Supplementary Appendix: Identifying Structural Vector Autoregression via Leptokurtic Economic Shocks

Markku Lanne
Faculty of Social Sciences, University of Helsinki

Keyan Liu
Faculty of Social Sciences, University of Helsinki

Jani Luoto*
Faculty of Social Sciences, University of Helsinki

*Corresponding author. E-mail address: jani.luoto@helsinki.fi

Appendix A Proof of Proposition 1

Suppose $\varepsilon_t = B^{-1}u_t$ satisfies Assumption 1, and assume that A in (11) solves the moment conditions (12)–(15). Letting $Q = AB$, and using $\varepsilon_t = B^{-1}u_t$, the unmixed innovations in (11) can be expressed as

$$e_t = Au_t = Q\varepsilon_t. \tag{A.1}$$

Now we are ready to show that if the conditions (12)–(15) (together with Assumption 1) are satisfied, then there exists a signed permutation matrix P such that $Q = P$. This together with $Q = AB$ implies that $e_t = P\varepsilon_t$ and $A = PB^{-1}$, and hence, B is globally identified up to signs and permutation of its columns (i.e., the structural shocks can be recovered from the estimated reduced form errors, but their signs and order remain unknown).

First, notice that conditions (12) and (13) together with (A.1) and Assumption 1 imply that Q is orthogonal: $I = E(e_t e_t') = QE(\varepsilon_t\varepsilon_t')Q' = QQ'$. The rest of the proof is straightforward and relies on the following results derived in Supplementary Appendix F:

$$E(e_{it}^2 e_{jt}^2) - 1 = \sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2 \Gamma_k = 0, \quad i > j \tag{A.2}$$

$$E(e_{it}^3 e_{jt}) = \sum_{k=1}^{n} Q_{ik}^3 Q_{jk} \Gamma_k = 0, \quad i \neq j. \tag{A.3}$$

where we denote by $\Gamma_i = E(\varepsilon_{it}^4) - 3$ the excess kurtosis of the structural shocks ε_{it}, $i = 1, \ldots, n$, and Q_{ij} are the (i,j)-elements, $i, j = 1, \ldots, n$, of Q in (A.1).

Now, based on (A.2) and (A.3), conditions (14) and (15) yield

$$Q_{11}^2 Q_{21}^2 \Gamma_1 + Q_{12}^2 Q_{22}^2 \Gamma_2 = 0, \tag{A.4}$$

$$Q_{11}^3 Q_{21} \Gamma_1 + Q_{12}^3 Q_{22} \Gamma_2 = 0. \tag{A.5}$$

According to Assumption 1(iv) at most one component of ε_t has zero excess kurtosis. Suppose first that both Γ_1 and Γ_2 are different from zero (i.e., both shocks have nonzero excess kurtosis). Then, multiplying (A.4) by Q_{11} and (A.5) by Q_{21} and subtracting the first resulting equation from the second yields

$$Q_{22} Q_{12}^2 (Q_{21} Q_{12} - Q_{11} Q_{22}) \Gamma_2 = 0. \tag{A.6}$$

Therefore, if both Q_{22} and Q_{12} are different from zero, we have

$$Q_{21}Q_{12} = Q_{11}Q_{22}. \tag{A.7}$$

This implies that $\det(Q) = 0$ but Q was shown to be orthogonal, a contradiction. Thus, either Q_{12} or Q_{22} is zero (they cannot both be zero because Q is orthogonal).

Suppose $Q_{12} = 0$, then the orthogonality of Q implies

$$QQ' = \begin{pmatrix} Q_{11}^2 & Q_{11}Q_{21} \\ Q_{11}Q_{21} & Q_{21}^2 + Q_{22}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{A.8}$$

which yields $Q_{11}^2 = 1$, $Q_{11}Q_{21} = 0$, and $Q_{21}^2 + Q_{22}^2 = 1$. Therefore,

$$Q = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}, \tag{A.9}$$

a signed permutation matrix. Similarly, if $Q_{22} = 0$, we have

$$Q = \begin{pmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{pmatrix}, \tag{A.10}$$

also a signed permutation matrix. As already discussed, this together with $Q = AB$ implies that $e_t = P\varepsilon_t$ and $A = PB^{-1}$, and hence, B is globally identified up to sign and permutation of its columns.

Suppose now that $\Gamma_1 = 0$ (i.e., ε_{1t} has zero excess kurtosis) and $\Gamma_2 \neq 0$. Then for (A.4) (and (A.5)) to hold, either Q_{12} or Q_{22} must be zero (they cannot both be zero because Q is orthogonal). If $Q_{12} = 0$, by the orthogonality of Q, Q is given by (A.9), and if $Q_{22} = 0$, for the same reason, Q is given by (A.10). Hence $Q = P$ implying that $e_t = P\varepsilon_t$ and $A = PB^{-1}$. Similarly, if $\Gamma_2 = 0$ and $\Gamma_1 \neq 0$, then (A.4) and (A.5) imply that either Q_{21} or Q_{11} must be zero. Using the same arguments as above, we have that in the former case Q is given by (A.9) and in the latter by (A.10), and therefore $Q = P$.

Finally, if Assumption 1(*iv*) does not hold, so that both ε_{1t} and ε_{2t} have zero excess kurtosis ($\Gamma_1 = \Gamma_2 = 0$), equations (A.4) and (A.5) are satisfied for any choice of Q. Thus, the structural errors ε_t cannot be expressed as a unique linear combination of the reduced-form errors u_t.

Appendix B Proof of Proposition 2

Suppose that $\varepsilon_t = B^{-1}u_t$ satisfies Assumption 2, and assume that A in (11) solves the moment conditions (16)–(18). Then, by conditions (16) and (17), and using $e_t = Q\varepsilon_t$ (see (A.1) in Appendix A) and Assumptions 2(ii)–(iii), we obtain $I = E(e_t e_t') = QE(\varepsilon_t \varepsilon_t')Q' = QQ'$. That is, Q is orthogonal.

From (A.2) in Appendix A, condition (18) yields

$$E(e_{it}^2 e_{jt}^2) - 1 = \sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2 \Gamma_k = 0, \quad i \neq j \tag{B.1}$$

where we use the fact that $E(e_{it}^2 e_{jt}^2) = E(e_{jt}^2 e_{it}^2)$. By Assumption 2($iv$) at most one component of ε_t has zero excess kurtosis, while 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).

Suppose first that all the shocks have positive (negative) excess kurtosis: $\Gamma_k > 0$ ($\Gamma_k < 0$) for all $k = 1, \ldots, n$. Then, by (B.1) it must be that

$$Q_{ik}^2 Q_{jk}^2 = 0. \quad i, j, k = 1, \ldots, n \tag{B.2}$$

This means that each column of Q has at most one nonzero element. Thus, by the orthogonality of Q, each column of Q has exactly one nonzero element, and for the same reason each row of Q has exactly one nonzero element equal to ± 1. Thus, $Q = P$, a signed permutation matrix.

Suppose next that only one component of ε_t, say, ε_{lt} has zero excess kurtosis (i.e., $\Gamma_l = 0$, $1 \leq l \leq n$). Then, by (B.1), $Q_{ik}^2 Q_{jk}^2 = 0$ for all $i, j, k = 1, \ldots, n$, $k \neq l$, and therefore we know that each column of Q except the lth has at most one nonzero element. By the orthogonality of Q, it thus follows that Q_k, the kth column of Q, ($k = 1, \ldots, n$, $k \neq l$) has exactly one nonzero element equal to ± 1. Similarly, because of the orthogonality of Q, $Q_i' Q_j = 0$ ($i \neq j$), and hence the $n \times (n-1)$ matrix Q_{-l}, obtained by dropping Q_l from Q, has exactly one zero row, and each of its remaining rows has exactly one nonzero element equal to ± 1. Therefore, from $Q_k' Q_l = 0$ for $k = 1, \ldots, n$, $k \neq l$, it follows that Q_l has at most one nonzero element (corresponding to the zero row of Q_{-l}), and as $Q_l' Q_l = 1$, this element must equal ± 1. Thus, $Q = P$, a signed permutation matrix.

Finally, it is important to realize that additional asymmetric co-kurtosis conditions of the form $E(e_{it}^3 e_{jt}) = 0$ $(i \neq j)$ do not destroy the global identification result. Since Q is a signed permutation matrix under Assumption 2, either Q_{ik} or Q_{jk} must equal zero for all $i, j, k = 1, \ldots, n$, and this is exactly what is required for

$$E(e_{it}^3 e_{jt}) = \sum_{k=1}^{n} Q_{ik}^3 Q_{jk} \Gamma_k = 0, \quad i \neq j,$$

to hold as well.

Appendix C Second-order local identification of B

We consider $\theta = \text{vec}(A)$ that is more convenient to work with than $\vartheta = \text{vec}(B)$. Recall that $A = PB^{-1}$, where P is a signed $(n \times n)$ permutation matrix. We begin by showing that first-order local identification fails. The necessary condition for first-order local identification of B is that the expectation of the Jacobian matrix $E[J_T(\theta_0)]$, evaluated at θ_0, the true value of θ, has full column rank k. Thus, we need to show that, for θ, the row vector of $k = n^2$ parameters to be estimated, and $f(u_t, \theta_0)$ the vector of population moment conditions (16)–(18), $\text{rank}(E[\partial f(u_t, \theta_0)/\partial\theta']) < k$. Because the row rank equals the column rank of a square matrix, it suffices to show that some rows of $M(\theta_0) \equiv E[\partial f(u_t, \theta_0)/\partial\theta']$ are linearly dependent. The $n^2 \times n^2$ matrix $M(\theta_0)$ is obtained by stacking the n components of the form (C.1), the $n(n-1)/2$ components of the form (C.2), and $n(n-1)/2$ components of the form (C.3):

$$E\left[\left.\frac{\partial e_{it}^2}{\partial\theta'}\right|_{\theta=\theta_0}\right] = 2(b'_{0i} \otimes \iota'_i), \quad i = 1, \ldots, n, \tag{C.1}$$

$$E\left[\left.\frac{\partial e_{it}e_{jt}}{\partial\theta'}\right|_{\theta=\theta_0}\right] = (b'_{0i} \otimes \iota'_j) + (b'_{0j} \otimes \iota'_i), \quad i > j, \tag{C.2}$$

$$E\left[\left.\frac{\partial e_{it}^2 e_{jt}^2}{\partial\theta'}\right|_{\theta=\theta_0}\right] = 2(b'_{0i} \otimes \iota'_i) + 2(b'_{0j} \otimes \iota'_j), \quad i > j, \tag{C.3}$$

where $i, j \in \{1, ..., n\}$, ι_i is the ith column of the $n \times n$ identity matrix, and b_{0i} is the ith column of B_0, the true value of B, (expressions (C.1)–(C.3) are derived in Supplementary Appendix G). It can clearly be seen that any of the rows given by (C.3) is obtained as a sum of two rows given by (C.1), and hence the Jacobian matrix is of reduced rank.

We now show that although local identification fails at the first order, it holds at the second order. To this end, we let $m(\theta) = E[f(u_t, \theta)]$ and

$$M_k^2(\theta_0) = E\left[\left.\frac{\partial^2 f_k(u_t, \theta)}{\partial\theta\partial\theta'}\right|_{\theta=\theta_0}\right], \quad k = 1, 2, \ldots, n^2, \tag{C.4}$$

where $f_k(u_t, \theta)$ is the kth element of $f(u_t, \theta)$. Second-order local identification is defined as follows (adapted from Dovonon and Hall (2018)):

Definition 1. *The moment condition $m(\theta) = 0$ locally identifies θ_0 up to the second order if :*

(a) $m(\theta_0) = 0$

(b) for all u in the range of $M(\theta_0)'$ and all v in the null space of $M(\theta_0)$, we have:

$$\Big(M(\theta_0)u + \big(v' M_k^2(\theta_0) v\big)_{1 \le k \le n^2} = 0\Big) \Rightarrow (u = v = 0). \tag{C.5}$$

Notice first that if $v = 0$, then by the sufficient condition in (C.5), u must be in the null space of $M(\theta_0)$: $M(\theta_0)u = 0$. On the other hand, the null space of $M(\theta_0)$ is the orthogonal complement of the column space (range) of $M(\theta_0)'$ (see, for example, the result (2.37) in Seber (2007)). Thus, u is in both the range and the null space of $M(\theta_0)'$, which means that it must be 0. Therefore, it suffices to show that $v = 0$ for all u in the range of $M(\theta_0)'$.

To show this, we first introduce the second-order derivatives $M_k^2(\theta_0)$ of the moment conditions (16)–(18) whose derivation is deferred to Supplementary Appendix G. The $n^2 \times n^2$ matrices $M_k^2(\theta_0)$ are given by

$$E\left[\left.\frac{\partial^2 e_{it}^2}{\partial\theta\partial\theta'}\right|_{\theta=\theta_0}\right] = 2(B_0 B_0') \otimes (\iota_i \iota_i'), \quad i = 1, \ldots, n, \tag{C.6}$$

$$E\left[\left.\frac{\partial^2 e_{it} e_{jt}}{\partial\theta\partial\theta'}\right|_{\theta=\theta_0}\right] = (B_0 B_0') \otimes (\iota_i \iota_j' + \iota_j \iota_i') \quad i > j,\ i, j = 1, \ldots, n, \tag{C.7}$$

$$\begin{aligned} E\left[\left.\frac{\partial^2 e_{it}^2 e_{jt}^2}{\partial\theta\partial\theta'}\right|_{\theta=\theta_0}\right] &= 2(B_0 \Xi_j B_0') \otimes (\iota_i \iota_i') + 2(B_0 \Xi_i B_0') \otimes (\iota_j \iota_j') \\ &\quad + 4(B_0 \Upsilon_{i,j} B_0') \otimes (\iota_i \iota_j' + \iota_j \iota_i'), \quad i > j,\ i, j = 1, \ldots, n, \end{aligned} \tag{C.8}$$

with $\Xi_i = E[\varepsilon_{it}^2(\varepsilon_t \varepsilon_t')]$ and $\Upsilon_{i,j} = E[\varepsilon_{it}\varepsilon_{jt}(\varepsilon_t \varepsilon_t')]$.

Because all n^2 equations in (C.5) must hold for all v in the null space of $M(\theta_0)$, we use only some of them to show that if they hold (and v is in the null space of $M(\theta_0)$), then v must be a zero vector. In particular, the equations in (C.5) we have in mind, are the ones associated with (C.6) and (C.8) (and (C.1) and (C.3)), and we use (C.6) (and (C.1)) multiple times.

We begin by substituting (C.1) for i and j and (C.3) for i, j $(i, j = 1, \ldots, n,\ j < i)$ into (C.5), and subtracting the first two rows of the resulting expression from its third row:

$$\begin{aligned}
&2(b_i' \otimes \iota_i' + b_j' \otimes \iota_j')u - 2(b_i' \otimes \iota_i')u - 2(b_j' \otimes \iota_j')u \\
&+v' \left[E\left[\frac{\partial^2 e_{it}^2 e_{jt}^2}{\partial\theta\partial\theta'}\bigg|_{\theta=\theta_0} \right] - E\left[\frac{\partial^2 e_{it}^2}{\partial\theta\partial\theta'}\bigg|_{\theta=\theta_0} \right] - E\left[\frac{\partial^2 e_{jt}^2}{\partial\theta\partial\theta'}\bigg|_{\theta=\theta_0} \right] \right] v \\
&= v' \left[E\left[\frac{\partial^2 e_{it}^2 e_{jt}^2}{\partial\theta\partial\theta'}\bigg|_{\theta=\theta_0} \right] - E\left[\frac{\partial^2 e_{it}^2}{\partial\theta\partial\theta'}\bigg|_{\theta=\theta_0} \right] - E\left[\frac{\partial^2 e_{jt}^2}{\partial\theta\partial\theta'}\bigg|_{\theta=\theta_0} \right] \right] v = 0. \qquad \text{(C.9)}
\end{aligned}$$

Substituting (C.6) and (C.8) into (C.9) yields

$$\begin{aligned}
&2v' \left[(B_0(\Xi_j - I_n)B_0') \otimes (\iota_i \iota_i') \right] v \\
&+ 2v' \left[(B_0(\Xi_i - I_n)B_0') \otimes (\iota_j \iota_j') \right] v \\
&+ 4v' \left[(B_0 \Upsilon_{i,j} B_0') \otimes (\iota_i \iota_j' + \iota_j \iota_i') \right] v = 0. \qquad \text{(C.10)}
\end{aligned}$$

By Assumption 2(ii) that the components of the error term ε_t have no excess co-kurtosis, Ξ_i is a $(n \times n)$ diagonal matrix with $E(\varepsilon_{it}^4)$ in the ith diagonal position and ones elsewhere in the diagonal. Therefore, we have $\Xi_i - I_n = (E(\varepsilon_{it}^4) - 1)(\iota_i \iota_i')$. Using this result with $B_0 \iota_i = b_{0i}$, (C.10) can be expressed as

$$\begin{aligned}
&2(E(\varepsilon_{jt}^4) - 1)v' \left[(b_{0j} b_{0j}') \otimes (\iota_i \iota_i') \right] v \\
&+ 2(E(\varepsilon_{it}^4) - 1)v' \left[(b_{0i} b_{0i}') \otimes (\iota_j \iota_j') \right] v \\
&+ 4v' \left[(B_0 \Upsilon_{i,j} B_0') \otimes (\iota_i \iota_j' + \iota_j \iota_i') \right] v = 0. \qquad \text{(C.11)}
\end{aligned}$$

Again, by Assumption 2(ii) that the components of the error term ε_t have no excess co-kurtosis, a typical element $E(\varepsilon_{it}\varepsilon_{jt}\varepsilon_{kt}\varepsilon_{lt})$ $(i > j,\ i, j, k, l = 1, \ldots, n)$ of the $(n \times n)$ matrix $\Upsilon_{i,j}$ is 1 when $i = k,\ j = l \neq k$ or $i = l,\ j = k \neq l$ and zero otherwise. As a result, $\Upsilon_{i,j} = \iota_i \iota_j' + \iota_j \iota_i'$, and the last term on the left hand side of (C.11) becomes

$$\begin{aligned}
4v' \left[(B_0 \Upsilon_{i,j} B_0') \otimes (\iota_i \iota_j' + \iota_j \iota_i') \right] v &= 4v' \left[(b_{0i} b_{0j}' + b_{0j} b_{0i}') \otimes (\iota_i \iota_j' + \iota_j \iota_i') \right] v \\
&= 4v'[(b_{0i} \otimes \iota_i)(b_{0j}' \otimes \iota_j') + (b_{0i} \otimes \iota_j)(b_{0j}' \otimes \iota_i') \\
&\quad + (b_{0j} \otimes \iota_i)(b_{0i}' \otimes \iota_j') + (b_{0j} \otimes \iota_j)(b_{0i}' \otimes \iota_i')]v \\
&= 4v' \left[(b_{0i} \otimes \iota_j)(b_{0j}' \otimes \iota_i') + (b_{0j} \otimes \iota_i)(b_{0i}' \otimes \iota_j') \right] v, \quad \text{(C.12)}
\end{aligned}$$

where in the last equality we use $v'(b_{0i}\otimes\iota_i)=0$ (recall that v is in the null space of $M(\theta_0)$, i.e., $M(\theta_0)v=0$, and cf. (C.1)). By adding and subtracting the term $4[(b_{0j}\otimes\iota_i)(b_{0j}'\otimes\iota_i')+(b_{0i}\otimes\iota_j)(b_{0i}'\otimes\iota_j')]$, we have

$$4\left[(b_{0i}\otimes\iota_j)(b_{0j}'\otimes\iota_i')+(b_{0j}\otimes\iota_i)(b_{0i}'\otimes\iota_j')\right]=4[(b_{0i}\otimes\iota_j+b_{0j}\otimes\iota_i)(b_{0i}\otimes\iota_j+b_{0j}\otimes\iota_i)' \\ -(b_{0j}\otimes\iota_i)(b_{0j}'\otimes\iota_i')-(b_{0i}\otimes\iota_j)(b_{0i}'\otimes\iota_j')].$$

Substituting the above into (C.12), and using $v'(b_{0i}\otimes\iota_j+b_{0j}\otimes\iota_i)=0$ (recall again that v is in the null space of $M(\theta_0)$), and cf. (C.2)) yields

$$4v'\left[(B_0\Upsilon_{i,j}B_0')\otimes(\iota_i\iota_j'+\iota_j\iota_i')\right]v=-4v'\left[(b_{0j}\otimes\iota_i)(b_{0j}'\otimes\iota_i')+(b_{0i}\otimes\iota_j)(b_{0i}'\otimes\iota_j')\right]v \\ =-4v'\left[(b_{0j}b_{0j}')\otimes(\iota_i\iota_i')+(b_{0i}b_{0i}')\otimes(\iota_j\iota_j')\right]v. \tag{C.13}$$

Substituting (C.13) into (C.11), we obtain

$$2(E(\varepsilon_{jt}^4)-3)v'\left[(b_{0j}b_{0j}')\otimes(\iota_i\iota_i')\right]v+2(E(\varepsilon_{it}^4)-3)v'\left[(b_{0i}b_{0i}')\otimes(\iota_j\iota_j')\right]v=0, \tag{C.14}$$

where the $n^2\times n^2$ matrices $(b_{0i}b_{0i}')\otimes(\iota_j\iota_j')$ and $(b_{0j}b_{0j}')\otimes(\iota_i\iota_i')$ are positive semi-definite (their only nonzero eigenvalues are given by $b_{0i}'b_{0i}$ and $b_{0j}'b_{0j}$, respectively), implying $v'[(b_{0i}b_{0i}')\otimes(\iota_j\iota_j')]v\geq 0$ and $v'[(b_{0j}b_{0j}')\otimes(\iota_i\iota_i')]v\geq 0$ ($i>j$, $i,j=1,\ldots,n$).

By Assumption 2(iv) at most one component of ε_t has zero excess kurtosis, while 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). Suppose first $E(\varepsilon_{it}^4)-3$ is positive (negative) for all $i=1,\ldots,n$. Then, by the results $v'[(b_{0i}b_{0i}')\otimes(\iota_j\iota_j')]v\geq 0$ and $v'[(b_{0j}b_{0j}')\otimes(\iota_i\iota_i')]v\geq 0$ obtained above together with $(b_{0i}b_{0i}')\otimes(\iota_j\iota_j')=(b_{0i}\otimes\iota_j)(b_{0i}'\otimes\iota_j')$, it follows from (C.14) that $v'(b_{0i}\otimes\iota_j)=0$ and $v'(b_{0j}\otimes\iota_i)=0$ ($i>j$, $i,j=1,\ldots,n$). Using $v'(b_{0i}\otimes\iota_i)=0$, we hence obtain

$$v'(b_{0k}\otimes\iota_l)=0,\ \text{ for all }\ k,l=1,\ldots,n, \tag{C.15}$$

a system of n^2 linear equations of the form $Cv=0$, where the $n^2\times n^2$ matrix C is obtained by stacking $b_{0k}'\otimes\iota_l'$ for $k,l=1,\ldots,n$. By the linear independence of b_{0i}s, C is of full rank, and hence, by the invertible matrix theorem, $Cv=0$ has only the trivial solution $v=0$.

Suppose next that $E(\varepsilon_{it}^4)-3=0$ for some i ($0\leq i\leq n$). Then, by the preceding discussion, (C.14) yields

$$v'(b_{0k}\otimes\iota_l)=0.\ \text{ for all }\ k,l=1,\ldots,n,\ \ k\neq i. \tag{C.16}$$

Because permuting the rows of a matrix does not change its rank and because (C.16) must hold for any permutation of the rows of ε_t, we may set $i = n$ without loss of generality. By the assumption in Definition 1(b) that v is in the null space of $M(\theta_0)$, it follows from (C.1) and (C.2) that

$$v'(b_{0n} \otimes \iota_n) = 0, \tag{C.17}$$

and

$$v'[(b_{0n} \otimes \iota_j) + (b_{0j} \otimes \iota_n)] = 0, \;\; j = 1, \ldots, n-1 \tag{C.18}$$

respectively. Combining (C.16)–(C.18), we have a system of n^2 linear equations of the form $Dv = 0$, where the $n^2 \times n^2$ matrix D is obtained by stacking $(b'_{0k} \otimes \iota'_l)$ for $k, l = 1, \ldots, n,\;\; k < n$, $(b'_{0n} \otimes \iota'_n)$, and $[(b'_{0n} \otimes \iota'_j) + (b'_{0j} \otimes \iota'_n)]$ for $j = 1, \ldots, n-1$. By some row-addition transformations, D can be transformed into C, and hence it is of full rank. Thus, by the invertible matrix theorem, $Dv = 0$ has only the trivial solution $v = 0$.

Appendix D Proof of Proposition 3

We consider $\theta = \text{vec}(A)$ that is more convenient to work with than $\vartheta = \text{vec}(B)$. Recall that $A = PB^{-1}$, where P is a signed $(n \times n)$ permutation matrix. The necessary condition for first-order local identification of B is that the expectation of the Jacobian matrix $E[J_T(\theta_0)]$, evaluated at θ_0, the true value of θ, has full column rank k. Thus, we need to show that $\text{rank}(E[\partial f(u_t, \theta_0)/\partial \theta']) = k$ for θ, the row vector of $k = n^2$ parameters to be estimated, and $f(u_t, \theta_0)$, the vector of population moment conditions, consisting of (16)–(18) and $n(n-1)/2$ asymmetric co-kurtosis conditions of the form $E(e_{it}^3 e_{jt}) = 0$ $(i \neq j)$. Because the row rank equals the column rank, it suffices to show that $k = n^2$ rows of $M(\theta_0) \equiv E[\partial f(u_t, \theta_0)/\partial \theta']$ are linearly independent. The $(n^2 + n(n-1)/2) \times n^2$ matrix $M(\theta_0)$ is obtained by stacking the n components of the form (D.1), the $n(n-1)/2$ components of the form (D.2), $n(n-1)/2$ components of the form (D.3), and $n(n-1)/2$ components of the form (D.4):

$$E\left[\left.\frac{\partial e_{it}^2}{\partial \theta'}\right|_{\theta=\theta_0}\right] = 2(b_{0i}' \otimes \iota_i'), \quad i = 1, \dots, n, \tag{D.1}$$

$$E\left[\left.\frac{\partial e_{it} e_{jt}}{\partial \theta'}\right|_{\theta=\theta_0}\right] = (b_{0i}' \otimes \iota_j') + (b_{0j}' \otimes \iota_i'), \quad i > j, \tag{D.2}$$

$$E\left[\left.\frac{\partial e_{it}^2 e_{jt}^2}{\partial \theta'}\right|_{\theta=\theta_0}\right] = 2(b_{0i}' \otimes \iota_i') + 2(b_{0j}' \otimes \iota_j'), \quad i > j, \tag{D.3}$$

$$E\left[\left.\frac{\partial e_{it}^3 e_{jt}}{\partial \theta'}\right|_{\theta=\theta_0}\right] = E(\varepsilon_{it}^4)(b_{0i}' \otimes \iota_j') + 3(b_{0j}' \otimes \iota_i'), \quad i \neq j, \tag{D.4}$$

where $i, j \in \{1, ..., n\}$, ι_i is the ith column of the $n \times n$ identity matrix, and b_{0i} is the ith column of B_0, the true value of B, (the expressions (D.1)–(D.3) are taken from (C.1)–(C.3) in Appendix C, and expression (D.4) is derived in the same manner as (C.1)–(C.3)).

Any of the rows given by (D.3) is obtained as a sum of two rows given by (D.1) (see the discussion Appendix C), and hence, only n of the rows corresponding to (D.1) and (D.3) are linearly independent. Also the $n + n(n-1)/2$ rows of the form (D.1) and (D.2) are readily seen to be linearly independent. Furthermore, they are also linearly independent of the $n(n-1)/2$ rows of the form (D.4), provided at most one of the components of ε_t is Gaussian (has zero excess kurtosis) and suitable asymmetric co-curtosis conditions are selected. To see this, suppose first that all n components of ε_t have positive excess kurtosis. In this case, generally $E(\varepsilon_{it}^4) \neq 3$ for all i, and it is not possible to express any of the $n(n-1)/2$ rows of the form (D.4) as a linear combination of the $n+n(n-1)/2$ rows given by (D.1) and (D.2), and because we thus have $n+n(n-1)/2+n(n-1)/2 = n^2 = k$ linearly independent rows in the Jacobian matrix $M(\theta_0)$, it is of full column rank. In contrast, if the ith component of ε_t is Gaussian, so that $E(\varepsilon_{it}^4) = 3$, one of the rows given by (D.4) may equal 3 times one of the rows given by (D.2). Then the Jacobian matrix is of reduced column rank. However, by inspecting (D.4), it is easy to see that if the asymmetric co-kurtosis conditions do not involve the third power of the element of ε_t that has zero excess kurtosis, the rows given by (D.2) and (D.4) are linearly independent, and the Jacobian matrix is of full column rank.

Appendix E Asymptotic distribution of the two-stage estimator

In this appendix, we derive the asymptotic distribution of the two-stage estimator of the parameters of the SVAR model, consisting of the OLS estimator $\hat{\pi}$ of $\pi = \text{vec}(\nu, A_1, \ldots, A_p)$ and the GMM estimator $\hat{\vartheta}$ of $\vartheta = \text{vec}(B)$ based on the OLS residuals u_t. The derivation is straightforward, with the most essential parts following Subsection 5.2 of the Supplementary Appendix to Gouriéroux et al. (2020).

Let us first write $T^{1/2}(\hat{\pi} - \pi_0)$ as

$$T^{1/2}(\hat{\pi} - \pi_0) = \left(\left(T^{-1} \sum_{t=1}^{T} Z_{t-1} Z_{t-1}' \right)^{-1} \otimes I_n \right) T^{-1/2} \sum_{t=1}^{T} \text{vec}(u_t Z_{t-1}'), \tag{E.1}$$

where $Z_t = (1, y_t', \ldots, y_{t-p+1}')'$, π_0 is the true value of π, and it is assumed that the probability limit F of $F_T = T^{-1} \sum_{t=1}^{T} Z_{t-1} Z_{t-1}'$ exists and is nonsingular. The above expression can be found in Lütkepohl (2005 Chapter 3.2).

The GMM estimator of ϑ given the OLS estimator $\hat{\pi}$ is given by

$$\hat{\vartheta} = \arg\min_{\vartheta} \frac{1}{T} \sum_{t=1}^{T} f(u_t(\hat{\pi}), \vartheta)' W_T \frac{1}{T} \sum_{t=1}^{T} f(u_t(\hat{\pi}), \vartheta) \tag{E.2}$$

with the positive semi-definite $(q \times q)$ matrix W_T (potentially dependent on the data) converging to a positive definite weighting matrix W. The two-stage GMM estimator in (E.2) is consistent and asymptotically normal under standard regularity conditions. It is important to realize, however, that its asymptotic covariance matrix depends on the OLS estimation error. It follows that the optimal weighting matrix also depends on the estimation error.

To derive the asymptotic covariance matrix and the optimal weighting matrix, we begin with the first order conditions for (E.2):

$$\frac{1}{T} \sum_{t=1}^{T} f(u_t(\hat{\pi}), \hat{\vartheta})' W_T \frac{1}{T} \sum_{t=1}^{T} \frac{\partial f(u_t(\hat{\pi}), \hat{\vartheta})}{\partial \vartheta'} = 0. \tag{E.3}$$

We simplify the presentation by defining $g_T(\pi, \vartheta) = T^{-1} \sum_{t=1}^{T} f(u_t(\pi), \vartheta)$, $G_T(\pi; \vartheta) = T^{-1} \sum_{t=1}^{T} \frac{\partial f(u_t(\pi), \vartheta)}{\partial \pi'}$, and $G_T(\vartheta; \pi) = T^{-1} \sum_{t=1}^{T} \frac{\partial f(u_t(\pi), \vartheta)}{\partial \vartheta'}$. The mean value theorem implies that

$$g_T(\hat{\pi}, \hat{\vartheta}) = g_T(\hat{\pi}, \vartheta_0) + G_T(\hat{\vartheta}, \vartheta_0, c_T^1; \hat{\pi})(\hat{\vartheta} - \vartheta_0), \tag{E.4}$$

and

$$g_T(\hat{\pi}, \vartheta_0) = g_T(\pi_0, \vartheta_0) + G_T(\hat{\pi}, \pi_0, c_T^2; \vartheta_0)(\hat{\pi} - \pi_0), \tag{E.5}$$

where ϑ_0 is the true value of ϑ, and the ith row of the $(q \times n^2)$ matrix $G_T(\hat{\vartheta}, \vartheta, c_T^1; \hat{\pi})$ is the corresponding row of $G_T(\bar{\vartheta}^{(i)}; \hat{\pi})$ with $\bar{\vartheta}^{(i)} = c_{T,i}^1 \vartheta_0 + (1 - c_{T,i}^1)\hat{\vartheta}$ for some $0 \leq c_{T,i}^1 \leq 1$ (see Hall (2005, Chapter 3.4.2)). Similaryly, the ith row of the $(q \times n(np+1))$ matrix $G_T(\hat{\pi}, \pi, c_T^2; \vartheta)$ is the corresponding row of $G_T(\bar{\pi}^{(i)}; \vartheta)$ with $\bar{\pi}^{(i)} = c_{T,i}^2 \pi_0 + (1 - c_{T,i}^2)\hat{\pi}$ for some $0 \leq c_{T,i}^2 \leq 1$. Here $c_T^1 = (c_{T,1}^1, \ldots, c_{T,q}^1)$ and $c_T^2 = (c_{T,1}^2, \ldots, c_{T,q}^2)$.

Substituting (E.5) into (E.4), premultiplying the resulting equation by $G_T(\hat{\vartheta}; \hat{\pi})' W_T$, and using (E.3), we obtain

$$\begin{aligned} T^{1/2}(\hat{\vartheta} - \vartheta_0) = &-[G_T(\hat{\vartheta}; \hat{\pi})' W_T G_T(\hat{\vartheta}, \vartheta_0, c_T^1; \hat{\pi})]^{-1} \\ &\times G_T(\hat{\vartheta}; \hat{\pi})' W_T [T^{1/2} g_T(\pi_0, \vartheta_0) + G_T(\hat{\pi}, \pi_0, c_T^2; \vartheta_0) T^{1/2}(\hat{\pi} - \pi_0)]. \end{aligned} \tag{E.6}$$

Based on the arguments in Hall (2005, Chapter 3.4.2), both $G_T(\hat{\vartheta}; \hat{\pi})$ and $G_T(\hat{\vartheta}, \vartheta, c_T^1; \hat{\pi})$ converge in probability to $G_\vartheta = E\left[\frac{\partial f(u_t(\pi_0), \vartheta_0)}{\partial \vartheta'}\right]$. Similarly, $G_T(\hat{\pi}, \pi_0, c_T^2; \vartheta_0)$ converges in probability to $G_\pi = E\left[\frac{\partial f(u_t(\pi_0), \vartheta_0)}{\partial \pi'}\right]$. Hence, (E.6) together with (E.1) implies that for $T \to \infty$,

$$T^{1/2} \begin{bmatrix} \hat{\pi} - \pi \\ \hat{\vartheta} - \vartheta \end{bmatrix} \approx \begin{bmatrix} F^{-1} \otimes I_n & 0 \\ -I_{\vartheta\vartheta}^{-1} I_{\vartheta\pi}(F^{-1} \otimes I_n) & -I_{\vartheta\vartheta}^{-1} G_\vartheta' W \end{bmatrix} \begin{bmatrix} T^{-1/2} \sum_{t=1}^T \operatorname{vec}(u_t Z_{t-1}') \\ T^{-1/2} \sum_{t=1}^T f(u_t, \vartheta_0) \end{bmatrix}. \tag{E.7}$$

where $I_{\vartheta\vartheta} = G_\vartheta' W G_\vartheta$, $I_{\vartheta\pi} = G_\vartheta' W G_\pi$ and $F = E\left(Z_{t-1} Z_{t-1}'\right)$.

As shown in Lütkepohl (2005, Chapter 3.2.2) and Hall (2005, Chapter 3.4.2), respectively, the zero mean vectors $T^{-1} \sum_{t=1}^T \operatorname{vec}(u_t Z_{t-1}')$ and $T^{-1} \sum_{t=1}^T f(u_t, \vartheta_0)$ in the latter matrix are asymptotically normally distributed under standard assumptions. Thus, based on the properties of the multivariate normal distribution, we obtain

$$T^{-1/2} \begin{bmatrix} \sum_{t=1}^T \operatorname{vec}(u_t Z_{t-1}') \\ \sum_{t=1}^T f(u_t, \vartheta) \end{bmatrix} \xrightarrow{d} N(0, H), \tag{E.8}$$

with

$$H = \lim_{T\to\infty} Var\left[T^{1/2}\begin{pmatrix} T^{-1}\sum_{t=1}^{T}\text{vec}(u_t Z_{t-1}') \\ T^{-1}\sum_{t=1}^{T} f(u_t,\vartheta_0)\end{pmatrix}\right], \tag{E.9}$$

the long-run covariance matrix of all moment conditions. Combining (E.7) and (E.8), we have (see, e.g., Hall (2005, Lemma 1.4))

$$T^{1/2}\begin{bmatrix}\hat{\pi}-\pi \\ \hat{\vartheta}-\vartheta\end{bmatrix} \xrightarrow{d} N(0,\Omega), \tag{E.10}$$

where

$$\Omega = \begin{bmatrix} F^{-1}\otimes I_n & 0 \\ -I_{\vartheta\vartheta}^{-1}I_{\vartheta\pi}(F^{-1}\otimes I_n) & -I_{\vartheta\vartheta}^{-1}G_\vartheta' W\end{bmatrix} H \begin{bmatrix} F^{-1\prime}\otimes I_n & -(F^{-1\prime}\otimes I_n)I_{\vartheta\pi}'I_{\vartheta\vartheta}^{-1} \\ 0 & -W'G_\vartheta I_{\vartheta\vartheta}^{-1}\end{bmatrix}. \tag{E.11}$$

Hence, according to (E.10) above

$$T^{1/2}(\hat{\vartheta}-\vartheta) \xrightarrow{d} N(0,\Omega_{\vartheta\vartheta}), \tag{E.12}$$

where

$$\begin{aligned}
\Omega_{\vartheta\vartheta} &= \left[I_{\vartheta\vartheta}^{-1}I_{\vartheta\pi}(F^{-1}\otimes I_n) \quad I_{\vartheta\vartheta}^{-1}G_\vartheta' W\right] H \left[I_{\vartheta\vartheta}^{-1}I_{\vartheta\pi}(F^{-1}\otimes I_n) \quad I_{\vartheta\vartheta}^{-1}G_\vartheta' W\right]' \\
&= I_{\vartheta\vartheta}^{-1}G_\vartheta'[WG_\pi(F^{-1}\otimes I_n) \quad W]H[WG_\pi(F^{-1}\otimes I_n) \quad W]'G_\vartheta I_{\vartheta\vartheta}^{-1} \\
&= I_{\vartheta\vartheta}^{-1}G_\vartheta' W H_0 W' G_\vartheta I_{\vartheta\vartheta}^{-1} \\
&= [G_\vartheta' W G_\vartheta]^{-1}G_\vartheta' W H_0 W' G_\vartheta [G_\vartheta' W G_\vartheta]^{-1}
\end{aligned} \tag{E.13}$$

with

$$H_0 = [G_\pi(F^{-1}\otimes I_n) \quad I_q]H[G_\pi(F^{-1}\otimes I_n) \quad I_q]' \tag{E.14}$$

The optimal weighting matrix is obtained by setting $W = H_0^{-1}$ (see Gouriéroux et al. (2020)), and, in this case, (E.13) reduces to

$$\Omega_{\vartheta\vartheta} = [G_\vartheta' H_0^{-1} G_\vartheta]^{-1}. \tag{E.15}$$

Finally, a consistent estimator of Ω, $\hat{\Omega}_T$, is obtained by replacing G_ϑ and G_π by their consistent estimators $G_T(\hat{\vartheta};\hat{\pi}) = T^{-1}\sum_{t=1}^{T}\frac{\partial f(u_t(\hat{\pi}),\hat{\vartheta})}{\partial\vartheta'}$ and $G_T(\hat{\pi};\hat{\vartheta}) = T^{-1}\sum_{t=1}^{T}\frac{\partial f(u_t(\hat{\pi}),\hat{\vartheta})}{\partial\pi'}$, respectively, and H is estimated consistently under quite general conditions discussed in Hall (2005 Chapter 3.5) by a member of the class of HAC covariance matrix estimators, including the Newey-West estimator. Also, F needs to be replaced by its consistent estimator $F_T = T^{-1}\sum_{t=1}^{T} Z_{t-1}Z_{t-1}'$.

Appendix F Derivation of equations (A.2) and (A.3)

We begin with (A.2), and then derive (A.3). Using $e_t = Q\varepsilon_t$ in (A.1), condition (18) can be written as

$$E(e_{it}^2 e_{jt}^2) - 1 = E[(Q_{i1}\varepsilon_{1t} + \cdots + Q_{in}\varepsilon_{nt})^2 (Q_{j1}\varepsilon_{1t} + \cdots + Q_{jn}\varepsilon_{nt})^2] - 1 = 0. \tag{F.1}$$

By straightforward manipulation of the squared quantities $(Q_{i1}\varepsilon_{1t} + \cdots + Q_{in}\varepsilon_{nt})^2$, we obtain

$$\left(\sum_{k=1}^{n} Q_{ik}\varepsilon_{kt}\right)^2 = \sum_{k=1}^{n} Q_{ik}^2 \varepsilon_{kt}^2 + \sum_{l=1}^{n} \sum_{m=1,\ldots,n;m\neq l} Q_{il}Q_{im}\varepsilon_{lt}\varepsilon_{mt}. \tag{F.2}$$

Using this result, by Assumption 1(ii) that the components of the error term ε_t have no excess co-kurtosis, we have that

$$\begin{aligned} E(e_{it}^2 e_{jt}^2) = {} & E\left[\left(\sum_{k=1}^{n} Q_{ik}^2 \varepsilon_{kt}^2\right)\left(\sum_{k=1}^{n} Q_{jk}^2 \varepsilon_{kt}^2\right)\right] \\ & + E\left[\left(\sum_{l=1}^{n} \sum_{m=1,\ldots,n;m\neq l} Q_{il}Q_{im}\varepsilon_{lt}\varepsilon_{mt}\right)\left(\sum_{l=1}^{n} \sum_{m=1,\ldots,n;m\neq l} Q_{jl}Q_{jm}\varepsilon_{lt}\varepsilon_{mt}\right)\right]. \end{aligned} \tag{F.3}$$

Again, from Assumption 2(ii), by adding and subtracting $\sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2$, the first term on the right hand side of (F.3) yields

$$\begin{aligned} E\left[\left(\sum_{k=1}^{n} Q_{ik}^2 \varepsilon_{kt}^2\right)\left(\sum_{k=1}^{n} Q_{jk}^2 \varepsilon_{kt}^2\right)\right] &= \sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2 E(\varepsilon_{kt}^4) - \sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2 \\ &\quad + \left(\sum_{k=1}^{n} Q_{ik}^2\right)\left(\sum_{k=1}^{n} Q_{jk}^2\right) \\ &= \sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2 E(\varepsilon_{kt}^4) - \sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2 + 1, \end{aligned} \tag{F.4}$$

where in the last equality we use the orthogonality of Q ($\sum_{k=1}^{n} Q_{ik}^2 = 1$).

We now turn to the latter term on the right hand side of (F.3), which we denote by K:

$$K = E\left[\left(\sum_{l=1}^{n} \sum_{m=1,\ldots,n;m\neq l} Q_{il}Q_{im}\varepsilon_{lt}\varepsilon_{mt}\right)\left(\sum_{l=1}^{n} \sum_{m=1,\ldots,n;m\neq l} Q_{jl}Q_{jm}\varepsilon_{lt}\varepsilon_{mt}\right)\right].$$

Assumptions 1(ii)–(iii) imply that $E(\varepsilon_{it}\varepsilon_{jt}\varepsilon_{kt}\varepsilon_{lt}) = 1$ when $i = k, j = l \neq k$ (or $i = l, j = k \neq l$ or $i = j \neq k = l$) and zero otherwise (excluding the case $i = j = k = l$). Therefore, we find that

$$
\begin{aligned}
K &= 2\sum_{l=1}^{n}\sum_{m=1,\ldots,n;m\neq l} Q_{il}Q_{im}Q_{jl}Q_{jm} \\
&= 2\sum_{l=1}^{n}\sum_{m=1}^{n} Q_{il}Q_{jl}Q_{im}Q_{jm} - 2\sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2 \\
&= 2\left(\sum_{m=1}^{n} Q_{im}Q_{jm}\right)^2 - 2\sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2 \\
&= -2\sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2, \qquad \text{(F.5)}
\end{aligned}
$$

where the last equality holds due to the orthogonality of Q ($\sum_{m=1}^{n} Q_{im}Q_{jm} = 0$).

Combining the above results, condition (18) can be expressed as

$$
\begin{aligned}
E(e_{it}^2 e_{jt}^2) - 1 &= \sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2 E(\varepsilon_{kt}^4) - 3\sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2 \\
&= \sum_{k=1}^{n} Q_{ik}^2 Q_{jk}^2 (E(\varepsilon_{kt}^4) - 3) = 0, \qquad \text{(F.6)}
\end{aligned}
$$

which is Equation (A.2) in Appendix A.

We now derive (A.3). Based on $e = Q\varepsilon$ (A.1), the asymmetric co-kurtosis conditions can be written as

$$
\begin{aligned}
E(e_{it}^3 e_{jt}) &= E[(Q_{i1}\varepsilon_{1t} + \cdots + Q_{in}\varepsilon_{nt})^3 (Q_{j1}\varepsilon_{1t} + \cdots + Q_{jn}\varepsilon_{nt})] \\
&= E\left[\left(\sum_{k=1}^{n} Q_{ik}\varepsilon_{kt}\right)^2 \left(\sum_{k=1}^{n} Q_{ik}\varepsilon_{kt}\right)\left(\sum_{k=1}^{n} Q_{jk}\varepsilon_{kt}\right)\right] \\
&= 0. \qquad \text{(F.7)}
\end{aligned}
$$

Straightforward computations yield

$$
\left(\sum_{k=1}^{n} Q_{ik}\varepsilon_{kt}\right)\left(\sum_{k=1}^{n} Q_{jk}\varepsilon_{kt}\right) = \sum_{k=1}^{n} Q_{ik}Q_{jk}\varepsilon_{kt}^2 + \sum_{l=1}^{n}\sum_{m=1,\ldots,n;m\neq l} Q_{il}Q_{jm}\varepsilon_{lt}\varepsilon_{mt}. \qquad \text{(F.8)}
$$

Multiplying the above equation by (F.2) and taking expectations result in

$$
\begin{aligned}
E(e_{it}^3 e_{jt}) = {} & E\left[\left(\sum_{k=1}^{n} Q_{ik}^2\varepsilon_{kt}^2\right)\left(\sum_{k=1}^{n} Q_{ik}Q_{jk}\varepsilon_{kt}^2\right)\right] \\
& + E\left[\left(\sum_{l=1}^{n}\sum_{m=1,\ldots,n;m\neq l} Q_{il}Q_{im}\varepsilon_{l}\varepsilon_{mt}\right)\left(\sum_{l=1}^{n}\sum_{m=1,\ldots,n;m\neq l} Q_{il}Q_{jm}\varepsilon_{lt}\varepsilon_{mt}\right)\right] \qquad \text{(F.9)}
\end{aligned}
$$

by Assumption 2(ii) that the components of the error term ε_t have no excess co-kurtosis. For the same reason, the first term on the right hand side of (F.9) yields

$$E\left[\left(\sum_{k=1}^{n} Q_{ik}^2 \varepsilon_{kt}^2\right)\left(\sum_{k=1}^{n} Q_{ik} Q_{jk} \varepsilon_{kt}^2\right)\right] = \sum_{k=1}^{n} Q_{ik}^3 Q_{jk} E(\varepsilon_{kt}^4) - \sum_{k=1}^{n} Q_{ik}^3 Q_{jk} + \sum_{l=1}^{n}\sum_{m=1}^{n} Q_{il}^2 Q_{im} Q_{jm}$$

$$= \sum_{k=1}^{n} Q_{ik}^3 Q_{jk} E(\varepsilon_{kt}^4) - \sum_{k=1}^{n} Q_{ik}^3 Q_{jk} + \left(\sum_{k=1}^{n} Q_{ik}^2\right)\left(\sum_{k=1}^{n} Q_{ik} Q_{jk}\right)$$

$$= \sum_{k=1}^{n} Q_{ik}^3 Q_{jk} E(\varepsilon_{kt}^4) - \sum_{k=1}^{n} Q_{ik}^3 Q_{jk}, \qquad \text{(F.10)}$$

where in the first equality we add and subtract the term $\sum_{k=1}^{n} Q_{ik}^3 Q_{jk}$, and in the last equality we use the orthogonality of Q.

We denote by M the last term on right hand side of (F.9):

$$M = E\left[\left(\sum_{l=1}^{n} \sum_{m=1,\dots,n; m\neq l} Q_{il} Q_{im} \varepsilon_{lt} \varepsilon_{mt}\right)\left(\sum_{l=1}^{n} \sum_{m=1,\dots,n; m\neq l} Q_{il} Q_{jm} \varepsilon_{lt} \varepsilon_{mt}\right)\right].$$

Now, recall that Assumptions 1(ii)–(iii) imply that $E(\varepsilon_{it}\varepsilon_{jt}\varepsilon_{kt}\varepsilon_{lt}) = 1$ when $i = k, j = l \neq k$ (or $i = l, j = k \neq l$ or $i = j \neq k = l$) and zero otherwise (excluding the case $i = j = k = l$). Therefore, by straightforward calculations, from the last term on right hand side of (F.9), we find that

$$M = 2\sum_{l=1}^{n} \sum_{m=1,\dots,n; m\neq l} Q_{il} Q_{im} Q_{il} Q_{jm}$$

$$= 2\sum_{l=1}^{n}\sum_{m=1}^{n} Q_{il}^2 Q_{im} Q_{jm} - 2\sum_{k=1}^{n} Q_{ik}^3 Q_{jk}$$

$$= 2\left(\sum_{k=1}^{n} Q_{ik}^2\right)\left(\sum_{k=1}^{n} Q_{ik} Q_{jk}\right) - 2\sum_{k=1}^{n} Q_{ik}^3 Q_{jk}$$

$$= -2\sum_{k=1}^{n} Q_{ik}^3 Q_{jk}, \qquad \text{(F.11)}$$

where in the last equality we use the orthogonality of Q.

Putting (F.9)–(F.11) together, we obtain

$$E(e_i^3 e_j) = \sum_{k=1}^{n} Q_{ik}^3 Q_{jk} E(\varepsilon_{kt}^4) - 3 \sum_{k=1}^{n} Q_{ik}^3 Q_{jk}$$
$$= \sum_{k=1}^{n} Q_{ik}^3 Q_{jk} (E(\varepsilon_{kt}^4) - 3) = 0, \qquad \text{(F.12)}$$

which is Equation (A.3) in Appendix A.

Appendix G Derivation of equations (C.1)–(C.3) and (C.6)–(C.8)

We first calculate the first-order derivatives of the functions e_{it}^2, $e_{it}e_{jt}$, and $e_{it}^2 e_{jt}^2$ in (16)–(18) with respect to $\theta = \text{vec}(A)$, and then their second-order derivatives. After computing the second-order derivatives, we evaluate them at the true parameter value θ_0, and take expectations using the fact that at the true parameter value $u_t = B_0\varepsilon_t$ and $e_t = \varepsilon_t$. Several results from Seber (2008) are frequently used below; for brevity, result 17.30(h), say, is referred to as S 17.30(h).

We start by defining $e_{it} = \iota_i' A u_t = (u_t' \otimes \iota_i')\text{vec}(A)$ with ι_i the ith unit vector. Straightforward differentiation based on S 17.20(a) yields

$$\left[\frac{\partial e_{it}^2}{\partial \theta}\right] = 2e_{it}(u_t \otimes \iota_i), \quad i = 1, \ldots, n. \tag{G.1}$$

While applying S 17.30(h) to the scalar functions and then using S 17.20(a) yield

$$\left[\frac{\partial e_{it}e_{jt}}{\partial \theta}\right] = e_{jt}(u_t \otimes \iota_i) + e_{it}(u_t \otimes \iota_j), \quad i > j,\ i,j = 1, \ldots, n. \tag{G.2}$$

Using the same two results from Seber (2008), we obtain

$$\left[\frac{\partial e_{it}^2 e_{jt}^2}{\partial \theta}\right] = 2e_{jt}^2 e_{it}(u_t \otimes \iota_i)) + 2e_{it}^2 e_{jt}(u_t \otimes \iota_j)). \quad i > j,\ i,j = 1, \ldots, n \tag{G.3}$$

The expressions (C.1)–(C.3) in Appendix C are obtained by evaluating (G.1)–(G.3) at the true parameter value θ_0, substituting $u_t = B_0\varepsilon_t$ and $e_t = \varepsilon_t$ into the resulting equations, taking expectations of both sides of these equations using Assumption 2(ii) that the components of the error term ε_t are orthogonal and have no excess co-kurtosis, and transposing.

Based on (G.1)–(G.3), let us next derive the second-second order derivatives. By direct computation using S 17.20(a), we obtain from (G.1)

$$\begin{aligned}\left[\frac{\partial e_{it}^2}{\partial \theta \partial \theta'}\right] &= 2(u_t \otimes \iota_i)(u_t' \otimes \iota_i') \\ &= 2(u_t u_t' \otimes \iota_i \iota_i'), \quad i = 1, \ldots, n,\end{aligned} \tag{G.4}$$

where the latter equality follows from S 11.11(a). In the same manner, from (G.2) we have

$$\left[\frac{\partial e_{it}e_{jt}}{\partial\theta\partial\theta'}\right] = (u_t \otimes \iota_i)(u_t' \otimes \iota_j') + (u_t \otimes \iota_j)(u_t' \otimes \iota_i')$$
$$= (u_t u_t' \otimes \iota_i \iota_j') + (u_t u_t' \otimes \iota_j \iota_i')$$
$$= (u_t u_t') \otimes (\iota_i \iota_j' + \iota_j \iota_i'), \quad i > j,\ i, j = 1, \ldots, n \tag{G.5}$$

where the last two equalities are obtained using S 11.11(a) and 11.10(b), respectively.

Applying S 17.30(h) to the scalar functions $e_{jt}^2 e_{it}$ and $e_{it}^2 e_{jt}$, and then using S 17.20(a), we obtain

$$\left[\frac{\partial e_{it}^2 e_{jt}^2}{\partial\theta\partial\theta'}\right] = 2(u_t \otimes \iota_i)\left[e_{jt}^2(u_t' \otimes \iota_i') + 2e_{it}e_{jt}(u_t' \otimes \iota_j')\right]$$
$$+ 2(u_t \otimes \iota_j)\left[2e_{it}e_{jt}(u_t' \otimes \iota_i') + e_{it}^2(u_t' \otimes \iota_j')\right]$$
$$= 2e_{jt}^2(u_t u_t' \otimes \iota_i \iota_i') + 2e_{it}^2(u_t u_t' \otimes \iota_j \iota_j')$$
$$+ 4e_{it}e_{jt}(u_t u_t') \otimes (\iota_i \iota_j' + \iota_j \iota_i'), \quad i > j,\ i, j = 1, \ldots, n \tag{G.6}$$

where the latter equality follows from S 11.11(a) and 11.10(b).

Evaluating (G.4) and (G.5) at the true parameter value θ_0, substituting $u_t = B_0\varepsilon_t$ and $e_t = \varepsilon_t$ into the resulting equations, and taking expectations, we obtain expressions (C.6)–(C.7) in Appendix C:

$$E\left[\frac{\partial e_{it}^2}{\partial\theta\partial\theta'}\right]_{\theta=\theta_0} = 2(B_0 E(\varepsilon_t \varepsilon_t') B_0' \otimes \iota_i \iota_i')$$
$$= 2(B_0 B_0') \otimes (\iota_i \iota_i') \quad i = 1, \ldots, n, \tag{G.7}$$

$$E\left[\frac{\partial e_{it}e_{jt}}{\partial\theta\partial\theta'}\right]_{\theta=\theta_0} = (B_0 E(\varepsilon_t \varepsilon_t') B_0') \otimes (\iota_i \iota_j' + \iota_j \iota_i')$$
$$= (B_0 B_0') \otimes (\iota_i \iota_j' + \iota_j \iota_i'), \quad i > j,\ i, j, k, l = 1, \ldots, n, \tag{G.8}$$

where the latter equalities hold due to Assumption 2(ii)–(iii): $E(\varepsilon_t \varepsilon_t') = I_n$. Similarly, evaluating (G.6) at θ_0, substituting $u_t = B_0\varepsilon_t$ and $e_t = \varepsilon_t$ into the resulting equation, and taking expectations, we obtain expression (C.8) in Appendix C:

$$\begin{aligned} E\left[\frac{\partial e_{it}^2 e_{jt}^2}{\partial\theta\partial\theta'}\right]_{\theta=\theta_0} &= 2(B_0E[\varepsilon_{jt}^2(\varepsilon_t\varepsilon_t')]B_0' \otimes \iota_i\iota_i') + 2(B_0E[\varepsilon_{it}^2(\varepsilon_t\varepsilon_t')]B_0' \otimes \iota_j\iota_j') \\ &\quad + 4(B_0E[\varepsilon_{it}\varepsilon_{jt}(\varepsilon_t\varepsilon_t')]B_0') \otimes (\iota_i\iota_j' + \iota_j\iota_i') \\ &= 2(B_0\Xi_jB_0') \otimes (\iota_i\iota_i') + 2(B_0\Xi_iB_0') \otimes (\iota_j\iota_j') \\ &\quad + 4(B_0\Upsilon_{i,j}B_0') \otimes (\iota_i\iota_j' + \iota_j\iota_i'), \quad i > j,\ i,j = 1,\ldots,n, \end{aligned} \tag{G.9}$$

where $\Xi_i = E[\varepsilon_{it}^2(\varepsilon_t\varepsilon_t')]$, and $\Upsilon_{i,j} = E[\varepsilon_{it}\varepsilon_{jt}(\varepsilon_t\varepsilon_t')]$.

References

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