Identifying Factor-Augmented Vector Autoregression Models by a Change in Shock Variances*

Yohei Yamamoto†

Hitotsubashi University

January 31, 2018
[**Preliminary: Prepared for IAAE submission**]

Abstract

In this study, we develop a new method of structural identification for the factor-augmented vector autoregression models by using a change in shock variances that occurred on a historical date. The proposed method can incorporate both observed and unobserved factors in the structural vector autoregression system and allows for the contemporaneous matrix to be fully unrestricted. We derive the asymptotic distribution of the impulse response estimator and consider a bootstrap inference method. A Monte Carlo simulation shows that the proposed method achieves a comparable accuracy with the existing methods even when the breaks are not very large. It outperforms the existing methods as the breaks become larger. Both the asymptotic and the bootstrap methods show a good coverage rate when a shock of observed factors is studied, although the latter will be more accurate when an unobserved factor shock is of interest.

JEL Classification Numbers: C14, C22

Keywords: factor-augmented vector autoregression, identification through heteroskedasticity, bootstrap, coverage rate, impulse response function

---

*The author acknowledges financial support for this work from the MEXT Grants-in-Aid for Scientific Research Nos. 16K03593 and 17H00985. I appreciate useful comments and suggestions from the participants of 2018 Kansai Econometrics Study Group Meeting.

†Hitotsubashi University, Department of Economics, 2-1 Naka, Kunitachi, Tokyo, Japan 186-8601 (yohei.yamamoto@econ.hit-u.ac.jp).

1 Introduction

One of the most challenging issues in the conventional structural VAR models[1] is that they can only accommodate a handful number of variables in the equation system and that the obtained empirical results may entail an omitted variable problem. A useful idea to address this concern is to incorporate a small number of unobserved variables that are extracted from a large macroeconomic panel data. This approach called the factor-augmented VAR (FAVAR) was initiated by Bernanke et al. (2005) and put forward in work summarized by Stock and Watson (2016), for example. It is well-known that such factors extracted by the principal component method are only identified up to a random rotation (Bai and Ng, 2013) and giving a particular interpretation to the estimated unobserved factors has been an empirical deadlock. Bai et al. (2016) address this concern by proposing three sets of exclusion restrictions to identify the unobserved factors, or more correctly speaking, they fix the random rotation matrix in the setup of FAVAR model of Bernanke et al. (2005).[2] However, the proposed exclusion restrictions assume no contemporaneous correlations between the sets of observed and unobserved factors that are relevant in many empirical settings. For instance, Brunnermeier et al. (2016) use the structural VAR to investigate the effect of as many as ten orthogonal shocks on macroeconomic variables and document that when the dimension is relatively large their contemporaneous ordering is not a trivial problem.

A popular approach of identifying simultaneous equations by using changes in variances in a demand-and-supply system was proposed by Rigobon (2003) and Rigobon and Sack (2003). This idea has been extended to the conventional small-scale structural VAR system by Lanne and Lütkepohl (2008), Lütkepohl (2012) and Lütkepohl and Netsunajev (2017).[3] The fundamental idea of this method is to increase the number of equations in the system by accounting for a change of variances[4] to achieve the order conditions. As a result, one can save agnostic zero restrictions that are imposed by particularly ordering the VAR variables or on the impulse responses. Apparently, one needs to pay the cost of finding an evidence of the variance change, however, one may utilize econometric methods of testing and estimating structural changes to check the validity of the identification assumptions in a relatively straightforward manner. See Casini and Perron (2018) for a survey of recent developments

[1] See Sims (1980) and vast amount work surveyed by Kilian and Lütkepohl (2018).

[2] Bai et al. (2016) consider the maximum likelihood estimation, however, the rotation problem remains the same as using the principal component estimation.

[3] See chapter 14 of Kilian and Lütkepohl (2018) for a survey.

[4] Importantly, the coefficient parameters are typically assumed constant.

in structural break analysis.[5]

This motivates us to develop a new method of structural identification of the impulse response functions in FAVAR models by using a change in shock variances that occurred on a historical date. Hence, the goal of this study is to identify the rotated version of the contemporaneous matrix and is fundamentally the same as the problem tackled by Bai et al. (2016), however, we allow for the contemporaneous matrix to be fully unrestricted. In particular, we propose a method and derive the asymptotic distributions of the identified structural coefficients and impulse response estimators under condition $\sqrt{T}/N \to 0$, where T and N are the time and cross-section dimensions, respectively, as $N, T \to \infty$. We also consider a bootstrap method to construct the confidence interval, by following seminal studies of Gonçalves and Perron (2014)[6] and Yamamoto (2017). A Monte Carlo experiment shows that, when T is small, the proposed identification method requires relatively a large variance change to achieve comparable performance to Bai et al.'s (2016) methods. However, as T gets larger, the proposed method outperforms theirs methods under relatively small variance changes. We also find that both the asymptotic and bootstrap confidence intervals yield a good finite sample properties when responses to a shock of an observed factor is studied. However, the latter provides more accurate coverage rate when a shock to the observed factor is of interest.

The rest of the paper is structured as follows. In section 2, we introduce the model and assumptions. In section 3, we propose a new identification method of FAVAR models and derive the asymptotic distributions of the estimators. We also design a bootstrap method to construct the confidence interval in the same section. In section 4, we examine finite sample properties of the proposed estimator by comparing with alternative methods proposed by Bai et al. (2016). Section 5 concludes the paper and the appendixes include technical derivations and details of computing estimators in study.

Throughout the paper, we use the following notation. The Euclidean norm of vector x is denoted by $\|x\|$. For matrices, the vector-induced norm is used. The symbols "$\xrightarrow{p}$" and "$\xrightarrow{d}$"

---

[5] Han and Inoue (2014) propose a test for changes in the factor loadings at a common time by investigating a change in variance of factors. Chen et al. (2013) consider the relationship between the estimated first factor and the rests of estimated factors. Breitung and Eickmeier (2011) and Yamamoto and Tanaka (2015) investigate a change in factor loadings of an individual response variable.

[6] Gonçalves and Perron (2014) allow serial correlations in the idiosyncratic errors and show that the bootstrap does not have to replicate them. This is extended to the case in which the errors in the factor-augmented model are serially correlated, introduced by Djogbenou et al. (2015), and to the case in which the idiosyncratic errors are cross-sectionally correlated, introduced by Gonçalves and Perron (2016). Gonçalves et al. (2017) consider the bootstrap prediction intervals in the factor-augmented model by explicitly accounting for the uncertainty of unobservable factors and coefficients.

represent convergence in probability under the probability measure P and convergence in distribution, respectively. Symbols $O_p(\cdot)$ and $o_p(\cdot)$ are the order of convergence under P. We use symbol "$A \approx B$" if, for two random matrices A and B, $\|A - B\| = o_p(1)$ as $N, T \to \infty$. Let L be the standard lag operator. The operator $vec(X)$ transforms an $m \times m$ matrix X into an $m^2 \times 1$ vector by stacking the columns, whereas $vech(X)$ stacks only the element on and above the main diagonal of a square matrix X.

2 Model and assumptions

2.1 Structural FAVAR models

We consider the structural factor-augmented vector autoregression (FAVAR) model of order p:

$$h_t^* = \sum_{j=1}^{p} A_j^* h_{t-j}^* + \varepsilon_t, \tag{1}$$

for $t = 1, \cdots T$, where h_t^* is an $r \times 1$ vector of factors and A_j^* is the VAR coefficient of the jth lag.[7] The factors can be a mixture of unobserved factors denoted by an $r_1 \times 1$ vector f_t^* and observed factors denoted by an $r_2 \times 1$ vector g_t^* where $h_t^* = [f_t^{*\prime}, g_t^{*\prime}]'$ and $r = r_1 + r_2$. The error term ε_t is considered structural shocks and is assumed to be serially uncorrelated with a diagonal covariance matrix $E(\varepsilon_t \varepsilon_t') = \Pi$. We also observe a large number of economic variables that are driven by the factors so that an $N \times 1$ vector x_t is generated by

$$\begin{aligned} x_t &= \Lambda^* f_t^* + \Gamma^* g_t^* + u_t^*, \\ &= C^* h_t^* + u_t^*, \end{aligned} \tag{2}$$

where Λ^* and Γ^* are $N \times r_1$ and $N \times r_2$ factor loading matrices of the unobserved and observed factors, respectively. Let $C^* = [\Lambda^*, \Gamma^*]$ denote the factor loadings. When we write c_i^*, it is an $r \times 1$ vector of factor loadings of the ith response variable for $i = 1, ..., N$. Similarly, λ_i^* and γ_i^* are $r_1 \times 1$ and $r_2 \times 1$ vectors of factor loadings attached to the unobserved and the observed factors, respectively. The error term u_t^* is an $N \times 1$ vector that is called idiosyncratic errors.

The structural VAR allows a contemporaneous correlations among the factors to pin down the causal effects of an economic shock. To this end, we define an estimable reduced-form models by plugging $h_t^* = B^{-1} h_t$ with a nonsingluar $r \times r$ matrix B in (1)

$$B^{-1} h_t = \sum_{j=1}^{p} A_j^* B^{-1} h_{t-j} + \varepsilon_t,$$

[7] This is a simplified model of Stock and Watson's (2016) formulation in that the number of shocks are the same as factors.

or equivalently

$$h_t = \sum_{j=1}^{p} BA_j^* B^{-1} h_{t-j} + B\varepsilon_t.$$

To simplify the notation, we define the reduced-form parameters and errors $A_j = BA_j^* B^{-1}$ and $e_t = B\varepsilon_t$ and write the reduced-form VAR model

$$h_t = \sum_{j=1}^{p} A_j h_{t-j} + e_t. \tag{3}$$

This is the same model as Bai et al. (2016). If we simplify the notation by letting $W = [\iota, H_{(-1)}, H_{(-2)}, \cdots, H_{(-p)}]$ be a $T \times (rp+1)$ matrix, where ι is a $T \times 1$ vector of ones and $H_{(-j)} = [h_{1-j}, \cdots, h_{T-j}]'$, and $A = [\nu, A_1, \cdots, A_p]'$ being an $(rp+1) \times r$ matrix. Then, (3) can equivalently be written as $H = WA + e$. The constant term in the model can be omitted for simplicity when the data is demeaned. The corresponding reduced-form factor model is

$$\begin{aligned} x_t &= \Lambda f_t + \Gamma g_t + u_t, \\ &= C h_t + u_t, \end{aligned} \tag{4}$$

where $C = C^* B^{-1}$ and $u_t^* = u_t$.

The goal of this paper is to develop a method to identify the structural FAVAR model through estimating matrix B without imposing any exclusion restrictions on it. Our approach follows the method of identification using changes in shock variances proposed by Rigobon (2003) and Rigobon and Sack (2003) in a demand-and-supply system and extended to the conventional small-scaled structural VAR models by Lanne and Lütkepohl (2008), Lütkepohl (2012) and Lütkepohl and Netsunajev (2017).[8] In particular, we define the structural impulse response functions of an observed variable $x_{i,t+h}$ to the structural shocks in ε_t at horizon h as

$$\Theta_{i,h} \equiv \frac{\partial x_{i,t+h}}{\partial \varepsilon_t'} = c_i^{*\prime} \Phi_h^* = c_i' \Phi_h B,$$

where Φ_h^* and Φ_h are the structural and the reduced-form vector moving-average coefficients, which are recursively defined by $\Phi_h^* = \sum_{j=1}^{h} \Phi_{h-j}^* A_j^*$ and $\Phi_h = \sum_{j=1}^{h} \Phi_{h-j} A_j$, respectively, with $\Phi_0^* = \Phi_0 = I_r$.

2.2 Assumptions

[8] The present method considers a discrete change in the unconditional variance of errors, while changes in conditional variances and various types of volatility changes are instudy. See chapter 14 of Kilian and Lütkepohl (2018). Brunnermeier et al. (2016) apply the method to measure the effects of as many as ten orthogonal shocks in financial markets on macroeconomic variables.

We consider the following assumptions on the models introduced in the previous subsection.

Assumption A (VARs)

(i) The structural VAR shocks ε_t are independent and identically distributed with mean zero and its covariance matrix has a change at a known date $T_b = [\kappa T]$ where $\kappa \in (0, 1)$. That is,

$$E(\varepsilon_t \varepsilon_t') = \begin{cases} \Pi_1 & \text{for } t = 1, ..., T_b \\ \Pi_2 & \text{for } t = T_b + 1, ..., T \end{cases},$$

where $\Pi_1 = diag[\pi_1^{(1)}, \pi_1^{(2)}, \cdots, \pi_1^{(r)}]$ and $\Pi_2 = diag[\pi_2^{(1)}, \pi_2^{(2)}, \cdots, \pi_2^{(r)}]$ are diagonal matrices with distinct positive diagonal elements.

(ii) The VAR variables in h_t^* are ordered according to the magnitudes of ratio $\pi_1^{(l)}/\pi_2^{(l)}$ for $l = 1, \cdots r$.

(iii) The $r \times r$ contemporaneous matrix B is nonsingular.

(iv) The roots of $\det(I_r - A_1 y - A_2 y^2 - \cdots - A_p y^p) = 0$ lie outside the unit circle.

Assumption B (idiosyncratic errors)

(i) $E(u_{it}) = 0$ and $E\left|u_{it}\right|^8 \leq M$, for all (i, t).

(ii) $N^{-1} \sum_{i=1}^{N} \sum_{k=1}^{N} |\tau_{ik}| \leq M$, where $\tau_{ik} = E(u_{it} u_{kt})$. $T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} |\gamma_{st}| \leq M$, where $\gamma_{st} = E(N^{-1} \sum_{i=1}^{N} u_{is} u_{it})$.

(iii) For every (s, t), $E\left|N^{-1/2} \sum_{i=1}^{N} [u_{is} u_{it} - E(u_{is} u_{it})]\right|^4 \leq M$.

(iv) For any i, $T^{-1/2} \sum_{t=1}^{T} h_t u_{it} \xrightarrow{d} N(0, \Omega_{Hui})$, where $\Omega_{Hui} \equiv \lim Var\left(T^{1/2} \sum_{t=1}^{T} h_t u_{it}\right)$.

Assumption C (factor model)

(i) The factor loadings of unobserved factors λ_i are deterministic and $\Lambda'\Lambda/N \to \Sigma_\Lambda$ a fixed positive definite matrix.

(ii) The eigenvalues of $r_1 \times r_1$ matrix $\Sigma_\Lambda \Sigma_F$, where $\Sigma_F = p\lim_{T\to\infty} F' M_G F/T$ and $M_G = I_T - G(G'G)^{-1}G'$, are distinct.

(iii) ε_t and u_{is} are mutually independent for all (i, s, t).

Assumption D (normalization through structural IRFs) The matrix of contemporaneous structural impulse responses of the first r variables has unit main diagonal elements.

The key assumptions of this paper are Assumptions A (i), (ii) and D. Assumptions A (i) and (ii) are indeed similar to that considered by Lütkepohl (2012) and Lütkepohl and

Netsunajev (2017). While they normalize the covariance matrix of one regime to be an identity matrix, our method does not do so but fixes the scale of eigenvectors by Assumption D. Assumption A (iii) assumes a known break date for simplicity, however, this could be relaxed by using the standard break date estimate such as Perron and Qu (2007). The standard break date fraction estimator would have a faster rate than $\sqrt{T}$ so that the estimation uncertainty pertaining to the break date fraction does not affect the asymptotic inference for the least square coefficient estimators. Assumptions A (iv) guarantees a stable VAR system and it also ensures that the probability limit of $F'M_GF/T$ becomes a fixed positive definite matrix which we use in Assumption C (ii). Assumption B is essentially the same as the standard literature of the factor model such as Bai (2003) and Bai and Ng (2006) and allows for idiosyncratic errors to have weak serial and cross-sectional dependence. Assumption C is the standard regularity condition and guarantees a statistical identification of the unobserved factors up to a random rotation as we discuss in the next section. Finally, by Assumption D, we can fix the scale of the unobserved economic shocks in an arbitrary way and this normalization depends on the researcher's discretion.

3 Estimation and inference

3.1 Estimation of the reduced-form models

We use the two-step principal component method to estimate the reduced-form models (3) and (4). We first obtain the regression residuals of x_{it} on the observed factors g_t. Then, the unobserved factors f_t are estimated as the r_1 principal components of the residuals. In other words, we obtain an estimate $\widetilde{F}$ as a $T \times r_1$ eigenvector matrix corresponding to the r_1 largest eigenvalues of $M_G XX' M_G/(TN)$ where $M_G \equiv I_T - G(G'G)^{-1}G'$ under normalization of $\widetilde{F}'\widetilde{F}/T = I_r$. We then construct a $T \times r$ matrix of factor estimate $\widetilde{H} \equiv [\widetilde{F}, G]$ and its lags up to pth order $\widetilde{W} \equiv [\iota, \widetilde{H}_{(-1)}, \cdots, \widetilde{H}_{(-p)}]$. The factor loadings are estimated using the least squares method $\widetilde{C} = X'\widetilde{H}(\widetilde{H}'\widetilde{H})^{-1}$ and the VAR coefficients are estimated by $\widetilde{A} = (\widetilde{W}'\widetilde{W})^{-1}\widetilde{W}'\widetilde{H}$. Note that if we let the random rotation matrix induced by the two-step principal component estimation be R, it is known that the least squares estimators $\widetilde{c}_i'$ and $\widetilde{A}_j$ are consistent estimators[9] for $c_i'R^{-1}$ and RA_jR^{-1} and the following asymptotic normality

[9]More precisely speaking, because R is a random matrix, $\left\|\widetilde{c}_i - R'^{-1}c_i\right\| \xrightarrow{p} 0$ and $\left\|\widetilde{A}_j - RA_jR^{-1}\right\| \xrightarrow{p} 0$.

results follow Bai and Ng (2006)[10]

$$\sqrt{T}(\widetilde{c}_i - R'^{-1}c_i) \xrightarrow{d} N(0, V_{c_i}),$$
$$\sqrt{T}[\widetilde{A} - (I_p \otimes R'^{-1})AR'] \xrightarrow{d} N(0, V_A),$$

where the asymptotic variances V_{c_i} and V_A are stated in Appendix A as Lemma A1. Hence, the reduced-form impulse response $c_i'\Phi_h$ is also estimated up to a rotation matrix R such that it is a consistent estimator for $c_i'\Phi_h R^{-1}$.

Remark 1 *Bai et al. (2016) consider the same problem without explicitly accounting for the structural identification, i.e. contemporaneous correlation matrix B based on the quasi maximum likelihood estimation. Hence, their setup is the same as our reduced-form model (3) and (4) and the goal is to identify R matrix. However, because we do not separately identify R and B, our problem is essentially the same as theirs. We will elaborate on their approaches in section 3.2 and Appendix B.*

3.2 Identification of the structural parameters

We here propose an identification method for the structural coefficients and impulse responses. The goal is to consistently estimate C^*, A^*, and $\Theta_{i,h}$ and the asymptotic distribution

results follow in this subsection.

Algorithm:

1. Estimate the reduced-form models (3) and (4) to obtain $\widetilde{A}$, $\widetilde{C}$, and $\widetilde{e}_t$. Let $\widetilde{C}_{1:r}$ be the first r rows of $\widetilde{C}$.

2. Construct the sample covariance matrices for the pre- and post- break VAR residuals, respectively, by

$$\frac{1}{T_b}\sum_{t=p+1}^{T_b} \widetilde{e}_t\widetilde{e}_t' = \widetilde{\Omega}_1,$$
$$\frac{1}{T-T_b}\sum_{t=T_b+1}^{T} \widetilde{e}_t\widetilde{e}_t' = \widetilde{\Omega}_2.$$

[10] In Proof of Theorem 1 in Appendix A, we show that

$$R = \begin{bmatrix} Q^{-1} & 0_{r_1 \times r_2} \\ (G'G)^{-1}G'F & I_{r_2} \end{bmatrix},$$

where $Q \equiv V^{-1}(\widetilde{F}'M_G F/T)(\Lambda'\Lambda/N)$ and V is a $r_1 \times r_1$ diagonal matrix with the main diagonal elements as the r_1 largest eigenvalues of $M_G XX'M_G/(TN)$.

3. Obtain an $r \times r$ eigenvector matrix of

$$\widetilde{S} = \widetilde{C}_{1:r}\widetilde{\Omega}_1\widetilde{\Omega}_2^{-1}\widetilde{C}_{1:r}^{-1}$$

in descending order of their associated eigenvalues. Let the kth eigenvector divided by its kth element be $\widetilde{\delta}_k$. Then, $\widetilde{\Delta} = [\,\widetilde{\delta}_1, \cdots, \widetilde{\delta}_r]$ is an $r \times r$ matrix.

4. Estimate the contemporaneous matrix B by

$$\widetilde{B} = \widetilde{C}_{1:r}^{-1}\widetilde{\Delta}.$$

Once we obtain $\widetilde{B}$, the structural parameters are conventionally constructed by combining it with the reduced-form estimates $\hat{C}^* \equiv \widetilde{C}\widetilde{B}$ and $\hat{A}_j^* \equiv \widetilde{B}^{-1}\widetilde{A}_j\widetilde{B}$ and the structural impulse response is constructed by $\hat{\Theta}_{ih} \equiv \widetilde{c}_i'\widetilde{\Phi}_h\widetilde{B}$. The following result is shown for them.

Theorem 1 *Under Assumptions A-D, the above algorithm yields:* $\hat{c}_i^* \xrightarrow{p} c_i^*$, $\hat{A}_j^* \xrightarrow{p} A_j^*$, *and* $\hat{\Theta}_{ih} \xrightarrow{p} \Theta_{ih}$ *for each* i *and* j, *uniformly in* h, *as* $N, T \to \infty$.

A formal proof is given in Appendix A, but an intuitive explanation of this identification mechanism is as follows. The goal is to obtain $\widetilde{B}$ that is asymptotically equivalent to the rotated version RB. To this end, in step 3, we obtain $\widetilde{S}$ such that

$$\begin{aligned}
\widetilde{S} &= \underbrace{\widetilde{C}_{1:r}}_{CR^{-1}}\underbrace{\widetilde{\Omega}_1}_{RB\Pi_1B'R'}\underbrace{\widetilde{\Omega}_2^{-1}}_{R'^{-1}B'^{-1}\Pi_2^{-1}B^{-1}R^{-1}}\underbrace{\widetilde{C}_{1:r}^{-1}}_{RC^{-1}}, \\
&\to_p (C_{1:r}B)\Pi_1\Pi_2^{-1}(B^{-1}C_{1:r}^{-1}), \\
&= C_{1:r}^*\Pi_1\Pi_2^{-1}C_{1:r}^{*-1}.
\end{aligned}$$

In the last line of the above equations, because $\Pi_1\Pi_2^{-1}$ is a diagonal matrix in descending order by Assumptions A (i) and (ii) and $C_{1:r}^*$ has unit diagonal elements by Assumption D, the eigenvector of $\widetilde{S}$ becomes a consistent estimate for $C_{1:r}^*$, which is $C_{1:r}B$. Then, in step 4,

$$\widetilde{B} = \widetilde{C}_{1:r}^{-1}\widetilde{\Delta} \approx (RC_{1:r}^{-1})(C_{1:r}B) = RB.$$

Furthermore, by plugging $\widetilde{B}$ in the structural parameters, we can show that R's are cancelled out in the structural parameter and impulse response estimates.

We next derive the asymptotic distributions. To this end, the following lemma is useful. Note that matrix $C_{1:r}$ is a submatrix of C but is used only for identification of RB, hence to avoid notational confusion, we denote $C_{1:r}$ by Ξ and $C_{1:r}^*$ by Ξ^* in the following part.

Lemma 1 *Under Assumptions A-D, the following holds:*

$$\sqrt{T} vec(\widetilde{B} - RB) \xrightarrow{d} N(0, V_B),$$

as N, $T \to \infty$ *under* $\sqrt{T}/N \to 0$,*where*

$$V_B \equiv (I_r \otimes \bar{R}\Xi^{-1})V_\Delta(I_r \otimes \Xi'^{-1}\bar{R}') + (\Delta'\Xi'^{-1}\bar{R}' \otimes \bar{R}\Xi^{-1})V_\Xi(\Delta'\Xi'^{-1}\bar{R}' \otimes \bar{R}\Xi^{-1}),$$

with $V_\Xi \equiv [(R\Sigma_H R')^{-1} \otimes \Sigma_u^{1:r}]$ *and* V_Δ *is given in Lemma A5.*

The asymptotic variance of $\widetilde{C}_{1:r}$ here is denoted by V_Ξ and is constructed by using the existing results of Bai (2003) and Bai and Ng (2006). The asymptotic variance of the eigenvector matrix $\widetilde{\Delta}$ denoted by V_Δ is detailed in Lemma A5 presented in Appendix A. We now derive the asymptotic distributions of the structural parameter and impulse response estimators in the following theorem.

Theorem 2 *Under Assumptions A-D, the above algorithm yields: (i) for the factor loadings:*

$$\sqrt{T}(\hat{c}_i^* - c_i^*) \xrightarrow{d} N(0, \Omega_{c_i}),$$

as $N, T \to \infty$ *and* $\sqrt{T}/N \to 0$ *for any* $i = 1, ..., N$, *where*

$$\Omega_{c_i} \equiv (\bar{R}'^{-1}c_i' \otimes I_r)K_{rr}V_B K_{rr}'(c_i\bar{R}^{-1} \otimes I_r) + B'\bar{R}'V_{c_i}\bar{R}B,$$

where V_{c_i} *is defined in Lemma A1 and* K_{rr} *is the* $r^2 \times r^2$ *commutation matrix. (ii) For the VAR coefficients:*

$$\sqrt{T} vec(\hat{A}_j^* - A_j^*) \xrightarrow{d} N(0, \Omega_{Aj}),$$

as $N, T \to \infty$ *and* $\sqrt{T}/N \to 0$ *for any* $j = 1, ..., p$, *where*

$$\Omega_{A_j} \equiv D_B V_B D_B' + (B'\bar{R}' \otimes B^{-1}\bar{R}^{-1})V_{A_j}(\bar{R}B \otimes \bar{R}'^{-1}B'^{-1}),$$

with

$$D_B \equiv [(I_r \otimes B^{-1}A_j\bar{R}^{-1}) - (B'A_j'B'^{-1} \otimes B^{-1}\bar{R}^{-1})]$$

(iii) For the impulse responses:

$$\sqrt{T}(\hat{\Theta}_{i,h} - \Theta_{i,h}) \xrightarrow{d} N(0, \Omega_{\Theta_{i,h}}),$$

as $N, T \to \infty$ *and* $\sqrt{T}/N \to 0$ *uniformly in* h *for any* $i = 1, ..., N$, *where*

$$\Omega_{\Theta_{i,h}} = \bar{R}\Phi_h B V_{c_i} B'\Phi_h'\bar{R}' + D_{A,h}V_A D_{A,h}' + (I_r \otimes c_i'\Phi_h\bar{R}^{-1})V_B(I_r \otimes \bar{R}'^{-1}\Phi_h' c_i),$$

with

$$D_{A,h} \equiv (B'\bar{R}' \otimes \bar{R}'^{-1}c_i)\left(\sum_{j=0}^{h-1} JA'^{h-1-j} \otimes \bar{R}\Phi_j\bar{R}^{-1}\right),$$

where $J = [I_r; 0_{r\times r(p-1)}]$ *and* A *is a companion form of* $[A_1, \cdots, A_p]$.

There are two remarks pertaining to this result. First, all of the above asymptotic variances can be consistently estimated by combining the reduced-form estimates, because all the entities in the expressions are the rotated version (some of them are cancelled out in the above expressions). For example $R'^{-1}c_i'$ is consistently estimated by $\widetilde{c}_i'$ and $B'R'$ is consistently estimated by $\widetilde{B}'$. Second, we only derive the asymptotic distribution under $\sqrt{T}/N \to 0$ in this paper. Gonçalves and Perron (2014) detailed that relaxing this condition in form of $\sqrt{T}/N \to c$ where $0 \leq c < \infty$ generates asymptotic bias of the estimators and the inference that explicitly deals with this asymptotic bias is of importance. However, we leave this issue as a future research agenda.

3.3 Bai et al.'s (2016) strategies

Bai et al. (2016) consider the identification strategies of FAVAR models that provide estimates for coefficients and impulse responses that are free from the random rotation. Their setup is the same as our reduced-form models (3) and (4), where

$$E(e_t e_t') = \Sigma_e = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}.$$

The following three sets of exclusion restrictions are in study.

- **IRa:** $\Omega_{11} = I_{r_1}$, $\Omega_{12} = \Omega_{21} = 0_{r_1 \times r_2}$, and $N^{-1}\Lambda'\Sigma_u^{-1}\Lambda$ is a diagonal matrix with its diagonal elements being distinct and arranged in descending order.
- **IRb:** $\Omega_{11} = I_{r_1}$, $\Omega_{12} = \Omega_{21} = 0_{r_1 \times r_2}$, and $\Lambda_{1:r} = [\lambda_1, \cdots, \lambda_{r_1}]$ is a lower triangular matrix.
- **IRc**: $\Omega_{12} = \Omega_{21} = 0_{r_1 \times r_2}$ and $\Lambda_{1:r} = [\lambda_1, \cdots, \lambda_{r_1}] = I_{r_1}$.

In the form of structural VAR models (1) and (2), the above restrictions are translated into those on matrix B through the relationship $\Sigma_e = BB'$. In other words, we can restrict matrix B so that $\Sigma_e = BB'$ satisfies the above assumptions. A common feature of the three methods is to assume that $\Omega_{12} = \Omega_{21} = 0_{r_1 \times r_2}$. Unlike our method, this is achieved by assuming that the same off-diagonal parts of B matrix are zeros, i.e., the unobserved and the observed factors are not contemporaneously correlated. In this sense, our method will complement their methods, because it can benefit by allowing for matrix B to be fully unrestricted.

Note that Bai et al. (2016) consider the quasi maximum likelihood method instead of the principal component method to estimate the unobserved factors. However, as for the model identification, their strategies go through with the principal component factor estimate and the steps after the factor extraction can be exactly followed. Hence, we later make a comparison between our method and theirs by accommodating their identification methods with the principal component estimate.

3.4 Bootstrap confidence intervals

Seminal studies of Gonçalves and Perron (2014) and Yamamoto (2017) show that the asymptotic inference of the factor augmented models may under-evaluate the sampling errors of the coefficient estimators, as it stands on the ground of asymptotic results that are free from factor estimation errors in the scenario of $\sqrt{T}/N \to 0$. They develop a bootstrap inference as an alternative method which is consistent even under $\sqrt{T}/N \to c$ $(0 \leq c < \infty)$. Here, we also consider the bootstrap method in the same line. To this end, we additionally assume independent[11] but allow for heteroskedasticity in the idiosyncratic errors u_{it}. The procedure is described as follows.

Bootstrap algorithm:

1. Estimate the reduced-form models (3) and (4) and obtain the parameter estimate $\widetilde{C}$, $\widetilde{A}$, $\widetilde{B}$ and the residuals $\widetilde{e}_t$ and $\widetilde{u}_{it}$. Construct the structural impulse response estimator $\hat{\Theta}_{i,h}$ by using the algorithm proposed in section 3.2.

2. Make sure that the pre- and post- break VAR residuals $\{\widetilde{e}_t\}_{t=1}^{T_b}$ and $\{\widetilde{e}_t\}_{t=T_b+1}^{T}$are demeaned, respectively, in time direction. Resample with replacements the pre-break residuals $\{\widetilde{e}_t\}_{t=1}^{T_b}$ as $r \times 1$ vectors in an i.i.d. fashion and label them $\left\{e_t^b\right\}_{t=1}^{T_b}$. Do the same for the post-break residuals and label them $\left\{e_t^b\right\}_{t=T_b+1}^{T}$. Generate the bootstrapped factors via $h_t^b = \sum_{j=1}^{p} \widetilde{A}_j h_{t-j}^b + e_t^b$, for $t = p+1, ..., T$.[12]

3. Demean the idiosyncratic residuals $\widetilde{u}_{it}$ in both time and cross-sectional directions. For each $i = 1, ..., N$, if $\{u_{it}\}_{t=1}^{T}$ are considered homoskedastic, we propose the i.i.d.

[11] Gonçalves and Perron (2014) show that the serial correlations in the idiosyncratic errors are irrelevant when we focus on the coefficients of the factor-autoregression model in which the dependent variable is observed. However, this is not the case when a part of the dependent variables are the estimated factors as in the FAVAR model.

[12] The bias-correction method discussed by Kilian (1998) can be applied. The bias is estimated by taking the average of $\hat{A}_j^* - R^* \hat{A}_j R^{*-1}$ in another bootstrap loop, where R^* is a bootstrap analogue of R and can be constructed by the original estimate and the bootstrap samples.

resampling of $\{\widetilde{u}_{it}\}_{t=1}^{T}$ to obtain $\{u_{it}^{b}\}_{t=1}^{T}$. If $\{u_{it}\}_{t=1}^{T}$ are concerned heteroskedastic, use the wild bootstrap $u_{it}^{b} = \widetilde{u}_{it}\eta_{it}$, where $\eta_{it} \sim i.i.d.(0,1)$ is an external random variable to obtain $\{u_{it}^{b}\}_{t=1}^{T}$. Generate the bootstrapped observations $x_{it}^{b} = \widetilde{c}_i h_t^b + u_{it}^b$ for $t = 1, ..., T$ and $i = 1, ..., N$.

4. Using the bootstrap observations x_{it}^{b}, obtain the parameter estimate $\widetilde{C}^b$, $\widetilde{A}^b$, and $\widetilde{B}^b$, by using the same estimation and identification methods proposed in section 3.2. This yields the bootstrap estimate of the structural impulse response $\hat{\Theta}_{i,h}^{b}$.

5. Repeat steps 2 to 4 n_b times and store the recentered statistic $s_{i,h} \equiv \hat{\Theta}_{i,h}^{b}\zeta - \hat{\Theta}_{i,h}\zeta$, where ζ is an $r \times 1$ vector of shock of interest.

6. Sort the statistics and pick the $100\frac{\alpha}{2}^{th}$ and $100(1-\frac{\alpha}{2})^{th}$ percentiles $[s_{i,h}^{(\alpha/2)}, s_{i,h}^{(1-\alpha/2)}]$. The resulting $100(1-\alpha)\%$ confidence interval for $\Theta_{i,h}\zeta$ is $[\hat{\Theta}_{i,h}\zeta - s_{i,h}^{(1-\alpha/2)}, \hat{\Theta}_{i,h}\zeta + s_{i,h}^{(\alpha/2)}]$.

The bootstrap consistency of the method under $\sqrt{T}/N \to c$ $(0 \leq c < \infty)$ would be similarly shown to that of Yamamoto (2017) by explicitly deriving the asymptotic bias of the original and bootstrap estimates for the impulse responses, however, we will leave this work for a future agenda.

4 Monte Carlo simulation

In this section, we investigate the finite sample properties of the impulse response estimator by using Monte Carlo simulations. We first compare the estimation accuracy of our identification method with a variant of the three methods based on IR_a, IR_b and IR_c assumptions proposed by Bai et al. (2016) in which the unobserved factors are estimated by the principal component method. We also investigate the finite sample coverage property of the asymptotic and the bootstrap confidence intervals when the proposed identification method is used. Throughout this section, we use the following reduced-form model as the data generating process (DGP):

$$
\begin{aligned}
h_t &= Ah_{t-1} + e_t, \\
x_{it} &= c_i' h_t + u_{it},
\end{aligned}
$$

with $h_t = [f_t', g_t']'$ where f_t is unobserved and g_t is observed factors, respectively. The number of unobserved factors is $r_1 = 2$ and that of observed factors is $r_2 = 1$, so that $r = 3$. The order of VAR model is 1 with the VAR coefficient matrix

$$A = \begin{bmatrix} 0.6 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.2 \end{bmatrix},$$

while the VAR errors are generated by $e_t = B\varepsilon_t$ with the structural shock $\varepsilon_t \sim i.i.d.N(0, \Pi_t)$. The shock covariance matrix has a change at the middle of the sample so that

$$\Pi_t = \begin{cases} diag(1+d, 1+d/2, 1), & \text{for } t = 1, ..., [T/2] \\ I_r & \text{for } t = [T/2]+1, ..., T \end{cases},$$

where d represents the magnitude of variance change and we specifically consider $d = 2, 5$ and 10 in this experiment. We arbitrarily set the unrestricted version of matrix B

$$B = \begin{bmatrix} 1.0 & 0.4 & 0.8 \\ 0.2 & 1.0 & -0.3 \\ -0.6 & 0.4 & 1.0 \end{bmatrix}, \tag{5}$$

unless otherwise stated. For the factor model, the factor loadings c_i are generated as an $r \times 1$ vector with each element independently drawn from the standard normal distribution in the outset of the Monte Carlo simulation and they are fixed over the replications. Furthermore, the idiosyncratic errors are generated from $u_{it} \sim i.i.d.N(0, 1)$. We consider the impulse responses of the Nth variable $\Theta_{N,h}$, but this choice does not lose generality, because x_{it} are homogeneously generated for every i.[13] The number of Monte Carlo replications is 3,000.

To get an intuition on how large the breaks of such values of d are, we consider the structural break test for the estimated covariance matrix at a known date.

$$W = Tvech(\widetilde{\Omega}_1 - \widetilde{\Omega}_2)\widetilde{\Sigma}^{-1}vech(\widetilde{\Omega}_1 - \widetilde{\Omega}_2)'(\frac{T_b}{T})(1 - \frac{T_b}{T}),$$

where $\widetilde{\Sigma}$ is a consistent estimate for the long-run variance

$$\Sigma \equiv \lim Var\left[vech\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\widetilde{e}_t\widetilde{e}_t' - \widetilde{\Omega}\right)\right],$$

and $\widetilde{\Omega}_1$, $\widetilde{\Omega}_2$ and $\widetilde{\Omega}$ are the sample covariance matrices of the pre-break, the post-break and the full-sample VAR residuals, respectively. The long-run variance is estimated using the

[13] The first r variables are used for identification.

Newey-West method. The test is computed in each replication and rejection frequencies at the nominal 5% level are reported in Table 1. It show that when $T = 50$, the change of $d = 2$ is only detected 16.5% and that of $d = 10$ does not reach one. When $T = 150$, the change of $d = 5$ and 10 are always detected but that of $d = 2$ may not always be detected.

4.1 Mean squared error of the impulse response estimator

We first examine the mean of squared error (MSE) of the impulse response estimator to one unit structural shock of the observed factor. In particular, we compare our method (in the column of "IR_v") with Bai et al.'s (2016) methods (in the columns of "IR_a", "IR_b" and "IR_c"). In doing so, matrix B of IR_a, IR_b and IR_c is restricted as follow. Let $\Omega = BB'$ where B is specified by (5) and, for IR_a and IR_b, we impose Ω_{11} part to be the identity matrix and, for IR_a, IR_b, and IR_c, we impose Ω_{12} and Ω_{21} parts to be zero matrices. This results in

$$\Omega = \begin{cases} \begin{bmatrix} I_{r_1} & 0 \\ 0 & \Omega_{22} \end{bmatrix} & \text{for } IR_a \text{ and } IR_b \\ \begin{bmatrix} \Omega_{11} & 0 \\ 0 & \Omega_{22} \end{bmatrix} & \text{for } IR_c \end{cases}, \tag{6}$$

where Ω_{22} is the lower right $r_2 \times r_2$ block of $\Omega = BB'$. After obtaining such Ω, we construct B matrix by imposing a recursive assumption such that $B = chol(\Omega)$ for IR_a, IR_b and IR_c.

If the DGPs are different when several identification methods are studied, one hardly tells whether the difference of performance comes from the DGPs or from the identification methods. To address this concern, we consider the following four scenarios. In scenario A, the DGPs conform with the identification assumptions that each estimation method is based on, hence, four methods are applied to different DGPs. In scenario B, the DGP uses the unrestricted B matrix (5), hence, the method based on IR_v is regarded as a correct method but none of IR_a, IR_b and IR_c are so because they suffer from misspecifications. In scenario C, the DGP uses B matrix of IR_a and IR_b and, in scenario D, the DGP uses B matrix of IR_c. As you can see in (6), scenario C is the most restrictive case and, in which case, all the four methods satisfy their assumptions. Scenario D is somewhat less restrictive than scenario C because Ω_{11} part is unrestricted, hence, the methods based on IR_a and IR_b assumptions involve a misspecification while the methods based on IR_c and IR_v are correctly specified.

In contrast, the identification assumptions on the factor model are correctly imposed in every case as follows. When we use the method based on IR_a, the first $[0.7N]$ response variables have zero second factor loadings and the last $[0.3N]$ response variables have zero first factor loadings, so that $\Lambda'\Sigma_u\Lambda/N$ becomes a diagonal matrix. When we use the method based on IR_b, we put three zeros in the factor loading matrix so that the first r_1 response variables $\Lambda_{1:r_1}$ becomes a lower triangular matrix. When we use the method based on IR_c, $\Lambda_{1:r_1}$ is the identity matrix. When we use the method based on IR_v, we follow Assumption D and sets the main diagonals of $C_{1:r}$ so that the diagonal elements of $C_{1:r}B$ are unity.[14] Because we use the impulse response of the Nth variable, the objectives are the same regardless of the assumptions except for IR_a under which the second factor loading is restricted to be zero. However, that has very little effects on the MSEs.

Table 2 reports the MSE of the impulse response estimator of the time horizon $h = 1$ in the left panels and those of $h = 5$ are in the right panels, respectively, for scenarios A to D.[15] In each table, we consider the sample size pairs $(T, N) = (50, 50), (50, 150), (150, 50)$ and $(150, 150)$. Throughout the four scenarios, the MSE of IR_v decreases as d becomes large because the asymptotic variance of eigenvectors becomes smaller as the corresponding eigenvalues get more distinct, which is shown in Lemma A5 of Appendix A. In contrast, the MSEs of IR_a, IR_b and IR_c increase as d gets large because the variance of the VAR errors becomes larger.

We now compare the MSE of the four methods in scenario A. With $h = 1$, when the change is small ($d = 2$), the MSE of the method based on IR_v is larger than the other methods in every sample size pair, however, it becomes comparable to the others when $d = 5$. Indeed, when $d = 5$, IR_v gives the smallest MSEs when we have a large T ($T = 150$). A larger N also reduces the MSE, however, its effect is more blurred than that of T.

We next examine the results under different scenarios. In scenario B, because matrix B is unrestricted, we expect that the MSEs of the methods based on IR_a, IR_b and IR_c get worse due to the misspecification, although the MSE of the method based on IR_v does not. Indeed, in scenario B, the method based on IR_v gives the smallest MSE even when $d = 2$ for the $T = 150$ case. It gives the smallest MSE when $d = 5$ for either $T = 50$ and $T = 150$ cases. It is also remarked that the MSE of the methods based on IR_a, IR_b and IR_c are very different in scenarios A and B, however, the MSE of the method based on IR_v is quite robust and does not change very much. In scenario C, matrix B has the most restrictions thus the relative merit of the method based on IR_v would not be as high as in scenario B.

[14] We first compute $C_{1:r}B$ and set its diagonal elements one and get back $C_{1:r}$ by right-multiplying B^{-1}.

[15] When the sample size and the break size are not large, the MSE of IR_v may be driven by a very few large errors and the results may be unstable. Hence, we report the MSEs after truncating the 1% largest squared errors. However, this does not qualitative results of the Monte Carlo simulation.

However, we find that the method based on IR_v gives the smallest MSE when $d = 5$ for the $T = 150$ case and this performance is similar to scenario A. Finally, the results under scenario D are in the same line with those under scenario C, hence, the relative benefit of IR_v is not weakened even if B is slightly more restricted. These results qualitatively remain the same when a longer horizon ($h = 5$) is considered.

In summary, the Monte Carlo results confirm that the impulse responses are more accurately estimated as the magnitude of the variance changes gets larger and as the sample size grows. More importantly, it suggests that the proposed method is robust to specification of B compared to the existing methods of Bai et al. (2016). When B is correctly restricted, the MSE of the proposed method is comparable with them under the change of variance that is not necessarily very large, that is to say, with which the standard structural change test shows 80-90% rejection rate in this Monte Carlo simulation.

4.2 Coverage properties of the confidence intervals

We next investigate the coverage property of the confidence intervals of the impulse responses. Table 2 presents the coverage rate and the median length of the 95% asymptotic confidence interval. We consider the horizon h up to 5 as it is sufficient considering the VAR model of order 1. The same set of sample size pairs $(T, N) = (50, 50), (50, 150), (150, 50)$ and $(150, 150)$ is used to study the effects. The upper table shows the results of the impulse responses to a unit shock of the observed factor and the lower table shows the impulse responses to a unit shock of the first unobserved factor. When we consider the shock to the observed factor, the coverage rate is very close to the nominal level 95% in every case in study, which validates Theorem 2. The median length is shortened as d becomes larger as suggested by the asymptotic variance of the eigenvectors. The interval will be tighter as T gets larger, which is consistent with the standard theory while the effect of a large N is more blurred. Importantly, when we consider the responses to an unobserved factor shock in the lower table, the asymptotic confidence interval tends to fall short of the nominal coverage rate. This undercoverage property is exacerbated when the sample size is small, especially when N is smaller than T ($T = 150, N = 50$).

It is evidenced that the bootstrap method can alleviate this problem by the aforementioned literature on bootstrapping factor models. To see this, Table 3 presents the coverage rate and the median length of the 95% bootstrap confidence interval proposed in section 3.3. We consider an i.i.d. bootstrap based on 599 repetitions. The upper table again shows the results of a unit shock of the observed factor and suggests that the coverage rate is very close to the nominal level 95% and the median length is somewhat longer than the asymptotic interval. The lower table shows the results of the impulse responses to a unit shock of the first unobserved factor. Clearly, the bootstrap interval gains a coverage rate that is very close to the nominal level at shorter horizons and significant improvement over the asymptotic interval. As is consistent with the seminal study of Gonçalves and Perron (2014), the bootstrap accounts for the effect of factor estimation errors that is neglected by the asymptotic interval under assumption $\sqrt{T}/N \to 0$. We have an additional caveat that the coverage rate may deteriorate as h becomes larger. This could potentially be fixed by introducing a bias correction procedure proposed by Kilian (1998).

5 Conclusion

In this study, we develop a new method of structural identification for the FAVAR models. The main idea is based on a growing literature that uses a change in shock variances that occurred on a historical date. Similar to Bai et al. (2016), the proposed method can identify the rotated version of the contemporaneous matrix, however, it has advantage as it allows for the contemporaneous matrix to be fully unrestricted. We derived the asymptotic distributions of the structural parameters and impulse responses under $\sqrt{T}/N \to 0$ as $N, T \to \infty$ and the asymptotic covariance matrices are easily constructed. A bootstrap method for the confidence interval is also studied. A Monte Carlo experiment suggests that the MSE of the proposed method becomes smaller as the change of the shock variances becomes larger. Compared to the exclusion restrictions of Bai et al. (2016), our method requires a relatively large break of the shock variances to achieve comparable performance when T is small, however, it easily outperforms as the break becomes larger. This advantage will be more distinct when we have a larger T. Both the asymptotic and bootstrap confidence intervals can provide a good coverage rate when the impulse response to a shock of the observed factor is investigated. The bootstrap interval gives more accurate coverage rate when a shock to an unobserved factor is studied, especially when N is relatively smaller than T, as suggested by the literature. Several important agendas remain for future work. The asymptotic theory under $\sqrt{T}/N \to c$ $(0 \leq c < \infty)$ has to be developed to justify the bootstrap methods. More importantly, the new identification strategy would have a large potential to apply to empirical data in face of growing literature of structural VARs reviewed by Stock and Watson (2016) and Kilian and Lütkepohl (2018), for example, especially in a data-rich environment.

Appendix A: Proof of Theorems

Proof of Theorem 1: We rewrite the reduced-form factor model (4) as

$$\begin{aligned} x_i &= F\lambda_i + G\gamma_i + u_i, \\ &= M_G F\lambda_i + G\theta_i + u_i, \end{aligned}$$

where $\theta_i \equiv [(G'G)^{-1}G'F\lambda_i + \gamma_i]$. Because $M_G F$ and G are orthogonal, by using Theorem 1 of Bai and Ng (2006), the reduced-form factor loading estimator has the property

$$\widetilde{c}_i \approx \begin{bmatrix} Q^{-1} & 0_{r_1 \times r_2} \\ 0_{r_2 \times r_1} & I_{r_2} \end{bmatrix} \begin{bmatrix} \lambda_i \\ \theta_i \end{bmatrix},$$

so that

$$\begin{aligned} \widetilde{c}_i &\approx \begin{bmatrix} Q^{-1} & 0_{r_1 \times r_2} \\ 0_{r_2 \times r_1} & I_{r_2} \end{bmatrix} \begin{bmatrix} I_{r_1} & 0_{r_1 \times r_2} \\ (G'G)^{-1}G'F & I_{r_2} \end{bmatrix} \begin{bmatrix} \lambda_i \\ \gamma_i \end{bmatrix}, \\ &= \begin{bmatrix} Q^{-1} & 0_{r_1 \times r_2} \\ (G'G)^{-1}G'F & I_{r_2} \end{bmatrix} c_i, \end{aligned}$$

where $Q \equiv V^{-1}(\widetilde{F}'M_G F/T)(\Lambda'\Lambda/N)$ and V is an $r_1 \times r_1$ diagonal matrix with the main diagonal elements as the r_1 largest eigenvalues of $M_G XX'M_G/(TN)$. This results in

$$\widetilde{\Xi} \equiv \widetilde{C}_{1:r} \xrightarrow{p} C_{1:r}\bar{R}^{-1} \equiv \Xi\bar{R}^{-1}, \tag{A.1}$$

where

$$R \equiv \begin{bmatrix} Q^{-1} & 0_{r_1 \times r_2} \\ (G'G)^{-1}G'F & I_{r_2} \end{bmatrix},$$

and $\bar{R} = p\lim(R)$, which is guaranteed to exist. This implies that the inverse has the probability limit

$$\widetilde{\Xi}^{-1} \xrightarrow{p} \bar{R}\Xi^{-1}. \tag{A.2}$$

Because the definition of the reduced-form VAR errors yields $E(e_t e_t') = B\Pi_j B'$ for $j = 1$ and 2 and Lemma 2 of Yamamoto (2017) gives $\left\| \frac{1}{T_j} \sum_j \hat{e}_t \hat{e}_t' - R\Pi_j R' \right\| = O_p(\frac{1}{\min\{N,T\}})$, the sample covariance matrix of the reduced-form VAR residuals for regime j has its limit

$$\widetilde{\Omega}_j \xrightarrow{p} \bar{R}B\Pi_j B'\bar{R}', \tag{A.3}$$

for $j = 1$ and 2.

Therefore, by using (A.1), (A.2) and (A.3),

$$\begin{aligned}\widetilde{S} &\equiv \widetilde{\Xi}\widetilde{\Omega}_1\widetilde{\Omega}_2^{-1}\widetilde{\Xi}^{-1} \xrightarrow{p} (\Xi\bar{R}^{-1})(\bar{R}B\Pi_1 B'\bar{R}')(\bar{R}B\Pi_2 B'\bar{R}')^{-1}(\bar{R}\Xi^{-1}),\\ &= \Xi B\Pi_{1/2}B^{-1}\Xi^{-1} \equiv S,\end{aligned}$$

where $\Pi_{1/2}$ is a diagonal matrix with the kth main diagonal element being $\pi_{1/2}^{(k)} \equiv \pi_1^{(k)}/\pi_2^{(k)}$ for $k = 1, ..., r$. Because $\widetilde{\Delta}$ is the eigenvectors of $\widetilde{S}$, its probability limit is the eigenvectors of

$$S = \Xi B\Pi_{1/2}B^{-1}\Xi^{-1},$$

and the diagonal elements of $\Xi B \equiv C_{1:r}^*$ are normalized by Assumption D, we obtain

$$\widetilde{\Delta} \xrightarrow{p} \Xi B, \tag{A.4}$$

and, by using (A.2) and (A.4),

$$\widetilde{B} = \widetilde{\Xi}^{-1}\hat{\Delta} \xrightarrow{p} \bar{R}\Xi^{-1}\Xi B = \bar{R}B.$$

Finally, since the reduced-form estimator satisfies $\widetilde{c}_i \approx R'^{-1}c_i$ for $i = 1, ..., N$ and $\widetilde{A}_j \approx RA_jR^{-1}$ for $j = 1, ..., p$,

$$\hat{c}_i = \widetilde{B}'\widetilde{c}_i \approx B'R'R'^{-1}c_i = B'c_i = c_i^*,$$

and

$$\hat{A}_j = \widetilde{B}^{-1}\widetilde{A}_j\widetilde{B} \approx B^{-1}R^{-1}RA_jR^{-1}RB = B^{-1}A_jB = A_j^*.$$

The results follow.■

In the following Lemmas A1 and A2, we only consider the case of spherical idiosyncratic errors to simplify the notation of the long-run variance. However, the results can be extended to the case of heteroskedasticity and standard type of serial correlations.

Lemma A1. Under Assumptions A-D, the following hold as $N, T \to \infty$ and $\sqrt{T}/N \to 0$:
(i)

$$\sqrt{T}(\widetilde{c}_i - R'^{-1}c_i) \xrightarrow{d} N(0, V_{c_i}),$$

where $V_{c_i} \equiv \sigma_i^2(\bar{R}\Sigma_H\bar{R}')^{-1}$.
(ii)

$$\sqrt{T}vec[\widetilde{A} - (I_p \otimes R'^{-1})AR'] \to N(0, V_A),$$

where

$$\begin{aligned}V_A &\equiv [I_r \otimes (I_p \otimes \bar{R})\Sigma_Z(I_p \otimes \bar{R}')]^{-1}\\ &\quad\times[(\bar{R}\Sigma_e\bar{R}') \otimes (I_p \otimes \bar{R})\Sigma_Z(I_p \otimes \bar{R}')]\\ &\quad\times[I_r \otimes (I_p \otimes \bar{R})\Sigma_Z(I_p \otimes \bar{R}')]'^{-1}.\end{aligned}$$

Proof of Lemma A1: Part (i) is a direct result of Theorem 1 of Bai and Ng (2006). Part (ii) is Theorem 1 of Yamamoto (2017).■

Lemma A2. Under Assumptions A-D, the following holds:

$$\sqrt{T}vec(\widetilde{\Xi} - \Xi R^{-1}) \xrightarrow{d} N(0, V_\Xi),$$

where

$$V_\Xi \equiv [(\bar{R}\Sigma_H\bar{R}')^{-1} \otimes \Sigma_u^{1:r}],$$

$$\Sigma_u^{1:r} \equiv \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1r} \\ \vdots & \ddots & \vdots \\ \sigma_{r1} & \cdots & \sigma_r^2 \end{bmatrix}, \ \sigma_i^2 = E(u_{it}^2) \text{ and } \sigma_{ij} = E(u_{it}u_{jt}).$$

Proof of Lemma A2: This is directly derived from Lemma A1 (i).■

Lemma A3. The following holds as $N, T \to \infty$ with $\sqrt{T}/N \to 0$:

$$\sqrt{T}vec(\widetilde{\Omega}_1\widetilde{\Omega}_2^{-1} - R\Omega_1\Omega_2^{-1}R^{-1}) \xrightarrow{d} N(0, V_{\Omega_{1/2}}),$$

where

$$\begin{aligned} V_{\Omega_{1/2}} \ \equiv \ & (\bar{R}'^{-1}\Omega_2^{-1}\bar{R}^{-1} \otimes I_r)V_{\Omega_1}(\bar{R}'^{-1}\Omega_2^{-1}\bar{R}^{-1} \otimes I_r) \\ & +(\bar{R}'^{-1}\Omega_2^{-1}\bar{R}^{-1} \otimes \bar{R}\Omega_1\Omega_2^{-1}\bar{R}^{-1})V_{\Omega_2}(\bar{R}'^{-1}\Omega_2^{-1}\bar{R}^{-1} \otimes \bar{R}\Omega_1\Omega_2^{-1}\bar{R}^{-1}), \end{aligned}$$

with

$$\begin{aligned} V_{\Omega_j} \ &\equiv \ 2P_D(\bar{R}\Omega_j\bar{R}' \otimes \bar{R}\Omega_j\bar{R}')P_D, \quad \text{for } j = 1 \text{ and } 2, \\ P_D \ &\equiv \ D_r(D_r'D_r)^{-1}D_r', \end{aligned}$$

and D_r being a duplication matrix of dimension r.

Proof of Lemma A3: We first obtain the following result for the asymptotic distribution of the sample covariance matrix of the VAR errors by using Lemma 4 of Yamamoto (2017)

$$\sqrt{T}vec(\widetilde{\Omega}_j - R\Omega_jR^{-1}) \xrightarrow{d} N(0, V_{\Omega_j}), \tag{A.5}$$

where

$$V_{\Omega_j} \equiv 2P_D(\bar{R}\Omega_j\bar{R}' \otimes \bar{R}\Omega_j\bar{R}')P_D, \tag{A.6}$$

for $j = 1, 2$ as $N, T \to \infty$ and $\sqrt{T}/N \to 0$. Then, we let $\bar{\Omega}_j \equiv R\Omega_jR'$ for $j = 1, 2$. The delta method yields

$$\sqrt{T}vec(\widetilde{\Omega}_1\widetilde{\Omega}_2^{-1} - R\Omega_1\Omega_2^{-1}R^{-1}) \xrightarrow{d} N(0, V_{\Omega_{1/2}}),$$

where

$$
\begin{aligned}
V_{\Omega_{1/2}} \equiv\ & \frac{\partial vec(\bar{\Omega}_1\bar{\Omega}_2^{-1})}{\partial vec(\bar{\Omega}_1)'} V_{\Omega_1} \frac{\partial vec(\bar{\Omega}_1\bar{\Omega}_2^{-1})'}{\partial vec(\bar{\Omega}_1)} \\
& + \frac{\partial vec(\bar{\Omega}_1\bar{\Omega}_2^{-1})}{\partial vec(\bar{\Omega}_2)'} V_{\Omega_2} \frac{\partial vec(\bar{\Omega}_1\bar{\Omega}_2^{-1})'}{\partial vec(\bar{\Omega}_2)}. \qquad \text{(A.7)}
\end{aligned}
$$

Furthermore,

$$
\frac{\partial vec(\bar{\Omega}_1\bar{\Omega}_2^{-1})}{\partial vec(\bar{\Omega}_1)'} = \bar{\Omega}_2^{-1} \otimes I_r = (\bar{R}'^{-1}\bar{\Omega}_2^{-1}\bar{R}^{-1}) \otimes I_r, \qquad \text{(A.8)}
$$

$$
\begin{aligned}
\frac{\partial vec(\bar{\Omega}_1\bar{\Omega}_2^{-1})}{\partial vec(\bar{\Omega}_2)'} &= -(I_r \otimes \bar{\Omega}_1)(\bar{\Omega}_2^{-1} \otimes \bar{\Omega}_2^{-1}), \\
&= -(\bar{\Omega}_2^{-1} \otimes \bar{\Omega}_1\bar{\Omega}_2^{-1}), \\
&= -(\bar{R}'^{-1}\Omega_2^{-1}\bar{R}^{-1} \otimes \bar{R}\Omega_1\bar{R}'\bar{R}'^{-1}\Omega_2^{-1}\bar{R}^{-1}). \qquad \text{(A.9)}
\end{aligned}
$$

Plugging (A.6), (A.8), and (A.9) in (A.7) yields the result. ■

Lemma A4. The following holds as $N, T \to \infty$ with $\sqrt{T}/N \to 0$:

$$
\sqrt{T} vec(\widetilde{Q} - \Xi\Omega_1\Omega_2^{-1}\Xi^{-1}) \xrightarrow{d} N(0, V_Q),
$$

where

$$
\begin{aligned}
V_Q \equiv\ & (\Xi'^{-1}\bar{R}' \otimes \Xi\bar{R}^{-1})V_{\Omega_{1/2}}(\bar{R}\Xi^{-1} \otimes \bar{R}'^{-1}\Xi') \\
& + (\Xi'^{-1}\bar{R}' \otimes \Xi\Omega_{1/2}\Xi^{-1})V_{\Xi}(\bar{R}\Xi^{-1} \otimes \Xi'^{-1}\Omega_{1/2}\Xi') \\
& + (\Xi'^{-1}\Omega_{1/2}\bar{R}^{-1} \otimes I_r)V_{\Xi}(\bar{R}'^{-1}\Omega_{1/2}\Xi^{-1} \otimes I_r),
\end{aligned}
$$

where $V_{\Omega_{1/2}}$and V_{Ξ} are given in Lemmas A3 and A2, respetively.

Proof of Lemma A4: We let the probability limit of $\widetilde{\Omega}_1\widetilde{\Omega}_2^{-1}$ be $\bar{\Omega}_{1/2} \equiv \bar{R}\Omega_1\bar{R}'(\bar{R}\Omega_2\bar{R}')^{-1} = \bar{R}\Omega_1\Omega_2^{-1}\bar{R}^{-1}$ and $\bar{\Xi} \equiv \Xi\bar{R}^{-1}$. Then,

$$
\begin{aligned}
\widetilde{Q} &\approx \bar{\Xi}\bar{\Omega}_{1/2}\bar{\Xi}^{-1} \\
&= (\Xi\bar{R}^{-1})(\bar{R}\Omega_1\Omega_2^{-1}\bar{R}^{-1})(\bar{R}\Xi^{-1}) \\
&= \Xi\Omega_1\Omega_2^{-1}\Xi^{-1} \equiv Q.
\end{aligned}
$$

Using Lemmas A2 and A3, the delta method yields

$$
\sqrt{T} vec(\widetilde{Q} - Q) = \sqrt{T} vec(\widetilde{Q} - \Xi\Omega_{1/2}\Xi^{-1}) \xrightarrow{d} N(0, V_Q),
$$

as $N, T \to \infty$ with $\sqrt{T}/N \to 0$, where

$$V_Q \equiv \frac{\partial vec(Q)}{\partial vec(\bar{\Omega}_{1/2})'} V_{\Omega_{1/2}} \frac{\partial vec(Q)'}{\partial vec(\bar{\Omega}_{1/2})} + \frac{\partial vec(Q)}{\partial vec(\bar{\Xi})'} V_{\Xi} \frac{\partial vec(Q)'}{\partial vec(\bar{\Xi})}. \tag{A.10}$$

Furthermore,

$$\begin{aligned} \frac{\partial vec(Q)}{\partial vec(\bar{\Omega}_{1/2})'} &= \frac{\partial vec(\bar{\Xi}\bar{\Omega}_{1/2}\bar{\Xi}^{-1})}{\partial vec(\bar{\Omega}_{1/2})'} \\ &= (\bar{\Xi}'^{-1} \otimes \bar{\Xi}) = (\Xi'^{-1} R' \otimes \Xi R^{-1}), \end{aligned} \tag{A.11}$$

and

$$\begin{aligned} \frac{\partial vec(Q)}{\partial vec(\bar{\Xi})'} &= \frac{\partial vec(\bar{\Xi}\bar{\Omega}_{1/2}\bar{\Xi}^{-1})}{\partial vec(\bar{\Xi})'} \\ &= -(I_r \otimes \bar{\Xi}\bar{\Omega}_{1/2})(\bar{\Xi}'^{-1} \otimes \bar{\Xi}^{-1}) + (\bar{\Xi}'^{-1}\bar{\Omega}'_{1/2} \otimes I_r), \\ &= -(\bar{\Xi}'^{-1} \otimes \bar{\Xi}\bar{\Omega}_{1/2}\bar{\Xi}^{-1}) + (\bar{\Xi}'^{-1}\bar{\Omega}'_{1/2} \otimes I_r), \\ &= -(\Xi'^{-1}\bar{R}' \otimes \Xi\Omega_{1/2}\Xi^{-1}) + (\Xi'^{-1}\Omega'_{1/2}\bar{R}^{-1} \otimes I_r). \end{aligned} \tag{A.12}$$

Hence, plugging (A.11) and (A.12) in (A.10) yields the result.■

Lemma A5. Under Assumptions A-D, for the normalized eigenvector matrix $\widetilde{\Delta} = [\widetilde{\delta}_1, \cdots, \widetilde{\delta}_k]$ of $\widetilde{Q}$, the following holds as $N, T \to \infty$ with $\sqrt{T}/N \to 0$:

$$\sqrt{T}(\widetilde{\Delta} - C^*_{1:r}) \xrightarrow{d} N(0, V_\Delta),$$

where

$$V_\Delta \equiv \begin{bmatrix} V_{\delta_{11}} & \cdots & V_{\delta_{1r}} \\ \vdots & \ddots & \vdots \\ V_{\delta_{r1}} & \cdots & V_{\delta_{rr}} \end{bmatrix},$$

with

$$\begin{aligned} V_{\delta_{kk}} &\equiv \textstyle\sum_{j=1, j\neq k}^{r} \dfrac{AVar(v_{ii})}{(\pi^{(k)}_{1/2} - \pi^{(j)}_{1/2})^2} c^*_{\bullet j} c^{*\prime}_{\bullet j}, \\ V_{\delta_{kl}} &\equiv -\dfrac{AVar(v_{kl})}{(\pi^{(k)}_{1/2} - \pi^{(j)}_{1/2})^2} c^*_{\bullet k} c^{*\prime}_{\bullet l}, \quad \text{for } k \neq l, \end{aligned}$$

v_{ij} is the (i, j)th element of $\sqrt{T}(\widetilde{Q} - Q)$ and $c^*_{\bullet j}$ is the jth column of $C^*_{1:r}$.

Proof of Lemma A5: To prove this lemma, we follow the derivation of asymptotic distribution of eigenvectors of sample covariance in Theorem 13.5.1 of Anderson (2003). However, we take eigenvectors of $\widetilde{Q}$ which is not symmetric in finite samples and its finite sample distribution is not necessarily Wishart distribution. However, the asymptotic distribution of $\widetilde{Q}$ is obtained in Lemma A4. We first consider the asymptotic distribution of the normalized eigenvector E of $C^{*-1}\widetilde{Q}C^* \equiv S$. Then, consider that of $\widetilde{\Delta} = C^*E$. This relation holds because if we let D be the eigenvalue matrix then

$$\begin{aligned} SE &= ED \Leftrightarrow C^*SC^{*-1}(C^*E) = (C^*E)D \\ &\Leftrightarrow \widetilde{Q}\widetilde{\Delta} = \widetilde{\Delta}D. \end{aligned}$$

Let $\sqrt{T}(S - \Pi_{1/2}) = V$ and $\sqrt{T}(E - I_r) = W$, then $SE = ED$ implies

$$\left(\Pi_{1/2} + \frac{1}{\sqrt{T}}V\right)\left(I + \frac{1}{\sqrt{T}}W\right) = \left(I + \frac{1}{\sqrt{T}}W\right)\left(\Pi_{1/2} + \frac{1}{\sqrt{T}}D\right),$$

so that

$$V = W\Pi_{1/2} - \Pi_{1/2}W + D + O_p(T^{-1/2}).$$

If we neglect terms of order $T^{-1/2}$, then

$$V = (W\Pi_{1/2} - \Pi_{1/2}W) + D.$$

It is shown that for $w_{ii} = 0$ because of the diagonal elements are known by Assumption D and for $i \neq j$

$$w_{ij} = \frac{v_{ij}}{\pi_{1/2}^{(j)} - \pi_{1/2}^{(i)}}.$$

Because $\widetilde{\delta}_k$ is the kth column of $C^*_{1:r}E$,

$$\sqrt{T}(\widetilde{\delta}_k - c^*_{\bullet j}) \xrightarrow{d} N\left(0, \textstyle\sum_{j=1,j\neq k}^r \frac{AVar(v_{ii})}{(\pi_{1/2}^{(k)} - \pi_{1/2}^{(j)})^2} c^*_{\bullet j}c^{*\prime}_{\bullet j}\right),$$

and the asymptotic covariance of $\widetilde{\delta}_k$ and $\widetilde{\delta}_l$ $(k \neq l)$ is given by

$$-\frac{AVar(v_{kl})}{(\pi_{1/2}^{(k)} - \pi_{1/2}^{(l)})^2} c^*_{\bullet k}c^{*\prime}_{\bullet l}.$$

The result follows.■

Proof of Theorem 2: We let $\bar{B} \equiv \bar{R}B$, $\bar{c}_i \equiv c_i\bar{R}^{-1}$, and $\bar{A}_j \equiv \bar{R}A_j\bar{R}^{-1}$. (i) The delta method yields

$$\Omega_{c_i} \equiv \frac{\partial \bar{B}'\bar{c}_i}{\partial \bar{c}_i'} V_{c_i} \frac{\partial \bar{c}_i'\bar{B}}{\partial \bar{c}_i} + \frac{\partial \bar{B}'\bar{c}_i}{\partial vec(\bar{B})'} V_B \frac{\partial \bar{c}_i'\bar{B}}{\partial vec(\bar{B})},$$

but

$$\frac{\partial \bar{B}'\bar{c}_i}{\partial \bar{c}_i'} = \bar{B}',$$

and

$$\begin{aligned}\frac{\partial \bar{B}'\bar{c}_i}{\partial vec(\bar{B})'} &= (\bar{c}_i' \otimes I_r)\frac{\partial vec(\bar{B}')}{\partial vec(\bar{B})'} \\ &= (\bar{c}_i' \otimes I_r)K_{rr} = (\bar{R}'^{-1}c_i' \otimes I_r)K_{rr},\end{aligned}$$

yield the results.

(ii) The delta method yields

$$\Omega_{A_j} = \frac{\partial vec(\bar{B}^{-1}\bar{A}_j\bar{B})}{\partial vec(\bar{A}_j)'}V_{A_j}\frac{\partial vec(\bar{B}^{-1}\bar{A}_j\bar{B})'}{\partial vec(\bar{A}_j)} + \frac{\partial vec(\bar{B}^{-1}\bar{A}_j\bar{B})}{\partial vec(\bar{B})'}V_{A_j}\frac{\partial vec(\bar{B}^{-1}\bar{A}_j\bar{B})'}{\partial vec(\bar{B})},$$

but

$$\begin{aligned}\frac{\partial vec(\bar{B}^{-1}\bar{A}_j\bar{B})}{\partial vec(\bar{A}_j)'} &= (\bar{B}' \otimes \bar{B}^{-1}), \\ &= (B'\bar{R}' \otimes B^{-1}\bar{R}^{-1})\end{aligned}$$

and

$$\begin{aligned}\frac{\partial vec(\bar{B}^{-1}\bar{A}_j\bar{B})}{\partial vec(\bar{B})'} &= (I_r \otimes \bar{B}^{-1}\bar{A}_j) + (\bar{B}'\bar{A}_j' \otimes I_r)\frac{\partial vec(\bar{B}^{-1})}{\partial vec(\bar{B})'}, \\ &= (I_r \otimes \bar{B}^{-1}\bar{A}_j) - (\bar{B}'\bar{A}_j' \otimes I_r)(\bar{B}'^{-1} \otimes \bar{B}^{-1}), \\ &= (I_r \otimes \bar{B}^{-1}\bar{A}_j) - (\bar{B}'\bar{A}_j'\bar{B}'^{-1} \otimes \bar{B}^{-1}), \\ &= (I_r \otimes B^{-1}A_j\bar{R}^{-1}) - (B'A_j'B'^{-1} \otimes B^{-1}\bar{R}^{-1}),\end{aligned}$$

yield the result.

(iii) The proof is essentially the same as Proposition 3.6 of Lütkepohl (2005) and thus omitted.■

Appendix B: Estimation under IR_a, IR_b and IR_c assumptions

We obtain $\widetilde{B}$ under the three sets of identification assumptions of Bai et al.'s (2016) as follows.

1. Estimate the reduced-form models to obtain $\widetilde{A}$, $\widetilde{C}$ and $\widetilde{e}_t$. Let $\widetilde{C}_{1:r}$ be the first r rows of C.

2. Construct the sample covariance matrices of the VAR residuals for the whole sample by

$$\widetilde{\Omega} = \frac{1}{T - pk - 1} \sum_{t=p+1}^{T} \widetilde{e}_t \widetilde{e}_t',$$

and obtain its partitions $\widetilde{\Omega}_{11}$, $\widetilde{\Omega}_{12}$, $\widetilde{\Omega}_{21}$ and $\widetilde{\Omega}_{22}$. Let

$$\widetilde{\Omega}_{11\cdot 2} = \widetilde{\Omega}_{11} - \widetilde{\Omega}_{12}\widetilde{\Omega}_{22}^{-1}\widetilde{\Omega}_{21}.$$

IRa Let V_a be the eigenvector matrix of

$$\widetilde{\Omega}_{11\cdot 2}^{1/2}(\widetilde{\Lambda}'\widetilde{\Sigma}_u^{-1}\widetilde{\Lambda}/N)\widetilde{\Omega}_{11\cdot 2}^{1/2},$$

in descending order of their associated eigenvalues. Calculate

$$\widetilde{B} = \begin{bmatrix} V_a'\widetilde{\Omega}_{11\cdot 2}^{-1/2} & -V_a'\widetilde{\Omega}_{11\cdot 2}^{-1/2}\widetilde{\Omega}_{12}\widetilde{\Omega}_{22}^{-1} \\ 0_{r_2\times r_1} & I_{r_2} \end{bmatrix}^{-1}.$$

IRb Let V_b be the same as V_a and Q_b be the QR decomposition of $\widetilde{\Omega}_{11\cdot 2}^{1/2}\widetilde{\Lambda}_{1:r}'$. Then,

$$\widetilde{B} = \begin{bmatrix} Q_b'\widetilde{\Omega}_{11\cdot 2}^{-1/2} & -Q_b'\widetilde{\Omega}_{11\cdot 2}^{-1/2}\widetilde{\Omega}_{12}\widetilde{\Omega}_{22}^{-1} \\ 0_{r_2\times r_1} & I_{r_2} \end{bmatrix}^{-1}.$$

IRc Let

$$\widetilde{B} = \begin{bmatrix} \widetilde{\Lambda}_{1:r} & -\widetilde{\Lambda}_{1:r}\widetilde{\Omega}_{12}\widetilde{\Omega}_{22}^{-1} \\ 0_{r_2\times r_1} & I_{r_2} \end{bmatrix}^{-1}.$$

References

Anderson, T.W., 2003, An introduction to multivariate statistical analysis: third edition, Wiley.

Bai, J., 2003, Inferential theory for factor models of large dimensions, *Econometrica* 71(1), 135-171.

Bai, J., K. Li, and L. Lu, 2016, Estimation and inference on FAVAR models, *Journal of Business and Economic Statistics* 34(4), 620-641.

Bai, J., and S. Ng, 2006, Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions, *Econometrica* 74(4), 1133-1150.

Bai, J., and S. Ng, 2013, Principal components estimation and identification of the factors, *Journal of Econometrics* 176, 18-29.

Bernanke, B.S., J. Boivin, and P. Eliasz, 2005, Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach, *Quarterly Journal of Economics* 120(1), 387-422.

Breitung, J. and S. Eickmeier, 2011, Testing for structural breaks in dynamic factor models, *Journal of Econometrics* 163, 71-84.

Brunnermeier, M., D. Palia, K.A. Sastry, and C.A. Sims, 2016, Feedbacks: Financial markets and economic activity.

Casini A. and P. Perron, 2017, Structural breaks in time series, Unpublished manuscript, Boston University.

Chen, L., J.J.Dolado. and J. Gonzalo, 2014, Detecting big structural breaks in large factor models, *Journal of Econometrics* 180, 30-48.

Djogbenou, A., S. Gonçalves, and B. Perron, 2015, Bootstrap inference in regressions with estimated factors and serial correlation, *Journal of Time Series Analysis* 36, 481-502.

Gonçalves, S., and B. Perron, 2014, Bootstrapping factor-augmented regression models, *Journal of Econometrics* 182, 156-173.

Gonçalves, S., and B. Perron, 2016, Bootstrapping factor models with cross-sectional dependence, mimeo.

Gonçalves, S., B. Perron, and A. Djogbenou, 2017, Bootstrap prediction intervals for factor models, *Journal of Business and Economic Statistics* 35(1), 53-69.

Han, X. and A. Inoue, 2015, Tests for parameter instability in dynamic factor models, *Econometric Theory* 31, 1117-1152.

Kilian, L., 1998, Small sample confidence intervals for impulse response functions, *Review of Economics and Statistics* 80(2), 218-230.

Kilian, L., and H. Lütkepohl, 2017, Structural vector autoregressive analysis, Cambridge University Press (forthcoming), http://www-personal.umich.edu/~lkilian/book.html

Lanne, M. and H. Lütkepohl, 2008, Identifying monetary policy shochs via changes in volatility, *Journal of Money, Credit and Banking* 40, 1131-1149.

Lütkepohl, H., 1990, Asymptotic distributions of impulse response functions and forecast error variance decompositions of vector autoregressive models, *Review of Economics and Statistics* 72(1), 116-125.

Lütkepohl, H., 2005, New introduction to multiple time series analysis, Springer.

Lütkepohl, H., 2013, Identifying structural vector autoregressions via changes in volatility, *Advances in Econometrics* 32, 169-203.

Lütkepohl, H. and A. Netsunajev, 2017, Structural vector autoregressions with heteroskedasticity: A review of different volatility models, *Economics and Statistics* 1, 2-18.

Perron, P., and Z. Qu, 2007, Estimating and testing structural changes in multivariate regressions, *Econometrica* 75(2), 459-502.

Rigobon, R., 2003, Identification through heteroskedasticity, *The Review of Economics and Statistics* 85(4), 777-792.

Rigobon, R. and Sack, B., 2003, Measuring the effect of monetary policy to the stock market, *Quarterly Journal of Economics* 118, 639-669.

Sims, C.A., 1980, Macroeconomic and reality, *Econometrica* 48(1), 1-48.

Stock, J.H., and M.W. Watson, 2016, Dynamic factor models. Oxford Handbook of Economic Forecasting, M.P. Clements and D.F. Hendry Eds. Oxford University Press.

Yamamoto, Y., 2017, Bootstrap inference for impulse response functions in factor-augmented vector autoregressions, HIAS Discussion Paper series E-4.

Yamamoto, Y., and S. Tanaka, 2015, Testing for factor loading structural change under common breaks, *Journal of Econometrics* 189(1), 187-206.

Table 1. Rejection frequencies of the structural break test

d	T=50, N=50	T=50, N=150	T=150, N=50	T=150, N=150
2	0.165	0.175	0.966	0.970
5	0.570	0.573	1.000	1.000
10	0.834	0.842	1.000	1.000

Table 2. Mean squared error of the impulse response estimator

Scenario A: Correct DGP

h=1	d	IRa	IRb	IRc	IRv
T=50, N=50	2	0.134	0.227	0.162	7.178
	5	0.149	0.307	0.198	0.561
	10	0.165	0.440	0.256	0.333
T=50, N=150	2	0.130	0.231	0.165	3.221
	5	0.145	0.295	0.199	0.620
	10	0.173	0.456	0.254	0.325
T=150, N=50	2	0.104	0.137	0.115	0.307
	5	0.109	0.165	0.124	0.105
	10	0.115	0.211	0.143	0.072
T=150, N=150	2	0.106	0.140	0.115	0.284
	5	0.110	0.160	0.126	0.101
	10	0.116	0.202	0.142	0.080

h=5	d	IRa	IRb	IRc	IRv
T=50, N=50	2	0.002	0.005	0.004	0.037
	5	0.004	0.009	0.006	0.024
	10	0.006	0.016	0.011	0.029
T=50, N=150	2	0.002	0.005	0.003	0.042
	5	0.004	0.009	0.006	0.022
	10	0.006	0.018	0.011	0.031
T=150, N=50	2	0.001	0.001	0.001	0.006
	5	0.001	0.003	0.002	0.007
	10	0.002	0.005	0.003	0.010
T=150, N=150	2	0.001	0.002	0.001	0.007
	5	0.001	0.003	0.002	0.007
	10	0.002	0.005	0.003	0.010

Scenario B: Unrestricted DGP

h=1	d	IRa	IRb	IRc	IRv
T=50, N=50	2	0.730	1.021	0.869	3.766
	5	1.160	1.635	1.314	0.575
	10	1.775	2.432	1.928	0.342
T=50, N=150	2	0.741	1.076	0.884	5.264
	5	1.257	1.687	1.401	0.616
	10	1.774	2.540	1.952	0.342
T=150, N=50	2	0.731	0.929	0.850	0.298
	5	1.257	1.589	1.379	0.103
	10	1.864	2.408	2.031	0.078
T=150, N=150	2	0.749	0.910	0.851	0.275
	5	1.235	1.540	1.348	0.100
	10	1.812	2.350	1.984	0.070

h=5	d	IRa	IRb	IRc	IRv
T=50, N=50	2	0.013	0.019	0.014	0.039
	5	0.022	0.034	0.023	0.025
	10	0.029	0.049	0.030	0.027
T=50, N=150	2	0.015	0.019	0.015	0.038
	5	0.021	0.033	0.022	0.023
	10	0.029	0.052	0.031	0.032
T=150, N=50	2	0.013	0.015	0.013	0.006
	5	0.020	0.026	0.021	0.006
	10	0.029	0.041	0.031	0.010
T=150, N=150	2	0.012	0.014	0.012	0.006
	5	0.021	0.026	0.021	0.007
	10	0.031	0.041	0.031	0.010

Table 2. Mean squared error of the impulse response estimator (continued)

Scenario C: DGP of IR_a and IR_b assumptions

h=1	d	IRa	IRb	IRc	IRv
T=50, N=50	2	0.129	0.219	0.146	3.986
	5	0.140	0.306	0.171	0.504
	10	0.175	0.441	0.213	0.320
T=50, N=150	2	0.131	0.232	0.146	3.405
	5	0.150	0.325	0.175	0.505
	10	0.176	0.466	0.213	0.304
T=150, N=50	2	0.107	0.145	0.112	0.240
	5	0.110	0.167	0.118	0.089
	10	0.117	0.209	0.130	0.068
T=150, N=150	2	0.098	0.132	0.103	0.231
	5	0.115	0.166	0.124	0.085
	10	0.113	0.208	0.126	0.064

h=5	d	IRa	IRb	IRc	IRv
T=50, N=50	2	0.002	0.005	0.003	0.025
	5	0.004	0.011	0.005	0.013
	10	0.007	0.019	0.008	0.016
T=50, N=150	2	0.002	0.005	0.002	0.029
	5	0.004	0.011	0.005	0.012
	10	0.007	0.017	0.008	0.015
T=150, N=50	2	0.001	0.002	0.001	0.003
	5	0.001	0.002	0.001	0.002
	10	0.002	0.005	0.003	0.004
T=150, N=150	2	0.001	0.002	0.001	0.003
	5	0.001	0.003	0.001	0.003
	10	0.002	0.005	0.003	0.004

Scenario D: DGP of IR_c assumptions

h=1	d	IRa	IRb	IRc	IRv
T=50, N=50	2	0.138	0.260	0.161	4.285
	5	0.154	0.383	0.195	0.621
	10	0.201	0.577	0.254	0.390
T=50, N=150	2	0.140	0.275	0.161	4.443
	5	0.165	0.406	0.200	0.629
	10	0.202	0.604	0.253	0.370
T=150, N=50	2	0.109	0.160	0.117	0.297
	5	0.114	0.194	0.126	0.109
	10	0.125	0.254	0.143	0.082
T=150, N=150	2	0.101	0.146	0.107	0.293
	5	0.120	0.193	0.131	0.104
	10	0.121	0.249	0.139	0.077

h=5	d	IRa	IRb	IRc	IRv
T=50, N=50	2	0.003	0.007	0.004	0.033
	5	0.006	0.015	0.007	0.017
	10	0.009	0.026	0.011	0.022
T=50, N=150	2	0.003	0.007	0.003	0.038
	5	0.005	0.015	0.007	0.017
	10	0.010	0.024	0.012	0.021
T=150, N=50	2	0.001	0.002	0.001	0.004
	5	0.002	0.003	0.002	0.003
	10	0.003	0.007	0.004	0.006
T=150, N=150	2	0.001	0.002	0.001	0.004
	5	0.002	0.004	0.002	0.004
	10	0.003	0.007	0.004	0.006

Table 3. Coverage rate and median length of the asymptotic confidence interval

Responses to a shock of the observed factor

		Coverage rate					Median length				
	d	h=1	2	3	4	5	h=1	2	3	4	5
T=50, N=50	2	94.8	96.2	96.1	94.1	92.2	3.27	2.70	2.21	1.76	1.37
	5	98.1	98.3	98.2	95.8	93.1	2.27	2.14	1.91	1.63	1.30
	10	99.2	98.2	97.3	96.8	94.8	1.98	2.18	2.08	1.77	1.44
T=50, N=150	2	91.3	95.8	96.6	95.0	91.5	7.20	4.05	2.67	1.98	1.42
	5	97.9	97.6	98.2	96.9	93.3	5.09	3.45	2.56	1.96	1.48
	10	99.3	98.0	98.3	96.9	94.7	4.46	3.57	2.78	2.10	1.54
T=150, N=50	2	97.6	97.6	97.5	96.5	95.4	1.26	1.10	0.95	0.80	0.64
	5	98.2	97.6	97.9	97.1	96.0	0.99	1.01	0.94	0.82	0.68
	10	99.2	97.8	98.2	97.4	96.2	0.87	1.04	1.04	0.92	0.77
T=150, N=150	2	95.4	96.5	98.1	98.2	95.9	3.16	1.76	1.19	0.88	0.67
	5	98.0	97.3	97.8	97.3	96.3	2.24	1.59	1.24	0.99	0.77
	10	98.5	97.3	98.5	97.9	96.9	1.89	1.70	1.44	1.18	0.95

Responses to a shock of the first unobserved factor

		Coverage rate					Median length				
	d	h=1	2	3	4	5	h=1	2	3	4	5
T=50, N=50	2	95.1	96.0	94.1	90.5	85.3	3.50	2.67	2.16	1.75	1.36
	5	96.2	97.0	92.8	88.2	81.9	2.25	1.88	1.63	1.36	1.09
	10	94.8	95.9	91.5	83.3	75.4	1.64	1.62	1.49	1.27	1.00
T=50, N=150	2	80.7	92.0	92.6	88.7	84.5	4.24	2.26	1.59	1.19	0.94
	5	79.1	90.3	88.4	83.0	78.9	2.88	1.63	1.21	0.95	0.72
	10	78.9	89.8	90.1	84.1	79.6	2.49	1.41	1.12	0.90	0.71
T=150, N=50	2	81.3	87.1	93.3	93.0	92.0	1.38	0.58	0.33	0.22	0.16
	5	85.9	89.0	90.7	89.3	86.8	1.09	0.48	0.29	0.20	0.14
	10	89.1	89.1	89.7	85.0	81.0	0.96	0.45	0.28	0.18	0.12
T=150, N=150	2	85.9	91.3	93.3	90.2	87.9	2.84	1.43	1.11	0.92	0.75
	5	88.2	91.3	93.1	90.2	86.3	2.16	1.13	0.94	0.79	0.65
	10	90.0	91.2	91.3	87.2	82.1	1.97	1.04	0.87	0.74	0.60

Table 3. Coverage rate and median length of the bootstrap confidence interval

Responses to a shock of the observed factor

	c	Coverage rate h=1	2	3	4	5	Median length h=1	2	3	4	5
T=50, N=50	2	98.6	97.5	96.7	93.7	90.6	6.02	4.23	2.74	1.73	1.17
	5	99.2	98.5	96.6	93.8	90.2	2.63	2.28	1.70	1.24	0.90
	10	98.2	98.1	95.9	92.1	88.3	2.03	2.08	1.59	1.19	0.85
T=50, N=150	2	98.1	98.3	98.0	96.7	94.6	13.55	5.88	3.20	1.98	1.35
	5	99.2	98.5	98.5	97.0	93.7	5.40	3.46	2.20	1.46	1.03
	10	98.0	98.1	98.2	97.8	94.1	4.01	3.26	2.11	1.45	1.00
T=150, N=50	2	99.2	97.8	97.0	95.4	92.4	1.39	1.17	0.91	0.68	0.52
	5	97.3	96.3	95.3	93.0	88.6	0.93	0.92	0.77	0.61	0.46
	10	94.8	96.4	94.5	92.5	86.9	0.85	0.99	0.84	0.68	0.52
T=150, N=150	2	98.9	98.3	97.9	95.3	91.0	3.65	1.81	1.10	0.74	0.54
	5	97.3	95.9	95.9	93.4	87.1	2.01	1.46	0.97	0.69	0.51
	10	95.7	95.1	95.8	94.1	88.1	1.65	1.60	1.11	0.80	0.58

Responses to a shock of the first unobserved factor

	c	Coverage rate h=1	2	3	4	5	Median length h=1	2	3	4	5
T=50, N=50	2	97.7	98.0	95.5	90.4	83.9	7.94	5.78	4.07	2.78	1.77
	5	99.1	99.4	97.0	92.0	85.5	5.29	4.22	3.15	2.25	1.53
	10	97.8	98.5	95.1	89.3	82.6	4.01	3.79	2.87	2.00	1.41
T=50, N=150	2	98.0	97.6	96.2	91.7	87.4	15.18	5.41	2.97	1.76	1.17
	5	98.0	97.5	94.3	89.6	82.8	12.85	3.92	2.40	1.44	0.96
	10	98.0	97.8	96.5	91.2	85.9	11.51	3.59	2.27	1.45	1.00
T=150, N=50	2	98.3	98.2	97.5	95.9	95.4	3.79	1.23	0.49	0.27	0.18
	5	99.1	97.9	95.5	95.1	93.9	2.40	0.84	0.40	0.25	0.17
	10	99.0	97.6	95.5	95.2	94.9	2.24	0.79	0.38	0.25	0.18
T=150, N=150	2	98.8	98.2	96.7	94.4	90.8	6.67	2.55	1.67	1.18	0.89
	5	99.4	99.1	97.1	93.9	90.6	4.70	1.95	1.34	1.02	0.80
	10	99.5	98.6	96.4	93.7	88.4	4.46	1.84	1.27	1.00	0.80