Econometric Policy Analysis under Uncertainty

Sargan Lectures of the Econometric Society - Sydney 2023

Raffaella Giacomini (UCL and Chicago Fed)[1]

[1]The lectures do not reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System

Recall: Uncertainties in causal effects estimation

- We will consider different kinds of scientific uncertainties
  - Impulse-response analysis
    - Uncertainty about identifying assumptions (Lecture 1)
    - Uncertainty about the measurement of shocks (Lecture 2)
  - Average treatment effects (Lecture 3)
    - Uncertainty about the validity of the control group
    - Uncertainty about model specification

Lecture 2. Uncertainty about measurement of shocks

Based on:

- Giacomini, R., T. Kitagawa and M. Read (2021), "Identification and inference under narrative restrictions"
- Giacomini, R., T. Kitagawa and M. Read (2022), "Narrative restrictions and proxies"

Measuring shocks

- In lecture 1. we considered impulse response analysis in SVARs identified by zero and sign restrictions, a prominent approach to identification
- Here we consider a major competing approach to identification, that attempts to measure the structural shock directly
- **Narrative measure of shocks:** Measure shocks by text analysis, changes in market expectations around policy announcements etc. (e.g., Romer and Romer, 04)
  - Pros: if these are the true shocks, can identify and estimate dynamic causal effects by local projections (point estimates)
  - Cons: these are probably not the true shocks
- We are thus uncertain about our ability to measure shocks, what can we do about it?

What to do about uncertain measurement of shocks?

- One partial answer can be given if we assume that the narrative time series is a noisy measurement of the true shocks
- **Narrative measures as instruments:** Treat the narrative time series as instruments/proxies, then use IV estimation (Mertens and Ravn, 13; Stock and Watson, 18)
  - Pros: don't need to assume shocks are correctly measured
  - Cons: instrument is probably invalid: weak and/or not exogenous (some historical episodes were not structural shocks, i.e., unanticipated changes)

What to do about uncertain measurement?

- In this lecture I will focus on robust methods when we are uncertain that a time series of narrative measures are truly shocks
- Idea: instead of using the whole time series of narrative shocks, only focus on a few time periods
- Further relax the requirement that we can measure the shock and assume we only have weaker information about the shock

Two approaches to uncertain measurement

1. **Narrative restrictions (NR):** focus on a few historical episodes where we know sign of shocks. Then impose these as identifying assumptions (Antolin Diaz and Rubio Ramirez, 18, Giacomini, Kitagawa and Read, 21)
    - Pros: robust as it imposes minimal assumptions and does not suffer from weak instrument issues
    - Cons: results in set estimates
2. **Narrative instruments/proxies (NP):** Consider the same 'minimal information' but use it to construct an instrument (Giacomini, Kitagawa and Read, 22)
    - Pros: gives point estimates
    - Cons: instrument has many zeros so very weak

Narrative restrictions

- Idea: consider only a small number of dates and transform the narrative into inequality restrictions on shocks in a SVAR
  - Different from the traditional SVARs we discussed in lecture 1, where we restrict *parameters*, not *shocks*
- Examples:
  - Shock-sign: there was a positive monetary shock on given dates, e.g., "Volcker shock" in Oct 1979 (Antolin-Diaz & Rubio-Ramirez, 2018 - AR18)
  - Historical decomposition: the change in interest rate on Oct 1979 was mostly due to a monetary shock (AR18)

The econometrics of narrative restrictions

- The problem is non-standard, in terms of identification and inference
- AR18 perform Bayesian inference
- We show that there are issues with the existing approach and propose an alternative

Bivariate example

Static SVAR + shock-sign narrative information

- For $t = 1, \ldots, T$

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}}_{y_t} = \underbrace{\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}}_{\varepsilon_t}, \quad \varepsilon_t \overset{iid}{\sim} N(0, I)$$

- NR: Assume $1 \leq K \leq T$ shock-sign restrictions for ε_{1t} imposed for the first K periods, i.e., $\text{sign}(\varepsilon_{1t})$ is observed for $t = 1, \ldots, K$

- NP: construct a proxy as

$$z_t = \begin{cases} \text{sign}(\varepsilon_{1t}) & \text{for } t = 1, \ldots, K \\ 0 & \text{for } t = K+1, \ldots, T \end{cases}$$

The NR approach

Bivariate example

- Same simplified static setup as in lecture 1

$$Ay_t = \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} N(0, I)$$

- Reduced form is $y_t \overset{iid}{\sim} N(0, \Sigma)$ with $\Sigma = A^{-1}(A^{-1})'$
- $\phi = \text{vech}(\Sigma_{tr})$ is the reduced-form parameter, where $\Sigma_{tr}\Sigma_{tr}' = \Sigma$ and

$$\Sigma_{tr} = \begin{bmatrix} \sigma_{11} & 0 \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

Identification problem

- Multiple structural parameters compatible with reduced-form parameter:

$$A^{-1} = \Sigma_{tr} Q$$

where Q is an orthonormal matrix

$$Q \in \left\{ \begin{bmatrix} \cos\delta & -\sin\delta \\ \sin\delta & \cos\delta \end{bmatrix} \right\} \cup \left\{ \begin{bmatrix} \cos\delta & \sin\delta \\ \sin\delta & -\cos\delta \end{bmatrix} \right\}$$

with $\delta \in [-\pi, \pi]$

- Object of interest α is the impulse-response of y_{2t} with respect to a one-unit shock to the first variable (unit-effect IR)

$$\alpha = \frac{(A^{-1})_{(2,1)}}{(A^{-1})_{(1,1)}} = \frac{\sigma_{21}}{\sigma_{11}} + \frac{\sigma_{22}}{\sigma_{11}} \tan\delta$$

'Traditional' restrictions

- Imposed on (functions of) structural parameters
- E.g, sign restrictions: $A_{12} \geq 0$ and $A_{21} \leq 0$ (+sign norm.)

$$\delta \in \left[0, \arctan\left(\frac{\sigma_{22}}{\sigma_{21}}\right)\right], \quad \alpha \in \left[\sigma_{11} \cos\left(\arctan\left(\frac{\sigma_{22}}{\sigma_{21}}\right)\right), \sigma_{11}\right]$$

- Induce a (set-valued) mapping from reduced-form to structural parameters

Shock-sign NR

- Imposing the shock-sign restriction $\varepsilon_{1k} \geq 0$ provides identifying information as it restricts the space of δ
- If $\sigma_{21} y_{1k} - \sigma_{11} y_{2k} < 0$ and $y_{1k} > 0$,

$$\delta \in \left[\arctan\left(\frac{\sigma_{22} y_{1k}}{\sigma_{21} y_{1k} - \sigma_{11} y_{2k}} \right), \arctan\left(\frac{\sigma_{22}}{\sigma_{21}} \right) \right]$$

- Data-dependent set-valued mapping from reduced-form to structural parameter

Implications for frequentists and Bayesians

- Frequentists: NR do not fit the standard identification analysis.
- Bayesians: apparently little practical difference once conditioning on the sample
  - But the approach in the literature (AR18) suffers from prior sensitivity issues

Existing standard Bayesian inference under NR

Flat likelihood

- Write the shock-sign restrictions as $\{N(\delta, \phi, y_t) \geq 0\}_{t=1}^{K}$
- Choose a prior for ϕ and a prior for δ
- Update priors using the unconditional likelihood:

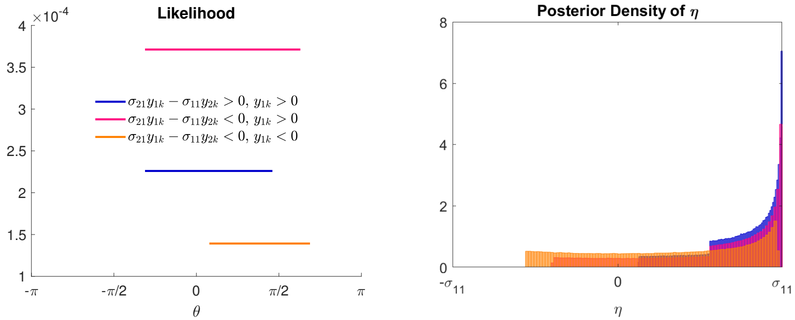
$$p(y^T, \{N(\delta, \phi, y_t) \geq 0\}_{t=1}^{K} | \delta, \phi)$$
$$= \underbrace{\prod_{t=1}^{T} (2\pi)^{-1} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} y_t' \Sigma^{-1} y_t\right)}_{f(y^T|\phi)} \cdot \prod_{t=1}^{K} 1\{N(\delta, \phi, y_t) \geq 0\}$$

- NR truncates likelihood so fixing ϕ it is flat for δ satisfying NR and is zero otherwise (and points of truncation depend on $\{y_t\}_{t=1}^{K}$)

Pitfalls of single-prior Bayesian inference

- Flat likelihood $\rightarrow$ posterior inference sensitive to the choice of prior for δ
- Similar to the case of traditional set-identifying restrictions that we saw in lecture 1
  - If you choose a single prior (even one that is uniform for the impulse responses), a component of this prior is never updated

Flat likelihood and posterior for impulse response



Solution: robust-Bayesian inference

- We propose robust-Bayesian inference for $\alpha = \alpha(\delta, \phi)$
  - Single prior for ϕ (revisable), multiple priors for δ (unrevisable)

Conditional Identified Set and Robust Credible region:

1. Estimate reduced-form VAR to obtain the posterior for ϕ
2. Draw ϕ from the posterior, and get the conditional identified set for α (here available analytically):

$$\begin{aligned} CIS_\alpha &\equiv \{\alpha(\delta, \phi) : N(\delta, \phi, y_t) \geq 0, t = 1, \ldots, K\} \\ &= [\ell(\phi), u(\phi)] \end{aligned}$$

3. Construct robust credible region using quantiles of posteriors of $\ell(\phi)$ and $u(\phi)$

Identification under NR

- In Giacomini, Kitagawa and Read (21) we show that NR create a non-standard identification problem
  - NR are in principle point-identifying if we assume that they hold in repeated samples...
  - ...but this does NOT give a way to construct point estimators, only an estimator of the conditional identified set CIS_α and associated credible region

Frequentist properties of robust Bayesian approach

- We also show that
  - If K is fixed as $T \to \infty$ (+ regularity assumptions), the robust credible region attains correct frequentist coverage for the true impulse response (coverage $\geq$ nominal)
  - If $K = T$ - so we know the sign of the shock in all periods - the conditional identified set CIS_α shrinks and converges to the true impulse response as $T \to \infty$

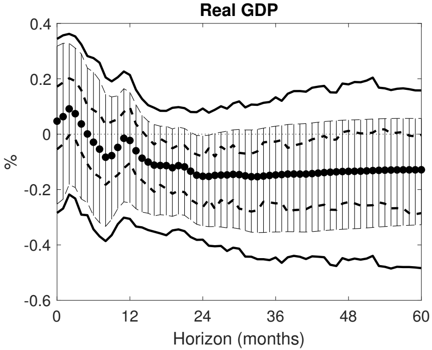
Robust-Bayesian inference under NR, in practice

- Report estimators of sets CIS_α and of the associated robust credible region, typically obtained numerically
- Note: for unit-effect impulse-responses, these sets can be the whole real line
- For impulse-responses to a one standard deviation shock, these sets are bounded

Robust-Bayesian inference under NR, in practice

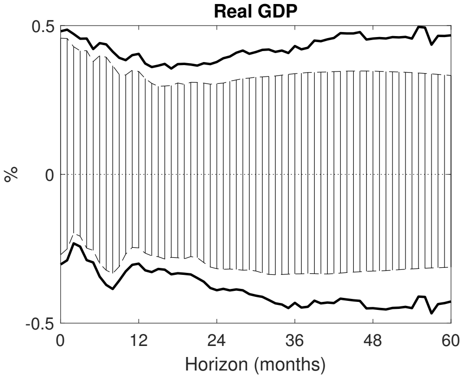
- SVAR in AR18: 6 variables, 12 lags, monthly 1965-2007, α: output response to unit s.d. monetary policy shock
- NR: Oct. 1979 mon. policy shock was positive and largest one

Figure: Impulse Responses to a Monetary Policy Shock



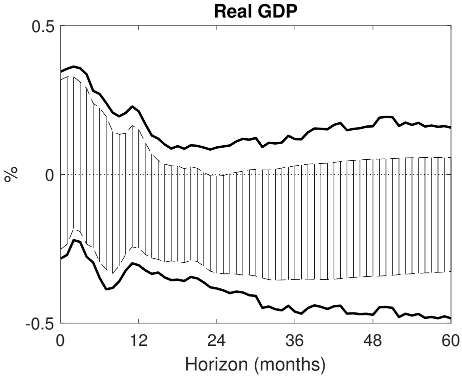
Shock-sign NR are not very informative

Figure: Sign + shock-sign NR



Historical decomposition NR are more informative

Figure: Sign + shock-sign NR + historical decomposition NR



The NP approach

NP approach

- Main idea: In the static example α can be identified in the regression

$$y_{2t} = \alpha y_{1t} + \varepsilon_{2t},$$

if can instrument y_{1t} by z_t such that $E(z_t \varepsilon_{1t}) \neq 0$ and $E(z_t \varepsilon_{2t}) = 0$

- Proxy/instrument

$$z_t = \begin{cases} \text{sign}(\varepsilon_{1t}) & \text{for } t = 1, \ldots, K \\ 0 & \text{otherwise} \end{cases}$$

satisfies both requirements

NP identification

Lemma: NP identification

Define

$$\gamma_1 = T^{-1} \sum_{t=1}^{T} E[z_t y_{1t}] = T^{-1} \sum_{t=1}^{K} E[\text{sign}(\varepsilon_{1t}) y_{1t}]$$

$$\gamma_2 = T^{-1} \sum_{t=1}^{T} E[z_t y_{2t}] = T^{-1} \sum_{t=1}^{K} E[\text{sign}(\varepsilon_{1t}) y_{2t}].$$

Then, the impulse response α can be identified by the Wald estimand:

$$\alpha = \frac{\gamma_2}{\gamma_1}$$

NP is a weak IV

- Under normality, the expected value of the covariance between NP and shock is proportional to $\frac{K}{T}$
  - Weak instrument if K small relative to T
  - If K fixed, covariance goes to zero at rate T (faster than $\sqrt{T}$ rate considered in weak-instrument literature)
- 2SLS estimator $\hat{\alpha} = \frac{\sum_{t=1}^{K} \text{sign}(\varepsilon_{1t}) y_{2t}}{\sum_{t=1}^{K} \text{sign}(\varepsilon_{1t}) y_{1t}}$ is biased, not consistent, the sampling distribution is far from normal
- Standard solution: weak-instrument robust inference, e.g., Anderson-Rubin (AR) confidence intervals

AR confidence intervals

- AR confidence intervals are constructed by inverting tests based on the Wald statistic:

$$W_T(\alpha) = \frac{T(\hat{\gamma}_2 - \alpha\hat{\gamma}_1)^2}{\widehat{Var}(\sqrt{T}(\hat{\gamma}_2 - \alpha\hat{\gamma}_1))},$$

where $\hat{\gamma}_1 = T^{-1}\sum_{t=1}^{T} z_t y_{1t}$ and $\hat{\gamma}_2 = T^{-1}\sum_{t=1}^{T} z_t y_{2t}$

- Under regularity conditions, $W_T(\alpha) \to_d \chi^2(1)$ under the null even if instrument is weak
  - For this to be true, you need the estimator of the covariance between the proxy and the data to be asymptotically normal (scaled by $\sqrt{T}$)
  - For NP this distribution can be degenerate (roughly speaking, it's because it depends on the ratio K/T)

Validity of AR is not guaranteed

Theorem

(i) With fixed K, $W_T(\alpha)$ does not converge to $\chi^2(1)$ asymptotically as $T \to \infty$

(ii) If $K \to \infty$ as $T \to \infty$, $W_T(\alpha) \to_d \chi^2(1)$ as $T \to \infty$

Monte Carlo

- $T = 1000$, $K = 1, ..., 1000$
- $A^{-1} = \begin{pmatrix} 0.5 & -0.5 \\ 0.2 & 1.8 \end{pmatrix}$, $\Sigma = \begin{pmatrix} 0.5 & -0.8 \\ -0.8 & 3.28 \end{pmatrix}$
- Confidence level 95% and 68%
- We report coverage of CI, proportion of unbounded CIs, median width CI

Table: Weak-proxy robust inference – conf. level = 0.95

K	Coverage prob.	Prop. unbounded	Median width
1	1.000	1.000	∞
2	1.000	1.000	∞
3	1.000	1.000	∞
4	0.996	0.985	∞
5	0.978	0.887	∞
10	0.958	0.549	∞
20	0.952	0.180	8.222
30	0.951	0.047	5.067
40	0.954	0.011	3.975
50	0.951	0.001	3.371
100	0.947	0.000	2.133
500	0.950	0.000	0.894
1000	0.950	0.000	0.625

Table: Weak-proxy robust inference – conf. level = 0.68

K	Coverage prob.	Prop. unbounded	Median width
1	0.000	0.000	0.000
2	0.493	0.376	7.824
3	0.581	0.360	8.549
4	0.603	0.311	7.008
5	0.627	0.276	6.075
10	0.654	0.109	3.634
20	0.654	0.020	2.386
30	0.667	0.003	1.895
40	0.674	0.000	1.640
50	0.674	0.000	1.457
100	0.682	0.000	1.008
500	0.687	0.000	0.448
1000	0.683	0.000	0.316

Table: Robust Bayesian inference – conf level = 0.95

K	Coverage prob.	Prop. unbounded	Median width
1	1.000	0.766	∞
2	1.000	0.596	∞
3	1.000	0.461	28.343
4	1.000	0.345	8.725
5	0.999	0.271	6.083
10	0.998	0.075	2.742
20	0.996	0.005	1.571
30	0.994	0.000	1.226
40	0.993	0.000	1.034
50	0.993	0.000	0.932
100	0.986	0.000	0.717
500	0.966	0.000	0.545
1000	0.955	0.000	0.522

Table: Robust Bayesian inference – conf level = 0.68

K	Coverage prob.	Prop. unbounded	Median width
1	1.000	0.757	∞
2	0.998	0.581	∞
3	0.997	0.442	19.893
4	0.995	0.330	7.615
5	0.994	0.256	5.422
10	0.985	0.066	2.436
20	0.970	0.004	1.302
30	0.955	0.000	0.969
40	0.947	0.000	0.778
50	0.932	0.000	0.678
100	0.882	0.000	0.466
500	0.757	0.000	0.296
1000	0.718	0.000	0.275

Conclusion

- So-called ‘narrative’ measures of shocks may not be credible (not real shocks)
- Account for uncertainty in measurement by focusing on only a few time periods where we have credible information
- Considered two ways of incorporating minimal narrative information for identification and inference for impulse responses
  - (Bayesian + set) Impose the restriction that we know the sign of some shocks + robust Bayesian analysis gives valid credible regions. Finite-sample coverage is conservative for few restrictions
  - (Frequentist + point) Use the same information to construct (weak) proxy and consider weak-proxy robust confidence intervals. Only valid when you have a “large enough” number of restrictions

References

- Antolin-Diaz, J. and J. F. Rubio-Ramirez (2018): Narrative Sign Restrictions for SVARs, American Economic Review
- Giacomini, R., T. Kitagawa, and M. Read (2021): Identification and Infer- ence Under Narrative Restrictions, arXiv: 2102.06456
- Giacomini, R., T. Kitagawa, and M. Read (2022): Narrative restrictions and proxies, Journal of Business and Economic Statistics
- Mertens, K. and M. O. Ravn (2013): The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States, American Economic Review
- Romer, C. and D. Romer (2004): A New Measure of Monetary Shocks: Derivation and Implications, American Economic Review
- Stock, J. and M. Watson (2018): Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments, The Economic Journal