SVAR Identification From Higher Moments: Has the Simultaneous Causality Problem Been Solved?

José Luis Montiel Olea — Columbia

Mikkel Plagborg-Møller — Princeton

Eric Qian — Princeton

Slides: `https://scholar.princeton.edu/mikkelpm`

January 7, 2022

SVAR identification from higher moments

- Two recent strands of the SVAR literature exploit higher moments of the data to achieve point ID without imposing economic restrictions.

  1. ID from non-Gaussianity. Gouriéroux, Monfort & Renne (2017); Lanne, Meitz & Saikkonen (2017)

  2. ID from heteroskedasticity. Sentana & Fiorentini (2001); Rigobon (2003); Lewis (2021a)

- Has the old simultaneous causality problem been solved? Why does the applied micro literature continue to look for IVs and quasi-experiments?

SVAR identification from higher moments

- Two recent strands of the SVAR literature exploit higher moments of the data to achieve point ID without imposing economic restrictions.
  1. ID from non-Gaussianity. Gouriéroux, Monfort & Renne (2017); Lanne, Meitz & Saikkonen (2017)
  2. ID from heteroskedasticity. Sentana & Fiorentini (2001); Rigobon (2003); Lewis (2021a)
- Has the old simultaneous causality problem been solved? Why does the applied micro literature continue to look for IVs and quasi-experiments?
- This paper: Critical review of the literature on higher-moment ID.
  - Higher-moment methods rely on stronger assumptions about the shock process.
  - These assumptions can and should be tested.
  - Weak ID issues should be given high priority.

Outline

Traditional SVAR identification from second moments

- Bivariate SVAR model with data $\mathbf{y}_t = (y_{1,t}, y_{2,t})$ and latent shocks $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \varepsilon_{2,t})$:

$$\mathbf{y}_t = \mathbf{c} + \sum_{\ell=1}^{p} \mathbf{A}_\ell \mathbf{y}_{t-\ell} + \mathbf{H}\boldsymbol{\varepsilon}_t, \quad \mathbf{H} = \begin{pmatrix} 1 & ? \\ ? & 1 \end{pmatrix}.$$

- Assumption: Shocks are orthogonal white noise.

$$\text{Cov}(\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_{t-\ell}) = \mathbf{0} \ \text{ for } \ell \geq 1, \quad \text{Cov}(\varepsilon_{1,t}, \varepsilon_{2,t}) = 0.$$

- White noise assumption only has implications for second moments:

$$\text{Var}(\boldsymbol{\eta}_t) = \mathbf{H} \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \mathbf{H}', \quad \boldsymbol{\eta}_t \equiv \mathbf{y}_t - \text{proj}(\mathbf{y}_t \mid \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots) = \mathbf{H}\boldsymbol{\varepsilon}_t.$$

- Simultaneous causality problem: 4 parameters, 3 non-redundant equations.

Economic content and robustness

- Traditional SVAR literature overcomes the simultaneous causality problem by imposing additional exclusion or sign restrictions.
  - Examples: External IV, timing restrictions, long-run neutrality, knowledge about policy rule.
  - Defended based on economic theory or institutional background.
  - Similar to applied micro. Nakamura & Steinsson (2018b); Stock & Watson (2018)
- Second-moment SVAR methods are usually robust to statistical properties of the data.
  - OLS estimator can be viewed as quasi-MLE of the SVAR model under the working assumptions that ε_t is i.i.d., homoskedastic, and Gaussian.
  - But none of these working as'ns are required for consistency of OLS. Gonçalves & Kilian (2004)

Outline

Mutual shock independence

- Borrowing from Independent Components Analysis (ICA) in statistics, recent SVAR papers have proposed strengthening the white noise assumption. Comon (1994); Hyvärinen, Karhunen & Oja (2001); Gouriéroux, Monfort & Renne (2017); Lanne, Meitz & Saikkonen (2017)
- Additional assumption: $\varepsilon_{1,t} \perp\!\!\!\perp \varepsilon_{2,t}$.
- Question: If $(C_{11}\eta_{1,t} + C_{12}\eta_{2,t}) \perp\!\!\!\perp (C_{21}\eta_{1,t} + C_{22}\eta_{2,t})$, must these linear combinations equal the true shocks $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ (up to scale and ordering)?
  - Answer, part 1: If $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are Gaussian, then no. There exist many linear combinations that are uncorrelated and thus independent.
  - Answer, part 2: If either of the shocks is non-Gaussian, then yes (Darmois-Skitovich Th'm).

Identification from mutual independence and non-Gaussianity

- Hence, if shocks are mutually independent and non-Gaussian, the SVAR model is point-identified (up to labeling of the shocks)!
- Intuitively, though second moments are not enough for point-ID, shock independence has implications for higher moments of the data.
  - Higher moments are redundant if shocks are Gaussian, but there is no reason to believe they would be *exactly* Gaussian in reality.
- Several SVAR-ICA estimation procedures have been proposed, such as quasi-MLE or method of moments. Do not require knowledge of shock distribution.
  Fiorentini & Sentana (2020); Lanne & Luoto (2021); Sims (2021)
- This approach achieves point-ID, not by exploiting economic knowledge, but by strengthening the statistical assumptions on the shock process.

How strong is the mutual independence assumption?

- Mutual independence rules out shared volatility, e.g.,

$$\varepsilon_{j,t} = \tau_t \zeta_{j,t}, \quad j = 1, 2, \tag{$\dagger$}$$

where τ_t is a shared scalar volatility factor, and $\zeta_{1,t} \perp\!\!\!\perp \zeta_{2,t}$.

- Doesn't the impulse-propagation paradigm for macro demand that shocks should be independent? Perhaps, but it doesn't follow that these should enter linearly in the SVAR.
- Example: In SVAR model with shock process ($\dagger$), one could think of the basic shocks as $\zeta_{1,t}$, $\zeta_{2,t}$, and shocks to τ_t.
  - Standard 2nd-moment methods estimate impulse responses wrt. $\varepsilon_{j,t}$, which is still meaningful.
  - Since $\varepsilon_{1,t} \not\perp\!\!\!\perp \varepsilon_{2,t}$, SVAR-ICA may fail to estimate anything with a structural interpretation.

Take-aways for future work

- Our opinion: If we insist on imposing shock independence, it seems necessary to consider nonlinear SVAR models.
  - At least allow shared (and persistent) volatility dynamics, consistent with the data.
  - Nonlinear ICA methods exist, but have not been adapted to macro applications. Hyvärinen, Karhunen & Oja (2001)
- Reminder: Choice is not between assuming Gaussian or non-Gaussian shocks!
  - Traditional 2nd-moment approaches do not require shocks to be Gaussian or mutually independent.

Testing independence: Empirical example

- The mutual shock independence assumption is testable: Just check whether the estimated shocks $(\hat{\varepsilon}_{1,t}, \hat{\varepsilon}_{2,t})' = \hat{\mathbf{H}}^{-1}\hat{\boldsymbol{\eta}}_t$ are in fact mutually independent.
  Matteson & Tsay (2017); Amengual, Fiorentini & Sentana (2021); Davis & Ng (2021)
- Empirical example:
  - SVAR in inflation, output gap, 3-month Treasury rate. GMR (2017)
  - Bootstrap test of $\text{Corr}(\varepsilon_{i,t}^2, \varepsilon_{j,t}^2) = 0$ for $i \neq j$:

Sample	Test statistic	5% CV	10% CV
1959–2019	0.055	0.121	0.091
1973–2019	**0.169**	0.129	0.108
1985–2019	**0.170**	0.130	0.116

$p = 6$ lags. OLS estimator of reduced-form VAR coef's. GMR (2017) quasi-MLE $\hat{\mathbf{H}}$.

  - Suggests that independence as'n should not be viewed as unobjectionable.

Outline

Conditional orthogonality

- Instead of assuming mutual shock independence, another strand of the literature has exploited heteroskedasticity in the data through the following assumption.
  Sentana & Fiorentini (2001); Rigobon (2003); Lewis (2021a); Sims (2021)

- Assumption: Shocks are conditionally unpredictable and orthogonal, given info set $\mathcal{I}_{t-1}$.

$$E(\boldsymbol{\varepsilon}_t \mid \mathcal{I}_{t-1}) = 0, \quad \text{Cov}(\varepsilon_{1,t}, \varepsilon_{2,t} \mid \mathcal{I}_{t-1}) = 0.$$

- Define $\sigma^2_{j,t-1} \equiv \text{Var}(\varepsilon_{j,t} \mid \mathcal{I}_{t-1})$ and $\boldsymbol{\Sigma}_{t-1} \equiv \text{Var}(\boldsymbol{\eta}_t \mid \mathcal{I}_{t-1})$. Then

$$\boldsymbol{\Sigma}_{t-1} = \mathbf{H} \begin{pmatrix} \sigma^2_{1,t-1} & 0 \\ 0 & \sigma^2_{2,t-1} \end{pmatrix} \mathbf{H}' \quad \Longrightarrow \quad \boldsymbol{\Sigma}_t \boldsymbol{\Sigma}_{t-1}^{-1} = \mathbf{H} \begin{pmatrix} \frac{\sigma^2_{1,t}}{\sigma^2_{1,t-1}} & 0 \\ 0 & \frac{\sigma^2_{2,t}}{\sigma^2_{2,t-1}} \end{pmatrix} \mathbf{H}^{-1}.$$

- Columns of $\mathbf{H}$ = eigenvectors of $\boldsymbol{\Sigma}_t \boldsymbol{\Sigma}_{t-1}^{-1}$. Unique if eigval's $\frac{\sigma^2_{j,t}}{\sigma^2_{j,t-1}}$ are distinct for $j = 1, 2$.

ID through heteroskedasticity

- ID through heterosk. exploits *conditional* second moments across distinct volatility regimes. We needed to strengthen the white noise as'n to conditional orthogonality.
  - ID argument applies both with Markov-switching vol jumps and continuous vol changes.
- Reminder: Choice is <u>not</u> between assuming homoskedastic or heteroskedastic shocks!
  - Traditional 2nd-moment approaches do not require shocks to be homoskedastic.
- ID argument suggests test of conditional orthogonality as'n: Check that eigenvectors of $\boldsymbol{\Sigma}_t \boldsymbol{\Sigma}_{t-1}^{-1}$ are constant over time (only eigval's may vary). Rigobon (2003)
- Incongruity in literature: Many procedures for ID through heterosk. impose a vol process such that $\varepsilon_{1,t} \not\perp\!\!\!\perp \varepsilon_{2,t}$, while many SVAR-ICA procedures impose homosk.

Outline

Sensitivity of higher-moment identification

- Higher-moment ID necessarily fails when shocks are Gaussian $\implies$ Potential weak ID.
- Relevant question is whether the shocks are *sufficiently* non-Gaussian relative to the (considerable) finite-sample estimation uncertainty in the higher moments.
- Higher-moment inference procedures likely to be more sensitive to minor perturbations of the data than second-moment procedures are. Lanne & Luoto (2021)
- Our opinion: Applied researchers should tackle potential weak ID head on.
  - Run simulation study calibrated to the application at hand.
  - Use weak-ID-robust procedures.
    Nakamura & Steinsson (2018a); Drautzburg & Wright (2021); Lee & Mesters (2021); Lewis (2021b)

Outline

Summary: Our views

1. ID from higher moments does not simply exploit more info in the data than traditional SVAR methods. It requires stronger assumptions on the shock process.
2. Applied work should routinely test the additional shock assumptions. Should also give high priority to issues of sensitivity and weak identification.
3. Second-moment ID methods remain relevant due to their robustness to statistical properties of the data. It is a virtue rather than a limitation that researchers must defend assumptions on economic grounds, instead of appealing to statistical assumptions.

Summary: Our views

1. ID from higher moments does not simply exploit more info in the data than traditional SVAR methods. It requires stronger assumptions on the shock process.
2. Applied work should routinely test the additional shock assumptions. Should also give high priority to issues of sensitivity and weak identification.
3. Second-moment ID methods remain relevant due to their robustness to statistical properties of the data. It is a virtue rather than a limitation that researchers must defend assumptions on economic grounds, instead of appealing to statistical assumptions.

Thank you!

Appendix

Bootstrap test of mutual shock independence

- Simple bootstrap test for SVAR with n variables:

1. Compare root mean squared correlation of squared estimated shocks

$$\hat{S} \equiv \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i} \widehat{\text{Corr}}(\hat{\varepsilon}_{i,t}^2, \hat{\varepsilon}_{j,t}^2)^2}$$

   to critical value from bootstrapped null distribution.

2. Generate bootstrap data from estimated SVAR model, where shocks $\{\varepsilon_{j,t}^*\}_t$ are drawn with replacement from estimated shocks $\{\hat{\varepsilon}_{j,t}\}_t$, independently across $j = 1, \dots, n$.