Moment tests of independent components

Dante Amengual
CEMFI

Gabriele Fiorentini
Università di Firenze

Enrique Sentana
CEMFI

9th Research Workshop
BANCO DE ESPAÑA – CEMFI

November 15, 2021

Motivation

- The literature on structural vector autoregressions (SVAR) is vast. Popular identification schemes include:
  - Short- and long-run homogenous restrictions (see, e.g., Sims (1980) and Blanchard and Quah (1989)),
  - Sign restrictions (see, e.g., Faust (1998) and Uhlig (2005)),
  - Time-varying heteroskedasticity (Sentana and Fiorentini (2001)) or
  - External instruments (see, e.g., Mertens and Ravn (2012), Stock and Watson (2018) or Dolado, Motyosvski and Pappa (2020)).
- Recently, identification through independent non-Gaussian shocks has become increasingly popular after Lanne, Meitz and Saikkonen (2017) and Gouriéroux, Monfort and Renne (2017).
  - The signal processing literature on Independent Component Analysis popularised by Comon (1994) shares the same identification scheme.

Motivation

- Specifically, if in a static model the $N \times 1$ observed random vector $\mathbf{y}$ is the result of an affine combination of N unobserved shocks $\boldsymbol{\varepsilon}$ whose mean and variance we can set to $\mathbf{0}$ and $\mathbf{I}_N$, namely

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{C}\boldsymbol{\varepsilon}^*, \tag{1}$$

then the matrix $\mathbf{C}$ of loadings can be identified from an *i.i.d.* sample of observations on $\mathbf{y}$ provided:

Assumption I (*Identification*):

I.1) The N shocks in (1) are cross-sectionally **independent**,

I.2) at least $N-1$ of them follow a **non-Gaussian** distribution, and

I.3) $\mathbf{C}$ is invertible.

- Failure of **Assumption I** results in an underidentified model e.g.:
  - *Gaussian* $\boldsymbol{\varepsilon}^*$: one can only identify $V(\mathbf{y}) = \mathbf{CC}'$ but not $\mathbf{C}$ despite the fact that the elements of $\boldsymbol{\varepsilon}^*$ are cross-sectionally independent; or
  - *Non-Gaussian spherical distribution for* $\boldsymbol{\varepsilon}^*$: lack of identifiability of $\mathbf{C}$ because linear combinations share not only their mean and covariance, but also their non-linear dependence.

What we do

- We propose –simple to implement and interpret– specification tests that check the **normality** of a single element of ε^* and the potential cross-sectional **dependence** among several of them.
  - We compare the integer (product) moments of the shocks in the sample with their population counterparts:
    1. In the **normality** tests we compare the marginal third and fourth moments of a single shock to 0 and 3, respectively.
    2. In the case of two or more shocks, we assess the statistical significance of their second, third and fourth cross-moments, which should be equal to the product of the corresponding marginal moments under **independence**.
- We focus on latent shocks rather than observed variables because **Assumption I** is written in terms of ε^* rather than $\mathbf{y}$.
  - If we knew the true values of $\boldsymbol{\mu}$ and $\mathbf{C}$ our tests would be straightforward.
  - In practice, though, both $\boldsymbol{\mu}$ and $\mathbf{C}$ are unknown, so we need to estimate them before computing our tests.

What we do

- Although many estimation procedures have been proposed, we consider discrete mixtures of normals-based pseudo maximum likelihood estimators [PMLEs, see Fiorentini and Sentana (2020)] because of:
  1. They are consistent for all model parameters (under standard regularity conditions) if **Assumption I** holds.
  2. We can easily compute in closed-form the influence functions on which they are based.
- We derive simple closed-form expressions for the asymptotic covariance matrices of the sample moments underlying our tests under the null adjusted for parameter uncertainty.
- We discuss some bootstrap procedures that seem to improve their reliability.
- Finally, we study the finite sample properties of our tests in several Monte Carlo exercises.

Set up

- Consider the following N-variate SVAR process of order p:

$$\mathbf{y}_t = \boldsymbol{\tau} + \sum_{j=1}^{p} \mathbf{A}_j \mathbf{y}_{t-j} + \mathbf{C}\boldsymbol{\varepsilon}_t^*, \quad \boldsymbol{\varepsilon}_t^* \sim i.i.d.\ (\mathbf{0}, \mathbf{I}_N),$$

  where $\boldsymbol{\varepsilon}_t^*$ denotes the (standardised) vector of "structural" shocks.
- Let $\boldsymbol{\varepsilon}_t = \mathbf{C}\boldsymbol{\varepsilon}_t^*$ denote the reduced form innovations, so that $\boldsymbol{\varepsilon}_t | I_{t-1} \sim i.i.d.\ (\mathbf{0}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = \mathbf{CC}'$.
- Let $\boldsymbol{\theta} = [\boldsymbol{\tau}', vec'(\mathbf{A}_1), \ldots, vec'(\mathbf{A}_p), vec'(\mathbf{C})]' = (\boldsymbol{\tau}', \mathbf{a}', \mathbf{c}')$ denote the structural parameters characterising the first two conditional moments of $\mathbf{y}_t$.
- In addition, we assume $\varepsilon_{it}^* | I_{t-1} \sim i.i.d.\ D(0, 1, \boldsymbol{\varrho}_i)$, where $\boldsymbol{\varrho}_i$ is a $q_i \times 1$ vector of variation-free shape parameters so that

$$l(\mathbf{y}_t; \boldsymbol{\phi}) = -\ln|\mathbf{C}| + \ln f[\varepsilon_{1t}^*(\boldsymbol{\theta}); \boldsymbol{\varrho}_1] + \ldots + \ln f[\varepsilon_{Nt}^*(\boldsymbol{\theta}); \boldsymbol{\varrho}_N],$$

  where $\boldsymbol{\phi} = (\boldsymbol{\theta}', \boldsymbol{\varrho}')'$, $f[\varepsilon_{it}^*(\boldsymbol{\theta}); \boldsymbol{\varrho}_i]$ is the log-likelihood for the i^{th} shock, $\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta}) = \mathbf{C}^{-1}\boldsymbol{\varepsilon}_t(\boldsymbol{\theta})$, and $\boldsymbol{\varepsilon}_t(\boldsymbol{\theta}) = \mathbf{y}_t - \boldsymbol{\tau} - \mathbf{A}_1\mathbf{y}_{t-1} - \ldots - \mathbf{A}_p\mathbf{y}_{t-p}$.

Set up

- Let $\boldsymbol{\phi}_\infty = (\boldsymbol{\theta}_0', \boldsymbol{\varrho}_\infty')'$ so that

$$\mathcal{A}(\boldsymbol{\phi}_\infty; \boldsymbol{\varphi}_0) = -E[\partial \mathbf{s}_{\boldsymbol{\phi} t}(\boldsymbol{\phi}_\infty)/\partial \boldsymbol{\phi}')|\boldsymbol{\varphi}_0]$$

  and

$$\mathcal{B}(\boldsymbol{\phi}_\infty; \boldsymbol{\varphi}_0) = V[\mathbf{s}_{\boldsymbol{\phi} t}(\boldsymbol{\phi}_\infty)|\boldsymbol{\varphi}_0]$$

  denote the (-) expected value of the log-likelihood Hessian and the variance of the score, respectively, where:
  - $\boldsymbol{\varrho}_\infty$ are the pseudo true values of the shape parameters of the distributions of the shocks assumed for estimation purposes, and
  - $\boldsymbol{\upsilon}$ contains the potentially infinite-dimensional shape parameters of the true distributions of the shocks, and
  - $\boldsymbol{\varphi} = (\boldsymbol{\theta}', \boldsymbol{\upsilon}')'$.
- Then, the asymptotic distribution of the PMLEs of $\boldsymbol{\phi}$, $\hat{\boldsymbol{\phi}}_T$, will be given by

$$\sqrt{T}(\hat{\boldsymbol{\phi}}_T - \boldsymbol{\phi}_\infty) \to N[\mathbf{0}, \mathcal{A}^{-1}(\boldsymbol{\phi}_\infty; \boldsymbol{\varphi}_0)\mathcal{B}(\boldsymbol{\phi}_\infty; \boldsymbol{\varphi}_0)\mathcal{A}^{-1}(\boldsymbol{\phi}_\infty; \boldsymbol{\varphi}_0)].$$

Specification tests based on integer product moments

- As is well known, stochastic **independence** between the elements of a random vector is equivalent to the **joint distribution being the product of the marginal ones**.
  - Therefore, a rather intuitive way of testing for independence without considering any specific parametric alternative can be based on individual moment conditions of the form

$$m_{\mathbf{h}}[\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})] = \prod_{i=1}^{N} \varepsilon_{it}^{*h_i}(\boldsymbol{\theta}) - \prod_{i=1}^{N} E[\varepsilon_{it}^{*h_i}(\boldsymbol{\theta}_0)],$$

  where $\mathbf{h} = \{h_1, ..., h_N\}$, with $h_i \in \mathbb{Z}_{0+}$, denotes the index vector characterising a specific product moment.
- Moreover, by setting all the elements of $\mathbf{h}$ but one to 0, we can also look at the **marginal moments of a single shock**.
  - We focus on $h_i = 3$ and 4 because most common departures from **normality** of the shocks will be reflected in coefficients of skewness or kurtosis different from 0 and 3, respectively.

Specification tests based on integer product moments

- As mentioned earlier, we explicitly consider the sampling variability resulting from using shocks computed with consistent parameter estimators.
- Our propositions report closed-form expressions for:
  - the covariance across influence functions,
$$cov\{m_{\mathbf{h}}[\varepsilon_t^*(\boldsymbol{\theta}_0)], m_{\mathbf{h}'}[\varepsilon_t^*(\boldsymbol{\theta}_0)]|\boldsymbol{\varphi}_0\},$$
  - their expected Jacobian,
$$E\left\{ \frac{\partial m_{\mathbf{h}}[\varepsilon_t^*(\boldsymbol{\theta}_0)]}{\partial \boldsymbol{\phi}} \middle| \boldsymbol{\varphi}_0 \right\},$$
  - and their covariance with the score,
$$cov\{m_{\mathbf{h}}[\varepsilon_t^*(\boldsymbol{\theta}_0)], \mathbf{s}_{\boldsymbol{\phi}t}(\boldsymbol{\phi}_\infty)|\boldsymbol{\varphi}_0\}.$$

Testing normality

- We can test the null hypothesis that **a single structural shock is Gaussian** by comparing its third and fourth sample moments with 0 and 3, respectively.
- Nevertheless, many authors (see, e.g., Bontemps and Meddahi (2005)) suggest looking at the sample averages of the third and fourth Hermite polynomials instead:
  - $\varepsilon_{it}^{*3} - 3\varepsilon_{it}^{*}$ and $\varepsilon_{it}^{*4} - 6\varepsilon_{it}^{*2} + 3$ rather than ε_{it}^{*3} and ε_{it}^{*4}.
- It turns out, though, that the usual implementation of the Jarque and Bera (1980) test, which simply looks at the sample averages of $\varepsilon_{it}^{*3}(\hat{\boldsymbol{\theta}}_T)$ and $\varepsilon_{it}^{*4}(\hat{\boldsymbol{\theta}}_T)$, yields numerically the same statistics as the tests based on the Hermite polynomials because

$$\frac{1}{T}\sum\nolimits_{t=1}^{T} \varepsilon_{it}^{*}(\hat{\boldsymbol{\theta}}_T) = 0 \quad \text{and} \quad \frac{1}{T}\sum\nolimits_{t=1}^{T} \varepsilon_{it}^{*2}(\hat{\boldsymbol{\theta}}_T) - 1 = 0$$

for all i when using discrete mixtures of normals-based PMLEs, regardless of the sample size.

Testing independence

- The first test for **independence** that we consider is based on the **second cross-moment** condition

$$E(\varepsilon_{it}^{*}\varepsilon_{i't}^{*}) = 0, \ \ i \neq i'$$

i.e. assessing if the sample correlation between the i^{th} and i'^{th} estimated shocks is significantly different from zero in the usual statistical sense.

- Nevertheless, we can also go beyond **linear independence**, and look at moments that characterise **co-skewness** across the structural shocks.
- These can be of two types:

$$E(\varepsilon_{it}^{*2}\varepsilon_{i't}^{*}) = 0, \ \ i \neq i',$$

and

$$E(\varepsilon_{it}^{*}\varepsilon_{i't}^{*}\varepsilon_{i''t}^{*}) = 0, \ \ i \neq i' \neq i'',$$

depending on whether they involve two or three different shocks.

Testing independence

- Finally, we can also look at the different **co-kurtosis** among the shocks, which may involve a pair of shocks, namely

$$E(\varepsilon_{it}^{*2}\varepsilon_{i't}^{*2}) - 1 = 0, \ i \neq i', \tag{2}$$

and

$$E(\varepsilon_{it}^{*3}\varepsilon_{i't}^{*}) = 0, \ i \neq i', \tag{3}$$

three shocks

$$E(\varepsilon_{it}^{*2}\varepsilon_{i't}^{*}\varepsilon_{i''t}^{*}) = 0, \ i \neq i' \neq i'',$$

and even four shocks

$$E(\varepsilon_{it}^{*}\varepsilon_{i't}^{*}\varepsilon_{i''t}^{*}\varepsilon_{i'''t}^{*}) = 0, i \neq i' \neq i'' \neq i'''.$$

  - Thus, we expand the set of moments relative to Hyvärinen (2013), who only suggested looking at the co-kurtosis terms in (2).
  - The above moment conditions also augment those considered by Lanne and Luoto (2021), who focus on (2) and (3), together with $E(\varepsilon_{it}^{*}) = 0$ and $E(\varepsilon_{it}^{*2}) = 1$ and $E(\varepsilon_{it}^{*}\varepsilon_{i't}^{*}) = 0, \ i \neq i'$.

Combining our tests

- We can use the expressions previously derived to prove that the two separate tests are asymptotically independent under the **joint null** hypothesis of mutually **independent** shocks and the **normality** of one of them, so effectively the joint test would simply be the sum of those two components.
- In addition, we can also prove that a test that jointly assessed the independence and normality of all the shocks would be asymptotically equivalent under the null to:
  - a multivariate Hermite-based test of multivariate normality (Amengual, Fiorentini and Sentana (2021)) applied to the reduced form residuals
  - once one eliminates the moment condition related to the covariance of the shocks, whose asymptotic variance when evaluated at the Gaussian–PMLEs would be zero under the null.

Monte Carlo evidence: Model & estimation

- We generate samples of size T from the following bivariate static process

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix} + \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} \varepsilon^*_{1t} \\ \varepsilon^*_{2t} \end{pmatrix} \tag{4}$$

  with $\tau_1 = 1$, $\tau_2 = -1$, $c_{11} = 1$, $c_{12} = .5$, $c_{21} = 0$ and $c_{22} = 2$.
- We also study dynamic as well as trivariate generalisations of (4).
- We consider both:
  - $T = 250$, realistic in most macro applications, and
  - $T = 1,000$, representative of financial applications.
- To estimate the parameters of model (4), we assume that ε^*_{1t} and ε^*_{2t} follow two serially and cross-sectionally independent standardised discrete mixture of two normals.

Monte Carlo evidence: DGPs design

- The DGPs for the standardised shocks that we consider under the null of **independence** are:
  1. A normal distribution and a discrete mixture of two normals with kurtosis coefficient 4 and skewness coefficients equal to $-.5$.
  2. Independent discrete mixtures of two normals with kurtosis coefficient 4 and skewness coefficients equal to .5 and $-.5$, respectively.
  3. A Student t with 10 degrees of freedom (and kurtosis coefficient equal to 4), and an asymmetric t with kurtosis and skewness coefficients equal to 4 and $-.5$, respectively.
- We also use them to evaluate the small sample size and power of the **normality** tests using the results from the simulation designs 1 (null), and 2 and 3 (alternative).

Monte Carlo evidence: DGPs design

- Additionally, we consider as alternative hypotheses to **independence**:
  4. Bivariate Student t with 6 degrees of freedom.
  5. Bivariate asymmetric t with skewness vector $\boldsymbol{\beta} = -5\ell_2$ and degrees of freedom parameter $\nu = 16$ (see Mencía and Sentana (2012) for details).
  6. Bivariate mixture of two zero-mean normal vectors with covariance matrices

$$\boldsymbol{\Omega}_1 = \begin{pmatrix} 1/[\lambda + \varkappa_1(1-\lambda)] & 0 \\ 0 & 1/[\lambda + \varkappa_2(1-\lambda)] \end{pmatrix},$$

and

$$\boldsymbol{\Omega}_2 = \begin{pmatrix} \varkappa_1/[\lambda + \varkappa_1(1-\lambda)] & 0 \\ 0 & \varkappa_2/[\lambda + \varkappa_2(1-\lambda)] \end{pmatrix},$$

where $\varkappa_1 = 0.1$, $\varkappa_2 = 0.2$ and $\lambda = 1/5$ (see Lanne and Lütkepohl (2010) for details).

- To gauge the small-sample properties of our tests, we generate $20,000$ ($5,000$) samples for each of the designs under the null (alternative).

Monte Carlo evidence: Bootstrap methods

- Regarding the **independence** tests, for each Monte Carlo sample, we can easily generate bootstrap samples of size T that impose the null by:
  - generating NT draws R_{is} from a discrete uniform distribution between 1 and T, which we then use to construct

$$\tilde{\mathbf{y}}_s = \hat{\boldsymbol{\tau}}_T + \hat{\mathbf{C}}_T \tilde{\boldsymbol{\varepsilon}}_s^*,$$

    where $\tilde{\varepsilon}_{is}^* = \hat{\varepsilon}_{iR_{is}}^*$ and $\hat{\boldsymbol{\varepsilon}}_t^* = \boldsymbol{\varepsilon}_t^*(\hat{\boldsymbol{\theta}}_T) = \hat{\mathbf{C}}_T^{-1}(\mathbf{y}_t - \hat{\boldsymbol{\tau}}_T)$ are the estimated residuals in any given sample.

- As for the **normality** tests, whose null is that a single shock i, ε_{it}^*, is Gaussian, we adopt a partially parametric resampling scheme in which:
  - $\tilde{\varepsilon}_{is}^*$'s are independently simulated from a $N(0,1)$ distribution,
  - while the remaining shocks $\tilde{\varepsilon}_{ks}^*$ $(k \neq i)$ are drawn nonparametrically as above.

Monte Carlo results: Normality tests

- Regarding **size**, the tests tend to be oversized at the usual nominal levels, especially for samples of length 250.
- However, the bootstrap version of our tests is pretty accurate for both the third and fourth moment tests.
- As for **power**, the null of normality is correctly rejected a large number of times when it does not hold, even in samples of length $T = 250$.
    - The only possible exception is the skewness component of the Jarque-Bera test when applied to the symmetric Student t.
    - Given that the population third moment is zero in this case, the only source of power is the fact that the sample variability of H_3 is larger for this shock than its theoretical value under Gaussianity.
- Finally, one could look at the univariate tests *for all* the shocks expecting at least $N - 1$ rejections.
    - We report contingency tables –characterising simultaneous rejections of the individual normality tests– which are quite informative in that regard.

Monte Carlo results: Independence tests

- We find some small to moderate finite sample **size** distortion when $T = 250$, although in several cases they are corrected by the bootstrap.
  - The only exception is when the shocks are independent but one of them is Gaussian, in which some small distortions remain even with this procedure.
- Finite sample **size** improves considerably for $T = 1,000$.
  - Indeed, the bootstrap versions of the tests seem unnecessary.
- Regarding **power**, we find that:
  - The power of our tests against Student t shocks is somewhat low.
  - Power is mostly coming from the co-skewness component in the case of asymmetric t shocks.
  - The co-kurtosis component is the most powerful moment test under the DGP á là Lanne and Lütkepohl.
  - Interestingly, the test based on the second moment also has non-negligible power.

Monte Carlo results: Assumption I matters

- When the shocks are **independent** but one of them is **Gaussian**, the sampling variability and the finite sample bias are noticeably larger than when both shocks are **non-Gaussian** and **independent** .
- The situation is completely different when the true shocks are cross-sectionally **dependent**.
- Failure of that condition results into significant biases, mostly in the off-diagonal terms of the impact multiplier matrix.
- In fact, the Monte Carlo variance of these estimators seems to increase with the sample size.
    - Remember that the elements of the $\mathbf{C}$ matrix are no longer point identified when the joint distribution of the true shocks is either a symmetric or asymmetric Student t.
    - This is confirmed by the fact that the bias of the estimators is lower for the DGP á là Lanne and Lütkepohl (2010), in which rotations of the shocks are not observationally equivalent.

Conclusions and directions for further research

- Identification through independent non-Gaussian shocks is a powerful result but not without concerns.
  - For that reason, it is desirable that empirical researchers checked the underlying assumptions to increase the credibility of their findings.
- The moment conditions that we consider for testing **independence** could form the basis of a GMM estimation procedure for the model parameters $\boldsymbol{\theta}$ along the lines of Lanne and Luoto (2021), although with a larger set of third and fourth cross-moments.
  - The overidentification restrictions tests obtained as a by product of this procedure could also be used as a specification test of the assumed independence-like restrictions.

Conclusions and directions for further research

- Our tests for **normality** only tackle a single shock at a time. The normality of two or more shocks would violate the second identification condition in **Assumption I**.
  - Such joint tests constitute a very interesting topic for further research.
  - The same applies to the limiting probability of finding $N - 1$ rejections of the univariate normality tests in those circumstances.
- Another important research topic would be the behaviour of the PMLEs of θ when **Assumption I** does not hold, either because two or more of the shocks are **Gaussian** or because they are not **independent**.

Conclusions and directions for further research

- Finally, our proposals may have little power against alternatives in which the dependence is mostly visible in certain regions of the domain of the random shocks.
- With this in mind, in "*Specification tests for non-Gaussian structural vector autoregressions*" (in progess) we have begun to study moment tests that look at the product of non-linear transformations of the shocks.