Consistent statistical identification of SVARs under (co-)heteroskedasticity of unknown form

Helmut Herwartz Shu Wang

Chair of Econometrics
Georg August University of Göttingen
Humboldtallee 3, Göttingen D-37073, Germany

EEA-ESEM 2024, Rotterdam
August 26, 2024

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

"Has the Simultaneous Causality Problem Been Solved?"

Simultaneous equations model:

$$\mathrm{x}_t = B\xi_t,\ \ \xi_t \sim (\mathbf{0}, I_N),\ \ t = 1, \ldots, T, \tag{1}$$

1. x_t: $N \times 1$ vector of observables
2. ξ_t: $N \times 1$ vector of latent structural shocks
3. B: $N \times N$ non-singular matrix
4. $\mathbb{E}[\mathrm{x}_t \mathrm{x}_t^\top] = BB' = \Omega$
5. B under-identified from the first two moments

Most relevant applications: Factor models, structural VAR

"Has the Simultaneous Causality Problem Been Solved?"

Assumption A

The random vector ξ_t contains N independent components and at most one component exhibits a Gaussian distribution.

Darmois-Skitovich characterization theorem: Under Assumption A, matrix B is uniquely identified up to column permutations and sign flips.

Identification by ICA:

1. ML methods relying on independence: Lanne et al. (2017), Gouriéroux et al. (2017), Fiorentini and Sentana (2023), Jarociński (2024)
2. GMM or CMD methods relying on (weaker) mean/moment independence: Lanne and Luoto (2021), Keweloh (2021), Guay (2021), Mesters and Zwiernik (2022)
3. CMD or HL methods relying on independence: Herwartz (2018), Drautzburg and Wright (2023), Herwartz and Wang (2023, 2024)

"Has the Simultaneous Causality Problem Been Solved?"

...but

Assumption A rules out an empirically relevant scenario where shocks exhibit common volatility processes, known as **co-heteroskedasticity** (Montiel Olea et al. 2022, Lewis 2024):

- *Great Moderation*
- co-movement of shock volatility during economic downturns and financial distress
- expect larger 'precautionary oil demand shock' if there is a larger oil supply shock, like geopolitical conflicts causing production disruptions or supply bottlenecks (sanctions)

$\Rightarrow$ shocks are dependent in their second-order moments

In this paper:

We develop a heteroskedasticity-consistent kernel-based maximum likelihood (HC-KML) approach for identifications of structural VARs, study its asymptotic and finite-sample properties, and apply it to a stylized US monetary policy model.

The HC-KML estimator:

- maximizes a log-likelihood function, constructed using non-parametric variance smoothing and kernel density estimates
- adjusts each orthogonalized residual by its corresponding variance, without imposing specific parametric forms for the variance process
- consistent regardless of shock distributions, and whether they are (co-)heteroskedastic or not
- inherits the favorable finite-sample properties of KML estimator (Hafner et al. 2024)

Content

1. Introduction
2. Identification
3. Estimation
4. Simulation
5. Application
6. Conclusion

Content

Structural VAR model

Structural VAR(p) and VMA(∞) models:

$$A(L)y_t \equiv \mathrm{x}_t = B\xi_t \tag{2}$$

$$y_t = A(1)^{-1}B\xi_t = \Phi(L)\xi_t \tag{3}$$

1. y_t: $N \times 1$ vector of observables
2. $A(z) = \left(I_N - \sum_{j=1}^{p} A_j z^j\right)$, $z \in \mathbb{C}$, is a reverse characteristic polynomial of order p
3. $\Phi(L) = \sum_{j=0}^{\infty} \Phi_j L^j$: IRFs with $\Phi_0 = B$ (structural impact multiplier)

Heteroskedasticity and dependence

Simultaneous equations model with heteroskedastic shocks:

$$\mathrm{x}_t = B\xi_t = B\Sigma_t^{1/2}\eta_t, \ \eta_t \overset{iid}{\sim} (\mathbf{0}, I_N), \tag{4}$$

- $\Sigma_t = \mathrm{diag}\left(\sigma_{1t}^2, \ldots, \sigma_{Nt}^2\right)$: marginal variances of the structural shocks

Co-heteroskedasticity:

$$\xi_{it} = \sigma_t \eta_{it} \quad \text{and} \quad \xi_{jt} = \sigma_t \eta_{jt}.$$

$$\mathbb{E}[\xi_{it}\xi_{jt}] = \mathbb{E}[\sigma_t^2]\mathbb{E}[\eta_{it}\eta_{jt}] = 0, \ \mathbb{E}[\xi_{it}|\xi_{jt}] = \sigma_t\mathbb{E}[\eta_{it}] = 0, \ \mathbb{E}[\xi_{it}^2|\xi_{jt}] = \sigma_t^2\mathbb{E}[\eta_{it}^2] = \sigma_t^2$$

Assumption A is violated, structural shocks ξ_{it} and ξ_{jt} under-identified.

Co-heteroskedasticity leads to dependence

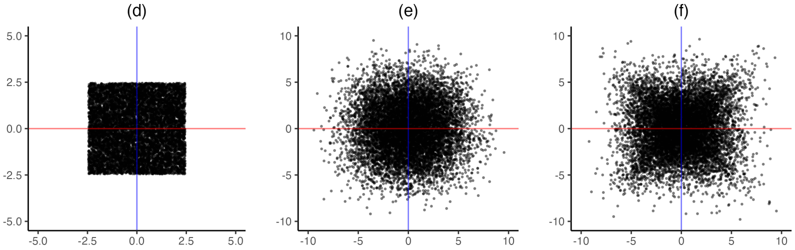


Figure 1: Joint distribution of $(\xi_{1t}, \xi_{2t})'$ with time-varying variances. Red and blue lines indicate orientations of marginals of ξ_{1t} and ξ_{2t}. Uniform shocks with χ^2 random variances. From left to right: Homoskedasticity, (independent) heteroskedasticity, and co-heteroskedasticity.

Identification based on time-varying volatility

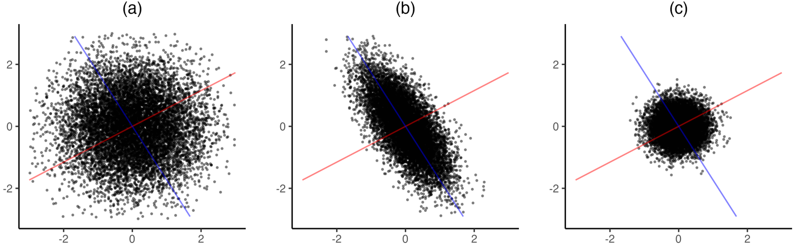


Figure 2: Joint distribution of structural shocks $(\xi_{1t}, \xi_{2t})'$ with variance changes. Red and blue lines indicate orientations of marginals of ξ_{1t} and ξ_{2t}. From left to right: Pre-change, non-proportional change, and proportional change (i.e., co-heteroskedasticity).

Using a selection of moment restrictions

- $\mathbb{E}[\xi_{it}^2 \xi_{jt}] = \mathbb{E}[\sigma_{it}^2 \sigma_{jt}]\mathbb{E}[\eta_{it}^2]\mathbb{E}[\eta_{jt}] = 0$
- $\mathbb{E}[\xi_{it}^3 \xi_{jt}] = \mathbb{E}[\sigma_{it}^3 \sigma_{jt}]\mathbb{E}[\eta_{it}^3]\mathbb{E}[\eta_{jt}] = 0$
- $\mathbb{E}[\xi_{it}^2 \xi_{jt}^2] = \mathbb{E}[\sigma_{it}^2 \sigma_{jt}^2]\mathbb{E}[\eta_{it}^2]\mathbb{E}[\eta_{jt}^2] \neq 0$
- symmetric moment independence conditions can be violated
- asymmetric moment/cumulant conditions can remain valid
- GMM or CMD approaches rely on a subset of valid moment conditions, like (Keweloh 2021, Guay 2021, Mesters and Zwiernik 2022), permit (co-)heteroskedasticity
- but high-order moment statistics might be sensitive to heavy-tailed source distributions

Content

Identification under (co-)heteroskedasticity

Allowing (co-)heteroskedasticity:

$$\mathrm{z}_t := \Omega^{-1/2}\mathrm{x}_t = Q\Sigma_t^{1/2}\eta_t, \tag{5}$$

where Q is orthogonal such that $QQ^\top = I_N$ and $|\det Q| = 1$.

Assumption B

The random vector η_t contains N independent components and at most one component exhibits a Gaussian distribution.

Theorem 1

Under Assumption B, the mixing matrix Q in (5) is identified up to a right-multiplication by $\mathcal{D}\mathcal{P}$, where $\mathcal{D}$ is diagonal and $\mathcal{P}$ is a permutation matrix.

Content

Estimation under (co-)heteroskedasticity: KML

Parametrization of Q

$$Q = Q(\theta) = \left(\prod_{i=1}^{N-1} \prod_{j=i+1}^{N} \mathcal{G}_{i,j}(\theta_n) \right)^\top,$$

- $\theta \in \boldsymbol{\Theta} \subset \mathbb{R}^{N(N-1)/2}$
- $\xi_t(\theta) := Q(\theta)^\top \mathrm{z}_t$, $\xi_{it}(\theta) := e_i^\top \xi_t(\theta)$, $\eta_{it}(\theta) = \xi_{it}(\theta)/\sigma_{it}$
- θ_0: true parameter value; $\eta_t := \eta_t(\theta_0)$: true independent components and f_{i,θ_0}: pdf of η_{it}.

ML estimator maximizes

$$l_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} \left(\log f_{i,\theta_0}\left(\eta_{it}(\theta)\right) - \log \sigma_{it} \right) \qquad (6)$$

Estimation under (co-)heteroskedasticity: KML

KML estimator maximizes

$$\tilde{l}_T(\theta) = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\left(\log \hat{f}_{i,\theta}\left(\eta_{it}(\theta)\right) - \log \sigma_{it}\right), \tag{7}$$

$$\hat{f}_{i,\theta}(\eta_{it}(\theta)) = \frac{1}{Th_f}\sum_{k=1}^{T} K_f\left(\frac{\eta_{it}(\theta) - \eta_{ik}(\theta)}{h_f}\right), \ \ i = 1, \ldots, N,$$

Theorem 2

Let $f_{i,\theta}(x)$ *be the marginal pdf of* $\eta_{it}(\theta)$, $\mathbb{E}_{i,\theta}[g(X)] = \int_{\mathbb{R}} g(x) f_{i,\theta}(x)dx$, $\eta_{it} := e_i^\top \eta_t(\theta_0)$, *suppose* $\theta_0 \in \Theta^*$, *for all* $\epsilon > 0$, $\exists \mathcal{B}(\theta_0, \epsilon)$ *s.t.,*

$$\sup_{\theta \in \Theta^* \cap \mathcal{B}(\theta_0,\epsilon)^c} \sum_{i=1}^{N} \mathbb{E}_{i,\theta}\left[\log f_{i,\theta}(\eta_{it}(\theta))\right] < \sum_{i=1}^{N} \mathbb{E}_{i,\theta_0}\left[\log f_{i,\theta_0}(\eta_{it})\right]. \tag{8}$$

Proof of Theorem 2 ($N = 2$)

Lemma 3

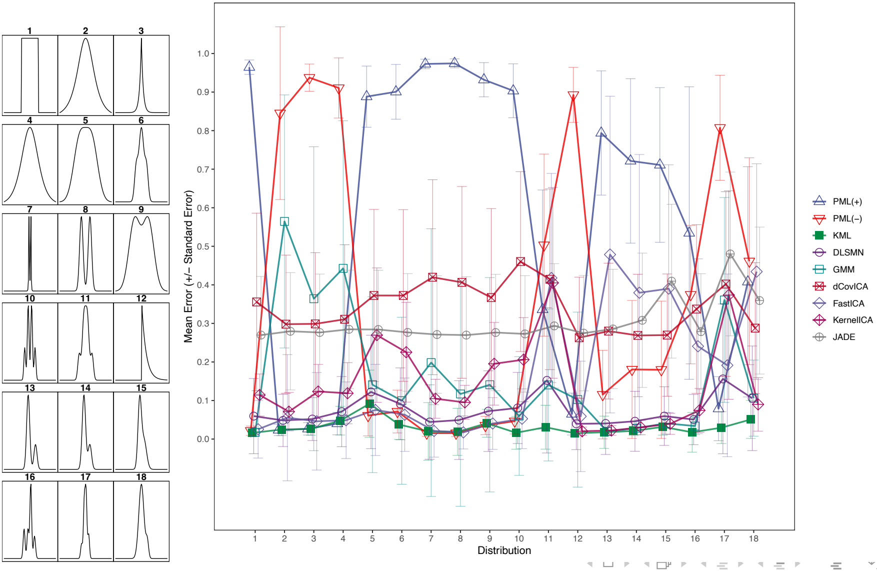
$$\mathbb{E}_{\theta}\left[\log f_{\theta}(\eta_t(\theta))\right] = \mathbb{E}_{\theta_0}\left[\log f_{\theta_0}(\eta_t)\right], \tag{9}$$

where $\mathbb{E}_{\theta}[g(\boldsymbol{X})] = \int_{\mathbb{R}^N} g(\boldsymbol{x}) f_{\theta}(\boldsymbol{x}) |d\boldsymbol{x}|$ *with* $|d\boldsymbol{x}| = dx_1 \dots dx_N$ *and* $f_{\theta}(x)$ *is the N-dimensional joint pdf of* $\eta_t(\theta)$.

Proof of Theorem 2

$$\begin{aligned}
\sum_{i=1}^{2} \mathbb{E}_{i,\theta}\left[\log f_{i,\theta}(\eta_{it}(\theta))\right] - \sum_{i=1}^{2} \mathbb{E}_{i,\theta_0}\left[\log f_{i,\theta_0}(\eta_{it})\right] &= \mathbb{E}_{\theta}\left[\log\left(f_{1,\theta}(\eta_1(\theta)) f_{2,\theta}(\eta_2(\theta))\right)\right] - \mathbb{E}_{\theta_0}\left[\log f_{\theta_0}(\eta_t)\right] \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \log\left(\frac{f_{1,\theta}(\eta_1) f_{2,\theta}(\eta_2)}{f_{\theta}(\eta_1, \eta_2)}\right) f_{\theta}(\eta_1, \eta_2) d\eta_1 d\eta_2 \\
&\leq \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\frac{f_{1,\theta}(\eta_1) f_{2,\theta}(\eta_2)}{f_{\theta}(\eta_1, \eta_2)} - 1\right) f_{\theta}(\eta_1, \eta_2) d\eta_1 d\eta_2 \\
&= \int_{\mathbb{R}} f_{2,\theta}(\eta_2) d\eta_2 \int_{\mathbb{R}} f_{1,\theta}(\eta_1) d\eta_1 - \int_{\mathbb{R}}\int_{\mathbb{R}} f_{\theta}(\eta_1, \eta_2) d\eta_1 d\eta_2 = 1 - 1 = 0.
\end{aligned}$$

KML (Hafner, Herwartz & Wang 2024)



Heteroskedasticity-consistent KML (HC-KML)

Problem: Σ_t unknown;

Assumption C

(i) $\exists p > 1$, $\mathbb{E}[\eta_{it}^p \eta_{jt}^p \eta_{kt}^p \eta_{lt}^p] \leq M_\eta < \infty$, $\mathbb{E}[\sigma_{it}^p \sigma_{jt}^p \sigma_{kt}^p \sigma_{lt}^p] \leq M_\sigma < \infty$
(ii) as $T \to \infty$, $\sigma_{i,\lfloor sT \rfloor} \Rightarrow \sigma_i(s)$, $\sigma_i(s)$ *continuous a.s. on* $\mathcal{C}[0,1]$,
$\sup_{s\in[0,1]} \sigma_i(s) \leq \sigma_+ < \infty$, $\inf_{s\in[0,1]} \sigma_i(s) \geq \sigma_- > 0$, $\forall s$ *and* $\int_0^1 \sigma_i^{-2}(s)ds < \infty$

$$\hat{\sigma}_{it}^2(\theta) := \frac{\sum_{k=1}^T K_\sigma((k-t)/h_\sigma)\xi_{ik}^2(\theta)}{\sum_{k=1}^T K_\sigma((k-t)/h_\sigma)},$$

- $K_\sigma : [-1,1] \mapsto [0,1]$ kernel s.t. $\int K_\sigma(x)dx > 0$, h_σ: bandwidth
- $\hat{\eta}_{it}(\theta) = \xi_{it}(\theta)/\hat{\sigma}_{it}(\theta)$

HC-KML estimator maximizes

$$\hat{l}_T(\theta) = \frac{1}{T}\sum_{t=1}^T \sum_{i=1}^N \left(\log \hat{f}_{i,\theta}(\hat{\eta}_{it}(\theta)) - \log \hat{\sigma}_{it}(\theta)\right), \tag{10}$$

Asymptotic properties of HC-KML

Assumption D

(i) The true parameter value $\theta_0 \in \mathbf{\Theta}$ *lies in the parameter space* $\mathbf{\Theta}$*, a compact subspace of* $\mathbb{R}^{N(N-1)/2}$*, on which the local identification condition holds.*
(ii) For all $i \in \{1, \ldots, N\}$*,* f_i *is uniformly continuous with bounded derivatives on its support, and* $\inf_x f_i(x) \geq \epsilon_1 > 0$*.*
(iii) The kernel function in the density estimator satisfies: $K_f(x) \geq 0$*,* $\int K_f(x)dx = 1$*,* $\int |K_f(x)|dx < \infty$*,* $|x|K_f(x) \to 0$ *as* $|x| \to \infty$*,* $\sup_x K_f(x) < \infty$*,* $\sup_x |dK_f(x)/dx| < \infty$ *and* $K_f(0) \geq \epsilon_2 > 0$*. The bandwidth parameter satisfies:* $h_f \to 0$ *and* $Th_f/\log(T) \to \infty$ *as* $T \to \infty$*.*
(iv) The kernel function in the volatility smoothing satisfies $\int K_\sigma(x)dx > 0$*. The bandwidth parameter satisfies:* $h_\sigma = \alpha T^\beta$ *for some* $0 < \alpha < \infty$ *and* $(4p)^{-1} < \beta < 1$*, where p is determined in Assumption C(i).*

Asymptotic properties of HC-KML

Theorem 4

Under Assumptions B, C and D,

$$\hat{\theta} \xrightarrow{p} \theta_0$$

Remark 1

If for some $i \in \{1, \ldots, N\}$, $\sigma_{it}^2 = \mathbb{E}[\sigma_{it}^2]$, *a.s.* $\forall t$, $\hat{\theta} \xrightarrow{p} \theta_0$.

Asymptotic properties of HC-KML

To ensure that the sequence of score estimates is bounded, we define the trimmed estimator

$$\hat{\psi}_{it}(\hat{\eta}_{it}) := \begin{cases} \tilde{\psi}_{it}(\hat{\eta}_{it}), & \text{if } \hat{\eta}_{it} \leq M_T^{\hat{\eta}},\ \hat{f}_i(\hat{\eta}_{it}) \geq d_T,\ \hat{f}_i'(\hat{\eta}_{it}) \leq M_T^{\psi} \hat{f}_i(\hat{\eta}_{it}) \\ 0, & \text{otherwise,} \end{cases}$$

where we assume for constant sequences $d_T \to 0$, $M_T^{\hat{\eta}} \to \infty$ and $M_T^{\psi} \to \infty$, as $T \to \infty$. This also ensures that the estimated variance sequence is bounded away from zero. Alternative trimming conditions for $\hat{\sigma}_{it}^2$ have been discussed in Hansen (1995).

Assumption E

(i) For all $i \in \{1, \ldots, N\}$, f_i is symmetric about zero.
(ii) The bandwidth parameter h_f satisfies $h_f M_T^{\psi} \to 0$, and $T^{-1} M_T^{\hat{\eta}} / h_f^3 \to 0$.
(iii) The true parameter value $\theta_0 \in$ interior$(\mathbf{\Theta})$.

Asymptotic properties of HC-KML

Theorem 5

Under Assumptions B, C, D and E, $\sqrt{T}(\hat{\theta} - \theta_0) \Rightarrow \mathcal{J}(\theta_0)^{-1} \left(\sum_{i=1}^{N} \int_0^1 \sigma_i^{-1}(s) d\boldsymbol{W}_i(s) \right)$, *where* $\boldsymbol{W}_i(s)$ *is a vector Brownian motion on* $\mathcal{C}^{N(N-1)/2}[0,1]$ *with covariance matrix* $\mathcal{I}_i(\theta_0)$ *for all* $i = 1, \ldots, N$ *and*

$$\mathcal{I}_i(\theta_0) = \mathbb{E}\left[\psi_{it}^2 \frac{\partial \xi_{it}(\theta_0)}{\partial \theta} \frac{\partial \xi_{it}(\theta_0)}{\partial \theta^\top} \right],$$

$$\mathcal{J}(\theta_0) = -\sum_{i=1}^{N} \mathbb{E}\left[\psi_{it} \frac{\partial^2 \xi_{it}(\theta_0)}{\partial \theta \partial \theta^\top} + \psi'_{it} \frac{\partial \xi_{it}(\theta_0)}{\partial \theta} \frac{\partial \xi_{it}(\theta_0)}{\partial \theta^\top} \right] \left(\int_0^1 \sigma_i^{-2}(s) ds \right)$$

Remark 2

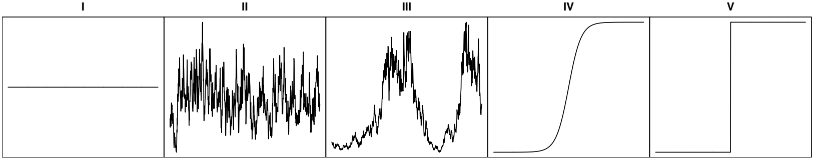
If $\sigma_{it}^2 = \mathbb{E}[\sigma_{it}^2]$, *a.s.* $\forall t, \forall i$, $\sqrt{T}(\hat{\theta} - \theta_0) \Rightarrow \mathcal{N}\left(0, \mathcal{J}(\theta_0)^{-1} \mathcal{I}(\theta_0) \mathcal{J}(\theta_0)^{-1}\right)$, *where* $\mathcal{I}(\theta_0) = \sum_{i=1}^{N} \mathcal{I}_i(\theta_0)$.

Content

Finite-sample properties of HC-KML

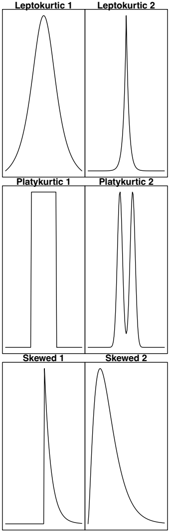
Data generating process

- $z_t = Q\xi_t = Q\Sigma_t^{1/2}\eta_t$, $\eta_t \sim$ leptokurtic/platykurtic/skewed distributions
- $Q = Q(\theta)$, $\theta \sim \text{unif}[-\pi; \pi]^{N(N-1)/2}$
- $\Sigma_t = \sigma_t^2 I_N$: common variance, five scenarios



Compare 7 competing approaches: i) KML (using Assumption A), ii) HC-KML, iii) TV-KML (infeasible), iv) DLSMN with two components, v) GMM using asymmetric third- and fourth-order moment conditions, vi) PML(+) and vii) PML(−) with Student-t and a Gaussian mixture pdf

Finite-sample properties of HC-KML



Finite-sample properties of HC-KML

Vola. Scenario	I	II	III	IV	V	I	II	III	IV	V
			$T = 50$					$T = 100$		
KML	**.094 (.115)**	.228 (.242)	.304 (.277)	.252 (.248)	.281 (.268)	**.056 (.063)**	.107 (.143)	.248 (.262)	.185 (.229)	.190 (.230)
HC-KML	.141 (.167)	**.171 (.192)**	**.163 (.183)**	**.146 (.170)**	**.156 (.183)**	.077 (.102)	**.090 (.110)**	**.090 (.112)**	**.075 (.088)**	**.082 (.109)**
TV-KML	.091 (.106)	.103 (.132)	.100 (.130)	.096 (.119)	.098 (.119)	.056 (.062)	.054 (.068)	.052 (.058)	.059 (.072)	.057 (.066)
DLSMN	.139 (.164)	.280 (.264)	.361 (.294)	.396 (.303)	.417 (.306)	.107 (.148)	.162 (.203)	.346 (.299)	.403 (.322)	.406 (.304)
GMM	.338 (.300)	.412 (.288)	.445 (.292)	.442 (.279)	.455 (.284)	.277 (.296)	.337 (.281)	.429 (.287)	.436 (.269)	.432 (.270)
PML(+)	.360 (.308)	.404 (.299)	.416 (.306)	.426 (.307)	.434 (.306)	.345 (.313)	.364 (.306)	.398 (.312)	.408 (.322)	.422 (.320)
PML(−)	.491 (.322)	.458 (.310)	.461 (.316)	.443 (.314)	.432 (.314)	.489 (.340)	.488 (.339)	.454 (.316)	.440 (.329)	.447 (.329)
			$T = 200$					$T = 1,000$		
KML	**.034 (.040)**	.048 (.060)	.198 (.255)	.101 (.175)	.123 (.208)	**.012 (.011)**	**.014 (.013)**	.143 (.269)	.029 (.096)	.033 (.112)
HC-KML	.041 (.040)	**.046 (.046)**	**.048 (.049)**	**.044 (.043)**	**.044 (.042)**	.014 (.012)	.015 (.013)	**.015 (.012)**	**.014 (.012)**	**.015 (.012)**
TV-KML	.034 (.032)	.032 (.029)	.033 (.030)	.032 (.034)	.033 (.030)	.012 (.011)	.013 (.012)	.012 (.010)	.012 (.010)	.012 (.010)
DLSMN	.074 (.121)	.088 (.130)	.304 (.293)	.389 (.333)	.376 (.323)	.048 (.106)	.053 (.125)	.252 (.316)	.346 (.368)	.377 (.367)
GMM	.216 (.278)	.225 (.255)	.402 (.293)	.444 (.261)	.436 (.263)	.093 (.207)	.092 (.200)	.335 (.291)	.389 (.255)	.382 (.255)
PML(+)	.315 (.321)	.345 (.327)	.377 (.325)	.405 (.342)	.426 (.339)	.307 (.351)	.320 (.356)	.346 (.364)	.410 (.393)	.404 (.388)
PML(−)	.518 (.371)	.490 (.364)	.476 (.334)	.435 (.341)	.420 (.335)	.493 (.414)	.496 (.414)	.494 (.369)	.389 (.337)	.394 (.333)

Finite-sample properties of HC-KML (estimated processes)

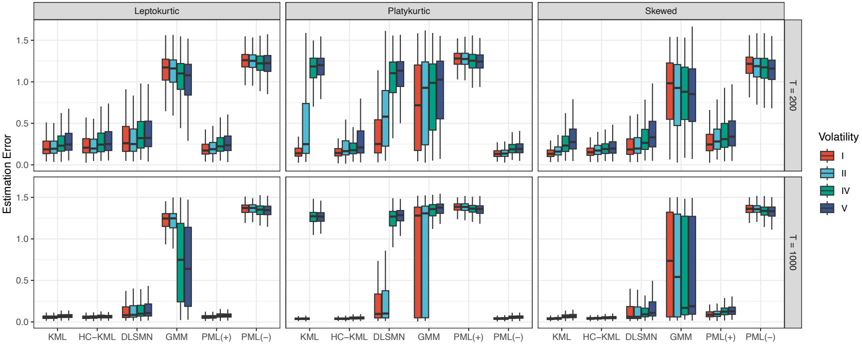
Data generating process

- stable VAR(1) with some persistence (An and Schorfheide 2007, Lütkepohl and Schlaak 2022)

$$y_t - \begin{bmatrix} 0.79 & 0.00 & 0.25 \\ 0.19 & 0.95 & -0.46 \\ 0.12 & 0.00 & 0.62 \end{bmatrix} y_{t-1} = B\Sigma_t^{1/2}\eta_t,$$

- roots $A(z) = 0$: 1.05, 1.11 and 1.95
- $B = \Omega^{1/2}Q(\theta)$, both correlation $\Omega^{1/2}$ and rotation matrices are randomly sampled
- $\Sigma_t = \sigma_t I_N$: common variance, four scenarios

Finite-sample properties of HC-KML (estimated processes)



Content

A small scale US monetary policy model

Empirical model: VARX(4)

- three endogenous: $y_t = ($ output gap$_t$, inflation of GPD depflator$_t$, 1Y T-bill rate$_t)'$, one exogenous: real price of crude oil
- Quarterly observations, 1975Qu4 – 2019Qu4, $T = 173$
- Implementation of HC-KML estimation:
  $K_f(x) = (2\pi)^{-1/2} \exp(-x^2/2)$, $h_f = 1.06\,T^{1/5}$,
  $K_\sigma(x) = 3/4(1 - x^2)\delta_{|x| \leq 1}$, $h_\sigma = 8 \approx 0.6\,T^{1/2}$
  results are robust for alternative kernel and bandwidths

$$\widehat{B} = \begin{bmatrix} \underset{(7.182)}{0.571} & \underset{(-0.607)}{-0.107} & \underset{(0.261)}{0.037} \\ \underset{(0.879)}{0.206} & \underset{(6.549)}{0.765} & \underset{(-0.017)}{-0.003} \\ \underset{(1.073)}{0.16} & \underset{(0.175)}{0.019} & \underset{(4.517)}{0.592} \end{bmatrix}$$

A small scale US monetary policy model

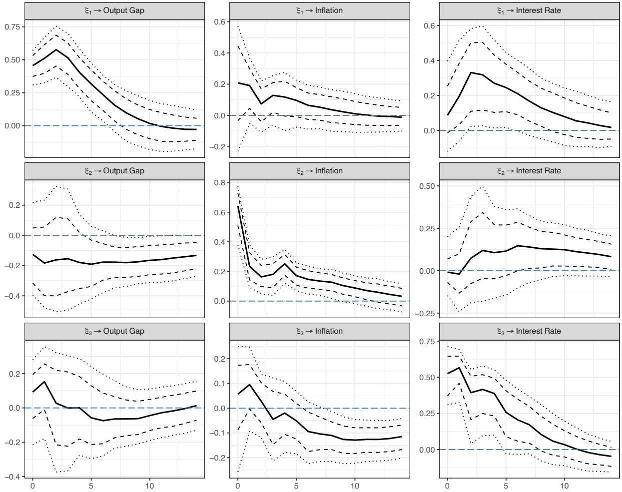


Figure 3: IRFs. CI from a moving block bootstrap with 1,000 reps.

A small scale US monetary policy model

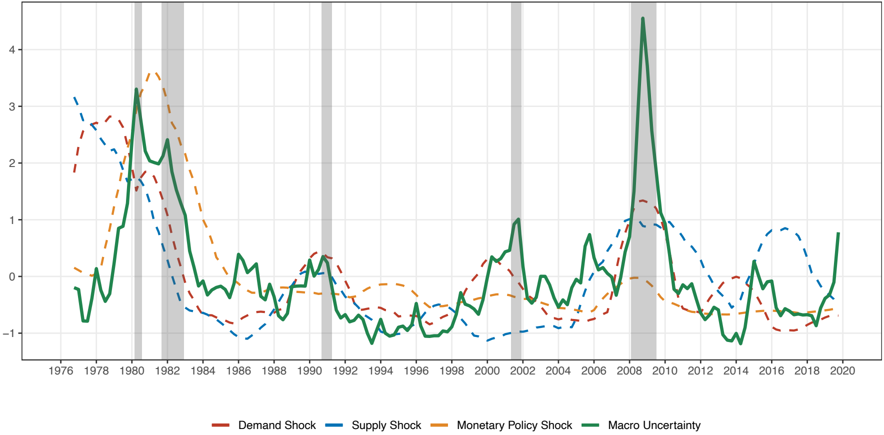


Figure 4: Variance estimates based on $h_\sigma = 8$ (dashed) compared to the macro uncertainty estimates of Jurado et al. (2015) (solid).

Content

Conclusion

In this paper,

1. we introduce a novel heteroskedasticity-consistent kernel-based maximum likelihood estimator (HC-KML) for independence-based identification of simultaneous equations models
2. we study its asymptotic and finite-sample properties:
   i) HC-KML is consistent under mild condition regardless of true shock distributions, permitting (co-)heteroskedastic shocks
   ii) HC-KML has favorable small-sample properties, robust under different shock distributions and volatility processes (incl. homoskedasticity)
3. we apply the estimator to a small scale US MP model, identify shocks with sound economic interpretations, and find co-movement in the shock volatilities

Thank you!

Contact:

Shu Wang,

e-mail: shu.wang@uni-goettingen.de

Chair of Econometrics

University of Göttingen

Humboldtallee 3, D-37073 Göttingen, Germany.

References

An, S. and F. Schorfheide (2007): "Bayesian analysis of DSGE models," *Econometric Reviews*, 26, 113–172.

Drautzburg, T. and J. H. Wright (2023): "Refining set-identification in VARs through independence," *Journal of Econometrics*, 235, 1827–1847.

Fiorentini, G. and E. Sentana (2023): "Discrete mixtures of normals pseudo maximum likelihood estimators of structural vector autoregressions," *Journal of Econometrics*, 235, 643–665.

Gouriéroux, C., A. Monfort, and J. Renne (2017): "Statistical inference for independent component analysis: Application to structural VAR models," *Journal of Econometrics*, 196, 111–126.

Guay, A. (2021): "Identification of structural vector autoregressions through higher unconditional moments," *Journal of Econometrics*, 225, 27–46, themed Issue: Vector Autoregressions.

Hafner, C., H. Herwartz, and S. Wang (2024): "Statistical identification of independent shocks with kernel-based maximum likelihood estimation and an application to the global crude oil market," *Journal of Business & Economic Statistics*.

Hansen, B. E. (1995): "Regression with nonstationary volatility," *Econometrica*, 63, 1113–1132.

Herwartz, H. (2018): "HodgesLehmann Detection of Structural Shocks An Analysis of Macroeconomic Dynamics in the Euro Area," *Oxford Bulletin of Economics and Statistics*, 80, 736–754.

Herwartz, H. and S. Wang (2023): "Point estimation in sign-restricted SVARs based on independence criteria with an application to rational bubbles," *Journal of Economic Dynamics and Control*, 151, 104630.

——— (2024): "Statistical identification in panel structural vector autoregressive models based on independence criteria," *Journal of Applied Econometrics*.

Jarociński, M. (2024): "Estimating the Feds unconventional policy shocks," *Journal of Monetary Economics*, 103548.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring uncertainty," *American Economic Review*, 105, 1177–1216.

Keweloh, S. A. (2021): "A generalized method of moments estimator for structural vector autoregressions based on higher moments," *Journal of Business & Economic Statistics*, 39, 772–782.

Lanne, M. and J. Luoto (2021): "GMM estimation of non-Gaussian structural vector autoregression," *Journal of Business & Economic Statistics*, 39, 69–81.

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

Lewis, D. J. (2024): "Identification based on higher moments," cemmap working paper CWP03/24, London.

Lütkepohl, H. and T. Schlaak (2022): "Heteroscedastic proxy vector autoregressions," *Journal of Business & Economic Statistics*, 40, 1268–1281.

Mesters, G. and P. Zwiernik (2022): "Non-independent components analysis," Working Papers 1358, Barcelona School of Economics.