Specification tests for non-Gaussian structural vector autoregressions

Dante Amengual
CEMFI

Gabriele Fiorentini
Università di Firenze

Enrique Sentana
CEMFI

ESEM 23

UPF, August 30, 2023

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), or Stock and Watson (2018)).
- Recently, **identification** through **independent, non-Gaussian shocks** has become increasingly popular after Lanne, Meitz and Saikkonen (2017) and Gouriéroux, Monfort and Renne (2017).

- Consider an N-variate SVAR process of order p:

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

where $\boldsymbol{\varepsilon}_t^*$ and $\mathbf{C}$ are the structural shocks and their impact multipliers, respectively, and I_{t-1} the information available at time $t-1$.

Identification Assumption

1) the N shocks in $\boldsymbol{\varepsilon}_t^*$ are cross-sectionally **independent**,

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

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

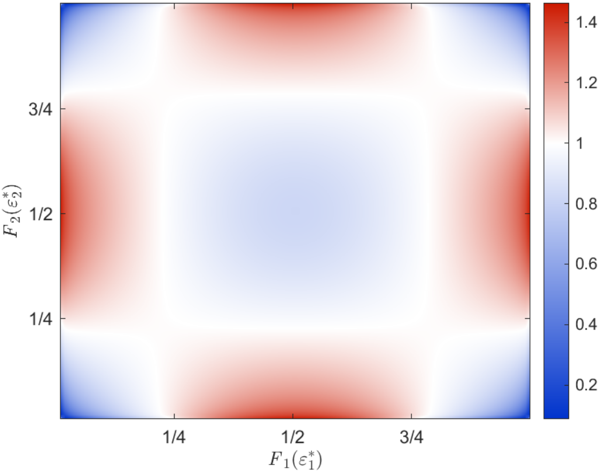
Motivation

- Failure of any of those conditions results in an underidentified model e.g.:
  - *Gaussian* ε^*: one can only identify $V(\mathbf{y}) = \mathbf{CC}'$ but not $\mathbf{C}$ despite the fact that the elements of ε^* are cross-sectionally independent.
  - *Non-Gaussian spherical distribution for* ε^*: lack of identifiability of $\mathbf{C}$ despite the non-normality because orthogonal rotations of the structural shocks share not only their mean and covariance, but also their non-linear dependence.
  - *Multivariate location-scale mixture of normals for* ε^*: Same reason.
- The signal processing literature on **Independent Component Analysis** popularised by Comon (1994) shares the same identification scheme.

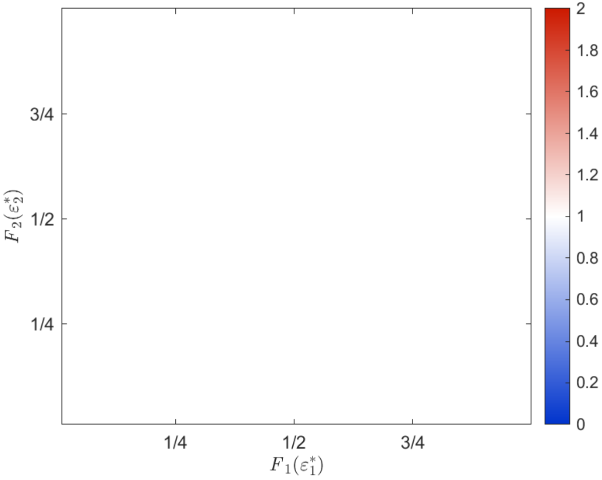
Motivation

- In Amengual, Fiorentini and Sentana (2022), we proposed to assess the potential cross-sectional **dependence** among two or more shocks by comparing the integer (product) moments of those shocks in the sample with their population counterparts.
- Specifically, we assessed the statistical significance of their third and fourth cross-moments, which should be equal to the product of the corresponding marginal moments under **independence**.
- The problem with these tests is twofold:
  1. Standard asymptotic theory provides poor finite sample approximations to moment tests based on higher-order moments, which are quite sensitive to outliers.
  2. Tests based on a finite number of moments are not consistent.

A distribution with zero co-skewness and co-kurtosis



The independence copula



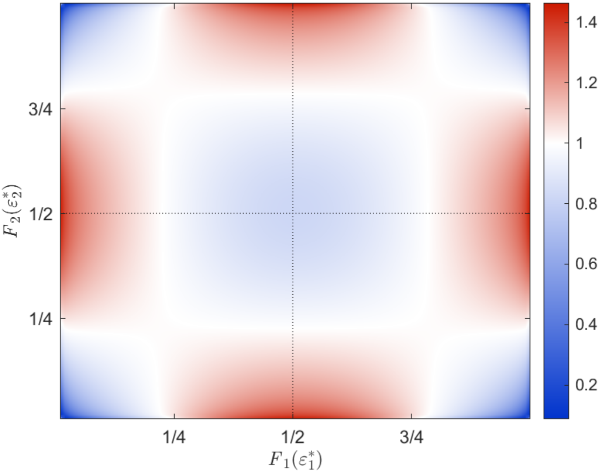
What we do

- In this paper, we propose to assess the potential cross-sectional **dependence** among two or more shocks by comparing the joint empirical cdf to the product of the marginal empirical cdfs.
- We do so for:
  1. A **discrete grid** of values of the arguments of the joint cdf, which provides the intuition for our approach.
  2. A **continuum grid** of values using the continuum of moments inference procedures in Carrasco and Florens (2000).

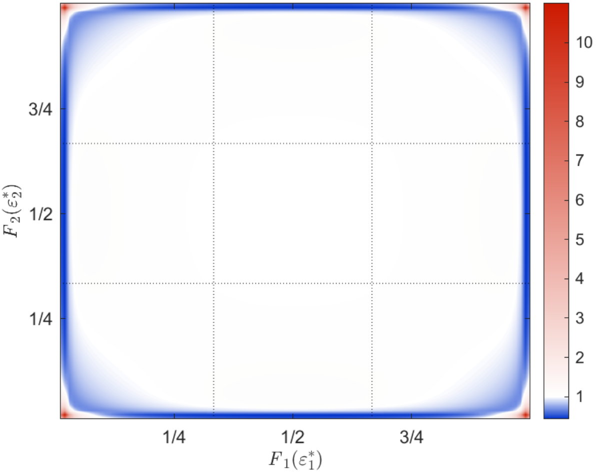
What we do

- Our discrete grid test, which is closely related to the classical independence test in contingency tables proposed by Pearson in 1900, is very simple to compute.
- The main difference is that we use shocks obtained with estimated parameter values, as well as intervals based on the empirical quantiles of those estimated shocks, so the asymptotic covariance matrix of the average influence functions must reflect these two different sources of sampling variability.
- Nevertheless, a finite grid test is not consistent either for any fixed partition of the domain of the shocks.
- In contrast, our continuum of grid values test is consistent.

A distribution with an equiprobable 2×2 contingency table



A distribution with an equiprobable 3 × 3 contingency table



What we do

- We focus on latent shocks rather than observed variables because the **Identification Assumption** is written in terms of ε^* rather than $\mathbf{y}$.
- If we knew the true values of $\boldsymbol{\mu}$ and $\mathbf{C}$ in the static case, or $\boldsymbol{\tau}$, $\mathbf{A}_j$ $(j = 1, \ldots, p)$ and $\mathbf{C}$ in the dynamic one, our tests would be straightforward.
- In practice, though, all those mean and variance parameters are unknown, so we need to estimate them before computing our tests.
- We follow Fiorentini and Sentana (2023), who showed that if the univariate log-likelihoods are based on an unrestricted **finite Gaussian mixture**, then all conditional mean and variance **parameters will be consistently estimated** (under standard regularity conditions) when the **Identification Assumption** holds.

Set up

- Let $\varepsilon_t = \mathbf{C}\varepsilon_t^*$ denote the reduced form innovations, so that $\varepsilon_t | I_{t-1} \sim i.i.d.\ (\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{\Sigma} = \mathbf{CC}'$.
- 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, $\varepsilon_t^*(\boldsymbol{\theta}) = \mathbf{C}^{-1}\varepsilon_t(\boldsymbol{\theta})$, and $\varepsilon_t(\boldsymbol{\theta}) = \mathbf{y}_t - \boldsymbol{\tau} - \mathbf{A}_1\mathbf{y}_{t-1} - \ldots - \mathbf{A}_p\mathbf{y}_{t-p}$.

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

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

Going beyond integer moments: an event-based approach

- To keep the notation simple, consider the **bivariate case**.
- For each i define H points, $k_1 < \cdots < k_h < \cdots < k_H$, so that we can form a partition of the support of ε_{it}^* into $H+1$ intervals after suitably defining $k_0 = -\infty$ and $k_{H+1} = +\infty$.
- Let $I_A(x)$ denote the usual indicator function for $x \in A$.
- **Independence** implies that the joint probability of any combined event is the product of the marginal probabilities.
- Specifically, if $\pi_{ih} = \Pr[I_{(k_{h-1},k_h)}(\varepsilon_{it}^*) = 1]$ for $i = 1, 2$, then

$$\Pr[I_{(k_{h-1},k_h)}(\varepsilon_{1t}^*) = 1, I_{(k_{h'-1},k_{h'})}(\varepsilon_{2t}^*) = 1] = \pi_{1h} \times \pi_{2'h'},$$

so that we can use as testing moment,

$$E[I_{(k_{h-1},k_h)}(\varepsilon_{1t}^*) I_{(k_{h'-1},k_{h'})}(\varepsilon_{2t}^*)] - \pi_{1h}\pi_{2h'} = 0$$

with the π_{ih}'s identified from $E[I_{(k_{h-1},k_h)}(\varepsilon_{it}^*)] - \pi_{ih} = 0$ for all i, h.

Going beyond integer moments: an event-based approach

- Intuitively, we will have the following contingency table:

$$
\begin{array}{c|c|cccc}
 & & \multicolumn{3}{c}{\varepsilon_2^*} & \\
 & & 1 & & H+1 & \\
\hline
 & 1 & n_{1,1} & \cdots & n_{1,H+1} & n_{1,\bullet} \\
\varepsilon_1^* & & \vdots & \ddots & \vdots & \vdots \\
 & H+1 & n_{H+1,1} & \cdots & n_{H+1,H+1} & n_{H+1,\bullet} \\
 & & n_{\bullet,1} & \cdots & n_{\bullet,H+1} & T
\end{array}
$$

where:

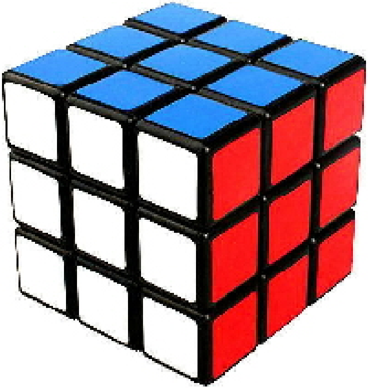
$$
\begin{aligned}
n_{i,j} &= \#\text{ of times that } \varepsilon_{1t} \in (k_{h_i-1}, k_{h_i}) \text{ and } \varepsilon_{2t} \in (k_{h_j-1}, k_{h_j}), \\
n_{i,\bullet} &= \#\text{ of times that } \varepsilon_{1t} \in (k_{h_i-1}, k_{h_i}), \text{ and} \\
n_{\bullet,j} &= \#\text{ of times that } \varepsilon_{2t} \in (k_{h_j-1}, k_{h_j}).
\end{aligned}
$$

Going beyond integer moments: an event-based approach

- In turn, in the **trivariate case** we will have...

Going beyond integer moments: an event-based approach

- In turn, in the **trivariate case** we will have... something like this:



Going beyond integer moments: an event-based approach

- Let $\mathbf{p}[\varepsilon_t^*(\boldsymbol{\theta})]$ denote a vector containing the collection of non-redundant influence functions.
- We can easily obtain the moment test statistic after providing expressions for $V\{\mathbf{p}[\varepsilon_t^*(\boldsymbol{\theta})]\}$ that account for estimation uncertainty about both π_{ih}'s and $\boldsymbol{\theta}$.
- Importantly, we show that we can write this Pearson-type test for independence as a moment test of independence of the cdf at a finite grid of points because the moment conditions of the latter are a simple full-rank linear transformation of the former with known coefficients.

Going beyond integer moments: a copula approach

- A practical issue is how to partition the support of ε^*_{it} for a given H.
- For that reason, we consider a variant of our independence test in which, instead of fixing arbitrary points of the range of each of the ε^*_i's, we look at fixed fractions of the observations.
- These tests can be understood as **independence copula tests**.
- We show that for a common *a priori* partition, **both tests are asymptotically equivalent** under the null and sequences of local alternatives.

Going beyond integer moments: about the partition

- The number of elements of the partition is crucial for both small sample performance and power considerations because:

  1. A fine partition relative to the sample size will lead to size distortions since the joint probability of some individual cells will be poorly estimated.
  2. Even in large samples, a fine partition will generate substantial correlation between the influence functions, potentially causing numerical instability.
  3. There is a power trade-off between the size of the non-centrality parameter and the number of degrees of freedom of the limiting distribution.

A continuous grid: extant tests

- As previously illustrated, the finite grid tests we have just discussed are not consistent for any specific finite partition of the domain of the shocks.

- Consistent tests for independence based on comparing the joint cdf to the product of the marginal ones go back at least to Hoeffding (1948), who considered a Cramér-von Mises type-test based on the integral of the square differences between the joint cdf and the product of the marginal cdfs.

- In turn, Blum, Kiefer and Rosenblat (1961) also considered Kolmogorov-Smirnov-type tests based on the maximum absolute discrepancy.

A continuous grid: exploiting correlation

- However, these tests rely on specific functionals of the difference, while our discrete grid tests also take into account not only the asymptotic variance of the influence functions for each value of the arguments, like an Anderson-Darling (1961) test would do, but also the covariance between those influence functions for different values of the arguments.
- In principle, we could try to find the limiting distribution of our discrete grid tests in a double asymptotic framework in which the partitions get finer and finer as the sample size increases.
- However, this is really unnecessary because the influence functions indexed with respect to the arguments of the joint cdf over the entire real line give rise to a continuum of moments.

A continuous grid: overidentification tests

- We can then follow Carrasco and Florens (2000) in constructing a J test for overidentifying restrictions based on these moments, but with a covariance operator playing the role of the usual covariance matrix.
- Specifically, define

$$p_t(\mathbf{x}) = I[\varepsilon_{it}(\boldsymbol{\theta}_0) \leq x_i, \varepsilon_{i't}(\boldsymbol{\theta}_0) \leq x_{i'}] - P[\varepsilon_{it}(\boldsymbol{\theta}_0) \leq x_i] P[\varepsilon_{i't}(\boldsymbol{\theta}_0) \leq x_{i'}]$$

so that

$$\bar{p}_T(\mathbf{x}) = \frac{1}{T} \sum_{t=1}^{T} p_t(\mathbf{x}),$$

can be regarded as the sample version of $p(\mathbf{x}) = F(\mathbf{x}) - \prod_{i=1}^{N} F_i(x_i)$.

- For any random vector with continuous joint cdf, $p(\mathbf{x})$ should be identically 0 for all $\mathbf{x}$ iff the underlying random variables are independent, which justifies consistency.

A continuous grid: empirical cdfs

- In practice, the marginal cdfs $F_i(x_i)$ will typically be unknown, but we can similarly estimate them from the continuum of influence functions

$$p_{it}(x_i) = I[\varepsilon_{it}(\boldsymbol{\theta}_0) \leq x_i] - P[\varepsilon_{it}(\boldsymbol{\theta}_0) \leq x_i],$$

  which effectively leads to the empirical cdf of the i^{th} shock.

- An extension of our earlier results implies that the continuum of moments test that looks at $p_T(\mathbf{x})$ over the entire real line will be numerically equivalent to the one that looks at the difference between the copula and the unit hyperplane over the unit hypercube.

- Like in the discrete grid test, though, we must correct for the sampling variability in estimating marginal cdfs as well as mean and variance parameters.
  - We do so by applying the Carrasco and Florens (2000) approach to first-order expansion of the influence functions previously discussed taken with respect to the off-diagonal elements of $\mathbf{C}$.

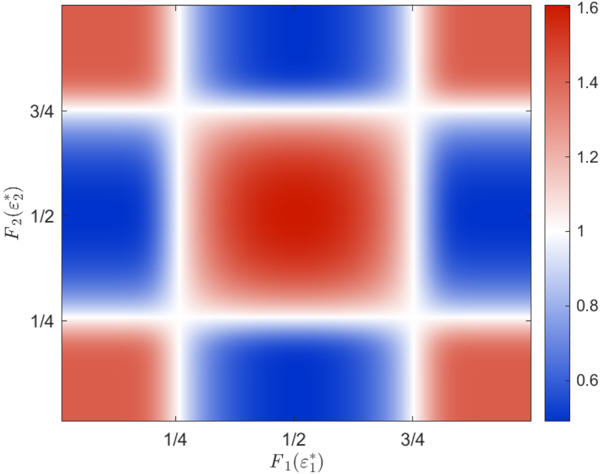
Monte Carlo evidence: model and estimation

- We look at bivariate and trivariate DGPs with VAR(1) dynamics.
- We consider both:
    - $T = 250$, realistic in most macro applications, and
    - $T = 1,000$, representative of financial applications.
- To estimate the model parameters, we assume that ε^*_{it}'s follow serially and cross-sectionally independent standardised discrete mixture of two normals, which provides consistent estimators under the null regardless of the true distribution of the shocks.
- We compare our tests not only to the co-skewness/co-kurtosis tests in Amengual, Fiorentini and Sentana (2022), but also to the Kolmogorov-Smirnov test that imposes the independent copula under the null, like in Blum, Kiefer and Rosenblat (1961), as well as to the Matteson and Tsay (2017) statistic based on the ICA criterion function.

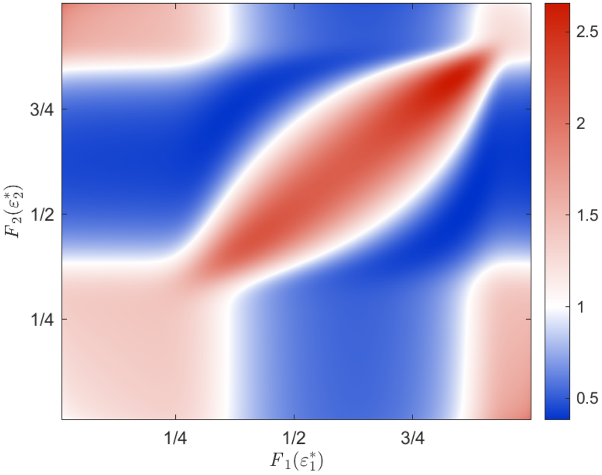
Monte Carlo evidence: DGPs design

- The bivariate DGP for the standardised shocks that we consider **under the null** of independence imposes that one of them follows a Student t with 10 degrees of freedom, while the other one an asymmetric t with kurtosis and skewness coefficients equal to 4 and $-.5$, respectively, while in the trivariate case, ε_{3t} follows an asymmetric t with the same kurtosis but opposite skewness coefficient as ε_{2t}.
- Additionally, we consider as **alternative hypotheses**:
  1. Standardised scale mixture of two zero mean normals with scalar covariance matrices.
  2. Standardised mixture of two normal vectors.
  3. Asymmetric t with skewness vector $\boldsymbol{\beta} = -10\ell_N$ and degrees of freedom parameter $\nu = 12$ (see Mencía and Sentana (2012) for details).
- To gauge the small-sample properties of our tests, we generate (10,000) 2,500 samples for each of the designs under the null (alternative), systematically using 99 additional samples that impose the null to obtain resampling critical values.

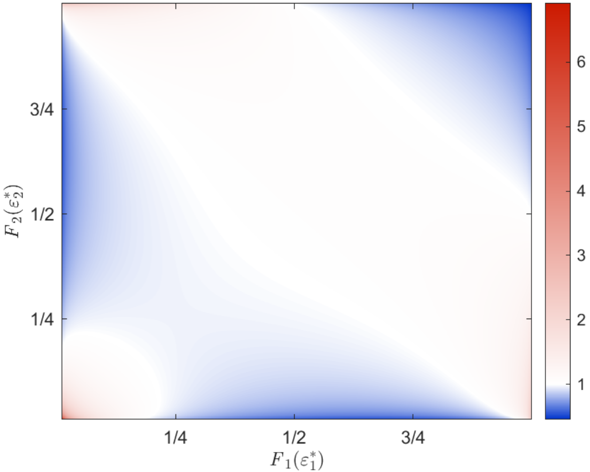
Scale mixture of normals



Finite normal mixture



Asymmetric Student t distribution



Monte Carlo evidence: Size properties

	Discrete Q tests							Continuous Q tests		
	Asymptotic critical values			Resampling critical values				Resampling critical values		
H	10%	5%	1%	10%	5%	1%	α	10%	5%	1%
				$N = 2,\ T = 250$						
3	8.4	4.0	0.9	9.2	4.5	0.8	10^{-6}	9.4	4.5	0.8
5	8.6	4.4	0.8	9.2	4.4	0.8	10^{-8}	9.3	4.5	0.9
				$N = 3,\ T = 250$						
3	8.2	3.8	0.8	9.2	4.3	0.9	10^{-6}	8.3	3.8	0.7
5	8.5	4.1	0.9	9.5	4.5	0.9	10^{-8}	8.0	4.1	0.7
				$N = 2,\ T = 1,000$						
3	9.2	4.3	0.9	9.4	4.8	0.9	10^{-6}	10.7	5.5	1.3
5	9.6	4.6	0.8	10.3	5.2	1.0	10^{-8}	10.6	5.4	1.2

Monte Carlo evidence: Power properties (N=2 & T=250)

	Scale mixture of two normals			Finite normal mixture			Asymmetric Student t		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
H				Discrete Q tests					
3	54.8	41.4	16.4	89.3	81.7	55.1	34.9	23.0	7.0
5	36.6	23.0	7.1	90.4	81.9	56.1	33.5	21.1	6.3
α				Continuous Q tests					
10^{-6}	46.5	33.0	11.7	95.0	90.1	69.8	43.9	30.1	12.1
10^{-8}	45.7	32.2	11.4	94.7	90.1	69.7	42.9	29.8	11.8
				Moment tests					
Skew	9.4	4.4	0.6	32.7	21.3	7.7	89.1	80.9	52.5
Kurt	14.3	8.0	1.7	28.8	17.8	5.2	60.4	48.8	23.0
				Other tests					
MT	15.1	7.2	1.0	89.7	79.9	48.2	70.4	58.3	29.6
KS	18.9	9.6	2.0	76.2	65.5	40.4	34.4	21.7	6.1

Monte Carlo evidence: Power properties (N=3 & T=250)

	Scale mixture of two normals			Finite normal mixture			Asymmetric Student t		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
H				Discrete Q tests					
3	90.8	84.3	61.4	98.1	96.2	86.3	19.6	11.2	2.8
5	61.6	46.2	22.8	98.4	96.1	83.0	16.8	9.6	1.2
α				Continuous Q tests					
10^{-6}	68.5	53.9	23.3	99.6	98.8	92.7	58.4	44.4	20.0
10^{-8}	66.9	51.9	21.6	99.3	98.6	91.9	58.4	45.2	21.2
				Moment tests					
Skew	9.6	4.7	1.0	38.8	25.8	9.4	94.8	88.4	69.2
Kurt	13.2	7.2	1.3	38.6	26.3	8.1	73.2	66.4	40.8
				Other tests					
MT	19.6	9.0	1.6	95.3	89.1	59.7	76.0	63.2	32.4
KS	11.3	5.7	1.0	28.6	16.8	4.5	32.8	15.6	2.8

Monte Carlo evidence: Power properties (N=2 & T=1,000)

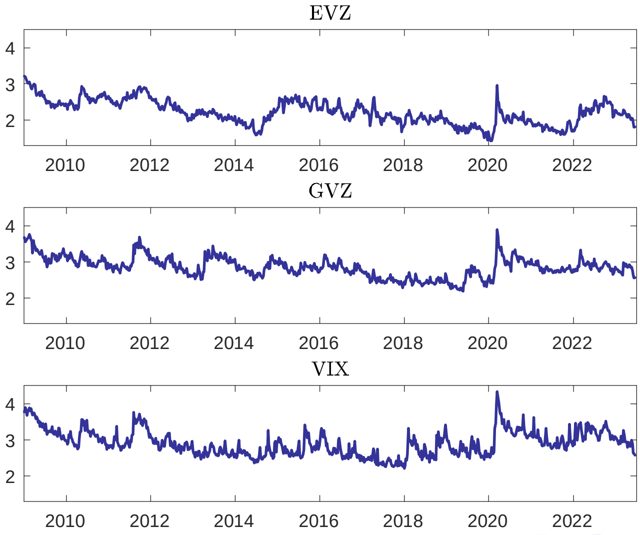
	Scale mixture of two normals			Finite normal mixture			Asymmetric Student t		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
H				Discrete Q tests					
3	100	99.9	99.2	100	100	100	95.1	90.2	72.6
5	99.8	99.7	96.7	100	100	100	96.6	93.6	78.4
α				Continuous Q tests					
10^{-6}	98.6	96.7	83.5	100	100	100	94.1	89.8	69.3
10^{-8}	97.4	94.8	78.4	100	100	100	93.5	88.8	67.0
				Moment tests					
Skew	11.6	6.0	1.5	84.4	76.0	53.4	100	100	99.6
Kurt	50.4	36.8	15.7	80.4	71.9	40.2	95.7	91.1	70.4
				Other tests					
MT	99.8	98.4	72.2	100	100	100	100	100	99.7
KS	82.0	69.9	36.1	99.9	99.9	98.5	97.2	92.4	67.7

Empirical application to volatility indices

- We consider three weekly series of market-based implied volatilities: the VIX index, the EVZ EuroCurrency ETF volatility index, and the GVZ Gold ETF volatility index.
- They represent three of the most actively traded asset classes: stocks, exchange rates and commodities.
- Our sample spans 2009/01/21 to 2023/06/21 (753 Wednesday observations).
- Weekly (log) volatility indices show mean reversion over the long run but persistent mean deviations during extended periods.
- But we are really interested in their dynamic linkages.

Empirical application to volatility indices

Volatility index series (logs)



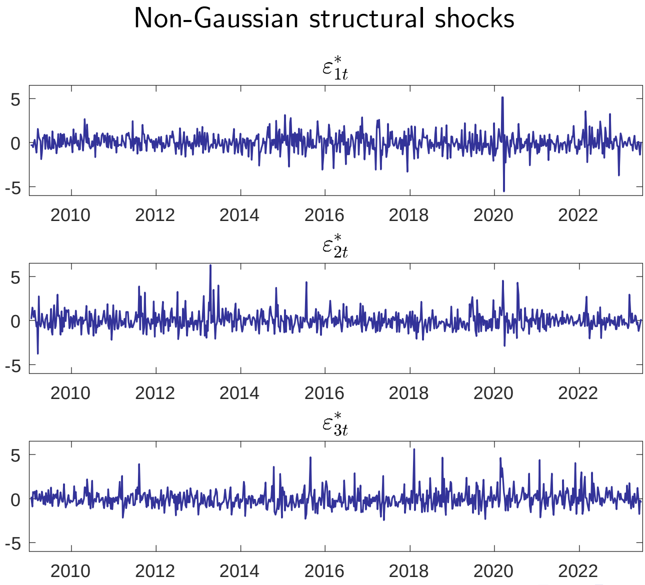
Empirical application to volatility indices

- We estimate the following trivariate SVAR:

$$\mathbf{x}_t = \boldsymbol{\tau} + \mathbf{A}_1\mathbf{x}_{t-1} + \cdots + \mathbf{A}_p\mathbf{x}_{t-p} + \mathbf{C}\varepsilon_t^*,$$

  selecting $p = 2$ by looking at the Akaike information criterion and the likelihood ratio test for lack of residual serial correlation.
- We consistently estimate $(\boldsymbol{\tau}, \mathbf{a}, \mathbf{c})$ jointly by PML assuming that $\varepsilon_{it}^* \sim i.i.d.\ DLSMN(\delta_i, \kappa_i, \lambda_i)$.
- For initial values, we run OLS regressions for each of the three variables, we apply fastICA routine to the OLS residuals and, finally, the EM algorithm for mixture parameters.

Empirical application to volatility indices

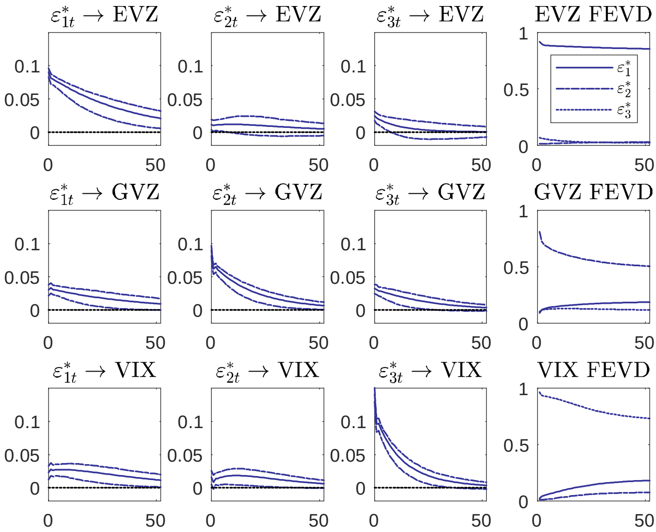


Empirical application: Results

- If we use the unmixing matrix $\mathbf{C}^{-1}$ to try and intepret the shocks as "long/short portfolios" of the one-period ahead prediction errors, we find that each of them "invests" approximately between 130 and 180% on one of the reduced form shocks and between -15 and -40% on each of the other two.
- We can get a more standard intepretation by looking at the IRFs up to a year ahead.
- The strong persistence implied by the $\text{SVAR}(2)$ parameter estimates implies that the IRFs decay slowly.
- As can be seen, each series reacts mostly to one shock.
- Nevertheless, they also react significantly to the other ones, especially in the case of the Gold volatility index and to some extent the VIX.

Empirical application to volatility indices

Impulse response functions and FEVDs



Empirical application: Specification tests

- The different test procedures that we have considered fail to reject the null hypothesis of stochastic independence of the structural shocks, with the exception of the co-skewness and co-kurtosis tests in Amengual, Fiorentini and Sentana (2022).
- Nevertheless, these rejections seem to be closely associated to the unusual behaviour of the three series at the onset of the COVID-19 pandemic.
- Specifically, if we remove two additive outliers from the observations of the three series for March 11 and 18, 2020, using the procedures in Chen and Liu (1993), the moment tests no longer reject.
- In contrast, the quantile-based independence tests that we have proposed in this paper, and indeed the MT and KS tests seem far more robust to the presence of these ususual observations.

Conclusions and directions for further research

- Identification through independent non-Gaussian shocks is a powerful result but not without concerns.
- For that reason, it would be desirable that empirical researchers checked the underlying assumptions to increase the credibility of their findings.
- Our proposed tests can be rather useful in this respect.
- An important question for further research is what would happen to our proposed tests when the true joint distribution of the shocks is Gaussian.
- Given that most theoretical models implicitly assume independent underlying economic shocks even though this is not exploited in estimation, it should also be of interest to apply our tests to the shocks of SVAR models identified using more traditional methods.