Journal of Business & Economic Statistics

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ubes20

Identifying Structural Vector Autoregression via Leptokurtic Economic Shocks

Markku Lanne, Keyan Liu & Jani Luoto

To cite this article: Markku Lanne, Keyan Liu & Jani Luoto (2023) Identifying Structural Vector Autoregression via Leptokurtic Economic Shocks, Journal of Business & Economic Statistics, 41:4, 1341-1351, DOI: 10.1080/07350015.2022.2134872

To link to this article: https://doi.org/10.1080/07350015.2022.2134872

JOURNAL OF BUSINESS & ECONOMIC STATISTICS
2023, VOL. 41, NO. 4, 1341–1351
https://doi.org/10.1080/07350015.2022.2134872

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS

Check for updates

Identifying Structural Vector Autoregression via Leptokurtic Economic Shocks

Markku Lanne, Keyan Liu, and Jani Luoto

Faculty of Social Sciences, University of Helsinki, Helsinki, Finland

ABSTRACT
We revisit the generalized method of moments (GMM) estimation of the non-Gaussian structural vector autoregressive (SVAR) model. It is shown that in the n-dimensional SVAR model, global and local identification of the contemporaneous impact matrix is achieved with as few as $n^2 + n(n-1)/2$ suitably selected moment conditions, when at least $n-1$ of the structural errors are all leptokurtic (or platykurtic). We also relax the potentially problematic assumption of mutually independent structural errors in part of the previous literature to the requirement that the errors be mutually uncorrelated. Moreover, we assume the error term to be only serially uncorrelated, not independent in time, which allows for univariate conditional heteroscedasticity in its components. A small simulation experiment highlights the good properties of the estimator and the proposed moment selection procedure. The use of the methods is illustrated by means of an empirical application to the effect of a tax increase on U.S. gasoline consumption and carbon dioxide emissions.

ARTICLE HISTORY
Received July 2021
Accepted October 2022

KEYWORDS
Carbon dioxide emissions; Gasoline tax; Generalized method of moments; Non-Gaussian time series; Structural vector autoregression

1. Introduction

Efficient use of the statistical properties of the data for identification has recently become increasingly popular in the literature on structural vector autoregressive (SVAR) models. In particular, heteroscedasticity and non-Gaussianity commonly encountered in economic data are features that facilitate statistical identification (see Kilian and Lütkepohl (2017, chap. 14) for a survey of some of the literature). While economic knowledge is still required to give statistically identified structural shocks an interpretation, the fact that a SVAR model is exactly identified by statistical properties is useful as it presumably enhances estimation accuracy and makes it possible to test any economic identifying restrictions, including those entertained in the previous literature.

In this article, we revisit the generalized method of moments (GMM) approach put forth by Lanne and Luoto (2021) and Keweloh (2021). Lanne and Luoto showed how local identification of the parameters of the SVAR model is achieved by a suitable selection of co-kurtosis conditions when the structural errors are orthogonal. Keweloh, in turn, pointed out that these conditions do not guarantee global identification, and introduced a much larger set of moment conditions that, under the stronger assumption of independent structural shocks, yields a locally and globally identified GMM estimator. However, as we show by a counterexample in Section 2.2, while his moment conditions are sufficient for identification, they are not all necessary, that is, global and local identification can be achieved by a considerably smaller set.

Lanne and Luoto (2021) and Keweloh (2021) assume that at most one of the structural errors has zero excess kurtosis. While we are able to demonstrate by a counterexample that not all of Keweloh's co-kurtosis conditions are necessarily needed for global identification under this assumption, it does not seem possible to show generally that our reduced set suffices to that end. However, we are able to show global identification under the assumption of all (but one) structural errors being either leptokurtic or platykurtic. In other words, in an n-dimensional system, it is required for global identification that at least $n-1$ structural errors have excess kurtosis of the same sign. This assumption is slightly stronger than the ubiquitous assumption of at least $n-1$ non-Gaussian shocks in the related literature. However, it seems innocuous, as platykurtic errors are extremely unlikely to be encountered in economic applications, while economic shocks can, in general, be expected to be leptokurtic (see, e.g., Acemoglu, Ozdaglar, and Tahbaz-Salehi (2017), who showed that the long tails of aggregate variables, such as the GDP, can be explained by sectoral heterogeneity as long as microeconomic shocks, such as total factor productivity shocks, in different sectors have heavy tails, which they found indeed to be the case, at least in U.S. data).

In contrast to some of the previous statistical identification literature, where mutual independence of the structural errors has been assumed, we make the milder assumption that the errors are mutually uncorrelated. Moreover, in line with Guay (2021), we assume that the error term is only serially uncorrelated, not independent in time, which allows for conditional heteroscedasticity in the individual components of the error term. Under these assumptions, we show that when the errors exhibit no excess co-kurtosis, only certain n^2 moment conditions are needed for global identification of the n^2 elements of contemporaneous response matrix as long as at least $n-1$ of the structural

CONTACT Jani Luoto jani.luoto@helsinki.fi Faculty of Social Sciences, University of Helsinki, Helsinki, Finland.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/UBES.

errors are all leptokurtic (or platykurtic). In addition, global and first-order local identification are achieved by augmenting the n^2 moment conditions by $n(n-1)/2$ asymmetric co-kurtosis conditions. This facilitates statistical inference, and while it is more than mere consistency calls for, the number of conditions required is still considerably smaller than required in Keweloh (2021). Our result is likely to be of importance especially in models with a large number of variables compared to the sample size. Keweloh, acknowledging the potential problems of a very large set of moment conditions, also introduced a fast SVAR-GMM estimator, which, however, is not asymptotically efficient. In contrast, the GMM estimator based on our moment conditions is not hampered by the latter property. Finally, we propose a procedure, based on a well-known moment selection criterion, for finding the optimal set of moment conditions among the sets of $n^2+n(n-1)/2$ moment conditions that guarantee global and first-order local identification.

In addition to Lanne and Luoto (2021) and Keweloh (2021), GMM estimation of SVAR models has recently been considered by Guay (2021) and Lewis (2021) (see also Herwartz and Plödt (2016), Braun (2021), Gouriéroux, Monfort, and Renne (2017), and Gouriéroux and Jasiak (2021) for closely related approaches). Guay derives conditions for partial local identification and derives procedures for testing for local identification of part of the structural shocks. His identification results are based on the assumption that the orthogonal structural errors exhibit zero co-skewness and no excess co-kurtosis, which comes very close to assuming independence of the structural errors. Although Guay concentrates on local identification, as a matter of fact, our results in Supplementary Appendix F can be used to show that his GMM estimator also achieves global identification when (all but one of) the structural errors exhibit (positive) excess kurtosis. Nevertheless, in the fully identified model, he suggests using the same large set of moment conditions as Keweloh. Lewis's approach is different in that instead of co-skewness and co-kurtosis conditions, he considers identification based on the autocovariance structure of the second moments of the errors implied by an arbitrary stochastic process for the error variances. However, without parametric assumptions his GMM estimator faces challenges in small samples encountered in typical applications, and he actually ends up recommending a specific parametric model to capture the conditional heteroscedasticity.

We illustrate the use of the methods in an empirical application to the effect of a tax increase on U.S. gasoline consumption and CO_2 emissions. In particular, we consider a bivariate SVAR specification introduced by Davis and Kilian (2011) that contains the percent change in the real gasoline consumption and the percent change in the inflation-adjusted gasoline tax. Unlike Davis and Kilian, in our setup, we are able to test their identification restriction, and it is not rejected at conventional significance levels. Furthermore, our estimates of the effects of a tax increase, obtained without imposing any additional restrictions, turn out to be quite close to those in Davis and Kilian.

The structure of the rest of the article is as follows. In Section 2, we introduce the SVAR model and detail the assumptions under which local and global identification of the GMM estimator can be shown. Specifically, in Section 2.1, we introduce the GMM estimator. For ease of exposition, we first concentrate on the two-stage estimator based on residuals of a vector autoregression estimated by ordinary least squares. In high-dimensional SVAR models it may also be the only feasible alternative although, in general, we recommend estimating all parameters of the SVAR model jointly. In Section 2.2, we first show by a counterexample that not all of Keweloh's (2021) moment conditions are necessary for global identification, and then derive the much smaller set of moment conditions required for global identification under the slightly stronger conditions that all (but one) structural errors be either leptokurtic or platykurtic. For local identification, and hence, asymptotic inference, additional moment conditions are needed, as we show in Section 2.3, where also the asymptotic distribution of the two-stage GMM estimator is derived. The asymptotic distribution of the joint GMM estimator is, in turn, derived in Section 2.4. The optimal selection of moment conditions is discussed in Section 2.5. Finally, in Section 2.6, the computation of impulse responses and their confidence intervals is considered. Section 3 contains the results of small simulation experiment to highlight the properties of the proposed GMM estimator. Section 4 contains the empirical application to the effect of a tax increase on U.S. gasoline consumption. Finally, Section 5 concludes. The proofs of the results in Section 2 are deferred to Supplementary Appendix.

2. Model

We consider the structural VAR model of order p,

$$y_t = \nu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + B\varepsilon_t, \quad t = 1, \ldots, T, \quad (1)$$

where y_t is the n-dimensional time series of interest, ν is an $(n \times 1)$ intercept term, and $A_1, \ldots, A_p$ and B are $(n \times n)$ parameter matrices. The $(n \times n)$ nonsingular matrix B defines the $(n \times 1)$ vector of reduced-form errors u_t as a linear combination of the structural errors ε_t with zero mean and identity covariance matrix, that is, $u_t = B\varepsilon_t$. We further assume y_t to be stable, that is,

$$\det A(z) \overset{def}{=} \det\left(I_n - A_1 z - \cdots - A_p z^p\right) \neq 0, \quad |z| \leq 1, \quad (2)$$

and weakly stationary.

An alternative SVAR formulation is obtained by left-multiplying (1) by the inverse of B:

$$A_0 y_t = \nu^* + A_1^* y_{t-1} + \cdots + A_p^* y_{t-p} + \varepsilon_t, \quad (3)$$

where $A_0 = B^{-1}$, $\nu^* = B^{-1}\nu$, and $A_j^* = B^{-1}A_j$ $(j = 1, \ldots, p)$. Typically, the diagonal elements of A_0 are normalized to unity, and the covariance matrix of ε_t is a diagonal matrix in this specification. In our empirical application in Section 4, we consider this formulation of the SVAR model.

Following Lanne and Luoto (2021) and Keweloh (2021), we assume that at most one of the elements of the structural error vector ε_t is Gaussian. Specifically, we make the following assumption:

Assumption 1.

(i) The error process $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})'$ is a sequence of serially uncorrelated random vectors.
(ii) The components $\varepsilon_{1t}, \ldots, \varepsilon_{nt}$ of ε_t are orthogonal and have no excess co-kurtosis.
(iii) Each component ε_{it}, $i = 1, \ldots, n$, has mean zero, variance unity, and finite third and fourth moments.
(iv) At most one component of ε_t has zero excess kurtosis.

The requirement of no excess co-kurtosis in Assumption 1(ii) means that, given Assumption 1(iii), $E(\varepsilon_{it}\varepsilon_{jt}\varepsilon_{kt}\varepsilon_{lt}) = 1$ when $i = k, j = l \neq k$ or $i = l, j = k \neq l$ or $i = j \neq k = l$ $(i, j, k, l = 1, \ldots, n)$ and zero otherwise (excluding univariate kurtosis $E(\varepsilon_i^4)$ for $i = 1, \ldots, n$, when $i = j = k = l$). These are the values that would prevail if the structural errors were independent. However, Assumption 1(*ii*) is clearly less restrictive than the independence assumption of Lanne, Meitz, and Saikkonen (2017), Keweloh (2021), and Herwartz (2018) since, in addition to no excess co-kurtosis, it only requires that the shocks be orthogonal, and hence does not impose any restrictions, for instance, on co-skewness. As pointed out by Kilian and Lütkepohl (2017, chap. 14), assuming independence is potentially problematic in that independent structural errors need not be obtained as linear transformations of the residuals of a reduced-form VAR model. Hence, the independence assumption may be more restrictive than appears at first sight. While our Assumption 1(ii) is stronger than mere orthogonality, the structural shocks are still linear combinations of the reduced-form residuals. Finally, it is worth pointing out that Assumption 1(i) allows each component ε_{it} to follow a univariate conditionally heteroscedastic process.

2.1. GMM Estimation

Recently, Lanne and Luoto (2021), and Keweloh (2021) have suggested estimation of B in (1) by the GMM, that is, by minimizing with respect to the n^2-dimensional vector of parameters of interest $\vartheta = \text{vec}(B)$ the objective function

$$Q_T(\vartheta) = T^{-1}\sum_{t=1}^{T} f(u_t, \vartheta)' W_T T^{-1} \sum_{t=1}^{T} f(u_t, \vartheta), \qquad (4)$$

where $u_t = y_t - \nu - A_1 y_{t-1} - \cdots - A_p y_{t-p}$. The positive semidefinite $(q \times q)$ matrix W_T (potentially dependent on the data) converging to a positive definite weighting matrix W contains the weights of the sample counterparts of the q-dimensional vector of moment conditions $E\left[f(u_t, \vartheta)\right] = 0$. Notice that (4) depends on only u_t and B. This is facilitated by the fact that the intercept ν and the autoregressive coefficient matrices $A_1, \ldots, A_p$ of the reduced-form VAR model (1) can be estimated consistently by ordinary least squares (OLS). Hence, in practice the estimation of B can be based on the reduced-form residuals, in line with much of the previous related literature. It must be borne in mind, however, that the asymptotic covariance matrix of this two-stage GMM estimator depends on the OLS estimation error (we thank an anonymous referee for pointing this out). When there is a need to make the dependence of the reduced form error vector u_t on $\pi = \text{vec}(\nu, A_1, \ldots, A_p)$ explicit, we denote $u_t(\pi)$.

Lanne and Luoto (2021) based estimation on a subset of the following moment conditions suggested by Assumption 1:

$$E\left(\varepsilon_{it}^2\right) - 1 = 0, \quad i = 1, \ldots, n \qquad (5)$$

$$E\left(\varepsilon_{it}\varepsilon_{jt}\right) = 0, \quad i > j,\ i, j = 1, \ldots, n \qquad (6)$$

$$E\left(\varepsilon_{it}^2\varepsilon_{jt}^2\right) - 1 = 0, \quad i > j,\ i, j = 1, \ldots, n \qquad (7)$$

$$E\left(\varepsilon_{it}^3\varepsilon_{jt}\right) = 0, \quad i \neq j \qquad (8)$$

whereas, in addition to (5) and (6), Keweloh (2021) considered all fourth-order co-moment conditions, including conditions (7) and (8). Moment conditions (5) and (6) naturally follow from the assumptions of zero mean, unit variance, and orthogonality of the components of ε_t. Under Gaussianity, uncorrelatedness implies independence, and hence the symmetric and asymmetric co-kurtosis conditions (7) and (8) are implied by conditions (5) and (6) and the zero-mean assumption, while otherwise they are informative in estimation.

It is, in general, desirable to use the efficient estimator with minimum asymptotic variance obtained by setting $W = H_0^{-1}$ with

$$H_0 = [G_\pi(F^{-1} \otimes I_n) \quad I_q] H [G_\pi(F^{-1} \otimes I_n) \quad I_q]', \qquad (9)$$

where $G_\pi = E\left[\frac{\partial f(u_t(\pi_0), \vartheta_0)}{\partial \pi'}\right]$, ϑ_0 and π_0 are the true values of ϑ and π, respectively, $F = E(Z_{t-1}Z_{t-1}')$ with $Z_t = (1, y_t', \ldots, y_{t-p+1}')'$, and H is the long-run covariance matrix of all moment conditions (see Supplementary Appendix E and Gourieroux et al. (2020, Supplementary Appendix, Section 5.2), for details). The form of (9) reflects the fact that in the case of the two-stage GMM estimator, the optimal weighting matrix $W = H_0^{-1}$ depends on the OLS estimation error.

The matrix H is given by (see, e.g., Hansen 1982)

$$H = \lim_{T\to\infty} \text{var}\left[T^{1/2}\begin{pmatrix} T^{-1}\sum_{t=1}^{T} \text{vec}(u_t(\pi_0)Z_{t-1}') \\ T^{-1}\sum_{t=1}^{T} f(u_t(\pi_0), \vartheta_0) \end{pmatrix}\right],$$

and it is estimated consistently (under regularity conditions, see Newey and West 1994) by the following heteroscedasticity and autocorrelation consistent covariance (HAC) matrix estimator:

$$\hat{H}_{\text{HAC}} = \hat{\Gamma}_0 + \sum_{i=1}^{T-1} \omega_{i,T}\left(\hat{\Gamma}_i + \hat{\Gamma}_i'\right), \qquad (10)$$

where $\hat{\Gamma}_i$ is a consistent estimator of Γ_i, the ith autocovariance matrix of the vector $[\text{vec}(u_t(\pi_0)Z_{t-1}')', f(u_t(\pi_0), \vartheta_0)']'$. A number of different kinds of weighting functions (or kernels) to compute the weights $\omega_{i,T}$ have been put forth in the GMM literature, including the Bartlett, Parzen and Quadratic Spectral kernels, but according to the simulation evidence of Newey and West (1994), the bandwidth (the number of autocovariance matrices included) is far more important for the finite-sample performance of the HAC estimator than the choice of the kernel, and they propose an automatic bandwidth selection procedure, which, coupled with the Bartlett kernel, we also employ in Sections 3 and 4.

In practice, estimation can be carried out in at least three different ways using numerical optimization methods. First, Hansen's (1982) two-step estimator is obtained by first minimizing (4) with W_T suboptimal (such as the identity matrix), and then re-estimating ϑ based on $\hat{H}_{\text{HAC}}$ computed using the first-step estimator of ϑ. Second, this procedure can be continued iteratively until the estimate of ϑ converges to obtain the iterated GMM estimator. Finally, the continuous updating estimator (CUE) of Hansen, Heaton, and Yaron (1996) acknowledges the dependence of the efficient weighting matrix on the parameters. It is based on directly minimizing with respect to ϑ the objective function

$$T^{-1}\sum_{t=1}^{T} f(u_t,\vartheta)' H_{0T}(\vartheta)^{-1} T^{-1}\sum_{t=1}^{T} f(u_t,\vartheta),$$

where $H_{0T}(\vartheta)$ is obtained by replacing G_π, F and H in (9) by their consistent estimators. The small-sample simulation results of Lanne and Luoto (2021) suggest the superiority of the two-step estimator, and we will use it throughout.

2.2. Global Identification

For consistency of the GMM estimator, global identification is required. There are in total $n(n+1)(n+2)(n+3)/24 - n$ symmetric and asymmetric co-kurtosis conditions such as (7) and (8), and Keweloh (2021) indeed claimed that under Assumption 1 (reinforced by the assumption of mutually independent structural errors), all of them (in addition to the unit-variance and zero-correlation conditions) are necessarily needed for global identification in the GMM estimation of the SVAR model. However, as we show by a counterexample below, only a subset of the co-kurtosis conditions suffice to globally identify B up to permutation and multiplication by -1 of its columns. That is, in contrast to Keweloh's claim, all the conditions in the much larger set that he put forth, are not necessary for global identification. Our counterexample involves a bivariate model where the result can be conveniently shown, but the conclusion should not depend on the dimension of the model.

Since the ordering and signs of the elements of ε_t are irrelevant from the viewpoint of estimating the SVAR model, instead of the structural errors, it is convenient to work with the vector of unmixed innovations that we define as

$$e_t(A) = Au_t, \tag{11}$$

where A is called the unmixing matrix. When there is no need to make the dependence of the vector of unmixed innovations on A explicit, we denote $e_t(A)$ by just e_t. Because the vector of the reduced-form errors $u_t = B\varepsilon_t$, $e_t = B^{-1}u_t = \varepsilon_t$ (the unmixed innovations equal the structural shocks) if $A = B^{-1}$. Reordering the columns of B and multiplying them by -1 only changes the order of the elements of e_t and their signs, respectively. Hence, the moment conditions (5)–(8) can be written in terms of the unmixed innovations instead of the structural errors by just replacing ε_{it} and ε_{jt} by e_{it} and e_{jt}, respectively.

Proposition 1 demonstrates that if in the bivariate model ($n = 2$), we include all but one of Keweloh's (2021) variance and co-kurtosis conditions, global identification is still reached. The omitted co-kurtosis condition is $E\left(e_{1t}e_{2t}^3\right) = 0$, but the conclusion holds for the condition $E\left(e_{1t}^3 e_{2t}\right) = 0$ as well. Hence, while his moment conditions are sufficient, they are not necessary for global identification. The proof of the proposition is found in Supplementary Appendix A.

Proposition 1. Let $\varepsilon_t = (\varepsilon_{1t},\varepsilon_{2t})'$, $u_t = (u_{1t},u_{2t})'$, $e_t = (e_{1t},e_{2t})'$, and $e_t = Au_t$. Suppose $\varepsilon_t = B^{-1}u_t$ satisfies Assumption 1, and assume that A in (11) solves the following moment conditions:

$$E\left(e_{it}^2\right) - 1 = 0, \quad i = 1,2 \tag{12}$$

$$E\left(e_{1t}e_{2t}\right) = 0, \tag{13}$$

$$E\left(e_{1t}^2 e_{2t}^2\right) - 1 = 0, \tag{14}$$

$$E\left(e_{1t}^3 e_{2t}\right) = 0. \tag{15}$$

Then, under Assumption 1, for some signed (2×2) permutation matrix P, $A = PB^{-1}$, and hence $e_t = P\varepsilon_t$ (given the signs, the vector of unmixed innovations equals the vector of structural shocks up to permutation of its rows).

While we are unable to prove global identification under Assumption 1 when some of the asymmetric co-kurtosis conditions are omitted in the general case with $n > 2$, this can be done under the following slightly stronger assumption:

Assumption 2.

(i) The error process $\varepsilon_t = (\varepsilon_{1t},\ldots,\varepsilon_{nt})'$ is a sequence of serially uncorrelated random vectors.
(ii) The components $\varepsilon_{1t},\ldots,\varepsilon_{nt}$ of ε_t are orthogonal and have no excess co-kurtosis.
(iii) Each component ε_{it}, $i = 1,\ldots,n$, has mean zero, variance unity, and finite third and fourth moments.
(iv) At most one component of ε_t has zero excess kurtosis, and the excess kurtosis of each of the remaining $n-1$ components has the same sign (i.e., these $n-1$ shocks are all either leptokurtic or platykurtic).

Parts (i)–(iii) of Assumption 2 are identical to those of Assumption 1, whereas part (iv) is replaced by a slightly stronger yet innocuous assumption that at most one component of ε_t has zero excess kurtosis, while the remaining $n-1$ components are all either leptokurtic or platykurtic (i.e., the excess kurtosis of each of them has the same sign). This assumption does not seem restrictive because platykurtic structural shocks are highly unlikely in economic applications, while leptokurtic shocks abound. In particular, macroeconomic and financial shocks closely associated with rare disasters such as the COVID-19 pandemic (see, e.g., Woodford 2020; Guerrieri et al. 2020) are leptokurtic by nature. Macroeconomic shocks are likely to be leptokurtic even more generally. For instance, when studying macro risks, Bekaert, Engstrom, and Ermolov (2021) show that U.S. aggregate demand and supply shocks are leptokurtic. In the same vein, Brunnermeir et al. (2021) found evidence in favor of leptokurtic structural shocks in a ten-variable SVAR model of the U.S. economy. Furthermore, nonlinear features of the data can generate leptokurtic shocks in a linear model. Namely, even if the shocks in a nonlinear dynamic stochastic general equilibrium (DSGE) model generating the data are not leptokurtic, nonlinearity tends to increase the kurtosis of macroeconomic and financial variables far above that of a Gaussian random variable, which shows up as leptokurtic shocks in a linear model, such as a linear DSGE or SVAR model (see, e.g., Andreasen 2012) for results based on a calibrated DSGE model). Moreover, conditionally heteroscedastic errors often encountered in macroeconomic and financial applications exhibit excess kurtosis that has been used in identification of SVAR models also in the previous literature (see, e.g., Kilian and Lütkepohl (2017, chap. 14), and the references therein).

Under Assumption 2, we are able to show that n variance, $n(n-1)/2$ covariance, and $n(n-1)/2$ co-kurtosis conditions of the form $E(\varepsilon_{it}^2\varepsilon_{jt}^2)-1=0$ $(i>j)$ are sufficient for the global identification of B. In other words, only n^2 moment conditions are needed for global identification of the n^2 elements of B (up to signs and permutation of its columns). While asymmetric co-kurtosis conditions of the form $E(\varepsilon_{it}^3\varepsilon_{jt})=0$ are not needed, including additional moment conditions does not destroy global identification. This result is stated as the following proposition (for a proof, see Supplementary Appendix B):

Proposition 2. Suppose $\varepsilon_t = B^{-1}u_t$ satisfies Assumption 2, and assume that A in (11) solves the following n^2 moment conditions:

$$E\left(e_{it}^2\right)-1=0, \quad i=1,\ldots,n \tag{16}$$

$$E\left(e_{it}e_{jt}\right)=0, \quad i>j,\ i,j=1,\ldots,n \tag{17}$$

$$E\left(e_{it}^2e_{jt}^2\right)-1=0, \quad i>j,\ i,j=1,\ldots,n. \tag{18}$$

Then, for some signed $(n\times n)$ permutation matrix P, $A=PB^{-1}$, and hence $e_t=P\varepsilon_t$ (given the signs, the vector of unmixed innovations equals the vector of structural shocks up to permutation of its rows).

Proposition 2 provides one solution to the identification problem. It states that if the $n(n-1)/2$ symmetric co-kurtosis conditions (18) (in addition to conditions (16)–(17)) are satisfied, the unmixed innovations and structural shocks are equal apart from signs and permutations of the components of ε_t. In terms of the computational burden, this result provides a considerable improvement on Keweloh's (2021, Proposition 3), according to which $n(n+1)(n+2)(n+3)/24-n$ co-kurtosis conditions (in addition to conditions (16)–(17)) are required for global identification under nonzero excess kurtosis. For instance, a five-dimensional SVAR model involves in total as many as 65 co-kurtosis conditions, while according to Proposition 2, only 10 co-kurtosis conditions are needed to guarantee global identification. Of course, our result is based on the slightly stronger assumption that structural shocks must all be either leptokurtic or platykurtic, but as discussed above, at least in macroeconomic and financial applications it is hard to imagine a platykurtic structural shock, so that this assumption can quite safely be assumed to hold.

2.3. Statistical Inference

In order to conduct statistical inference on the two-stage estimator ϑ, we need to show its consistency and find its asymptotic distribution. When the global identification result in Proposition 2 holds, under standard regularity conditions, including strict stationarity and ergodicity, and mild technical conditions concerning the GMM objective function $Q_T(\vartheta)$, the parameter space $\Theta \subset \mathbb{R}^k$ $(k=n^2)$ and the weighting matrix W_T, the GMM estimator $\hat{\vartheta}$ is consistent for any fixed signed permutation matrix P (see, e.g., Dovonon and Hall 2018, Proposition 1).

Unfortunately, moment conditions (16)–(18) are not, in general, sufficient to establish the asymptotic normality of the GMM estimator because they do not guarantee the expectation of the Jacobian matrix $J_T(\vartheta_0)=\partial f(u_t,\vartheta)/\partial\vartheta'|_{\vartheta=\vartheta_0}$, evaluated at ϑ_0, the true value of ϑ, to be of full column rank n^2 (see, e.g., Hall (2005, chap. 3.4.2) for a discussion on this requirement). However, although this standard condition of first-order local identification fails, Dovonon and Hall's (2018) condition for second order local-identification holds (for a proof, see Supplementary Appendix C), and, thus, their Theorem 1 could be used to establish the asymptotic distribution of the GMM estimator. That result is, alas, in practice rarely useful in our setup because they concentrate on the only analytically tractable case where the rank of the expectation of $J_T(\vartheta_0)$ is reduced by one, and with moment conditions (16)–(18), this additional condition is satisfied only in a bivariate SVAR model.

In view of the latter conclusion, we turn to finding sets of moment conditions that guarantee both global and first-order local identification, and, thus, asymptotic normality. According to Proposition 2, conditions (16)–(18), potentially augmented with a number of asymmetric co-kurtosis conditions, suffice for global identification. On the other hand, according to Proposition 1 of Lanne and Luoto (2021), conditions (16)–(17) coupled with $n(n-1)/2$ asymmetric co-kurtosis conditions are, in general, sufficient for local identification. Based on their idea, it can be shown that first-order local identification is achieved by augmenting conditions (16)–(18) with $n(n-1)/2$ asymmetric co-kurtosis conditions. Because including additional asymmetric co-kurtosis conditions does not destroy global identification, global and first-order local identification are achieved by augmenting conditions (16)–(18) with $n(n-1)/2$ asymmetric co-kurtosis conditions, as stated in the following Proposition 3 (for a proof, see Supplementary Appendix D):

Proposition 3 (Global and first-order local identification). Suppose all n components of ε_t have positive (or negative) excess kurtosis. Then moment conditions (16)–(18), and $n(n-1)/2$ asymmetric co-kurtosis conditions of the form $E(e_{it}^3e_{jt})=0$ $(i\neq j)$ suffice to globally and locally identify B characterized by a given permutation and signs of its columns. If one of the components of ε_t has zero excess kurtosis, the asymmetric co-kurtosis conditions must not involve its third power.

The augmented set comprises $n^2+n(n-1)/2$ moment conditions, which is more than consistency calls for, but still considerably fewer than required by Keweloh's (2021, Proposition 3). For instance, in a five-dimensional SVAR model, the total number of moment conditions required is 35, while according to Keweloh, 80 conditions are needed.

When the assumptions of Proposition 3 hold, the first-order local identification condition that $\text{rank}\{E[J_T(\vartheta_0)]\}=n^2$ is satisfied. However, if one of the components of ε_t has zero excess kurtosis, the locally identifying asymmetric co-kurtosis conditions must be such that they do not involve its third power. Nevertheless, it is straightforward to show that first-order local identification can be achieved by a suitable selection of asymmetric co-kurtosis conditions even if any one of the shocks is Gaussian. In practice, it is virtually never known whether one of the structural shocks is Gaussian, and therefore, in Section 2.5, we propose a procedure, based on a well-known moment selection criterion, to find the optimal set of moment conditions among the sets of $n^2+n(n-1)/2$ moment conditions that guarantee global and first-order local identification.

It is important to realize that Proposition 3 only applies to a given SVAR model characterized by a given ordering and signs of the columns of B. In particular, if one permutation of the columns of B asymptotically satisfies the moment conditions, so do all permutations. Thus, in order to facilitate standard asymptotic inference, additional restrictions are needed to pinpoint a particular permutation (and signs) of the columns. These restrictions are not really restrictive, however, because any permutation of the columns of B produces the same shocks (reordered) and impulse responses. To this end, there are many alternative restriction schemes entertained in the previous literature on statistical identification that could be employed (see Lanne, Meitz, and Saikkonen (2017), and the references therein). In this article, we use the permutation convention of Pham and Garat (1987) which entails picking the permutation that maximizes the absolute value of the product of the diagonal elements of B, and restricts the diagonal elements of B positive.

As already discussed, under standard regularity conditions, including strict stationarity and ergodicity (see, e.g., Hall 2005, chap. 3 and 5.3), in addition to global identification, the GMM estimator $\hat{\vartheta}_T$ is a consistent estimator of ϑ_0. This holds for all two-step, iterated and continuous updating GMM estimators that are asymptotically equivalent although they may behave differently in finite samples. Moreover, when also the condition of local identification is satisfied, the efficient GMM estimator is asymptotically normally distributed with covariance matrix $[G_\vartheta' H_0^{-1} G_\vartheta]^{-1}$, where $G_\vartheta = E\left[\frac{\partial f(u_t(\pi_0),\vartheta_0)}{\partial \vartheta'}\right]$, and H_0 is given in (9) (see Supplementary Appendix E for details). Because the SVAR model is statistically identified, additional restrictions on the elements of B can be tested by standard Wald and likelihood ratio (LR) type tests, once G_ϑ and H_0 are replaced by their consistent estimators. However, it must be borne in mind that any test on the parameters of the impact matrix only pertains to the particular ordering of its columns. Therefore, any hypothesis on the parameters is, in general, meaningful only once the shocks pertaining to those parameters have an economic interpretation (see the empirical example in Section 4).

2.4. Joint GMM Estimator

So far, we have considered the two-stage GMM estimator, where the impact matrix B is estimated based on the residuals of the reduced-form VAR(p) model. An alternative to the two-stage estimator, feasible in models involving a small number of variables in relation to the number of observations, is to estimate ν, $A_1,\ldots,A_p$, and B jointly by augmenting the set of moment conditions discussed above by $E[\varepsilon_t \otimes (1, y_{t-1}',\ldots,y_{t-p+1}')'] = 0$. Also in this case, any one of the two-step, iterated and continuous-updating estimators can be employed. The theoretical results derived for the two-stage estimator generalize in a straightforward manner to cover this case. In particular, the conditions for local and global identification of B in Sections 2.2 and 2.3 do not depend on whether all parameters are estimated jointly or in two stages. In our empirical application involving only bivariate models in Section 4, we estimate all the parameters of the SVAR model jointly.

The joint GMM estimator of $\phi = \text{vec}(\pi', \vartheta')'$ where $\pi = \text{vec}(\nu, A_1,\ldots,A_p)$ and $\vartheta = \text{vec}(B)$, is consistent and asymptotically normally distributed:

$$T^{1/2}(\phi - \hat{\phi}_T) \xrightarrow{d} N(0,\Omega), \tag{19}$$

where Ω and its consistent estimator $\hat{\Omega}_T$ are obtained as above by replacing ϑ_0 by ϕ_0 and $\hat{\vartheta}_T$ by $\hat{\phi}_T$ in the expressions above. This follows from the fact that $\hat{\vartheta}_T$ is asymptotically normal when $f(u_t(\pi),\vartheta)$ contains the augmented set of the $n(n-1)/2+n^2$ moment conditions. Because the global and local identification results in Propositions 2 and 3 hold for any $u_t = y_t - \nu - A_1 y_{t-1} - \cdots - A_p y_{t-p}$, also the GMM estimator of ϕ based on the condition $E[\varepsilon_t \otimes (1, y_{t-1}',\ldots,y_{t-p+1}')'] = 0$ in addition to those discussed in Section 2.3 is asymptotically normally distributed.

Finally, the efficient estimator with minimum asymptotic variance is obtained by setting the weighting matrix W in GMM estimation at S^{-1}, the inverse of the long-run covariance matrix of the moment conditions, $S = \lim_{T\to\infty} \text{var}\left[T^{1/2}\left(T^{-1}\sum_{t=1}^T f(u_t(\pi_0),\vartheta_0)\right)\right]$. The latter is estimated consistently by (10), where $\hat{\Gamma}_i$ is a consistent estimator of Γ_i, the ith autocovariance matrix of $f(u_t(\pi_0),\vartheta_0)$.

2.5. Moment Selection

While Propositions 2 and 3 imply global and first-order local identification when the set of moment conditions contains $n(n-1)/2$ asymmetric co-kurtosis conditions of the form $E\left(e_{it}^3 e_{jt}\right) = 0$ $(i \neq j)$ in addition to conditions (16)–(18), they provide no guidance on selecting the asymmetric co-kurtosis conditions to include. However, the number of different combinations of $n(n-1)/2$ asymmetric co-kurtosis conditions is large and increasing rapidly in n. As a matter of fact, there are $n(n-1)!/\{[n(n-1)/2]!\}^2$ such combinations, and they are not all likely to be equally informative for the parameters of interest. Therefore, it is desirable to be able to select the combination that is the most informative and agrees with the data. To this end, we suggest using the relevant moment selection criterion (RMSC) of Hall et al. (2007), and using Hansen's (1982) well-known J test of over-identifying restrictions as a diagnostic check to confirm the accordance of the selected moment conditions with the data.

The RMSC for the joint GMM estimator of $\phi = \text{vec}(\pi', \vartheta')'$ is given by

$$\text{RMSC}(c) = \ln(\det[V_{\hat{\phi},T}(c)]) + (q-k)\ln[(T/b_T)^{1/2}](T/b_T)^{-1/2} \tag{20}$$

where the sets of asymmetric co-kurtosis conditions are indexed by c, $V_{\hat{\phi},T}(c)$ is a consistent estimator of the covariance matrix $\left(G_0' S^{-1} G_0\right)^{-1}$ of the GMM estimator $\hat{\phi}$, $G_0 = E[J_T(\phi_0)]$, and $S = \lim_{T\to\infty} \text{var}\left[T^{1/2}\left(T^{-1}\sum_{t=1}^T f(u_t(\pi_0),\vartheta_0)\right)\right]$. The

RMSCs can be computed for the two-stage estimators analogously (with obvious modifications). The bandwidth parameter b_T of the $\hat{S}_{\text{HAC}}$ estimator accounts for its rate of convergence. The value of the RMSC based on each set of moment conditions is computed, and the set minimizing it (or equivalently maximizing estimation accuracy) is selected. It is important to use Pham and Garat (1987) permutation convention or an analogous procedure to fix the order and signs of the elements of the structural error vector ε_t, so the same SVAR model is estimated in each case. In small samples, the GMM estimate depends on the particular moment conditions selected although, in our experience, the estimates are quite robust with respect to the particular moment conditions. However, as pointed out by Hall (2005, chap. 7), under general conditions, including global identification, using the RMSC to select the most informative moment conditions does not affect the asymptotic properties of the GMM estimator.

It should be borne in mind that if the ith shock has zero excess kurtosis, and only $n(n-1)/2$ asymmetric co-kurtosis conditions are entertained, then for the sets of moment conditions containing the condition $E\left(e_{it}^3 e_{jt}\right) = 0$, $\text{rank}(E[J_T(\phi_0)]) < n(np+1) + n^2$, implying that $G_0' S^{-1} G_0$ is noninvertible (see, e.g., result 3.19 of Seber 2008). We therefore recommend calculating $\det(V_{\hat{\phi},T}) = (\det[G_T(\hat{\phi})'\hat{S}_T^{-1} G_T(\hat{\phi})])^{-1}$ instead of $\det([G_T(\hat{\phi})'\hat{S}_T^{-1} G_T(\hat{\phi})]^{-1})$, where $G_T(\hat{\phi})$ and $\hat{S}_T$ are the consistent estimators of G_0 and S, respectively. If $G_T(\hat{\phi})$ is near rank-deficient for some set of moment conditions, then the corresponding value of $(\det[G_T(\hat{\phi})'\hat{S}_T^{-1} G_T(\hat{\phi})])^{-1}$ is very large, and, as a result, the set should not be selected (at least asymptotically). That is, we expect to find the minimum of the RMSC among the first-order locally identifying sets. A similar remark applies to the two-stage GMM estimators.

Estimating the model based on all $n(n-1)!/\{[n(n-1)/2]!\}^2$ combinations of $n(n-1)/2$ asymmetric co-kurtosis conditions can be computationally burdensome if the dimension of the model n is large. For instance, with $n = 4$, there are already 924 such combinations. Hence, in the case of a high-dimensional model, a viable alternative is to use all $n(n-1)$ asymmetric co-kurtosis conditions in estimation. Notice that even if all $n(n-1)$ asymmetric co-kurtosis conditions of the form $E\left(e_{it}^3 e_{jt}\right) = 0$ $(i \neq j)$ are entertained, the augmented set comprises only $n^2 + n(n-1)$ moment conditions, which is considerably fewer than required by Keweloh's (2021, Proposition 3). In a five-dimensional SVAR model, for instance, the total number of moment conditions required is 45, while according to Keweloh, 80 conditions are needed (in a 10-dimensional model, the corresponding figures are 190 and 760, respectively).

As already discussed, we recommend checking the moment conditions by Hansen's J test of over-identifying restrictions that follows the χ^2 distribution with $n(n-1)/2$ degrees of freedom under the null hypothesis. If it rejects, the set of moment conditions producing the second smallest value of the RMSC is considered next, and this is repeated until an acceptable set of moment conditions is found. Checking for misspecification by the J test prior to conducting tests on the elements of B is important because standard Wald and LR type tests have also power against misspecification, as shown by Hall and Inoue (2003). Therefore, they may reject because the moment conditions are violated even if the restrictions actually being tested hold.

2.6. Impulse Response Analysis

Once the model has been estimated, the effects of the structural shocks can be studied by means of impulse response analysis in the usual way. Impulse responses are obtained based on the moving average (MA) representation of y_t,

$$y_t = \mu + \sum_{k=0}^{\infty} C_k B \varepsilon_{t-k}, \tag{21}$$

where μ is the unconditional expectation of y_t, C_0 is the identity matrix, and C_k, $k = 1, 2, \ldots$, are obtained recursively as $C_k = \sum_{l=1}^{k} C_{k-l} A_l$ by setting $A_k = 0$ for $k > p$. The jth column of $C_k B$, $k = 0, 1, \ldots$, contains the impulse responses of the jth structural shock ε_{jt}, $j = 1, \ldots, n$, and its (i,j) element is the response of $y_{i,t+k}$ to a one-unit change in ε_{jt}. That is,

$$\iota_i' C_k B \iota_j = \frac{\partial y_{i,t+k}}{\partial \varepsilon_{jt}}, \tag{22}$$

with ι_i the ith unit vector. We denote this structural impulse response coefficient by $\lambda_{k,i,j}(\pi, \vartheta)$, where $\pi = \text{vec}(\nu, A_1, \ldots, A_p)$ and $\vartheta = \text{vec}(B)$. Thus, in the notation $\lambda_{k,i,j}(\pi, \vartheta)$, we ignore the fact that (22) does not depend on the parameter vector ν. A consistent estimator of $\lambda_{k,i,j}(\pi, \vartheta)$, denoted by $\hat{\lambda}_{k,i,j}(\hat{\pi}, \hat{\vartheta})$, is obtained by replacing π and ϑ by their consistent estimators, $\hat{\pi}$ and $\hat{\vartheta}$, respectively.

In order to compute the confidence intervals of the impulse response functions $\lambda_{k,i,j}(\pi, \vartheta)$, we derive their asymptotic distribution, and to that end, following Montiel Olea, Stock, and Watson (2021), we use the delta-method. Let us first consider the case where all the parameters of the SVAR model are estimated jointly. It was shown in Section 2.4 that $T^{1/2}(\phi - \hat{\phi}) \xrightarrow{d} N(0, \Omega)$. Hence, a delta-method calculation yields

$$T^{1/2}[\hat{\lambda}_{k,i,j}(\hat{\pi}, \hat{\vartheta}) - \lambda_{k,i,j}(\pi, \vartheta)] \xrightarrow{d} N(0, \sigma_{k,i,j}^2), \tag{23}$$

where

$$\sigma_{k,i,j}^2 = \frac{\partial \lambda_{k,i,j}(\pi, \vartheta)}{\partial(\pi', \vartheta')} \Omega \frac{\partial \lambda_{k,i,j}(\pi, \vartheta)}{\partial(\pi', \vartheta')'}.$$

The asymptotic confidence interval of $\lambda_{k,i,j}(\pi, \vartheta)$ is obtained by replacing ϑ, π and Ω by their consistent estimators. Alternatively, a suitable bootstrap procedure can be employed.

Result (23) applies also when the two-stage estimation procedure is used, where π is first estimated by OLS and ϑ then by the GMM based on the OLS residuals. However, in that case, Ω is given by

$$\Omega = \begin{bmatrix} F^{-1} \otimes I_n & 0 \\ -I_{\vartheta\vartheta}^{-1} I_{\vartheta\pi}(F^{-1} \otimes I_n) & -I_{\vartheta\vartheta}^{-1} G_\vartheta' W \end{bmatrix}$$
$$H \begin{bmatrix} F^{-1'} \otimes I_n & -(F^{-1'} \otimes I_n) I_{\vartheta\pi}' I_{\vartheta\vartheta}^{-1} \\ 0 & -W' G_\vartheta I_{\vartheta\vartheta}^{-1} \end{bmatrix}, \tag{24}$$

where $I_{\vartheta\vartheta} = G_\vartheta' W G_\vartheta$, $I_{\vartheta\pi} = G_\vartheta' W G_\pi$, $G_\vartheta = E\left[\frac{\partial f(u_t(\pi_0), \vartheta_0)}{\partial \vartheta'}\right]$, $G_\pi = E\left[\frac{\partial f(u_t(\pi_0), \vartheta_0)}{\partial \pi'}\right]$, $F = E(Z_{t-1} Z_{t-1}')$ with $Z_t = (1, y_t', \ldots, y_{t-p+1}')'$, and H is the long-run covariance matrix of all moment conditions (see Supplementary Appendix E for details). A consistent estimator of Ω is obtained by replacing G_ϑ, G_π, F, and H by their consistent estimators.

3. Simulation Results

In order to gauge the properties of the GMM estimator in finite samples, we conduct a small Monte Carlo simulation experiment. In particular, we are interested in finding out about the potential effects of relying on fewer asymmetric co-kurtosis conditions (either given or selected by the RMSC) than recommended by Keweloh (2021). To facilitate comparison the results in the previous literature, we generate data from the bivariate SVAR(0) model also considered by Gouriéroux, Monfort, and Renne (2017) and Keweloh (2021):

$$y_t = B\varepsilon_t,$$

where B is an orthogonal matrix dependent on a single parameter, that is,

$$B = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

with $\theta = -\pi/5$.

Because all elements of B depend only on θ, it suffices to concentrate on the estimates of just one element, say $B_{11} = \cos(-\pi/5) \approx 0.809$. We consider two cases: one where both independent elements of ε_t follow Student's t distribution with 12 degrees of freedom, and another where they are $t(48)$-distributed. The error terms are standardized to have variance unity. In both cases, B is identified, but with the greater degree-of-freedom parameter, the error distribution is closer to normality, which is expected to show up as a deteriorating performance of the GMM estimator.

We report results based on three different kinds of sets of moment conditions, all of which contain the following conditions: $E(\varepsilon_{1t}^2) = E(\varepsilon_{2t}^2) = 1$, $E(\varepsilon_{1t}\varepsilon_{2t}) = 0$, and $E(\varepsilon_{1t}^2\varepsilon_{2t}^2) = 1$. In addition, the first set contains both asymmetric co-kurtosis conditions, $E(\varepsilon_{1t}^3\varepsilon_{2t}) = 0$ and $E(\varepsilon_{1t}\varepsilon_{2t}^3) = 0$, while the second set contains only the latter. Finally, in the third set, either one of the asymmetric co-kurtosis conditions is selected by the RMSC. In view of the simulation results of Lanne and Luoto (2021), only results based on the two-step GMM estimator are reported.

The averages of bias and standard deviation of the GMM estimate of B_{11} and the rejection rate of the nominal 5%-level J-test are reported in Table 1 for samples of 250, 500, and 1000 observations. Recall that B is identified only up to permutation and multiplication by -1 of its columns. Therefore, the estimator of B_{11} may estimate either B_{11}, $-B_{11}$, B_{12} or $-B_{12}$, and the measures of bias and standard deviation are based on a transformation of the estimate of B_{11} that is closest to the true value of B_{11} (minimizing the squared deviation; see, Gouriéroux, Monfort, and Renne 2017, sec. 2.7).

In estimation, the true values of the parameters were used as initial estimates. However, to find out whether the performance of the estimator depends on the initial estimate, we also estimated B_{11} using a large grid of all permissible values of $\cos(\theta)$, and found no indication of the deterioration of the finite-sample performance of the estimator based on a subset of Keweloh's (2021) co-kurtosis conditions; the behavior of estimators compared was actually very similar.

As expected, bias and standard deviation systematically decrease with increasing sample size irrespective of the asymmetric co-kurtosis conditions included when the errors follow the $t(12)$ distribution. Moreover, the differences with respect to the set of moment conditions are minor. In all cases, the GMM estimator is less accurate when the errors follow the less leptokurtic $t(48)$ distribution. Indeed, in this case, the simulated bias does not decrease, and compared to the case of $t(12)$-distributed errors with strong identification, the simulated standard deviation hardly decreases with the sample size. These findings are a manifestation of lack of point identification when the error terms only slightly deviate from normality.

Interestingly, the rejections rates of the J-test vary considerably across the sets of moment conditions, with either the set containing only the asymmetric co-kurtosis condition $E(\varepsilon_{1t}\varepsilon_{2t}^3) = 0$ or co-kurtosis conditions selected by the RMSC in all cases being the winner when the errors follow the $t(12)$-distribution, while in the case of both co-kurtosis conditions included, the test strongly over-rejects. When identification is weaker with the errors following the $t(48)$-distribution, the J-test is clearly undersized in the first two cases, which is expected in correctly specified unidentified models.

All in all, the simulation results indicate that in terms of estimation accuracy, it is not necessary to include all asymmetric co-kurtosis conditions in the GMM estimation, and the RMSC works quite well. Moreover, when identification is strong, the J-test clearly over-rejects if all asymmetric co-kurtosis conditions are included, while its empirical size quite closely corresponds to its nominal size.

Table 1. Simulation results of the two-step GMM estimator of the SVAR(0) model.

T	Asymmetric Co-Kurtosis conditions	$DF = 12$ Bias	Std.	J test	$DF = 48$ Bias	Std.	J test
	$E(\varepsilon_{1t}^3\varepsilon_{2t}) = E(\varepsilon_{1t}\varepsilon_{2t}^3) = 0$	0.039	0.100	0.098	0.067	0.101	0.052
250	$E(\varepsilon_{1t}\varepsilon_{2t}^3) = 0$	0.041	0.095	0.051	0.059	0.098	0.038
	Selected by RMSC	0.039	0.095	0.058	0.060	0.098	0.048
	$E(\varepsilon_{1t}^3\varepsilon_{2t}) = E(\varepsilon_{1t}\varepsilon_{2t}^3) = 0$	0.023	0.083	0.098	0.065	0.094	0.044
500	$E(\varepsilon_{1t}\varepsilon_{2t}^3) = 0$	0.029	0.083	0.052	0.060	0.092	0.034
	Selected by RMSC	0.027	0.084	0.054	0.064	0.093	0.035
	$E(\varepsilon_{1t}^3\varepsilon_{2t}) = E(\varepsilon_{1t}\varepsilon_{2t}^3) = 0$	0.007	0.069	0.106	0.062	0.092	0.041
1000	$E(\varepsilon_{1t}\varepsilon_{2t}^3) = 0$	0.013	0.074	0.054	0.058	0.089	0.029
	Selected by RMSC	0.013	0.074	0.053	0.061	0.090	0.031

NOTE: The results for the two-step GMM estimator is based on 5000 simulated samples of T = 250, 500, and 1000 observations. The components of the error term $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$, are first generated from independent t distributions with 12 and 48 degrees of freedom (DF). The errors are centered and standardized to have variance unity. Then the data y_t are computed from $y_t = B\varepsilon_t$, where the entries of B are $B_{11} = \cos(\theta)$, $B_{12} = \sin(\theta)$, $B_{21} = -\sin(\theta)$, and $B_{22} = \cos(\theta)$ with $\theta = -\pi/5$. In addition to the asymmetric co-kurtosis conditions listed in the second column, in each case, the set of moment conditions contains the following conditions: $E(\varepsilon_{1t}^2) = E(\varepsilon_{2t}^2) = 1$, $E(\varepsilon_{1t}\varepsilon_{2t}) = 0$, and $E(\varepsilon_{1t}^2\varepsilon_{2t}^2) = 1$. In both panels, columns labeled "Bias" and "Std." contain the average bias and standard deviation of the GMM estimate of B_{11}, respectively. The column labeled "J test" contains the rejection rate of the J-test of over-identifying restrictions at the 5% nominal significance level.

4. Empirical Application

We illustrate the methods in an empirical application to the effect of a tax increase on the U.S. monthly gasoline consumption, with the ultimate goal of estimating the effect of a gasoline tax increase on carbon dioxide (CO_2) emissions. In particular, we consider Davis and Kilian's (2011) bivariate SVAR specification containing the percent change in the real gasoline consumption (Δx_t) and the percent change in the inflation-adjusted gasoline tax (Δtax_t). Davis and Kilian (2011) considered also a model for after-tax price of gasoline and the percent change in the real gasoline consumption that provides a crude estimate of the response of gasoline consumption to a tax increase. However, because the response of consumption to a price change due to a tax increase is probably different from that due to other reasons, the specification for $(\Delta tax_t, \Delta x_t)'$ considered here is likely to yield a more direct estimate of the tax elasticity. Nevertheless, it should be pointed out that by relying on this simple bivariate model, we abstract from changes in gasoline storage (see, Coglianese et al. 2017). The series cover the period from January 1989 to March 2008. (For a detailed discussion of the variables, see Davis and Kilian (2011). The data were downloaded from *http://qed.econ.queensu.ca/jae/2011-v26.7/davis-kilian/*.)

The starting point of the analysis is the following stylized structural model

$$\Delta tax_t = \alpha \Delta x_t + \varepsilon_{1t}$$
$$\Delta x_t = \beta \Delta tax_t + \varepsilon_{2t}$$

with β the tax elasticity of interest. Here ε_{1t} and ε_{2t} are mutually orthogonal structural errors that have no excess co-kurtosis. This structural model is a special case of the more general SVAR(p) model in (3) with $y_t = (\Delta tax_t, \Delta x_t)'$ and

$$A_0 = \begin{bmatrix} 1 & -\alpha \\ -\beta & 1 \end{bmatrix}.$$

In order to estimate the effect of a tax increase on gasoline consumption, we need to identify a shock that has a positive effect on tax and a negative effect on consumption on impact, assuming gasoline is a normal good. To that end, Davis and Kilian (2011) (DK henceforth) ruled out the contemporaneous feedback from Δx_t to Δtax_t by setting $\alpha = 0$, such that only ε_{1t} can be the desired shock. DK were unable to test this identification restriction, but in our setup, it is an over-identifying restriction, and hence, testable conditional on the maintained assumptions.

Following DK, we start by estimating a reduced-form VAR(12) model with an intercept and seasonal dummies for $(\Delta \text{tax}_t, \Delta x_t)'$. For identification, non-Gaussianity of at least one of the structural errors is crucial, and because the structural errors are linear combinations of the reduced-form residuals, this condition is satisfied if at least one of them is non-Gaussian in the bivariate model. To check for their normality, we employ the bootstrap version of the Jarque-Bera test for VAR models proposed by Kilian and Demiroglu (2000) with 10,000 bootstrap replications to construct the bootstrap critical values. The null hypothesis of joint normality of the residuals is rejected at the 1% significance level based on the empirical distribution of the Jarque-Bera test statistic. This suggests non-Gaussianity of the reduced-form residual series, and, hence, lends strong support to the necessary condition for identification.

In order to find the optimal set of moment conditions, we estimate the parameters of the SVAR(12) model (3) using the different combinations of the $n(n-1)/2 + n^2$ conditions discussed in Section 2.5. All parameters are estimated jointly, so we also include the orthogonality conditions $E[\varepsilon_t \otimes Z_{t-1}] = 0$, where Z_{t-1} contains the lagged variables as well as an intercept

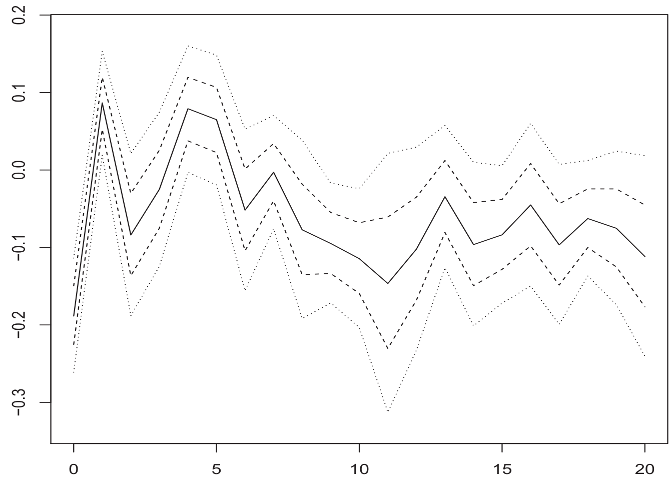

Figure 1. Impulse responses of U.S. aggregate gasoline consumption to a positive gasoline tax shock. The dotted and dashed lines are, respectively, the 95% and 68% confidence bands obtained by the delta-method.

term and seasonal dummies. In all cases, conditions (16)–(18) are included, and the most informative combination of the $n(n-1)/2$ asymmetric co-kurtosis conditions is selected by the RMSC.

The asymmetric moment condition minimizing the RMSC is $E(\varepsilon_{1t}^3 \varepsilon_{2t}) = 0$, and the p-value of the related J-test is 0.43. We also estimated the parameters of the SVAR(12) model using both asymmetric co-kurtosis conditions, with only a marginal change in the parameter estimates. For the latter set of moment conditions, the value of the RMSC is slightly higher than that based on $E(\varepsilon_{1t}\varepsilon_{2t}^3) = 0$ only. The GMM estimates of α and β are, respectively, 0.07 (0.037) and –0.19 (0.037), where the figures in parentheses are asymptotic standard errors. DK identified the model by setting $\alpha = 0$, and according to the asymptotic Wald test, this restriction cannot be rejected at 5% significance level (the p-value is 0.065). The estimate –0.19 of β is close to DK's estimate of –0.14.

In line with these estimates, ε_{1t} turns out to be the only shock having impact effects on consumption and taxes of opposite signs (the first column of the inverse of estimated A_0 matrix is $(0.99, -0.19)'$), and following DK, we label it the gasoline tax shock. The impulse responses of a positive gasoline tax shock causing a 1% tax increase on impact along with the 95% and 68% confidence bands based on the delta-method are depicted in Figure 1. The asymptotic covariance matrix Ω of the joint GMM estimator of $\phi = \text{vec}(\pi', \vartheta')'$ is estimated using the Newey-West HAC estimator with the automatic bandwidth selection procedure proposed by Newey and West (1994). Our impulse response function resembles that of DK's, which is not surprising given that the data lend support to their identification restriction. In particular, we also find a significant impact effect (albeit somewhat greater than DK's estimate), and impulse responses that are insignificant up to the seven-month horizon. Following Montiel Olea, Stock, and Watson (2021), we also considered a bootstrap-like method, where the empirical distribution of $\hat{\lambda}_{k,i}(\hat{\pi}, \hat{\vartheta})$ is obtained using an estimated normal distribution of $\hat{\phi}$ (recall that $\hat{\phi} = (\hat{\pi}', \hat{\vartheta}')'$ is asymptotically normally distributed). The results remain intact irrespective of the method used.

Table 2. The predicted effect of a 10-cent gasoline tax increase on gasoline consumption and CO_2 emissions based on the 12-month tax elasticity (%).

Gasoline consumption	−2.63(1.71)
CO_2 emissions in the United States	−0.89(0.58)
CO_2 emissions in the OECD	−0.40(0.26)
CO_2 emissions worldwide	−0.19(0.12)

NOTE: The figures in parentheses are standard errors.

4.1. Policy Implications

In this section, we report estimates of a gasoline tax increase on gasoline consumption and carbon dioxide emissions. To facilitate comparison, we follow DK in computing the predicted percent change in consumption by multiplying the estimated 12-month tax elasticity −0.10 by a given relative tax increase (a change of tax divided by the mean tax level in cents)

$$-0.10\left(\frac{\tau}{\text{tax}}\right)100.$$

The tax elasticity −0.10 is the estimated 12-month impulse response depicted in Figure 1. We then evaluate the effects for a 10-cent tax increase ($\tau = 10$) at the mean volume-weighted tax level of 38.4 cents ($\text{tax} = 38.4$) in March 2008. A tax increase of this size lies well within the normal historical range. The estimated effect on gasoline consumption, reported on the first row of Table 2 is of similar magnitude as that of DK.

Following DK, the estimated changes in the CO_2 emissions for the U.S. are obtained by multiplying the estimated change in gasoline consumption by 0.338, the assumed share of carbon dioxide emissions derived from the transportation sector. The estimates for the OECD and the world are obtained by assuming that the U.S. represents 44.7% of total OECD CO_2 emissions and 21.0% of world emissions. DK report only estimates based on instrumental variables regression on state-level panel data, but they are not that different from ours.

5. Conclusion

In this article, we have revisited GMM estimation of the SVAR model by moment conditions implied by the assumption commonly made in the statistical identification literature that at most one of the structural errors is Gaussian. Our new results complement those of Lanne and Luoto (2021) and Keweloh (2021). In particular, we have shown that not all of Keweloh's moment conditions are necessarily needed for global and local identification, while the moment conditions of Lanne and Luoto only suffice for local identification. By slightly modifying the common assumption of at most one Gaussian structural error, we have derived the necessary and sufficient condition for global and local identification. Under the modified assumption, all but one of the structural errors are either leptokurtic or platykurtic, which is innocuous from the practical point of view as platykurtic shocks are extremely unlikely to occur in economic applications. Our set of moment conditions is smaller than that required by Keweloh's result, which is likely to be important in models where the number of variables relative to the sample size is large.

We have also relaxed the assumptions of mutually independent structural errors made in much of the previous literature, including Keweloh (2021). The independence assumption is problematic in that it may not be possible to obtain such errors as linear transformations of the reduced-form residuals, as pointed out by Kilian and Lütkepohl (2017, chap. 14). Instead, we show that identification is achieved under the milder assumption that the structural errors exhibit no excess co-kurtosis. In addition, we assume the error term to be only serially uncorrelated, not independent in time, which allows its components to follow univariate conditionally heteroscedastic processes.

If the SVAR model is large, GMM estimation of all the parameters simultaneously may not be feasible. To that end, we have proposed a two-stage estimator, where estimation of the impact matrix is based on reduced-form residuals obtained by OLS estimation, and derive its asymptotic distribution. We have also derived the asymptotic distribution of the impulse response function based on both the simultaneous and two-stage estimators.

According to a small simulation experiment, the accuracy of the GMM estimator based on the reduced set of moment conditions is comparable to that based on the full set of moment conditions suggested by Keweloh (2021). Moreover, with the reduced set of moment conditions, the rejection rate of the J-test of over-identifying restrictions is found to closely correspond to its nominal size, while with the full set of moment conditions, the test tends to severely over-reject. This finding is important since the J-test plays a central role in our procedure for selecting the optimal set of moment conditions.

We have illustrated the use of the methods in an empirical application to the effect of a tax increase on U.S. gasoline consumption and CO_2 emissions. We were unable to reject the identification restriction imposed by Davis and Kilian (2011) at the 5% significance level. Consequently, our estimates of the effects of the tax increase turned out to be quite close to those in Davis and Kilian.

Here we concentrate on the case where the SVAR model is completely statistically identified when at most one of the structural errors is Gaussian, while the rest are leptokurtic (or platykurtic). However, in economic applications, more than one of the shocks may be Gaussian, and it seems that in such a case, some of the shocks may still be statistically identified. In the related literature, the latter case of partial identification has been entertained by Guay (2021), and Bertsche and Braun (2022), and our framework should also lend itself to a similar extension. We leave this issue for future research.

Supplementary Materials

The online supplement comprises six appendices. Appendices A, B and D contain the proofs of Propositions 1, 2 and 3, respectively. In Appendix C, second-order local identification of the contemporaneous impact matrix of the SVAR model under the conditions of Proposition 2 is shown. In Appendix E, the asymptotic distribution of the two-stage estimator of the SVAR model is derived. Finally, Appendices F and G contain auxiliary results needed in the proofs of Propositions 1, 2 and 3. The data and code for all computations are available via the online supplement.

Acknowledgments

The authors would like to thank Jianqing Fan (the editor), Associate Editor, and three anonymous referees for useful comments and suggestions.

Disclosure Statement

The authors report there are no competing interests to declare.

Funding

Financial support from the Academy of Finland (grant 347986) is gratefully acknowledged. Open access funded by Helsinki University Library.

References

Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2017), "Microeconomic Origins of Macroeconomic Tail Risks," *American Economic Review*, 107, 54–108. [1341]

Andreasen, M. M. (2012), "On the Effects of Rare Disasters and Uncertainty Shocks for Risk Premia in Non-linear DSGE Models," *Review of Economic Dynamics*, 15, 295–315. [1345]

Bekaert, G., Engstrom, E., and Ermolov, A. (2021), "Macro Risks and the Term Structure of Interest Rates," *Journal of Financial Economics*, 141, 479–504. [1344]

Bertsche, D., and Braun, R. (2022), "Identification of Structural Vector Autoregressions by Stochastic Volatility," *Journal of Business and Economic Statistics*, 40, 328–341. [1350]

Braun, R. (2021), "The Importance of Supply and Demand for Oil Prices: Evidence from non-Gaussianity," Bank of England, Staff working paper no. 957. [1342]

Brunnermeir, M., Palia, D., Sastry, K. A., and Sims, C. A. (2021), "Feedbacks: Financial Markets and Economic Activity," *American Economic Review*, 111, 1845–1879. [1344]

Coglianese, J., Davis, L. W., Kilian, L., and Stock, J. H. (2017), "Anticipation, Tax Avoidance, and the Price Elasticity of Gasoline Demand," *Journal of Applied Econometrics*, 32, 1–15. [1349]

Davis, L. W., and Kilian, L. (2011), "Estimating the Effect of a Gasoline Tax on Carbon Emissions," *Journal of Applied Econometrics*, 26, 1187–1214. [1342,1348,1349,1350]

Dovonon, P., and Hall, A. R. (2018), "The Asymptotic Properties of GMM and Indirect Inference under Second-order Identification," *Journal of Econometrics*, 205, 76–111. [1345]

Guay, A. (2021), "Identification of Structural Vector Autoregressions through Higher Unconditional Moments," *Journal of Econometrics*, 225, 27–46. [1341,1342,1350]

Guerrieri, V., Lorenzoni, G., Straub, L., and Werning (2020), "Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages?" NBER working paper no. 26918, April. [1344]

Gouriéroux, C., Monfort, A., and Renne, J.-P. (2017), "Statistical Inference for Independent Component Analysis: Application to Structural VAR Models," *Journal of Econometrics*, 196, 111–126. [1342,1348]

——— (2020), "Identification and Estimation in Non-Fundamental Structural VARMA Models," *The Review of Economic Studies*, 87, 1915–1953. [1343]

Gouriéroux, C., and Jasiak, J. (2021), "Generalized Covariance Estimator," Papers 2107.06979, arXiv.org. [1342]

Hall, A. R. (2005), *Generalized Method of Moments*, New York: Oxford University Press. [1345,1346,1347]

Hall, A. R., and Inoue, A. (2003), "The Large Sample Behavior of the Generalized Method of Moments Estimator in Misspecified Models," *Journal of Econometrics*, 114, 361–394. [1347]

Hall, A. R., Inoue, A., Jana, K., and Shin, C. (2007), "Information in Generalized Method of Moments Estimation and Entropy-Based Moment Selection," *Journal of Econometrics*, 138, 488–512. [1346]

Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054. [1343,1344,1346]

Hansen, L. P., Heaton, J., and Yaron, A. (1996), "Finite Sample Properties of Some Alternative GMM Estimators Obtained from Financial Market Data," *Journal of Business and Economic Statistics*, 14, 262–280. [1344]

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, 736–754. [1343]

Herwartz, H., and Plödt, M. (2016), "The Macroeconomic Effects of Oil Price Shocks: Evidence from a Statistical Identification Approach, ," *Journal of International Money and Finance*, 61, 30–44. [1342]

Keweloh, S. A. (2021), "A Generalized Method of Moments Estimator for Structural Vector Autoregressions Based on Higher Moments," *Journal of Business and Economic Statistics*, 39, 772–782. [1341,1342,1343,1344,1345,1347,1348,1350]

Kilian, L., and Demiroglu, U. (2000), "Residual-Based Tests for Normality in Autoregressions: Asymptotic Theory and Simulation Evidence," *Journal of Business and Economic Statistics*, 18, 40–50. [1349]

Kilian, L., and Lütkepohl, H. (2017), *Structural Vector Autoregressive Analysis*, Cambridge: Cambridge University Press. [1341,1343,1345,1350]

Lanne, M., and Luoto, J. (2021), "GMM Estimation of Non-Gaussian Structural Vector Autoregression," *Journal of Business and Economic Statistics*, 39, 69–81. [1341,1342,1343,1344,1345,1348,1350]

Lanne, M., Meitz, M., and Saikkonen, P. (2017), "Identification and Estimation of Non-Gaussian Structural Vector Autoregressions," *Journal of Econometrics*, 196, 288–304. [1343,1346]

Lewis, D. J. (2021), "Identifying Shocks via Time-Varying Volatility," *Review of Economic Studies*, 88, 3086–3124. [1342]

Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Berlin: Springer.

Montiel Olea, J. L., Stock, J. H., and Watson, M. W. (2021), "Inference in Structural Vector Autoregressions Identified with an External Instrument," *Journal of Econometrics*, 225, 74–87. [1347,1349]

Newey, W. K., and West, K. D. (1994), "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631–653. [1343,1349]

Pham, D. T., and Garat, P. (1987), "Blind Separation of Mixture of Independent Sources through a Quasi-maximum Likelihood Approach," *IEEE Transactions on Signal Processing*, 45, 1712–1725. [1346,1347]

Seber, G. A. F. (2008), "A Matrix Handbook for Statisticians," Hoboken, New Jersey: Wiley. [1347]

Woodford, M. (2020), "Effective Demand Failures and the Limits of Monetary Stabilization Policy," NBER working paper no. 27768, September. [1344]