Online Appendix

Non-Gaussian Structural Dynamic Factor Models:
Identification and Estimation

Andrzej Kocięcki, Christian Matthes, and Michele Piffer

O.1 Posterior simulation

This appendix gives the conditional-regression updates for the Student-t specification in Section 5. Write $\Lambda = (\Lambda_0, \ldots, \Lambda_s)$, $\Pi = (\Pi_1, \ldots, \Pi_l)$, $\Phi = (\Phi_1, \ldots, \Phi_p)$, and $D = (D_1, \ldots, D_T)$, and let f collect the factor path and its required initial lags. A vertical bar followed by $\bullet$ denotes conditioning on all other blocks. The regression formulas below use the observations for which the requisite lags are available.

O.1.1 Sampling the dynamic factors $f \mid \bullet$

Substituting the observation equation into the VAR for the Gaussian component gives

$$\widetilde{y}_t = \sum_{a=0}^{s+l} K_a f_{t-a} + e_t, \tag{O.1}$$

$$\widetilde{y}_t = y_t - \sum_{h=1}^{l} \Pi_h y_{t-h}, \qquad K_a = \Lambda_a - \sum_{h=1}^{l} \Pi_h \Lambda_{a-h},$$

where $\Lambda_a = 0$ outside $0 \le a \le s$. Let $d_f = \max\{p, s+l+1\}$ and $s_t = (f_t', \ldots, f_{t-d_f+1}')'$. With zero padding of missing coefficient blocks, the state-space system is

$$s_t = A_f s_{t-1} + J_f \varepsilon_t^*, \qquad \varepsilon_t^* \sim N(0, B D_t B'), \tag{O.2}$$

$$\widetilde{y}_t = C_f s_t + e_t, \qquad e_t \sim N(0, \Omega), \tag{O.3}$$

Here A_f is the companion matrix with first block row $(\Phi_1, \ldots, \Phi_p, 0, \ldots, 0)$. The input and observation matrices are

$$J_f = (I_q, 0, \ldots, 0)', \qquad C_f = (K_0, \ldots, K_{s+l}, 0, \ldots, 0).$$

We draw the states jointly using the multimove method of Carter and Kohn (1994) and a square-root Kalman filter, as in Bai and Wang (2015). The initial state has the fixed Gaussian distribution specified in footnote 14.

O.1.2 Sampling factor loadings $\Lambda \mid \bullet$

Conditional on the factors and Π, define

$$W_t^a = f'_{t-a} \otimes I_n - \sum_{h=1}^{l} f'_{t-a-h} \otimes \Pi_h, \qquad a = 0, \ldots, s.$$

Then $\widetilde{y}_t = \sum_{a=0}^{s} W_t^a \operatorname{vec}(\Lambda_a) + e_t$. Stack these equations over the usable dates. Any linear normalization can be written as $\operatorname{vec}(\Lambda) = s_\lambda + S_\lambda \lambda$, where λ contains the free loading coefficients. With $\mathcal{W}$ denoting the stacked unrestricted design matrix, set $Y_\lambda = \widetilde{y} - \mathcal{W} s_\lambda$, $W = \mathcal{W} S_\lambda$, and $V = I \otimes \Omega$. Under the prior $\lambda \sim N(\mu_\lambda, \Sigma_\lambda)$,

$$V_\lambda = (W' V^{-1} W + \Sigma_\lambda^{-1})^{-1}, \tag{O.4}$$

$$m_\lambda = V_\lambda (W' V^{-1} Y_\lambda + \Sigma_\lambda^{-1} \mu_\lambda), \tag{O.5}$$

$$\lambda \mid \bullet \sim N(m_\lambda, V_\lambda). \tag{O.6}$$

The named factor normalization fixes the first q rows of Λ_0 to I_q and leaves the remaining coefficients free.

O.1.3 Sampling the Gaussian VAR coefficients $\Pi \mid \bullet$

Conditional on the dynamic factors and Λ, recover $u_t = y_t - \sum_{a=0}^{s} \Lambda_a f_{t-a}$. Write the regression as $U = \Pi X + E$, where $U = (u_{l+1}, \ldots, u_T)$ and each column of X stacks the l lags of u_t. Let $N_e = T - l$ and $k = nl$. The likelihood is

$$p(U \mid \Pi, \Omega) \propto |\Omega|^{-N_e/2} \exp\{-\tfrac{1}{2} \operatorname{tr}[\Omega^{-1}(U - \Pi X)(U - \Pi X)']\}. \tag{O.7}$$

For the unrestricted normal-prior specification, use the joint prior kernel

$$p(\Pi, \Omega) \propto IW(\Omega; \bar{Q}, \bar{\nu}) MN(\Pi; M_0, \Omega, V_0) \mathbf{1}_{\mathcal{S}}(\Pi), \tag{O.8}$$

where $\mathcal{S}$ is the set of stable coefficient matrices. The matrix-normal density has covariance $V_0 \otimes \Omega$ for $\operatorname{vec}(\Pi)$. We set $M_0 = 0$ and $V_0 = 0.1 I_k$. The stability restriction applies to the joint density, so the marginal prior for Ω after imposing this restriction need not be inverse Wishart. The conditional update is matrix normal with row covariance Ω and

$$C = (XX' + V_0^{-1})^{-1}, \qquad M = (UX' + M_0 V_0^{-1}) C, \tag{O.9}$$

as column covariance and mean, respectively. Reject coefficient draws outside $\mathcal{S}$.

O.1.4 Sampling the Gaussian innovation covariance $\Omega \mid \bullet$

Let $E = U - \Pi X$. Combining the likelihood with the joint prior gives

$$p(\Omega \mid \bullet) \propto |\Omega|^{-(N_e+\bar{\nu}+k+n+1)/2} \times \exp\{-\tfrac{1}{2}\operatorname{tr}(\Omega^{-1}Q_*)\}, \tag{O.10}$$

$$Q_* = \bar{Q} + EE' + (\Pi - M_0)V_0^{-1}(\Pi - M_0)'. \tag{O.11}$$

Hence

$$\Omega \mid \bullet \sim IW(Q_*, \bar{\nu} + N_e + k). \tag{O.12}$$

The coefficient prior contributes both its quadratic term and k degrees of freedom because $|V_0 \otimes \Omega| = |V_0|^n|\Omega|^k$. Conditioning on Π does not remove its prior density's dependence on Ω. The joint stability indicator is constant in Ω at the current admissible Π draw. When $l = 0$, there is no coefficient-prior contribution. A flat coefficient prior would omit the quadratic term and the additional k degrees of freedom; the reported unrestricted runs use the conditional normal prior.

In the diagonal-noise specification, Π_h and Ω are diagonal. Each series is updated as a separate stationary univariate autoregression with a normal coefficient prior conditional on its innovation variance and an inverse-gamma variance prior. Its variance conditional includes the corresponding coefficient-prior quadratic and determinant terms.

O.1.5 Sampling the factor VAR coefficients $\Phi \mid \bullet$

Let $\bar{f}_{t-1} = (f'_{t-1}, \ldots, f'_{t-p})'$ and $X_t = \bar{f}'_{t-1} \otimes I_q$. Then

$$f_t = X_t \operatorname{vec}(\Phi) + BD_t^{1/2}\varepsilon_t. \tag{O.13}$$

Write any coefficient restrictions as $\operatorname{vec}(\Phi) = s_\phi + S_\phi\phi$; under the named factor normalization $s_\phi = 0$ and $S_\phi = I$. Define $Y_t = f_t - X_t s_\phi$, $W_t = X_t S_\phi$, and $Q_t = BD_tB'$. For a normal prior $\phi \sim N(\mu_\phi, V_\phi)$, the Gaussian conditional has covariance and mean

$$V_* = \left(V_\phi^{-1} + \sum_t W_t'Q_t^{-1}W_t\right)^{-1}, \tag{O.14}$$

$$m_* = V_*\left(V_\phi^{-1}\mu_\phi + \sum_t W_t'Q_t^{-1}Y_t\right). \tag{O.15}$$

Reject draws that imply unstable factor dynamics. All reported runs use $\mu_\phi = 0$ and a diagonal Minnesota-type prior covariance with standard deviations $0.5/h^2$ for own-lag

coefficients and $0.25\widehat{\sigma}_i/(h^2\widehat{\sigma}_j)$ for the lag-h coefficient on factor j in equation $i \neq j$. Here $\widehat{\sigma}_i$ is the residual standard deviation from a univariate autoregression for the i-th named observable. The scales are fixed before posterior simulation.

O.1.6 Sampling $B \mid \bullet$, $D \mid \bullet$, and $\nu \mid \bullet$

Conditional on the factors and Φ, the residuals $f_t - \sum_{h=1}^{p} \Phi_h f_{t-h}$ give the model studied by Kocięcki et al. (2025). Let $\epsilon_t = B^{-1}(f_t - \sum_{h=1}^{p} \Phi_h f_{t-h})$. The independent scale updates are

$$d_{jt} \mid \bullet \sim IG\left(\frac{\nu_j + 1}{2}, \frac{\nu_j - 2 + \epsilon_{jt}^2}{2}\right). \tag{O.16}$$

We discretize the support of ν_j above four and use the Griddy-Gibbs method of Ritter and Tanner (1992). For B, the factorization $K = B^{-1} = \Gamma LU$ in Kocięcki et al. (2025) uses $\Gamma = \operatorname{diag}(\gamma_1, \ldots, \gamma_q)$ and unit-diagonal lower- and upper-triangular matrices L and U.

The prior flat in B induces $p(K) \propto |\det K|^{-2q}$. The Jacobian for the second transformation is

$$\left|\frac{\partial \operatorname{vec} K}{\partial(\gamma, L, U)}\right| = \prod_{j=1}^{q} |\gamma_j|^{q-1}, \qquad p(\gamma, L, U) \propto \prod_{j=1}^{q} |\gamma_j|^{-q-1},$$

where derivatives with respect to L and U refer to their free entries. Let $v_t = f_t - \sum_{h=1}^{p} \Phi_h f_{t-h}$ and let N denote the number of usable transition observations. Conditional on the scales D_t, the likelihood contributes $\prod_j |\gamma_j|^N$. Consequently,

$$\gamma_j^2 \mid \bullet \sim \text{Gamma}\left(\frac{N - q}{2}, \frac{1}{2}\sum_{t=1}^{N} \frac{[(LUv_t)_j]^2}{d_{jt}}\right), \tag{O.17}$$

in shape–rate form, with independent equiprobable signs for γ_j. The induced prior is constant in the free entries of L and U; their conditional updates follow Kocięcki et al. (2025). The generalized likelihood-preserving normalization in (32) resolves the remaining signs and permutations. Online Appendix O.4 describes the reference impact matrices and simulation lengths used in each experiment.

O.1.7 Shock normalization and the reference impact matrix

To solve (32), we match rows and columns of $\widehat{B}^{-1}B$ to maximize the sum of absolute matched entries, with each row and column used once. We then choose signs so that the matched entries are positive. This is a linear assignment problem. The resulting signed permutation is applied jointly to B and all shock-specific quantities. With the named factor normalization

and static loadings, the first q standardized observables have impact matrix B. Their true impact block therefore supplies the Monte Carlo reference in factor coordinates.

In empirical work, we construct $\widehat{B}$ from a preliminary sampler run under the named factor normalization. Let $\widehat{f}_t$ and $\widehat{\Phi}_h$ denote posterior means from that run. We maximize the independent, unit-variance Student-t likelihood of the residuals $\widehat{f}_t - \sum_h \widehat{\Phi}_h \widehat{f}_{t-h}$ over an unrestricted impact matrix and the degrees of freedom. This is a quasi-maximum-likelihood estimate based on the estimated factor residuals. We orient the estimated impact matrix once using (32) with reference I_q, then hold the resulting $\widehat{B}$ fixed throughout the final sampler run. The identity reference fixes signs and ordering without imposing zero restrictions.

O.2 Proofs of the rank and minimality results

O.2.1 The proof of Proposition 1

First we prove implication for Assumption 6. Suppose $[\Lambda_0 \vdots \Lambda_1 \vdots \ldots \vdots \Lambda_s]$ has full column rank and Λ_0 has no full column rank i.e. $rank(\Lambda_0) < q$. Since $q(s+1) = rank[\Lambda_0 \vdots \Lambda_1 \vdots \ldots \vdots \Lambda_s] \leq rank[\Lambda_0] + rank[\Lambda_1 \vdots \ldots \vdots \Lambda_s] \leq rank[\Lambda_0] + sq < q + sq = q(s+1)$ i.e. contradiction. We conclude that if $[\Lambda_0 \vdots \Lambda_1 \vdots \ldots \vdots \Lambda_s]$ has full column rank it follows Λ_0 has full column rank. Q.E.D.

We turn to proving implication for Assumption 5. For further analysis we need to deal with nonsingular $I_{\tau q} - zA$. To this end we have to exclude $z \in \mathbb{C}$ such that $det(I_{\tau q} - zA) = 0$. The latter is a finite set and will be denoted as $\gamma(A)$ (these are the reciprocals of eigenvalues of A except possible zero eigenvalue). Recall the system matrix (12). We need the instrumental lemma:

Lemma O.1. *For all $z \in \mathbb{C} \,|\gamma(A)$, $rank(M(z)) = \tau q + rank(\Psi(z))$.*

Proof: Let us introduce the matrix decomposition, valid for all $z \in \mathbb{C} \backslash \gamma(A)$:

$$\begin{aligned} \begin{bmatrix} I_{\tau q} - zA & -\mathbf{B} \\ C & 0_{n \times q} \end{bmatrix} \begin{bmatrix} I_{\tau q} & (I_{\tau q} - zA)^{-1}\mathbf{B} \\ 0_{q \times \tau q} & I_q \end{bmatrix} &= \begin{bmatrix} I_{\tau q} - zA & 0_{\tau q \times q} \\ C & C(I_{\tau q} - zA)^{-1}\mathbf{B} \end{bmatrix} \\ &:= \begin{bmatrix} I_{\tau q} - zA & 0_{\tau q \times q} \\ C & \Psi(z) \end{bmatrix} \end{aligned} \tag{O.18}$$

The second matrix on the left is nonsingular. Hence for all $z \in \mathbb{C} \backslash \gamma(A)$ we have:

$$rank\begin{bmatrix} I_{\tau q} - zA & -\mathbf{B} \\ C & 0_{n\times q} \end{bmatrix} = rank\begin{bmatrix} I_{\tau q} - zA & 0_{\tau q\times q} \\ C & \Psi(z) \end{bmatrix} \tag{O.19}$$
$$= rank(I_{\tau q} - zA) + rank(\Psi(z)) = \tau q + rank(\Psi(z))$$

The last line follows since $I_{\tau q} - zA$ is nonsingular for $z \in \mathbb{C}\backslash\gamma(A)$. Q.E.D.

To proceed further we consider two cases:

a) $\tau = max\{p, s+1\} = s+1 \geq p$. Then $M(z) = \begin{bmatrix} I_{(s+1)q} - zA & -\mathbf{B} \\ [\Lambda_0 \vdots \Lambda_1 \vdots \ldots \vdots \Lambda_s] & 0_{n\times q} \end{bmatrix}$. Since B is nonsingular and $\Lambda = [\Lambda_0 \vdots \Lambda_1 \vdots \ldots \vdots \Lambda_s]$ has full column rank, $M(z)$ has full column rank for all $z \in \mathbb{C}$, see e.g. Abadir and Magnus (2005), pp. 121-122. Hence by Lemma O.1, if Λ has full column rank then $\Psi(z)$ has full column rank for all $z \in \mathbb{C}\backslash\gamma(A)$ i.e. $normrank(\Psi(z)) = q$. Further, $I_{\tau q} - zA$ is nonsingular for all $z \in \mathbb{C}$ such that $|z| = 1$ (see footnote 8 in the main text). It means that all $z \in \mathbb{C}$ such that $|z| = 1$ are excluded from $\gamma(A)$. Hence $\Psi(z)$ has full column rank at $|z| = 1$.

b) $\tau = max\{p, s+1\} = p > s+1$. We have:

$$I_{pq} - zA = \begin{bmatrix} I_q - z\Phi_1 & -z\Phi_2 & -z\Phi_3 & \ldots & -z\Phi_{p-1} & -z\Phi_p \\ -zI_q & I_q & 0 & \cdots & 0 & 0 \\ 0 & -zI_q & I_q & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -zI_q & I_q \end{bmatrix}$$

We first assume that $s > 0$. Completing the full $M(z)$ but interchanging its rows (which does not change the rank), we can form the modified matrix:

$$\begin{bmatrix} -zI_q & I_q & 0 & \cdots & 0 & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & -zI_q & I_q & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -zI_q & I_q & 0 & 0 & \ldots & 0 & 0 \\ \Lambda_0 & \Lambda_1 & \Lambda_2 & \ldots & \Lambda_{s-1} & \Lambda_s & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & -zI_q & I_q & 0 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 & -zI_q & I_q & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & I_q & 0 \\ I_q - z\Phi_1 & -z\Phi_2 & -z\Phi_3 & \ldots & -z\Phi_s & -z\Phi_{s+1} & -z\Phi_{s+2} & -z\Phi_{s+3} & \ldots & -z\Phi_p & -B \end{bmatrix}$$

since B is nonsingular, this matrix has full column rank if and only if

$$\begin{bmatrix} -zI_q & I_q & 0 & \cdots & 0 & 0 \\ 0 & -zI_q & I_q & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & I_q & 0 \\ 0 & 0 & 0 & \dots & -zI_q & I_q \\ \Lambda_0 & \Lambda_1 & \Lambda_2 & \dots & \Lambda_{s-1} & \Lambda_s \end{bmatrix}$$

has full column rank, see e.g. Abadir and Magnus (2005), pp. 119-120. If $[\Lambda_0 \vdots \Lambda_1 \vdots \dots \vdots \Lambda_s]$ has full column rank, the latter matrix has full column rank for all $z \in \mathbb{C}$. Then, as explained in case a), $\Psi(z)$ has full column rank for all $z \in \mathbb{C} \backslash \gamma(A)$ i.e. $normrank(\Psi(z)) = q$ and $rank(\Psi(z)) = q$ at $|z| = 1$.

On the other hand, assume that $s = 0$. Then interchanging the rows in $M(z)$ such that the first block is $[\Lambda_0 \vdots 0]$, since B is nonsingular, one may easily show that such a modified $M(z)$ has full column rank for all $z \in \mathbb{C}$ if and only if Λ_0 has full column rank. The remaining conclusion follows exactly as in case a).

O.2.2 The proof of Proposition 2

The case $s = 0$ can be tackled exactly like in the last paragraph in Appendix O.2.1. Assume that $s > 0$. Using the results and technique from Appendix O.2.1 i.e. interchanging the rows in $M(z)$, it can be shown that provided that B is nonsingular, the system matrix (12) has full column rank if and only if:

$$G(z) := \begin{bmatrix} -zI_q & I_q & 0 & \cdots & 0 & 0 \\ 0 & -zI_q & I_q & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & I_q & 0 \\ 0 & 0 & 0 & \dots & -zI_q & I_q \\ \Lambda_0 & \Lambda_1 & \Lambda_2 & \dots & \Lambda_{s-1} & \Lambda_s \end{bmatrix} \tag{O.20}$$

has full column rank. Let us apply the elementary column operations to $G(z)$ (which do not change the rank). First, let us multiply the last block of q columns by z and add them to the previous block of q columns. Then multiply the resulting block of q columns by z and add it to the third-to-last block of q columns, and so on. Continuing in this way, we will obtain the modified $G(z)$, which has the following structure:

$$\begin{bmatrix} 0 & I \\ \Lambda_0 + z\Lambda_1 + z^2\Lambda_2 + \cdots + z^s\Lambda_s & \times \end{bmatrix} \tag{O.21}$$

By standard results, see e.g. Abadir and Magnus (2005), pp. 121-122, we conclude that the above matrix has full column rank if and only if $\Lambda_0 + z\Lambda_1 + z^2\Lambda_2 + \cdots + z^s\Lambda_s$ has full column rank. Hence the last one condition is also necessary and sufficient for the full column rank of the system matrix (12).

O.2.3 The proof of Proposition 3

If $\tau = max\{p, s+1\} = s+1 \geq p$, then $\mathcal{O} = \begin{bmatrix} [\Lambda_0 \vdots \Lambda_1 \vdots \ldots \vdots \Lambda_s] \\ \times \end{bmatrix}$. If $[\Lambda_0 \vdots \Lambda_1 \vdots \ldots \vdots \Lambda_s]$ has full column rank, $\mathcal{O}$ also has full column rank.

O.2.4 The proof of Proposition 4

In what follows we assume $\tau = max\{p, s+1\} = p > s+1$. We employ the Popov-Belevitch-Hautus (PBH) observability criterion, according to which the pair (A, C) is observable (i.e. $\mathcal{O}$ has full column rank) if and only if there exists no nonzero vector $v \in \mathbb{C}^{qp}$ satisfying

$$Av = zv, \qquad Cv = 0$$

for any eigenvalue z of A. Recall that the companion matrix A is nonsingular if and only if Φ_p is nonsingular (see, e.g., Abadir and Magnus (2005), pp. 121–122). Since Φ_p is nonsingular by assumption, zero cannot be an eigenvalue of A. It therefore suffices to consider its nonzero eigenvalues. Let $z \neq 0$ be an arbitrary eigenvalue of A, and let $v \in \mathbb{C}^{qp} \setminus \{0\}$ be a corresponding eigenvector satisfying

$$Av = zv$$

Partition v into q-dimensional blocks:

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{bmatrix}, \qquad v_i \in \mathbb{C}^q$$

The last $p-1$ block equations of the eigenvalue equation imply

$$v_i = zv_{i+1}, \qquad i = 1, \ldots, p-1$$

Setting $x = v_p$ and iterating these relations yields

$$v = \begin{bmatrix} z^{p-1}x \\ z^{p-2}x \\ \vdots \\ zx \\ x \end{bmatrix}, \qquad x \in \mathbb{C}^q$$

In particular, $x \neq 0$, since otherwise the block relations would imply $v = 0$, contradicting the definition of an eigenvector.

Using the structure of the matrix C, we have

$$\begin{aligned} Cv &= \Lambda_0 z^{p-1}x + \Lambda_1 z^{p-2}x + \cdots + \Lambda_s z^{p-1-s}x \\ &= z^{p-1-s}L(z)x, \end{aligned}$$

where

$$L(z) = \Lambda_0 z^s + \Lambda_1 z^{s-1} + \cdots + \Lambda_s$$

Observe that $L(z)$ admits the factorization

$$L(z) = \Lambda V(z),$$

where

$$V(z) = \begin{bmatrix} z^s I_q \\ z^{s-1} I_q \\ \vdots \\ z I_q \\ I_q \end{bmatrix} \in \mathbb{C}^{q(s+1)\times q}$$

and $\Lambda = [\Lambda_0 \Lambda_1 ... \Lambda_s]$. By assumption, Λ has full column rank. Therefore, there exists a left inverse $G \in \mathbb{R}^{q(s+1)\times n}$ such that

$$G\Lambda = I_{q(s+1)}$$

It follows that

$$GL(z) = G\Lambda V(z) = V(z)$$

Consequently,

$$\operatorname{rank} L(z) \geq \operatorname{rank}\Big(GL(z)\Big) = \operatorname{rank} V(z) = q$$

where the last equality follows from the fact that the last block of $V(z)$ is I_q. Since $L(z)$ has

exactly q columns, we obtain

$$\operatorname{rank} L(z) = q \qquad \text{for every } z \in \mathbb{C}$$

In particular, because $x \neq 0$, we get

$$L(z)x \neq 0$$

Since $z \neq 0$, it yields

$$Cv = z^{p-1-s}L(z)x \neq 0$$

Thus, no nonzero eigenvector of A belongs to the null space of C. By the PBH observability criterion, the pair (A, C) is observable.

O.3 Extensions of the identification result

O.3.1 State-space version of the identification result

This appendix states the state-space implication of the moving-average identification argument used in the proof of Theorem 1. Consider

$$x_t = Ax_{t-1} + R\epsilon_t, \tag{O.22}$$

$$y_t = Cx_t + u_t, \tag{O.23}$$

where $x_t \in \mathbb{R}^d$, $y_t \in \mathbb{R}^n$, and $\epsilon_t \in \mathbb{R}^m$. The Gaussian component has the normalized invertible representation $u_t = \Upsilon(L)e_t$, with $e_t \sim N(0, \Omega)$, Ω positive definite, and $\Upsilon_0 = I_n$. Define

$$\Psi(z) = C(I - zA)^{-1}R.$$

Theorem O.1 (State-space transfer-function identification)**.** *Let barred and unbarred parameterizations of the state-space model satisfy the following conditions: A and $\bar{A}$ are stable; the components of ϵ_t and $\bar{\epsilon}_t$ are mutually independent, i.i.d. over time, mean zero, variance one, have no nondegenerate independent Gaussian summand as in Assumption 1d, and satisfy the common-order condition in Assumption 2; each structural-shock process is independent of its entire Gaussian innovation process; $\Psi(z)$ and $\bar{\Psi}(z)$ have normal rank m, have rank m on the unit circle, and have full-column-rank impact matrices CR and $\bar{C}\bar{R}$; and the Gaussian components use normalized causal, stable, invertible spectral factors. If the two parameterizations generate the same distribution for $\{y_t\}$, then there exists a signed*

permutation matrix P_s such that

$$\bar{C}\bar{A}^h\bar{R} = CA^hRP_s, \qquad h = 0, 1, 2, \ldots, \tag{O.24}$$

and

$$\bar{\epsilon}_t \overset{d}{=} P_s'\epsilon_t.$$

The Gaussian spectral density is also identified; under the normalized invertible representation, $\bar{\Upsilon}(L) = \Upsilon(L)$ and $\bar{\Omega} = \Omega$.

If, in addition, (A, R, C) and $(\bar{A}, \bar{R}, \bar{C})$ are minimal realizations of their transfer functions, then there exists a nonsingular matrix H such that

$$\bar{A} = HAH^{-1}, \tag{O.25}$$

$$\bar{R} = HRP_s, \tag{O.26}$$

$$\bar{C} = CH^{-1}. \tag{O.27}$$

For a fixed pair of equivalent minimal realizations, H is unique.

Proof. Iterating the state equation gives the causal moving-average representation

$$y_t = \sum_{h=0}^{\infty} CA^hR\epsilon_{t-h} + \Upsilon(L)e_t.$$

Thus the state-space model is a special case of the representation used in Proposition 6, with moving-average coefficients $\Psi_h = CA^hR$. Proposition 6 therefore implies (O.24), the signed-permutation equivalence of the structural shocks, and identification of the normalized Gaussian spectral factor.

It remains to translate transfer-function identification into a statement about state-space matrices. Equality of the Markov parameters in (O.24) is equality of the transfer functions after applying the signed permutation to the structural shocks. If both realizations are minimal, standard realization theory implies that two minimal realizations of the same rational transfer function are related by a nonsingular linear transformation of the latent state; see, for example, Kailath (1980), Hannan and Deistler (2012), and Trentelman et al. (2001). Hence there exists a nonsingular H satisfying the displayed transformation equations. The same realization result gives uniqueness of H for the fixed pair of minimal coordinate systems. Q.E.D.

This result separates two roles of minimality. The distributional and rank conditions identify the structural transfer function, or equivalently the sequence of observable impulse responses CA^hR, without choosing state coordinates. Minimality is needed only to obtain

uniqueness of its state-space realization up to a nonsingular linear transformation of the latent state. This claim concerns the realization of the non-Gaussian transfer $\Psi(z)$. Without minimality, one can append stable state components that are not reachable from ϵ_t, not observable from y_t, or both, without changing the distribution of observables.

The left-invertibility condition in the theorem implies $m \leq n$. Thus the distribution-free result does not cover overcomplete ICA cases with more non-Gaussian shocks than observables. With independent Student-t shocks, higher-order cumulant tensors could in principle identify overcomplete mixtures under additional moment and tensor-uniqueness restrictions. Such a route is closer to parametric overcomplete ICA than to the distribution-free argument used here, and we therefore leave it as an extension.

O.3.2 Additional non-Gaussian components

The Gaussianity restriction on u_t applies to the component left outside the factor vector. An additional non-Gaussian component with a finite-dimensional dynamic representation can be absorbed into the factors if its innovations satisfy the same independence, Gaussian-summand exclusion, and moment restrictions as the structural shocks and the expanded system satisfies the required stationarity, rank, and observability conditions.

Suppose, for example, that the observation equation contains an additional finite-dimensional non-Gaussian component v_t:

$$y_t = \sum_{i=0}^{s} \Lambda_i f_{t-i} + \sum_{i=0}^{r} \Gamma_i v_{t-i} + u_t^G, \tag{O.28}$$

$$f_t = \sum_{j=1}^{p} \Phi_j f_{t-j} + B\epsilon_t, \tag{O.29}$$

$$v_t = \sum_{\ell=1}^{a} A_\ell v_{t-\ell} + C\xi_t, \tag{O.30}$$

where u_t^G is a stationary invertible Gaussian VARMA process. Let $v_t, \xi_t \in \mathbb{R}^m$, and assume that C is nonsingular. Define

$$\tilde{f}_t = \begin{bmatrix} f_t \\ v_t \end{bmatrix}, \qquad \tilde{\epsilon}_t = \begin{bmatrix} \epsilon_t \\ \xi_t \end{bmatrix}.$$

Let $\tilde{s} = \max\{s, r\}$ and $\tilde{p} = \max\{p, a\}$. Set missing lag matrices equal to zero: $\Lambda_i = 0$ for $i > s$, $\Gamma_i = 0$ for $i > r$, $\Phi_j = 0$ for $j > p$, and $A_j = 0$ for $j > a$. Then the system can be

written as

$$y_t = \sum_{i=0}^{\tilde{s}} \tilde{\Lambda}_i \tilde{f}_{t-i} + u_t^G, \tag{O.31}$$

$$\tilde{f}_t = \sum_{j=1}^{\tilde{p}} \tilde{\Phi}_j \tilde{f}_{t-j} + \tilde{B}\tilde{\epsilon}_t, \tag{O.32}$$

with

$$\tilde{\Lambda}_i = \begin{bmatrix} \Lambda_i & \Gamma_i \end{bmatrix}, \qquad \tilde{\Phi}_j = \begin{bmatrix} \Phi_j & 0 \\ 0 & A_j \end{bmatrix}, \qquad \tilde{B} = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}.$$

Thus the model is again an NG-SDFM, now with $\tilde{q} = q + m$ non-Gaussian innovations. Theorem 1 applies to this expanded system if the expanded shocks $\tilde{\epsilon}_t$ are mutually independent across components, i.i.d. over time, mean zero, variance one, independent as a process of the entire Gaussian innovation process, have no nondegenerate independent Gaussian summand in any marginal as required by Assumption 1d, and satisfy the common cumulant and moment condition in Assumption 2. The expanded system must also satisfy the corresponding stationarity, impact-rank, transfer-rank, and observability conditions: $\tilde{B}$ is nonsingular, the expanded impact matrix $\tilde{\Lambda}_0\tilde{B}$ has full column rank, the expanded transfer function $\tilde{\Psi}(z)$ has full column rank on the unit circle, and the expanded state-space realization is minimal.

The relevant size restriction is

$$\tilde{q} = q + m \leq n.$$

This restriction counts the number of independent non-Gaussian innovations, not the number of lags in the finite VAR for v_t. If $q + m > n$, the theorem developed here no longer applies.

O.4 Monte Carlo design and additional results

O.4.1 Calibration, priors, and scoring

All designs use static observation loadings, a factor VAR(4), horizons 0–20, and a Gaussian VAR(1) component. Each specification has 100 independent datasets. Simulation discards 200 initial observations; the estimation sample has 250 observations except in the noisy monetary proxy design, which has 500. Series are demeaned and divided by their own sample standard deviations. Structural shocks have unit unconditional variance.

In the two-factor panel designs the impact matrix and VAR coefficients are

$$B = \begin{pmatrix} 1 & .35 \\ -.20 & 1 \end{pmatrix}, \qquad \Phi_1 = \begin{pmatrix} .55 & .10 \\ -.05 & .45 \end{pmatrix}, \qquad \Phi_2 = \begin{pmatrix} .12 & .02 \\ .01 & .10 \end{pmatrix},$$

$$\Phi_3 = \begin{pmatrix} .04 & 0 \\ 0 & .03 \end{pmatrix}, \qquad \Phi_4 = \begin{pmatrix} .02 & -.01 \\ .01 & .02 \end{pmatrix}.$$

The first two loading rows form I_2. For $i = 3, \ldots, n$, row i is $(.65\cos a_i + .20,\ .55\sin a_i - .10)$, where $a_i = 2\pi(i-3)/(n-2)$. The Gaussian component follows $u_t = \rho u_{t-1} + e_t$, with $\mathrm{Var}(e_t) = \sigma^2[(1-r)I_n + r\mathbf{1}\mathbf{1}']$. The strong Gaussian noise specification has $n = 10$, $(\rho, \sigma, r) = (.8, 1, .6)$, and $\nu = 8$. Moderate Gaussian noise uses $(.4, .6, .3)$, $n = 10$, and $\nu = 5, 20$.

The noisy monetary proxy design has Gaussian noise only in the proxy, with unconditional variance .25 and AR coefficient .3. Its innovation variance is therefore $.25(1 - .3^2)$. Its macro impact matrix is

$$\begin{pmatrix} 1.6 & -.5 & -.4 & -.1 \\ 1.1 & -.5 & -.3 & .4 \\ 1.2 & -.7 & .8 & -.7 \end{pmatrix},$$

and its proxy impact row is $(\sqrt{c}, 0, \sqrt{1-c}, 0)$.

The macro VAR coefficients are

$$\Phi_{x,1} = \begin{pmatrix} .50 & .12 & -.08 \\ -.06 & .45 & .10 \\ .04 & -.05 & .40 \end{pmatrix}, \qquad \Phi_{x,2} = .10I_3, \quad \Phi_{x,3} = .04I_3, \quad \Phi_{x,4} = .02I_3.$$

Shock 3 is monetary and shock 1 is the contaminating information shock. The proxy does not enter the macro equations. Its own equation is (34), so the fourth row of the augmented lag-one coefficient matrix is $(.10, -.05, .08, .25)$ and its remaining lag rows are zero. The first three macro responses to the monetary shock are therefore identical across the four proxy specifications. The replication files record all simulation seeds.

The prior on factor-VAR coefficient (i, j) at lag k is independent, zero-centered Normal with standard deviation

$$\frac{.5}{k^2}\frac{\widehat{\sigma}_i}{\widehat{\sigma}_j}\begin{cases} 1, & i = j, \\ .5, & i \neq j, \end{cases}$$

where $\widehat{\sigma}_i$ are univariate autoregressive innovation scales for the first q observables used in the loading normalization. Free loadings have independent $N(0,1)$ priors; the impact matrix has a flat prior in B coordinates. For unrestricted noise dynamics, $\mathrm{vec}(\Pi) \mid \Omega \sim N(0, .1I_{n\ell} \otimes \Omega)$.

In these unrestricted-noise runs, the inverse-Wishart prior for Ω has $n+2$ degrees of freedom and diagonal scale formed from the squared univariate innovation scales. Degrees of freedom lie on the grid $[4.01, 100]$ with exponential-prior rate .1. We impose stable dynamics for both the factors and the Gaussian component.

For every design, the reference impact matrix in the likelihood-preserving normalization is the true impact matrix for the first q standardized observables. Thus its entries equal the corresponding raw impact responses divided by each observable's sample standard deviation. We keep this matrix fixed within each replication and apply the signed-permutation criterion in (32). Every chain uses 2,000 discarded iterations and 2,000 further iterations thinned by four, giving 500 retained draws.

Let $\widehat{\Psi}_{h,ij}^{(r)}$ be the posterior median response and $\Psi_{h,ij}^{(r)}$ its truth in replication r. The tables report the root mean squared error (RMSE), defined as

$$\text{RMSE} = \left\{ \frac{1}{R(H+1)nq} \sum_{r,h,i,j} (\widehat{\Psi}_{h,ij}^{(r)} - \Psi_{h,ij}^{(r)})^2 \right\}^{1/2}, \tag{O.33}$$

with $R = 100$ and $H = 20$. The sum includes all variables, shocks, and horizons 0–20, except for the proxy comparison described below. Responses are measured in each dataset's standardized-variable units per one-standard-deviation structural shock. An RMSE of 0.05 is therefore an error of five hundredths of an observable's sample standard deviation, pooled over these responses.

Figures 2 and 4 reverse sample standardization and plot all 100 posterior-median paths per specification in common raw units. Their dispersion describes sampling variation, not posterior uncertainty. Coverage uses the pointwise 5th and 95th percentiles within each chain. The tables' shock correlations use posterior means, with dates aligned to the estimator's lags. Figure 3 instead plots posterior medians in replication 1 of the moderate-noise specification with $\nu = 5$. Its correlations use all 249 available dates; the figure shows the first 100, with pointwise 5th–95th posterior percentile bands. Main-text Table 1 averages the posterior-mean shock correlations across all shocks within each dataset before summarizing the 100 datasets for each of the seven main-text specifications. The replication files retain these summaries and all 700 dataset-level correlations. Neither path is rescaled to match sample variances. Correlation alone does not establish correct shock amplitudes.

O.4.2 Response recovery in the core designs

Table O.1 shows that, with ten observables and moderate noise, IRF RMSE is 0.04 at $\nu = 5$, rising to 0.05 at $\nu = 20$. With strong Gaussian noise, RMSE is 0.07. Pointwise 90% interval

Table O.1: Recovery of true impulse responses in the three panel specifications

Calibration	IRF RMSE	90% coverage
Moderate Gaussian noise ($n = 10$)		
$\nu = 5$	0.04 (0.00)	0.84 (0.02)
$\nu = 20$	0.05 (0.00)	0.77 (0.02)
Strong Gaussian noise ($\nu = 8$)		
$n = 10$	0.07 (0.00)	0.84 (0.01)

Each row uses 100 replications. Parentheses contain Monte Carlo standard errors across independent datasets, using the delta method for RMSE. Standard errors displayed as 0.00 reflect rounding, not zero uncertainty. RMSE uses standardized-variable units for a one-standard-deviation shock. Coverage pools horizons 0–20, variables, and shocks within each dataset before averaging across datasets. It refers to pointwise credible intervals.

Table O.2: One-SD response RMSE under strong Gaussian noise

Calibration	NG-SDFM	Student-t two-step	Recursive PCA
Strong Gaussian noise ($\nu = 8$)			
$n = 10$	0.07	0.11	0.11

Both comparison estimators use PCA factors and an OLS VAR(4). Losses pool all variables, shocks, horizons 0–20, and 100 datasets in standardized-variable units. The comparison estimators supply no interval estimates.

coverage is 83.98%, 77.16%, and 84.10% for the moderate-noise specifications with $\nu = 5, 20$ and the strong-noise specification, respectively.

O.4.3 Comparison estimators

For the strong-noise design, the comparison estimators extract the leading principal components and fit a VAR(4) by OLS with an intercept. One identifies a rotation of the whitened residuals using independent Student-t densities; the other keeps the recursive Cholesky rotation in principal-component order. These are two-step point estimators, not joint Student-t likelihood estimates. Signs and columns are matched ex post to true observable impacts; this matching cannot correct a wrong continuous rotation. Table O.2 reports their response losses without interpreting the comparison as a common-prior or common-restriction experiment.

In the proxy design the competitor retains its first, proxy-ordered recursive shock; it does not search across columns to maximize agreement with the true monetary shock. Table O.3's loading-label frequency averages a draw-level classification; it is not the fraction of datasets with a correctly selected final label. The proxy design's reported coverage in the archived summary covers all variables and shocks, whereas its comparison loss uses only the three macro responses to the monetary shock.

Table O.3: Monetary-response recovery with a noisy, contaminated proxy

	One-SD IRF RMSE		Label
(ν, c)	NG-SDFM	VAR	frequency
$\nu = 5,\ c = 0.0$	0.03	0.03	0.97
$\nu = 5,\ c = 0.2$	0.02	0.08	0.95
$\nu = 5,\ c = 0.4$	0.03	0.11	0.70
$\nu = 8,\ c = 0.4$	0.04	0.11	0.57

Errors use the three macro responses to the monetary shock, horizons 0–20, and 100 replications. One-SD responses use standardized-variable units. NG-SDFM errors are scored using the true monetary column. Label frequency is the mean fraction of retained draws in which the largest squared proxy loading selects that column. VAR denotes the proxy-first recursive VAR.

Table O.4: Core response errors and coverage by horizon

	$h = 0$		$h = 1$–4		$h = 5$–20	
Calibration	RMSE	Coverage	RMSE	Coverage	RMSE	Coverage
Moderate Gaussian noise ($n = 10$)						
$\nu = 5$	0.10	0.88	0.05	0.90	0.02	0.82
$\nu = 20$	0.17	0.88	0.07	0.87	0.02	0.74
Strong Gaussian noise ($\nu = 8$)						
$n = 10$	0.24	0.64	0.11	0.76	0.03	0.87

RMSE uses standardized-variable units. Coverage refers to pointwise 90% intervals. Each entry pools the stated horizons, variables, shocks, and 100 replications. Replication-level standard errors for these entries are supplied in the numerical archive.

O.4.4 Horizon and numerical diagnostics

Table O.4 separates impact and short-run performance from the many smaller responses at later horizons. Monte Carlo standard errors use independent datasets as the sampling units. Let x_r denote the mean squared IRF error over the scored responses in dataset r. Pooled RMSE is $\sqrt{\bar{x}}$; its delta-method standard error is $s_x/(2\sqrt{\bar{x}}\sqrt{100})$, where s_x is the sample standard deviation of the 100 values of x_r. Coverage standard errors are the sample standard deviation of each dataset's coverage fraction divided by $\sqrt{100}$. These quantify finite-replication uncertainty, not MCMC error.

These checks use existing single chains and can reveal persistence or drift; they cannot certify convergence or exclude unexplored modes. No independent-chain convergence statistic is available for this batch. The comparison optimizers likewise use one deterministic start and do not retain convergence flags. All 700 replications completed, but completion is distinct from numerical accuracy.

Table O.5: Descriptive retained-chain checks for replication 1

Specification	Lag-1 corr.	Batch SE / SD	Half difference / SD
Moderate Gaussian noise ($n = 10$)			
$\nu = 5$	0.92	0.20	0.27
$\nu = 20$	0.82	0.17	0.37
Strong Gaussian noise			
$n = 10$	0.90	0.20	0.94
Noisy monetary proxy			
$\nu = 8,\ c = 0.4$	0.95	0.20	1.32

Each entry is the maximum over B_{11}, $(\Phi_1)_{11}$, Ω_{11}, and ν_1 in the predetermined first replication. Lag 1 refers to successive retained draws. Batch SE uses 20 contiguous batches of 25 retained draws and is divided by the scalar's posterior sample SD. The final column divides the absolute difference between the first and second 250-draw means by that same SD.

O.5 Additional details for the FRED-QD application

O.5.1 Data and transformations

The application in Section 7 uses 22 quarterly series from the December 2025 vintage of FRED-QD (McCracken and Ng, 2020). The panel covers output and expenditure, employment and hours, housing, productivity, prices, interest rates, credit spreads, equities, and money. We select these headline series directly; the panel is not selected using the database's factor flags or a test of non-Gaussianity. Table O.6 gives the series in estimation order.

We apply the transformations in the table, omit the first two raw observations, and retain dates before 2020. Complete-case selection gives a prepared panel of 242 observations from 1959Q3 through 2019Q4. The Gaussian VAR(1) filter uses the first observation as a lag, leaving 241 filtered observation equations. We demean each transformed series and divide by its sample standard deviation. No outlier replacement is applied. The introductory Gaussianity exercise uses a different panel: 125 factor-flagged series with the database's default transformations and outlier procedure.

Table O.6: FRED-QD series, transformations, and economic units

Series	Description	Original code	Transformation	Units
GDPC1	Real GDP	5	$\Delta \log$	pp growth
CPIAUCSL	Consumer price index	6	$\Delta \log$	pp inflation
FEDFUNDS	Federal funds rate	2	Level	pp
BAA10YM	Baa corporate yield less ten-year Treasury yield	1	Level	pp
UNRATE	Civilian unemployment rate	2	Level	pp
S&P 500	S&P 500 stock price index	5	$\Delta \log$	pp growth
PCECC96	Real personal consumption expenditures	5	$\Delta \log$	pp growth
GPDIC1	Real gross private domestic investment	5	$\Delta \log$	pp growth
GCEC1	Real government consumption and gross investment	5	$\Delta \log$	pp growth
INDPRO	Industrial production index	5	$\Delta \log$	pp growth
PAYEMS	Total nonfarm payroll employment	5	$\Delta \log$	pp growth
CE16OV	Civilian employment	5	$\Delta \log$	pp growth
AWHMAN	Average weekly hours, manufacturing	1	Level	Hours
HOUST	Housing starts	5	$\Delta \log$	pp growth
HOANBS	Hours worked, nonfarm business	5	$\Delta \log$	pp growth
OPHNFB	Output per hour, nonfarm business	5	$\Delta \log$	pp growth
COMPRNFB	Real compensation per hour, nonfarm business	5	$\Delta \log$	pp growth
PCECTPI	Personal consumption expenditure price index	6	$\Delta \log$	pp inflation
PPIACO	Producer price index, all commodities	6	$\Delta \log$	pp inflation
TB3MS	Three-month Treasury bill rate	2	Level	pp
GS10	Ten-year Treasury yield	2	Level	pp
M2REAL	Real M2 money stock	5	$\Delta \log$	pp growth

Original codes refer to the December 2025 FRED-QD vintage: 1 denotes levels, 2 first differences, 5 first log differences, and 6 second log differences. The transformation column gives the specification used here. Economic units undo standardization and express log differences in percent; pp denotes percentage points. Growth and inflation equal $100\Delta \log$ and are quarterly, without annualization. The first four rows have loading matrix I_4. All other contemporaneous loading coefficients are estimated.

O.5.2 Specification, priors, and normalization

We estimate four factors with four autoregressive lags, static loadings, and an unrestricted first-order Gaussian VAR:

$$\begin{aligned} y_t &= \Lambda_0 f_t + u_t, \\ f_t &= \sum_{h=1}^{4} \Phi_h f_{t-h} + B\epsilon_t, \qquad u_t = \Pi_1 u_{t-1} + e_t. \end{aligned} \tag{O.34}$$

Here y_t contains the standardized transformed observations. The components of ϵ_t are independent, unit-variance Student-t shocks, and $e_t \sim N(0, \Omega)$ is independent of the structural shocks. The matrix Ω is unrestricted. Both autoregressive systems are restricted to be stable.

The loading rows for GDP growth, CPI inflation, the federal funds rate, and the Baa–Treasury spread equal I_4. This normalization fixes the factor coordinates. It does not fix B or assign economic interpretations to its columns. The free loadings have independent $N(0, 1)$ priors. Factor VAR coefficients have the zero-centered Minnesota prior in (27): the lag-h own-coefficient standard deviation is $0.50/h^2$, and a cross-coefficient standard deviation is $0.25(\widehat{\sigma}_i/\widehat{\sigma}_j)/h^2$. We obtain the fixed scales $\widehat{\sigma}_i$ from univariate AR(4) regressions of the standardized named observables.

For the Gaussian VAR we use the joint prior in (28), with $M_0 = 0$, $V_0 = 0.10 I_{22}$, $\bar{\nu} = 24$, and $\bar{Q} = \operatorname{diag}(\widehat{\sigma}_1^2, \ldots, \widehat{\sigma}_{22}^2)$. These scales come from univariate AR(4) regressions of all standardized observables. The stability restriction applies to the joint prior for (Π_1, Ω). We use a flat prior on B and independent degrees-of-freedom priors with weights proportional to $\exp(-0.10\nu_j)$ on $4.01, 4.02, \ldots, 100$.

The reference impact matrix $\widehat{B}$ for the likelihood-preserving normalization is estimated from a preliminary sampler run. Using that run's posterior means, we form residuals $\widehat{f}_t - \sum_{h=1}^{4} \widehat{\Phi}_h \widehat{f}_{t-h}$ and estimate an unrestricted impact matrix and degrees of freedom by Student-t quasi-maximum likelihood. We align this estimate once toward I_4 by signed permutation, then hold the resulting $\widehat{B}$ fixed in (32). The identity loading block fixes factor coordinates; the separate alignment toward I_4 fixes only shock signs and ordering. The structural impact matrix B remains unrestricted. Online Appendix O.1.7 gives the matching algorithm. The aggregate Gaussian and non-Gaussian covariance matrices are invariant to the signs and ordering of the non-Gaussian shocks.

Both the preliminary and normalized stages use 20,000 burn-in iterations followed by 2,000,000 iterations, retaining every 200th iteration. All reported posterior summaries use the 10,000 retained draws from the normalized stage. The random seed is 73.

O.5.3 Unconditional variance and covariance

The decomposition uses the population covariance implied by each parameter draw. It concerns the stationary variables entering estimation, including GDP growth and inflation, and uses variation at all frequencies. It is not a decomposition of the sample covariance or of cumulated log levels.

Let $x_t = (f_t', f_{t-1}', f_{t-2}', f_{t-3}')'$, let A be the factor companion matrix, and put $C = [\Lambda_0\ 0\ 0\ 0]$. Write r_j for column j of $[B'\ 0_{4\times 12}]'$. The state covariance for structural shock j and the Gaussian covariance solve the discrete Lyapunov equations

$$P_j = AP_jA' + r_jr_j', \qquad j = 1, \ldots, 4, \tag{O.35}$$

$$P_G = \Pi_1 P_G \Pi_1' + \Omega. \tag{O.36}$$

Independence gives

$$\Sigma_{NG} = \sum_{j=1}^{4} CP_jC', \qquad \Sigma_G = P_G, \qquad \Sigma = \Sigma_{NG} + \Sigma_G. \tag{O.37}$$

Stability and unit structural-shock variances ensure finite covariances. We integrate out the Student-t scale variables, whose expectation is one, and solve these equations without horizon truncation for all 10,000 draws, including those with roots close to one.

For component $c \in \{G, NG\}$, define its variance share s_i^c and its contribution C_{ij}^c to the total unconditional correlation by

$$s_i^c = \frac{(\Sigma_c)_{ii}}{\Sigma_{ii}}, \qquad C_{ij}^c = \frac{(\Sigma_c)_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}. \tag{O.38}$$

The same total standard deviations appear in both contributions, so $C_{ij}^G + C_{ij}^{NG} = \rho_{ij}$ and $C_{ii}^c = s_i^c$. For positive component variances, let ρ_{ij}^c denote the correlation within component c. Then

$$C_{ij}^c = \rho_{ij}^c \sqrt{s_i^c s_j^c}, \qquad |C_{ij}^c| \le \sqrt{s_i^c s_j^c}. \tag{O.39}$$

Variance shares bound each contribution, but its sign and magnitude also depend on correlation within the component. Component correlations themselves need not add to the total correlation.

We summarize draw-level shares and correlation contributions by posterior means and 5th and 95th percentiles. Means preserve additivity; marginal interval endpoints need not. Restoring economic units scales component and total covariances alike, leaving shares and correlation contributions unchanged.

O.5.4 Results for all variables

Table O.7 extends Table 2 to all 22 observables. At the posterior mean, the non-Gaussian component accounts for most of the unconditional variation in the three interest rates. The Gaussian share is larger for unemployment and most growth rates, including output, expenditure, employment, and productivity, as well as the credit spread. Both components can generate comovement across observables, so the decomposition does not separate common variation from variable-specific noise. The classification concerns the joint Gaussian process; it does not identify separate Gaussian structural shocks or equate Gaussian variation with measurement error.

Table O.7: Gaussian and non-Gaussian shares of unconditional variance: all variables

Observable	Non-Gaussian (%)	Gaussian (%)
Real GDP growth	20.54 [8.11, 42.22]	79.46 [57.78, 91.89]
CPI inflation	60.87 [39.32, 84.14]	39.13 [15.86, 60.68]
Federal funds rate	78.51 [58.69, 94.65]	21.49 [5.35, 41.31]
Baa–Treasury spread	33.58 [14.73, 63.89]	66.42 [36.11, 85.27]
Unemployment	7.13 [1.33, 19.33]	92.87 [80.67, 98.67]
S&P 500 log return	17.80 [4.03, 42.87]	82.20 [57.13, 95.97]
Real consumption growth	35.93 [19.09, 61.70]	64.07 [38.30, 80.91]
Real investment growth	4.20 [0.62, 11.76]	95.80 [88.24, 99.38]
Real government spending growth	35.83 [13.67, 63.16]	64.17 [36.84, 86.33]
Industrial production growth	7.79 [1.65, 20.65]	92.21 [79.35, 98.35]
Payroll employment growth	12.72 [3.89, 31.08]	87.28 [68.92, 96.11]
Household employment growth	14.22 [4.75, 32.79]	85.78 [67.21, 95.25]
Manufacturing weekly hours	51.35 [28.44, 78.08]	48.65 [21.92, 71.56]
Housing starts growth	11.39 [1.72, 30.02]	88.61 [69.98, 98.28]
Nonfarm business hours growth	10.47 [2.03, 28.48]	89.53 [71.52, 97.97]
Labor productivity growth	33.60 [15.53, 60.91]	66.40 [39.09, 84.47]
Real compensation per hour growth	43.93 [24.95, 69.82]	56.07 [30.18, 75.05]
PCE inflation	64.18 [43.10, 86.10]	35.82 [13.90, 56.90]
Producer-price inflation	57.00 [36.04, 80.42]	43.00 [19.58, 63.96]
Three-month Treasury rate	80.32 [61.42, 95.27]	19.68 [4.73, 38.58]
Ten-year Treasury yield	90.31 [78.49, 98.24]	9.69 [1.76, 21.51]
Real M2 growth	55.33 [33.56, 79.99]	44.67 [20.01, 66.44]

Entries are posterior mean shares in percent, with 5th and 95th percentiles in brackets. All variables retain their estimation transformations in Table O.6; no responses are cumulated and no frequency band is selected. Shares are computed within each of the 10,000 retained parameter draws and then summarized. The non-Gaussian share sums the four structural-shock contributions before taking posterior quantiles. The Gaussian component includes its full cross-sectional covariance and serial dependence.

References

Karim M. Abadir and Jan R. Magnus. *Matrix Algebra*. Cambridge University Press, 2005.

Jushan Bai and Peng Wang. Identification and Bayesian Estimation of Dynamic Factor Models. *Journal of Business & Economic Statistics*, 33(2):221–240, 2015. doi: 10.1080/07350015.2014.941467. URL https://doi.org/10.1080/07350015.2014.941467.

C. K. Carter and R. Kohn. On Gibbs sampling for state space models. *Biometrika*, 81(3): 541–553, 09 1994.

Edward J. Hannan and Manfred Deistler. *The Statistical Theory of Linear Systems*. SIAM, 2012.

Thomas Kailath. *Linear Systems*. Prentice-Hall, Inc., 1980.

Andrzej Kocięcki, Christian Matthes, and Michele Piffer. A Scalable Framework for Statistical Identification in Structural VARs. Working paper, 2025.

Michael W. McCracken and Serena Ng. Fred-qd: A quarterly database for macroeconomic research. Working Paper 2020-005B, Federal Reserve Bank of St. Louis, March 2020. URL https://doi.org/10.20955/wp.2020.005. Revision date: March 2020.

Christian Ritter and Martin A. Tanner. Facilitating the Gibbs Sampler: The Gibbs Stopper and the Griddy-Gibbs Sampler. *Journal of the American Statistical Association*, 87(419): 861–868, 1992.

Harry L. Trentelman, Anton A. Stoorvogel, and Malo Hautus. *Control Theory for Linear Systems*. Springer London, 2001.