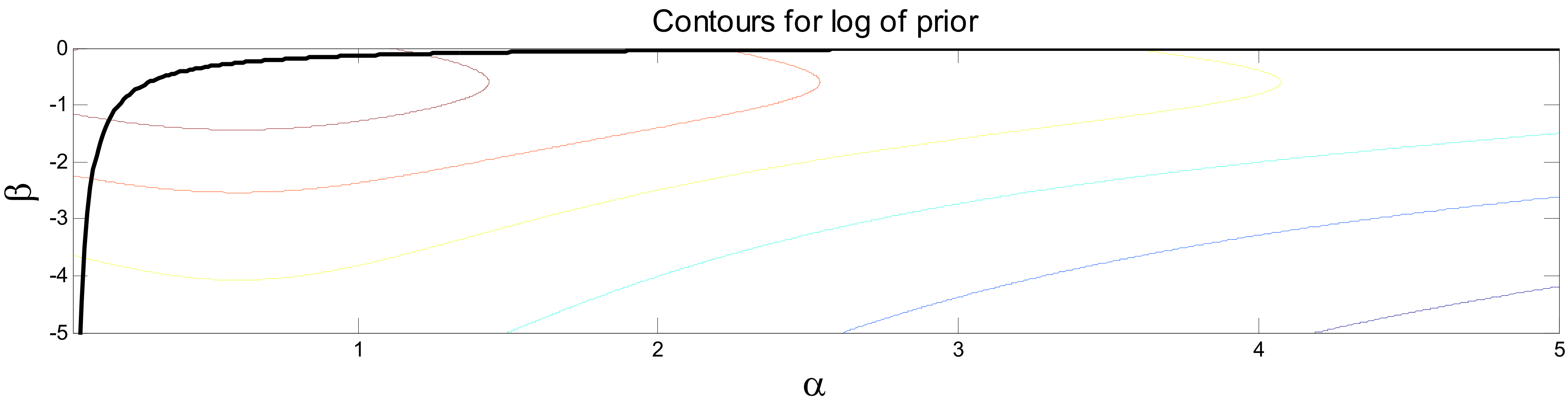
$\hat{\alpha}_{MLE} > h_L = \hat{\omega}_{12}/\hat{\omega}_{11}$

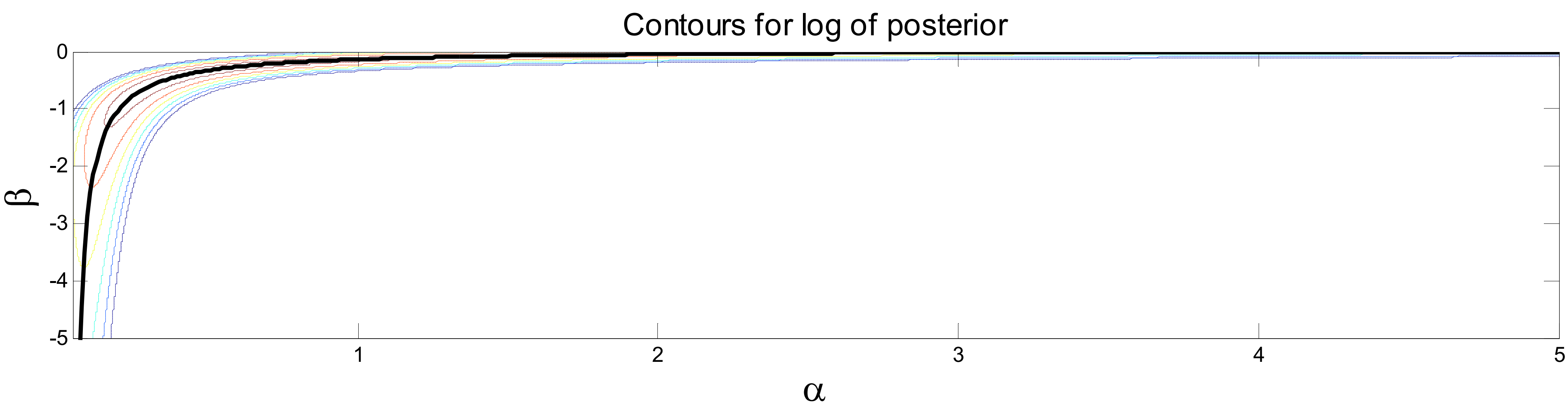
given $\hat{\omega}_{12} > 0$



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

- Student *t* priors for labor demand and supply elasticity
- Prior induces ranking in posterior





Moving to the Disaggregate Level

- So far, described the behavior of the world oil market
  BUT oil market is composed of many individual
  producers and consumers

- Differences between individual observations and aggregate outcomes can be an important source of identification.

Examples:

- Bartik instruments
  (Bartik, 1991; Blanchard et al., 1992; Goldsmith-Pinkham et al., 2020; Carlino and Drautzburg, 2020)
- Granular instrumental variables
  (Gabaix and Koijen, 2020; Qian, 2023; Banafti and Lee, 2022; Sarto, 2022; Caldara, Cavallo and Iacoviello, 2019)

A structural model of the world oil market

- Data from 1973M1 to 2023M2 (drop COVID period)

q_{it} = growth rate of country i oil production

s_{qi} = share of country i in world total

$\sum_{i=1}^{n} s_{qi} q_{it}$ = approximate growth in global oil production

- Our empirical analysis will use the three biggest producers (U.S., Saudi Arabia, Russia) plus the rest of the world $(n = 4)$

A structural model of the world oil market

c_{jt} = growth rate of country j oil consumption

s_{cj} = share of country j in world total

$\sum_{j=1}^{m} s_{cj} c_{jt}$ = approximate growth in global oil consumption

- Our empirical analysis will use the three biggest consumers (U.S., Japan, Europe) plus the rest of the world ($m = 4$)

Supply curve of country *i*

$$q_{it} = \phi_{qi} p_t + \mathbf{b}'_{qi} \mathbf{x}_{t-1} + u_{qit} + u_{\chi it}$$

$\phi_{qi} =$ country i short-run supply elasticity

$\mathbf{x}_{t-1}$ contains 12 lags production and consumption of every country in world plus 12 lags of world price

$u_{qit} =$ supply shock for country i

$u_{\chi it} =$ error in measuring country i production

Demand curve of country *j*

$$c_{jt} = \phi_{cj} p_t + \mathbf{b}'_{cj} \mathbf{x}_{t-1} + u_{cjt} + u_{\psi jt}$$

ϕ_{cj} = country j short-run demand elasticity

u_{cjt} = demand shock for country j

$u_{\psi jt}$ = error in measuring country j consumption

Inventory demand

$$v_t = \phi_v p_t + \mathbf{b}_v' \mathbf{x}_{t-1} + u_{vt}$$

This equals difference between correctly measured production and consumption

$$v_t = \sum_{i=1}^{n} s_{qi}(q_{it} - u_{\chi it}) - \sum_{j=1}^{m} s_{cj}(c_{jt} - u_{\psi jt})$$

Structural model

$$q_{it} = \phi_{qi} p_t + \mathbf{b}'_{qi} \mathbf{x}_{t-1} + u_{qit} + u_{\chi it} \qquad i = 1, \ldots, n$$

$$\text{or} \quad \underset{(n\times 1)}{\mathbf{q}_t} = \underset{(n\times 1)}{\boldsymbol{\phi}_q} \; p_t + \underset{(n\times k)}{\mathbf{B}_q} \; \mathbf{x}_{t-1} + \underset{(n\times 1)}{\mathbf{u}_{qt}} + \underset{(n\times 1)}{\mathbf{u}_{\chi t}}$$

$$c_{jt} = \phi_{cj} p_t + \mathbf{b}'_{cj} \mathbf{x}_{t-1} + u_{cjt} + u_{\psi jt} \qquad j = 1, \ldots, m$$

$$\text{or} \quad \underset{(m\times 1)}{\mathbf{c}_t} = \underset{(m\times 1)}{\boldsymbol{\phi}_c} \; p_t + \underset{(m\times k)}{\mathbf{B}_c} \; \mathbf{x}_{t-1} + \underset{(m\times 1)}{\mathbf{u}_{ct}} + \underset{(m\times 1)}{\mathbf{u}_{\psi t}}$$

$$(\mathbf{s}'_q \boldsymbol{\phi}_q - \mathbf{s}'_c \boldsymbol{\phi}_c - \phi_v) p_t =$$

$$(\mathbf{s}'_c \mathbf{B}_c - \mathbf{s}'_q \mathbf{B}_q + \mathbf{b}'_v) \mathbf{x}_{t-1} + \mathbf{s}'_c \mathbf{u}_{ct} - \mathbf{s}'_q \mathbf{u}_{qt} + u_{vt}$$

These are $n + m + 1$ equations to determine the $n + m + 1$ variables $\mathbf{q}_t, \mathbf{c}_t, p_t$ in terms of the structural shocks $(\mathbf{u}_{qt}, \mathbf{u}_{\chi t}, \mathbf{u}_{ct}, \mathbf{u}_{\psi t}, u_{vt})$.

It is possible to estimate structural parameters like $\boldsymbol{\phi}_q$ and $\boldsymbol{\phi}_c$ if we make assumptions about the correlations between these structural shocks.

A less restrictive model

- We allow for $\mathbf{u}_{qt}$ and $\mathbf{u}_{ct}$ to be correlated through common dependence on a single factor f_t.
- We also allow $\mathbf{u}_{ct}$ to depend on a global demand shock f_{ct} with different loadings for each country.
- We also allow $\mathbf{u}_{ct}$ to depend on idiosyncratic factors.

$$\mathbf{u}_{ct} = \mathbf{h}_c f_t + \boldsymbol{\gamma}_c f_{ct} + \boldsymbol{\eta}_{ct}$$

- Reduced form is a VAR(12)

$$\underset{(9\times 1)}{\mathbf{y}_t} = (\mathbf{q}_t', \mathbf{c}_t', p_t)'$$

$$\mathbf{y}_t = \Pi \mathbf{x}_{t-1} + \boldsymbol{\epsilon}_t$$

$$\mathbf{x}_{t-1} = (1, \mathbf{y}_{t-1}', \ldots, \mathbf{y}_{t-12}')'$$

$$\begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{nm} & -\boldsymbol{\phi}_q \\ \mathbf{0}_{mn} & \mathbf{I}_m & -\boldsymbol{\phi}_c \\ \mathbf{0}_{1n} & \mathbf{0}_{1m} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{\epsilon}_{qt} \\ \boldsymbol{\epsilon}_{ct} \\ \varepsilon_{pt} \end{bmatrix} =$$

$$\begin{bmatrix} \mathbf{h}_q f_t + \boldsymbol{\gamma}_q f_{qt} + \boldsymbol{\eta}_{qt} + \mathbf{u}_{\chi t} \\ \mathbf{h}_c f_t + \boldsymbol{\gamma}_c f_{ct} + \boldsymbol{\eta}_{ct} + \mathbf{u}_{\psi t} \\ \alpha \mathbf{s}_c'(\mathbf{h}_c f_t + \boldsymbol{\gamma}_c f_{ct} + \boldsymbol{\eta}_{ct}) - \alpha \mathbf{s}_q'(\mathbf{h}_q f_t + \boldsymbol{\gamma}_q f_{qt} + \boldsymbol{\eta}_{qt}) + \alpha u_{vt} \end{bmatrix}$$

$$\mathbf{A}\boldsymbol{\epsilon}_t = \mathbf{u}_t$$

$$E(\mathbf{u}_t \mathbf{u}_t') = \mathbf{D}$$

Assumptions

- Idiosyncratic shocks are uncorrelated

$$E\begin{bmatrix} \boldsymbol{\eta}_{qt} \\ \boldsymbol{\eta}_{ct} \end{bmatrix}\begin{bmatrix} \boldsymbol{\eta}'_{qt} & \boldsymbol{\eta}'_{ct} \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{\Sigma}_q & \mathbf{0}_{nm} \\ \mathbf{0}_{mn} & \boldsymbol{\Sigma}_c \end{bmatrix} \text{ (diagonal)}$$

Assumptions

- Factor normalization

$$E\begin{bmatrix} f_t \\ f_{qt} \\ f_{ct} \end{bmatrix}\begin{bmatrix} f_t & f_{qt} & f_{ct} \end{bmatrix} = \mathbf{I}_3$$

- Measurement errors have common variance

$$E\begin{bmatrix} \mathbf{u}_{\chi t} \\ \mathbf{u}_{\psi t} \end{bmatrix}\begin{bmatrix} \mathbf{u}'_{\chi t} & \mathbf{u}'_{\psi t} \end{bmatrix} = \begin{bmatrix} \sigma_\chi^2 \mathbf{I}_n & \mathbf{0}_{nm} \\ \mathbf{0}_{mn} & \sigma_\psi^2 \mathbf{I}_m \end{bmatrix}$$

Implied variance-covariance matrix

$$\mathbf{D} = \begin{bmatrix} \mathbf{h}_q\mathbf{h}_q' + \gamma_q\gamma_q' + \mathbf{\Sigma}_q + \sigma_\chi^2\mathbf{I}_n & \mathbf{h}_q\mathbf{h}_c' & -\alpha\left(\mathbf{h}_q\mathbf{h}_q' + \gamma_q\gamma_q' + \mathbf{\Sigma}_q\right)\mathbf{s}_q + \alpha\mathbf{h}_q\mathbf{h}_c'\mathbf{s}_c \\ \mathbf{h}_c\mathbf{h}_q' & \mathbf{h}_c\mathbf{h}_c' + \gamma_c\gamma_c' + \mathbf{\Sigma}_c + \sigma_\psi^2\mathbf{I}_m & -\alpha\mathbf{h}_c\mathbf{h}_q'\mathbf{s}_q + \alpha(\mathbf{h}_c\mathbf{h}_c' + \gamma_c\gamma_c' + \mathbf{\Sigma}_c)\mathbf{s}_c \\ -\alpha\mathbf{s}_q'\left(\mathbf{h}_q\mathbf{h}_q' + \gamma_q\gamma_q' + \mathbf{\Sigma}_q\right) + \alpha\mathbf{s}_c'\mathbf{h}_c\mathbf{h}_q' & -\alpha\mathbf{s}_q'\mathbf{h}_q\mathbf{h}_c' + \alpha\mathbf{s}_c'(\mathbf{h}_c\mathbf{h}_c' + \gamma_c\gamma_c' + \mathbf{\Sigma}_c) & \alpha^2\left[\mathbf{s}_q'\left(\mathbf{h}_q\mathbf{h}_q' + \gamma_q\gamma_q' + \mathbf{\Sigma}_q\right)\mathbf{s}_q - 2\mathbf{s}_c'\mathbf{h}_c\mathbf{h}_q'\mathbf{s}_q + \mathbf{s}_c'(\mathbf{h}_c\mathbf{h}_c' + \gamma_c\gamma_c' + \mathbf{\Sigma}_c)\mathbf{s}_c + \sigma_v^2\right] \end{bmatrix}$$

- Model has 15 testable overidentifying assumptions.
- These are not rejected in our dataset.

Maximum likelihood estimates of elasticities and their standard errors

U.S. supply	0.019	(0.017)
Saudi supply	0.259	(0.056)
Russia supply	0.029	(0.011)
ROW supply	0.043	(0.029)
U.S. demand	-0.094	(0.031)
Japan demand	-0.018	(0.037)
Europe demand	-0.225	(0.045)
ROW demand	-0.161	(0.045)
Inventory demand	-0.314	(0.060)

Global supply elasticity: 0.064 (0.021)

Global demand elasticity: -0.139 (0.037)

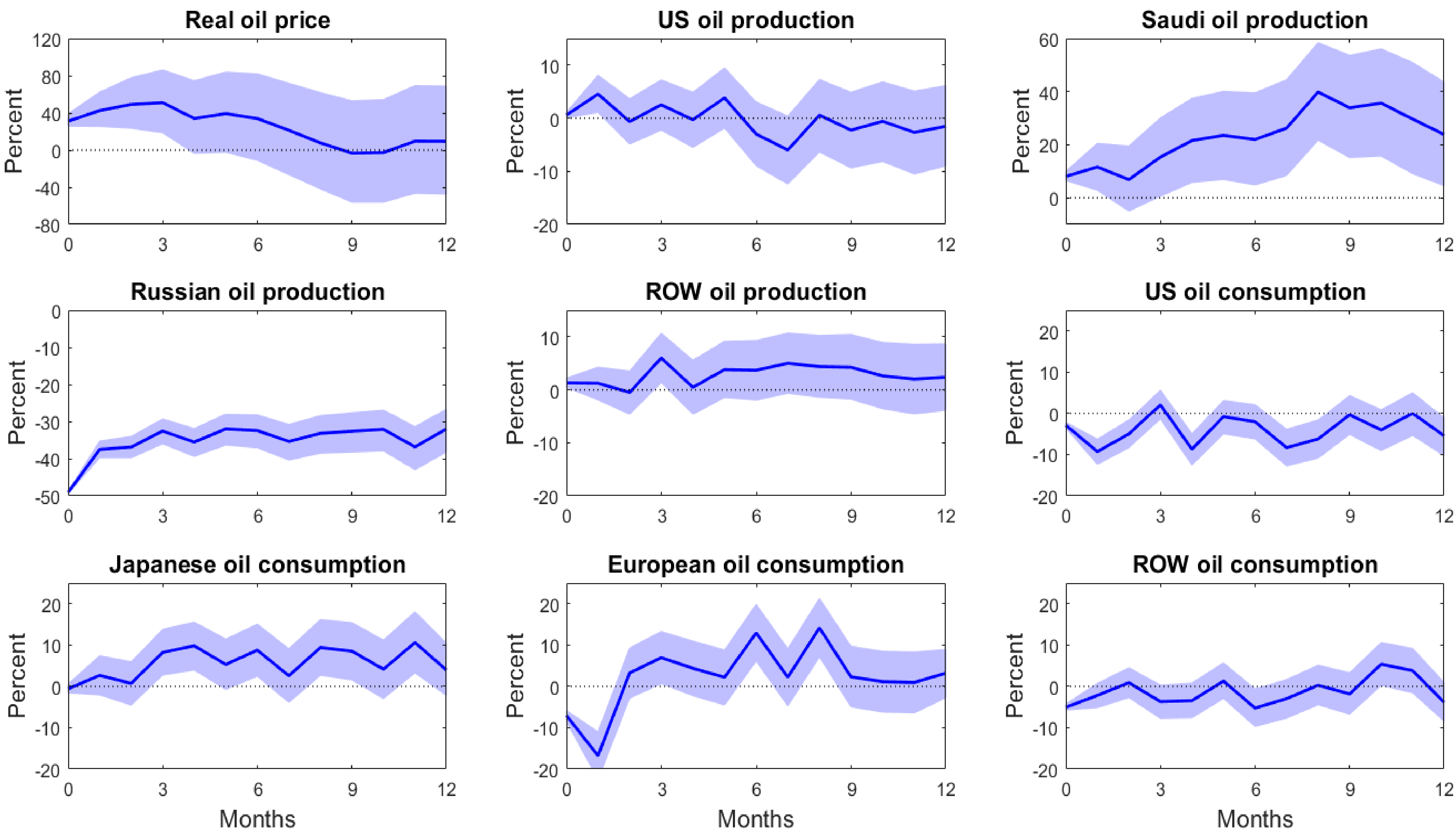
Effects of a 50% decrease in Russian oil production

- Suppose geopolitical developments lead to a cut in Russian production of 5.25 mb/d
- For this analysis we impose that inventory sales can not be used to mitigate production shortfall ($\phi_v = 0$)

Impact effects of 50% cut in Russian oil production

Variable	as % of country			in mb/d
	direct effect (5)	response to price (6)	net effect (7)	net effect (8)
p	31.185			
q_{US}	0	0.580	0.580	0.072
q_{Saudi}	0	8.084	8.084	0.798
q_{Russia}	-50	0.909	-49.091	-5.252
q_{ROW}	0	1.345	1.345	0.664
q				-3.718
c_{US}	0.000	-2.919	-2.919	-0.480
c_{Japan}	0.000	-0.569	-0.569	-0.019
c_{Europe}	0.000	-7.011	-7.011	-0.289
c_{ROW}	0.000	-5.015	-5.015	-2.930
c				-3.718

Dynamic effects of 50% cut in Russian oil production



Key Takeaway

- Identifying assumptions are not a “necessary evil” to recover structural magnitudes of interest like elasticities.

- Rich identifying information is available:
  - Bayesian approach: Think carefully about what prior evidence and economic theory tells us about the structure and use it coherently.
  - Disaggregated approach: Exploit interaction of individual units that determine aggregate economic outcomes for identifying structural magnitudes.