Sign and Zero Restrictions

Martin Bruns

University of St Andrews

Reduced Form Bayesian VAR models

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_{t-1} + \boldsymbol{u}_t \tag{1}$$

$$\boldsymbol{u}_t \sim N(\boldsymbol{0}, \Sigma), \tag{2}$$

where $\Pi = [\boldsymbol{c}; \Pi_1; ...; \Pi_p]$ and $\boldsymbol{w}_t = [1; \boldsymbol{y}_{t-1}; ...; \boldsymbol{y}_{t-p}]$

Reduced Form Bayesian VAR models

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_{t-1} + \boldsymbol{u}_t \tag{1}$$

$$\boldsymbol{u}_t \sim N(\boldsymbol{0}, \Sigma), \tag{2}$$

where $\Pi = [\boldsymbol{c}; \Pi_1; ...; \Pi_p]$ and $\boldsymbol{w}_t = [1; \boldsymbol{y}_{t-1}; ...; \boldsymbol{y}_{t-p}]$

We know

- How to set priors for Π and Σ
- How to draw from the posterior of Π and Σ
- How to generate forecasts $\boldsymbol{y}_{t+1|t}$

Structural VAR models

Reduced Form

- Generate forecasts
- Statements about correlations among variables

Structural VAR models

Reduced Form

- Generate forecasts
- Statements about correlations among variables

Structural Form

- What are the dynamic causal effects of...
  - An exogenous increase in the policy rate?
  - An oil supply shock?
  - A labour demand shock?

Structural VAR models

Reduced Form

- Generate forecasts
- Statements about correlations among variables

Structural Form

- What are the dynamic causal effects of...
  - An exogenous increase in the policy rate?
  - An oil supply shock?
  - A labour demand shock?
- How does the Central Bank react to deviations of inflation/output from their targets?
- What is the fiscal policy rule?

Structural VAR models (cont.)

Structural form

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_{\boldsymbol{t-1}} + B\boldsymbol{\epsilon}_{\boldsymbol{t}}$$
$$\boldsymbol{\epsilon}_{\boldsymbol{t}} \sim N(\mathbf{0}, I)$$

Structural VAR models (cont.)

Structural form

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_{\boldsymbol{t-1}} + B\boldsymbol{\epsilon}_{\boldsymbol{t}}$$
$$\boldsymbol{\epsilon}_{\boldsymbol{t}} \sim N(\mathbf{0}, I)$$

- Compute structural impulse response functions (IRFs)

Structural VAR models (cont.)

Structural form

$$\boldsymbol{y_t} = \Pi \boldsymbol{w_{t-1}} + B\boldsymbol{\epsilon_t}$$
$$\boldsymbol{\epsilon_t} \sim N(\mathbf{0}, I)$$

- Compute structural impulse response functions (IRFs)
- Effects of structural shocks $\boldsymbol{\epsilon_t}$
  - Mutually uncorrelated (diagonal covariance matrix)
  - Distinct economic interpretation
  - Not directly observable
  - Allows statements about **causal** effects of variations in ϵ_t on the data y_t

1. Introduction: Sign-restricted SVARs
2. Arias et al. (2018)
3. Direct Approach
   - Baumeister & Hamilton (2015)
   - Bruns & Piffer (2023)
4. Combining Sign and Zero Restrictions
   - Excursion: Importance Sampling
5. Summary

References and Further Codes

References

- Notation and illustrative example taken from Kilian & Lütkepohl (2017)
- Data and labour market example from Baumeister & Hamilton (2015)
- Codes based on Baumeister & Hamilton (2015), Arias et al. (2018) and Bruns and Piffer (2019)

References and Further Codes

References

- Notation and illustrative example taken from Kilian & Lütkepohl (2017)
- Data and labour market example from Baumeister & Hamilton (2015)
- Codes based on Baumeister & Hamilton (2015), Arias et al. (2018) and Bruns and Piffer (2019)

Further Codes

- BEAR toolbox
- Andrew Blake and Haroon Mumtaz
- Gary Koop
- Lutz Kilian and Helmut Lütkepohl

Structural VAR models (example)

A model of supply and demand

$$\boldsymbol{u}_t = B\epsilon_t \tag{3}$$

$$\begin{bmatrix} u_t^q \\ u_t^p \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \epsilon_t^{supply} \\ \epsilon_t^{demand} \end{bmatrix} \tag{4}$$

Structural VAR models (example)

A model of supply and demand

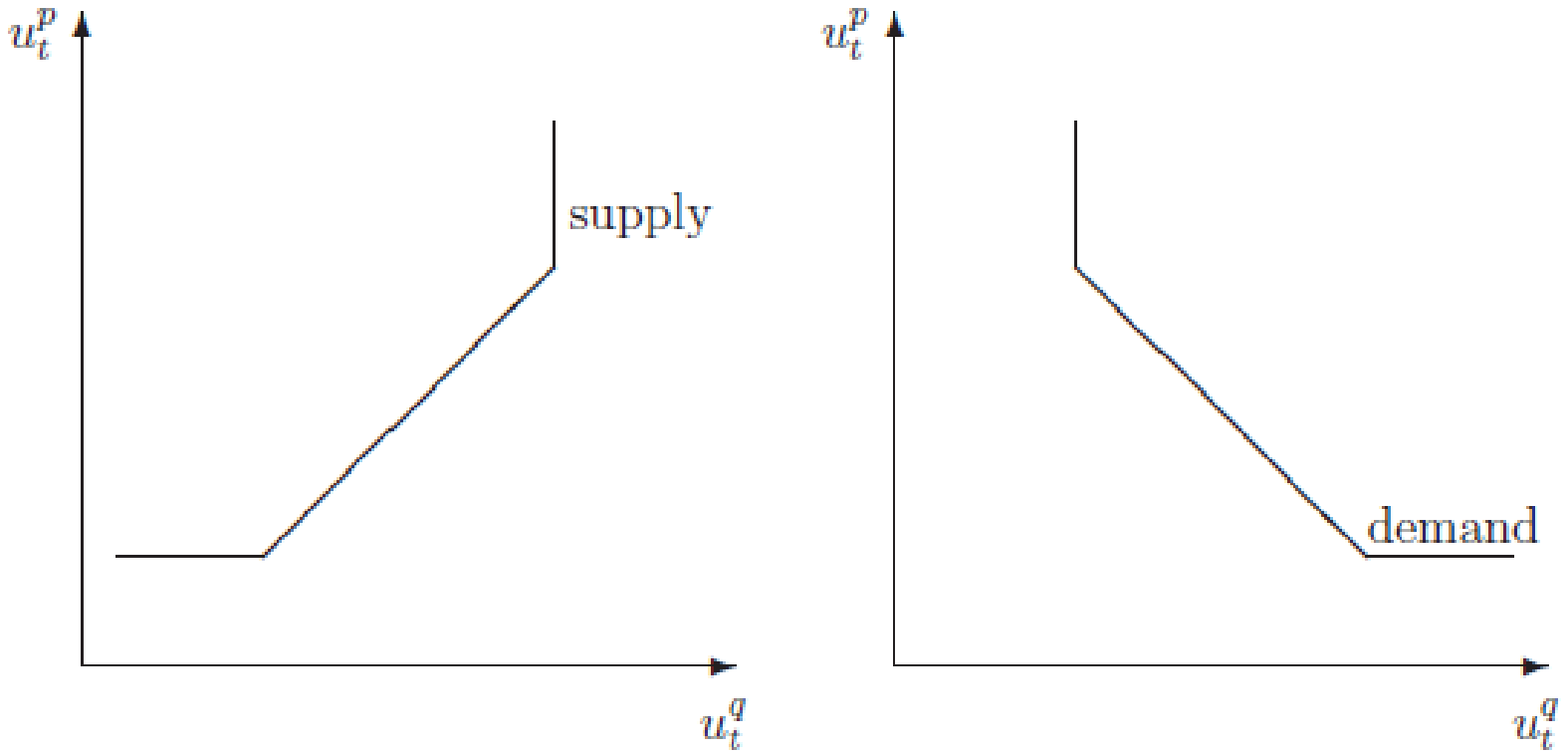
$$\boldsymbol{u}_t = B\epsilon_t \tag{3}$$

$$\begin{bmatrix} u_t^q \\ u_t^p \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \epsilon_t^{supply} \\ \epsilon_t^{demand} \end{bmatrix} \tag{4}$$

- b_{ij} : By how much does a shift in the supply/demand curve affect quantity/price?

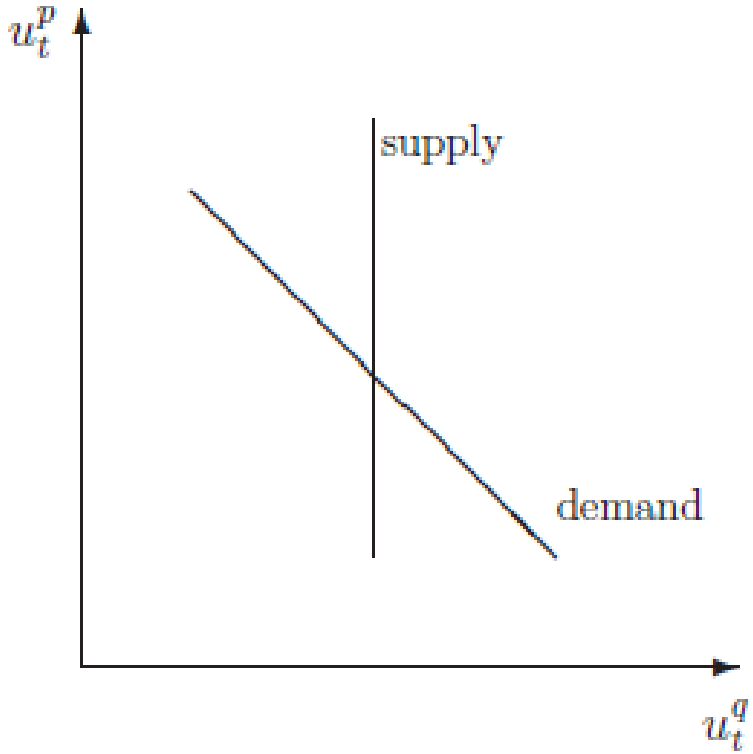
Structural VAR models (example, cont.)

$$\begin{bmatrix} u_t^q \\ u_t^p \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \epsilon_t^{supply} \\ \epsilon_t^{demand} \end{bmatrix}$$



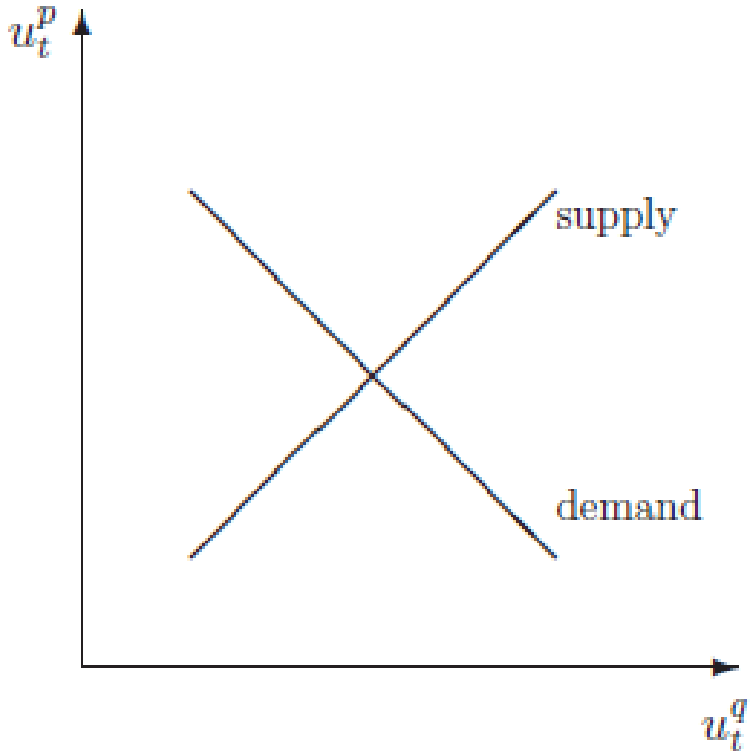
Recursive Identification

$$\begin{bmatrix} u_t^q \\ u_t^p \end{bmatrix} = \begin{bmatrix} b_{11} & \mathbf{0} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \epsilon_t^{supply} \\ \epsilon_t^{demand} \end{bmatrix}$$



Sign Restrictions

$$\begin{bmatrix} u_t^q \\ u_t^p \end{bmatrix} = \begin{bmatrix} + & + \\ - & + \end{bmatrix} \begin{bmatrix} \epsilon_t^{supply} \\ \epsilon_t^{demand} \end{bmatrix}$$



Takeaway

- We need additional restrictions on B to distinguish supply from demand shocks

Takeaway

- We need additional restrictions on B to distinguish supply from demand shocks
- If we have knowledge about the timing: zero restrictions

Takeaway

- We need additional restrictions on B to distinguish supply from demand shocks
- If we have knowledge about the timing: zero restrictions
- If we have knowledge about the slope of demand/ supply curves: sign restrictions

Takeaway

- We need additional restrictions on B to distinguish supply from demand shocks
- If we have knowledge about the timing: zero restrictions
- If we have knowledge about the slope of demand/ supply curves: sign restrictions
- Next steps:
  - What does the data tell us?

Takeaway

- We need additional restrictions on B to distinguish supply from demand shocks
- If we have knowledge about the timing: zero restrictions
- If we have knowledge about the slope of demand/ supply curves: sign restrictions
- Next steps:
  - What does the data tell us?
  - How many restrictions do we need?

Takeaway

- We need additional restrictions on B to distinguish supply from demand shocks
- If we have knowledge about the timing: zero restrictions
- If we have knowledge about the slope of demand/ supply curves: sign restrictions
- Next steps:
  - What does the data tell us?
  - How many restrictions do we need?
  - How do we combine information from the data and economic restrictions?

Observational Equivalence

$$\begin{aligned} \boldsymbol{y}_t &= \Pi \boldsymbol{w}_{t-1} + \boldsymbol{u}_t & \boldsymbol{u}_t &\sim N(\boldsymbol{0}, \Sigma) \\ &= \Pi \boldsymbol{w}_{\boldsymbol{t-1}} + B\boldsymbol{\epsilon}_{\boldsymbol{t}} & \boldsymbol{\epsilon}_{\boldsymbol{t}} &\sim N(\boldsymbol{0}, I) \end{aligned}$$

Observational Equivalence

$$\begin{aligned} \boldsymbol{y}_t &= \Pi \boldsymbol{w}_{t-1} + \boldsymbol{u}_t & \boldsymbol{u}_t &\sim N(\boldsymbol{0}, \Sigma) \\ &= \Pi \boldsymbol{w}_{\boldsymbol{t-1}} + B\boldsymbol{\epsilon}_{\boldsymbol{t}} & \boldsymbol{\epsilon}_{\boldsymbol{t}} &\sim N(\boldsymbol{0}, I) \end{aligned}$$

Covariance restriction

$$BB' = \Sigma \tag{5}$$

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \tag{6}$$

$$-\Sigma_{ii}^{0.5} \leq b_{ij} \leq \Sigma_{ii}^{0.5} \tag{7}$$

Observational Equivalence

$$BB' = \Sigma \tag{8}$$

Observational equivalence

- K^2 elements in B (K: number of variables)

Observational Equivalence

$$BB' = \Sigma \tag{8}$$

Observational equivalence

- K^2 elements in B (K: number of variables)
- $\frac{K(K+1)}{2}$ distinct elements in Σ

Observational Equivalence

$$BB' = \Sigma \tag{8}$$

Observational equivalence

- K^2 elements in B (K: number of variables)
- $\frac{K(K+1)}{2}$ distinct elements in Σ
- $\frac{K(K-1)}{2}$ restrictions missing for point identification

$$\{B\} \text{ s.t. } BB' = \Sigma \tag{9}$$

Point and set identification

Several approaches to pin down B

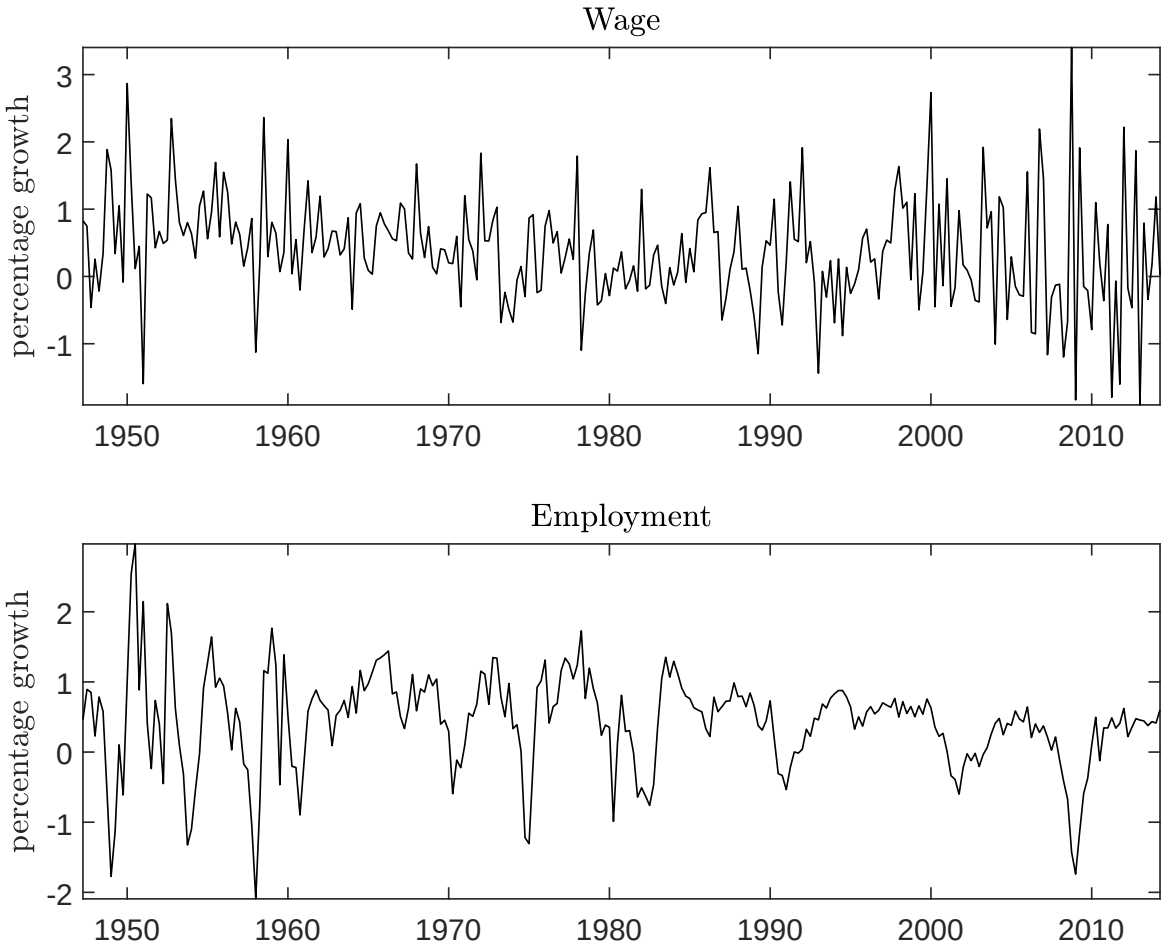
1. Timing restrictions (**recursive identification**)
2. Long-run restrictions
3. **Sign restrictions**
4. External instruments
5. Statistical identification (Heteroskedasticity, non-Gaussianity)

Point and set identification

Several approaches to pin down B

1. Timing restrictions (**recursive identification**)
2. Long-run restrictions
3. **Sign restrictions**
4. External instruments
5. Statistical identification (Heteroskedasticity, non-Gaussianity)

CODE 1 in MATLAB


Wage
percentage growth
3
2
1
0
-1
1950
1960
1970
1980
1990
2000
2010
Employment
percentage growth
2
1
0
-1
-2
1950
1960
1970
1980
1990
2000
2010

Timing restrictions: Recursive identification

$$\begin{bmatrix} Wage \\ Empl \end{bmatrix}_t = \Pi \begin{bmatrix} Wage \\ Empl \end{bmatrix}_{t-1} + \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \tag{10}$$

- Data from Baumeister & Hamilton (2015)
- Shock labelling/ economic interpretation?

Timing restrictions: Recursive identification

$$\begin{bmatrix} Wage \\ Empl \end{bmatrix}_t = \Pi \begin{bmatrix} Wage \\ Empl \end{bmatrix}_{t-1} + \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \tag{10}$$

- Data from Baumeister & Hamilton (2015)
- Shock labelling/ economic interpretation?

CODE 2 in MATLAB


shock 1 → Wage
percent
0.8
0.6
0.4
0.2
0
0
2
4
6
shock 2 → Wage
percent
0.05
0
-0.05
-0.1
0
2
4
6
shock 1 → Employment
percent
0.1
0.08
0.06
0.04
0
2
4
6
shock 2 → Employment
percent
0.3
0.25
0.2
0.15
0.1
0
2
4
6

Sign Restrictions

- Back to the problem: observational equivalence

$$\{B\} \textbf{ s.t. } BB' = \Sigma$$

- Convenient way of exploring the set of observational equivalence: rotation matrices Q s.t. $QQ' = I$

Sign Restrictions

- Back to the problem: observational equivalence

$$\{B\} \textbf{ s.t. } BB' = \Sigma$$

- Convenient way of exploring the set of observational equivalence: rotation matrices Q s.t. $QQ' = I$

$$BQQ'B' = \Sigma \tag{11}$$
$$B_1 = BQ_1 \tag{12}$$
$$B_2 = BQ_2 \tag{13}$$
$$B_3 = BQ_3 \tag{14}$$
$$\dots \tag{15}$$

Sign Restrictions

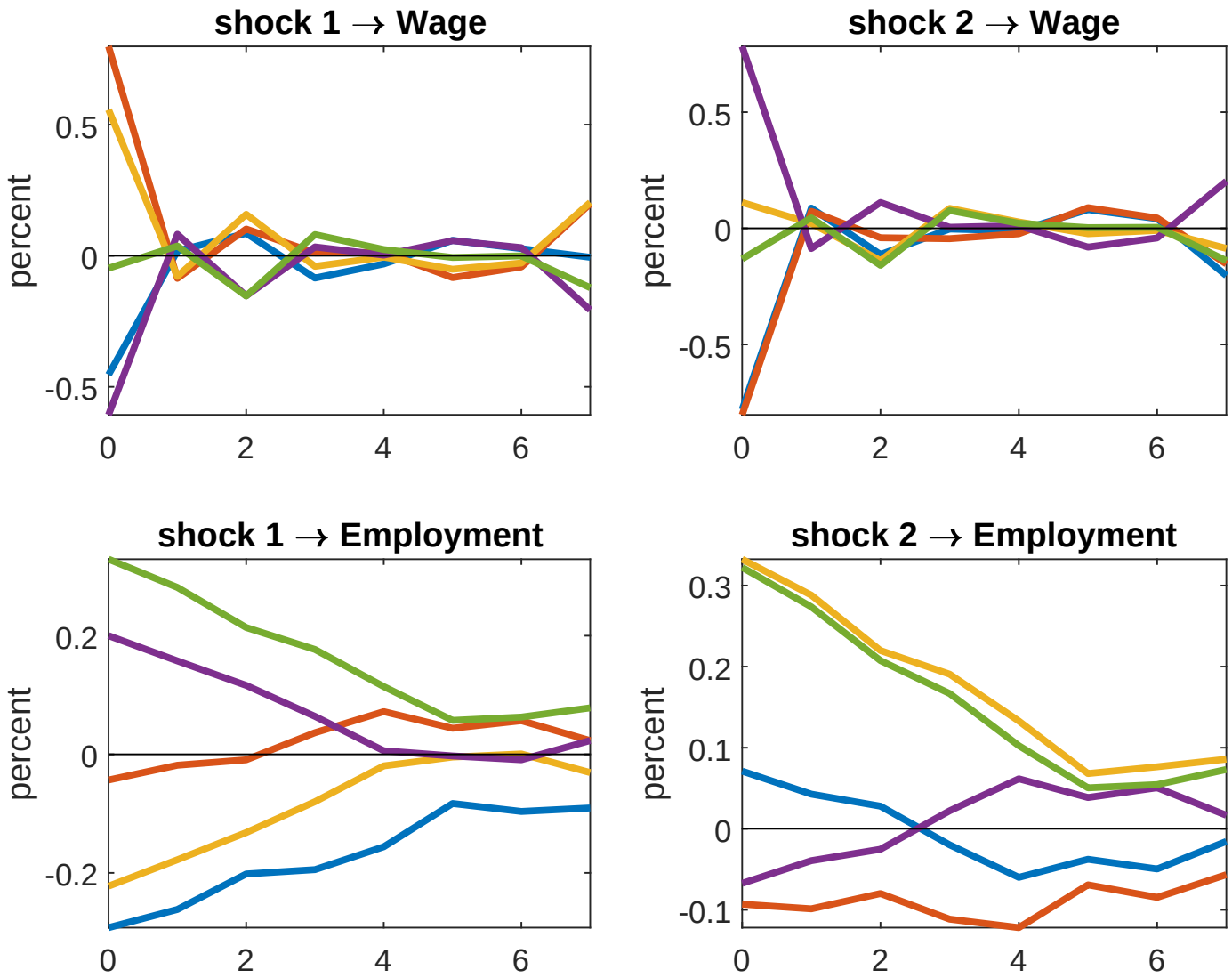
- Back to the problem: observational equivalence

$$\{B\} \textbf{ s.t. } BB' = \Sigma$$

- Convenient way of exploring the set of observational equivalence: rotation matrices Q s.t. $QQ' = I$

$$BQQ'B' = \Sigma \tag{11}$$
$$B_1 = BQ_1 \tag{12}$$
$$B_2 = BQ_2 \tag{13}$$
$$B_3 = BQ_3 \tag{14}$$
$$\dots \tag{15}$$

CODE 3 in MATLAB


shock 1 → Wage
shock 2 → Wage
shock 1 → Employment
shock 2 → Employment
percent
0.5
0
-0.5
0.2
-0.2
0.3
0.1
-0.1
0
2
4
6

Sign Restrictions

- Instead of finding **one** B, is there a way of shrinking the set $\{B\}$?

Sign Restrictions

- Instead of finding **one** B, is there a way of shrinking the set $\{B\}$?
- Economic restrictions:
- **Restriction I:** A labour supply shock moves wages and employment in the **same** direction within the quarter

Sign Restrictions

- Instead of finding **one** B, is there a way of shrinking the set $\{B\}$?
- Economic restrictions:
- **Restriction I:** A labour supply shock moves wages and employment in the **same** direction within the quarter
- **Restriction II:** A labour demand shock moves wages and employment in the **opposite** direction within the quarter

Sign restrictions (cont.)

$$\{B\} \textbf{ s.t. } (\text{I}) \; BB' = \Sigma \tag{16}$$

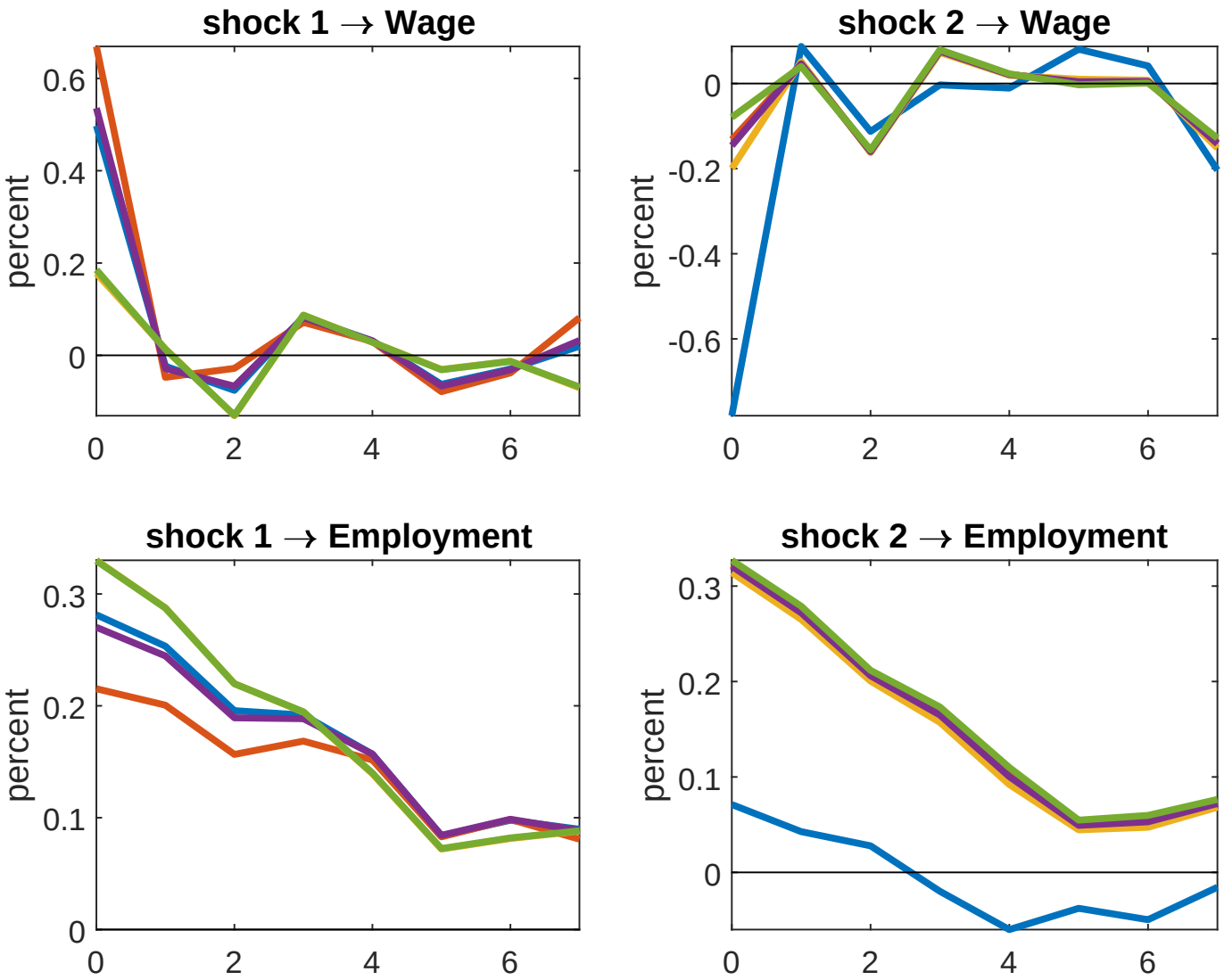
$$(\text{II}) \; B = \begin{bmatrix} + & - \\ + & + \end{bmatrix} \tag{17}$$

Sign restrictions (cont.)

$$\{B\} \textbf{ s.t. } (\text{I}) \; BB' = \Sigma \tag{16}$$

$$(\text{II}) \; B = \begin{bmatrix} + & - \\ + & + \end{bmatrix} \tag{17}$$

CODE 4 in MATLAB


shock 1 → Wage
percent
0.6
0.4
0.2
0
0
2
4
6
shock 2 → Wage
percent
0
-0.2
-0.4
-0.6
0
2
4
6
shock 1 → Employment
percent
0.3
0.2
0.1
0
0
2
4
6
shock 2 → Employment
percent
0.3
0.2
0.1
0
0
2
4
6

Takeaways

- Set of observational equivalence is potentially too large to draw meaningful economic conclusions

Takeaways

- Set of observational equivalence is potentially too large to draw meaningful economic conclusions
- It can be explored using rotation matrices

Takeaways

- Set of observational equivalence is potentially too large to draw meaningful economic conclusions
- It can be explored using rotation matrices
- Sign restrictions shrink the set of observational equivalence

Takeaways

- Set of observational equivalence is potentially too large to draw meaningful economic conclusions
- It can be explored using rotation matrices
- Sign restrictions shrink the set of observational equivalence
- Consensus that IRFs which do not satisfy the restrictions are excluded from the posterior

Takeaways

- Set of observational equivalence is potentially too large to draw meaningful economic conclusions
- It can be explored using rotation matrices
- Sign restrictions shrink the set of observational equivalence
- Consensus that IRFs which do not satisfy the restrictions are excluded from the posterior

Takeaways (cont.)

- Until now:
  - Condition on $\Sigma^{OLS}, \Pi^{OLS} \rightarrow$ no estimation uncertainty

Takeaways (cont.)

- Until now:
  - Condition on $\Sigma^{OLS}, \Pi^{OLS} \rightarrow$ no estimation uncertainty
  - No probabilistic statements
    - IRFs incompatible with sign restrictions are excluded
    - BUT: are the ones we keep equally likely a priori?
    - How do we analyse their distribution a posteriori?

Takeaways (cont.)

- Until now:
  - Condition on $\Sigma^{OLS}, \Pi^{OLS} \rightarrow$ no estimation uncertainty
  - No probabilistic statements
    - IRFs incompatible with sign restrictions are excluded
    - BUT: are the ones we keep equally likely a priori?
    - How do we analyse their distribution a posteriori?
- Next steps:
- How to implement sign restrictions:
  - Importance of prior distribution:
  - The shape of the prior influences the posterior
  - Even for $T \rightarrow \infty$ (set identification)

1. Introduction: Sign-restricted SVARs
2. Arias et al. (2018)
3. Direct Approach
   - Baumeister & Hamilton (2015)
   - Bruns & Piffer (2023)
4. Combining Sign and Zero Restrictions
   - Excursion: Importance Sampling
5. Summary

Sign restrictions

Different approaches:

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_{t-1} + \boldsymbol{u}_t \qquad \boldsymbol{u}_t \sim N(\boldsymbol{0}, \Sigma) \tag{18}$$

$$B = chol(\Sigma)Q \qquad QQ' = I \tag{19}$$

Sign restrictions

Different approaches:

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_{t-1} + \boldsymbol{u}_t \qquad \boldsymbol{u}_t \sim N(\boldsymbol{0}, \Sigma) \tag{18}$$

$$B = chol(\Sigma)Q \qquad QQ' = I \tag{19}$$

- Normal inverse-Wishart Uniform approach
  - Uhlig (2005), Rubio-Ramirez et al. (2010), Arias et al. (2018)
  - Prior on Π, Σ, Q

Sign restrictions

Different approaches:

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_{t-1} + \boldsymbol{u}_t \qquad \boldsymbol{u}_t \sim N(\boldsymbol{0}, \Sigma) \tag{18}$$

$$B = chol(\Sigma)Q \qquad QQ' = I \tag{19}$$

- Normal inverse-Wishart Uniform approach
  - Uhlig (2005), Rubio-Ramirez et al. (2010), Arias et al. (2018)
  - Prior on Π, Σ, Q
- Direct approach
  - Baumeister & Hamilton 2015, Baumeister & Hamilton 2019
  - Bruns & Piffer (2023)
  - Prior on Π, B

The NiWU approach

The reduced form model

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_{t-1} + \boldsymbol{u}_t \qquad \boldsymbol{u}_t \sim N(\boldsymbol{0}, \Sigma) \tag{20}$$

Mapping

$$B = chol(\Sigma)Q \qquad QQ' = I \tag{21}$$

The NiWU approach

The reduced form model

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_{t-1} + \boldsymbol{u}_t \qquad \boldsymbol{u}_t \sim N(\boldsymbol{0}, \Sigma) \tag{20}$$

Mapping

$$B = chol(\Sigma)Q \qquad QQ' = I \tag{21}$$

Prior and posterior

$$\begin{aligned} vec(\Pi) &\sim N & vec(\Pi)|\Sigma, Y &\sim N \\ \Sigma &\sim iW & \Sigma|\Pi, Y &\sim iW \\ Q|\Sigma &\sim U + \text{restr.} & Q|\Sigma, Y &\sim U + \text{restr.} \end{aligned}$$

The NiWU approach (cont.)

Pros

- Known (conditional) posterior distributions

The NiWU approach (cont.)

Pros

- Known (conditional) posterior distributions
- Amenable to Gibbs sampler

The NiWU approach (cont.)

Pros

- Known (conditional) posterior distributions
- Amenable to Gibbs sampler
- Efficient sampling, convergence

The NiWU approach (cont.)

Pros

- Known (conditional) posterior distributions
- Amenable to Gibbs sampler
- Efficient sampling, convergence
- Well-established guidance on choice of reduced form prior parameters (Litterman 1986, Kadiyala & Karlsson 1997,Kilian & Lütkepohl 2017)

The NiWU approach (cont.)

Cons

- Priors on structural parameters $p(B)$ matter (set identification)

The NiWU approach (cont.)

Cons

- Priors on structural parameters $p(B)$ matter (set identification)
- But: they are **implicit** in the NiWU approach

$$p^{i}_{NiWU}(B)$$

The NiWU approach (cont.)

Cons

- Priors on structural parameters $p(B)$ matter (set identification)
- But: they are **implicit** in the NiWU approach

$$p^{i}_{NiWU}(B)$$

- The NiWU potentially gives (too much) probability mass in unreasonable regions of the structural parameters

The NiWU approach (cont.)

- Note that the covariance restriction implies

$$\Sigma = BB' \tag{22}$$

$$-\Sigma_{ii}^{0.5} \leq b_{ij} \leq \Sigma_{ii}^{0.5} \tag{23}$$

The NiWU approach (cont.)

- Note that the covariance restriction implies

$$\Sigma = BB' \tag{22}$$

$$-\Sigma_{ii}^{0.5} \leq b_{ij} \leq \Sigma_{ii}^{0.5} \tag{23}$$

- Data variation constrains the impact effects of structural shocks

The NiWU approach (cont.)

- Note that the covariance restriction implies

$$\Sigma = BB' \tag{22}$$

$$-\Sigma_{ii}^{0.5} \leq b_{ij} \leq \Sigma_{ii}^{0.5} \tag{23}$$

- Data variation constrains the impact effects of structural shocks
- Implicit prior shape within these constraints (i.e. conditional on Σ^{OLS})?

The NiWU approach (cont.)

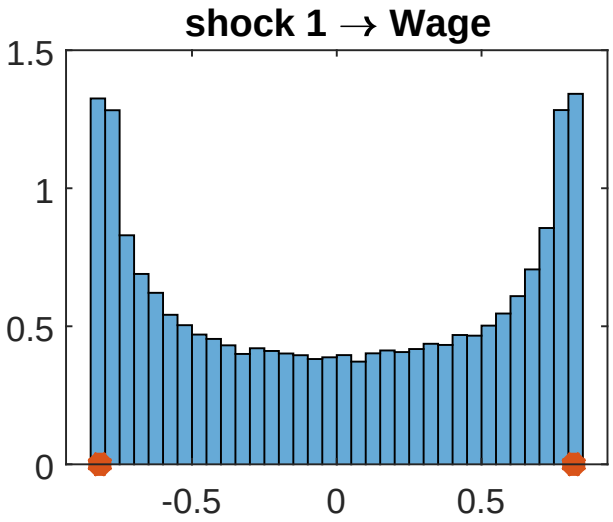
- Note that the covariance restriction implies

$$\Sigma = BB' \tag{22}$$

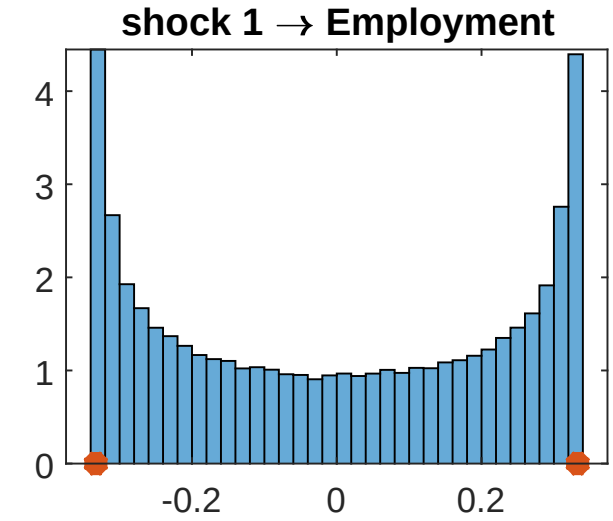
$$-\Sigma_{ii}^{0.5} \leq b_{ij} \leq \Sigma_{ii}^{0.5} \tag{23}$$

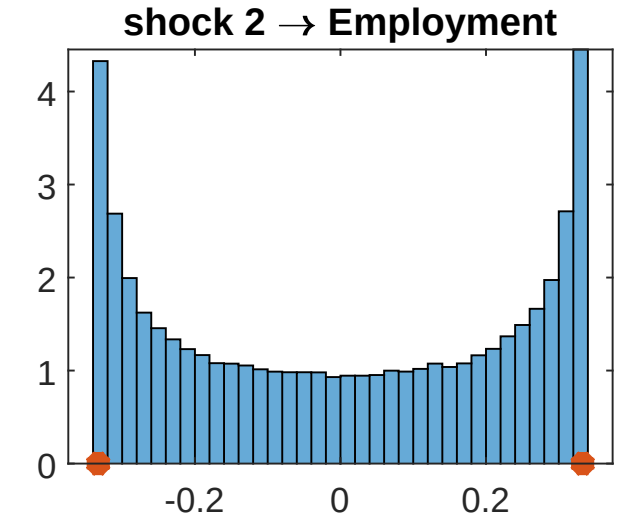
- Data variation constrains the impact effects of structural shocks
- Implicit prior shape within these constraints (i.e. conditional on Σ^{OLS})?
- CODE 5 in MATLAB

Implicit prior on B (no sign restrictions)









Implicit prior on B (sign restrictions)



The NiWU approach: algorithm

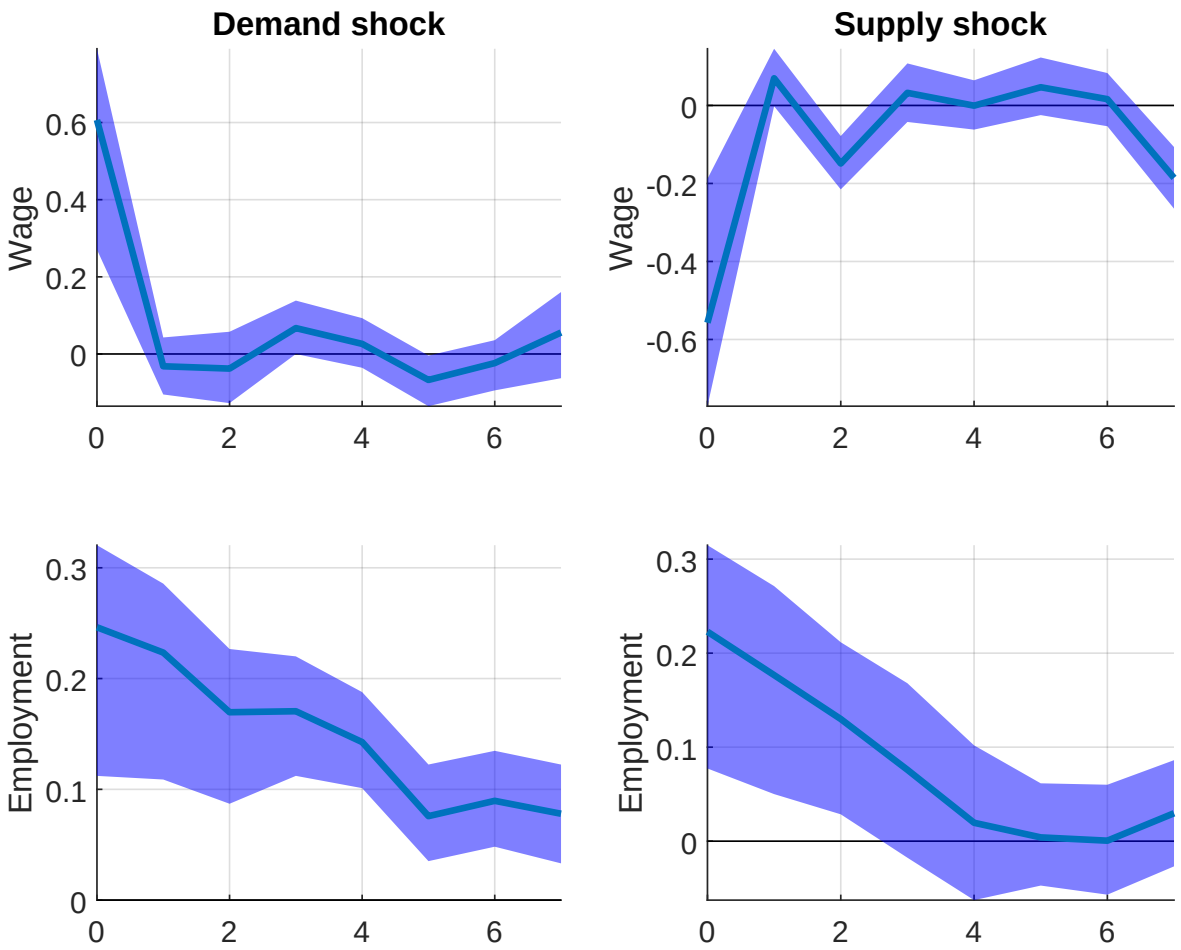
NiWU algorithm:

1. run a Gibbs sampler to explore $p(\pi, \Sigma | Y)$
2. randomly extract (π, Σ)
3. draw an orthogonal matrix Q
   - 4a. if (π, Σ, Q) satisfy the sign restrictions, store and proceed to Step 5;
   - 4b. if (π, Σ, Q) do not satisfy the sign restrictions, repeat Step 3 up to n_3 times. Otherwise discard (π_d, Σ_d) and move back to Step 2;
5. repeat Steps 2 to 4 a desired number of times

Posterior of B



Posterior IRFs



Open Questions for the NiWU approach

1 Horizon to be restricted
   - Sign restrictions can be imposed on impact (horizon 0)
   - but some applications impose the restrictions for longer (horizon ≥ 0)
   - Uhlig (2005) imposes that a monetary policy shock
     1. Raises the federal funds rate
     2. Reduces nonborrowed reserves
     3. Lowers prices

     for half a year (horizons 0 - 5)

Open Questions for the NiWU approach (cont.)

2 Combination of sign and zero restrictions
- Arias et al. (2014) consider cases where some elements of B (or A) are sign-restricted and others are 0-restricted
- The algorithm needs to be adapted
- An importance sampler is needed

1. Introduction: Sign-restricted SVARs
2. Arias et al. (2018)
3. Direct Approach
   - Baumeister & Hamilton (2015)
   - Bruns & Piffer (2023)
4. Combining Sign and Zero Restrictions
   - Excursion: Importance Sampling
5. Summary

The direct approach

The structural model

$$A\boldsymbol{y}_t = \tilde{\Pi}\boldsymbol{w}_{t-1} + \boldsymbol{\epsilon}_t \qquad \boldsymbol{\epsilon}_t \sim N(\boldsymbol{0}, D) \tag{24}$$

$$\boldsymbol{y}_t = \Pi\boldsymbol{w}_{t-1} + B\boldsymbol{\epsilon}_t \tag{25}$$

The direct approach

The structural model

$$A\boldsymbol{y}_t = \tilde{\Pi}\boldsymbol{w}_{t-1} + \boldsymbol{\epsilon}_t \qquad \boldsymbol{\epsilon}_t \sim N(\boldsymbol{0}, D) \tag{24}$$

$$\boldsymbol{y}_t = \Pi\boldsymbol{w}_{t-1} + B\boldsymbol{\epsilon}_t \tag{25}$$

Prior and posterior

$$p(A) \qquad p(A|Y)$$

$$d_{ii}|A \sim IG \qquad d_{ii}|A, Y \sim IG$$

$$vec(\tilde{\Pi})|A, D \sim N \qquad vec(\tilde{\Pi})|A, D, Y \sim N$$

The direct approach (cont.)

Pros

- Expresses priors over structural parameters with an economic interpretation

The direct approach (cont.)

Pros

- Expresses priors over structural parameters with an economic interpretation
  - Is their really no prior probability that the oil supply elasticity is above 0.0258? (Kilian & Murphy 2012, Baumeister & Hamilton 2019)

The direct approach (cont.)

Pros

- Expresses priors over structural parameters with an economic interpretation
  - Is their really no prior probability that the oil supply elasticity is above 0.0258? (Kilian & Murphy 2012, Baumeister & Hamilton 2019)
  - How likely is a short run wage elasticity of labour demand above 4? (Baumeister & Hamilton 2015)

Cons

- Drawing from the posterior involves a Metropolis-Hastings step

The direct approach (cont.)

Pros

- Expresses priors over structural parameters with an economic interpretation
  - Is their really no prior probability that the oil supply elasticity is above 0.0258? (Kilian & Murphy 2012, Baumeister & Hamilton 2019)
  - How likely is a short run wage elasticity of labour demand above 4? (Baumeister & Hamilton 2015)

Cons

- Drawing from the posterior involves a Metropolis-Hastings step
- Potentially problematic if
  - model is large
  - posterior is irregularly shaped (multimodal, flat in certain regions)

The direct approach (cont.)

Direct algorithm:

1. run a Metropolis-Hastings sampler to explore $p(A|Y)$. Store the draws in $\{A_d\}_{d=1}^{n}$;
   1a. Generate candidate $A_{cand} = A_{d-1} + v_t$
   1b. Set

$$A_d = \begin{cases} A_{cand} \text{ with probability } \alpha \\ A_{d-1} \text{ else,} \end{cases}$$

   where $\alpha = \frac{p(A_{cand}|Y)}{p(A_{d-1}|Y)}$
2. draw the elements $d_{ii,d}$ of $D_d|A_d, Y$ from their posterior inverse Gamma distributions;
3. draw $vec(\tilde{\Pi}_d)|A_d, D_d, Y$ from its posterior normal distribution
4. repeat Steps 2 to 3 n times.

The direct approach (cont.)

Direct algorithm:

1. run a Metropolis-Hastings sampler to explore $p(A|Y)$. Store the draws in $\{A_d\}_{d=1}^{n}$;
   1a. Generate candidate $A_{cand} = A_{d-1} + v_t$
   1b. Set

$$A_d = \begin{cases} A_{cand} \text{ with probability } \alpha \\ A_{d-1} \text{ else,} \end{cases}$$

   where $\alpha = \frac{p(A_{cand}|Y)}{p(A_{d-1}|Y)}$
2. draw the elements $d_{ii,d}$ of $D_d|A_d, Y$ from their posterior inverse Gamma distributions;
3. draw $vec(\tilde{\Pi}_d)|A_d, D_d, Y$ from its posterior normal distribution
4. repeat Steps 2 to 3 n times.

CODE BHcode in MATLAB

1. Introduction: Sign-restricted SVARs
2. Arias et al. (2018)
3. Direct Approach
   - Baumeister & Hamilton (2015)
   - Bruns & Piffer (2023)
4. Combining Sign and Zero Restrictions
   - Excursion: Importance Sampling
5. Summary

New Approach for Priors on IRFs

The reduced form model

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_{t-1} + B\boldsymbol{\epsilon}_t \qquad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, I) \tag{26}$$

New Approach for Priors on IRFs

The reduced form model

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_{t-1} + B\boldsymbol{\epsilon}_t \qquad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, I) \tag{26}$$

Prior and posterior

$$p(B) \qquad p(B|Y)$$
$$vec(\Pi) \sim N \qquad vec(\Pi)|Y, B \sim N$$

New Approach for Priors on IRFs (cont.)

Pros

- Expresses (general) priors over structural parameters with an economic interpretation

New Approach for Priors on IRFs (cont.)

Pros

- Expresses (general) priors over structural parameters with an economic interpretation
- We propose a specific prior informed by the volatility of the data

New Approach for Priors on IRFs (cont.)

Pros

- Expresses (general) priors over structural parameters with an economic interpretation
- We propose a specific prior informed by the volatility of the data
- Posterior sampler (importance sampler) is efficient and just as tractable as the NiWU prior

New Approach for Priors on IRFs (cont.)

Pros

- Expresses (general) priors over structural parameters with an economic interpretation
- We propose a specific prior informed by the volatility of the data
- Posterior sampler (importance sampler) is efficient and just as tractable as the NiWU prior
- Sharpens posterior inference
- Allows to explore the role of the prior explicitly

Main idea behind the prior

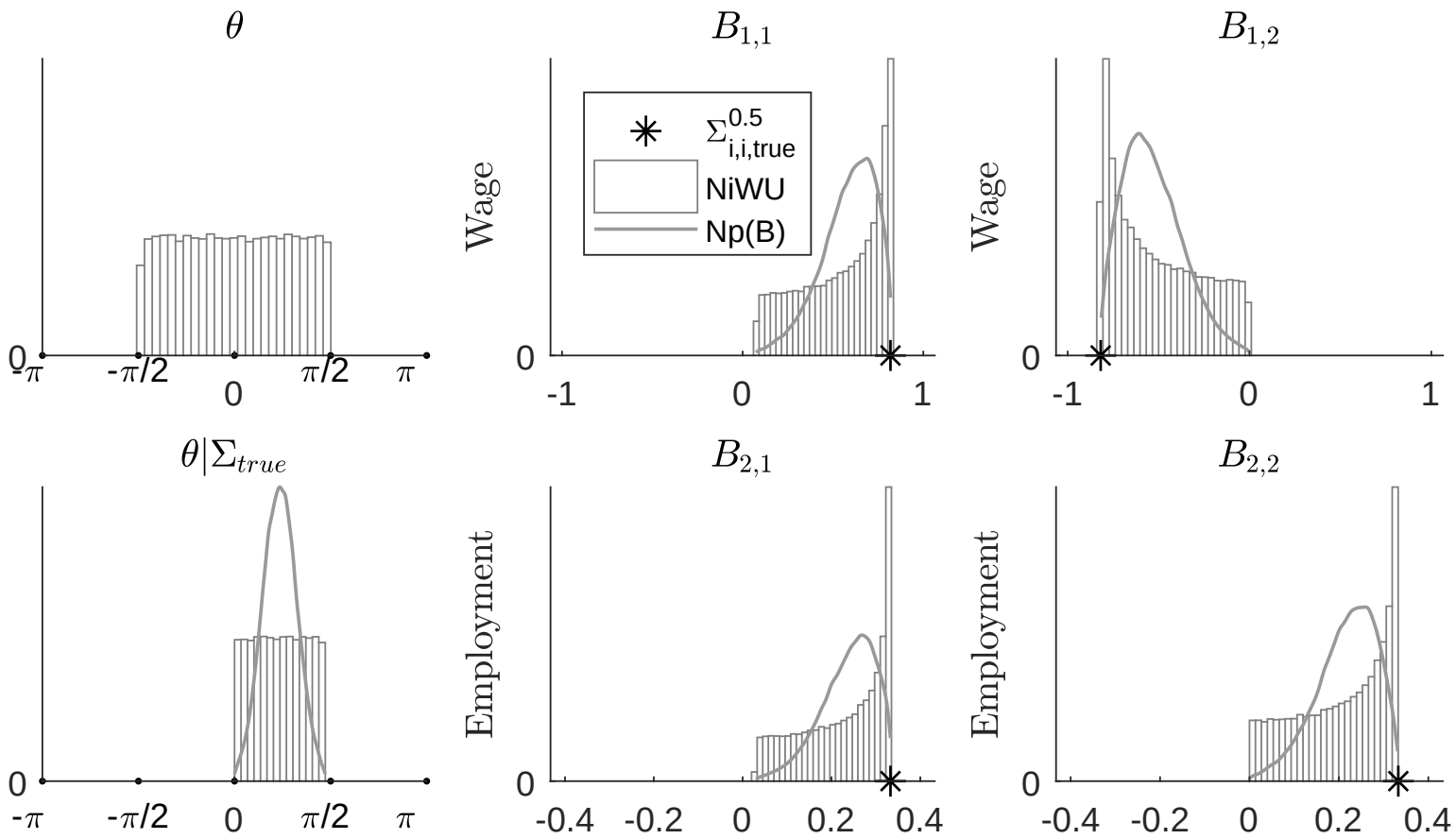
- Recall covariance restriction

$$\Sigma = BB' \tag{27}$$

$$-\Sigma_{ii}^{0.5} \leq b_{ij} \leq \Sigma_{ii}^{0.5} \tag{28}$$

- $\gamma_i = \hat{\Sigma}_{ii,training}^{0.5}$
- $b_{ij} \sim N(\mu_{ij})$ s.t. 95% prob. mass in $[0; \pm\psi_2\gamma_i]$

Illustrative example



Main idea behind the sampler

- Generate proposal draws from NiWU (computationally efficient, tractable)

Main idea behind the sampler

- Generate proposal draws from NiWU (computationally efficient, tractable)
- Map them into draws from our posterior

Main idea behind the sampler

- Generate proposal draws from NiWU (computationally efficient, tractable)
- Map them into draws from our posterior
- Uses an importance sampler

Main idea behind the sampler

- Generate proposal draws from NiWU (computationally efficient, tractable)
- Map them into draws from our posterior
- Uses an importance sampler
- Efficient, independent draws

1. Introduction: Sign-restricted SVARs
2. Arias et al. (2018)
3. Direct Approach
   - Baumeister & Hamilton (2015)
   - Bruns & Piffer (2023)
4. Combining Sign and Zero Restrictions
   - Excursion: Importance Sampling
5. Summary

Combining Sign and Zero Restrictions

- So far we have considered cases where either zero restrictions or pure sign restrictions are employed
- Instead we could try to combine the two approaches
- Going back to the link between reduced form and structural form: $u_t = B\varepsilon_t$
- Normalising $Var(\varepsilon_t) = I_K$
- and noting that $Var(u_t) = \Sigma$
- we have that $BB' = \Sigma$.

Combining Sign and Zero Restrictions

- The idea behind pure sign restrictions as in Arias et al. (2018):
- Use a rotation matrix Q with $QQ' = I_K$
- to generate a candidate $\tilde{B} = BQ$
- which does not change the likelihood since $\tilde{B}\tilde{B}' = BQQ'B' = BB' = \Sigma$
- then we check if the entries of $\tilde{B}$ satisfy the sign restrictions and decide whether to accept/ reject the candidate $\tilde{B}$

Combining Sign and Zero Restrictions

- Now impose an additional 0 restriction on B, e.g. $b_{12} = 0$
- In words: The second shock does not have a contemporaneous effect on the first variable
- The accept/ reject algorithm needs to be adapted
- Why? Because it is a zero probability event to draw a matrix Q such that $\tilde{B}$ satisfies the sign restriction
- Solution: draw columns q_j of Q directly ensuring that the 0 restriction holds

Algorithm 2: zero restrictions

1: Draw reduced-form parameters:

$$(\Pi, \Sigma) \sim p(\Pi, \Sigma \mid Y).$$

2: **for** $j = 1, \ldots, n$ **do**

3: Build the constraint matrix:

$$M_j = \begin{bmatrix} q_1' \\ \vdots \\ q_{j-1}' \\ F_j(\Pi, \Sigma) \end{bmatrix}.$$

4: Compute an orthonormal basis K_j for $\mathcal{N}(M_j)$.

5: Draw w_j uniformly on the unit sphere.

6: Set $q_j = K_j w_j$.

7: **end for**

8: Set $Q = [q_1, \ldots, q_n]$.

9: Construct $A_0^{-1} = P(\Sigma)Q$.

What Algorithm 2 achieves

- Algorithm 2 draws structural models that satisfy zero restrictions by construction.
- This avoids the problem:

$$\Pr(\text{IRF}_{ij}(h) = 0) = 0$$

under unconstrained random rotations.

Sign and zero restrictions: Labour market example

- Going back to the bivariate labour market example by Baumeister & Hamilton (2015)
- Reconsider the previous identifying restriction:

$$\begin{bmatrix} u_t^q \\ u_t^p \end{bmatrix} = \begin{bmatrix} b_{11} & \mathbf{0} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \epsilon_t^{supply} \\ \epsilon_t^{demand} \end{bmatrix}$$

- With this restrictions, the model is point identified and we obtain
- $B = \begin{bmatrix} 0.809 & 0 \\ 0.034 & 0.333 \end{bmatrix}$
- The only admissible rotation matrix is $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
- Any other rotation matrix would violate the zero restriction

Sign and zero restrictions: Labour market example

- Extend the bivariate model by including real GDP growth as a third variable:

$$\begin{bmatrix} Wage \\ Empl \\ GDP \end{bmatrix}_t = \Pi \begin{bmatrix} Wage \\ Empl \\ GDP \end{bmatrix}_{t-1} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \epsilon^{supply} \\ \epsilon^{demand} \\ \epsilon^{3} \end{bmatrix}_t \tag{29}$$

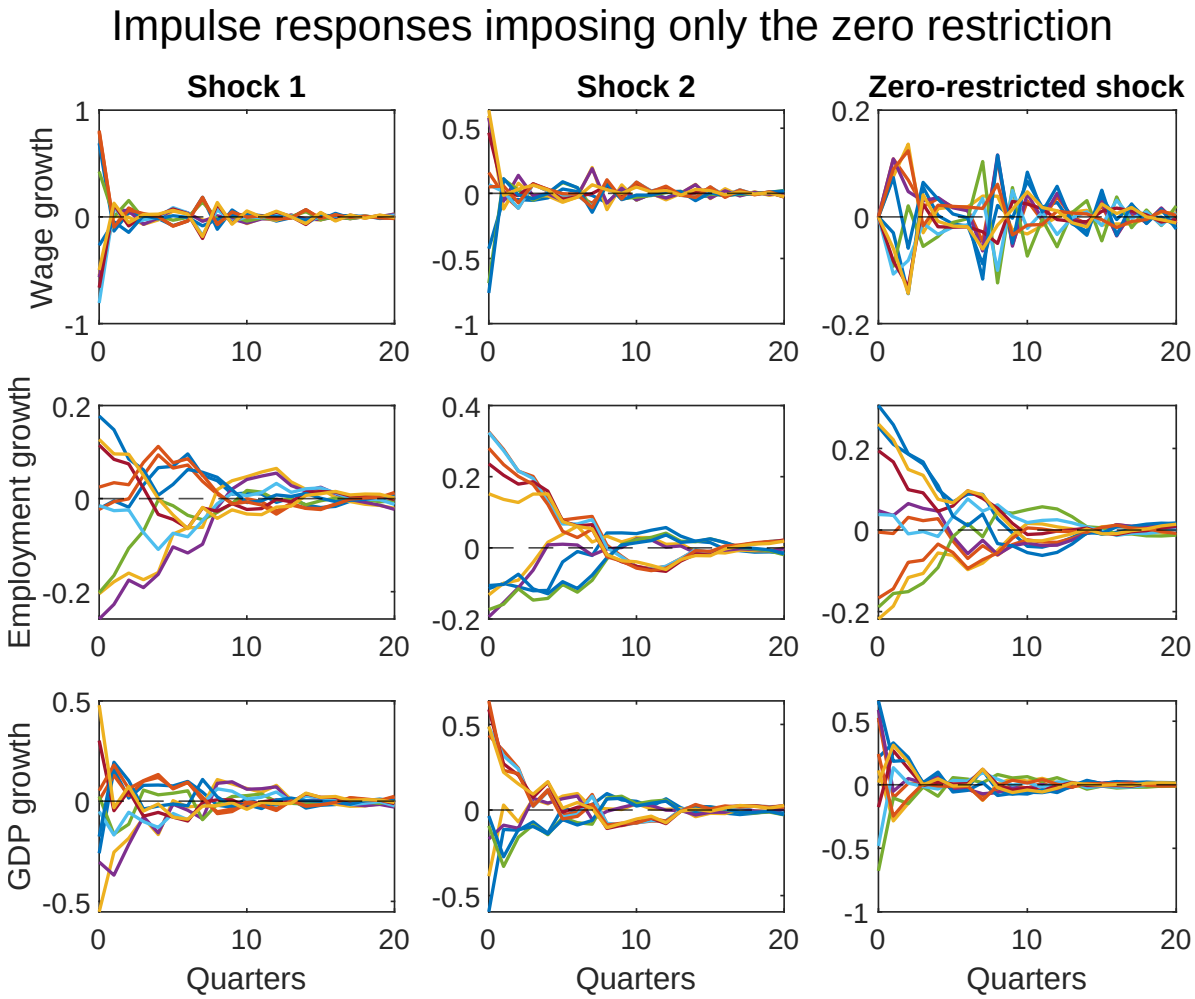
- And consider the following restrictions:

	Wage growth	Employment growth	Real GDP growth
ϵ^{supply}	$-$	$+$	$+$
ϵ^{demand}	$+$	$+$	$+$
ϵ^{3}	0	unrestricted	unrestricted

Sign and zero restrictions: Labour market example

- Now we can illustrate how to draw a set of 10 candidate rotation matrices Q satisfying the zero restriction
- CODE SignZeroRestrictions_BH2015.m in MATLAB

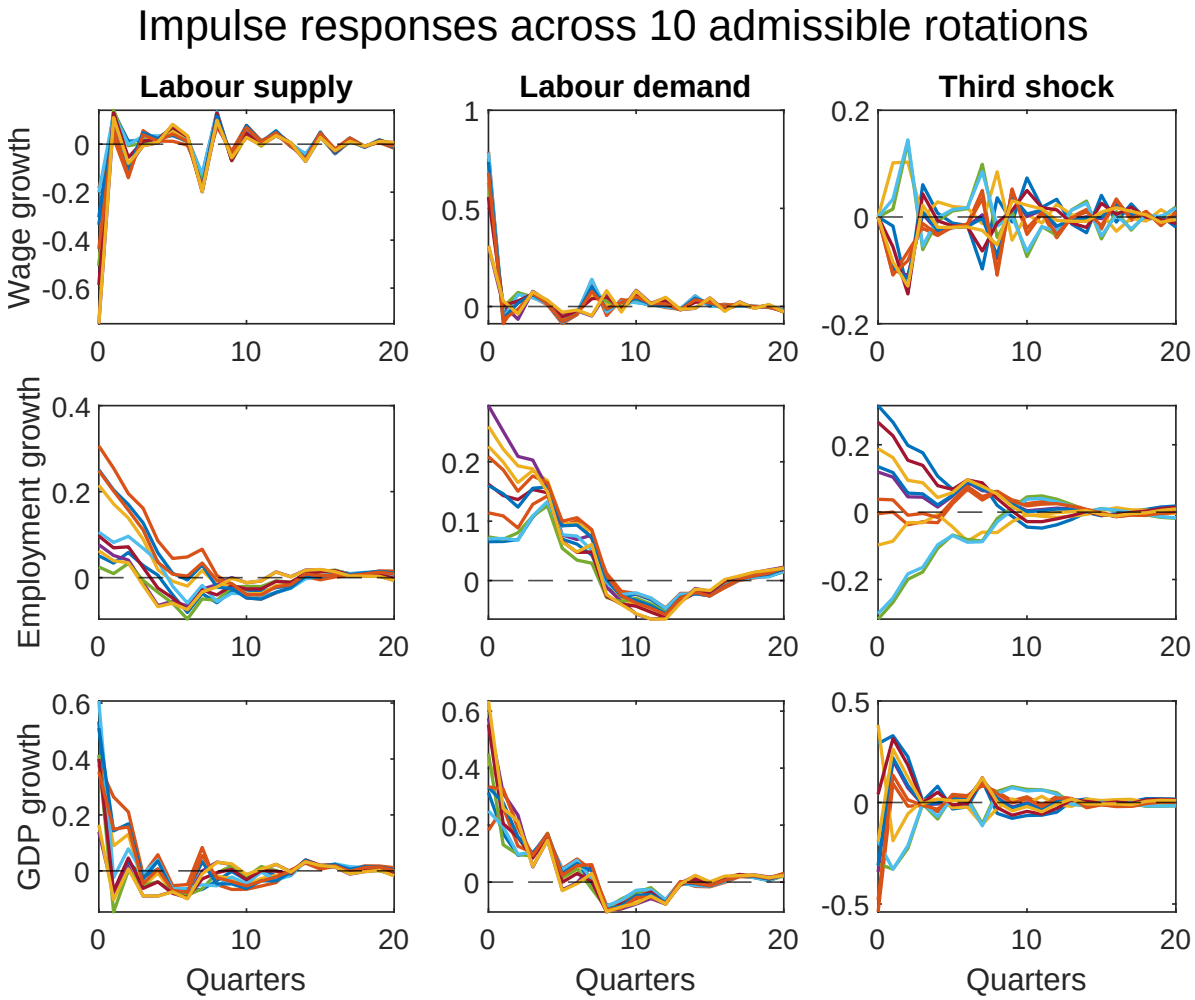
Labour market example: Only zero restriction



Sign and zero restrictions: Labour market example

- Out of the set of Q matrices satisying the zero restriction we can draw a set of 10 candidate rotation matrices Q satisfying
  - the zero restrictions
  - the sign restrictions

Sign and zero restrictions: Labour market example



Combining signs and zeros

- For now we haven't made any probabilistic statements
- We have
  - conditioned on Σ^{OLS},
  - constructed rotation matrices Q which satisfy the 0 restriction and
  - accepted/ rejected based on the sign restrictions
- One approach could be to adapt the NiWU algorithm (step 3) to only draw matrices Q satisfying the zero restrictions
- However, it can be shown that then the induced distribution over structural models is not generally the desired posterior distribution
- Instead Arias et al. (2018) use Algorithm 2 as a proposal mechanism and then apply an importance-sampling correction.
- This gives Algorithm 3 below

1. Introduction: Sign-restricted SVARs
2. Arias et al. (2018)
3. Direct Approach
   - Baumeister & Hamilton (2015)
   - Bruns & Piffer (2023)
4. Combining Sign and Zero Restrictions
   - Excursion: Importance Sampling
5. Summary

Importance weights: intuition

The importance weight adjusts for the fact that:

$$\text{proposal distribution} \neq \text{target posterior}.$$

High-weight draws are under-represented by the proposal.
Low-weight draws are over-represented by the proposal.

After weighting and resampling, the retained draws represent the desired posterior.

Excursion: Importance Sampling

- Suppose we cannot directly draw from $p(\boldsymbol{x})$
- But we have access to a proposal function $p^{propr}(\boldsymbol{x})$ which is easy to sample from
- $p^{propr}(\boldsymbol{x})$ covers the full support of $p(\boldsymbol{x})$
- We can use Importance Sampling
- We can draw from $p^{propr}(\boldsymbol{x})$ and reweight these draws to make the representative of $p(\boldsymbol{x})$
- Generates independent draws

Excursion: Importance Sampling

$$\boldsymbol{x} \sim N(\boldsymbol{\mu}, \Sigma), \tag{30}$$

where $\boldsymbol{\mu}$ is of dimension 2×1 and Σ is of dimension 2×2.
The probability density function is

$$p(\boldsymbol{x}) = (2\pi)^{-0.5k} |\Sigma|^{-0.5} exp\big(-0.5(\boldsymbol{x} - \boldsymbol{\mu})' \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \big). \tag{31}$$

Importance Sampling (cont.)

- Generate $x_1^{prop}, ..., x_i^{prop}, ...x_n^{prop}$ from $p^{prop}(x)$.
- Compute weights $\tilde{w}_i = \frac{p(x_i)}{p^{prop}(x_i)}$.
- Normalise these weights such that they sum up to one as $w_i = \frac{\tilde{w}_i}{\sum_{i=1}^{n} \tilde{w}_i}$.
- Generate $x_1, ..., x_i, ...x_n$ by resampling with replacement from $x_1^{prop}, ..., x_i^{prop}, ...x_n^{prop}$ according to the weights w_i.

ImportanceSampling.m

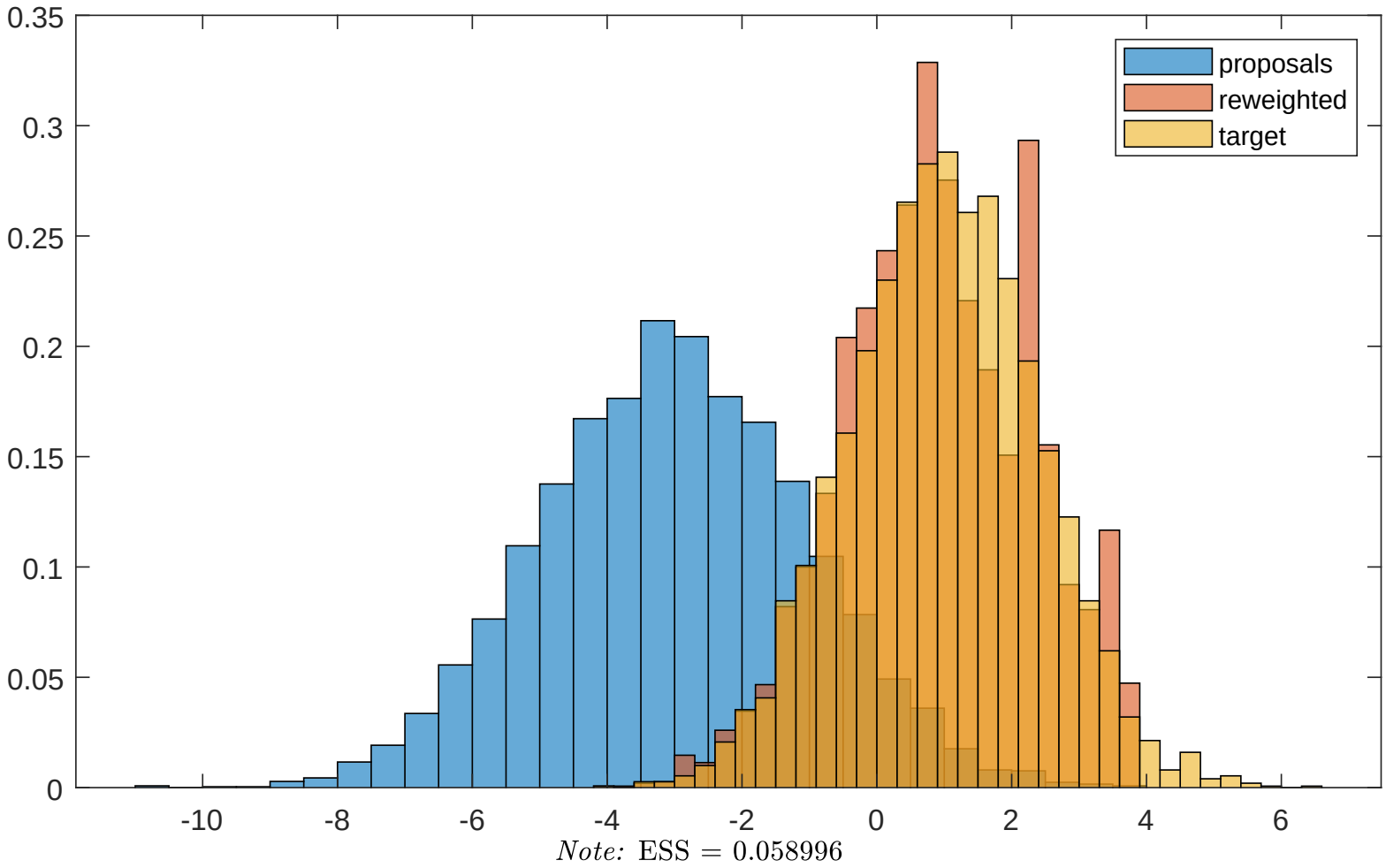
Effective sample size

- If the proposal draws are closely centred around the target draws, then the weights will be roughly equally distributed, i.e. close to $1/n$
- This ensures a good number of usable proposal draws
- Instead, if the proposal distribution is poorly chosen, i.e. too far from the target density, then weights could be concentrated:
- Some weights might be close to 1, while the majority is close to 0
- Then, few of the proposals draws are usable

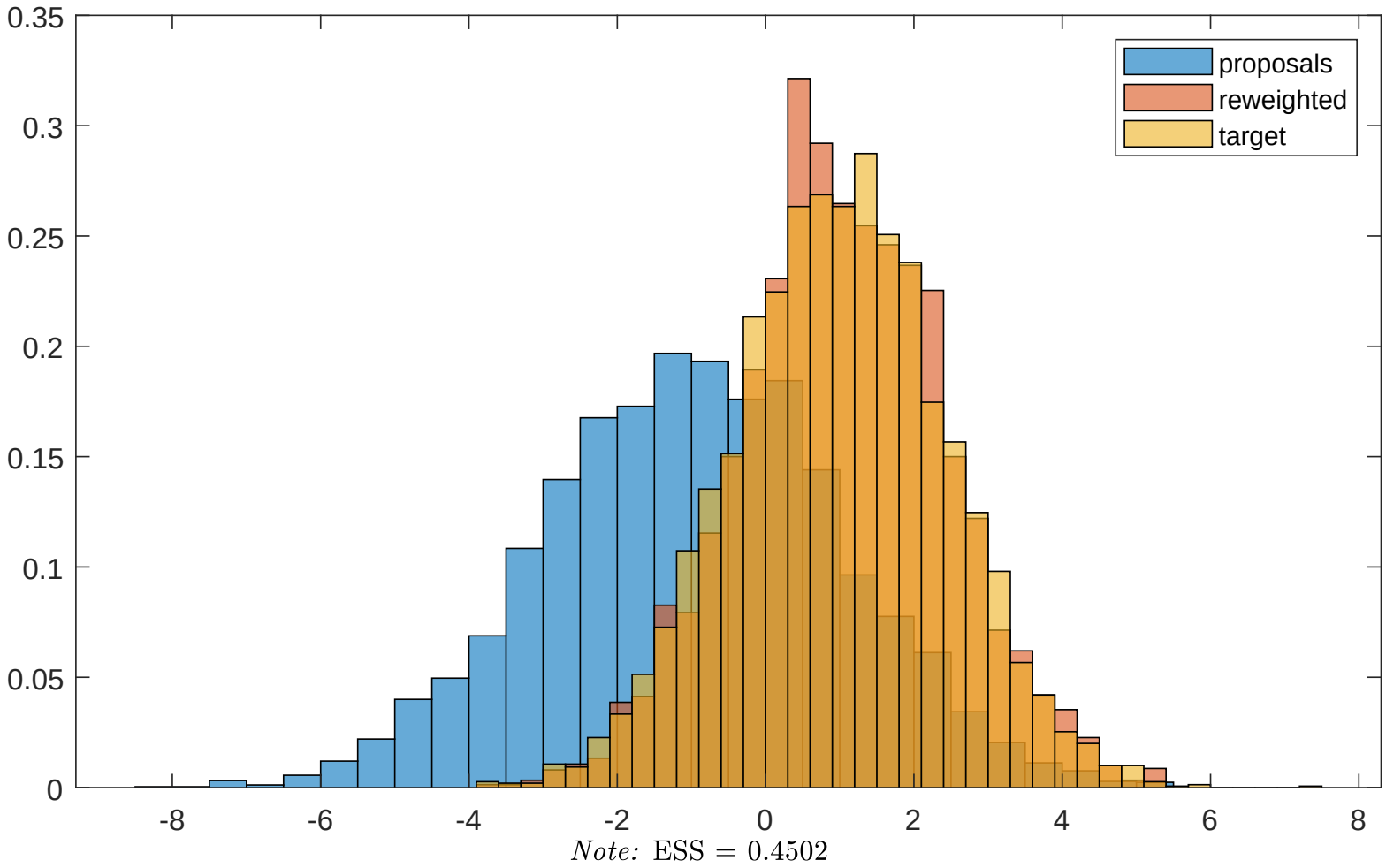
Effective sample size

- A measure for the number of usable draws is the absolute effective sample size (ESS)
- $ESS_{abs} = \frac{1}{\sum_i w_i^2}$
- or as a share of proposal draws: $ESS = ESS_{abs}/n$
- In the example, the true mean is $\mu_{true} = 1$
- We will choose a proposal centred around $\mu_{prop} = -3$, expecting a low ESS
- and a proposal centred around $\mu_{prop} = -1$, expecting the ESS to increase

Importance Sampling: $\mu_{prop} = -3$



Importance Sampling: $\mu_{prop} = -1$



Algorithm 3: sign and zero restrictions

1: Use Algorithm 2 to draw a proposal structural model satisfying zero restrictions.
2: Compute impulse responses.
3: **if** the sign restrictions fail **then**
4: set importance weight $w = 0$.
5: **else**
6: compute the ARW importance weight $w > 0$.
7: **end if**
8: Repeat for many proposal draws.
9: Normalize the weights:

$$\tilde{w}_s = \frac{w_s}{\sum_r w_r}.$$

10: Resample draws with probabilities $\tilde{w}_s$.

Algorithm 3: output

- Algorithm 3 produces unweighted posterior draws from:

$$p(\Pi, \Sigma, Q \mid Y, \text{signs}, \text{zeros}).$$

- These draws can then be used for:
  - impulse responses;
  - historical decompositions;
  - forecast error variance decompositions;

1. Introduction: Sign-restricted SVARs
2. Arias et al. (2018)
3. Direct Approach
   - Baumeister & Hamilton (2015)
   - Bruns & Piffer (2023)
4. Combining Sign and Zero Restrictions
   - Excursion: Importance Sampling
5. Summary

Summary

- Moving from reduced form VAR to structural VAR requires additional restrictions (observational equivalence)

Summary

- Moving from reduced form VAR to structural VAR requires additional restrictions (observational equivalence)
- Sign restrictions shrink the set of observational equivalence $\rightarrow$ set identification

Summary

- Moving from reduced form VAR to structural VAR requires additional restrictions (observational equivalence)
- Sign restrictions shrink the set of observational equivalence $\rightarrow$ set identification
- Shape of prior matters even for large samples

Summary

- Moving from reduced form VAR to structural VAR requires additional restrictions (observational equivalence)
- Sign restrictions shrink the set of observational equivalence $\rightarrow$ set identification
- Shape of prior matters even for large samples
- Debate about
  - Which priors to use
  - Where to impose them
  - How to sample from the posterior

Summary

- Moving from reduced form VAR to structural VAR requires additional restrictions (observational equivalence)
- Sign restrictions shrink the set of observational equivalence $\rightarrow$ set identification
- Shape of prior matters even for large samples
- Debate about
  - Which priors to use
  - Where to impose them
  - How to sample from the posterior
- Here: three approaches
  - NiWU
    - Arias et al. (2018)

Summary

- Moving from reduced form VAR to structural VAR requires additional restrictions (observational equivalence)
- Sign restrictions shrink the set of observational equivalence $\rightarrow$ set identification
- Shape of prior matters even for large samples
- Debate about
  - Which priors to use
  - Where to impose them
  - How to sample from the posterior
- Here: three approaches
  - NiWU
    - Arias et al. (2018)
  - Direct
    - Baumeister & Hamilton (2015)
    - Bruns & Piffer (2023)

Arias, J. E., Rubio-Ramírez, J. F. & Waggoner, D. F. (2018), 'Inference based on structural vector autoregressions identified with sign and zero restrictions: Theory and applications', *Econometrica* **86**(2), 685–720.

Arias, J., Rubio-Ramirez, J. F. & Waggoner, D. F. (2014), 'Inference based on SVARs identified with sign and zero restrictions: Theory and applications'.

Baumeister, C. & Hamilton, J. D. (2015), 'Sign restrictions, structural vector autoregressions, and useful prior information', *Econometrica* **83**(5), 1963–1999.

Baumeister, C. J. & Hamilton, J. D. (2019), 'Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks', *The American Economic Review* **109**(5), 1873–1910.

Bruns, M. & Piffer, M. (2023), 'A new posterior sampler for bayesian structural vector autoregressive models', *Quantitative Economics* **14**(4), 1221–1250.

Kadiyala, K. R. & Karlsson, S. (1997), 'Numerical methods for estimation and inference in Bayesian VAR-models', *Journal of Applied Econometrics* pp. 99–132.

Kilian, L. & Lütkepohl, H. (2017), *Structural vector autoregressive analysis*, Cambridge University Press.

Kilian, L. & Murphy, D. P. (2012), 'Why agnostic sign restrictions are not enough: understanding the dynamics of oil market VAR models', *Journal of the European Economic Association* **10**(5), 1166–1188.

Litterman, R. B. (1986), 'Forecasting with Bayesian vector autoregressions: five years of experience', *Journal of Business & Economic Statistics* **4**(1), 25–38.

Rubio-Ramirez, J. F., Waggoner, D. F. & Zha, T. (2010), 'Structural vector autoregressions: Theory of identification and algorithms for inference', *The Review of Economic Studies* **77**(2), 665–696.

Uhlig, H. (2005), 'What are the effects of monetary policy on output? Results from an agnostic identification procedure', *Journal of Monetary Economics* **52**(2), 381–419.