SFB 823

A generalized method of moments estimator for structural vector autoregressions based on higher moments

Sascha Alexander Keweloh

Nr. 36/2018

Discussion Paper

SFB 823

A Generalized Method of Moments Estimator for Structural Vector Autoregressions Based on Higher Moments

Sascha Alexander Keweloh* **

TU Dortmund University

September 11, 2019

I propose a generalized method of moments estimator for structural vector autoregressions with independent and non-Gaussian shocks. The shocks are identified by exploiting information contained in higher moments of the data. Extending the standard identification approach, which relies on the covariance, to the coskewness and cokurtosis allows to identify and estimate the simultaneous interaction without any further restrictions. I analyze the finite sample properties of the estimator and apply it to illustrate the simultaneous interaction between economic activity, oil and stock prices.

JEL Codes: C32, G12, Q43

Keywords: Non-Gaussian, independence, identification, oil prices, stock returns

---

* 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 acknowledge.

** I gratefully thank the associate editor and the two anonymous referees for their comments and recommendations.

1 Introduction

One of the most important tools to estimate the effects of economic shocks on a set of variables is the structural vector autoregressive model (SVAR). Due to the simultaneous nature of the interaction among economic variables, the identification of the underlying structural shocks generally requires the researcher to impose restrictions on the system. A variety of identifying restrictions have been proposed in the SVAR literature. The one thing all identification approaches have in common is the assumption of uncorrelated structural shocks. Unfortunately, uncorrelatedness is not sufficient to identify the simultaneous interaction.

A large part of the SVAR literature eliminates this lack of identification with short-run or long-run restrictions. For example, the often used recursive SVAR employs short-run zero restrictions between the included variables. However, these restrictions are often difficult to find, or hardly justifiable based on economic theory. Therefore, a number of proposals have been made to avoid these restrictions. The general idea is to exploit the independence of the shocks and not merely their uncorrelatedness. With independent and non-Gaussian shocks, results from the independent component analysis (ICA) literature can be applied to identify the SVAR.

In this paper, I present a generalized method of moments (GMM) estimator for non-Gaussian SVAR models with independent shocks. The identification is derived as a straightforward extension of traditional approaches relying on the assumption of uncorrelated shocks to independent shocks. The approach is purely data driven and does not require any assumptions or restrictions apart from independent and non-Gaussian shocks (more precisely: at most one shock is allowed to have zero skewness or zero excess kurtosis). In macroeconomic applications, where restrictions are scarce and traditional identification approaches fail, the proposed estimator allows to identify and estimate a given SVAR by exploiting information contained in moments beyond the variance.

Independence has rarely been used to identify SVAR models. A few authors use independent shocks to evaluate the fit of different causal orders (Hyvärinen et al. (2010) or Moneta et al. (2013)). More recently, independence has been used to identify SVAR models without restrictions on the interaction of the included variables. Lanne et al. (2017) and Gouriéroux et al.

(2017) propose a maximum likelihood (ML) and a pseudo maximum likelihood (PML) estimator for non-Gaussian SVAR models. Lanne and Luoto (2019) use cokurtosis conditions to derive a GMM estimator for non-Gaussian SVAR models. The authors relax the assumption of independent structural shocks and instead assume uncorrelated shocks with a few shocks additionally satisfying cokurtosis restrictions. Herwartz (2018) proposes a method to find the least dependent shocks, measured by the difference between the empirical copula and the copula under independence. Herwartz and Plödt (2016) apply the method and analyze the interaction of real economic activity, oil production and the real price of oil.

The SVAR-GMM estimator proposed in this paper requires no distributional assumptions, apart from independent and non-Gaussian shocks. In contrast to that, the estimators proposed by Lanne et al. (2017) and Gouriéroux et al. (2017) require to specify the distribution of the structural shocks a priori. In macroeconomic applications, distributional restrictions are probably even harder to derive from economic theory than traditional short-run or long-run restrictions. The PML estimator proposed by Gouriéroux et al. (2017) is to some extent robust to distributional misspecification. Based on a Monte Carlo study, I show that misspecifying the distribution can lead to a serious deterioration of the finite sample performance of the PML estimator. I find that the SVAR-GMM estimator performs more robustly across different error term specifications and it performs better than the misspecified PML estimator.

The moment conditions derived in this paper ensure global identification up to sign and permutation. However, the number of moment conditions increases quickly with the dimension of the SVAR, which makes the estimator computationally expensive in large models. Lanne and Luoto (2019) propose a GMM estimator which estimates the simultaneous relations based on a subset of the moment conditions derived in this paper. Relying only on a subset of the moment conditions yields a computationally cheap estimator, but it also destroys the global identification result. Therefore, the estimator proposed by Lanne and Luoto (2019) is only locally identified, which hinders asymptotic inference. I propose an alternative way to decrease the computational burden of the estimator. By sacrificing the asymptotic efficiency of the estimator, one can gain a consistent, globally identified and computationally cheap approximation, which is denoted as the fast SVAR-GMM estimator. Based on a Monte Carlo study, I show that even though the fast SVAR-GMM estimator is asymptotically not optimal, it performs well and robustly across various specifications and sample sizes. In small samples, it often even performs better than the two-step SVAR-GMM estimator, which estimates the optimal weighting matrix based on the standard two-step GMM procedure.

I find that neither the GMM estimator proposed by Lanne and Luoto (2019) nor the PML estimator proposed by Gouriéroux et al. (2017) (with a pseudo distribution equal to a t-distribution) are able to exploit information contained in the skewness of the structural shocks. Instead, both estimators primarily rely on the excess kurtosis of the shocks. The estimator proposed in this paper is more general and can use information contained in the skewness and excess kurtosis. I provide empirical evidence that macroeconomic variables like economic activity, oil or stock prices are driven by skewed shocks. The Monte Carlo study shows that estimators based on the skewness have desirable small sample properties. In particular, I find that the small sample bias and standard deviation of an estimator based on the skewness are driven by the relative skewness of the shocks and I find no deterioration with a decreasing sample size.

In an empirical application, the estimator is applied to analyze the interaction of economic activity, oil and stock prices. SVAR models with oil and stock prices have often been identified by imposing a recursive order on both variables, see Sadorsky (1999) or Kilian and Park (2009). I challenge this practice and provide evidence that no zero restrictions on the simultaneous relationship between oil and stock prices are feasible.

The remainder of the paper is organized as follows. Section 2 presents the SVAR model and derives the identification problem. Section 3 illustrates how the usual SVAR identification approach relying on uncorrelated shocks can be extended to independent shocks. Section 4 introduces the notation. Section 5 derives the identification of the SVAR model based on higher moments and introduces the SVAR-GMM estimator. Section 6 derives the fast SVAR-GMM estimator and Section 7 analyzes the finite sample properties of the estimators in a Monte Carlo study. The estimator is applied in Section 8 to study the interaction of economic activity, oil and stock prices. Section 9 concludes.

Throughout the paper real numbers are denoted by $\mathbb{R}$, natural numbers are denoted by $\mathbb{N}$ and the identity matrix is denoted by I. Moreover, the function $vec(.)$ denotes the vectorization of a matrix and $det(.)$ denotes the determinant of a matrix.

2 SVAR model

This section briefly summarizes the identification problem in SVAR models. A detailed explanation can be found in Kilian and Lütkepohl (2017). The SVAR is given by

$$A_0 y_t = A_1 y_{t-1} + ... + A_p y_{t-p} + \epsilon_t, \tag{1}$$

with constant parameter matrices $A_0, ..., A_p \in \mathbb{R}^{n \times n}$, the n-dimensional vector of time series $y_t = [y_{1,t}, ..., y_{n,t}]'$ and the vector of structural shocks $\epsilon_t = [\epsilon_{1,t}, ..., \epsilon_{n,t}]'$. The structural shocks are supposed to satisfy the following assumptions.

Assumption 1. *(i) ϵ_t is an i.i.d. vector of random variables.*
(ii) ϵ_t has mutually independent components, meaning that $\epsilon_{i,t}$ is independent of $\epsilon_{j,t}$ for $i \neq j$.
(iii) Each component of ϵ_t has zero mean, unit variance and finite third and fourth moments.
(iv) At most one component of ϵ has zero skewness and/or at most one component of ϵ has zero excess kurtosis.

The parameter matrix governing the simultaneous interaction is assumed to be invertible.

Assumption 2. $A_0 \in \mathcal{A} := \{A \in \mathbb{R}^{n \times n} | det(A) \neq 0\}$.

Equation (1) cannot be estimated consistently by OLS since a non-diagonal matrix A_0 leads to endogenous regressors. The reduced form vector autoregression (VAR) is given by

$$y_t = C_1 y_{t-1} + ... + C_p y_{t-p} + u_t. \tag{2}$$

The reduced form shocks u_t are i.i.d., and the VAR can be estimated by OLS. However, the estimated reduced form parameters and the reduced form shocks are of limited interest for the structural analysis, which focuses on the structural parameters and the structural shocks. The reduced form shocks can be written as a linear combination of the structural shocks

$$u_t = A_0^{-1} \epsilon_t. \tag{3}$$

However, neither the parameters of the matrix governing the simultaneous interaction nor the structural shocks are known. Put differently, the structural shocks cannot be directly recovered from the estimated reduced form VAR, leaving us with an identification problem.

Define the unmixed innovations as the vector of random variables obtained by unmixing the reduced form shocks by a matrix $A \in \mathcal{A}$ as

$$e_t(A) := A u_t. \tag{4}$$

If the unmixing matrix A is equal to A_0, the unmixed innovations are equal to the structural shocks. I show how to derive a system of moment conditions which globally identifies the matrix governing the simultaneous interaction and the structural shocks up to sign and permutation. The identification requires independent and non-Gaussian structural shocks. Intuitively, if the unmixed innovations and the structural shocks have the same covariance, coskewness and cokurtosis, then the unmixed innovations and structural shocks are equal up to sign and permutation. Note that equations (3) and (4) contain no lag structure and the shocks are i.i.d. over time. Therefore, the time index is suppressed whenever possible.

3 Illustration: Identification and higher moments

This section uses a bivariate SVAR to illustrate the intuition behind the identification approach presented in Section 5. The approach is a straightforward extension of the standard identification scheme relying on uncorrelated shocks to higher moments and independent shocks. In a nutshell, the shocks ϵ are supposed to be mutually independent and thus uncorrelated, which allows to postulate moment conditions. However, these second order moment conditions are not sufficient for identification. Exploiting the implications of independent shocks concerning higher co-moments allows to postulate additional moment conditions and to identify the SVAR.

Consider a bivariate SVAR such that the unmixed innovations are given by

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{5}$$

One can now use Assumption 1 to derive the stochastic properties of the unknown structural shocks and choose an unmixing matrix such that the unmixed innovations fulfill the same stochastic properties. Basically every SVAR identification approach exploits the second order properties of the structural shocks. In particular, the structural shocks have unit variance and thus the unmixed innovations should satisfy the following variance conditions

$$1 \overset{!}{=} E\left[e_1^2\right] = E\left[(a_{11}u_1 + a_{12}u_2)^2\right] \tag{6}$$

$$1 \overset{!}{=} E\left[e_2^2\right] = E\left[(a_{21}u_1 + a_{22}u_2)^2\right]. \tag{7}$$

Moreover, the components ϵ_1 and ϵ_2 are assumed to be (second order) uncorrelated and therefore satisfy the covariance condition

$$0 \overset{!}{=} E\left[e_1 e_2\right] = E\left[(a_{11}u_1 + a_{12}u_2)(a_{21}u_1 + a_{22}u_2)\right]. \tag{8}$$

Exploiting all second order properties yields three equations in the four unknown coefficients of A. Therefore, infinitely many unmixing matrices A generate unmixed innovations satisfying the second order properties and thus second order statistics are not sufficient to identify A_0.

If one coefficient of A_0 is known a priori, the corresponding coefficient of the unmixing matrix A can be restricted and the system is just identified, see e.g. Rubio-Ramírez et al. (2010). Using short-run restrictions of this kind is probably the most common way to solve the identification problem. However, the approach requires the researcher to know half of the simultaneous structure a priori and by relying solely on second moments, these restrictions cannot be tested. Technically, short-run restrictions reduce the number of unknowns. Alternatively, one could try to increase the number of equations while keeping the number of unknowns constant. Increasing the number of equations until no short-run restrictions are required seems appealing since one does not need to restrict the simultaneous structure a priori. Additional equations can only be generated by imposing more structure on the stochastic properties of structural shocks. The following argument shows how independence accomplishes that.

So far, only second order properties of the structural shocks have been used. Independent and non-Gaussian structural shocks allow to exploit information contained in moments beyond the variance and covariance. For independent structural shocks, it is fairly straightforward to generate as many equations as desired. In particular, independence implies third order uncorrelatedness, $E\left[\epsilon_1^2\epsilon_2\right] = E\left[\epsilon_1^2\right]E\left[\epsilon_2\right] = 0$ and analogously $E\left[\epsilon_1\epsilon_2^2\right] = 0$, and therefore the unmixed innovations should satisfy the coskewness conditions

$$0 \overset{!}{=} E\left[e_1^2e_2\right] = E\left[(a_{11}u_1 + a_{12}u_2)^2(a_{21}u_1 + a_{22}u_2)\right] \tag{9}$$

$$0 \overset{!}{=} E\left[e_1e_2^2\right] = E\left[(a_{11}u_1 + a_{12}u_2)(a_{21}u_1 + a_{22}u_2)^2\right]. \tag{10}$$

Thus, independence allows to generate further moment conditions analogously to the usual approach based on second moments. If the shocks are non-Gaussian, these moment conditions contain further information which allow to identify the SVAR. The system of equations (6) - (10) now contains five equations in the four unknowns. Note that the system contains nonlinear equations and thus it is not obvious whether the system globally identifies the SVAR. Section 5 shows that the system indeed identifies the SVAR up to sign and permutation, given that the structural shocks are independent and non-Gaussian.

4 Notation

The identification approach requires to calculate (co-)moments of order two, three and four. The following notation yields a short expression to collect all (co-)moments of a given order r. For an n-dimensional random variable x define a moment generating index $W = [w_1, ..., w_r] \in \{1, ..., n\}^r$ with $r \in \mathbb{N}$ and let

$$x_W := [x_{w_1}, ..., x_{w_r}]. \tag{11}$$

Let the expected value be denoted by $E[x_W] := E[x_{w_1}...x_{w_r}]$ and let $E_T[x_W]$ denote the respective sample mean. Furthermore, define a moment generating set $\mathcal{W} = \{W_1, ..., W_l\}$ with the moment generating indices $W_i \in \{1, ..., n\}^r$ for $i = 1, ..., l$ and define

$$x_{\mathcal{W}} := \begin{bmatrix} x_{W_1} \\ ... \\ x_{W_l} \end{bmatrix}, \; E[x_{\mathcal{W}}] := \begin{bmatrix} E[x_{W_1}] \\ ... \\ E[x_{W_l}] \end{bmatrix} \text{ and } E_T[x_{\mathcal{W}}] := \begin{bmatrix} E_T[x_{W_1}] \\ ... \\ E_T[x_{W_l}] \end{bmatrix}. \tag{12}$$

This notation can be used to generate a vector containing the r-th (co-)moments of an n-dimensional random variable.

For $r \in \mathbb{N}$, define the r-th moments generating set as

$$\mathcal{M}(r) = \{[m_1, ..., m_r] \in \{1, ..., n\}^r | m_1 = ... = m_r\}. \tag{13}$$

The set contains n elements and the vector $E\left[\epsilon_{\mathcal{M}(r)}\right]$ contains all r-th moments of the structural shocks and the vector $E\left[e_{\mathcal{M}(r)}(A)\right]$ contains all r-th moments of the unmixed innovations. In the bivariate example, the second moments generating set is equal to $\mathcal{M}(2) = \{[1,1], [2,2]\}$ and the variance of the structural shocks is given by

$$E\left[\epsilon_{\mathcal{M}(2)}\right] = E\begin{bmatrix}\epsilon_1\epsilon_1\\ \epsilon_2\epsilon_2\end{bmatrix}. \tag{14}$$

For $r \in \mathbb{N}$, define the r-th co-moments generating set as

$$\mathcal{C}(r) = \{[c_1, ..., c_r] \in \{1, ..., n\}^r | [c_1, ..., c_r] \notin \mathcal{M}(r) \text{ and } c_i \leq c_j \text{ for } i < j\}. \tag{15}$$

The set contains $\frac{(n+r-1)!}{(n-1)!r!} - n$ elements and can be used to generate the corresponding co-moments of an n-dimensional random variable. In particular, the vector $E\left[\epsilon_{\mathcal{C}(r)}\right]$ contains all r-th co-moments of the structural shocks and the vector $E\left[e_{\mathcal{C}(r)}(A)\right]$ contains all r-th co-moments of the unmixed innovations. In the bivariate example, the second co-moments generating set is equal to $\mathcal{C}(2) = \{[1,2]\}$ and the covariance of the structural shocks is given by

$$E\left[\epsilon_{\mathcal{C}(2)}\right] = E\left[\epsilon_1\epsilon_2\right]. \tag{16}$$

Analogously, the third co-moment generating set is equal to $\mathcal{C}(3) = \{[1,1,2], [1,2,2]\}$ and the coskewness of the structural shocks is given by

$$E\left[\epsilon_{\mathcal{C}(3)}\right] = E\begin{bmatrix}\epsilon_1\epsilon_1\epsilon_2\\ \epsilon_1\epsilon_2\epsilon_2\end{bmatrix}. \tag{17}$$

Additionally, for $X = [x_1, ..., x_r] \in \{1, ..., n\}^r$ define the index counting function

$$\#X := \left[\sum_{x\in X} 1_{x=1}, ..., \sum_{x\in X} 1_{x=n}\right], \tag{18}$$

where $1_{x=1} = \begin{cases} 1 & \text{, if } x = 1 \\ 0 & \text{, else} \end{cases}$ such that the i-th element of $\#X$ counts how often the index i appears in X.

5 SVAR-GMM estimator

This section generalizes the identification technique sketched in Section 3 to an n-dimensional non-Gaussian SVAR. I first derive a system of variance, covariance, coskewness and cokurtosis conditions, which globally identify the non-Gaussian SVAR up to sign and permutations. The SVAR is then estimated by matching the moments via a GMM estimator.

First, the (co-)moments of the unknown structural shocks need to be derived. The (co-)moments follow from Assumption 1.

Proposition 1. *Let ϵ satisfy Assumption 1. It holds that*

1. *For* $[m_1, m_2] \in \mathcal{M}(2)$*:* $E\left[\epsilon_{[m_1,m_2]}\right] = 1$
2. *For* $[c_1, c_2] \in \mathcal{C}(2)$*:* $E\left[\epsilon_{[c_1,c_2]}\right] = 0$
3. *For* $[c_1, c_2, c_3] \in \mathcal{C}(3)$*:* $E\left[\epsilon_{[c_1,c_2,c_3]}\right] = 0$
4. *For* $[c_1, c_2, c_3, c_4] \in \mathcal{C}(4)$*:* $E\left[\epsilon_{[c_1,c_2,c_3,c_4]}\right] = \begin{cases} 1, \text{ if } c_1 = c_2 \text{ and } c_3 = c_4 \\ 0, \text{ else} \end{cases}$

Proof. Independence embedded in Assumption 1 implies that for $C = [c_1, ..., c_r]$

$$E\left[\epsilon_C\right] = \prod_{i=1}^{n} E\left[\epsilon_i^{\#C_i}\right], \tag{19}$$

where $\#C_i$ is the i-th element of the index counting function, which counts how often the index i appears in C. All statements follow by plugging in the value of each factor implied by Assumption 1 (iii). □

One can now match the (co-)moments of the structural shocks with the (co-)moments of the unmixed innovations, which yields n variance conditions

$$E\left[e_{\mathcal{M}(2)}(A)\right] = E\left[\epsilon_{\mathcal{M}(r)}\right] \iff E\left[e_{\mathcal{M}(2)}(A)\right] - 1 = 0, \tag{20}$$

$\frac{n(n+1)}{2} - n$ covariance conditions

$$E\left[e_{\mathcal{C}(2)}(A)\right] = E\left[\epsilon_{\mathcal{C}(2)}\right] \iff E\left[e_{\mathcal{C}(2)}(A)\right] = 0, \tag{21}$$

$\frac{n(n+1)(n+2)}{6} - n$ coskewness conditions

$$E\left[e_{\mathcal{C}(3)}(A)\right] = E\left[\epsilon_{\mathcal{C}(3)}\right] \iff E\left[e_{\mathcal{C}(3)}(A)\right] = 0 \tag{22}$$

and $\frac{n(n+1)(n+2)(n+3)}{24} - n$ cokurtosis conditions

$$E\left[e_{\mathcal{C}(4)}(A)\right] = E\left[\epsilon_{\mathcal{C}(4)}\right], \tag{23}$$

with $E\left[\epsilon_{\mathcal{C}(4)}\right]$ as defined in Proposition 1.

The SVAR will only be identified up to sign and permutations. Let $\mathcal{P}$ be the set containing all $n \times n$ signed permutation matrices. For any signed permutation matrix $P \in \mathcal{P}$, the shocks $\tilde{\epsilon} := P\epsilon$ and the mixing matrix $\tilde{A}_0 := PA_0$ generate the same reduced form shocks as the shocks ϵ and the mixing matrix A_0. This can easily be verified since $u = A_0^{-1}\epsilon = A_0^{-1}P^{-1}P\epsilon = \tilde{A}_0^{-1}\tilde{\epsilon}$. Moreover, since $\tilde{\epsilon}$ is only a signed permutation of ϵ, both vectors of shocks share the same dependence structure and hence the identification approach cannot identify the correct sign and permutation. Identification up to permutation is equivalent to the problem of labeling the structural shocks, see Lanne et al. (2017) or Gouriéroux et al. (2017). Labeling and thus attaching a meaning to the shocks cannot be done statistically, but remains to the researcher. Since the identification approach cannot identify the correct sign and permutation, I redefine the problem such that the indeterminacy of sign and permutation does no longer appear in the new identification problem. Obviously, the indeterminacy is only removed from the statistical side of the problem and the researcher still needs to label the shocks. Note that the indeterminacy of scaling the shocks is already excluded from the identification problem by assuming shocks with unit variance. Define the set of sign-permutation representatives analogously to the set guaranteeing global identification in Lanne et al. (2017) as

$$\mathcal{A}^* := \{A \in \mathcal{A} | \forall i,\, a_{ii} > 0 \text{ and } \forall i < j,\, |a_{ii}| \geq |a_{ji}|\}. \tag{24}$$

An element $A \in \mathcal{A}^*$ is called a unique sign-permutation representative if for any signed permutation matrix $P \in \mathcal{P}$ with $P \neq I$ it holds that $PA \notin \mathcal{A}^*$. The set $\mathcal{A}^*$ fulfills the following properties:

Proposition 2. *Almost all elements $A \in \mathcal{A}^*$ are unique sign-permutation representatives. For any matrix $A \in \mathcal{A}$ there exists at least one signed permutation matrix P with $PA \in \mathcal{A}^*$.*

Proof. An inner point $A \in \mathcal{A}^*$ satisfies that $\forall i < j$, $|a_{ii}| > |a_{ji}|$. Let $A \in \mathcal{A}^*$ be an inner point. For any $P \in \mathcal{P}$ with $P \neq I$ it holds that for $\tilde{A} := PA$ there exist indices $i < j$ with $|\tilde{a}_{ii}| < |\tilde{a}_{ji}|$ and thus $\tilde{A} \notin \mathcal{A}^*$. Therefore, an inner point of $\mathcal{A}^*$ is a unique sign-permutation representative. Only the boundary of $\mathcal{A}^*$ contains elements which are not unique sign-permutation representatives. However, the boundary of the n^2 dimensional manifold $\mathcal{A}^*$ has dimension $n^2 - 1$ and is thus a null set in $\mathcal{A}^*$. The second statement is trivial. □

Since A_0 almost surely has a unique representative in $\mathcal{A}^*$, I replace Assumption 2.

Assumption 3. *$A_0 \in \mathcal{A}^*$ is a unique sign-permutation representative.*

The following proposition is based on Comon (1994) and shows that the variance, covariance, coskewness and cokurtosis conditions globally identify the SVAR.

Proposition 3. *Let $\epsilon = A_0 u$ satisfy Assumption 1 and 3. For $A \in \mathcal{A}^*$ it holds that*

$$E \begin{bmatrix} e_{\mathcal{M}(2)}(A) - 1 \\ e_{\mathcal{C}(2)}(A) \\ e_{\mathcal{C}(3)}(A) \\ e_{\mathcal{C}(4)}(A) - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{bmatrix} = 0 \iff A = A_0, \tag{25}$$

with $E\left[\epsilon_{\mathcal{C}(4)}\right]$ as defined in Proposition 1.

Proof. Let $\tilde{A} \in \mathcal{A}^*$ solve the moment conditions. Let $\tilde{Q} := \tilde{A}A_0^{-1}$ and thus $\tilde{e} = \tilde{A}u = \tilde{Q}\epsilon$. Assumption 1 implies that $I = E\left[\epsilon\epsilon'\right]$ and since $\tilde{e}$ solves the variance and covariance condition, it follows that $\tilde{Q}$ is orthogonal. The coskewness and cokurtosis conditions then imply that all third and fourth order cross-cumulants of $\tilde{e}$ are zero. Assumption 1 ensures that the shocks are non-Gaussian and have finite moments up to order four. One can thus apply Comon (1994) Theorem 16 and Comon (1994) equation (3.10), which yields that Q is in $\mathcal{P}$ and thus $\tilde{A}$ is a signed permutation of A_0. With Assumption 3, it follows that $\tilde{A} = A_0$. The other direction is trivial. □

If only the first [second] part of Assumption 1 (iv) is fulfilled, the cokurtosis [coskewness] conditions can be dropped and the SVAR is still globally identified. However, even if the cokurtosis [or alternatively the coskewness] conditions are dropped, there are still more moment conditions than unknown parameters. Importantly, dropping additional moment conditions immediately destroys the global identification result, see Appendix B. Lanne and Luoto (2019) basically identify the simultaneous interaction with a subsystem of the moment conditions used in Proposition 3. In particular, their system contains the variance, covariance and a subset of the cokurtosis conditions. Therefore, their system is only locally identified. Without the global identification result, it is difficult to derive the consistency and asymptotic normality of the estimator. In fact, the authors argue that the asymptotic properties can be derived under standard assumptions. However, one of these standard assumptions is a globally identified system and hence it is not obvious why the estimator proposed by Lanne and Luoto (2019) should satisfy the claimed asymptotic properties.

With Proposition 3, the matrix A_0 can be estimated by the SVAR-GMM estimator

$$\hat{A}_T(W) := \arg \min_{A \in \mathcal{A}^*} J_T(A, W), \tag{26}$$

where W is a positive semidefinite weighting matrix and

$$J_T(A, W) := \begin{pmatrix} E_T\left[e_{\mathcal{M}(2)}(A)\right] - 1 \\ E_T\left[e_{\mathcal{C}(2)}(A)\right] \\ E_T\left[e_{\mathcal{C}(3)}(A)\right] \\ E_T\left[e_{\mathcal{C}(4)}(A)\right] - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{pmatrix}' W \begin{pmatrix} E_T\left[e_{\mathcal{M}(2)}(A)\right] - 1 \\ E_T\left[e_{\mathcal{C}(2)}(A)\right] \\ E_T\left[e_{\mathcal{C}(3)}(A)\right] \\ E_T\left[e_{\mathcal{C}(4)}(A)\right] - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{pmatrix}. \tag{27}$$

If the estimator only contains the variance, covariance and coskewness conditions, it is denoted as the SVAR-GMM estimator based on the coskewness. The SVAR-GMM estimator based on the cokurtosis is defined analogously. The asymptotic properties of the estimator follow from standard arguments. Given standard assumptions (which in particular include global identification ensured by Proposition 3) the estimator $\hat{A}_T$ is a consistent estimator for A_0 and it is asymptotically normally distributed with $\sqrt{T}\left(vec(\hat{A}_T) - vec(A_0)\right) \xrightarrow{d} N(0, MSM')$, where M and S are defined as usual, see Hall (2005). Additionally, the weighting matrix $W = S^{-1}$ yields the estimator with the minimum asymptotic variance, see Hall (2005). The two-step SVAR-GMM estimator denotes the SVAR-GMM estimator under the standard two-step GMM procedure. Moreover, parameter hypothesis tests can be performed as usual, see Hall (2005). Of course, the interpretation of a test is always subject to a specific labeling, see Lanne et al. (2017). Bonhomme and Robin (2009) note that the asymptotic standard errors depend on the variances of third-order and fourth-order moments. In small samples these moments are difficult to estimate and the authors suggest to use bootstrap based confidence intervals instead of the estimated asymptotic standard errors.

To summarize, if the structural shocks are independent and non-Gaussian, the identification problem can be solved by extending the identification approach from covariance to coskewness and cokurtosis conditions. The SVAR can then be estimated by the SVAR-GMM estimator. However, the number of moment conditions used in Proposition 3 increases quickly with the dimension of the SVAR. Therefore, the computational burden of the SVAR-GMM estimator can be high in large models. The next section shows how an asymptotically not optimal weighting scheme can be used to derive a computationally cheap approximation of the SVAR-GMM estimator. Using an asymptotically not optimal weighting matrix leads to an increase of the asymptotic variance of the estimator, however, it has no impact on the global identification or consistency of the estimator. Thus in large SVAR models, one can trade asymptotic efficiency against a computationally cheap approximation.

6 Fast SVAR-GMM estimator

This section derives the fast SVAR-GMM estimator, which is a consistent, globally identified and computationally cheap estimator. I show that the fast SVAR-GMM estimator is approximately equal to the SVAR-GMM estimator proposed in Section 5 when a specific and asymptotically not optimal weighting matrix is used. Therefore, the asymptotic efficiency of the estimator can be sacrificed to reduce the computational burden. In the next section, I present evidence that the asymptotic efficiency loss is rather small. Additionally, in small samples the fast SVAR-GMM estimator often even performs better than the two-step SVAR-GMM estimator.

The fast SVAR-GMM estimator is a whitened estimator, meaning that it yields unmixed innovations, which by construction are uncorrelated and have unit variance. I first derive the whitened SVAR-GMM estimator and then show how a specific weighting matrix allows to decrease the computational costs of the estimator. Let $\mathcal{O}$ denote the set of orthogonal matrices and let $VV' = E[uu']$ be the Cholesky decomposition of the variance-covariance matrix of the reduced form shocks. Then define the whitened set of sign-permutation representatives as

$$\mathcal{A}_V^* := \{A \in \mathcal{A}^* | AV \in \mathcal{O}\}. \tag{28}$$

The set is constructed such that the unmixed innovations $e(A)$ with $A \in \mathcal{A}_V^*$ always fulfill the variance and covariance conditions in Proposition 3. Moreover, if Assumptions 1 and 3 are fulfilled, it holds that A_0 is a unique sign-permutation representative in $\mathcal{A}_V^*$. The SVAR is then globally identified with the following proposition based on Comon (1994).

Proposition 4. *Let $\epsilon = A_0 u$ satisfy Assumption 1 and 3. For $A \in \mathcal{A}_V^*$ it holds that*

$$E\begin{bmatrix} e_{\mathcal{C}(3)}(A) \\ e_{\mathcal{C}(4)}(A) - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{bmatrix} = 0 \iff A = A_0, \tag{29}$$

with $E\left[\epsilon_{\mathcal{C}(4)}\right]$ as defined in Proposition 1.

Proof. Assumption 1 ensures that the shocks are non-Gaussian and have finite moments up to order four. The coskewness and cokurtosis conditions imply that all third and fourth order cross-cumulants of $e(A)$ are equal to zero. The equivalence then follows from Comon (1994) Theorem 16 and Comon (1994) equation (3.10). □

Define the whitened SVAR-GMM estimator as

$$\hat{A}_T^V(W) := \arg \min_{A \in \mathcal{A}_V^*} K_T(A, W), \tag{30}$$

with a positive semidefinite weighting matrix W and

$$K_T(A,W) := \begin{pmatrix} E_T\left[e_{\mathcal{C}(3)}(A)\right] \\ E_T\left[e_{\mathcal{C}(4)}(A)\right] - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{pmatrix}' W \begin{pmatrix} E_T\left[e_{\mathcal{C}(3)}(A)\right] \\ E_T\left[e_{\mathcal{C}(4)}(A)\right] - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{pmatrix}. \tag{31}$$

The consistency of the estimator again follows from standard arguments. In practice, the matrix V is replaced by its consistent estimator $\hat{V}\hat{V}' = E_T\left[uu'\right]$, which is the Cholesky decomposition of the sample variance-covariance matrix of the reduced form shocks.

In comparison to the SVAR-GMM estimator proposed in Section 5, the whitened SVAR-GMM estimator does not require to calculate variance or covariance conditions, but it instead requires to optimize subject to the orthogonality constraints embedded in the set $\mathcal{A}_V^*$. In practice, it is not necessary to keep track of these constraints. Note that the whitened SVAR-GMM estimator can be calculated as

$$\hat{A}_T^V(W) = \left(\arg \min_{O \in \mathcal{O}_V^*} K_T(OV^{-1}, W) \right) V^{-1}, \tag{32}$$

with $\mathcal{O}_V^* = \{O \in \mathcal{O} | OV^{-1} \in \mathcal{A}_V^*\}$. An optimization problem of the form

$$\min_{O \in \mathcal{SO}} f(O) \tag{33}$$

can be pulled back to

$$\min_{s \in \mathfrak{so}} f(exp(s)), \tag{34}$$

where $\mathcal{SO}$ denotes the set of special orthogonal matrices, $\mathfrak{so}$ denotes the set of skew-symmetric matrices and $exp(.)$ is the matrix exponential, see e.g. Lezcano-Casado and Martínez-Rubio (2019). Therefore, instead of optimizing over the set of special orthogonal matrices, one can optimize over the set of skew-symmetric matrices, which is computationally less demanding. In this case, assuming that $O_0 := A_0 V$ is a special orthogonal matrix is not very restrictive, since by construction O_0 is an orthogonal matrix and one can find a singed permutation matrix $P \in \mathcal{P}$ such that PO_0 is a special orthogonal matrix.

The whitened SVAR-GMM estimator minimizes the dependency of the unmixed innovations measured by the weighted coskewness and cokurtosis conditions. It thus contains $q_3 := \frac{n(n+1)(n+2)}{6} - n$ coskewness conditions and $q_4 := \frac{n(n+1)(n+2)(n+3)}{24} - n$ cokurtosis conditions. When a specific weighting matrix is used, this minimization problem is equivalent to a computationally less demanding maximization problem involving only $2n$ instead of $q_3 + q_4$ moments. The equivalence between both optimization problems is related to a fundamental idea of ICA; minimizing the dependence means to maximize the non-Gaussianity. Define the $q_3 + q_4 \times q_3 + q_4$ dimensional fast weighting matrix W^{fast} as the diagonal matrix where the diagonal element corresponding to the co-moment $C \in \mathcal{C}(r)$ with $r \in \{3,4\}$ is defined as

$$w_r(C) := \binom{r}{\#C} = \frac{r!}{\prod_{i=1}^{n} \#C_i!}, \tag{35}$$

where $\#C_i$ denotes the i-th element of the counting function $\#C$. In a bivariate SVAR, the objective function of the whitened SVAR-GMM estimator $\hat{A}_T^{\hat{V}}(W^{fast})$ is equal to

$$\begin{aligned}
K_T(A, W^{fast}) =& \frac{3!}{2!1!} E_T\left[e_1(A)^2 e_2(A)\right]^2 && (36)\\
&+\frac{3!}{1!2!} E_T\left[e_1(A) e_2(A)^2\right]^2 && (37)\\
&+\frac{4!}{3!1!} E_T\left[e_1(A)^3 e_2(A)\right]^2 && (38)\\
&+\frac{4!}{2!2!} E_T\left[e_1(A)^2 e_2(A)^2 - 1\right]^2 && (39)\\
&+\frac{4!}{1!3!} E_T\left[e_1(A) e_2(A)^3\right]^2 . && (40)
\end{aligned}$$

This can easily be verified since for $\mathcal{C}(3) = \{[1,1,2],[1,2,2]\}$, the counting function yields $\#[1,1,2] = [2,1]$, $\#[1,2,2] = [1,2]$ and for $\mathcal{C}(4) = \{[1,1,1,2],[1,1,2,2],[1,2,2,2]\}$, the counting function yields $\#[1,1,1,2] = [3,1]$, $\#[1,1,2,2] = [2,2]$ and $\#[1,2,2,2] = [1,3]$. The fast weighting matrix W^{fast} can now be used to derive a computationally cheap expression for the estimator $\hat{A}_{T,r}^{\hat{V}}(W^{fast})$.

Proposition 5. *The estimator* $\hat{A}_T^{\hat{V}}(W^{fast})$ *is equal to*

$$\hat{A}_T^{\hat{V}}(W^{fast}) = arg\max_{A \in \mathcal{A}^*_{\hat{V}}} H_T(A), \tag{41}$$

with

$$H_T(A) := \sum_{M \in \mathcal{M}(3)} E_T\left[e_M(A)\right]^2 + \sum_{M \in \mathcal{M}(4)} \left(E_T\left[e_M(A)\right] - 3\right)^2 . \tag{42}$$

Proof. The weights are constructed, such that for a given sample of size T there exists a constant

$\Omega_T \in \mathbb{R}$ with $H_T(A) + K_T(A, W^{fast}) = \Omega_T$ for all $A \in \mathcal{A}^*_{\hat{V}}$, which is the sample analogue of equation (3.10) in Comon (1994). Rearranging yields $K_T(A, W^{fast}) = \Omega_T - H_T(A)$, which proves the Proposition. The proof is written down in more detail in Appendix A. $\square$

The whitened SVAR-GMM estimator with the fast weighting matrix can thus be calculated by equation (30) or by equation (41). Equation (30) minimizes the dependency of the unmixed innovations measured by the $q_3 + q_4$ squared coskewness and cokurtosis contentions. Equation (41) maximizes the non-Gaussianity of the unmixed innovations measured by the $2n$ squared skewness and excess kurtosis coefficients. If the fast weighting matrix is used, both optimization problems are equivalent. However, in large SVAR models the computation of the latter problem is considerably less demanding than the computation of the former problem. Proposition 5 thus yields a computationally cheap way to calculate $\hat{A}^{\hat{V}}_T(W^{fast})$, which is henceforth denoted as the fast SVAR-GMM estimator. Analogously to the SVAR-GMM estimator, one can define the fast SVAR-GMM estimator based on the skewness [or based on the kurtosis], which only contains the skewness [kurtosis] coefficients in equation (42) and the coskewness [cokurtosis] conditions in equation (30).

In the bivariate SVAR, the fast SVAR-GMM estimator $\hat{A}^{\hat{V}}_T(W^{fast})$ can thus be calculated as the maximum of the objective function $H_T(A)$ which is given by

$$H_T(A) = E_T\left[e_1(A)^3\right]^2 + E_T\left[e_2(A)^3\right]^2 + E_T\left[e_1(A)^4 - 3\right]^2 + E_T\left[e_2(A)^4 - 3\right]^2. \tag{43}$$

The fast SVAR-GMM estimator can be derived as the limiting case of the SVAR-GMM estimator proposed in Section 5, when a specific and asymptotically inefficient weighting matrix is used. Consider the SVAR-GMM estimator $\hat{A}_T(W_m)$, where the weighting matrix W_m puts a weight m on the variance and covariance conditions and uses the fast weighting matrix W^{fast} for the coskewness and cokurtosis conditions. The weighting matrix W_m is thus defined as

$$W_m := \begin{bmatrix} mI & 0 \\ 0 & W^{fast} \end{bmatrix}, \tag{44}$$

where I is an $\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}$ dimensional identity matrix and W^{fast} is the fast weighting matrix defined in equation (35). With the weighting matrix W_m, the objective function of the SVAR-GMM estimator $\hat{A}_T(W_m)$ is equal to

$$J_T(A, W_m) = \sum_{M \in \mathcal{M}(2)} m \left(E_T\left[e_M(A)\right] - 1\right)^2 + \sum_{C \in \mathcal{C}(2)} m \left(E_T\left[e_C(A)\right]\right)^2 \tag{45}$$

$$+ K_T(A, W^{fast}). \tag{46}$$

Therefore, the weighting matrix W_m splits the objective function of the SVAR-GMM estimator into two terms. The first term contains the variance and covariance conditions with the weight m. The second term is equal to the objective function of the fast SVAR-GMM estimator and contains the coskewness and cokurtosis conditions. The following proposition shows that when the weight m of the variance and covariance conditions tends to infinity, the SVAR-GMM estimator $\hat{A}_T(W_m)$ tends to the fast SVAR-GMM estimator $\hat{A}_T^{\hat{V}}(W^{fast})$.

Proposition 6.

$$\underset{m \to \infty}{lim}\ \hat{A}_T(W_m) = \hat{A}_T^{\hat{V}}(W^{fast}). \tag{47}$$

Proof. For $A \in \mathcal{A}^*$ it holds that

$$J_T(A, W_m) \overset{m \to \infty}{\to} \begin{cases} \infty & , \text{for } A \notin \mathcal{A}^*_{\hat{V}} \\ K_T(A, W^{fast}) & , \text{for } A \in \mathcal{A}^*_{\hat{V}} \end{cases}. \tag{48}$$

It thus follows that

$$\hat{A}_T(W_m) = \arg \min_{A \in \mathcal{A}^*} J_T(A, W_m) \overset{m \to \infty}{\to} \arg \min_{A \in \mathcal{A}^*_{\hat{V}}} K_T(A, W^{fast}) = \hat{A}^{\hat{V}}_T(W^{fast}). \tag{49}$$

□

For a sufficiently large weight m, the SVAR-GMM estimator can be approximated by the fast SVAR-GMM estimator. The weighting matrix W_m is asymptotically not optimal in terms of the asymptotic variance of the SVAR-GMM estimator. Therefore, in large SVAR models, when the computational burden of the SVAR-GMM estimator becomes too high, one can sacrifice the asymptotic efficiency and gain a computationally cheap approximation of the SVAR-GMM estimator by the fast SVAR-GMM estimator. The next section provides evidence that the fast SVAR-GMM estimator performs robustly in finite samples and that the efficiency loss in large samples is rather small.

7 Finite sample properties

This section analyzes the finite sample performance of the SVAR-GMM estimators and compares it to the GMM estimator proposed by Lanne and Luoto (2019) and to the PML estimator proposed by Gouriéroux et al. (2017). I show that the performance of the PML estimator deteriorates with the degree of misspecification, while the SVAR-GMM estimators perform more robustly throughout different specifications. In the empirical application in Section 8, I find structural shocks with non-zero skewness and a positive excess kurtosis. If the shocks are skewed and exhibit an excess kurtosis, the Monte Carlo study shows that the two-step SVAR-GMM estimator performs best and it clearly outperforms the PML estimator based on a t-distribution. Moreover, I find that the fast SVAR-GMM estimator performs well and robustly across specifications. It thus yields a viable, computationally cheap alternative to the asymptotically optimal two-step SVAR-GMM estimator. Additionally, the Monte Carlo simulation sheds light on the impact of the degree of non-Gaussianity on the finite sample performance of SVAR-GMM estimators. I find that estimators based on the skewness have desirable small sample properties, as the bias and standard deviation of the estimators is found to be almost solely determined by the relative skewness and shows no deterioration with a decreasing sample size.

The setup of the Monte Carlo study is similar to the setup in Gouriéroux et al. (2017) with $u = A_0^{-1}\epsilon$ and

$$A_0^{-1} = \begin{bmatrix} cos(\phi) & -sin(\phi) \\ sin(\phi) & cos(\phi) \end{bmatrix}, \tag{50}$$

where $\phi = -\pi/5$. The goal of the Monte Carlo study is to evaluate the performance of the estimators depending on the degree of non-Gaussianity and the sample size. The shocks are drawn from a distribution of the Pearson distribution system with mean zero, unit variance and different skewness/kurtosis parameters. The degree of non-Gaussianity is measured in two dimensions, by the relative skewness and relative excess kurtosis. A low relative skewness [excess kurtosis] is defined as the sample skewness [excess kurtosis] a standard normally distributed shock will not exceed in a sample of size T with a probability of 90%. A medium relative skewness [excess kurtosis] is defined analogously for a probability of 99% and a high relative skewness [excess kurtosis] is defined for a probability of 99.99%. The values are calculated by bootstrap and are shown in Table 1. Defining the non-Gaussianity in relative terms allows to compare the impact of the skewness and excess kurtosis. Moreover, it allows to disentangle the effects of the non-Gaussianity and the sample size.

Table 1: Relative non-Gaussianity

	rel. skewness			rel. excess kurtosis		
	low	medium	high	low	medium	high
$T = 200$	0.22	0.4	0.68	0.4	0.97	2.33
$T = 500$	0.14	0.26	0.41	0.27	0.6	1.25
$T = 5000$	0.04	0.08	0.13	0.09	0.17	0.3

The table shows the quantiles (low is 0.9*, medium is* 0.99*, high is* 0.9999*) of the sample skewness and sample excess kurtosis of standard normally distributed shocks in a sample of size T in a Monte Carlo simulation with* 5000000 *simulated samples for each sample size.*

The first Monte Carlo study contains three different specifications. In the first specification, both structural shocks have zero skewness and a high excess kurtosis. In the second specification, the first structural shock is Gaussian and the second shock has zero skewness and a high excess kurtosis. In the third specification, the first shock has a high skewness and zero excess kurtosis and the second shock has a high skewness and a high excess kurtosis. Table 2 shows the mean bias and standard deviation of different estimators. The first row contains the SVAR-GMM estimator with an identity weighting matrix, the second row shows the two-step SVAR-GMM estimator and the third row shows the fast SVAR-GMM estimator. The GMM estimator proposed by Lanne and Luoto (2019) is shown in the fourth row and uses the cokurtosis condition $E\left[\epsilon_1\epsilon_2^3\right] = 0$. The PML estimator proposed by Gouriéroux et al. (2017) is shown in the fifth row and assumes a t-distribution with twelve degrees of freedom. The results are largely unaffected by the chosen degree of freedom.

In the first specification, the PML estimator performs better than its competitors. This is not too surprising, since the PML estimator is essentially correctly specified (the shocks are drawn from a Pearson Type VII distribution, which contains the t-distribution). However, the advantage of the PML estimator compared to the fast SVAR-GMM estimator is small and vanishes with an increasing sample size. In the second specification, the GMM estimator proposed by Lanne and Luoto (2019) performs best, but is closely followed by the fast SVAR-GMM estimator. In the application presented in Section 8, the reduced form shocks are skewed and have a positive excess kurtosis. Thus the shocks simulated in the third specification are closest to the shocks in the application. In the third specification, the two-step SVAR-GMM estimator outperforms its competitors. Neither the GMM estimator proposed by Lanne and Luoto (2019) nor the PML estimator can exploit the non-Gaussianity introduced by the skewed structural shocks. However, the advantage of the two-step SVAR-GMM estimator compared to the fast SVAR-GMM estimators is again small. Overall, each estimator has its strength and weaknesses in different specifications. In the specification closest to the application presented in Section 8, the two-step SVAR-GMM estimator performs best. The fast SVAR-GMM estimator yields the most robust performance across specifications and sample sizes. The fast SVAR-GMM estimator is asymptotically inefficient, however, in small samples it often performs better than the two-step SVAR-GMM estimator. This finding can be explained by the fact that estimating the optimal weighting matrix requires to estimate the variances of third-order and fourth-order moments, which is a difficult task in small samples, compare to Bonhomme and Robin (2009).

Table 2: Finite sample performance

	$T=200$			$T=500$			$T=5000$		
Specification	1	2	3	1	2	3	1	2	3
$\hat{A}(I)$	-0.03 (0.14)	-0.04 (0.18)	-0.02 (0.11)	-0.02 (0.11)	-0.03 (0.15)	-0.02 (0.1)	-0.01 (0.07)	-0.02 (0.13)	-0.01 (0.08)
$\hat{A}(W^{\text{2-step}})$	-0.04 (0.18)	-0.04 (0.18)	-0.02 (0.1)	-0.02 (0.13)	-0.03 (0.15)	-0.01 (0.09)	-0.01 (0.08)	-0.01 (0.11)	-0.01 (0.07)
$\hat{A}^{\hat{V}}(W^{fast})$	-0.02 (0.14)	-0.03 (0.16)	-0.02 (0.12)	-0.01 (0.11)	-0.02 (0.14)	-0.01 (0.11)	-0.01 (0.07)	-0.01 (0.1)	-0.01 (0.08)
GMM_{LL}	-0.03 (0.15)	-0.03 (0.14)	-0.03 (0.14)	-0.03 (0.13)	-0.02 (0.12)	-0.02 (0.12)	-0.02 (0.11)	-0.01 (0.09)	-0.01 (0.1)
PML	-0.02 (0.13)	-0.04 (0.17)	-0.04 (0.18)	-0.01 (0.09)	-0.03 (0.15)	-0.03 (0.17)	-0.01 (0.07)	-0.02 (0.14)	-0.02 (0.14)

Monte Carlo simulation with sample sizes 200, 500 and 5000 each with 10000 iterations. For an estimator $\hat{A}$ of A_0, define the estimator $\hat{B} := \hat{A}^{-1}$ of $B := A_0^{-1}$. Each entry shows the $\underset{(\text{standard deviation})}{\text{mean bias}}$ of the estimated element $\hat{b}_{1,1}$, which is the upper left element of $\hat{B}$. The mean bias is calculated as $E\left[\hat{b}_{1,1} - b_{1,1}\right]$ and the standard deviation is calculated as the square root of $E\left[\left(\hat{b}_{1,1} - b_{1,1}\right)^2\right]$, where $b_{1,1} = cos(-\pi/5)$ is the upper left element of B. The distribution of the shocks in a given specification are explained in the text and use the relative skewness and relative excess kurtosis values shown in Table 1. The SVAR-GMM estimator denoted by $\hat{A}(I)$ uses an identity weighting matrix, the estimator $\hat{A}(W^{\text{2-step}})$ is the two-step SVAR-GMM estimator and the fast SVAR-GMM estimator is denoted by $\hat{A}^{\hat{V}}(W^{fast})$. The GMM estimator proposed by Lanne and Luoto (2019) is denoted by GMM_{LL} and uses the cokurtosis condition $E\left[\epsilon_1 \epsilon_2^3\right] = 0$. The PML estimator proposed by Gouriéroux et al. (2017) is denoted by PML and assumes a t-distribution with twelve degrees of freedom.

In a second Monte Carlo simulation I analyze the impact of the degree of non-Gaussianity and the sample size on the SVAR-GMM estimators. I find that the results do not depend on the weighting matrix and I thus only report the results for the fast SVAR-GMM estimator. The first part of Table 3 shows the impact of the relative skewness on the fast SVAR-GMM estimators based on coskewness conditions and the second part of the table shows the impact of the relative excess kurtosis on the fast SVAR-GMM estimators based on cokurtosis conditions. In the first part of the table the structural shocks have zero excess kurtosis and a low, medium or high relative skewness. In the second part of the table the shocks have zero skewness and a low, medium or high relative excess kurtosis.

Unsurprisingly, an increase in the degree of non-Gaussianity has a positive impact on the finite sample properties of the estimators. Intuitively, this finding is comparable to the strength or weakness of an instrument in an instrumental variables estimation. In the largest simulated sample, the impact of the relative skewness on the estimator based on the coskewness and the impact of the relative excess kurtosis on the estimator based on the cokurtosis are almost identical. However, decreasing the sample size reveals an important difference between both estimators. The bias and standard deviation of the estimator based on the coskewness appear to be entirely determined by the relative skewness and do not vary across sample sizes. In contrast to that, the bias and standard deviation of the estimator based on the cokurtosis increase with a decreasing sample size. This finding can be explained by the sample variance of the moment conditions, which increases with an increase of the excess kurtosis. The effect is more pronounced in small samples and partly offsets the positive impact of a higher excess kurtosis.

Table 3: Finite sample performance: The impact of non-Gaussianity

	$T = 200$			$T = 500$			$T = 5000$		
Skewness	low	med	high	low	med	high	low	med	high
Exc. kurtosis	zero	zero	zero	zero	zero	zero	zero	zero	zero
$\hat{A}^{\hat{V}}_{r=3}(W^{fast})$	-0.05 (0.2)	-0.02 (0.13)	-0.01 (0.07)	-0.05 (0.21)	-0.02 (0.13)	-0.01 (0.07)	-0.05 (0.21)	-0.02 (0.13)	-0.01 (0.07)
Skewness	zero	zero	zero	zero	zero	zero	zero	zero	zero
Exc. kurtosis	low	med	high	low	med	high	low	med	high
$\hat{A}^{\hat{V}}_{r=4}(W^{fast})$	-0.06 (0.23)	-0.04 (0.18)	-0.02 (0.14)	-0.05 (0.21)	-0.03 (0.16)	-0.01 (0.11)	-0.04 (0.2)	-0.02 (0.13)	-0.01 (0.07)

Monte Carlo simulation with sample sizes 200, 500 and 5000 each with 10000 iterations. For an estimator $\hat{A}$ of A_0, define the estimator $\hat{B} := \hat{A}^{-1}$ of $B := A_0^{-1}$. Each entry shows the $\underset{(standard\ deviation)}{mean\ bias}$ of the estimated element $\hat{b}_{1,1}$, which is the upper left element of $\hat{B}$. The mean bias is calculated as $E\left[\hat{b}_{1,1} - b_{1,1}\right]$ and the standard deviation is calculated as the square root of $E\left[\left(\hat{b}_{1,1} - b_{1,1}\right)^2\right]$, where $b_{1,1} = cos(-\pi/5)$ is the upper left element of B. The distributions of the shocks in a given specification are explained in the text and use the relative skewness and relative excess kurtosis values shown in Table 1. The fast SVAR-GMM estimator based on the skewness is denoted by $\hat{A}^{\hat{V}}_{r=3}(W^{fast})$ and the fast SVAR-GMM estimator based on the kurtosis is denoted by $\hat{A}^{\hat{V}}_{r=4}(W^{fast})$.

To sum up, I find that the SVAR-GMM estimators perform more robustly than the PML estimator proposed by Gouriéroux et al. (2017) and the GMM estimator proposed by Lanne and Luoto

(2019). In particular when the structural shocks exhibit a non-zero skewness, the performance of the SVAR-GMM estimators is superior to the performance of the two alternatives. Moreover, the Monte Carlo study reveals some desirable small sample properties of estimators based on the skewness. Last but not least, I find that the fast SVAR-GMM estimator performs well and that the efficiency loss compared to the two-step SVAR-GMM estimator is small. Moreover, in small samples the fast SVAR-GMM estimator often performs better than the two-step SVAR-GMM estimator.

8 Economic activity, oil and stock prices

This section applies the SVAR-GMM estimator to analyze the simultaneous relationship between economic activity, oil and stock prices. It might be convincing to argue that the real economy behaves sluggishly and does not respond to stock and oil price shocks contemporaneously. However, there aren't to many good arguments how to contemporaneously restrict the stock and oil market. Nevertheless, in applications (e.g. Sadorsky (1999), Kilian and Park (2009) or Apergis and Miller (2009)) stock market shocks have been restricted to have no simultaneous impact on oil prices. I provide evidence that the real economy might indeed behave sluggishly, but oil and stock prices do not.

The SVAR is estimated with monthly US data from 1990 to 2018 and contains three variables - a measure of real economic activity (EA), monthly real S&P 500 returns (SP) and the monthly growth rates of real oil prices (OP). Real economic activity is measured as 100 times the log difference of the monthly US Industrial Production Index. Real S&P 500 returns are calculated as 100 times the log difference of the the S&P 500 closing price deflated by the US CPI. The monthly growth rates of real oil prices are calculated as 100 times the log difference of the crude oil composite acquisition cost by refiners deflated by the US CPI. See Appendix D for more information on the data sources.

The SVAR is given by

$$A_0 \begin{bmatrix} EA_t \\ SP_t \\ OP_t \end{bmatrix} = \alpha + \sum_{i=1}^{p} A_i \begin{bmatrix} EA_{t-i} \\ SP_{t-i} \\ OP_{t-i} \end{bmatrix} + \begin{bmatrix} \epsilon_t^{EA} \\ \epsilon_t^{SP} \\ \epsilon_t^{OP} \end{bmatrix}. \tag{51}$$

Based on the AIC criterion, I estimate the reduced form with a lag length of $p = 4$ months. The moments of the residuals and the Jarque-Bera test for normality are shown in Table 4. The Jarque-Bera test indicates non-Gaussian residuals with non-zero skewness and a positive excess kurtosis. Based on the Monte Carlo simulation, I use the two step SVAR-GMM estimator

Table 4: Reduced form residuals

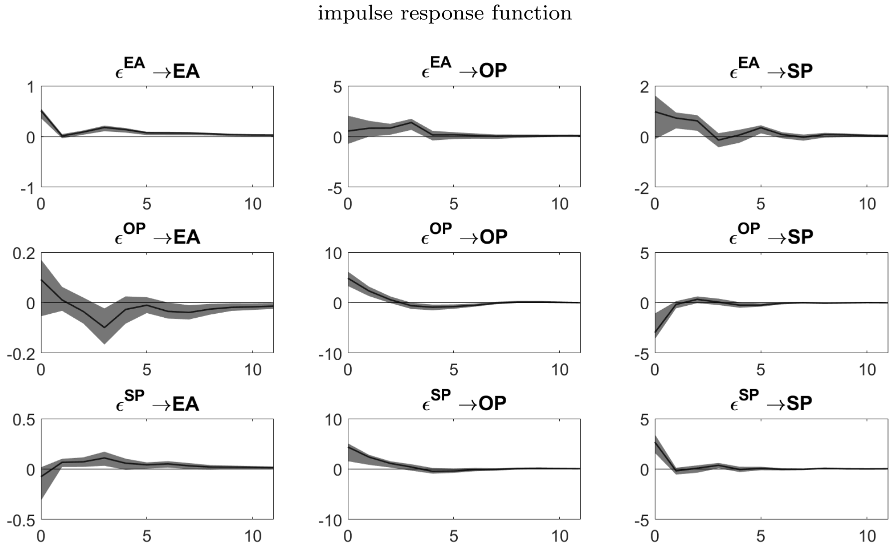
	Variance	Skewness	Kurtosis	JB-Test
u^{EA}	0.28	-0.84	10.71	0.00
u^{OP}	44.49	0.14	3.73	0.02
u^{SP}	15.53	-0.37	3.55	0.01

The JB-Test shows the p-value of the Jarque-Bera test for normality.

to estimate the simultaneous relationship. The results presented below are robust to different specifications and estimators, see Appendix D.

The impulse response function (IRF) is shown in Figure 1. The shocks are labeled such that economic activity shocks have a positive impact on oil and stock prices, oil price shocks have a long-run negative impact on economic activity, and the remaining shock is labeled as the stock market shock. According to the IRF, oil price shocks lead to a lagged decrease of economic activity and to an immediate decrease of stock returns. Stock market shocks lead to a long-run increase of economic activity and to an immediate increase of oil prices. Therefore, real economic activity behaves sluggishly with no simultaneous response to stock or oil market shocks. However, the stock and oil market are found to interact simultaneously and no recursive order of both variables is viable.

Figure 1: Impulse Response Function



Confidence intervals are calculated by bootstrap with 1000 *replications and the interval show the upper* 0.9 *and lower* 0.1 *percentiles. The reduced form VAR is estimated with four lags and the simultaneous interaction is estimated by the two step SVAR-GMM estimator.*

9 Conclusion

This paper proposes an identification approach based on higher moments, which is derived as a straightforward extension of the usual SVAR identification approach relying on uncorrelated shocks to independent shocks. Exploiting the skewness and excess kurtosis of the shocks allows to identify a non-Gaussian SVAR with independent structural shocks. The identification result is used to derive an SVAR-GMM estimator. The SVAR-GMM estimator becomes computationally expensive in large SVAR models. I show that by sacrificing the asymptotic efficiency of the SVAR-GMM estimator, one can derive a consistent, globally identified and computationally cheap approximation, which is denoted as the fast SVAR-GMM estimator.

In a Monte Carlo simulation, I show that even though the fast SVAR-GMM estimator is asymptotically inefficient, it performs well and robustly across different specifications. In the simulation closest to the empirical application, I find that the two step SVAR-GMM estimator performs best. In particular, the two step SVAR-GMM estimator outperforms the GMM estimator proposed by

Lanne et al. (2017) and the PML estimator with a t-distribution proposed by Gouriéroux et al. (2017). The advantage of the SVAR-GMM estimator is related to its ability to use information contained in the skewness. Additionally, the Monte Carlo simulation reveals desirable small sample properties of estimators based on the skewness.

Finally, the empirical applications analyzes the interaction between real economic activity, stock and oil markets. I find that stock and oil prices interact simultaneously, while the real economy appears to behave sluggish with no contemporaneous reaction to oil and stock market shocks. The application thus illustrates how an SVAR can be estimated based on higher moments without relying on incredible short-run restrictions.

A Appendix - Proofs

Let x be a an n-dimensional random variable with $E[x] = 0$ and $E[xx'] = I$. For $C = [c_1, c_2, c_3] \in \{1, ..., n\}^3$ the third order (cross-)cumulant of x is equal to

$$Cum(x_C) = Cum(x_{c_1}, x_{c_2}, x_{c_3}) = E[x_{c_1} x_{c_2} x_{c_3}]. \tag{52}$$

For $c = [c_1, c_2, c_3, c_4] \in \{1, ..., n\}^4$ the fourth order (cross-)cumulant of x is equal to

$$\begin{aligned} Cum(x_C) = Cum(x_{c_1}, x_{c_2}, x_{c_3}, x_{c_4}) =& E[x_{c_1} x_{c_2} x_{c_3} x_{c_4}] & (53) \\ &- E[x_{c_1} x_{c_2}] E[x_{c_3} x_{c_4}] & (54) \\ &- E[x_{c_1} x_{c_3}] E[x_{c_2} x_{c_4}] & (55) \\ &- E[x_{c_1} x_{c_4}] E[x_{c_2} x_{c_3}]. & (56) \end{aligned}$$

Consider a sample of the random variable x for which $E_T[x] = 0$ and $E_T[xx'] = I$, then the same equalities hold for the sample analogue $Cum_T(.)$ and $E_T(.)$.

Lemma 1. *Let x be an n-dimensional random variable with $E\left[x_{\mathcal{M}(1)}\right] = 0$, $E\left[x_{\mathcal{M}(2)}\right] = 1$, $E\left[x_{\mathcal{C}(2)}\right] = 0$ and $E\left[x_{\mathcal{M}(3)}\right] < \infty$. Let ϵ be as defined in Assumption 1.*

1) For $c \in \mathcal{C}(3)$ it holds that $E[x_c] = Cum(x_c)$.

2) For $c \in \mathcal{C}(4)$ it holds that $E[x_c] - E[\epsilon_c] = Cum(x_c)$.

3) For $m \in \mathcal{M}(3)$ it holds that $E[x_m] = Cum(x_m)$.

4) For $m \in \mathcal{M}(4)$ it holds that $E[x_m] - 3 = Cum(x_m)$.

Consider a sample of the random variable x with $E_T\left[x_{\mathcal{M}(1)}\right] = 0$, $E_T\left[x_{\mathcal{M}(2)}\right] = 1$, $E_T\left[x_{\mathcal{C}(2)}\right] = 0$ and $E_T\left[x_{\mathcal{M}(3)}\right] < \infty$, then the same statements hold for the sample analogue $E_T[.]$ and $Cum_T(.)$.

Proof. Statements 1) and 3) are trivial. Statement 2) holds since for $C = [c_1, c_2, c_3, c_4] \in \mathcal{C}(4)$

$$E[\epsilon_c] = \begin{cases} 1 & \text{, if } c_1 = c_2 \text{ and } c_3 = c_4 \\ 0 & \text{, else} \end{cases} \tag{57}$$

and

$$Cum(x_c) = \begin{cases} E[x_c] - 1 & \text{, if } c_1 = c_2 \text{ and } c_3 = c_4 \\ E[x_c] & \text{, else} \end{cases}. \tag{58}$$

Statement 4) holds since for $C = [c_1, c_2, c_3, c_4] \in \mathcal{M}(4)$

$$Cum(x_c) = E[x_c] - 3 \tag{59}$$

The sample analogue can be shown analogously. □

Proof of Proposition 5. Define $\mathcal{O}^*_{\hat{V}} = \{O \in \mathcal{O} | O\hat{V}^{-1} \in \mathcal{A}^*_{\hat{V}}\}$ and note that

$$\arg\max_{A \in \mathcal{A}^*_{\hat{V}}} H_T(A) = \left(\arg\max_{O \in \mathcal{O}^*_{\hat{V}}} \tilde{H}_T(O) \right) \hat{V}^{-1}, \tag{60}$$

with $\tilde{H}_T(O) = H_T(O\hat{V}^{-1})$ and

$$\arg\min_{A \in \mathcal{A}^*_{\hat{V}}} K_T(A, W) = \left(\arg\min_{O \in \mathcal{O}^*_{\hat{V}}} \tilde{K}_T(O, W) \right) \hat{V}^{-1}, \tag{61}$$

with $\tilde{K}_T(O, W) = K_T(O\hat{V}^{-1}, W)$. Therefore, Proposition 5 requires to show that

$$\arg\min_{O \in \mathcal{O}^*_{\hat{V}}} \tilde{K}_T(O, W^{fast}) = \arg\max_{O \in \mathcal{O}^*_{\hat{V}}} \tilde{H}_T(O). \tag{62}$$

For $O \in \mathcal{O}^*_{\hat{V}}$ define

$$\tilde{e}(O) := e(O\hat{V}^{-1}) = O\hat{V}^{-1}u = O\hat{V}^{-1}A_0^{-1}\epsilon = O\tilde{\epsilon}, \tag{63}$$

with $\hat{V}^{-1}A_0^{-1}\epsilon = \tilde{\epsilon}$ and by construction $I = E_T\left[\tilde{\epsilon}\tilde{\epsilon}'\right]$. Then the objective function $\tilde{K}_T(O, W^{fast})$ can be written as

$$\tilde{K}_T(O, W^{fast}) = \sum_{C \in \mathcal{C}(3)} \binom{3}{\#C} E_T\left[\tilde{e}_C(O)\right]^2 + \sum_{C \in \mathcal{C}(4)} \binom{4}{\#C} \left(E_T\left[\tilde{e}_C(O)\right] - E\left[\epsilon_C\right]\right)^2 \tag{64}$$

and the objective function $\tilde{H}_T(O)$ can be written as

$$\tilde{H}_T(O) = \sum_{M \in \mathcal{M}(3)} \binom{3}{\#M} E_T\left[\tilde{e}_M(O)\right]^2 + \sum_{M \in \mathcal{M}(4)} \binom{4}{\#M} \left(E_T\left[\tilde{e}_M(O) - 3\right]\right)^2, \tag{65}$$

since $\binom{r}{\#M} = 1$ for $M \in \mathcal{M}(r)$ and $r \in \{3, 4\}$. With Lemma 1 it follows that

$$\tilde{K}_T(O, W^{fast}) = \sum_{C \in \mathcal{C}(3)} \binom{3}{\#C} Cum_T(\tilde{e}_C(O))^2 + \sum_{C \in \mathcal{C}(4)} \binom{4}{\#C} Cum_T(\tilde{e}_C(O))^2 \tag{66}$$

and

$$\tilde{H}_T(O) = \sum_{M \in \mathcal{M}(3)} \binom{3}{\#M} Cum_T(\tilde{e}_M(O))^2 + \sum_{M \in \mathcal{M}(4)} \binom{4}{\#M} Cum_T(\tilde{e}_M(O))^2. \tag{67}$$

The weights are constructed such that

$$\tilde{K}_T(O, W^{fast}) + \tilde{H}_T(O) = \sum_{s_1,\dots,s_3=1}^{n} Cum_T(\tilde{e}_{s_1}(O), ..., \tilde{e}_{s_3}(O))^2 \tag{68}$$

$$+ \sum_{s_1,\dots,s_4=1}^{n} Cum_T(\tilde{e}_{s_1}(O), ..., \tilde{e}_{s_4}(O))^2. \tag{69}$$

Equation (3.10) in Comon (1994) states that on the pollination level, the right and side of the equation (68) is invariant with respect to $O \in \mathcal{O}$. However, the same statement also holds for the sample analogue and thus the right hand side of the equation (68) is invariant with respect to $O \in \mathcal{O}$. Let $\Omega_T := \sum_{s_1,...,s_3=1}^{n} Cum_T(\tilde{e}_{s_1}(O), ..., \tilde{e}_{s_3}(O))^2 + \sum_{s_1,...,s_4=1}^{n} Cum_T(\tilde{e}_{s_1}(O), ..., \tilde{e}_{s_4}(O))^2$ and thus

$$\tilde{K}_T(O, W^{fast}) = \Omega_T - \tilde{H}_T(O). \tag{70}$$

Therefore equation (62) holds, which proves the proposition. □

B Appendix - Notes on identification

This section uses a bivariate example to illustrate why global identification up to sign and permutation requires to include all coskewness conditions. Analogously, one can construct an example based on cokurtosis conditions and show that including only $n(n-1)/2$ cokurtosis conditions (as proposed by Lanne and Luoto (2019)) does not globally identify the SVAR.

Let $E[\epsilon_i] = 0$, $E[\epsilon_i^2] = 1$ and $E[\epsilon_i^3] = 1$ for $i \in \{1, 2\}$ and let

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} = A_0 \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \tag{71}$$

with $A_0 = I$. Define the unmixed innovations as

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} 1 & a_{1,2} \\ a_{2,1} & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{72}$$

To simplify calculations, the moment condition $E\left[e^2\right] = 1$ has been replaced by the assumption $a_{1,1} = a_{2,2} = 1$. The covariance and coskewness conditions are given by

$$0 = E\left[e_1 e_2\right] \tag{73}$$

$$0 = E\left[e_1^2 e_2\right] \tag{74}$$

$$0 = E\left[e_1 e_2^2\right]. \tag{75}$$

The system thus contains three equations in two unknowns (e_1 and e_2 are functions of $a_{1,2}$ and $a_{2,1}$). Omitting one of the coskewness conditions leads to a system not identifying A_0 up to sign and permutations. Consider the system of two equations and two unknowns

$$0 = E\left[e_1 e_2\right] \tag{76}$$

$$0 = E\left[e_1^2 e_2\right]. \tag{77}$$

The first equation yields

$$0 = E\left[e_1 e_2\right] = E\left[(\epsilon_1 + a_{1,2}\epsilon_2)(a_{2,1}\epsilon_1 + \epsilon_2)\right] \tag{78}$$

$$\implies 0 = a_{2,1} + a_{1,2}. \tag{79}$$

The second equation yields

$$0 = E\left[e_1^2 e_2\right] = E\left[(\epsilon_1 + a_{1,2}\epsilon_2)^2(a_{2,1}\epsilon_1 + \epsilon_2)\right] \tag{80}$$

$$\implies 0 = a_{2,1} + a_{1,2}^2. \tag{81}$$

Therefore, $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ solves the system and thus $A_0 = I$ is not identified up to sign and permutations.

C Appendix - Notes on PML

Also the PML estimator proposed by Gouriéroux et al. (2017) is closely related to the maximization of certain moments. To see this, consider the PML estimator with the pseudo distribution of the i-th shock being equal to $g_i \sim t(v)$, where $t(v)$ denotes a t-distribution with v degrees of freedom. The PML estimator is given by

$$\hat{O}^{PML} = \arg\max_{O \in \mathcal{O}} \sum_{t=1}^{T} \sum_{i=1}^{n} log\ g_i(e_{t,i}) \tag{82}$$

$$= \arg\max_{O \in \mathcal{O}} \sum_{t=1}^{T} \sum_{i=1}^{n} -\frac{1-v}{2} log\left(1 + \frac{e_{t,i}^2}{v-2}\right) \tag{83}$$

Ignoring the weighting implied by the degrees of freedom and using $log(1+x) \approx x - \frac{x^2}{2}$ yields

$$\hat{O}^{PML} \approx \arg\max_{O \in \mathcal{O}} \sum_{t=1}^{T} \sum_{i=1}^{n} -e_{t,i}^2 + \frac{e_{t,i}^4}{2} \tag{84}$$

$$= \arg\max_{O \in \mathcal{O}} T \sum_{i=1}^{n} \left(-T^{-1} \sum_{t=1}^{T} e_{t,i}^2\right) + T \sum_{i=1}^{n} \left(\frac{1}{2} T^{-1} \sum_{t=1}^{T} e_{t,i}^4\right) \tag{85}$$

The shocks are normalized and thus $T^{-1} \sum_{t=1}^{T} e_{t,i}^2 = 1$ for $i = 1, ..., n$. It follows that the PML estimator can be written as

$$\hat{O}^{PML} \approx \arg\max_{O \in \mathcal{O}} \sum_{i=1}^{n} \left(T^{-1} \sum_{t=1}^{T} e_{t,i}^4\right) \tag{86}$$

$$= \arg\max_{O \in \mathcal{O}} \sum_{M \in \mathcal{M}(4)} E_T \left[e_M(O)\right] \tag{87}$$

Therefore, maximizing the pseudo log likelihood function is approximately equal to maximizing the kurtosis of the unmixed innovations. Moreover, Equation (87) shows that the PML estimator based on a t-distribution cannot utilize any information contained in third moments.

D Appendix - SVAR: Data and robustness checks

U.S. Industrial Production Index

Source: Board of Governors of the Federal Reserve System (US)

Retrieved from: FRED, Federal Reserve Bank of St. Louis

Link: `https://fred.stlouisfed.org/series/INDPRO`, August 25, 2018

Crude oil composite acquisition cost by refiner

Source: U.S. Energy Information Administration

Link: `https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=R0000____3&f=M`, August 12, 2018

S&P 500

Source: Yahoo! Finance

Link: `https://finance.yahoo.com/quote/%5EGSPC?p=%5EGSPC`, August 12, 2018

U.S. CPI

Source: U.S. Bureau of Labor Statistics

Retrieved from: FRED, Federal Reserve Bank of St. Louis

Link: `https://fred.stlouisfed.org/series/CPIAUCSL`, August 12, 2018

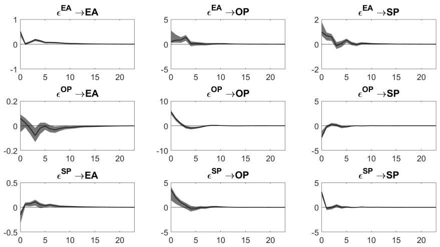
Table 5: Descriptive statistics

	Mean	Median	Mode	Std. deviation	Variance	Skewness	Kurtosis	Range
EA	0.13	0.15	0.41	0.58	0.34	-1.97	15.11	6.15
OP	0.03	0.14	-10.98	1.89	3.56	-1.35	9.42	16.48
SP	0.44	0.82	-17.7	4.12	17	-0.75	4.66	27.99

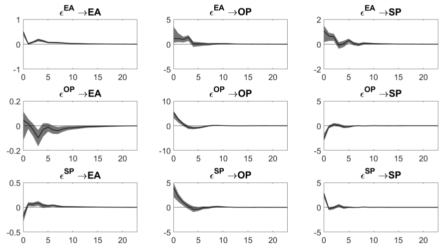
Real economic activity is measured as 100 times the log difference of the monthly US Industrial Production Index. Real S&P 500 returns are calculated as 100 times the log difference of the the S&P 500 closing price deflated by the US CPI. The monthly growth rates of real oil prices are calculated as 100 times the log difference of the crude oil composite acquisition cost by refiners deflated by the US CPI.

Figure 2: Impulse Response Functions - Robustness checks

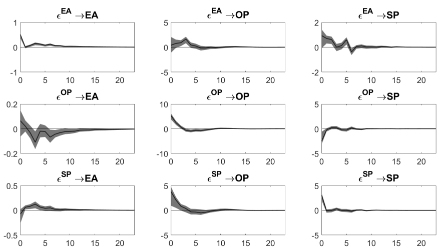
1) Fast SVAR-GMM



2) PML



3) SVAR-GMM (2-step) with six lags



The first and second set of IRFs are based on the reduced form described in Section 8. The first set of IRFs is based on the fast SVAR-GMM estimator and the second set of IRFs is based on the PML estimator proposed by Gouriéroux et al. (2017) and uses a t-distribution with twelve degrees of freedom. The third set of IRFs estimates the reduced form with six lags and the simultaneous relationship is estimated by the two-step SVAR-GMM estimator. Confidence intervals are calculated by bootstrap with 1000 *replications and the interval show the upper* 0.9 *and lower* 0.1 *percentiles.*

References

Apergis, N. and S. M. Miller (2009). Do Structural Oil-Market Shocks Affect Stock Prices? *Energy Economics 31*(4), 569–575.

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

Comon, P. (1994). Independent Component Analysis, a New Concept? *Signal processing 36*(3), 287–314.

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.

Hall, A. R. (2005). *Generalized Method of Moments*. Oxford university press.

Herwartz, H. (2018). Hodges–Lehmann Detection of Structural Shocks – An Analysis of Macroeconomic Dynamics in the Euro Area. *Oxford Bulletin of Economics and Statistics*.

Herwartz, H. and M. Plödt (2016). The Macroeconomic Effects of Oil Price Shocks: Evidence from a Statistical Identification Approach. *Journal of International Money and Finance 61*, 30–44.

Hyvärinen, A., K. Zhang, S. Shimizu, and P. O. Hoyer (2010). Estimation of a Structural Vector Autoregression Model Using Non-Gaussianity. *Journal of Machine Learning Research 11*(May), 1709–1731.

Kilian, L. and H. Lütkepohl (2017). *Structural Vector Autoregressive Analysis*. Cambridge University Press.

Kilian, L. and C. Park (2009). The Impact of Oil Price Shocks on the US Stock Market. *International Economic Review 50*(4), 1267–1287.

Lanne, M. and J. Luoto (2019). GMM Estimation of Non-Gaussian Structural Vector Autoregression. *Journal of Business & Economic Statistics* (just-accepted), 1–35.

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

Lezcano-Casado, M. and D. Martínez-Rubio (2019). Cheap Orthogonal Constraints in Neural Networks: A Simple Parametrization of the Orthogonal and Unitary Group. *arXiv preprint arXiv:1901.08428*.

Moneta, A., D. Entner, P. O. Hoyer, and A. Coad (2013). Causal Inference by Independent Component Analysis: Theory and Applications. *Oxford Bulletin of Economics and Statistics 75*(5), 705–730.

Rubio-Ramírez, J. F., D. F. Waggoner, and T. Zha (2010). Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. *The Review of Economic Studies 77*(2), 665–696.

Sadorsky, P. (1999). Oil Price Shocks and Stock Market Activity. *Energy economics 21*(5), 449–469.