SFB 823

A feasible approach to incorporate information in higher moments in structural vector autoregressions

Sascha Alexander Keweloh

Nr. 22/2021

Discussion Paper

A Feasible Approach to Incorporate Information in Higher Moments in Structural Vector Autoregressions.

Sascha Alexander Keweloh*
TU Dortmund University

October 19, 2021

Generalized method of moments and continuous updating estimators based on second- to fourth-order moment conditions can be used to solve the identification problem and estimate non-Gaussian structural vectorautoregressions. However, estimating the asymptotically optimal weighting matrix and the asymptotic variance of the estimators is challenging in small samples. I show that this can lead to a severe bias, large variance, and inaccurate inference in small samples. I propose to use the assumption of independent structural shocks not only to derive moment conditions but also to derive alternative estimators for the asymptotically optimal weighting matrix and the asymptotic variance of the estimator. I demonstrate that these estimators greatly improve the performance of the generalized method of moments and continuous updating estimators in terms of bias, variance, and inference.

JEL Codes: C12, C32, C51

Keywords: structural vector autoregression, higher moments, non-Gaussian, independent, GMM

---

* TU Dortmund University / RGS Econ, E-Mail: sascha.keweloh@tu-dortmund.de, Postal Address: TU Dortmund, Department of Economics, D-44221 Dortmund. Financial support from the German Science Foundation, DFG - SFB 823, is gratefully acknowledged.

1 Introduction

In a non-Gaussian structural vectorautoregression (SVAR) independent structural shocks imply higher-order moment conditions which identify the simultaneous relationship without any restrictions on the simultaneous interaction. These higher-order moment conditions can be used to estimate the SVAR with a generalized of moments (GMM) or continuous updating estimator (CUE), see, e.g., Lanne and Luoto (2021), Keweloh (2021), or Guay (2021). However, with higher-order moment conditions the long-run covariance matrix of the sample average of the moment conditions is difficult to estimate in small samples. Nevertheless, an accurate estimation of the covariance matrix is crucial for the estimation of the asymptotically optimal weighting matrix, the estimation of the asymptotic variance, and inference.

This study analyzes the small sample behavior of CUE and GMM estimators with higher-order moment conditions in SVAR models. I find that standard approaches to estimate the long-run covariance matrix lead to volatile and biased CUE and GMM estimators with distorted J and Wald test statistics. Moreover, the performance of the estimators decreases with the model size, to the point of limiting the usefulness of the approach for specifications usually considered in macroeconometrics. I propose to use the assumption of mutually independent structural shocks not only to derive moment conditions but also to estimate the asymptotically efficient weighting matrix and the asymptotic variance. I demonstrate that this simple modification substantially increases the small sample performance of the estimators.

The small sample behavior of CUE and GMM estimators in general has been studied extensively. GMM estimators are known to exhibit a small sample bias and the CUE is associated with a smaller bias, see, e.g., Hansen et al. (1996), Donald and Newey (2000), Han and Phillips (2006), or Newey and Windmeijer (2009). Moreover, the inability to precisely estimate the asymptotic variance leads to oversized Wald test statistics, see Burnside and Eichenbaum (1996). Therefore, Burnside and Eichenbaum (1996) propose to use restrictions implied by the underlying model to calculate test statistics. In the context of SVAR models, Bonhomme and Robin (2009), Keweloh

(2021), and Guay (2021) recognize that due to higher-order moment conditions, the long-run covariance matrix of the sample average of the moment conditions is particularly difficult to estimate. For example, the covariance of cokurtosis moment conditions is of order eight and therefore, difficult to estimate in samples with a few hundred observations. I show that exploiting the assumption of mutually independent structural shocks simplifies the problem of estimating the covariance of higher-order moment conditions. In particular, with independent structural shocks, higher-order moments of the covariance matrix can be calculated as products of lower-order moments. Therefore, I propose the SVAR CUE-MI and SVAR GMM-MI estimators, which are GMM and CUE estimators exploiting the assumption of mutually independent structural shocks to estimate the asymptotically optimal weighting matrix and the asymptotic variance. A Monte Carlo simulation demonstrates that the SVAR CUE-MI and SVAR GMM-MI outperform SVAR CUE and SVAR GMM estimators, which are not exploiting the assumption of mutually independent shocks to estimate the optimal weighting and asymptotic variance.

It is well known that the number coskewness and cokurtosis conditions implied by independent structural shocks increases quickly with the dimension of the SVAR. For example, with $n = 2$ variables, independent structural shocks imply two variance, one covariance, two coskewness and three cokurtosis conditions, and with $n = 4$ variables independent structural shocks imply four variance, six covariance, 16 coskewness and 31 cokurtosis conditions, see Keweloh (2021).[1] While the possible number of moment conditions increases quickly with the dimension of the SVAR, the number of moments contained in the covariance matrix of the moment conditions increases even more rapidly. In particular, for $n = 2$ variables the covariance matrix of all second- to fourth-order moment conditions implied by mutually independent shocks is a 8×8 matrix with five co-moments of order four, six co-moments of order five, seven co-moments of order six, eight co-moments of order seven, and nine co-moments of order eight. However, for $n = 4$ variables the covariance matrix of all second- to fourth-order moment conditions implied by mutually independent shocks

[1] Lanne and Luoto (2021) and Keweloh (2021) propose GMM estimators which minimize the second- and higher-order dependencies of the unmixed innovations. In contrast to that, the GMM estimator proposed by Guay (2021) minimizes the distance of the second- and higher-order co-moments of the reduced form shocks to the second- and higher-order co-moments implied by a mixture of independent structural shocks. The GMM estimator proposed by Guay (2021) has even more higher-order moment conditions.

is a 57×57 matrix with 35 co-moments of order four, 56 co-moments of order five, 84 co-moments of order six, 120 co-moments of order seven, and 165 co-moments of order eight. Exploiting the assumption of mutually independent shocks allows to estimate the covariance matrix of all second- to fourth-order moment conditions implied by mutually independent shocks as a product of n moments of order one, n moments of order two, n moments of order three, n moments of order four, n moments of order five, and n moments of order six. Therefore, with mutually independent shocks the researcher only needs to estimate moments up to order six instead of order eight and the number of these moments increases linearly in the dimension of the SVAR.

Mutually independent structural shocks simplify the estimation of the asymptotically optimal weighting matrix and the asymptotic variance. In many cases, mutually independent structural shocks are no additional assumption but assumed anyway to derive the identifying higher-order moment conditions, see, e.g., Keweloh (2021) and Guay (2021).[2] However, some authors argue that the assumption of mutually independent shocks is too strong, see, e.g., Kilian and Lütkepohl (2017, Chapter 14), Lewis (2021), or Lanne and Luoto (2021). In particular, independence implies that the volatility processes of the shocks are independent. Lanne and Luoto (2021) show that a suitable subset of $n(n-1)/2$ asymmetric cokurtosis conditions is sufficient to ensure local identification in a non-Gaussian SVAR.[3] These asymmetric cokurtosis conditions can be motivated by mutually mean independent shocks, which allows a dependence of the volatility processes. Therefore, I derive analogous results for the estimation of the weighting matrix and variance depending on mutually mean independent shocks. More generally, the approach proposed in this study does not rely on a specific set of moment conditions. Instead, I argue that the same statistical properties used to derive the moment conditions should be used to estimate the asymptotically optimal weighting matrix and the asymptotic variance of the estimator.

The remainder of this article is organized as follows. Section 2 summarizes the SVAR model and

[2]Note that also the (pseudo) maximum-likelihood estimators proposed by Lanne et al. (2017) and Gouriéroux et al. (2017) or the Bayesian approaches proposed by Lanne and Luoto (2020) and Anttonen et al. (2021) assume independent shocks to ensure identification.

[3]Additionally, Lanne et al. (2021) show that the second-order moment conditions together with $n(n-1)/2$ symmetric cokurtosis conditions can be sufficient to ensure global identification. However, the result only holds if the structural shocks satisfy all cokurtosis moment conditions implied by independent shocks.

the main assumptions. Section 3 defines the GMM estimator for SVAR models based on higher-order moment conditions. Section 4 proposes novel estimators for the long-run covariance matrix and the asymptotic variance by exploiting mutually independent shocks. Section 5 demonstrates the advantages of the proposed estimators over traditional estimators in a Monte Carlo simulation. Section 6 concludes.

2 SVAR models

This section briefly explains the identification problem and common identification approaches of SVAR models. A detailed overview can be found in Kilian and Lütkepohl (2017). Consider the SVAR $y_t = \sum_{p=1}^{P} A_p y_{t-p} + u_t$ with an n-dimensional vector of observable variables $y_t = [y_{1,t}, ..., y_{n,t}]'$, the reduced form shocks $u_t = [u_{1,t}, ..., u_{n,t}]'$, and

$$u_t = B_0 \epsilon_t \tag{1}$$

describing the impact of an n-dimensional vector of unknown structural shocks $\epsilon_t = [\epsilon_{1,t}, ..., \epsilon_{n,t}]'$. The matrix $B_0 \in \mathbb{R}^{n \times n}$ governs the simultaneous interaction and is assumed to be invertible.

Assumption 1. $B_0 \in \mathbb{B} := \{B \in \mathbb{R}^{n \times n} | det(B) \neq 0\}$.

The reduced form shocks can be estimated consistently, and for the sake of simplicity, I focus on the simultaneous interaction in Equation (1) and treat the reduced form shocks as observable random variables. The identically distributed structural shocks satisfy the following assumptions.

Assumption 2. *ϵ_t is serially independent (ϵ_t is independent of $\epsilon_{\tilde{t}}$ for $t \neq \tilde{t}$)*

Assumption 3. *ϵ_t has mutually uncorrelated components ($\epsilon_{i,t}$ is uncorrelated with $\epsilon_{j,t}$ for $i \neq j$).*

Assumption 4. *Each component of ϵ_t has zero mean, unit variance, and finite third- and fourth-order moments.*

Based on the assumptions used so far, neither the matrix B_0 nor the structural shocks ϵ_t are identified. Several identifying assumptions have been proposed in the literature ranging from short- or long-run restrictions on the interaction of the variables (see, e.g., Sims (1980), or Blanchard (1989)), over Proxy-SVAR models (see, e.g., Mertens and Ravn (2014)) up to sign restrictions (see, e.g., Uhlig (2005) or Peersman (2005)). A novel branch of the SVAR identification literature uses non-Gaussian and independent shocks to identify the SVAR, see, e.g., Lanne et al. (2017), Gouriéroux et al. (2017), Herwartz (2018), Lanne and Luoto (2021), Keweloh (2021), or Guay (2021)). These data driven identification schemes do not require to impose any short- or long-run restrictions on the interaction of the variables, instead, identification is based on statistical properties of the shocks. The most commonly used statistical property is the assumption of mutually independent structural shocks.

Assumption 5. *ϵ_t has mutually independent components ($\epsilon_{i,t}$ is independent of $\epsilon_{j,t}$ for $i \neq j$).*

The independence assumption can be used to derive moment conditions, see, e.g., Keweloh (2021) or Guay (2021). However, the assumption of mutually independent shocks has been criticized by several authors for being too restrictive, see, e.g., Kilian and Lütkepohl (2017, Chapter 14), Lewis (2021), or Lanne and Luoto (2021). In particular, it appears plausible that multiple macroeconomic shocks are driven by the same volatility process, which is not possible with mutually independent shocks. The independence assumption can be relaxed to the weaker assumption of mutually mean independent shocks.

Assumption 6. *ϵ_t has mutually mean independent components ($E\left[\epsilon_{i,t}|\epsilon_{-i,t}\right] = 0$ for $i = 1, ..., n$).*

If the shocks are mutually mean independent, no shock contains any information on the mean of another shock, however, shocks can contain information on the variance of other shocks and thus, the assumption may be more plausible in some applications. Note that mutually mean independent structural shocks are sufficient to derive the identifying coskewness conditions proposed in Keweloh (2021) and the identifying asymmetric cokurtosis conditions used in Lanne and Luoto (2021). For the sake of simplicity, this study focuses on mutually independent shocks. Results under the weaker assumption of mutually mean independent shocks can be obtained analogously and are briefly sketched.

The assumption of mutually (mean) independent shocks can be used to generate higher-order moment condition, however, these conditions are only informative if the structural shocks are non-Gaussian. Therefore, identification requires non-Gaussian shocks embedded in the following assumption.

Assumption 7. *At most one component of ϵ_t is Gaussian.*

Note that depending on the particular identification approach a slight modification of Assumption 7 is required. For example, the GMM estimators proposed by Lanne and Luoto (2021), Keweloh (2021) or Guay (2021) require that the third- and/or fourth-order moments of at most one structural shock is equal to the corresponding moment of a Gaussian shock.

3 SVAR GMM with higher-order moment conditions

This section briefly summarizes the SVAR estimators based on higher-order moment conditions derived from mutually independent structural shocks. A detailed description can be found in Lanne and Luoto (2021), Keweloh (2021), or Guay (2021).

The reduced form shocks are equal to an unknown mixture of the unknown structural shocks, $u_t = B\epsilon_t$. Reversing this relationship yields the unmixed innovations $e(B)_t$, defined as the innovations obtained by unmixing the reduced form shocks with some invertible matrix B

$$e(B)_t := B^{-1}u_t. \tag{2}$$

If B is equal to the true mixing matrix B_0, the unmixed innovations are equal to the structural shocks. Assumption 5 and 4 can be used to derive moment conditions. In particular, the structural shocks are uncorrelated with unit variance and therefore, the unmixing matrix B should yield uncorrelated unmixed innovations with unit variance, see Table 1. Moreover, independent structural shocks yield coskewness or third-order moment conditions and cokurtosis or fourth-order moment conditions, see Table 1.

Table 1: Illustration of moment conditions

covariance / second-order conditions			coskewness / third-order conditions		
$E[\epsilon_{1,t}^2] = 1$	$\Longrightarrow$	$E[e(B)_{1,t}^2] \overset{!}{=} 1$	$E[\epsilon_{1,t}^2 \epsilon_{2,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t}^2 e(B)_{2,t}] \overset{!}{=} 0$
	$\vdots$		$E[\epsilon_{1,t} \epsilon_{2,t}^2] = 0$	$\Longrightarrow$	$E[e(B)_{1,t} e(B)_{2,t}^2] \overset{!}{=} 0$
$E[\epsilon_{n,t}^2] = 1$	$\Longrightarrow$	$E[e(B)_{n,t}^2] \overset{!}{=} 1$	$E[\epsilon_{1,t} \epsilon_{2,t} \epsilon_{3,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t} e(B)_{2,t} e(B)_{3,t}] \overset{!}{=} 0$
$E[\epsilon_{1,t} \epsilon_{2,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t} e(B)_{2,t}] \overset{!}{=} 0$		$\vdots$	
	$\vdots$				

cokurtosis / fourth-order conditions		
$E[\epsilon_{1,t}^3 \epsilon_{2,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t}^3 e(B)_{2,t}] \overset{!}{=} 0$
$E[\epsilon_{1,t}^2 \epsilon_{2,t}^2] = 1$	$\Longrightarrow$	$E[e(B)_{1,t}^2 e(B)_{2,t}^2] \overset{!}{=} 1$
$E[\epsilon_{1,t} \epsilon_{2,t}^3] = 0$	$\Longrightarrow$	$E[e(B)_{1,t} e(B)_{2,t}^3] \overset{!}{=} 0$
$E[\epsilon_{1,t}^2 \epsilon_{2,t} \epsilon_{3,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t}^2 e(B)_{2,t} e(B)_{3,t}] \overset{!}{=} 0$
$E[\epsilon_{1,t} \epsilon_{2,t} \epsilon_{3,t} \epsilon_{4,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t}^2 e(B)_{2,t} e(B)_{3,t} e(B)_{4,t}] \overset{!}{=} 0$
	$\vdots$	

In general, all variance, covariance, coskewness, and cokurtosis moment conditions derived from independent structural shocks embedded in Assumption 5 can be written as

$$E[f_m(B, u_t)] = 0 \quad \text{with} \quad f_m(B, u_t) := \prod_{i=1}^{n} e(B)_{i,t}^{m_i} - m_0, \tag{3}$$

where $f_m(B, u_t)$ contains all variance and covariance conditions for $m \in \mathbf{2}$, all coskewness conditions for $m \in \mathbf{3}$, and all cokurtosis conditions for $m \in \mathbf{4}$ with

$$\mathbf{2} := \{[m_0, m_1, ...m_n] \in \{\{0,1\}, \{0,1,2\}^n\} | \sum_{i=1}^{n} m_i = 2 \text{ and} \tag{4}$$

$$m_0 = \begin{cases} 0, & \text{if } \exists m_i = 1 \text{ for } m_i \in m_1, ...m_n \\ 1, & \text{otherwise} \end{cases} \},$$

$$\mathbf{3} := \{[m_0, m_1, ...m_n] \in \{\{0,1\}, \{0,1,2\}^n\} | \sum_{i=1}^{n} m_i = 3 \text{ and } m_0 = 0\}, \tag{5}$$

$$\mathbf{4} := \{[m_0, m_1, ...m_n] \in \{\{0,1\}, \{0,1,2,3\}^n\} | \sum_{i=1}^{n} m_i = 4 \text{ and} \tag{6}$$

$$m_0 = \begin{cases} 0, & \text{if } \exists m_i = 1 \text{ for } m_i \in m_1, ...m_n \\ 1, & \text{otherwise} \end{cases} \}.$$

Note that all moment conditions except the symmetric cokurtosis conditions $E[\epsilon_{i,t}^2 \epsilon_{j,t}^2] = 1$ for $i \neq j$ can be derived from Assumption 6 and therefore, only require mutually mean independent shocks.

Based on all or a subset of the moment conditions presented above, Lanne and Luoto (2021) and Keweloh (2021) provide different local and global identification results for the following SVAR GMM estimator

$$\hat{B}_T := \arg\min_{B \in \mathbb{B}} g_T(B)' W g_T(B), \tag{7}$$

where $g_T(B) = \frac{1}{T} \sum_{t=1}^{T} f(B, u_t)$, and W is a positive semi-definite weighting matrix. Suppose that $f(B, u_t)$ contains all or a subset of the moment conditions $f_m(B, u_t)$ with $m \in \mathbf{2} \cup \mathbf{3} \cup \mathbf{4}$ sucht that the GMM estimator (7) is identified. Consistency and asymptotic normality of the estimator follow from standard assumptions

$$\begin{aligned} \hat{B}_T &\xrightarrow{p} B_0 \\ \sqrt{T}(\hat{B}_T - B_0) &\xrightarrow{d} \mathcal{N}(0, MSM') \end{aligned} \quad \text{with} \quad \begin{aligned} M &:= \left(G'S^{-1}G\right)^{-1} G'W \\ G &:= E\left[\frac{\partial f(B_0,u_t)}{\partial vec(B)'}\right] \\ S &:= \lim_{T\to\infty} E\left[T g_T(B_0) g_T(B_0)'\right], \end{aligned} \tag{8}$$

see Hall et al. (2005). In particular, asymptotic normality requires that the matrix S exists and is finite. For the SVAR GMM estimator based on second- to fourth-order moment conditions this holds if ϵ_t has finite moments up to order eight. The weighting matrix $W^* := S^{-1}$ leads to the estimator $\hat{B}_T^*$ with the asymptotic variance $\sqrt{T}(\hat{B}_T^* - B_0) \xrightarrow{d} \mathcal{N}(0, (G'S^{-1}G)^{-1})$, which is the lowest possible asymptotic variance, see Hall et al. (2005). Han and Phillips (2006) proposed the continuous updating estimator estimator (CUE)

$$\hat{B}_T := \arg\min_{B\in\mathbb{B}} g_T(B)'\hat{W}(B)g_T(B), \tag{9}$$

where $\hat{W}(B)$ is a consistent estimator for the asymptotically optimal weighting matrix W^*.

Han and Phillips (2006) and Newey and Windmeijer (2009) show that for i.i.d. observations and a nonrandom weighting matrix W the expected value of a GMM objective function is equal to

$$E\left[g_T(B)'Wg_T(B)\right] = E\left[\sum_{t\neq\tilde{t}} f(B,u_t)'Wf(B,u_{\tilde{t}})\right] + E\left[\sum_t f(B,u_t)'Wf(B,u_t)\right] \tag{10}$$

$$= (1-T^{-1})E[f(B,u_t)]'WE[f(B,u_t)] + trace(WS(B))/T, \tag{11}$$

where the second equality uses the nonrandom weighting matrix W and $S(B) := E[f(B,u_t)f(B,u_t)']$. The first term in Equation (11) is called signal term and is minimized at B_0 since $E\left[f(B_0,u_t)\right] = 0$. The second term in Equation (11) is called noise term and is not minimized at B_0. The impact of the noise term vanishes with $T\to\infty$. Nevertheless, in a finite sample the noise term can dominate the signal term and lead to a bias, especially in large SVAR models with many moment conditions. If the weighting matrix $W(B)$ is equal to $S(B)^{-1}$, the noise term in Equation (11) collapses to m/T, where m is equal to the number of moment conditions. Therefore, the noise term no longer depends on B and hence leads to no bias. The CUE is a feasible version of this approach and replaces $W(B) = S(B)^{-1}$ with some estimator $\hat{W}(B) = \hat{S}(B)^{-1}$.

4 Estimating S and G

In practice, S and G are unknown and need to be estimated for inference and asymptotically optimal weighting. These matrices can be difficult to estimate in small samples. In a GMM setup not related to SVAR models Burnside and Eichenbaum (1996) propose to impose restrictions of the underlying economic model on the estimator for S and G. They show that exploiting additional information on S and G can largely improve the rejection rates of Wald tests in small samples. This section shows how the structure of the SVAR can be used to improve the estimation of S and G. In particular, I propose to exploit the assumption of serially and mutually independent structural shocks. In the SVAR with second- to fourth-order moment conditions, estimation of the long-run covariance matrix S is particularly difficult, since it requires to estimate moments up to order eight. Bonhomme and Robin (2009), Keweloh (2021), and Guay (2021) recognize that the presence of these higher-order moments makes it difficult to estimate the asymptotically optimal weighting matrix and the asymptotic variance of the estimator. I show that exploiting the assumption of serially and mutually independent shocks largely simplifies the estimation of the long-run covariance matrix S and yields more precise estimates of the asymptotically optimal weighting matrix and the asymptotic variance. Additionally, for the first step GMM estimator, I propose an approximation of the asymptotically optimal weighting matrix based on the assumption of mutually independent shocks not requiring any prior estimates of B_0.

In the SVAR, the long-run covariance matrix of two arbitrary moment conditions $f_m(B, u_t) = \prod_{i=1}^{n} e(B)_{i,t}^{m_i} - m_0$ and $f_{\tilde{m}}(B, u_t) = \prod_{i=1}^{n} e(B)_{i,t}^{\tilde{m}_i} - \tilde{m}_0$ with $m, \tilde{m} \in \mathbf{2} \cup \mathbf{3} \cup \mathbf{4}$ at $B = B_0$ is equal to

$$S_{m,\tilde{m}} := \lim_{T\to\infty} E\left[T\left(\frac{1}{T}\sum_{t=1}^{T} f_m(B_0,u_t)\right)\left(\frac{1}{T}\sum_{t=1}^{T} f_{\tilde{m}}(B_0,u_t)\right)\right] \tag{12}$$

$$= \lim_{T\to\infty} E\left[T\left(\frac{1}{T}\sum_{t=1}^{T}\prod_{i=1}^{n} e(B_0)_{i,t}^{m_i} - m_0\right)\left(\frac{1}{T}\sum_{t=1}^{T}\prod_{i=1}^{n} e(B_0)_{i,t}^{\tilde{m}_i} - \tilde{m}_0\right)\right] \tag{13}$$

$$\begin{aligned} =&E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i+\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i}\right] + m_0\tilde{m}_0 \\ &+\sum_{j=1}^{\infty} E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i}\epsilon_{i,t-j}^{\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i}\right] + m_0\tilde{m}_0 \\ &+\sum_{j=1}^{\infty} E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{m_i}\epsilon_{i,t}^{\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{m_i}\right] + m_0\tilde{m}_0, \end{aligned} \tag{14}$$

where the last equality follows from identically distributed shocks and $e(B_0)_t = \epsilon_t$. Therefore, with fourth order moments $m, \tilde{m} \in \mathbf{4}$ such that $\sum_{i=1}^{n} m_i = \sum_{i=1}^{n} \tilde{m}_i = 4$, the long-run covariance matrix $S_{m,\tilde{m}}$ contains co-moments of the structural shocks up to order eight. In practice, $S_{m,\tilde{m}}$ in Equation (14) can be estimated by replacing ϵ_t with $e(B)_t$ and some initial estimate or guess B of B_0 and a heteroscedasticity and autocorrelation consistent covariance (HAC) estimator, see Newey and West (1994).

However, with serially independent structural shocks implied by Assumption 2 the expression of $S_{m,\tilde{m}}$ simplifies to[4]

$$S_{m,\tilde{m}}^{SI} = E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i+\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i}\right] + m_0\tilde{m}_0, \tag{18}$$

[4]To see this note that for $j > 1$

$$E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i}\epsilon_{i,t-j}^{\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{m_i}\right] + m_0\tilde{m}_0 \tag{15}$$

$$= E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i}\right] E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{m_i}\right] + m_0\tilde{m}_0 \tag{16}$$

$$= E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i} - m_0\right] E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{\tilde{m}_i} - \tilde{m}_0\right] = 0, \tag{17}$$

where the first equality follows from serially independent shocks.

where the superscript SI indicates that the equality $S_{m,\tilde{m}} = S^{SI}_{m,\tilde{m}}$ only holds for serially independent shocks. Let S^{SI} denote S under the assumption of serially independent shocks. Based on Equation (18), the long-run covariance under serially independent shocks can be estimated by

$$\hat{S}^{SI}_{m,\tilde{m}}(B) := \frac{1}{T}\sum_{t=1}^{T}\left[\prod_{i=1}^{n} e(B)_{i,t}^{m_i+\tilde{m}_i}\right] - m_0\frac{1}{T}\sum_{t=1}^{T}\left[\prod_{i=1}^{n} e(B)_{i,t}^{\tilde{m}_i}\right] - \tilde{m}_0\frac{1}{T}\sum_{t=1}^{T}\left[\prod_{i=1}^{n} e(B)_{i,t}^{m_i}\right] + m_0\tilde{m}_0, \tag{19}$$

where B is an initial guess or a consistent estimator for B_0. Let $\hat{S}^{SI}(B)$ denote the estimator where each element of the long-run covariance matrix is estimated by Equation (18). Note that serially independent shocks imply that $S^{SI} = E\left[f(B_0,u_t)f(B_0,u_t)'\right]$ and $\hat{S}^{SI}(B) = \frac{1}{T}\sum_{t=1}^{T} f(B,u_t)f(B,u_t)'$, which corresponds to the estimator for S under the frequently used assumption of serially uncorrelated moment conditions.

With serially independent structural shocks the expression of the long-run covariance matrix S simplifies to the covariance matrix S^{SI}. Nevertheless, the covariance matrix $S^{SI}_{m,\tilde{m}}$ of two fourth-order moments $m, \tilde{m} \in \mathbf{4}$ is still of order eight and remains difficult to estimate in small samples. Analogously, one can now exploit that the shocks are mutually independent to further simplify the estimation of S. Note that many non-Gaussian identification approaches rely on the assumption of mutual independent shocks to ensure identification, see, e.g., Lanne et al. (2017), Gouriéroux et al. (2017), or Keweloh (2021). In this case, the researcher already relies on the assumption of mutual independent shocks and may thus as well use it to simplify the estimation of S.

With serially and mutually independent shocks implied by Assumption 2 and 5 the expression of $S_{m,\tilde{m}}$ simplifies to

$$S^{SMI}_{m,\tilde{m}} = \prod_{i=1}^{n} E\left[\epsilon_{i,t}^{m_i+\tilde{m}_i}\right] - m_0\prod_{i=1}^{n} E\left[\epsilon_{i,t}^{\tilde{m}_i}\right] - \tilde{m}_0\prod_{i=1}^{n} E\left[\epsilon_{i,t}^{m_i}\right] + m_0\tilde{m}_0, \tag{20}$$

where the superscript SMI indicates that the equality $S_{m,\tilde{m}} = S^{SMI}_{m,\tilde{m}}$ only holds for serially and mutually independent shocks. Let S^{SMI} denote S under the assumption of serially and mutually independent shocks. Based on Equation (20), the long-run covariance under serially and mutually independent shocks can be estimated by

$$\hat{S}^{SMI}_{m,\tilde{m}}(B) := \prod_{i=1}^{n} \frac{1}{T} \sum_{t=1}^{T} \left[e(B)_{i,t}^{m_i+\tilde{m}_i} \right] - m_0 \prod_{i=1}^{n} \frac{1}{T} \sum_{t=1}^{T} \left[e(B)_{i,t}^{\tilde{m}_i} \right] - \tilde{m}_0 \prod_{i=1}^{n} \frac{1}{T} \sum_{t=1}^{T} \left[e(B)_{i,t}^{m_i} \right] + m_0 \tilde{m}_0, \tag{21}$$

where B is an initial guess or a consistent estimator for B_0. [5] Let $\hat{S}^{SMI}(B)$ denote the estimator where each element of the long-run covariance matrix is estimated by Equation (21).

Exploiting mutually independent shocks allows to transform higher-order co-moments into a product of lower-order moments. For example, consider the two moment conditions $E[\epsilon_{1,t}^3 \epsilon_{2,t}] = 0$ and $E[\epsilon_{3,t}^3 \epsilon_{4,t}] = 0$, such that the covariance of both moment conditions is equal to $E[\epsilon_{1,t}^3 \epsilon_{2,t} \epsilon_{3,t}^3 \epsilon_{4,t}]$, a co-moment of order eight. However, with mutually independent shocks the covariance is equal to $E[\epsilon_{1,t}^3]E[\epsilon_{2,t}]E[\epsilon_{3,t}^3]E[\epsilon_{4,t}]$, a product of moments of order one and three. In general, the covariance matrix under serially independent shocks S^{SI} requires to calculate co-moments of ϵ_t of order four to eight and the covariance matrix under serially and mutually independent shocks S^{SMI} requires to estimate moments of ϵ_t of order one to six. Table 2 shows the number of co-moments of ϵ_t contained in S^{SI} and S^{SMI} of a SVAR GMM estimator using all second- to

[5] Consistency of $\hat{S}^{SMI}_{m,\tilde{m}}(\hat{B}_T) \xrightarrow{p} S^{SMI}_{m,\tilde{m}}$ for $\hat{B}_T \xrightarrow{p} B_0$ follow as usual from continuity of $e(B)_{i,t}$, consistency of $\hat{B}_T$ which implies

$$P\left(|E\left[e(\hat{B}_T)_{i,t}^s \right] - E\left[\epsilon_{i,t}^s \right] | > \gamma/2 \right) \to 0 \tag{22}$$

and uniform convergence of $\frac{1}{T}\sum_{t=1}^{T} e(B)_{i,t}^s$, such that $\underset{B\in}{SIp} |\frac{1}{T}\sum_{t=1}^{T} e(B)_{i,t}^s - E\left[e(B)_{i,t}^s \right] | \xrightarrow{p} 0$ which implies

$$P\left(|\frac{1}{T}\sum_{t=1}^{T} e(\hat{B}_T)_{i,t}^s - E\left[e(\hat{B}_T)_{i,t}^s \right] | > \gamma/2 \right) \to 0 \tag{23}$$

and therefore

$$P\left(|\frac{1}{T}\sum_{t=1}^{T} e(\hat{B}_T)_{i,t}^s - E\left[\epsilon_{i,t}^s \right] | > \gamma \right) \to 0. \tag{24}$$

Therefore, if the structural shocks are serially and mutually independent, it holds that $S^{SMI}_{m,\tilde{m}} = S_{m,\tilde{m}}$ and hence $\hat{S}^{SMI}_{m,\tilde{m}}(\hat{B}_T)$ is also a consistent estimator for $S_{m,\tilde{m}}$.

fourth- order moment conditions. The number of higher-order co-moments increases quickly with the dimension of ϵ_t. For example, in an SVAR with $n = 2$ variables S^{SI} requires to estimate nine co-moments of order eight, however, in an SVAR with $n = 4$ this number grows to 156, and with $n = 6$ variables S^{SI} require to estimate 1287 co-moments of order eight. In contrast to that, the number of higher-order moments in S^{SMI} grows linearly in n. Therefore, using mutually independent shocks to estimate S appears particularly beneficial in larger SVARs.

Table 2: Number of moments

		$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
Number of GMM moment conditions:	second-order	3	6	10	10	15
	third-order	2	7	16	30	50
	fourth-order	3	12	31	65	120
	S dimension	8×8	25×25	57×57	105×105	185×185
Number of co-moments in S^{SI}:	fourth-order	5	15	35	70	126
	fifth-order	6	21	56	126	252
	sixth-order	7	28	84	210	462
	seventh-order	8	36	120	330	792
	eighth-order	9	45	156	495	1287
Number of moments in S^{SMI}:	first-order	2	3	4	5	6
	second-order	2	3	4	5	6
	third-order	2	3	4	5	6
	fourth-order	2	3	4	5	6
	fifths-order	2	3	4	5	6
	sixth-order	2	3	4	5	6

The table shows the number of GMM moment conditions implied by mutually independent shocks and the number of co-moments of ϵ_t contained in S^{SI} and S^{SMI} in a SVAR with two, three, and four variables.

In this study, I simultaneously use the assumption of serially and mutually independent shocks to estimate S. However, one could also directly exploit mutually independent shocks to simplify $S_{m,\tilde{m}}$ in Equation (14) without assuming serially independent shocks. Moreover, the weaker assumption of mutually mean independent shocks can also be used to simplify the estimation of S. In particular, with serially independent and mutually mean independent shocks implied by

Assumption 2 and 6 the expression of $S_{m,\tilde{m}}$ simplifies to

$$S^{SMMI}_{m,\tilde{m}} = E\left[\prod_{\substack{i=1\\ m_i+\tilde{m}_i\neq 1}}^{n} \epsilon_{i,t}^{m_i+\tilde{m}_i}\right] \prod_{\substack{i=1\\ m_i+\tilde{m}_i=1}}^{n} E\left[\epsilon_{i,t}^{m_i+\tilde{m}_i}\right] - m_0 E\left[\prod_{\substack{i=1\\ \tilde{m}_i\neq 1}}^{n} \epsilon_{i,t}^{\tilde{m}_i}\right] \prod_{\substack{i=1\\ \tilde{m}_i=1}}^{n} E\left[\epsilon_{i,t}^{\tilde{m}_i}\right] \qquad (25)$$

$$-\tilde{m}_0 E\left[\prod_{\substack{i=1\\ m_i\neq 1}}^{n} \epsilon_{i,t}^{m_i}\right] \prod_{\substack{i=1\\ m_i=1}}^{n} E\left[\epsilon_{i,t}^{m_i}\right] + m_0\tilde{m}_0,$$

where the superscript $SMMI$ indicates that the equality $S_{m,\tilde{m}} = S^{SMMI}_{m,\tilde{m}}$ only holds for serially independent and mutually mean independent shocks. For the sake of simplicity, the remainder of the paper focuses on the assumption of mutually independent shocks.

The assumption of mutually independent structural shocks can also be used to estimate G required to estimate the asymptotic variance of the CUE or GMM estimator. For an arbitrary moment condition $f_m(B,u_t) = \prod_{i=1}^{n} e(B)_{i,t}^{m_i} - m_0$ with $m \in \mathbf{2}\cup\mathbf{3}\cup\mathbf{4}$ the derivative with respect to b_{pq} the element at row p and column q of B evaluated at $B = B_0$ corresponds to an element of G and is equal to

$$G_{m,b_{ql}} := E\left[\frac{\partial f_m(B_0,u_t)}{\partial b_{pq}}\right] \qquad (26)$$

$$= \sum_{j=1, j\neq q}^{n} -m_j a_{jp} E\left[\epsilon_{j,t}^{m_j-1}\epsilon_{q,t}^{m_q+1} \prod_{i=1, i\neq j,q}^{n} \epsilon_{i,t}^{m_i}\right] - m_q a_{qp} E\left[\prod_{i=1}^{n} \epsilon_{i,t}^{m_i}\right], \qquad (27)$$

with $A = B_0^{-1}$ and a_{jp} are the elements of A. The equality follows from $e(B_0)_t = \epsilon_t$, the product rule, and $\frac{\partial e(B_0)_{i,t}}{\partial b_{pq}} = -a_{ip}\epsilon_{q,t}$. Again, for mutually independent structural shocks implied by Assumption 5 if follows

$$G^{MI}_{m,b_{ql}} = \sum_{j=1, j\neq q}^{n} -m_j a_{jp} E\left[\epsilon_{j,t}^{m_j-1}\right] E\left[\epsilon_{q,t}^{m_q+1}\right] \prod_{i=1, i\neq j,q}^{n} E\left[\epsilon_{i,t}^{m_i}\right] - m_q a_{qp} \prod_{i=1}^{n} E\left[\epsilon_{i,t}^{m_i}\right], \qquad (28)$$

where the superscript MI indicates that the equality $G_{m,b_{ql}} = G^{MI}_{m,b_{ql}}$ only holds for mutually independent shocks. Let G^{MI} denote G under the assumption of mutually independent shocks and let $\hat{G}$ and $\hat{G}^{MI}$ denote the corresponding estimators. Again, mutual independence allows to calculate higher-order co-moments as a product of lower-order moments. For example, in a SVAR with $n = 4$ variables and a GMM estimator including all 4 variance, 6 covariance, 16 coskewness, and 31 cokurtosis conditions implied by independent shocks with unit variance, the covariance matrix G is a 57×16 dimensional matrix containing 10 co-moments of order two, 16 co-moments of order three, and 31 co-moments of order four. In contrast to that, G^{MI} contains four moments of order one, four moments of order two, four moments of order three, four moments of order four.

The assumption of serially and mutually independent shocks can also be used to derive a guess for the optimal weighting matrix W^* without requiring an initial guess or estimate of the unknown simultaneous interaction B_0. Instead, the researcher can guess the distribution of each structural shock $\epsilon_{i,t}$ for $i = 1, ..., n$ and if the guess is correct Equation (20) directly yields the true covariance matrix S, which can be used to calculate the optimal weighting matrix. In practice, I recommend starting with the assumption of t- or normally distributed shocks to approximate S and hence W^*. I find that even if the initially assumed distributions are incorrect, the corresponding one-step GMM estimator performs similarly in terms of bias and interquartile range to the one-step GMM estimator using the true asymptotically optimal weighting matrix. This might be related to the fact that due to the normalization to mean zero and unit variance shocks, the guess of higher-order moments is irrelevant for many moments. For example the two moment conditions $E[\epsilon_{1,t}^3 \epsilon_{2,t}] = 0$ and $E[\epsilon_{1,t}^3 \epsilon_{3,t}] = 0$ require to estimate $E[\epsilon_{1,t}^6 \epsilon_{2,t} \epsilon_{3,t}]$. However, this co-moment of order eight is equal to zero for all independent shocks with mean zero and finite second to six moments.

5 Monte Carlo Simulation

This section compares the impact of the estimates $\hat{S}^{SI}$, $\hat{S}^{SMI}$, $\hat{G}$, and $\hat{G}^{MI}$ on the finite sample performance of CUE and GMM estimators. I simulate a SVAR $u_t = B_0 \epsilon_t$ with $n = 2$ and $n = 4$

variables with

$$B_0 = \begin{bmatrix} 1 & 0 \\ 0.5 & 1 \end{bmatrix} \text{ and } B_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 1 & 0 & 0 \\ 0.5 & 0.5 & 1 & 0 \\ 0.5 & 0.5 & 0.5 & 1 \end{bmatrix}. \tag{29}$$

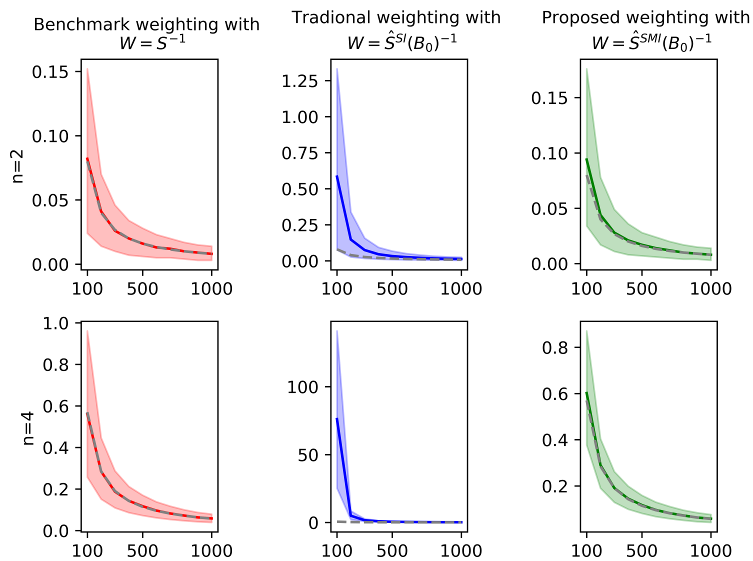
The structural shocks are drawn from a mixture of Gaussian distributions with mean zero, unit variance, skewness equal to 0.89 and an excess kurtosis of 2.35. In particular, the shocks satisfy

$$\varepsilon_i = z\phi_1 + (1-z)\phi_2 \text{ with } \phi_1 \sim \mathcal{N}(-0.2, 0.7),\ \phi_2 \sim \mathcal{N}(0.75, 1.5),\ z \sim \mathcal{B}(0.79), \tag{30}$$

where $\mathcal{B}(p)$ indicates a Bernoulli distribution and $\mathcal{N}(\mu, \sigma^2)$ indicates a normal distribution. Simulations based on t-distributed shocks are shown in the appendix.

Before turning to different CUE and GMM estimators, I analyze the impact of the estimated asymptotically efficient weighting matrix on the GMM loss. Figure 1 compares the average and quantiles of the GMM objective function $g_T(B)'Wg_T(B)$ at $B = B_0$ with all second- to fourth-order moment conditions implied by mutually independent shocks for different weighting matrices. The red loss serves as a benchmark and uses the true but in practice unknown asymptotically efficient weighting matrix, $W^* = S^{-1}$. The blue loss uses the traditional estimator for the asymptotically efficient weighting matrix relying on serially uncorrelated moment conditions equivalent to serially independent shocks, $\hat{W}^{SI} = \hat{S}^{SI}(B_0)^{-1}$. The green loss corresponds to the proposed estimator for the asymptotically efficient weighting matrix using serially and mutually independent shocks, $\hat{W}^{SMI} = \hat{S}^{SMI}(B_0)^{-1}$. The simulation shows that the standard estimator for the asymptotically efficient weighting matrix $\hat{W}^{SI}$ is ill suited to approximate the asymptotically efficient weighting matrix. In the small SVAR with $n = 2$ and the smallest sample size $T = 100$, the average GMM loss based on $\hat{W}^{SI}$ is approximately seven times larger than the average GMM loss based on the asymptotically efficient weighting matrix W^* and in the large

Figure 1: GMM loss at B_0 for different weighting schemes



Average, 10%, and 90% quantiles of the GMM loss $g_T(B)'Wg_T(B)$ with all second-, third-, and fourth-order moment conditions implied by mutually independent shocks at $B = B_0$ for $W = S^{-1}$ in red, $W = \hat{S}_T^{SI}(B_0)^{-1}$ in blue, and $W = \hat{S}_T^{SMI}(B_0)^{-1}$ in green with 5000 simulations and sample sizes $T = 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000$. The dotted gray line shows the expected value of the GMM objective function at B_0 and $W = S^{-1}$ which is equal to m/T where m denotes the number of moment conditions, compare Han and Phillips (2006) and Newey and Windmeijer (2009).

SVAR with $n = 4$ the average loss is approximately 132 times larger. In contrast to that, the weighting scheme proposed in this study, $\hat{W}^{SMI}$, which exploits the mutual independence of the structural shocks closely approximates the infeasible asymptotically efficient weighting scheme with $W^* = S^{-1}$.

In the following, I analyze the impact of the weighting scheme on the following asymptotically efficient estimators:

- GMM*: A one-step GMM estimator with the asymptotically efficient weighting matrix $W = S^{-1}$.
- GMM: A two-step GMM estimator with $W = I$ in the first step and $W = \hat{S}^{SI}(\hat{B})^{-1}$ in the second step.
- CUE: A continuous updating estimator with $W(B) = \hat{S}^{SI}(B)^{-1}$.
- GMM-MI: A two-step GMM estimator with $W = S_{Norm}^{SMI^{-1}}$ in the first step, $W = \hat{S}^{SI}(\hat{B})^{-1}$ in the second step and S_{Norm}^{SMI} denotes the long-run covariance matrix under serially and mutually independent Gaussian shocks.
- CUE-MI: A continuous updating estimator with $W(B) = \hat{S}^{SMI}(B)^{-1}$.

The estimator GMM* is infeasible since it uses the unknown asymptotically efficient weighting matrix $W = S^{-1}$ and it serves as a benchmark. The CUE and GMM estimators only use the assumption of serially independent shocks or serially uncorrelated moment conditions to estimate S and therefore, represent the traditional estimation approaches. The CUE-MI and GMM-MI estimators are the novel estimators proposed in this study and rely on the assumption of serially and mutually independent shocks to estimate S. All estimators use all second-, third-, and fourth-order moment conditions implied by mutually independent shocks. In particular, for $n = 2$ the estimators use three second-, two third-, and three fourth-order moment conditions and for $n = 4$ the estimators use 10 second-, 16 third-, and 31 fourth-order moment conditions. Note that all analyzed estimators have the same asymptotic variance and are asymptotically efficient. The appendix contains further results for the one-step GMM estimator using the identity weighting matrix $W = I$, the approximation of the asymptotically efficient weighting matrix based on serially and mutually independent Gaussian shocks $W = S_{Norm}^{SMI^{-1}}$, and the white fast weighting matrix proposed in Keweloh (2021).

Firstly, I analyze the impact of the weighting scheme on the finite sample bias and the interquartile range. The elements of B are equal on the diagonal, upper- and lower- triangular. Therefore, I summarizes the results for all elements on the diagonal, upper- and lower- triangular of B. Figure 2 shows the average of the median absolute bias and Figure 3 shows the average of the interquartile range (IQR) of the estimated elements on the diagonal, upper- and lower- triangular matrix $\hat{B}$. The results for all individual elements of B for $T = 100$ and $T = 1000$ are shown in the appendix. The computation of the average of the median absolute bias and the average of the interquartile range of the diagonal, upper-, lower-triangular elements is shown in the description of the corresponding figure.

In small samples, all estimators are biased and the bias increases with the dimension of the SVAR. The elements on the diagonal show the largest bias and the elements in the upper triangular have the smallest bias. This pattern can be explained by a bias due to scaling, meaning that for small T the GMM loss $E\left[g_T(B)'Wg_T(B)\right]$ is not minimized at $B = B_0$ but at $B = DB_0$ where $D = diag(d_1, ..., d_n)$ is a scaling matrix.[6] Moreover, Figure 2 shows that the CUE estimator has the largest bias. In contrast to that, the CUE-MI estimator performs notably better. In fact, in

[6]For example with $W = I$ the noise term in Equation 11 is equal to

$$\frac{1}{T}E\left[f(B_0,u_t)'Wf(B_0,u_t)\right] = \frac{1}{T}E\left[f(B_0,u_t)'f(B_0,u_t)\right], \tag{31}$$

which is the sum of the variances of the variance, covariance, and coskewness conditions. The variance of any moment condition m of the type $\prod_{i=1}^{n} e(B)_{i,t}^{m_i}$ at B_0 is equal to

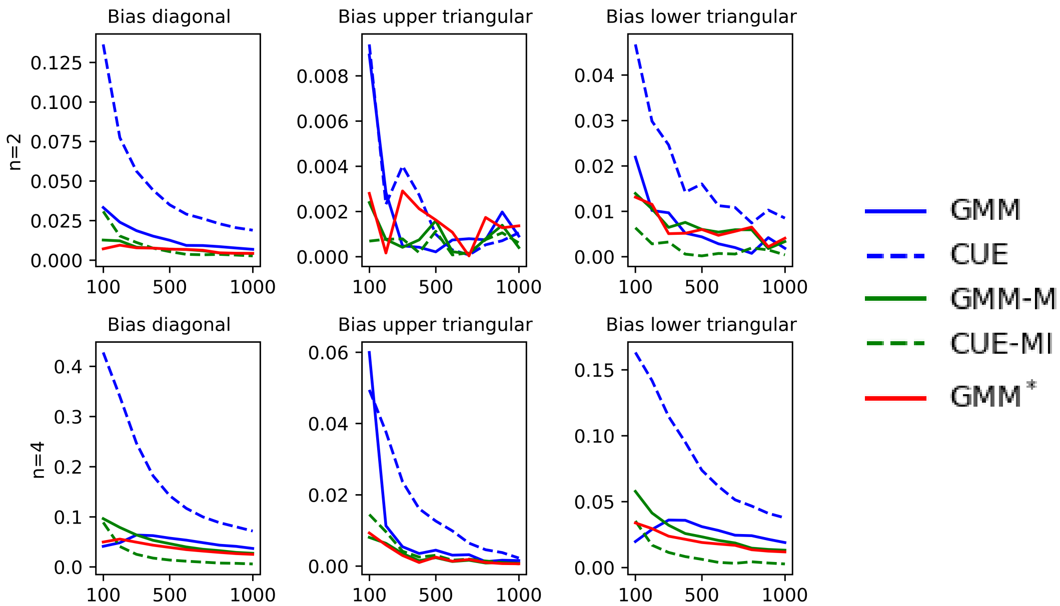
$$E\left[\prod_{i=1}^{n} e(B_0)_{i,t}^{2m_i})\right] = E\left[\prod_{i=1}^{n} \epsilon_{i,t}^{2m_i}\right]. \tag{32}$$

Let $S = diag(d_1, ..., d_n)$ be a scaling matrix and note that $De(B_0) = DB_0u = e(DB_0)$, such that

$$E\left[\prod_{i=1}^{n} e_{i,t}^{2m_i}(SB_0)\right] = E\left[\prod_{i=1}^{n} \frac{\epsilon_{i,t}^{2m_i}}{d_i^{2m_i}}\right] = \frac{1}{d_1^{2m_1}...d_n^{2m_n}} E\left[\prod_{i=1}^{n} \epsilon_{i,t}^{2m_i}\right]. \tag{33}$$

It is easy to see that at B_0 an increase in the scaling parameter d_i, which corresponds to a decrease of the sample variance of the i-th estimated structural shock, decreases the noise term and therefore, leads to a bias.

Figure 2: Average of the median absolute bias of the elements on the diagonal, upper- and lower-triangular of $\hat{B}$



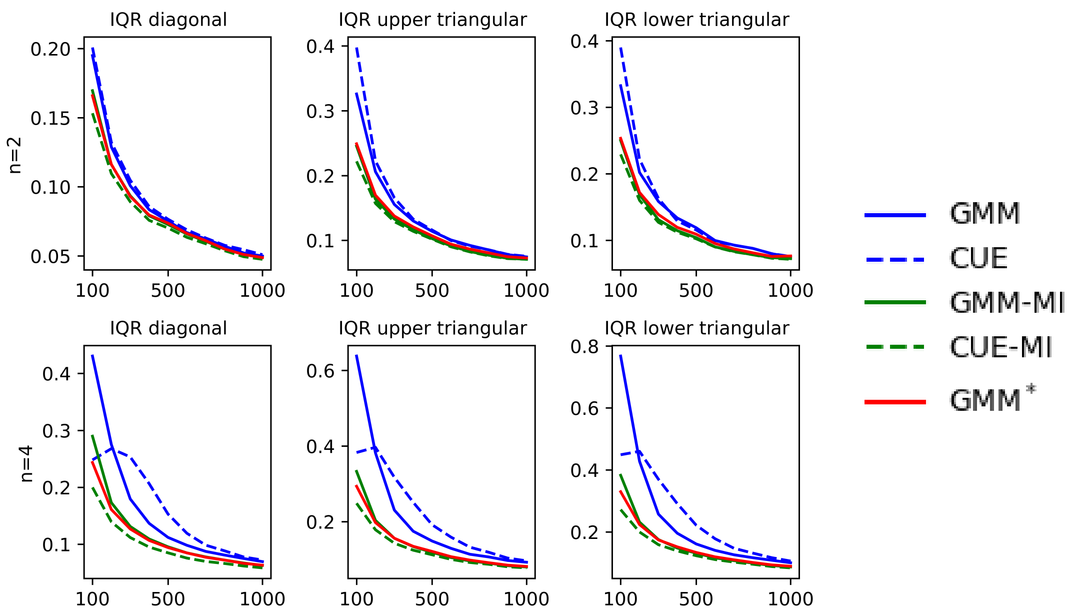
Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000$. The figure shows the average of the median absolute bias of the elements on the diagonal, upper-, lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. Let $\bar{B}^m := abs(\hat{B}^m - B_0)$ be the absolute bias in simulation m. The median absolute bias over all simulations M is then denoted by $bias := med(\bar{B}^m)$, which is a $n \times n$ matrix containing the median absolute bias over all simulations for each element $\hat{B}_{ij}$. The average of the median absolute bias of the elements on the diagonal, upper-, lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal, upper-, lower-triangular of $bias$.

the large SVAR it has the lowest bias on the diagonal and lower triangular of all estimators for sample sizes with more than 100 observations. This behavior can be explained by Figure 1, which suuggests that in small samples, the estimator $\hat{S}^{SI}(B_0)$ poorly approximates S while $\hat{S}^{SMI}(B_0)$ yields a much better approximation. Additionally, the derivation of the signal term in Equation (11) requires a nonrandom weighting matrix W, which is not satisfied by the CUE estimator. For a CUE estimator, the signal term can only be written as $E\left[\sum_{t\neq\tilde{t}} f(B,u_t)'\hat{S}(B)^{-1}f(B,u_{\tilde{t}})\right]$, which is not necessarily minimized at B_0. Therefore, the CUE can be biased since it searches for solutions which minimize the GMM loss by manipulating the weights $W(B) = \hat{S}(B)^{-1}$.

Figure 3 shows that the average IQR of the estimated elements increases with the dimension of the SVAR. Again, the GMM-MI and CUE-MI estimators perform better than the GMM and CUE estimator. In particular, for $n = 4$ and $T = 100$ the average IQR of the GMM estimator is two to three times larger than the average IQR of the CUE-MI estimator.

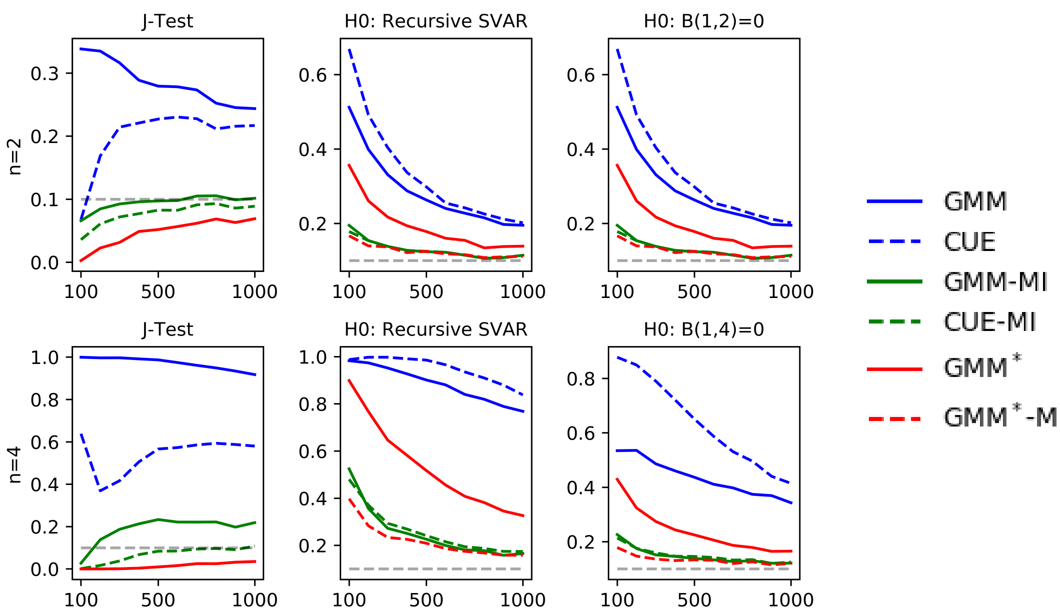
Secondly, I analyze the impact of the weighting scheme and the estimation of the asymptotic variance on the rejection frequencies for different tests. Figure 4 shows the rejection frequencies at the 10% nominal level for a J-Test, the Wald test with $H_0 : B_{i,j} = 0$ for $j > i$ testing the null hypothesis of a recursive SVAR, and a Wald test with the null hypothesis $H_0 : B_{1,n} = 0$ for $n = 2$ and $n = 4$. The Wald tests require an estimate of the asymptotic variance. The estimators GMM*, GMM, and CUE use the standard estimators $\hat{S}^{SI}_T(\hat{B})$ and $\hat{G}(\hat{B})$ and the estimators GMM*-MI, GMM-MI, and CUE-MI use the proposed estimators $\hat{S}^{SMI}_T(\hat{B})$ and $\hat{G}^{MI}(\hat{B})$ to estimate the asymptotic variance. The GMM*-MI estimator is equal to the GMM* estimator using the true asymptotically optimal weighting matrix $W = S^{-1}$. The tests for GMM* and GMM*-MI only differ in the way the asymptotic variance is estimated. Figure 4 shows that the distortion of the tests increases with an increase of the dimension of the SVAR, a decrease of the sample size, and an increase of the number of hypothesis being jointly tested. The GMM and CUE estimators have the largest distortions and they decrease only slowly with an increase of the sample size. For example, even in the largest sample with $T = 1000$ observations the CUE and GMM estimator reject the null hypothesis of a recursive SVAR in roughly 80% of

Figure 3: Average of the interquartile range of the elements on the diagonal, upper- and lower-triangular of $\hat{B}$



Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000$. The figure shows the average of the interquartile range (IQR) of the elements on the diagonal, upper-, lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. The interquartile range over all simulations M is then denoted by $iqr := quartile(\hat{B}^m, 0.75) - quartile(\hat{B}^m, 0.25)$, which is a $n \times n$ matrix containing the interquartile range over all simulations for each element $\hat{B}_{ij}$. The average of the interquartile range of the elements on the diagonal, upper-, lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal, upper-, lower- triangular of iqr.

Figure 4: Rejection rate at $\alpha = 10\%$ for J-Test, recursive SVAR Wald test, and Wald test with $H_0 : B_{1,n} = 0$ for $n = 2$ and $n = 4$



Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000$.

all cases. The rejection rates of GMM-MI and CUE-MI estimators are also distorted in small samples, however, the distortion is much smaller and decreases more quickly with an increase of the sample size. GMM* and GMM*-MI both use the same estimator, the tests only differ due to the estimated asymptotic variance. The smaller distortions of the GMM*-MI tests compared to the GMM* tests can solely be attributed to the impact of exploiting the assumption of mutually independent shocks to estimate the asymptotic variance.

6 Conclusion

This paper argues that the assumption of mutually independent shocks should be used to estimate the asymptotically efficient weighting matrix and the asymptotic variance of SVAR CUE and GMM estimators based on higher-order moment conditions. Without exploiting the assumption of mutually independent shocks, estimating the covariance of fourth-order moment conditions requires to estimate co-moments of order eight. This leads to biased and volatile estimates and oversize Wald test statistics in finite samples. With mutually independent shocks, the covariance of the higher-order moment conditions can be estimated as the product of moments up to order six. A Monte Carlo simulation demonstrates that the propose approach improves the finite sample performance of the CUE and GMM estimators, especially in larger SVAR models.

References

Anttonen, J., M. Lanne, and J. Luoto (2021). Statistically identified svar model with potentially skewed and fat-tailed errors. *Available at SSRN 3925575*.

Blanchard, O. J. (1989). A traditional interpretation of macroeconomic fluctuations. *The American Economic Review*, 1146–1164.

Bonhomme, S. and J.-M. Robin (2009). Consistent Noisy Independent Component Analysis. *Journal of Econometrics 149*(1), 12–25.

Burnside, C. and M. Eichenbaum (1996). Small-sample properties of gmm-based wald tests. *Journal of Business & Economic Statistics 14*(3), 294–308.

Donald, S. G. and W. K. Newey (2000). A jackknife interpretation of the continuous updating estimator. *Economics Letters 67*(3), 239–243.

Gouriéroux, C., A. Monfort, and J.-P. Renne (2017). Statistical Inference for Independent Component Analysis: Application to Structural VAR Models. *Journal of Econometrics 196*(1), 111–126.

Guay, A. (2021). Identification of structural vector autoregressions through higher unconditional moments. *Journal of Econometrics 225*(1), 27–46.

Hall, A. R. et al. (2005). *Generalized method of moments.* Oxford university press.

Han, C. and P. C. Phillips (2006). Gmm with many moment conditions. *Econometrica 74*(1), 147–192.

Hansen, L. P., J. Heaton, and A. Yaron (1996). Finite-sample properties of some alternative gmm estimators. *Journal of Business & Economic Statistics 14*(3), 262–280.

Herwartz, H. (2018). Hodges–lehmann detection of structural shocks–an analysis of macroeconomic dynamics in the euro area. *Oxford Bulletin of Economics and Statistics 80*(4), 736–754.

Keweloh, S. A. (2021). A generalized method of moments estimator for structural vector autoregressions based on higher moments. *Journal of Business & Economic Statistics 39*(3), 772–782.

Kilian, L. and H. Lütkepohl (2017). *Structural vector autoregressive analysis.* Cambridge University Press.

Lanne, M., K. Liu, and J. Luoto (2021). Identifying structural vector autoregression via large economic shocks. *Available at SSRN 3910532*.

Lanne, M. and J. Luoto (2020). Identification of economic shocks by inequality constraints in bayesian structural vector autoregression. *Oxford Bulletin of Economics and Statistics 82*(2), 425–452.

Lanne, M. and J. Luoto (2021). Gmm estimation of non-gaussian structural vector autoregression. *Journal of Business & Economic Statistics 39*(1), 69–81.

Lanne, M., M. Meitz, and P. Saikkonen (2017). Identification and Estimation of Non-Gaussian Structural Vector Autoregressions. *Journal of Econometrics 196*(2), 288–304.

Lewis, D. J. (2021, 05). Identifying Shocks via Time-Varying Volatility. *The Review of Economic Studies*. rdab009.

Mertens, K. and M. O. Ravn (2014). A reconciliation of svar and narrative estimates of tax multipliers. *Journal of Monetary Economics 68*, S1–S19.

Newey, W. K. and K. D. West (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies 61*(4), 631–653.

Newey, W. K. and F. Windmeijer (2009). Generalized method of moments with many weak moment conditions. *Econometrica 77*(3), 687–719.

Peersman, G. (2005). What caused the early millennium slowdown? evidence based on vector autoregressions. *Journal of Applied Econometrics 20*(2), 185–207.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, 1–48.

Uhlig, H. (2005). What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics 52*(2), 381–419.

A Appendix

A.1 Simulation with mixture of Gaussian distributions

This section supplements the Monte Carlo simulation shown in Section 5. Table 3 and 4 show the median (*med*), 0.25 quantile (*q*25), 0.75 quantile (*q*75), the 0.25 confidence level (*c*25), and the 0.75 confidence level (*c*75).

Table 3: Finite sample performance $n = 2$

	Element	$T = 100$	$T = 1000$
GMM^*	(1,1)	1.0 0.93 1.09 0.93 1.07	1.01 0.99 1.03 0.98 1.02
	(1,2)	−0.01 −0.13 0.12 −0.11 0.11	0.0 −0.04 0.04 −0.03 0.03
	(2,1)	0.52 0.39 0.64 0.39 0.61	0.51 0.47 0.54 0.47 0.53
	(2,2)	1.02 0.92 1.1 0.92 1.08	1.01 0.98 1.04 0.97 1.03
GMM	(1,1)	0.97 0.88 1.06 0.93 1.07	1.0 0.97 1.02 0.98 1.02
	(1,2)	0.0 −0.17 0.17 −0.11 0.11	−0.0 −0.04 0.04 −0.03 0.03
	(2,1)	0.48 0.31 0.65 0.39 0.61	0.5 0.46 0.54 0.47 0.53
	(2,2)	0.98 0.86 1.09 0.92 1.08	1.0 0.97 1.03 0.97 1.03
CUE	(1,1)	0.84 0.75 0.94 0.93 1.07	0.99 0.96 1.01 0.98 1.02
	(1,2)	−0.01 −0.2 0.2 −0.11 0.11	0.0 −0.04 0.04 −0.03 0.03
	(2,1)	0.46 0.26 0.65 0.39 0.61	0.49 0.46 0.53 0.47 0.53
	(2,2)	0.88 0.76 0.97 0.92 1.08	0.99 0.96 1.01 0.97 1.03
$GMM - MI$	(1,1)	1.01 0.93 1.1 0.93 1.07	1.01 0.99 1.03 0.98 1.02
	(1,2)	−0.0 −0.13 0.12 −0.11 0.11	0.0 −0.03 0.04 −0.03 0.03
	(2,1)	0.52 0.39 0.64 0.39 0.61	0.51 0.47 0.54 0.47 0.53
	(2,2)	1.02 0.93 1.11 0.92 1.08	1.01 0.98 1.04 0.97 1.03
$CUE - MI$	(1,1)	0.97 0.89 1.04 0.93 1.07	1.0 0.98 1.03 0.98 1.02
	(1,2)	−0.0 −0.12 0.11 −0.11 0.11	0.0 −0.03 0.04 −0.03 0.03
	(2,1)	0.5 0.38 0.61 0.39 0.61	0.5 0.47 0.54 0.47 0.53
	(2,2)	0.98 0.89 1.06 0.92 1.08	1.0 0.98 1.03 0.97 1.03

Monte Carlo simulation with 5000 iterations and sample sizes. Median (*med*), 0.25 quantile (*q*25), 0.75 quantile (*q*75), the 0.25 confidence level (*c*25), and the 0.75 confidence level (*c*75), where the confidence levels are calculated according to $B_{i,j} \pm z^* \frac{\sigma_{i,j}}{\sqrt{T}}$ with $z^* = 0.67$ and $\sigma_{i,j}$ is the square root of the variance of the element i, j according to $\sqrt{T}(\hat{B}_T - B_0) \xrightarrow{d} \mathcal{N}(0, MSM')$. For each element the data is shown as $\underset{\substack{q25\ q75 \\ c25\ c75}}{med}$.

Table 4: Finite sample performance $n = 4$

	$T = 100$				$T = 1000$			
GMM^*	1.04 0.95 1.14 0.93 1.07	0.0 −0.12 0.14 −0.11 0.11	−0.0 −0.13 0.13 −0.11 0.11	−0.02 −0.14 0.13 −0.11 0.11	1.02 1.0 1.05 0.98 1.02	−0.0 −0.04 0.04 −0.03 0.03	−0.0 −0.04 0.04 −0.03 0.03	0.0 −0.04 0.04 −0.03 0.03
	0.53 0.39 0.66 0.39 0.61	1.05 0.93 1.16 0.92 1.08	−0.01 −0.16 0.15 −0.12 0.12	−0.02 −0.17 0.15 −0.12 0.12	0.51 0.48 0.55 0.47 0.53	1.03 1.0 1.05 0.97 1.03	0.0 −0.04 0.04 −0.04 0.04	−0.0 −0.04 0.04 −0.04 0.04
	0.54 0.37 0.68 0.38 0.62	0.53 0.37 0.69 0.38 0.62	1.05 0.92 1.18 0.9 1.1	−0.02 −0.19 0.16 −0.14 0.14	0.51 0.47 0.56 0.46 0.54	0.51 0.47 0.55 0.46 0.54	1.02 0.99 1.06 0.97 1.03	−0.0 −0.05 0.05 −0.04 0.04
	0.54 0.35 0.71 0.36 0.64	0.54 0.35 0.72 0.36 0.64	0.54 0.35 0.7 0.36 0.64	1.05 0.89 1.19 0.89 1.11	0.51 0.46 0.56 0.46 0.54	0.51 0.46 0.56 0.46 0.54	0.51 0.46 0.56 0.46 0.54	1.03 0.99 1.07 0.96 1.04
GMM	1.0 0.85 1.19 0.93 1.07	0.13 −0.16 0.44 −0.11 0.11	0.05 −0.24 0.37 −0.11 0.11	0.01 −0.3 0.34 −0.11 0.11	0.96 0.94 0.99 0.98 1.02	−0.0 −0.04 0.04 −0.03 0.03	−0.0 −0.04 0.04 −0.03 0.03	0.0 −0.04 0.04 −0.03 0.03
	0.46 0.12 0.81 0.39 0.61	1.0 0.83 1.21 0.92 1.08	0.1 −0.2 0.44 −0.12 0.12	0.02 −0.31 0.37 −0.12 0.12	0.48 0.44 0.52 0.47 0.53	0.96 0.93 1.0 0.97 1.03	−0.0 −0.05 0.05 −0.04 0.04	−0.0 −0.05 0.05 −0.04 0.04
	0.47 0.1 0.84 0.38 0.62	0.51 0.13 0.89 0.38 0.62	0.98 0.77 1.21 0.9 1.1	0.05 −0.26 0.39 −0.14 0.14	0.48 0.44 0.53 0.46 0.54	0.48 0.43 0.53 0.46 0.54	0.96 0.92 1.0 0.97 1.03	−0.0 −0.06 0.05 −0.04 0.04
	0.46 0.06 0.85 0.36 0.64	0.5 0.11 0.91 0.36 0.64	0.5 0.08 0.92 0.36 0.64	0.87 0.58 1.13 0.89 1.11	0.48 0.43 0.53 0.46 0.54	0.48 0.43 0.54 0.46 0.54	0.48 0.42 0.53 0.46 0.54	0.96 0.92 1.01 0.96 1.04
CUE	0.57 0.5 0.66 0.93 1.07	0.1 −0.08 0.26 −0.11 0.11	0.05 −0.13 0.22 −0.11 0.11	0.01 −0.19 0.21 −0.11 0.11	0.93 0.9 0.95 0.98 1.02	0.0 −0.04 0.04 −0.03 0.03	0.0 −0.04 0.04 −0.03 0.03	−0.0 −0.04 0.04 −0.03 0.03
	0.32 0.13 0.5 0.39 0.61	0.6 0.5 0.7 0.92 1.08	0.09 −0.1 0.28 −0.12 0.12	0.02 −0.19 0.22 −0.12 0.12	0.46 0.42 0.51 0.47 0.53	0.93 0.89 0.96 0.97 1.03	−0.0 −0.05 0.05 −0.04 0.04	−0.0 −0.05 0.05 −0.04 0.04
	0.33 0.11 0.53 0.38 0.62	0.35 0.13 0.57 0.38 0.62	0.59 0.45 0.72 0.9 1.1	0.04 −0.17 0.24 −0.14 0.14	0.46 0.41 0.51 0.46 0.54	0.46 0.41 0.51 0.46 0.54	0.93 0.89 0.97 0.97 1.03	−0.01 −0.06 0.05 −0.04 0.04
	0.33 0.09 0.56 0.36 0.64	0.35 0.1 0.6 0.36 0.64	0.33 0.08 0.58 0.36 0.64	0.53 0.33 0.7 0.89 1.11	0.46 0.4 0.52 0.46 0.54	0.46 0.41 0.52 0.46 0.54	0.46 0.41 0.52 0.46 0.54	0.93 0.88 0.97 0.96 1.04
$GMM - MI$	1.1 0.98 1.22 0.93 1.07	0.0 −0.14 0.16 −0.11 0.11	−0.0 −0.15 0.15 −0.11 0.11	−0.01 −0.16 0.14 −0.11 0.11	1.03 1.0 1.05 0.98 1.02	−0.0 −0.04 0.04 −0.03 0.03	−0.0 −0.04 0.04 −0.03 0.03	0.0 −0.04 0.04 −0.03 0.03
	0.56 0.39 0.71 0.39 0.61	1.1 0.96 1.24 0.92 1.08	−0.0 −0.18 0.18 −0.12 0.12	−0.01 −0.19 0.17 −0.12 0.12	0.52 0.48 0.55 0.47 0.53	1.03 1.0 1.06 0.97 1.03	0.0 −0.04 0.04 −0.04 0.04	−0.0 −0.04 0.04 −0.04 0.04
	0.56 0.38 0.73 0.38 0.62	0.56 0.38 0.74 0.38 0.62	1.1 0.94 1.25 0.9 1.1	−0.02 −0.2 0.19 −0.14 0.14	0.51 0.47 0.56 0.46 0.54	0.51 0.47 0.55 0.46 0.54	1.03 0.99 1.06 0.97 1.03	−0.0 −0.05 0.05 −0.04 0.04
	0.56 0.34 0.77 0.36 0.64	0.56 0.35 0.77 0.36 0.64	0.55 0.34 0.75 0.36 0.64	1.09 0.91 1.25 0.89 1.11	0.51 0.46 0.56 0.46 0.54	0.51 0.47 0.56 0.46 0.54	0.51 0.47 0.56 0.46 0.54	1.03 0.99 1.07 0.96 1.04
$CUE - MI$	0.91 0.83 0.99 0.93 1.07	−0.0 −0.11 0.11 −0.11 0.11	−0.01 −0.12 0.11 −0.11 0.11	−0.01 −0.13 0.1 −0.11 0.11	0.99 0.97 1.01 0.98 1.02	−0.0 −0.04 0.04 −0.03 0.03	−0.0 −0.04 0.03 −0.03 0.03	0.0 −0.03 0.04 −0.03 0.03
	0.46 0.34 0.57 0.39 0.61	0.91 0.82 1.0 0.92 1.08	−0.01 −0.14 0.13 −0.12 0.12	−0.02 −0.15 0.11 −0.12 0.12	0.5 0.46 0.53 0.47 0.53	0.99 0.97 1.02 0.97 1.03	0.0 −0.04 0.04 −0.04 0.04	−0.0 −0.04 0.04 −0.04 0.04
	0.47 0.33 0.59 0.38 0.62	0.46 0.32 0.59 0.38 0.62	0.92 0.8 1.02 0.9 1.1	−0.03 −0.17 0.12 −0.14 0.14	0.5 0.46 0.54 0.46 0.54	0.5 0.45 0.54 0.46 0.54	0.99 0.96 1.03 0.97 1.03	−0.0 −0.05 0.04 −0.04 0.04
	0.48 0.33 0.62 0.36 0.64	0.47 0.31 0.6 0.36 0.64	0.45 0.31 0.6 0.36 0.64	0.91 0.78 1.02 0.89 1.11	0.5 0.45 0.54 0.46 0.54	0.5 0.45 0.54 0.46 0.54	0.5 0.45 0.54 0.46 0.54	0.99 0.96 1.03 0.96 1.04

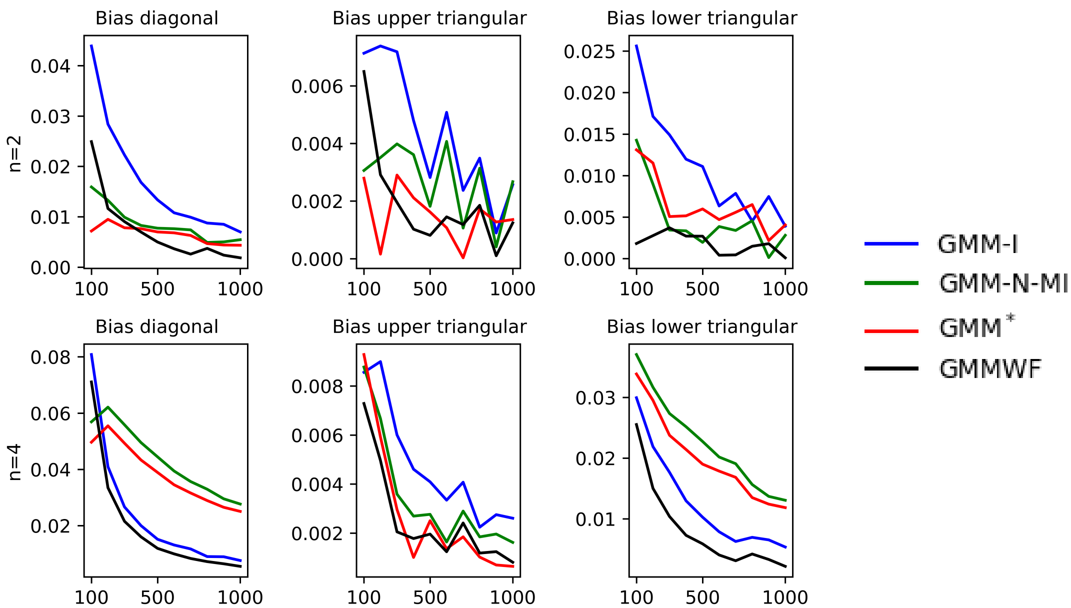
Monte Carlo simulation with 5000 iterations and sample sizes. Median (med), 0.25 quantile ($q25$), 0.75 quantile ($q75$), the 0.25 confidence level ($c25$), and the 0.75 confidence level ($c75$), where the confidence levels are calculated according to $B_{i,j} \pm z^* \frac{\sigma_{i,j}}{\sqrt{T}}$ with $z^* = 0.67$ and $\sigma_{i,j}$ is the square root of the variance of the element i, j according to $\sqrt{T}(\hat{B}_T - B_0) \xrightarrow{d} \mathcal{N}(0, MSM')$. For each element the data is shown as $\underset{\substack{q25\ q75 \\ c25\ c75}}{med}$.

Figure 5, 6, and 7 show the average of the median absolute bias, the average of the interquartile range (IQR) of the estimated elements, and the rejection frequencies for the hypothesis tests. The Figures contain results for the following estimators:

- GMM*: A one-step GMM estimator with the asymptotically efficient weighting matrix $W = S^{-1}$.
- GMM-I: A one-step GMM estimator with $W = I$..
- GMM-N-MI: A one-step GMM estimator with $W = S_{Norm}^{SMI^{-1}}$, where S_{Norm}^{SMI} denotes the long-run covariance matrix under serially and mutually independent Gaussian shocks.
- GMMWF: A one-step GMM estimator with the fast weighting matrix proposed in Keweloh (2021).

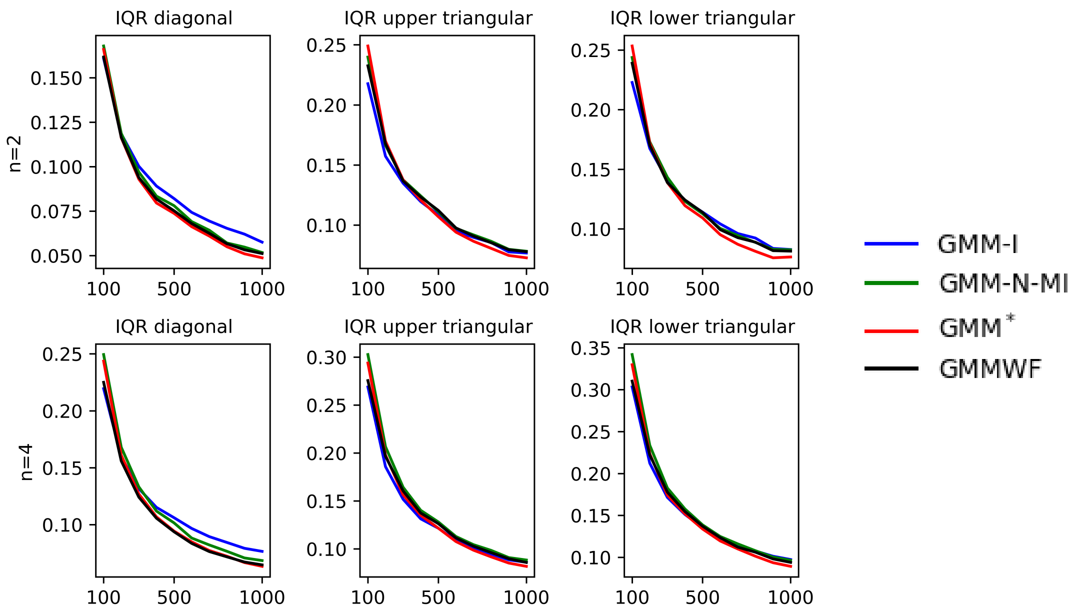
All estimators use all second- to fourth-order moment conditions implied by independent shocks. The Wald tests require an estimate of the asymptotic variance. The estimators GMM* and GMM-I use the standard estimators $\hat{S}_T^{SI}(\hat{B})$ and $\hat{G}(\hat{B})$ and the estimators GMM*-MI, GMM-N-MI, and GMMWF use the proposed estimators $\hat{S}_T^{SMI}(\hat{B})$ and $\hat{G}^{MI}(\hat{B})$ to estimate the asymptotic variance. The GMM*-MI estimator is equal to the GMM* estimator using the true asymptotically optimal weighting matrix $W = S^{-1}$. The tests for GMM* and GMM*-MI only differ in the way the asymptotic variance is estimated.

Figure 5: Average of the median absolute bias of the elements on the diagonal, upper- and lower-triangular of $\hat{B}$ - One-step estimators



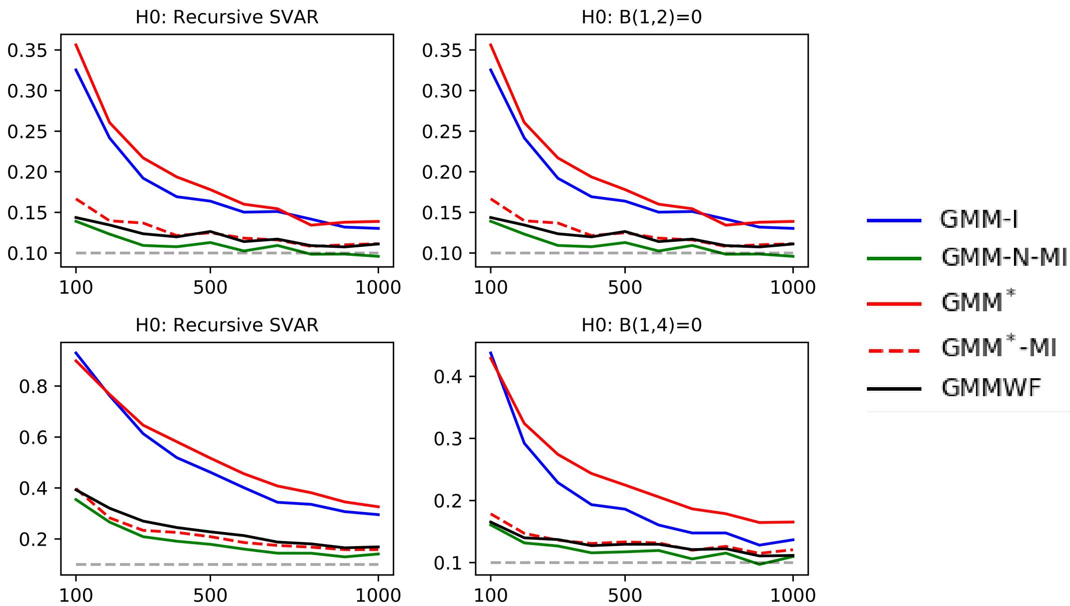
Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000$. The figure shows the average of the median absolute bias of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. Let $\bar{B}^m := abs(\hat{B}^m - B_0)$ be the absolute bias in simulation m. The median absolute bias over all simulations M is then denoted by $bias := med(\bar{B}^m)$, which is a $n \times n$ matrix containing the median absolute bias over all simulations for each element $\hat{B}_{ij}$. The average of the median absolute bias of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal/upper-/lower-triangular of $bias$.

Figure 6: Average of the interquartile range of the elements on the diagonal, upper- and lower-triangular of $\hat{B}$ - One-step estimators



Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000$. The figure shows the average of the interquartile range (IQR) of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. The interquartile range over all simulations M is then denoted by $iqr := quartile(\hat{B}^m, 0.75) - quartile(\hat{B}^m, 0.25)$, which is a $n \times n$ matrix containing the interquartile range over all simulations for each element $\hat{B}_{ij}$. The average of the interquartile range of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal/upper-/lower-triangular of iqr.

Figure 7: Rejection rate at $\alpha = 10\%$ for J-Test, recursive SVAR Wald test, and Wald test with $H_0 : B_{1,n} = 0$ for $n = 2$ and $n = 4$ - One-step estimators

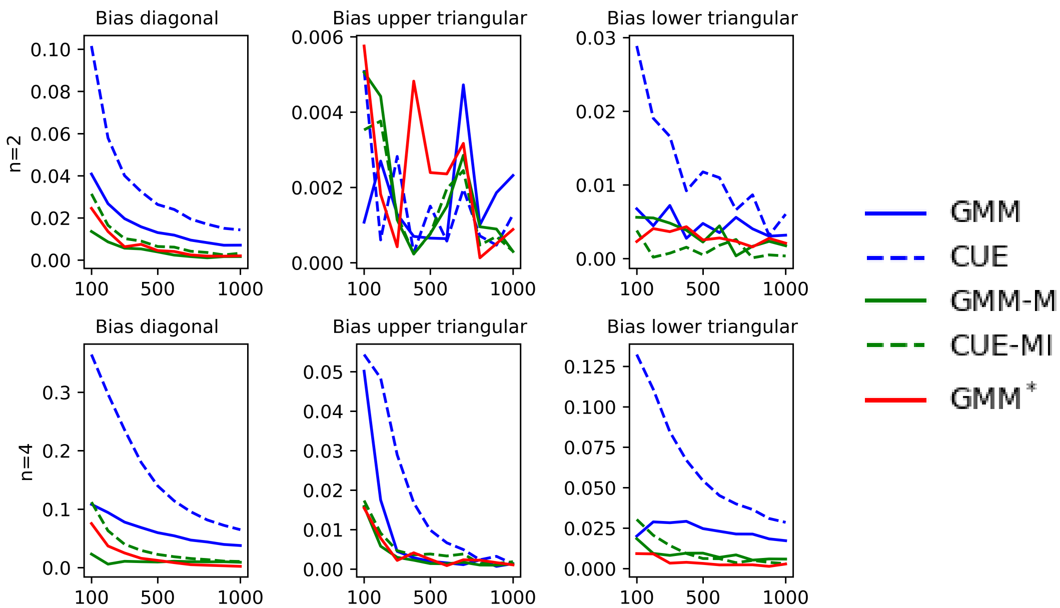


Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000$.

A.2 Simulation with t-distributed shocks

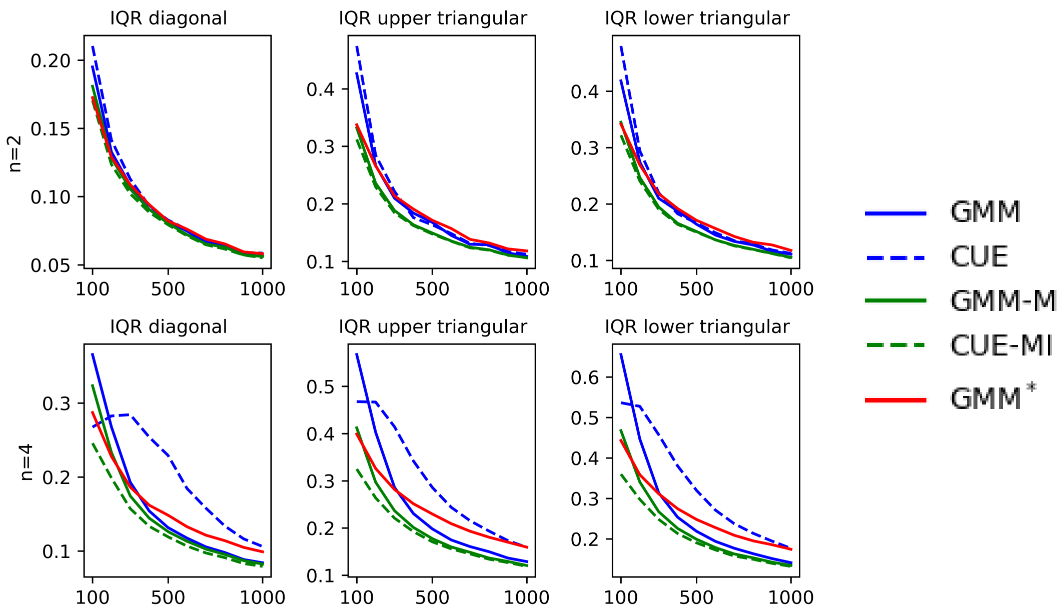
This section shows analogous results to the Monte Carlo simulation in Section 5. However, the structural shocks are drawn from a t-distribution with seven degrees of freedom. Additionally, the shocks have been normalized to unit variance by multiplying each shock with $1/\sqrt{(v/(v-2))}$ and $v = 7$.

Figure 8: Average of the median absolute bias of the elements on the diagonal, upper- and lower-triangular of $\hat{B}$ - t distributed shocks



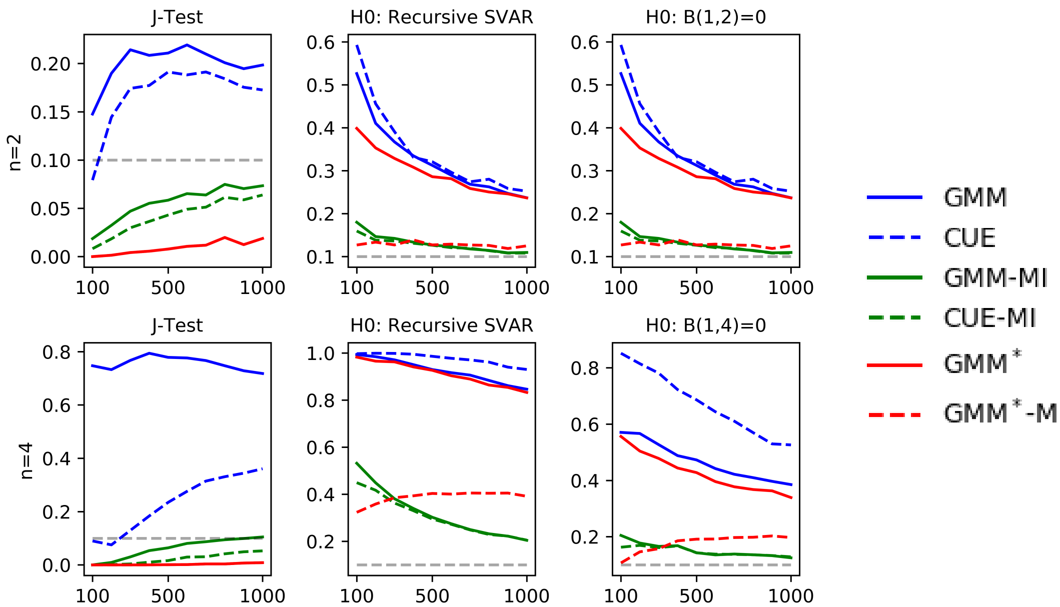
Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000$. The figure shows the average of the median absolute bias of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. Let $\bar{B}^m := abs(\hat{B}^m - B_0)$ be the absolute bias in simulation m. The median absolute bias over all simulations M is then denoted by $bias := med(\bar{B}^m)$, which is a $n \times n$ matrix containing the median absolute bias over all simulations for each element $\hat{B}_{ij}$. The average of the median absolute bias of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal/upper-/lower-triangular of $bias$.

Figure 9: Average of the interquartile range of the elements on the diagonal, upper- and lower-triangular of $\hat{B}$ - t distributed shocks



Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000$. The figure shows the average of the interquartile range (IQR) of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. The interquartile range over all simulations M is then denoted by $iqr := quartile(\hat{B}^m, 0.75) - quartile(\hat{B}^m, 0.25)$, which is a $n \times n$ matrix containing the interquartile range over all simulations for each element $\hat{B}_{ij}$. The average of the interquartile range of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal/upper-/lower-triangular of iqr.

Figure 10: Rejection rate at $\alpha = 10\%$ for J-Test, recursive SVAR Wald test, and Wald test with $H_0 : B_{1,n} = 0$ for $n = 2$ and $n = 4$ - t distributed shocks



Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000$.