

Supplement to “Asymmetric conjugate priors for large Bayesian VARs”

(*Quantitative Economics*, Vol. 13, No. 3, July 2022, 1145–1169)

JOSHUA C. C. CHAN

Department of Economics, Purdue University

1. PROOFS OF PROPOSITION AND COROLLARIES

This section provides proofs of Proposition 1 and two corollaries in the main text. We first record in the following lemma the determinant of the Jacobian of transformation from the structural-form parameterization to the reduced-form parameterization.¹ This lemma was proved in Chan and Jeliazkov (2009), and we include it here for convenience. The proof uses the differential forms approach that is equivalent to calculating the Jacobian (see, e.g., Theorem 2.1.1 in Muirhead (1982)).

LEMMA 1. *Suppose \mathbf{W} is a $n \times n$ positive definite matrix and $\mathbf{W} = \mathbf{T}'\tilde{\mathbf{T}}\mathbf{T}$, where \mathbf{T} is a lower triangular matrix with ones on the main diagonal and $\tilde{\mathbf{T}}$ is a diagonal matrix with positive diagonal elements. Denote the lower diagonal elements of \mathbf{T} by t_{ij} , $1 \leq j < i \leq n$, and the diagonal elements of $\tilde{\mathbf{T}}$ by t_{ii} , $i = 1, \dots, n$. Let $(d\mathbf{W})$ denote the differential form $(d\mathbf{W}) \equiv \wedge_{i \geq j} dw_{ij}$ and similarly define $(d\mathbf{T}) \equiv \wedge_{i \geq j} dt_{ij}$. Then we have*

$$(d\mathbf{W}) = \prod_{i=1}^n t_{ii}^{i-1} (d\mathbf{T}).$$

In other words, the determinant of the Jacobian of the transformation from $\mathbf{T}'\tilde{\mathbf{T}}\mathbf{T}$ to \mathbf{W} is $\prod_{i=1}^n t_{ii}^{-i+1}$.

PROOF OF THE LEMMA. By the definition $\mathbf{W} = \mathbf{T}'\tilde{\mathbf{T}}\mathbf{T}$, we have

$$\begin{pmatrix} w_{11} & w_{21} & \dots & w_{n1} \\ w_{21} & w_{22} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{pmatrix} = \begin{pmatrix} 1 & t_{21} & \dots & t_{n1} \\ 0 & 1 & \dots & t_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} t_{11} & 0 & \dots & 0 \\ 0 & t_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & t_{nn} \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ t_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \dots & 1 \end{pmatrix}.$$

Joshua C. C. Chan: joshuacc.chan@gmail.com

¹For a different structural-form parameterization where the error covariance matrix is the identity matrix and the impact matrix is triangular with free diagonal elements, Zha (1999) derives the determinant of the Jacobian of transformation from the reduced-form parameterization.

Hence, we can express each w_{ij} in terms of $\{t_{ij}\}$:

$$w_{ii} = t_{ii} + \sum_{j=i+1}^n t_{ji}^2 t_{jj}, \quad i = 1, \dots, n, \quad (1)$$

$$w_{ij} = t_{ij} t_{ii} + \sum_{k=i+1}^n t_{ki} t_{kj} t_{kk}, \quad 1 \leq j < i \leq n. \quad (2)$$

Next, we take differentials of these two equations so that we can write the differential form $(d\mathbf{W})$ in terms of $(d\mathbf{T})$. Since we are going to take the exterior product of these differentials and the exterior products of repeated differentials are zero, there is no need to keep track of differentials in t_{ij} that have previously occurred. Therefore, we take differentials of (1) and (2) and ignore those that have previously occurred:

$$\begin{aligned} dw_{nn} &= dt_{nn} \\ dw_{n,n-1} &= dt_{nn} dt_{n,n-1} + \dots \\ &\vdots \\ dw_{n1} &= t_{nn} dt_{n1} + \dots \\ dw_{n-1,n-1} &= dt_{n-1,n-1} + \dots \\ &\vdots \\ dw_{11} &= dt_{11} + \dots \end{aligned}$$

Finally, taking exterior products gives

$$\bigwedge_{i \geq j} dw_{ij} = t_{nn}^{n-1} t_{n-1,n-1}^{n-2} \cdots t_{22} \bigwedge_{i \geq j} dt_{ij}$$

as claimed. \square

PROOF OF PROPOSITION 1. Assume the same notation as in Lemma 1. To prove Proposition 1, we consider the case where

$$t_{ii} \sim \mathcal{G}\left(\frac{\nu_0 + i - n}{2}, \frac{s_i^2}{2}\right), \quad i = 1, \dots, n, \quad (3)$$

$$(t_{ij} | t_{ii}) \sim \mathcal{N}\left(0, \frac{t_{ii}^{-1}}{s_j^2}\right), \quad 1 \leq j < i \leq n, i = 2, \dots, n. \quad (4)$$

More specifically, we will show that the density of $\mathbf{W} = \mathbf{T}'\widetilde{\mathbf{T}}\mathbf{T}$ is the same as that of the Wishart distribution $\mathcal{W}(\nu, \mathbf{S}^{-1})$, where $\mathbf{S} = \text{diag}(s_1^2, \dots, s_n^2)$. Then, if we let $t_{ii} = 1/\sigma_i^2$ and $A_{i,j} = t_{ij}$, we have $\widetilde{\Sigma}^{-1} = \mathbf{A}'\Sigma^{-1}\mathbf{A} \sim \mathcal{W}(\nu_0, \mathbf{S}^{-1})$.

To prove the proposition, we first compute the determinant of \mathbf{W} and the trace $\text{tr}(\mathbf{S}\mathbf{W})$. Since the determinant of \mathbf{T} is 1, we have

$$|\mathbf{W}| = |\tilde{\mathbf{T}}| = \prod_{i=1}^n t_{ii}.$$

Next, using (1), we have

$$\begin{aligned} \text{tr}(\mathbf{S}\mathbf{W}) &= \sum_{i=1}^n w_{ii} s_i^2 \\ &= \sum_{i=1}^n t_{ii} s_i^2 + \sum_{i=1}^n \sum_{j=i+1}^n t_{ji}^2 t_{jj} s_i^2 \\ &= \sum_{i=1}^n t_{ii} s_i^2 + \sum_{j=2}^n \sum_{i=1}^{j-1} t_{ji}^2 t_{jj} s_i^2 \\ &= \sum_{i=1}^n t_{ii} s_i^2 + \sum_{i=2}^n \sum_{j=1}^{i-1} t_{ij}^2 t_{ii} s_j^2, \end{aligned}$$

where we change the order of the double summations in the third equality and interchange the dummy indices i and j in the last equality.

Now, it follows from the distributional assumptions in (3) and (4) that the kernel of the joint density of \mathbf{T} and $\tilde{\mathbf{T}}$ is

$$\begin{aligned} &\prod_{i=1}^n t_{ii}^{\frac{\nu_0+i-n}{2}-1} e^{-\frac{s_i^2}{2} t_{ii}} \times \prod_{i=2}^n t_{ii}^{\frac{i-1}{2}} e^{-\frac{1}{2} \sum_{j=1}^{i-1} t_{ij}^2 t_{ii} s_j^2} \\ &= \left(\prod_{i=1}^n t_{ii}^{\frac{\nu_0-n-1}{2}+(i-1)} \right) e^{-\frac{1}{2} (\sum_{i=1}^n t_{ii} s_i^2 + \sum_{i=2}^n \sum_{j=1}^{i-1} t_{ij}^2 t_{ii} s_j^2)}. \end{aligned}$$

Next, we derive the kernel of the density of \mathbf{W} . By the lemma, the determinant of the Jacobian is $\prod_{i=1}^n t_{ii}^{-i+1}$. Substituting $\text{tr}(\mathbf{W})$ and $|\mathbf{W}|$ into the above expression and multiplying the determinant of the Jacobian, we obtain the kernel of the density of \mathbf{W} :

$$|\mathbf{W}|^{\frac{\nu_0-n-1}{2}} e^{-\frac{1}{2} \text{tr}(\mathbf{S}\mathbf{W})},$$

which is the kernel of the Wishart density $\mathcal{W}(\nu_0, \mathbf{S}^{-1})$. \square

PROOF OF COROLLARY 1. Here, we use the same notation as in Proposition 1. Assume $\mathbf{W} \sim \mathcal{W}(\nu, \mathbf{S}^{-1})$, and let $\mathbf{W} = \mathbf{T}'\tilde{\mathbf{T}}\mathbf{T}$, where \mathbf{T} and $\tilde{\mathbf{T}}$ are given in Lemma 1. If we can show that t_{ii} and $(t_{ij} \mid t_{ii})$ follow the same normal-gamma distributions given in (3) and (4), respectively, then we are done. Since the transformation between \mathbf{W} and $\mathbf{T}'\tilde{\mathbf{T}}\mathbf{T}$ is one-to-one, the proof essentially just “reverses” the equalities given in Proposition 1. More

specifically, in the proof of Proposition 1 we showed that $|\mathbf{W}| = \prod_{i=1}^n t_{ii}$ and

$$\text{tr}(\mathbf{S}\mathbf{W}) = \sum_{i=1}^n t_{ii}s_i^2 + \sum_{i=2}^n \sum_{j=1}^{i-1} t_{ij}^2 t_{ii} s_j^2.$$

Also, by Lemma 1, the determinant of the Jacobian of transformation is $\prod_{i=1}^n t_{ii}^{i-1}$. Hence, the kernel of the joint distribution of t_{ii} , $i = 1, \dots, n$, and t_{ij} , $1 \leq j < i \leq n$, $i = 2, \dots, n$ is given by

$$\begin{aligned} & |\mathbf{W}|^{\frac{\nu_0 - n - 1}{2}} e^{-\frac{1}{2} \text{tr}(\mathbf{S}\mathbf{W})} \times \prod_{i=1}^n t_{ii}^{i-1} \\ &= \left(\prod_{i=1}^n t_{ii}^{\frac{\nu_0 - n - 1}{2} + (i-1)} \right) e^{-\frac{1}{2} (\sum_{i=1}^n t_{ii} s_i^2 + \sum_{i=2}^n \sum_{j=1}^{i-1} t_{ij}^2 t_{ii} s_j^2)} \\ &= \prod_{i=1}^n t_{ii}^{\frac{\nu_0 + i - n}{2} - 1} e^{-\frac{s_i^2}{2} t_{ii}} \times \prod_{i=2}^n t_{ii}^{\frac{i-1}{2}} e^{-\frac{1}{2} \sum_{j=1}^{i-1} t_{ij}^2 t_{ii} s_j^2}. \end{aligned}$$

It follows that t_{ii} , $i = 1, \dots, n$, are independent gamma random variables given in (3). Moreover, conditional on t_{ii} , t_{ij} , $1 \leq j < i$, are independent normal variables given in (4). \square

PROOF OF COROLLARY 3. Suppose $\tilde{\Sigma} \sim \mathcal{IW}(\nu_0, \mathbf{R})$, where \mathbf{R} is a symmetric positive definite matrix. Factor $\mathbf{R}^{-1} = \mathbf{L}'\mathbf{S}^{-1}\mathbf{L}$, where \mathbf{L} is lower triangular with ones on the main diagonal and \mathbf{S} is diagonal. Since $\tilde{\Sigma}^{-1} \sim \mathcal{W}(\nu_0, \mathbf{R}^{-1})$, by the properties of the Wishart distribution, we have $(\mathbf{L}')^{-1} \tilde{\Sigma}^{-1} \mathbf{L}^{-1} \sim \mathcal{W}(\nu_0, \mathbf{S}^{-1})$. Now, applying Corollary 1, we obtain $(\mathbf{L}')^{-1} \tilde{\Sigma}^{-1} \mathbf{L}^{-1} = \mathbf{A}'\tilde{\Sigma}^{-1}\mathbf{A}$, where \mathbf{A} is lower triangular with ones on the main diagonal and $\tilde{\Sigma}$ is diagonal. The diagonal elements of $\tilde{\Sigma}$ and the lower triangular elements of \mathbf{A} follow the normal-inverse-gamma distributions:

$$\begin{aligned} \sigma_i^2 &\sim \mathcal{IG}\left(\frac{\nu_0 + i - n}{2}, \frac{s_i^2}{2}\right), \quad i = 1, \dots, n, \\ (A_{i,j} | \sigma_i^2) &\sim \mathcal{N}\left(0, \frac{\sigma_i^2}{s_j^2}\right), \quad 1 \leq j < i \leq n, i = 2, \dots, n. \end{aligned}$$

Letting $\mathbf{C} = \mathbf{A}\mathbf{L}$, we can write $\tilde{\Sigma}^{-1} = \mathbf{C}'\tilde{\Sigma}^{-1}\mathbf{C}$. Since both \mathbf{A} and \mathbf{L} are lower triangular with ones on the main diagonal, so is \mathbf{C} . It remains to show that \mathbf{c}_i , the free elements of the i th row of \mathbf{C} , follows the normal distribution in (9). Since $\mathbf{C}' = \mathbf{L}'\mathbf{A}'$, we can write \mathbf{c}_i in terms of \mathbf{A} and \mathbf{L} as

$$\mathbf{c}_i = \mathbf{l}_i + \mathbf{L}'_{1:i-1} \mathbf{a}_i,$$

where \mathbf{l}_i and \mathbf{a}_i are respectively the free elements of the i th row of \mathbf{L} and \mathbf{A} , and $\mathbf{L}'_{1:i-1}$ is the $(i-1) \times (i-1)$ matrix that consists of the first $(i-1)$ rows and columns of \mathbf{L} . Since \mathbf{c}_i is an affine transformation of the normal vector \mathbf{a}_i , conditional on σ_i^2 , \mathbf{c}_i is

normally distributed with mean vector \mathbf{l}_i and covariance matrix $\sigma_i^2 \mathbf{L}'_{1:i-1} \mathbf{S}_{1:i-1}^{-1} \mathbf{L}_{1:i-1}$, where $\mathbf{S}_{1:i-1}$ is the submatrix consisting of the first $(i-1)$ rows and columns of $\mathbf{S} = \text{diag}(s_1^2, \dots, s_n^2)$. \square

2. DERIVATION OF THE IMPLIED STRUCTURAL-FORM HYPERPARAMETERS

In this section, we derive the implied prior hyperparameters of the structural-form VAR coefficients from the hyperparameters of the reduced-form VAR coefficients. Let $(B_k)_{ij}$ and $(\tilde{B}_k)_{ij}$ denote the (i, j) -th element of the structural-form coefficient matrix \mathbf{B}_k and the reduced-form coefficient matrix $\tilde{\mathbf{B}}_k$, respectively, for $i, j = 1, \dots, n, k = 1, \dots, p$. Further, let $\mathbf{m}_{\tilde{\boldsymbol{\beta}}, i}$ and $\sigma_i^2 \mathbf{V}_{\tilde{\boldsymbol{\beta}}, i}$ denote the prior mean vector and covariance matrix of the reduced-form parameters $\tilde{\boldsymbol{\beta}}_i = (\tilde{b}_i, (\tilde{B}_1)_{i1}, \dots, (\tilde{B}_1)_{in}, \dots, (\tilde{B}_p)_{i1}, \dots, (\tilde{B}_p)_{in})'$. The hyperparameters $\mathbf{m}_{\tilde{\boldsymbol{\beta}}, i}$ and $\mathbf{V}_{\tilde{\boldsymbol{\beta}}, i}$ can be elicited in a standard way, for example, following the Minnesota prior in Doan, Litterman, and Sims (1984) and Litterman (1986). We further assume that $\mathbf{V}_{\tilde{\boldsymbol{\beta}}, i}$ is diagonal as is usually done in the literature. The goal is to derive the structural-form hyperparameters $\mathbf{m}_{\boldsymbol{\beta}, i}$ and $\mathbf{V}_{\boldsymbol{\beta}, i}$ given $\mathbf{m}_{\tilde{\boldsymbol{\beta}}, i}$ and $\mathbf{V}_{\tilde{\boldsymbol{\beta}}, i}$.

To that end, recall that the structural-form and reduced-form coefficients are related via

$$(B_k)_{ij} = (\tilde{B}_k)_{ij} + \sum_{l=1}^{i-1} A_{i,l} (\tilde{B}_k)_{lj}.$$

Using the law of iterated expectations and the assumption that the prior means of $\boldsymbol{\alpha}_i = (A_{i,1}, \dots, A_{i,i-1})'$ are zero (see Proposition 1), we obtain

$$\mathbb{E}(B_k)_{ij} = \mathbb{E}[\mathbb{E}((B_k)_{ij} \mid \tilde{\mathbf{B}}_k)] = \mathbb{E}(\tilde{B}_k)_{ij}.$$

Hence, we have $\mathbf{m}_{\boldsymbol{\beta}, i} = \mathbf{m}_{\tilde{\boldsymbol{\beta}}, i}$.

Next, we elicit the prior covariance matrix $\mathbf{V}_{\boldsymbol{\beta}, i}$, which is in general a full matrix. Here, for simplicity we ignore the correlations between the structural-form coefficients and set $\mathbf{V}_{\boldsymbol{\beta}, i}$ to be diagonal. One could work out all the covariances following a similar derivation presented below. Now, we derive the variance of a generic element $(B_k)_{ij}$. Let $f(j, k) = (k-1)n + j + 1$ be an integer-value function keeping track of the indices. Further, let $(V_{\tilde{\boldsymbol{\beta}}, i})_{f(j,k)}$ and $(m_{\tilde{\boldsymbol{\beta}}, i})_{f(j,k)}$ denote the $f(j, k)$ -th diagonal element of $\mathbf{V}_{\tilde{\boldsymbol{\beta}}, i}$ and the $f(j, k)$ -th element of $\mathbf{m}_{\tilde{\boldsymbol{\beta}}, i}$, respectively. Then we have

$$\begin{aligned} \text{Var}((B_k)_{ij}) &= \mathbb{E}[\text{Var}((B_k)_{ij} \mid \mathbf{A})] + \text{Var}[\mathbb{E}((B_k)_{ij} \mid \mathbf{A})] \\ &= \mathbb{E} \left[\text{Var}((\tilde{B}_k)_{ij}) + \sum_{l=1}^{i-1} A_{i,l}^2 \text{Var}((\tilde{B}_k)_{lj}) \right] + \text{Var} \left[\mathbb{E}(\tilde{B}_k)_{ij} + \sum_{l=1}^{i-1} A_{i,l} \mathbb{E}(\tilde{B}_k)_{lj} \right] \\ &= \text{Var}((\tilde{B}_k)_{ij}) + \sum_{l=1}^{i-1} \text{Var}(A_{i,l}) \text{Var}((\tilde{B}_k)_{lj}) + \sum_{l=1}^{i-1} \text{Var}(A_{i,l}) (\mathbb{E}(\tilde{B}_k)_{lj})^2 \\ &= \sigma_i^2 (V_{\tilde{\boldsymbol{\beta}}, i})_{f(j,k)} + \sum_{l=1}^{i-1} \frac{\sigma_i^2}{s_l^2} [\sigma_l^2 (V_{\tilde{\boldsymbol{\beta}}, l})_{f(j,k)} + (m_{\tilde{\boldsymbol{\beta}}, l})_{f(j,k)}^2]. \end{aligned}$$

Note that this variance involves the error variances of other equations, namely, $\sigma_1^2, \dots, \sigma_{i-1}^2$. To avoid this issue, we replace σ_i^2 by its prior mean s_i^2 . Using this simplification, we obtain

$$\text{Var}((B_k)_{ij}) \approx \sigma_i^2 \left[(V_{\tilde{\beta},i})_{f(j,k)} + \sum_{l=1}^{i-1} ((V_{\tilde{\beta},l})_{f(j,k)} + s_l^{-2} (m_{\tilde{\beta},l})_{f(j,k)}^2) \right].$$

Given this approximation, we set the diagonal element in $\mathbf{V}_{\beta,i}$ associated with $(B_k)_{ij}$ to be $(V_{\tilde{\beta},i})_{f(j,k)} + \sum_{l=1}^{i-1} ((V_{\tilde{\beta},l})_{f(j,k)} + s_l^{-2} (m_{\tilde{\beta},l})_{f(j,k)}^2)$.

3. COMPARISON WITH INDEPENDENT NORMAL AND INVERSE-WISHART PRIOR

This section reports impulse responses of the 6-variable VAR described in Section 4 in the main text under the independent normal and inverse-Wishart priors. More specifically, we consider a standard Minnesota prior on the VAR coefficients with hyperparameters $\tilde{\kappa}_1 = \tilde{\kappa}_2 = 1$ and an inverse-Wishart prior on the reduced-form error covariance matrix $\tilde{\Sigma} \sim \mathcal{IW}(\nu_0, \mathbf{S})$. Given the posterior draws of the structural-form parameters, we transform them to the corresponding reduced-form parameters. We then use the algorithm described in Rubio-Ramirez, Waggoner, and Zha (2010) to incorporate the sign restrictions to construct impulse responses.

Figure 1 reports the impulse responses to a one-standard-deviation financial shock. These results are almost identical to those obtained under the asymmetric conjugate prior with the same hyperparameters $\tilde{\kappa}_1 = \tilde{\kappa}_2 = 1$ reported in the main text.

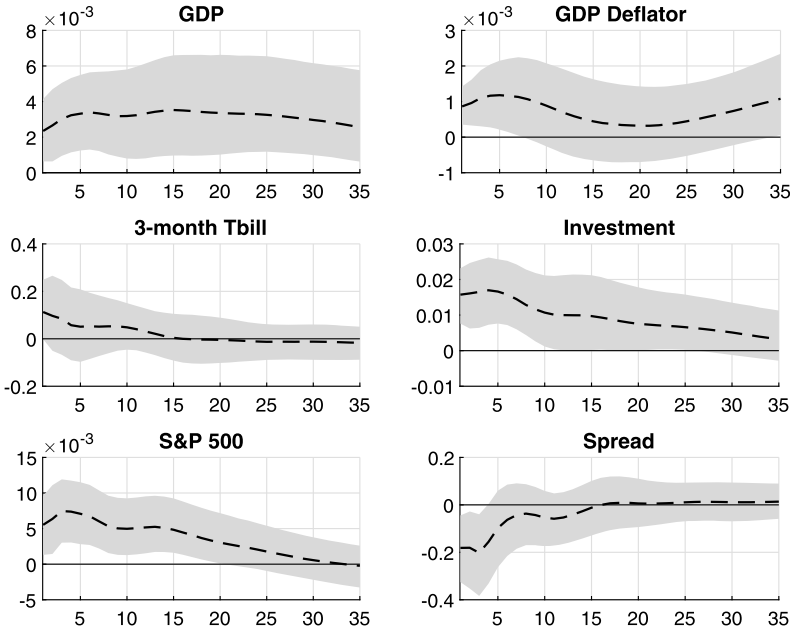


FIGURE 1. Impulse responses from a 6-variable VAR with the independent normal and inverse-Wishart priors to a one-standard-deviation financial shock. The shaded region represents the 16th and 84th percentiles.

REFERENCES

- Chan, Joshua C. C. and Ivan Jeliazkov (2009), “MCMC estimation of restricted covariance matrix.” *Journal of Computational and Graphical Statistics*, 18, 457–480. [1]
- Doan, Thomas, Robert Litterman, and Christopher Sims (1984), “Forecasting and conditional projection using realistic prior distributions.” *Econometric reviews*, 3 (1), 1–100. [5]
- Litterman, Robert B. (1986), “Forecasting with Bayesian vector autoregressions — five years of experience.” *Journal of Business and Economic Statistics*, 4, 25–38. [5]
- Muirhead, Robb J. (1982), *Aspects of Multivariate Statistical Theory*. Wiley. [1]
- Rubio-Ramirez, Juan F, Daniel F Waggoner, and Tao Zha (2010), “Structural vector autoregressions: Theory of identification and algorithms for inference.” *The Review of Economic Studies*, 77 (2), 665–696. [6]
- Zha, Tao (1999), “Block recursion and structural vector autoregressions.” *Journal of Econometrics*, 90 (2), 291–316. [1]

Co-editor Tao Zha handled this manuscript.

Manuscript received 3 July, 2019; final version accepted 30 October, 2021; available online 3 December, 2021.