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

By JOSÉ LUIS MONTIEL OLEA, MIKKEL PLAGBORG-MØLLER, AND ERIC QIAN*

Two recent strands of the literature on Structural Vector Autoregressions (SVARs) use higher moments for identification. One of them exploits independence and non-Gaussianity of the shocks; the other, stochastic volatility (heteroskedasticity). These approaches achieve point identification without imposing exclusion or sign restrictions. We review this work critically, and contrast its goals with the separate research program that has pushed for macroeconometrics to rely more heavily on credible economic restrictions and institutional knowledge, as is the standard in microeconometric policy evaluation. Identification based on higher moments imposes substantively stronger assumptions on the shock process than standard second-order SVAR identification methods do. We recommend that these assumptions be tested in applied work. Even when the assumptions are not rejected, inference based on higher moments necessarily demands more from a finite sample than standard approaches do. Thus, in our view, weak identification issues should be given high priority by applied users.

I. Identification From Second Moments

We first review the well-known issues in identification of SVARs from second moments of the data. Suppose the data $\mathbf{y}_t = (y_{1,t}, y_{2,t})'$ follows the SVAR process

$$\mathbf{y}_t = \mathbf{c} + \sum_{\ell=1}^{p} \mathbf{A}_\ell \mathbf{y}_{t-\ell} + \mathbf{H}\boldsymbol{\varepsilon}_t, \tag{1}$$

where the vector of unobserved shocks $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$ is assumed to be orthogonal white noise, that is, serially and mutually uncorrelated: $\mathrm{Cov}(\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_{t-\ell}) = 0$ for $\ell \geq 1$ and $\mathrm{Cov}(\varepsilon_{1,t}, \varepsilon_{2,t}) = 0$. The diagonal entries of $\mathbf{H}$ are normalized to 1.

Due to the familiar simultaneous causality problem, the basic SVAR model is not identified. The autoregressive matrices $\mathbf{A}_\ell$ in (1) can be identified through a projection of $\mathbf{y}_t$ on its lags (assuming stationarity); let $\boldsymbol{\eta}_t = (\eta_{1,t}, \eta_{2,t})' = \mathbf{H}\boldsymbol{\varepsilon}_t$ be the reduced-form residuals from this projection. If we knew $\mathbf{H}$, we could identify the dynamic effects of the shocks $\boldsymbol{\varepsilon}_t = \mathbf{H}^{-1}\boldsymbol{\eta}_t$. Unfortunately, it is not possible to identify $\mathbf{H}$ without further assumptions. Since the white noise assumption on the shocks only has implications for first and second moments of the data, the only exploitable information about $\mathbf{H}$ comes from the equation

$$\mathrm{Var}(\boldsymbol{\eta}_t) = \mathbf{H} \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \mathbf{H}', \tag{2}$$

where $\sigma_j^2 = \mathrm{Var}(\varepsilon_{j,t})$. The nonlinear equation (2) has multiple solutions. Intuitively, the left-hand side of (2) only has three non-redundant (consistently estimable) reduced-form parameters due to symmetry, while the right-hand side has four unknowns (the off-diagonal elements of $\mathbf{H}$ and variances of the shocks).

To overcome the simultaneous causality issue, a vast literature on SVAR identification has proposed to exploit various additional economic restrictions, such as exclusion or sign restrictions (Kilian and Lütkepohl, 2017). These restrictions are usually motivated using economic theory and/or the institutional background for the application at hand. For example, in recent years, it has become popular to achieve identification by constructing instrumental variables (also known as proxies) that are assumed to correlate only with the shock of interest but not other shocks. The practice of explicitly defending these in-

* Montiel Olea: Department of Economics, Columbia University, 1022 International Affairs Building New York, NY 10027, montiel.olea@gmail.com. Plagborg-Møller: Department of Economics, Princeton University, Julis Romo Rabinowitz Building, Princeton, NJ 08544, mikkelpm@princeton.edu. Qian: Department of Economics, Princeton University, Julis Romo Rabinowitz Building, Princeton, NJ 08544, eric-qian@princeton.edu. Plagborg-Møller acknowledges that this material is based upon work supported by the NSF under Grant #1851665.

strument exclusion restrictions brings identification in macroeconomics closer to the modus operandi in applied microeconomics (Nakamura and Steinsson, 2018*b*; Stock and Watson, 2018).

SVAR estimators based on second moments are usually highly robust to statistical properties of the data. For example, the usual Ordinary Least Squares (OLS) estimator can be viewed as the quasi-maximum likelihood estimator (MLE) of the model (1) under the working assumptions that the shocks ε_t are i.i.d., Gaussian, and homoskedastic. But importantly, the OLS estimator is consistent regardless of whether any of these working assumptions hold (e.g., Gonçalves and Kilian, 2004).[1]

II. Identification From Mutual Independence and Non-Gaussianity

Beginning with Lanne and Lütkepohl (2010), several recent papers have achieved SVAR identification by importing techniques from the statistics literature on Independent Components Analysis (ICA, see the review by Hyvärinen, Karhunen and Oja, 2001). The key assumptions in this literature are that the shocks $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are mutually independent and non-Gaussian. Given these assumptions, it is natural to ask:

Question: If we find two linear combinations $C_{11}\eta_{1,t} + C_{12}\eta_{2,t}$ and $C_{21}\eta_{1,t} + C_{22}\eta_{2,t}$ of the reduced-form residuals that are mutually independent, can we conclude that these linear combinations must equal the true shocks $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ (up to scale and ordering)?[2]

Answer, part 1: If the shocks were Gaussian, it is well-known that the answer is *no*. There exist multiple linear combinations that are uncorrelated and, due to Gaussianity, therefore also independent.

Answer, part 2: If at least one of the shocks $\varepsilon_{1,t}$ or $\varepsilon_{2,t}$ is *not* Gaussian, then the answer is *yes*. This follows from a mathematical result known as the Darmois-Skitovich Theorem (Comon, 1994): Two nontrivial linear combinations of independent random variables cannot themselves be independent, unless the variables are all Gaussian.

Intuitively, though second moments of the data do not suffice for identification (cf. Section I), higher moments provide further restrictions to pin down the structural parameters.[3] Third or higher moments are redundant if the shocks are exactly Gaussian, but since there is no reason to believe that this is the case in practice, the identification result appears very attractive.

Following the ICA literature, several SVAR estimators based on non-Gaussian identification are available, such as estimators that exploit restrictions on third and fourth moments implied by independence (Lanne and Luoto, 2021) or quasi-MLE (Gouriéroux, Monfort and Renne, 2017; Lanne, Meitz and Saikkonen, 2017; Fiorentini and Sentana, 2020; Sims, 2021). Many of these estimators are semiparametrically consistent, i.e., without requiring the functional form of the shock distribution to be known.

As is clear, the SVAR-ICA approach overcomes the simultaneous causality problem, not by exploiting economic theory or institutional details, but by imposing the strong statistical assumption that the shocks are mutually independent.[4] This substantially strengthens the orthogonal white noise assumption in standard SVAR analysis. For example, it rules out the white noise shock process

$$\varepsilon_{j,t} = \tau_t \zeta_{j,t}, \quad j = 1, 2, \tag{3}$$

where τ_t is a shared scalar volatility process that is independent of the i.i.d., independent, and mean-zero processes $\zeta_{1,t}$ and $\zeta_{2,t}$. This process is consistent with historical episodes such as the Great Moderation.

One could argue that the independence assumption is innocuous because it can be justified from the definition of structural shocks in the impulse-propagation paradigm for macroeconomic dynamics. While this is a reasonable argument, the SVAR-ICA literature further assumes that these independent shocks enter ad-

[1] Another attractive feature of the Gaussian quasi-MLE is that the asymptotic limit of the resulting impulse response estimates can be interpreted as arising from Local Projections, which have a transparent interpretation regardless of whether the true data generating process is in fact linear as assumed in equation (1) (Plagborg-Møller and Wolf, 2021).

[2] I.e., are there other choices than $\mathbf{C} = \mathbf{H}^{-1}$ (up to column permutations and scale changes) that yield independence?

[3] See Hyvärinen, Karhunen and Oja (2001, Ch. 7) for further graphical intuition.

[4] A modicum of economic reasoning is needed to label the estimated shocks, for example by inspecting impulse responses.

ditively in the SVAR model (1). To spell this out in an example, consider the SVAR model with stochastic volatility in equations (1) and (3). One could argue that the basic shocks in this model are $\zeta_{1,t}$ and $\zeta_{2,t}$ (as well as shocks to τ_t), which are indeed mutually independent. Traditional linear SVAR approaches based on second moments will not estimate impulse responses with respect to $\zeta_{1,t}$ and $\zeta_{2,t}$, but if their identifying restrictions are satisfied, they will consistently estimate impulse responses with respect to the orthogonal white noise disturbances $\varepsilon_{1,t} = \tau_t\zeta_{1,t}$ and $\varepsilon_{2,t} = \tau_t\zeta_{2,t}$, which are arguably still interesting structural objects. In contrast, since $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are not independent in this example, ICA approaches may fail to estimate any object with a structural interpretation.

If one insists on estimating impulse responses with respect to independent shocks, it seems necessary to consider classes of nonlinear models that—at a minimum—are rich enough to allow for shared (and persistent) volatility factors in the residuals, such as (3). Nonlinear variants of ICA exist (Hyvärinen, Karhunen and Oja, 2001, Ch. 17), but these have not yet been adapted to macroeconomic applications.

We stress that the choice between identification based on second moments vs. higher-order moments is *not* a choice between Gaussian shocks vs. non-Gaussian shocks. Second-moment SVAR estimators require neither Gaussianity, nor mutual shock independence, as discussed in Section I.

III. Testing Mutual Independence

As pointed out by Matteson and Tsay (2017) and Amengual, Fiorentini and Sentana (2021), the assumption of mutual shock independence is testable in the data. Given any ICA estimator $\hat{\mathbf{H}}$ that is consistent under independence, one simply needs to check whether the estimated shocks $(\hat{\varepsilon}_{1,t}, \hat{\varepsilon}_{2,t})' = \hat{\mathbf{H}}^{-1}\hat{\boldsymbol{\eta}}_t$ are in fact independent of each other. In the Online Appendix we implement a particular version of this idea, using a bootstrap to account for estimation error. We reject independence in some specifications of a canonical monetary SVAR model applied to U.S. data, though not in all specifications. This suggests that shock independence should not be viewed as an unobjectionable assumption, but rather as a hypothesis to be tested in every application. It is an interesting area of research to further develop and analyze such tests.

IV. Identification From Heteroskedasticity

A different approach to higher-moment identification exploits heteroskedasticity (i.e., time-varying conditional volatility) in the data, following the seminal paper of Rigobon (2003). See Brunnermeier et al. (2021), Lewis (2021*a*), and Sims (2021) for recent results and additional references. Instead of assuming serial and mutual shock independence, these papers make the weaker assumption that the shocks are nonlinearly unpredictable and conditionally (not just unconditionally) orthogonal:

$$E(\varepsilon_t \mid \mathcal{I}_{t-1}) = 0, \ \mathrm{Cov}(\varepsilon_{1,t}, \varepsilon_{2,t} \mid \mathcal{I}_{t-1}) = 0,$$

for all t, where $\mathcal{I}_{t-1}$ is an information set available to the econometrician at time $t-1$. Denote the conditional variance of shock j by $\sigma^2_{j,t-1} = \mathrm{Var}(\varepsilon_{j,t} \mid \mathcal{I}_{t-1})$. Then, similar to equation (2),

$$\mathrm{Var}(\boldsymbol{\eta}_t \mid \mathcal{I}_{t-1}) = \mathbf{H}\begin{pmatrix}\sigma^2_{1,t-1} & 0 \\ 0 & \sigma^2_{2,t-1}\end{pmatrix}\mathbf{H}'.$$

Denote the left-hand side by Σ_{t-1}; note that this conditional reduced-form variance-covariance matrix is in principle estimable in the data. Then it follows that

$$\Sigma_t\Sigma_{t-1}^{-1} = \mathbf{H}\begin{pmatrix}\frac{\sigma^2_{1,t}}{\sigma^2_{1,t-1}} & 0 \\ 0 & \frac{\sigma^2_{2,t}}{\sigma^2_{2,t-1}}\end{pmatrix}\mathbf{H}^{-1}. \tag{4}$$

In words, the columns of the matrix $\mathbf{H}$ are the eigenvectors of the matrix $\Sigma_t\Sigma_{t-1}^{-1}$. These eigenvectors are uniquely determined (given the normalization $H_{11} = H_{22} = 1$), provided that the eigenvalues $\frac{\sigma^2_{1,t}}{\sigma^2_{1,t-1}}$ and $\frac{\sigma^2_{2,t}}{\sigma^2_{2,t-1}}$ are distinct, i.e., if the volatilities of the two shocks jump by different amounts from time $t-1$ to time t. This rules out a model with a single shared volatility factor, such as (3).

In summary, by imposing the stronger assumption that the shocks are conditionally orthogonal, we are able to identify the SVAR model by essentially exploiting conditional second moments in multiple volatility "regimes". This is still a "higher-moment" procedure in the sense of unconditional moments.

We stress, once again, that the choice is *not* between assuming heteroskedasticity or homoskedasticity. Traditional (unconditional) second-moment procedures often use a homoskedastic Gaussian quasi-MLE, but homoskedasticity is not required for either consistency or identification.

The above identification argument immediately suggests a way to test the identifying assumptions: Check whether the eigenvectors of $\Sigma_t \Sigma_{t-1}^{-1}$ are indeed constant at each point in time t (only the eigenvalues may vary), as predicted by equation (4).[5] This general idea applies to many different volatility models that have been considered in the literature, such as discrete Markov-switching volatility jumps and continuous volatility changes (Lewis, 2021*a*).

Popular SVAR procedures that exploit heteroskedasticity rely on assumptions that are completely at odds with popular SVAR-ICA procedures. Whereas the former estimate time-varying volatilities, the latter usually assume that shocks are i.i.d. over time (though this is not required for identification). There does not appear to be any reason why the two different kinds of estimation procedures could not be made mutually consistent.[6] We hope that future research will pursue this, while addressing the shortcomings we have mentioned previously.

V. Sensitivity of Higher-Moment Identification

Any identification approach that relies on third or higher moments is at risk of suffering from weak identification. Recall that these approaches must necessarily fail when the data is (unconditionally) Gaussian, since then higher moments provide no additional information. It is of no comfort that macroeconomic data is unlikely to be exactly Gaussian; the relevant question is whether the data generating process is sufficiently non-Gaussian relative to the estimation error in the higher moments. As is well known, the estimation error in higher moments is typically large for moderate sample sizes.

The potential for weak identification suggests that higher-moment procedures could be more sensitive to minor perturbations of the data than conventional second-moment procedures are. Indeed, Lanne and Luoto (2021, Sec. 3.4) demonstrate in a model with t-distributed shocks that the performance of SVAR-ICA estimators and tests deteriorates as the degrees of freedom increase.

We therefore think it is important for the literature on higher-moment identification to discuss more seriously potential weak identification issues. A theoretical treatment of this problem seems non-trivial. However, applied researchers could easily report empirically calibrated simulations to verify that the higher-moment procedures behave as expected, given the actual sample size and potential non-Gaussianity of the residuals. Ideally, in the long-run, any reported inference results would be complemented by weak identification robust procedures, such as those that have been recently developed both for identification by ICA (Lee and Mesters, 2021) and by heteroskedasticity (Nakamura and Steinsson, 2018*a*, Appendix C; Lewis, 2021*b*).

VI. Conclusion

This paper has argued the following:

- Identification from higher moments does not simply exploit more information in the data than traditional SVAR methods; it requires stronger assumptions on the shocks than second-moment methods do.
- Applied work should routinely test the additional shock assumptions. It should also give high priority to issues of sensitivity and weak identification, since these are likely more acute when exploiting higher moments in macroeconomic data.
- Second-moment identification remains relevant due to its robustness to statistical properties of the data. Moreover, we believe it is a feature rather than a bug that researchers are forced to defend their identifying restrictions based on substantive economic or institutional knowledge, rather than appealing to statistical assumptions about the shock process.

[5]This is related to the over-identification test in Rigobon (2003).

[6]In fact, identification through non-Gaussianity is in some sense a special case of identification through heteroskedasticity, as pointed out by Lanne and Lütkepohl (2010): Most non-Gaussian shock distributions can be well approximated by a discrete mixture of Gaussian distributions, and the components of this mixture can be viewed as separate volatility regimes.

REFERENCES

Amengual, Dante, Gabriele Fiorentini, and Enrique Sentana. 2021. "Moment tests of independent components." Manuscript, CEMFI.

Brunnermeier, Markus, Darius Palia, Karthik A. Sastry, and Christopher A. Sims. 2021. "Feedbacks: Financial Markets and Economic Activity." *American Economic Review*, 111(6): 1845–79.

Comon, Pierre. 1994. "Independent component analysis, a new concept?" *Signal Processing*, 36: 287–314.

Fiorentini, Gabriele, and Enrique Sentana. 2020. "Discrete mixtures of normals pseudo maximum likelihood estimators of structural vector autoregressions." CEPR Discussion Paper DP15411.

Gonçalves, Sílvia, and Lutz Kilian. 2004. "Bootstrapping autoregressions with conditional heteroskedasticity of unknown form." *Journal of Econometrics*, 123(1): 89–120.

Gouriéroux, Christian, Alain Monfort, and Jean-Paul Renne. 2017. "Statistical inference for independent component analysis: Application to structural VAR models." *Journal of Econometrics*, 196(1): 111–126.

Hyvärinen, Aapo, Juha Karhunen, and Erkki Oja. 2001. *Independent Component Analysis.* John Wiley & Sons, Inc.

Kilian, Lutz, and Helmut Lütkepohl. 2017. *Structural Vector Autoregressive Analysis.* Cambridge University Press.

Lanne, Markku, and Helmut Lütkepohl. 2010. "Structural Vector Autoregressions With Nonnormal Residuals." *Journal of Business & Economic Statistics*, 28(1): 159–168.

Lanne, Markku, and Jani Luoto. 2021. "GMM Estimation of Non-Gaussian Structural Vector Autoregression." *Journal of Business & Economic Statistics*, 39(1): 69–81.

Lanne, Markku, Mika Meitz, and Pentti Saikkonen. 2017. "Identification and estimation of non-Gaussian structural vector autoregressions." *Journal of Econometrics*, 196(2): 288–304.

Lee, Adam, and Geert Mesters. 2021. "Robust Non-Gaussian Inference for Linear Simultaneous Equations Models." Manuscript, Universitat Pompeu Fabra.

Lewis, Daniel J. 2021*a*. "Identifying Shocks via Time-Varying Volatility." *Review of Economic Studies*. Forthcoming.

Lewis, Daniel J. 2021*b*. "Robust Inference in Models Identified via Heteroskedasticity." *Review of Economics and Statistics*. Forthcoming.

Matteson, David S., and Ruey S. Tsay. 2017. "Independent Component Analysis via Distance Covariance." *Journal of the American Statistical Association*, 112(518): 623–637.

Nakamura, Emi, and Jón Steinsson. 2018*a*. "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect." *The Quarterly Journal of Economics*, 133(3): 1283–1330.

Nakamura, Emi, and Jón Steinsson. 2018*b*. "Identification in Macroeconomics." *Journal of Economic Perspectives*, 32(3): 59–86.

Plagborg-Møller, Mikkel, and Christian K. Wolf. 2021. "Local Projections and VARs Estimate the Same Impulse Responses." *Econometrica*, 89(2): 955–980.

Rigobon, Roberto. 2003. "Identification Through Heteroskedasticity." *The Review of Economics and Statistics*, 85(4): 777–792.

Sims, Christopher A. 2021. "SVAR Identification Through Heteroskedasticity With Misspecified Regimes." Manuscript, Princeton University.

Stock, James H., and Mark W. Watson. 2018. "Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments." *The Economic Journal*, 128(610): 917–948.