

HIAS-E-26

# Bootstrap Inference for Impulse Response Functions in Factor-Augmented Vector Autoregressions

Yohei Yamamoto

*Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo, Japan 186-8601*

May 28, 2016



Hitotsubashi Institute for Advanced Study, Hitotsubashi University  
2-1, Naka, Kunitachi, Tokyo 186-8601, Japan  
tel: +81 42 580 8604    <http://hias.ad.hit-u.ac.jp/>

HIAS discussion papers can be downloaded without charge from:  
<http://hdl.handle.net/10086/27202>  
<https://ideas.repec.org/s/hit/hiasdp.html>

All rights reserved.

# Bootstrap Inference for Impulse Response Functions in Factor-Augmented Vector Autoregressions

Yohei Yamamoto\*

Hitotsubashi University

May 28, 2016

## Abstract

In this paper, we consider residual-based bootstrap methods à la Gonçalves and Peron (2014) to construct the confidence interval for structural impulse response functions in factor-augmented vector autoregressions. In particular, we compare the bootstrap with factor estimation (Procedure A) with the bootstrap without factor estimation (Procedure B). In theory, both procedures are asymptotically valid under a condition  $\sqrt{T}/N \rightarrow 0$ , where  $N$  and  $T$  are the cross-sectional dimension and the time dimension, respectively. Even when  $\sqrt{T}/N \rightarrow 0$  is irrelevant, Procedure A still accounts for the effect of the factor estimation errors on the impulse response function estimate and it achieves good coverage rates in most cases. On the contrary, Procedure B is invalid in such cases and tends to undercover if  $N$  is much smaller than  $T$ . However, Procedure B is implemented more straightforwardly from the standard structural VARs and the length of the confidence interval is shorter than that of Procedure A in finite samples. Given that Procedure B still gives a satisfactory coverage rate unless  $N$  is very small, it remains in consideration of empirical use, although using Procedure A is safer as it correctly accounts for the effect of the factor estimation errors.

**JEL Classification Numbers:** C14, C22

**Keywords:** factor-augmented vector autoregression, structural identification, coverage rate, impulse response function

---

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

## 1 Introduction

Factor-augmented vector autoregressions (FAVARs), introduced by Bernanke et al. (2005) and further explored by Stock and Watson (2005, 2010) in addition to many other empirical studies, have various attractive features for empirical researchers. For example, the factors essentially reduce the data dimension, enabling a conventional small-scale VAR framework to accommodate the vast amount of information contained in a large panel data set. In addition, if we consider that certain concepts in economic models such as productivity and inflation are better captured by latent factors measured by using multiple indicators rather than by a single specific series (e.g., Sargent and Sims, 1977; Boivin and Giannoni, 2006), FAVAR is a suitable specification to address such measurement issues. Therefore, the literature on the empirical applications of FAVARs is rapidly growing across various topics and a current but non-comprehensive list includes Ang and Piazzesi (2003), Giannone et al. (2005), Boivin et al. (2009, 2013), Acconcia and Simonelli (2008), Moench (2008), Ludvigson and Ng (2009a, 2009b), and Gilchrist et al. (2009), among others.

In this paper, we discuss methods to construct the confidence interval for structural impulse response functions (IRFs) in FAVARs when the latent factors are, in particular, estimated by using the principal component (PC) estimation method. The seminal work by Bai and Ng (2006) develops inferential theory for the coefficients in factor-augmented regression models including FAVARs by using a large dimensional factor model. They show that under certain conditions including  $\sqrt{T}/N \rightarrow 0$  as  $N, T \rightarrow \infty$ , where  $N$  and  $T$  are the cross-sectional dimension and the time dimension, respectively, the effect of the factor estimation errors on their coefficient estimates becomes negligible. This result implies that latent factors can be replaced with their PC estimate in VARs and that the asymptotic confidence interval for the coefficients can be constructed in the usual manner. Furthermore, this approach is straightforwardly applied to IRFs because they are a function of the coefficients in the FAVAR model.

On the contrary, the effect of the factor estimation errors on the coefficient estimates can be relevant when the data set involves  $N$  much smaller than  $T$  such that the condition  $\sqrt{T}/N \rightarrow 0$  is considered to be inappropriate. Therefore, bootstrap methods are a strong alternative to Bai and Ng's (2006) asymptotic confidence intervals. To that effect, Gonçalves and Perron (2014) and Shintani and Guo (2015) study the theoretical properties of residual-based bootstrap confidence intervals for the coefficients in factor-augmented regression models, showing their asymptotic validity under the more general framework of

$\sqrt{T}/N \rightarrow c$  with  $0 \leq c < \infty$ . Yet, their models are reduced form, and in this paper we extend them to models that involve structural identification schemes in VARs. In particular, we consider the recursive restriction scheme with sign restrictions as well as two major impulse response-based identifications: the short-run and the long-run restrictions considered by Stock and Watson (2005) via Monte Carlo simulations. In doing so, we focus on comparisons between the confidence intervals produced by the method à la Gonçalves and Perron (2014) and an alternative method that may be used in empirical studies. The former method, as recommended by Gonçalves and Perron (2014), implements factor estimations in the bootstrap repetitions. This procedure is called Procedure A. The alternative method does not estimate factors in the bootstrap replications, treating the original estimate as the observed processes. This procedure is called Procedure B and its algorithm is implemented more straightforwardly from the standard structural VAR. In essence, Procedure A accounts for the uncertainty associated with the factor estimation in the confidence interval, while Procedure B does not. By contrast, Procedure B may give a shorter confidence interval than Procedure A. As such, we discuss the advantages and disadvantages of both procedures.

Our findings are summarized as follows. In theory, both procedures are asymptotically valid when  $\sqrt{T}/N \rightarrow 0$  holds. However, Procedure A better mimics the factor estimation errors even when  $\sqrt{T}/N \rightarrow 0$  is not relevant. In finite samples, Procedure A provides a good coverage rate in most cases, while Procedure B tends to undercover if  $N$  is very small compared with  $T$ . On the contrary, the length of the confidence interval of Procedure B is shorter than that of Procedure A. Therefore, we conclude that given that Procedure B still provides a satisfactory coverage rate unless  $N$  is very small compared with  $T$ , it is in consideration of empirical use, although using Procedure A is safer as it correctly accounts for the effect of the factor estimation errors. Finally, the asymptotic confidence interval tends to provide erratic coverage rates, especially when the errors are not normally distributed.

The rest of the paper is organized as follows. Section 2 introduces the structural models and their reduced-form counterparts. The basic assumptions of the models are introduced as well. Section 3 discusses the assumptions and method for structural identification and asymptotic properties of the IRF estimator. In section 4, we propose bootstrap inference procedures and investigate their asymptotic properties. Section 5 assesses the finite sample properties of the suggested procedures via Monte Carlo simulations, using artificial data, along with calibrated models of US macroeconomic data. Section 6 provides concluding remarks. Finally, the appendices include technical derivations.

Throughout the paper, the following notations are used. The Euclidean norm of vector

$x$  is denoted by  $\|x\|$ . For matrices, the vector-induced norm is used. The symbols “ $\xrightarrow{p}$ ” and “ $\xrightarrow{d}$ ” represent convergence in probability under the probability measure  $P$  and convergence in distribution, respectively.  $O_p(\cdot)$  and  $o_p(\cdot)$  are the order of convergence in probability under  $P$ . We define  $P^*$  as the bootstrap probability measure, conditional on the original sample. For any bootstrap statistic  $T^*$ , we write  $T^* \xrightarrow{p^*} 0$ , in probability, or  $T^* = o_{p^*}(1)$ , in probability, when for all  $\epsilon > 0$ ,  $P^*(|T^*| > \epsilon) = o_p(1)$ . We write  $T^* = O_{p^*}(\cdot)$ , in probability, when for all  $\epsilon > 0$  there exists  $M(\epsilon) < \infty$  such that  $\lim_{N,T \rightarrow \infty} P[P^*(|T^*| > M(\epsilon)) > \epsilon] = 0$ . We also write  $T^* \xrightarrow{d^*} D$ , in probability, if, conditional on a sample with a probability that converges to one,  $T^*$  converges in distribution to  $D$  under  $P^*$ . Let  $\delta = \min \left\{ \sqrt{N}, \sqrt{T} \right\}$  and  $L$  be the standard lag operator.  $Chol(X)$  denotes the Cholesky decomposition of a positive definite matrix  $X$  returning a lower triangular matrix  $W$  with positive diagonal elements such that  $W'W = X$ . The operator  $vec(X)$  transforms an  $m \times m$  matrix  $X$  into an  $m^2 \times 1$  vector by stacking the columns and  $vech(X)$  stacks only the element on and above the main diagonal of a square matrix  $X$ .

## 2 Models and assumptions

### 2.1 Models

Consider the following factor model:

$$X_t = \mu^s + \Lambda^s F_t^s + u_t^s, \quad t = 1, \dots, T, \quad (1)$$

where  $X_t$  is an  $N \times 1$  vector of the response variables and  $N$  is the (typically large) number of equations. The large set of response variables  $X_t$  are assumed to be driven by much lower dimensional unobservable factors  $F_t^s$  ( $r \times 1$ ) with time-invariant unobservable factor loadings  $\Lambda^s = [\lambda_1^s, \dots, \lambda_N^s]'$  ( $N \times r$ ).  $u_t^s = [u_{1t}^s, \dots, u_{Nt}^s]'$  is an  $N \times 1$  vector of the idiosyncratic errors and  $\mu^s$  is an  $N \times 1$  vector of the constants. Of particular interest, the model allows the latent factors  $F_t^s$  to be fundamental or to have structural interpretations. Hence, they are designed to form the structural VAR of order  $p$  with  $r \times r$  coefficient parameters  $\Phi_j^s$  ( $j = 1, \dots, p$ ), an  $r \times 1$  constant vector  $v^s$ , and structural shocks  $e_t^s$  ( $r \times 1$ ) so that

$$F_t^s = v^s + \sum_{j=1}^p \Phi_j^s F_{t-j}^s + e_t^s, \quad (2)$$

where the order  $p$  and number of factors  $r$  are assumed to be fixed and known for simplicity and  $E(e_t^s e_t^{s'}) = I_r$ . When the variables are written without their associated  $t$  subscript, they

denote the entire matrix of observations, for example,  $X = [X_1, \dots, X_T]'$  is a  $T \times N$  matrix and  $F^s = [F_1^s, \dots, F_T^s]'$  is a  $T \times r$  matrix. The goal is to establish a valid inference procedure on the structural IRFs generated from these models. Given that invertibility is assured, models (1) and (2) can be rewritten in vector moving-average form such that

$$X_{it} = \mu^s + \lambda_i^{s'} \Psi^s(L) e_t^s + u_{it}, \quad \text{for } i = 1, \dots, N$$

where  $\Psi^s(L) \equiv \sum_{j=0}^{\infty} \Psi_j^s L^j$  with  $\Psi_0^s = I_r$  and  $\Psi^s(L) = \left[ I_r - \sum_{j=1}^p \Phi_j^s L^j \right]^{-1}$ . Let the structural IRF of the response variable  $X_{it}$  at time horizon  $h$  ( $h = 0, 1, 2, \dots$ ) be  $\varphi_{ih}$ . Then,

$$\varphi_{ih} \equiv \frac{\partial X_{it+h}}{\partial e_t^s} = \lambda_i^{s'} \Psi_h^s.$$

This class of models is called the FAVARs, introduced by Bernanke et al. (2005) and further explored by Stock and Watson (2005). Bernanke et al. (2005) include an observable structural factor as well as unobservable factors in  $F_t^s$ . Under the assumption that the observable and unobservable factors are separately identified, this extension does not add any complications to the analysis, and the results in this paper hold. Second, in many empirical applications, accounting for individual serial correlations in  $X_{it}$  is desired so that  $X_t$  can be replaced by  $A(L)X_t$  in the above models, where  $A(L) = [A_1(L), \dots, A_N(L)]'$  and  $A_i(L) = 1 + A_{i,1}L + \dots + A_{i,p_i}L^{p_i}$ . However, consistency and asymptotic normality of the coefficients  $A_{i,j}$  ( $i = 1, \dots, N$ , and  $j = 1, \dots, p_i$ ) are obtained by using the least squares of the current response variable  $X_{it}$  on the factors and the past response variables  $X_{i,t-j}$ . Hence, we use the simple models (1) and (2) throughout the paper.

While these models can be estimated by using the maximum likelihood principle with respect to the identifying assumptions, the maximum likelihood estimators for models of such a large dimension may have questionable finite sample properties. Thus, we instead consider the less computationally demanding PC estimation studied by Bai and Ng (2006). To this end, we state the estimable reduced-form representations of (1) and (2) following the conventional strategy in the VAR literature. Rewrite the models by using an  $r \times r$  invertible matrix  $B$  that captures the contemporaneous correlations among  $F_t^s$ , so that  $F_t = BF_t^s$ . It then follows that  $\Lambda = \Lambda^s B^{-1}$ ,  $\Phi_j = B\Phi_j^s B^{-1}$ ,  $e_t = Be_t^s$ , and  $v = Bv^s$ . The constant and idiosyncratic errors in (1) are not affected by the factor structure, so  $\mu = \mu^s$  and  $u_t = u_t^s$ . The reduced-form models are

$$X_t = \mu + \Lambda F_t + u_t, \tag{3}$$

$$F_t = v + \sum_{j=1}^p \Phi_j F_{t-j} + e_t. \tag{4}$$

To simplify the notation, let  $Z = [\iota, F_{(-1)}, F_{(-2)}, \dots, F_{(-p)}]$  be a  $T \times (rp + 1)$  matrix, where  $\iota$  is a  $T \times 1$  vector of ones and  $F_{(-j)} = [F_{1-j}, \dots, F_{T-j}]'$ , and let  $\Phi = [\nu, \Phi_1, \dots, \Phi_p]'$  be an  $(rp + 1) \times r$  matrix so that (4) can equivalently be written as  $F = Z\Phi + e$ . The constant terms in the models can be omitted if the data are demeaned.<sup>1</sup>

The estimation procedure is as follows. First, the factors are extracted from (3) by using the PC method to find the solution of

$$(\hat{F}, \hat{\Lambda}) = \arg \min_{\Lambda, F} \sum_{i=1}^N \sum_{t=1}^T (X_{it} - \lambda_i' F_t)^2, \quad (5)$$

with normalization  $\hat{F}'\hat{F}/T = I_r$ . Second, we estimate the VAR equation (4) by adopting the least squares approach using  $\hat{F}_t$ . As pointed out by Bai and Ng (2013), when estimating models by using the PC method, a well-known rotation problem arises, implying that the factors are not individually identified in the reduced form. Define the rotation matrix pertaining to the PC estimation as

$$H_{NT} = V_{NT}^{-1}(\hat{F}'F/T)(\Lambda'\Lambda/N), \quad (6)$$

where  $V_{NT}$  is an  $r \times r$  diagonal matrix whose main diagonal elements are the  $r$  largest eigenvalues of  $XX'/(TN)$  in descending order and whose probability limit is  $Q^{-1} = p \lim H'_{NT}$  as  $N, T \rightarrow \infty$ , given it exists.

Finally, let the reduced-form IRF of the  $i$ th response variable  $X_i$  at horizon  $h$  ( $h = 0, 1, 2, \dots$ ) be  $\psi_{ih}$ . Then,

$$\psi_{ih} \equiv \frac{\partial X_{it+h}}{\partial e_t} = \lambda_i' \Psi_h.$$

Since the statistical properties of the IRF estimates are often discussed in terms of the reduced-form IRFs, it is important to note that, as in standard structural VAR analysis, the structural IRF can be represented by its reduced-form counterpart,

$$\varphi_{ih} = \psi_{ih}B = \lambda_i' \Psi_h B, \quad (7)$$

by using the reduced-form parameters.

---

<sup>1</sup>For simplicity, the theoretical derivations in this paper do not include the constant term, assuming that the data are demeaned. In practice, when the model does not include a constant term and demeaned data are used, the researcher should make sure that the residuals are demeaned in the bootstrap procedures. See section 4.1.

## 2.2 Assumptions

We discuss the standard regularity conditions for the remainder of the analysis. First, let the data-generating processes above be defined on a probability space  $(\Omega, F, P)$  and let  $M < \infty$  be a generic constant.

### Assumption A.

- (a) The loadings  $\lambda_i$  are either deterministic such that  $\|\lambda_i\| \leq M$  or stochastic such that  $E\|\lambda_i\|^4 \leq M$ . In either case,  $\Lambda'\Lambda/N \rightarrow \Sigma_\Lambda$ , where  $\Sigma_\Lambda$  is an  $r \times r$  positive definite matrix  $\Sigma_\Lambda$ .
- (b) The eigenvalues of the  $r \times r$  matrix  $\Sigma_\Lambda \Sigma_F$  are distinct.
- (c)  $\lambda_i$ ,  $u_{js}$  and  $e_t$  are mutually independent for all  $(i, j, s, t)$ .

### Assumption B.

- (a)  $E(u_{it}) = 0$  and  $E|u_{it}|^8 \leq M$  for all  $(i, t)$ .
- (b)  $N^{-1} \sum_{i=1}^N \sum_{j=1}^N |\tau_{ij}|^2 \leq M$ , where  $\tau_{ij} = E(u_{it}u_{jt})$ .  $T^{-1} \sum_{s=1}^T \sum_{t=1}^T |\gamma_{ts}|^2 \leq M$ , where  $\gamma_{ts} = E(N^{-1} \sum_{i=1}^N u_{it}u_{is})$ .
- (c) For every  $(t, s)$ ,  $E \left| N^{-1/2} \sum_{i=1}^N [u_{is}u_{it} - E(u_{is}u_{it})] \right|^4 \leq M$ .
- (d) For each  $i$ ,  $T^{-1/2} \sum_{t=1}^T F_t u_{it} \xrightarrow{d} N(0, \Theta_i)$ , where  $\Theta_i \equiv Var \left( T^{-1/2} \sum_{t=1}^T F_t u_{it} \right)$ .

### Assumption C.

- (a)  $E(e_t) = 0$ ,  $E(e_t e_t') = \Sigma_e$  an  $r \times r$  positive definite matrix, and  $E(e_t e_s') = 0_{r \times r}$  for all  $t \neq s$ .
- (b)  $E|e_{lt}|^4 \leq M$  for all  $l = 1, \dots, r$ , and  $t$ .
- (c) For  $h = 1, 2, \dots$ ,  $T^{-1/2} \sum_{t=p+1}^T vec(Z_t e_t') \xrightarrow{d} N(0, \Sigma_Z \otimes \Sigma_e)$  with  $\Sigma_Z \equiv p \lim_{T \rightarrow \infty} Z'Z/T$ .
- (d)  $T^{-1/2} \sum_{t=1}^T (e_t e_t' - \Sigma_e) \xrightarrow{d} N(0, \Sigma_e \otimes \Sigma_e)$ .

- (e) The roots of  $\det(I_r - \Phi_1 z - \Phi_2 z^2 - \dots - \Phi_p z^p) = 0$  lie outside the unit circle.
- (f) The  $r \times r$  matrix  $B$  has full rank.

Assumptions A and B have been discussed in Bai and Ng (2006). In particular, Assumption A(b) guarantees the uniqueness of the limit of the factor rotation matrix. Assumption A(c) imposes the independence of the factors, factor loadings, and idiosyncratic errors as in Bai and Ng (2006), and it greatly simplifies the moment conditions without losing much substance in our context. Assumption B(b) allows for idiosyncratic errors to have weak serial and cross-sectional dependence. Assumption C consists of the basic assumptions in the VAR literature. It enforces a stable system that is estimable by using the least squares method. Assumption C(a) imposes a white noise property with a fixed covariance matrix on  $\{e_t\}$  to obtain a structural identification using up to the second moments of the residuals. Finally, these assumptions are on the parameters and variables  $\{F_t, \Lambda, \Phi, u_t, e_t\}$  in the reduced-form representation rather than the structural entities  $\{F_t^s, \Lambda^s, \Phi^s, u_t^s, e_t^s\}$ . This is mainly for notational simplicity in the technical proofs; however, researchers may need to validate these assumptions for the structural entities, in which case, all of these assumptions are stated by using the structural counterparts. For example, Assumption C(a) can be replaced by the same assumption on  $e_t^s$  with a small notational change from  $\Sigma_e$  to  $B^{-1}\Sigma_e B'^{-1}$ . This assumption holds since  $e_t^s = B^{-1}e_t$  and  $B$  is a fixed full rank matrix.

### 3 Structural inference

Once the reduced-form models have been estimated, structural parameter estimates can be obtained by using the contemporaneous coefficient matrix  $\hat{B}$ . Several identifying methods are applied in empirical research; however, we particularly consider the popular recursive scheme by imposing the following assumption.

**Assumption D.**  $Q^{-1'}B$  is an upper or lower triangular matrix and the signs of its diagonal elements are known.

A glitch in the FAVAR context arises when  $\hat{B}$  is obtained through the Cholesky decomposition. This is because the recursive restrictions must be imposed not on matrix  $B$  itself but on its rotation  $Q^{-1'}B$  as the latter is the asymptotic counterpart of  $\hat{B}$ . This is the case of the B-estimator of Lütkepohl (2005) and the structural IRF is only locally identified. In order

to obtain the unique estimate, we fix the sign of  $\hat{\varphi}_{ih}$  if the sign of  $\hat{\varphi}_{i0}$  is not desired. This restriction includes the limit of the rotation matrix  $Q$  and it is not straightforward to interpret restrictions on such non-structural entities. Hence, we further break down Assumption D into the following set of sufficient conditions to ascertain its feasibility.<sup>2</sup>

**Assumption D'.** The following three restrictions imply Assumption D:

- (a)  $\Sigma_F$  is diagonal;
- (b)  $\Sigma_\Lambda$  is diagonal;
- (c)  $B$  is an upper triangular matrix and the signs of  $\lambda'_i B = \varphi_{i0}$  are known.

The first two parts of Assumption D' imply that the model involves orthogonal factors and loadings in its reduced form, which are statistical assumptions. Given these two statistical restrictions, we are able to impose a recursive structure directly on  $B$  as in a conventional structural VAR. Finally, the signs of the elements of  $\lambda'_i B$  can be deduced from the underlying economic model. The identification is implemented as follows:

**ID (recursive identification with sign restriction):**

1. Obtain  $\hat{B}$  such that  $\hat{B} = Chol(\hat{e}'\hat{e}/T)$ .
2. Change the signs of  $\hat{\varphi}_{ih}$  ( $h = 1, \dots$ ), if the signs of  $\hat{\varphi}_{i0}$  are not desired.

Under these identifying assumptions and the method, we obtain the following theorems about the consistency and asymptotic normality of the structural parameters and the structural IRF.

**Theorem 1** (*Consistency of the structural parameters*) Under Assumptions A, B, C, and D,  $\hat{\lambda}_i^s - \lambda_i^s \xrightarrow{p} 0$ ,  $\hat{\Phi}^s - \Phi^s \xrightarrow{p} 0$ , and  $\hat{\varphi}_{ih} - \varphi_{ih} \xrightarrow{p} 0$ , for all  $i$  as  $N, T \rightarrow \infty$ .

---

<sup>2</sup>Lemma 3 in Appendix A provides a proof that Assumption D' implies Assumption D.

**Theorem 2** (*Asymptotic distribution of structural IRFs*) Under Assumptions A, B, C, and D,

$$\sqrt{T}(\hat{\varphi}_{ih} - \varphi_{ih}) \xrightarrow{d} N(0, \Omega_{\varphi_{ih}}),$$

$\forall i$  as  $T, N \rightarrow \infty$  and  $\sqrt{T}/N \rightarrow 0$  provided  $\partial\varphi_{ih}/\partial\theta \neq 0$ , where  $\theta = [\lambda_i, \text{vec}(\Phi)', \text{vech}(B)']'$ ,

$$\Omega_{\varphi_{ih}} = \frac{\partial\varphi'_{ih}}{\partial\theta'} \Omega_{\theta_i} \frac{\partial\varphi_{ih}}{\partial\theta},$$

and  $\Omega_{\theta_i} = \text{diag}(\Omega_{\lambda_i}, \Omega_{\Phi}, \Omega_B)$  with  $\Omega_{\lambda_i} \equiv Q^{-1'}\Theta_i Q^{-1}$ ,

$$\begin{aligned} \Omega_{\Phi} &\equiv [(I_p \otimes Q^{-1'})\Sigma_Z(I_p \otimes Q^{-1})]^{-1} \\ &\times [(I_p \otimes Q^{-1'})\Sigma_Z(I_p \otimes Q^{-1})] \otimes [Q^{-1'}\Sigma_e Q^{-1}] \\ &\times [(I_p \otimes Q^{-1'})\Sigma_Z(I_p \otimes Q^{-1})]^{-1'}, \end{aligned}$$

and  $\Omega_B$  is defined in Appendix A.

There are two remarks on these results. First, despite the implication that the structural IRFs only involve structural parameters, when we consider the distribution of the IRF estimate, we present the expression in terms of the reduced-form parameters as in Lütkepohl (1990, 2005), so that the asymptotic covariance estimates are easily constructed. Second, more importantly, the asymptotic approximation is reasonable when  $\sqrt{T}/N \rightarrow 0$  is relevant. However, if  $N$  is smaller than  $T$  and  $\sqrt{T}/N \rightarrow c$  ( $c > \infty$ ) is more appropriate, then the parameter estimates suffer from asymptotic bias, as suggested by Ludvigson and Ng (2009b) and Gonçalves and Perron (2014).

**Remark 1** *Stock and Watson (2005) and other empirical studies also consider valid impulse response-based identification schemes such as short-run and long-run restrictions. The finite sample properties when using these identifying schemes are discussed via the Monte Carlo simulations in section 5.*

## 4 Bootstrap inference

### 4.1 Procedures

This section outlines the residual-based bootstrap algorithms used to construct the confidence interval for the IRFs. We consider the independent and identically distributed (i.i.d.)

bootstrap for the VAR shocks in order to make use of the white noise property of  $\{e_t\}$ . This treatment is not as trivial regarding the idiosyncratic errors in the factor model. Here, we also illustrate the i.i.d. resampling for the homoskedastic idiosyncratic errors  $\{u_{it}\}$ ; however, the effects of the potential heteroskedasticity in  $\{u_{it}\}$  on the IRF estimates are accounted for by employing the wild bootstrap procedure as suggested by Gonçalves and Perron (2014). The same can be said for the serial correlations in the idiosyncratic errors by employing a relevant bootstrap method to account for the dependence in the original idiosyncratic errors.

In this paper, we compare two alternative algorithms that are straightforwardly implemented in empirical studies. The first method, as recommended by Gonçalves and Perron (2014), implements factor estimation in the bootstrapping repetitions. This is called Procedure A. The second method does not estimate factors in the bootstrap repetitions, treating the original factor estimate as the observed processes. Therefore, the implementation is no more involved than standard small-scale structural VAR models. This is called Procedure B. We first outline Procedure A as follows.

#### **Procedure A: Bootstrap with factor estimation**

1. Estimate the model by using the PC procedure, obtaining the parameter estimates  $\hat{\Lambda}$ ,  $\hat{\Phi}$ , and  $\hat{B}$  as well as the residuals  $\hat{u}_t$  and  $\hat{e}_t$ . Obtain the structural IRF estimate  $\hat{\varphi}_i$ .
2. Make sure that the residuals  $\hat{u}_{it}$  and  $\hat{e}_t$  are demeaned. Resample the residuals  $\hat{e}_t$  with replacements and label them  $e_t^*$ . Generate the bootstrapped sample  $F_t^*$  by using  $F_t^* = \hat{v} + \sum_{j=1}^p \hat{\Phi}_j F_{t-j}^* + e_t^*$ . In addition, resample the residuals  $\hat{u}_t$  with replacements and label them  $u_t^*$ .<sup>3</sup> Generate the bootstrapped response variables  $X^*$  by using  $X_t^* = \hat{\mu} + \hat{\Lambda} F_t^* + u_t^*$ .<sup>4</sup>
3. By using the bootstrapped response variables  $X_t^*$ , estimate  $(\hat{F}^*, \hat{\Lambda}^*)$  following the first step of the PC procedure. Then, estimate the VAR equation of  $\hat{F}_t^*$  to obtain the bootstrapped estimates  $\hat{\Phi}^*$  and  $\hat{B}^*$  using the second step of the PC procedure. This yields the bootstrap IRF estimates  $\hat{\varphi}_i^*$ .
4. Repeat steps 2 and 3  $R$  times.

---

<sup>3</sup>Here, one can implement another bootstrap scheme that mimics the properties of the idiosyncratic errors.

<sup>4</sup>The bias-correction method discussed by Kilian (1998) is applied. The bias is estimated by taking the average of  $\hat{\Phi}_j^* - H^* \hat{\Phi}_j H^{*-1}$  in another bootstrap loop in Procedure A and by  $\hat{\Phi}_j^{**} - \hat{\Phi}_j$  in Procedure B.

5. Store the recentered statistic  $s \equiv \hat{\varphi}_{ih}^* - \hat{\varphi}_{ih}$ . Sort the statistics and select the  $100 \cdot \alpha / 2^{th}$  and  $100 \cdot (1 - \alpha / 2)^{th}$  percentiles ( $s^{(\alpha/2)}, s^{(1-\alpha/2)}$ ). The resulting  $100 \cdot (1 - \alpha)\%$  confidence interval for  $\varphi_{ih}$  is  $[\hat{\varphi}_{ih} - s^{(1-\alpha)}, \hat{\varphi}_{ih} - s^{(\alpha)}]$  for  $h = 0, 1, \dots$

In step 2, the bootstrap sample  $X_t^*$  shares the same data-generating process as the original sample  $X_t$ . In step 3, the bootstrap IRF estimates involve the same identification and estimation methods as the original IRF estimate, suggesting that the dispersions of all the bootstrap estimates can mimic the sampling distribution of the original estimates.

### **Procedure B: Bootstrap without factor estimation**

Procedure B is the same as Procedure A except for a change in step 3, which is formalized as follows:

3. By using the bootstrapped observations  $X_t^*$  and factors  $F_t^*$ , estimate  $\hat{\Lambda}^{**}$ ,  $\hat{\Phi}^{**}$  and  $\hat{B}^{**}$ . This yields the bootstrap IRF estimates  $\hat{\varphi}_i^{**}$ .

Procedure B is a natural and simple extension of the methods conducted in standard structural VAR analyses since it generates the VAR variables as if they were observable. Hence, it does not involve factor estimation in the procedure, and  $\hat{F}_t^*$  is not generated. Finally, the bootstrap interval specified in step 5 is known as Hall's (1992) percentile intervals. One can alternatively use what is commonly called Efron's percentile method by storing  $s \equiv \hat{\varphi}_{ih}^*$  and constructing  $[s^{(\alpha)}, s^{(1-\alpha)}]$ ; however, this method is not asymptotically centered (see Efron and Tibshirani, 1994). It is found that the IRFs are likely to produce non-centered distributions because of structural identifications and factor estimation errors. Another popular choice is the percentile- $t$  interval. However, it is observed that the percentile- $t$  is inaccurate for IRF estimates, especially over long horizons when the sample size is small (Kilian, 1999).

## **4.2 Asymptotic validity**

This section studies the asymptotic validity of the proposed procedures. The asymptotic validity of Procedures A and B is given in Theorems 3 and 4, respectively, under  $\sqrt{T}/N \rightarrow 0$ , with remarks on higher-order correctness following. For Procedure A, we obtain the following theorem.

**Theorem 3** (*Asymptotic validity of Procedure A*) Under Assumptions A, B, C, and D,

$$\sup_{x \in \mathbf{R}} |P^* [(\hat{\varphi}_{ih}^* - \hat{\varphi}_{ih}) \leq x] - P [(\hat{\varphi}_{ih} - \varphi_{ih}) \leq x]| \xrightarrow{p} 0,$$

for all  $i$  as  $N, T \rightarrow \infty$ , and  $\sqrt{T}/N \rightarrow 0$ .

This result is of the first order; therefore, to better understand the finite sample properties of inference, higher-order terms in the estimation errors must be investigated. For simplicity, let us consider the case of no serial correlations in the idiosyncratic errors  $\{u_{it}\}$ , i.e.  $\gamma_{ts} = 0$  for all  $t \neq s$  in Assumption B(b). The errors in the original structural parameter estimation can be expanded into three components: errors pertaining to the contemporaneous coefficient matrix  $\hat{B}$ , factor estimation errors, and combinations of these errors. Let us now take the structural IRF at horizon 0 as an example. The expansion of the original estimate is

$$\begin{aligned} \sqrt{T}(\hat{\varphi}_{i0} - \varphi_{i0}) &= T^{-1/2} B' H_{NT}' H_{NT} F' u_i + T^{1/2} \varepsilon' H_{NT}'^{-1} \lambda_i \\ &\quad + \underbrace{T^{-1/2} \varepsilon' H_{NT} F' u_i}_{\text{(I): errors in } \hat{B}} \\ &\quad + \underbrace{T^{-1/2} B' H_{NT}' \hat{F}' (F - \hat{F} H_{NT}'^{-1}) \lambda_i + T^{-1/2} B' H_{NT}' (\hat{F} - F H_{NT}')' u_i}_{\text{(II): factor estimation errors}} \\ &\quad + \underbrace{T^{-1/2} \varepsilon' \hat{F}' (F - \hat{F} H_{NT}'^{-1}) \lambda_i + T^{-1/2} \varepsilon' (\hat{F} - F H_{NT}')' u_i}_{\text{(III): (I) and (II)}}. \end{aligned} \quad (8)$$

where  $\varepsilon = \hat{B} - H_{NT} B$ . In the original estimate, the terms in (I), (II), and (III) are of  $o_p(1)$ ,  $O_p(\sqrt{T}/\delta^2)$ , and  $o_p(1)$ , respectively. Hence, the term of  $O_p(\sqrt{T}/\delta^2) = o_p(1)$  under  $\sqrt{T}/N \rightarrow 0$ . When we use Procedure A, the bootstrap parameter estimates take the same form in the bootstrap space, so that

$$\begin{aligned} \sqrt{T}(\hat{\varphi}_{i0}^* - \hat{\varphi}_{i0}) &= T^{-1/2} \hat{B}' H_{NT}'^* H_{NT}^* F^{*'} u_i^* + T^{1/2} \varepsilon^{*'} H_{NT}'^{*-1} \hat{\lambda}_i \\ &\quad + \underbrace{T^{-1/2} \varepsilon^{*'} H_{NT}^* F^{*'} u_i^*}_{\text{(I): errors in } \hat{S}^*} \\ &\quad + \underbrace{T^{-1/2} \hat{B}' H_{NT}'^* \hat{F}^{*'} (F^* - \hat{F}^* H_{NT}^{*'}{}^{-1}) \hat{\lambda}_i + T^{-1/2} \hat{B}' H_{NT}'^* (\hat{F}^* - F^* H_{NT}^{*'})' u_i^*}_{\text{(II): factor estimation errors}} \\ &\quad + \underbrace{T^{-1/2} \varepsilon^{*'} \hat{F}^{*'} (F^* - \hat{F}^* H_{NT}^{*'}{}^{-1}) \hat{\lambda}_i + T^{-1/2} \varepsilon^{*'} (\hat{F}^* - F^* H_{NT}^{*'})' u_i^*}_{\text{(III): (I) and (II)}}. \end{aligned} \quad (9)$$

with

$$H_{NT}^* = \hat{V}^{-1} (\hat{F}^{*'} F^* / T) (\hat{\Lambda}' \hat{\Lambda} / N) \quad (10)$$

where  $\hat{V}$  is a diagonal matrix whose elements are eigenvalues of  $X^*X^{*'}/(TN)$  in descending order. The validity follows under the stated conditions, which guarantee that all the terms in (9) are of the same stochastic order under  $P^*$  as those in (8) under  $P$ . Hence, (I) and (III) disappear as  $N, T \rightarrow \infty$ , and so does (II) with the additional condition  $\sqrt{T}/N \rightarrow 0$ . We now prove the validity of Procedure B.

**Theorem 4** (*Asymptotic validity of Procedure B*) *Under Assumptions A, B, C, and D,*

$$\sup_{x \in \mathbf{R}} \left| P^* [(\hat{\varphi}_{i,h}^{**} - \hat{\varphi}_{i,h}) \leq x] - P [(\hat{\varphi}_{i,h} - \varphi_{i,h}) \leq x] \right| \xrightarrow{P} 0,$$

for all  $i$  as  $N, T \rightarrow \infty$ ,  $\sqrt{T}/N \rightarrow 0$ .

In Procedure B, the bootstrap estimate of the structural IRF at the contemporaneous timing of the shock is expanded in the bootstrap space as follows:

$$\begin{aligned} \sqrt{T}(\hat{\varphi}_{i0}^{**} - \hat{\varphi}_{i0}) &\equiv T^{-1/2} \hat{B}' F^{*'} u_i^* + T^{1/2} \varepsilon^{**'} \hat{\lambda}_i \\ &\quad + \underbrace{T^{-1/2} \varepsilon^{**'} F^{*'} u_i^*}_{\text{(I): errors in } \hat{B}^{**}}, \end{aligned} \tag{11}$$

with  $\varepsilon^{**} = \hat{B}^{**} - B$ . The higher-order terms associated with factor estimation errors (II) and (III) in (9) do not appear in (11). Hence, we expect the confidence intervals constructed by Procedure B to be generally tighter than those constructed by Procedure A because of the factor estimation errors. On the contrary, the interval given by Procedure B captures sampling variations less accurately than Procedure A does, especially when  $N$  is significantly smaller than  $T$  ( $\sqrt{T}/N \rightarrow 0$  does not hold), because the terms in (II) that are not present in Procedure B are relevant. When the errors in the contemporaneous matrix estimate  $\varepsilon$  are not small, the terms in (I) and (III) can play a significant role. This leads to coverage errors over short horizons that depend on the structural identification scheme; however, the effect of  $\varepsilon$  does diminish over long horizons. Finally, the normal approximation neither accounts for the factor estimation errors (II) and (III) nor captures the effect of (I).

## 5 Finite sample properties

### 5.1 Monte Carlo simulations with artificial data

This section reports the Monte Carlo simulation results to assess the finite sample properties of the bootstrap procedures. We consider a simple VAR(1) model so that the response

variables  $x_{i,t}$  are generated by  $x_{i,t} = \lambda_i f_t + u_{i,t}$ , with the factors ( $f_t : r \times 1$ ) evolving such that  $f_t = \Phi f_{t-1} + e_t$  ( $i = 1, \dots, N$  and  $t = 1, \dots, T$ ), where  $\lambda_i = [\lambda_{i,1}, \dots, \lambda_{i,r}]'$  is an  $r \times 1$  vector,  $\Phi$  is an  $r \times r$  matrix, and  $e_t = B e_t^s$ , where  $e_t^s = [e_{t,1}^s, \dots, e_{t,r}^s]'$ . We set  $r = 2$  and the particular choice of the parameter values is

$$\Phi = \begin{bmatrix} 0.4 & 0 \\ 0 & 0.4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix},$$

unless otherwise stated. The factor loadings  $\lambda_{i,j}$  ( $j = 1, 2$ ) are generated from i.i.d. standard normal distributions.

We consider the following three patterns of the error terms. (1)  $e_{t,j}^s$  and  $u_{i,t}$  follow i.i.d. standard normal distributions and this case is called ‘‘Gaussian,’’ (2)  $e_{t,j}^s$  and  $u_{i,t}$  follow i.i.d.-centered chi-square distributions with one degree of freedom with unit variance and this case is called ‘‘Chi,’’ and (3)  $e_{t,j}^s$  follows i.i.d. standard normal distributions and  $u_{i,t}$  follows the independent normal distribution of random variance generated by the uniform distribution  $U[0.5, 1.5]$  and this case is called ‘‘Hetero.’’

We report the result for the IRF of the first response variable to a structural shock of the first factor. Because the estimation errors in  $\hat{B}$  can affect the results, as we conjectured in section 4.2, we study the following three popular structural identification schemes. The first scheme is the recursive identification described in Section 3. This is called ‘‘ID1.’’ The second structural identification scheme is the short-run restriction and this is called ‘‘ID2.’’ This method is based on a different assumption from the recursive scheme and assumes that the IRF of the second response variable to the structural shock of the first factor at time  $h = 0$  is zero. Without loss of generality, take the IRFs of the first  $r$  response variables to the  $r$  structural shocks. Then, we have an  $r \times r$  matrix of the structural IRFs  $\varphi_{1:r,h}$  and the reduced-form IRFs  $\psi_{1:r,h}$  at time horizon  $h$ , respectively. The matrix consists of the first  $r$  response variables stacked in rows and  $r$  shocks in columns. Following Stock and Watson (2005), the  $B$  matrix is estimated by

$$\hat{B} = \hat{\psi}_{1:r,0}^{-1} \hat{\varphi}_{1:r,0},$$

where

$$\hat{\varphi}_{1:r,0} = \text{Chol} \left[ \hat{\psi}_{1:r,0} (\hat{e}' \hat{e} / T) \hat{\psi}_{1:r,0}' \right],$$

$\hat{\psi}_{1:r,0} = [\hat{\lambda}_1, \dots, \hat{\lambda}_r]' \equiv \hat{\Lambda}_{1:r}$ , and  $\hat{e}$  is a  $T \times r$  matrix of the reduced-form VAR residuals. In addition, we consider the long-run restriction based on the identifying assumption that the

same IRF at the infinite horizon is zero. This is called “ID3.” In practice, the  $B$  matrix is obtained by

$$\hat{B} = \hat{\psi}_{1:r,\infty}^{-1} \hat{\varphi}_{1:r,\infty},$$

where

$$\hat{\varphi}_{1:r,\infty} = \text{Chol} \left[ \hat{\psi}_{1:r,\infty} (\hat{e}' \hat{e} / T) \hat{\psi}_{1:r,\infty}' \right],$$

and  $\hat{\psi}_{1:r,\infty} = \hat{\Lambda}_{1:r} \left[ I_r - \hat{\Phi} \right]^{-1}$ . In our data-generating process ( $r = 2$ ), these identifying assumptions are achieved by additionally imposing zero on the (2,1) element of the  $N \times 2$  factor loadings matrix  $\Lambda = [\lambda_1, \dots, \lambda_N]'$ .<sup>5</sup>

Since the effect of the sample size on the inference results is of major interest, we compare the results of the four  $(T, N)$  combinations  $(T, N) = (40, 50)$ ,  $(40, 200)$ ,  $(120, 50)$ , and  $(120, 200)$ . The experiment is based on 3,000 replications with 399 bootstrap repetitions. The results for equal-sided confidence intervals of the 95% nominal level are reported. By default, the bias correction of Kilian (1998) is applied where the bias for  $\hat{\Phi}$  is estimated by using another 300 bootstrap replications. The coverage rates and median of the lengths of the confidence intervals for the structural IRFs up to five periods are reported. As a comparison, the results obtained by using the asymptotic normal approximations are reported as well (denoted by “Normal”).

Tables 1 and 2 show the results of Procedures A, B, and the asymptotic approximations under “Gaussian” and “Chi,” respectively. In Table 1, we observe that the coverage rate of Procedure A is close to the nominal level of 95% in most cases. We see errors in the coverage rate at  $h = 0$  when ID2 or ID3 is used; however, the coverage rate improves as the sample size increases. Procedure B also reports a satisfactory coverage rate, but it tends to undercover in many cases compared with Procedure A because the interval of Procedure B does not account for the factor estimation errors. In theory, the undercoverage of Procedure B would be most distinct in the case of a larger  $T$  and a smaller  $N$ , where the condition  $\sqrt{T}/N \rightarrow 0$  is less appropriate. This is most clearly evidenced when ID2 or ID3 is used or over longer horizons in ID1. On the contrary, the length of Procedure B is always shorter than that of Procedure A as our theory predicts. Table 2 shows that the findings of both bootstrap procedures are robust even when the errors follow the chi-square distribution. Finally, the asymptotic normal interval tends to show the coverage errors as the time horizon increases (see Tables 1 and 2). Further, the normal approximation becomes exacerbated when the errors are not

---

<sup>5</sup>More precisely, by doing this, the  $2 \times 2$  matrices of the short-run structural IRF  $\varphi_{1:2,0} = \Lambda_{1:2} B$  and the long-run structural IRF  $\varphi_{1:2,\infty} = \Lambda_{1:2} [I_2 - \Phi]^{-1} B$  are upper triangular matrices.

normally distributed (see Table 2). Table 3 reports the results under “Hetero.” Here, we use the wild bootstrap of Gonçalves and Perron (2014) by modifying step 2 of the algorithms in section 4.1 such that  $e_{k,t}^* = \hat{e}_{k,t}\eta_{k,t}^e$  and  $u_{i,t}^* = \hat{u}_{i,t}\eta_{i,t}^u$ , where  $\eta_{k,t}^e$  and  $\eta_{i,t}^u$  are independent draws from the standard normal distribution at all  $k$ ,  $i$ , and  $t$ . The results in Table 3 are similar to those in Tables 1 and 2, although when ID2 or ID3 is used, the coverage rate at  $h = 0$  is unstable. In Table 4, we further investigate the cases of small  $N$  to assess how differently the two bootstrap procedures behave. In particular, we set  $(T, N) = (120, 10)$ ,  $(120, 30)$ ,  $(240, 10)$ , and  $(240, 30)$ . We focus on the Gaussian case. We now observe more significant undercoverage in Procedure B, while Procedure A can mitigate the problem.

## 5.2 Monte Carlo simulation using empirical data

Finally, we present an empirical experiment to ascertain the robustness of the obtained results by using actual economic data. To this end, we use 129 monthly US macroeconomic time series investigated by McCracken and Ng (2015).<sup>6</sup> Although the data span the period from January 1960 to December 2014,<sup>7</sup> we use the data up to December 2007 to exclude the financial crisis period in which some series exhibit explosive behaviors.<sup>8</sup> All series are transformed into stationary processes following them, and outliers are removed by using the method proposed by Yamamoto (2015).<sup>9</sup> In addition, the data are demeaned and standardized to have unit standard deviations. We use a model with two factors and the lag order four, although slight variations in the number of factors and lag order do not affect the qualitative results. We also find that the first factor is closely related to real economic activity measures (e.g., production) and that the second factor has a stronger correlation with the price variables. This finding is consistent with those of Sargent and Sims (1977) and Stock and Watson (2005). Hence, for identification, we assume that the producer price index is contemporaneously affected by the second factor but not by the first factor. We select the order of the VAR to be four. The models are then identical to those described in the previous subsection, except with a higher lag order.

---

<sup>6</sup>These data can be downloaded from the author’s website. The original data consist of 133 series; however, four series (real manufacturing and trade industries sales, new orders for consumer goods, new orders for nondefense capital goods, and consumer sentiment index) are not available for the full sample and thus they are omitted from our data set.

<sup>7</sup>The original data set starts from January 1959, but five series of new private housing permits are available from January 1960; hence, we use all other series from January 1960 as well.

<sup>8</sup>If we include the financial crisis period, the stationary assumption of the VAR model is violated.

<sup>9</sup>This method corrects outliers that exceed five times the standard error estimate in the filtered individual series.

The aim here is to evaluate the coverage properties of Procedures A and B. However, the coverage probabilities of the confidence intervals constructed from the actual data cannot be calculated. Hence, we use the following calibration experiment in order to replicate an approximation of the actual data-generating process.

1. Estimate the model and obtain the coefficient estimates and residuals.
2. Generate the quasi-observations from the calibrated model with the error terms re-sampled from  $\{\hat{e}_t\}$  and  $\{\hat{u}_t\}$  with replacement. The  $\{\hat{e}_t\}$  are orthonormalized by  $e_t^s = \hat{e}_t \hat{\Sigma}_e^{-1/2}$ , where  $\hat{\Sigma}_e^{-1/2}$  is the Cholesky-decomposed covariance matrix of  $\hat{e}_t$ . This allows  $e_t^s$  to be interpreted as a structural innovation.
3. By using each generated data set, construct 95% confidence intervals for the IRFs using the proposed bootstrap procedure with 599 repetitions and see if the true (calibrated) IRFs are included in the estimated interval. In doing so, we allow for heteroskedasticity in  $\{\hat{u}_t\}$  and use the wild bootstrap.
4. Repeat steps 2 and 3 3,000 times to evaluate the coverage rates.

We consider the IRFs of inflation (CPI: all items), short-term interest rate (3-month Treasury bill), a production index (Industrial Production Index), and the unemployment rate (Civilian Unemployment Rate) to the first factor shock. Table 5-a provides the coverage rates for the IRFs for 8 periods using Procedures A and B and Table 5-b shows the median length of the confidence intervals over the replications. Procedure A yields coverage rates very close to the 95% nominal level for all four variables. Therefore, the good finite sample properties of this bootstrap procedure are confirmed by this calibrated experiment. The coverage rates of Procedure B tend to be below the nominal level, as the theory suggests. Hence, the results are also consistent with the theoretical findings in this empirical example.

## 6 Conclusions

This paper discussed how to design residual-based bootstrap methods for the confidence interval of structural IRFs in FAVARs. The methods are based on the bootstrap procedure for factor-augmented models suggested by Gonçalves and Perron (2014) and are extended into the context of structural FAVARs. We compared procedures with and without factor estimation in the bootstrap repetitions. In theory, both procedures are asymptotically valid

under a condition  $\sqrt{T}/N \rightarrow 0$ , where  $N$  and  $T$  are the cross-sectional dimension and the time dimension, respectively. Even when  $\sqrt{T}/N \rightarrow 0$  is irrelevant, Procedure A still accounts for the effect of the factor estimation errors on the IRF estimate and achieves good coverage rates in most cases. On the contrary, Procedure B is invalid in such cases and tends to undercover if  $N$  is much smaller than  $T$ . However, Procedure B is implemented more straightforwardly from the standard structural VARs and the length of the confidence interval is shorter than that of Procedure A in finite samples. Given that Procedure B still gives a satisfactory coverage rate unless  $N$  is very small, it remains in consideration of empirical use, although using Procedure A is safer as it correctly accounts for the effect of the factor estimation errors.

## Appendix A: Proof of Theorems

In this appendix, the subscript  $NT$  for the PC rotation matrix  $H_{NT}$  is suppressed and it is denoted by  $H$ .

**Lemma 1.** Under Assumptions A, B, and C, the following hold for  $j = 0, 1, \dots, p$ .

- (a)  $T^{-1} \sum_{t=j+1}^T \left\| \hat{F}_t - HF_t \right\| \left\| \hat{F}_{t-j} - HF_{t-j} \right\| = O_p(\delta^{-2})$ ,
- (b)  $T^{-1} \sum_{t=j+1}^T (\hat{F}_t - HF_t)' F_{t-j} = O_p(\delta^{-2})$ ,
- (c)  $T^{-1} \sum_{t=j+1}^T (\hat{F}_{t-j} - HF_{t-j})' F_t = O_p(\delta^{-2})$ ,
- (d)  $T^{-1} \sum_{t=j+1}^T (\hat{F}_t - HF_t)' \hat{F}_{t-j} = O_p(\delta^{-2})$ ,
- (e)  $T^{-1} \sum_{t=j+1}^T (\hat{F}_{t-j} - HF_{t-j})' \hat{F}_t = O_p(\delta^{-2})$ ,
- (f)  $T^{-1} \sum_{t=j+1}^T (\hat{F}_{t-j} - HF_{t-j})' e_t = O_p(\delta^{-2})$ ,
- (g)  $T^{-1} \sum_{t=j+1}^T (\hat{F}_t - HF_t)' u_t = O_p(\delta^{-2})$ .

**Proof of Lemma 1:** Part (a) is a direct consequence of Lemma A1(a) of Bai and Ng (2006). For  $j = 1, \dots, p$ , the Cauchy-Schwarz inequality yields

$$\sum_{t=j+1}^T \left\| \hat{F}_t - HF_t \right\| \left\| \hat{F}_{t-j} - HF_{t-j} \right\| \leq \sqrt{\sum_{t=j+1}^T \left\| \hat{F}_t - HF_t \right\|^2} \sqrt{\sum_{t=j+1}^T \left\| \hat{F}_{t-j} - HF_{t-j} \right\|^2},$$

but this is  $O_p(T\delta^2)$ . For parts (b) and (c), Lemma A1(b) of Bai and Ng (2006) is also directly applied to get

$$\begin{aligned} T^{-1} \sum_{t=j+1}^T (\hat{F}_t - HF_t)' F_{t-j} &= O_p(\delta^{-2}) \text{ for } j = 0, 1, \dots, p, \\ T^{-1} \sum_{t=j+1}^T (\hat{F}_{t-j} - HF_{t-j})' F_t &= O_p(\delta^{-2}) \text{ for } j = 1, \dots, p. \end{aligned}$$

For parts (d) and (e), if we apply parts (a), (b), and (c) to the equations

$$\begin{aligned} \sum_{t=j+1}^T (\hat{F}_t - HF_t)' \hat{F}_{t-j} &= \sum_{t=j+1}^T (\hat{F}_t - HF_t)' F_{t-j} - \sum_{t=j+1}^T (\hat{F}_t - HF_t) (\hat{F}_{t-j} - HF_{t-j}), \\ \sum_{t=j+1}^T (\hat{F}_{t-j} - HF_{t-j})' \hat{F}_t &= \sum_{t=j+1}^T (\hat{F}_{t-j} - HF_{t-j})' F_t - \sum_{t=j+1}^T (\hat{F}_{t-j} - HF_{t-j}) (\hat{F}_t - HF_t), \end{aligned}$$

we obtain the results. Part (f) directly follows from Lemma A1 (iv) of Bai and Ng (2006). Part (g) is the same as Bai's (2003) Lemma B1, which is shown under weaker assumptions than ours.

**Lemma 2.** Under Assumptions A, B, and C,  $\frac{\hat{e}'\hat{e}}{T} - Q^{-1'}\Sigma_e Q^{-1} = O_p(\delta^{-2})$ .

**Proof of Lemma 2:** First, we expand the residuals  $\hat{e}$ .

$$\begin{aligned}
\hat{e} &= \hat{F} - \hat{Z}\hat{\Phi} = FH' + (\hat{F} - FH') - \hat{Z}\hat{\Phi}, \\
&= eH' + Z\Phi H' + (F - FH') - \hat{Z}\hat{\Phi}, \\
&= eH' + (\hat{F} - FH') + \left[ Z(I_p \otimes H') - \hat{Z} \right] (I_p \otimes H'^{-1})\Phi H' \\
&\quad + \hat{Z} \left[ (I_p \otimes H'^{-1})\Phi H' - \hat{\Phi} \right].
\end{aligned} \tag{A.1}$$

Then,

$$\begin{aligned}
\hat{e}'\hat{e}/T &= H(e'e/T)H' + T^{-1}(\hat{F} - FH')'(\hat{F} - FH') \\
&\quad + H\Phi'(I_p \otimes H'^{-1})'T^{-1} \left[ (Z(I_p \otimes H') - \hat{Z})'(Z(I_p \otimes H') - \hat{Z}) \right] (I_p \otimes H'^{-1})\Phi H' \\
&\quad + \left[ (I_p \otimes H'^{-1})\Phi H' - \hat{\Phi} \right]' T^{-1}\hat{Z}'\hat{Z} \left[ (I_p \otimes H'^{-1})\Phi H' - \hat{\Phi} \right] + \text{cross terms}, \\
&= H(e'e/T)H' + I + II + III + \text{cross terms}.
\end{aligned} \tag{A.2}$$

By using Lemma 1(a),  $\|I\| = O_p(\delta^{-2})$  and  $\|II\| = O_p(\delta^{-2})$ . Further,  $\|III\| = O_p(T^{-1})$  by using (A.5). Note that (A.5) is shown later without using this result. The cross terms are  $C + C'$ , where

$$\begin{aligned}
C &= \underbrace{He'(\hat{F} - FH')/T}_{O_p(\delta^{-2})} + H \left( \underbrace{e' \left[ Z(I_p \otimes H') - \hat{Z} \right] / T}_{O_p(\delta^{-2})} \right) (I_p \otimes H'^{-1})\Phi H' \\
&\quad + \underbrace{H \left( \underbrace{e'\hat{Z}/T}_{O_p(T^{-1/2})} \right)}_{O_p(T^{-1})} \left[ \underbrace{(I_p \otimes H'^{-1})\Phi H' - \hat{\Phi}}_{O_p(T^{-1/2})} \right] \\
&\quad + \underbrace{\left( (\hat{F} - FH')' \left[ Z(I_p \otimes H') - \hat{Z} \right] / T \right)}_{O_p(\delta^{-2})} (I_p \otimes H'^{-1})\Phi H' \\
&\quad + \underbrace{\left[ (\hat{F} - FH')\hat{Z}/T \right]}_{O_p(\delta^{-2})} \left[ \underbrace{(I_p \otimes H'^{-1})\Phi H' - \hat{\Phi}}_{O_p(T^{-1/2})} \right] \\
&\quad + H\Phi'(I_p \otimes H'^{-1})' \left( \underbrace{\left[ Z(I_p \otimes H') - \hat{Z} \right]' \hat{Z}/T}_{O_p(\delta^{-2})} \right) \left[ \underbrace{(I_p \otimes H'^{-1})\Phi H' - \hat{\Phi}}_{O_p(T^{-1/2})} \right],
\end{aligned}$$

so that  $C = O_p(\delta^{-2})$ .

**Lemma 3:** Under Assumptions A, B, C, and D',  $\hat{B}$  constructed by (ID) satisfies  $\left\| \hat{B} - HB \right\| \xrightarrow{p} 0$  as  $N, T \rightarrow \infty$ .

**Proof of Lemma 3:** We first show that Assumption D' implies Assumption D. To this end, we show that under Assumptions D' (a) and (b),  $Q^{-1}$  becomes a diagonal matrix. Combining this fact and the fact that  $B$  is triangular according to Assumption D' (c) yields the result. The proof is similar to Bai and Ng's (2013) PC1 condition. We start with

$$\hat{F}'F/T = (\hat{F} - FH')'F + HF'F/T = HF'F/T + O_p(\delta^{-2}).$$

Define  $H_F = (F'F/T)H$ . From the definition of  $H$ ,

$$H = V_{NT}^{-1}(\hat{F}'F/T)(\Lambda'\Lambda/N) = V_{NT}^{-1}H'_F(\Lambda'\Lambda/N) + O_p(\delta^{-2}),$$

so that multiplying by  $F'F/T$  on both sides gives

$$H_F = V_{NT}^{-1}H'_F(\Lambda'\Lambda/N)(F'F/T) + O_p(\delta^{-2}).$$

Multiplying by  $V_{NT}$  on each side and taking the transpose gives

$$(F'F/T)(\Lambda'\Lambda/N)H_F = H_F V_{NT} + O_p(\delta^{-2}).$$

Denote  $Q_F^{-1} = \Sigma_F Q^{-1}$  where  $\Sigma_F = p \lim_{T \rightarrow \infty} F'F/T$ . In the limit,

$$\Sigma_F \Sigma_\Lambda Q_F^{-1} = Q_F^{-1} V,$$

with  $V \equiv p \lim_{N, T \rightarrow \infty} V_{NT}$ . This equation suggests that  $Q_F^{-1}$  is a matrix consisting of the eigenvectors of  $\Sigma_F \Sigma_\Lambda$ . Since  $\Sigma_F \Sigma_\Lambda$  is diagonal according to Assumptions D' (a) and (b) and it has distinct eigenvalues under Assumption A(b), each eigenvalue is associated with a unique eigenvector, and each eigenvector has a single nonzero element. This implies that  $Q_F^{-1}$  is diagonal. Now, since  $\Sigma_F$  is diagonal and so is  $Q_F^{-1}$ , as shown above,  $Q^{-1} = \Sigma_F Q_F^{-1}$  is a diagonal matrix. Finally, it is straightforward to obtain

$$\begin{aligned} \left\| \hat{B} - HB \right\| &= \left\| Chol(\hat{e}'\hat{e}/T) - HB \right\|, \\ &\rightarrow_p \left\| Chol(Q^{-1'}BB'Q^{-1}) - Q^{-1'}B \right\|, \\ &= 0, \end{aligned}$$

where the second line uses Lemma 2 and  $\Sigma_e = BB'$ . The third line uses the triangularity of  $Q^{-1'}B$  and the sign restriction.

**Proof of Theorem 1:** We show the results for the individual structural parameters  $\lambda_i^s$  and  $\Phi^s$ . First, the reduced-form estimate  $\hat{\lambda}_i$  is expanded into the following form (see Bai, 2003, proof of Theorem 2):

$$\hat{\lambda}_i = H^{-1}\lambda_i + T^{-1}HF'u_i + T^{-1}\hat{F}'(F - \hat{F}H^{-1})\lambda_i + T^{-1}(\hat{F} - FH')'u_i. \quad (\text{A.3})$$

Lemma 3 tells us that  $\varepsilon = \hat{S} - HS \xrightarrow{p} 0$ , which in turn implies that  $\xi = \hat{S}^{-1} - S^{-1}H^{-1} \xrightarrow{p} 0$ . Then, the estimate for the structural parameter  $\lambda_i^s = S'\lambda_i$  is given by

$$\begin{aligned} \hat{\lambda}_i^s &= \hat{S}'\hat{\lambda}_i = S'\lambda_i + \varepsilon'H^{-1}\lambda_i + T^{-1}S'H'HF'u_i + T^{-1}\varepsilon'H'F'u_i \\ &\quad + T^{-1}\hat{S}'\hat{F}'(F - \hat{F}H^{-1})\lambda_i + T^{-1}\hat{S}'(\hat{F} - FH')'u_i. \end{aligned} \quad (\text{A.4})$$

Rearranging the terms in (A.4) gives

$$\begin{aligned}
\hat{\lambda}_i^s - \lambda_i^s &= T^{-1}S'H'HF'u_i + \varepsilon'H'^{-1}\lambda_i + T^{-1}\varepsilon'HF'u_i \\
&\quad + T^{-1}\hat{S}'\hat{F}'(F - \hat{F}H'^{-1})\lambda_i + T^{-1}\hat{S}'(\hat{F} - FH')'u_i, \\
&= I + II + III + IV + V.
\end{aligned}$$

Since  $I = O_p(T^{-1/2})$ ,  $II = o_p(1)$ ,  $III = o_p(T^{-1/2})$  according to Assumption B(e) and Lemma 3 and  $IV, V = O_p(\delta^{-2})$  from Lemma 1 (d) and (g), it is shown that  $\hat{\lambda}_i^s - \lambda_i^s \xrightarrow{p} 0$ . For  $\Phi^s$ , the least squares estimate for  $\Phi$  is given by

$$\begin{aligned}
\hat{\Phi} &= (\hat{Z}'\hat{Z})^{-1}(\hat{Z}'\hat{F}), \\
&= (\hat{Z}'\hat{Z})^{-1}(\hat{Z}'FH') + (\hat{Z}'\hat{Z})^{-1}\hat{Z}'(\hat{F} - FH'), \\
&= (\hat{Z}'\hat{Z})^{-1}\left[\hat{Z}'Z(I_p \otimes H')(I_p \otimes H'^{-1})\Phi H' + \hat{Z}'eH'\right] + (\hat{Z}'\hat{Z})^{-1}\hat{Z}'(\hat{F} - FH'), \\
&= (\hat{Z}'\hat{Z})^{-1}\left[\hat{Z}'\hat{Z}(I_p \otimes H'^{-1})\Phi H'\right] + (\hat{Z}'\hat{Z})^{-1}(\hat{Z}'eH') \\
&\quad + (\hat{Z}'\hat{Z})^{-1}\left[\hat{Z}'(Z(I_p \otimes H') - \hat{Z})(I_p \otimes H'^{-1})\Phi H'\right] + (\hat{Z}'\hat{Z})^{-1}\hat{Z}'(\hat{F} - FH'), \\
&= (I_p \otimes H'^{-1})\Phi H' + (T^{-1}\hat{Z}'\hat{Z})^{-1}\left[T^{-1}(I_p \otimes H)Z'eH'\right] \\
&\quad + (T^{-1}\hat{Z}'\hat{Z})^{-1}\left[T^{-1}(\hat{Z} - Z(I_p \otimes H'))'eH'\right] \\
&\quad + (T^{-1}\hat{Z}'\hat{Z})^{-1}\left[T^{-1}\hat{Z}'(Z(I_p \otimes H') - \hat{Z})(I_p \otimes H'^{-1})\Phi H'\right] \\
&\quad + (T^{-1}\hat{Z}'\hat{Z})^{-1}\left[T^{-1}\hat{Z}'(\hat{F} - FH')\right]. \tag{A.5}
\end{aligned}$$

Since  $T^{-1}\hat{Z}'\hat{Z} = O_p(1)$ , the last three terms in (A.5) are  $O_p(\delta^{-2})$  using Lemma 1. The estimate for the structural parameter  $\Phi^s = (I_p \otimes B')\Phi B'^{-1}$  is then

$$\begin{aligned}
(I_p \otimes \hat{B}')\hat{\Phi}\hat{B}'^{-1} &= (I_p \otimes B')\Phi B'^{-1} + (I_p \otimes B'H')(T^{-1}\hat{Z}'\hat{Z})^{-1}\left[T^{-1}(I_p \otimes H)Z'eB'^{-1}\right] \\
&\quad + (I_p \otimes B')\left[\Phi H' + (T^{-1}\hat{Z}'\hat{Z})^{-1}T^{-1}(I_p \otimes H)Z'eH'\right]\xi' \\
&\quad + (I_p \otimes \varepsilon')\left[(I_p \otimes H'^{-1})\Phi + (T^{-1}\hat{Z}'\hat{Z})^{-1}T^{-1}(I_p \otimes H)Z'e\right]B'^{-1} \\
&\quad + (I_p \otimes \varepsilon')\left[(I_p \otimes H'^{-1})\Phi + (T^{-1}\hat{Z}'\hat{Z})^{-1}T^{-1}(I_p \otimes H)Z'eH'\right]\xi' \\
&\quad + (I_p \otimes B)\left[T^{-1}\hat{Z}'(\hat{F} - FH')\right]B'^{-1}, \tag{A.6}
\end{aligned}$$

or

$$\begin{aligned}
\hat{\Phi}^s - \Phi^s &= (I_p \otimes B'H')(T^{-1}\hat{Z}'\hat{Z})^{-1} (T^{-1}(I_p \otimes H) Z'eB'^{-1}) \\
&\quad + (I_p \otimes B') \left[ \Phi H' + (T^{-1}\hat{Z}'\hat{Z})^{-1}T^{-1}(I_p \otimes H)Z'eH' \right] \xi' \\
&\quad + (I_p \otimes \varepsilon')[(I_p \otimes H'^{-1})\Phi + (T^{-1}\hat{Z}'\hat{Z})^{-1}T^{-1}(I_p \otimes H)Z'e]B'^{-1} \\
&\quad + (I_p \otimes \varepsilon') \left[ (I_p \otimes H'^{-1})\Phi + (T^{-1}\hat{Z}'\hat{Z})^{-1}T^{-1}(I_p \otimes H)Z'eH' \right] \xi' \\
&\quad + (I_p \otimes B)[T^{-1}\hat{Z}'(\hat{F} - FH')]B'^{-1}, \\
&= I + II + III + IV + V.
\end{aligned}$$

Since  $I = O_p(T^{-1/2})$ ,  $II, III, IV = o_p(1)$  according to Lemma 3, and  $V = O_p(\delta^{-2})$  from Lemma 1(a), we obtain  $\hat{\Phi}^s - \Phi^s \xrightarrow{p} 0$ . These imply the result for the structural IRF estimate  $\hat{\varphi}_{i,h} - \varphi_{i,h} \xrightarrow{p} 0$ , for any  $i$ .

**Lemma 4:** Let  $\hat{\sigma} = \text{vech}(\frac{\hat{\varepsilon}'\hat{\varepsilon}}{T})$  and  $\sigma = \text{vech}(Q^{-1}SS'Q^{-1})$ . Then,

$$\sqrt{T}(\hat{\sigma} - \sigma) \xrightarrow{d} N(0, \Sigma_\sigma)$$

as  $N, T \rightarrow \infty$  and  $\sqrt{T}/N \rightarrow 0$ .

**Proof of Lemma 4:** This result is the same as the conventional structural VAR proven by Lütkepohl (2005) based on the Cholesky identification. The difference in this context is that  $\hat{\varepsilon}$  involves factor estimation errors, meaning that the expression of  $\frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sqrt{T}}$  is (A.2), but, as in Lemma 2, it is shown that the terms associated with the factor estimation errors are at most  $O_p(\sqrt{T}\delta^{-2})$  and are negligible when  $\sqrt{T}/N \rightarrow 0$  under our assumptions. Hence, we can apply this result, and the explicit form of  $\Sigma_\sigma$  is given by

$$\Sigma_\sigma = 2D_r^+ [(Q^{-1}SS'Q^{-1}) \otimes (Q^{-1}SS'Q^{-1})] D_r^{+'},$$

where  $D_r^+ = (D_r'D_r)^{-1}D_r'$  is the Moore-Penrose inverse of the duplication matrix  $D_r$ .

**Proof of Theorem 2:** Given that the reduced-form parameter estimates  $\hat{\lambda}_i$  and  $\hat{\Phi}$  are asymptotically normal as  $N, T \rightarrow \infty$  and  $\sqrt{T}/N \rightarrow 0$  under Assumptions A and B (see Bai and Ng, 2006) up to rotation and  $\hat{B}$  is asymptotically normal according to Lemma 4, the delta method implies that the structural IRF is asymptotically normal with the variance given in the theorem.

Theorems 3 and 4 are shown by using Lemmas BT1 to BT4. These lemmas are proven by more primitive conditions than in Lemmas BT-B and BT-C, which are shown to hold in Appendix B under Procedures A and B.

**Lemma BT-B:** Suppose Assumptions A, B, and C hold. Under the proposed bootstrap Procedures A and B, the following holds:

- (a)  $E^*(u_{it}^*) = 0$  for all  $(i, t)$ .
- (b)  $T^{-1} \sum_{t=1}^T \sum_{s=1}^T |\gamma_{st}^*|^2 = O_p(1)$ , where  $\gamma_{st}^* = E^*(N^{-1} \sum_{i=1}^N e_{is}^* e_{it}^*)$ .  $N^{-1} \sum_{i=1}^N \sum_{j=1}^N |\tau_{ij}^*|^2 = O_p(1)$ , where  $\tau_{ij}^* = E^*(e_{it}^* e_{jt}^*)$ .
- (c)  $T^{-2} \sum_{t=1}^T \sum_{s=1}^T E^* \left| N^{-1/2} \sum_{i=1}^N [u_{is}^* u_{it}^* - E^*(u_{is}^* u_{it}^*)] \right|^2 = O_p(1)$ .
- (d) For each  $i$ ,  $\Theta_i^{*-1/2} T^{-1/2} \sum_{t=1}^T F_t^* u_{it}^* \xrightarrow{d} N(0, I_r)$ , in probability, where  $\Theta_i^* \equiv Var^* \left( T^{-1/2} \sum_{t=1}^T F_t^* u_{it}^* \right)$  and  $p \lim \Theta_i^* = Q^{-1'} \Theta_i Q^{-1}$ .

**Lemma BT-C:** Suppose Assumptions A, B, and C hold. Under the proposed bootstrap Procedures A and B, the following holds:

- (a)  $E^*(e_t^*) = 0$ ,  $E^* \|e_t^*\|^2 = O_p(1)$ , and  $e_t^*$  and  $e_s^*$  are independent for all  $t \neq s$ .
- (b)  $\Sigma^{*-1/2} T^{-1/2} \sum_{t=p+1}^T vec(Z_t^* e_t^{*'}) \xrightarrow{d} N(0, I_{pr^2})$ , in probability, where  $\Sigma^* \equiv Var^* \left( \frac{1}{\sqrt{T}} \sum_{t=1}^T Z_t^* e_t^{*'} \right)$ , where  $p \lim \Sigma^* = [(I_p \otimes Q^{-1'}) \Sigma_Z (I_p \otimes Q^{-1})] \otimes (Q^{-1'} \Sigma_e Q^{-1})$ , in probability.
- (c) The roots of  $\det(I_r - \hat{\Phi}_1 z - \hat{\Phi}_2 z^2 - \dots - \hat{\Phi}_p z^p) = 0$  lie outside the unit circle.

**Lemma BT1.** Under Assumptions A and B and Conditions BT1 and BT2, the following conditions hold for  $j = 0, 1, \dots, p$ :

- (a)  $T^{-1} \sum_{t=j+1}^T \left\| \hat{F}_t^* - H^* F_t^* \right\| \left\| \hat{F}_{t-j}^* - H^* F_{t-j}^* \right\| = O_{p^*}(\delta^{-2})$ .
- (b)  $T^{-1} \sum_{t=j+1}^T (\hat{F}_t^* - H^* F_t^*) \hat{F}_{t-j}^{*'} = O_{p^*}(\delta^{-2})$ .
- (c)  $T^{-1} \sum_{t=j+1}^T (\hat{F}_{t-j}^* - H^* F_{t-j}^*) \hat{F}_t^{*'} = O_{p^*}(\delta^{-2})$ .
- (d)  $T^{-1} \sum_{t=j+1}^T (\hat{F}_t^* - H^* F_t^*) F_{t-j}^{*'} = O_{p^*}(\delta^{-2})$ .
- (e)  $T^{-1} \sum_{t=j+1}^T (\hat{F}_{t-j}^* - H^* F_{t-j}^*) F_t^{*'} = O_{p^*}(\delta^{-2})$ .
- (f)  $T^{-1} \sum_{t=j+1}^T (\hat{F}_{t-j}^* - H^* F_{t-j}^*) e_t^{*'} = O_{p^*}(\delta^{-2})$ .
- (g)  $T^{-1} \sum_{t=1}^T (\hat{F}_t^* - H^* F_t^*) u_{it}^* = O_{p^*}(\delta^{-2})$ , in probability.

**Proof of Lemma BT1:** As shown in Lemma A1, we only prove the case of  $j = 0$ . We have

$$\begin{aligned} \hat{F}_t^* - H^* F_t^* &= V^{*-1} \left( T^{-1} \sum_{s=1}^T \hat{F}_s^* \gamma_{st}^* + T^{-1} \sum_{s=1}^T \hat{F}_s^* \zeta_{st}^* + T^{-1} \sum_{s=1}^T \hat{F}_s^* \eta_{st}^* + T^{-1} \sum_{s=1}^T \hat{F}_s^* \eta_{ts}^* \right), \\ &= V^{*-1} (A_{1t}^* + A_{2t}^* + A_{3t}^* + A_{4t}^*), \end{aligned}$$

where  $V^* = O_{p^*}(1)$ ,

$$\begin{aligned}\gamma_{st}^* &= E^* \left( N^{-1} \sum_{i=1}^N u_{is}^* u_{it}^* \right), \\ \zeta_{st}^* &= N^{-1} \sum_{i=1}^N [u_{is}^* u_{it}^* - E(u_{is}^* u_{it}^*)], \\ \eta_{st}^* &= N^{-1} \sum_{i=1}^N \hat{\lambda}_i' F_s^* u_{jt}^*.\end{aligned}$$

For part (a),

$$T^{-1} \sum_{t=1}^T \left\| \hat{F}_t^* - H^* F_t^* \right\|^2 \leq 4T^{-1} \sum_{k=1}^4 \sum_{t=1}^T \|A_{kt}\|^2.$$

For the term of  $A_{1t}^*$ ,

$$\begin{aligned}T^{-1} \sum_{t=1}^T \|A_{1t}^*\|^2 &\leq T^{-1} (T^{-1} \sum_{s=1}^T \left\| \hat{F}_s^* \right\|^2) (T^{-1} \sum_{t=1}^T \sum_{s=1}^T \gamma_{st}^{*2}), \\ &= T^{-1} \times I \times II,\end{aligned}$$

but  $I$  is  $r$  according to the property of PC and  $II$  is  $O_p(1)$  from Lemma BT-B(b). Hence, this is  $O_p(T^{-1})$ . For the term of  $A_{2t}^*$ ,

$$\begin{aligned}T^{-1} \sum_{t=1}^T \|A_{2t}^*\|^2 &\leq (T^{-1} \sum_{s=1}^T \left\| \hat{F}_s^* \right\|^2) (T^{-2} \sum_{t=1}^T \sum_{s=1}^T \zeta_{st}^{*2}), \\ &= I \times II,\end{aligned}$$

but  $I$  is  $r$  and  $II$  is  $O_{p^*}(N^{-1})$  from Lemma BT-B(c). Hence, this is  $O_{p^*}(N^{-1})$ . For the term of  $A_{3t}^*$ ,

$$\begin{aligned}T^{-1} \sum_{t=1}^T \|A_{3t}^*\|^2 &= T^{-1} \sum_{t=1}^T T^{-2} \left\| \sum_{s=1}^T \hat{F}_s^* F_s^{*'} \frac{\hat{\Lambda}' u_t^*}{N} \right\|^2, \\ &\leq \left( T^{-1} \sum_{t=1}^T \left\| \frac{\hat{\Lambda}' u_t^*}{N} \right\|^2 \right) \left\| T^{-1} \sum_{s=1}^T \hat{F}_s^* F_s^{*'} \right\|, \\ &= I \times II,\end{aligned}$$

but  $I = O_{p^*}(N^{-1})$  according to the independence between  $\hat{\lambda}_i$  and  $u_{it}^*$  and  $II$  is  $O_p(1)$  and  $III = O_{p^*}(1)$ . Hence, this is  $O_{p^*}(N^{-1})$ . For parts (b) and (c),

$$T^{-1} \sum_{t=1}^T (\hat{F}_t^* - H^* F_t^*) \hat{F}_t^{*'} = \sum_{k=1}^4 b_k^*,$$

where

$$\begin{aligned}b_1^* &= T^{-1} \sum_{t=1}^T A_{1t} \hat{F}_t^{*'} \leq \underbrace{(T^{-1} \sum_{t=1}^T \|A_{1t}^*\|^2)^{1/2}}_{=O_p(T^{-1})} \underbrace{(T^{-1} \sum_{t=1}^T \left\| \hat{F}_t^* \right\|^2)^{1/2}}_{=r}, \\ b_2^* &= T^{-1} \sum_{t=1}^T A_{2t} \hat{F}_t^{*'} \leq \underbrace{(T^{-1} \sum_{t=1}^T \|A_{2t}^*\|^2)^{1/2}}_{O_p(N^{-1})} \underbrace{(T^{-1} \sum_{t=1}^T \left\| \hat{F}_t^* \right\|^2)^{1/2}}_{=r}, \\ b_3^* &= T^{-1} \sum_{t=1}^T A_{3t} \hat{F}_t^{*'} \leq \underbrace{(T^{-1} \sum_{t=1}^T \|A_{3t}^*\|^2)^{1/2}}_{=O_p(N^{-1})} \underbrace{(T^{-1} \sum_{t=1}^T \left\| \hat{F}_t^* \right\|^2)^{1/2}}_{=r},\end{aligned}$$

and the order of  $b_4^*$  is the same as that of  $b_3^*$ . Therefore,  $T^{-1} \sum_{t=1}^T (\hat{F}_t^* - H^* F_t^*) \hat{F}_t^{*'} = O_p(T^{-1}) + O_p(N^{-1})$  and the result follows. The proof of parts (d) and (e) is the same as that of parts (b) and (c) but changes  $T^{-1} \sum_{t=1}^T \left\| \hat{F}_t^* \right\|^2 = r$  into  $T^{-1} \sum_{t=1}^T \|F_t^*\|^2 = O_{p^*}(1)$ , which is satisfied by Lemma BT-C(c). Part (f) is shown by using  $T^{-1} \sum_{t=1}^T \|e_t^*\|^2 = O_{p^*}(1)$  according to Lemma BT-C(a). Part (g) is shown by using  $T^{-1} \sum_{t=1}^T u_{it}^{*2} = O_{p^*}(1)$  from Lemma BT-B(a).

**Lemma BT2.** Suppose Assumptions A, B, C, and D hold. Then,

- (a)  $\left\| \hat{B}^* - H^* \hat{B} \right\| \xrightarrow{p^*} 0$  in probability as  $N, T \rightarrow \infty$ ;  
(b)  $\left\| \hat{B}^{**} - \hat{B} \right\| \xrightarrow{p^*} 0$  in probability as  $N, T \rightarrow \infty$ .

**Proof of Lemma BT2:** This proof requires two steps. Step 1 shows that the identifying restrictions hold in the bootstrap space, i.e. the same restrictions are relevant as in the limits of the original estimates. Step 2 confirms the convergence results used to prove Lemma 2 in the bootstrap space.

**Step 1:** We show that the following conditions analogous to Assumption D' hold: 1.  $F^{*'} F^* / T \xrightarrow{p^*} I_r$  in probability; 2.  $\hat{\Lambda}' \hat{\Lambda} / N$  is a diagonal matrix; 3.  $\hat{B}$  is an upper or lower triangular matrix, and the signs of the diagonal elements of  $Q^{*-1'} \hat{B}$  are known. 1 is because  $F^{*'} F^* / T \xrightarrow{p^*} \hat{F}' \hat{F} / T = I_r$ . 2 is because  $\hat{\Lambda}' \hat{\Lambda} / N = V_{NT}$  by construction of the PCs and  $V_{NT}$  is a diagonal matrix. The triangularity of  $\hat{B}$  in 3 holds because  $\hat{B} = Chol\left(\frac{\hat{e}' \hat{e}}{T}\right)$ . Finally, the signs are deduced in the bootstrap replications from the signs of the structural IRF estimates as in the original estimates. Hence, Assumption D holds in the bootstrap space. This gives the fact that  $Q^{*-1'} \hat{B}$  is a triangular matrix and the signs of its diagonal matrix are known.

**Step 2 (Procedure A):** We first ensure that

$$\|\hat{e}^{*'} \hat{e}^* / T - H^* (e^{*'} e^* / T) H^{*'}\| = O_p(\delta^{-2}),$$

by following the same discussion as Lemma 2 given Lemma BT1 and (A.11). Then, we obtain

$$\left\| \hat{B}^* - H^{*'} \hat{B} \right\| = \underbrace{\| Chol\left(\underbrace{\hat{e}^{*'} \hat{e}^* / T}_{\rightarrow Q^{*-1'} \hat{B} \hat{B}' Q^{*-1}}\right) - \underbrace{H^{*'} \hat{S}}_{\rightarrow Q^{*-1'} \hat{B}} \|}_{Q^{*-1'} \hat{B}} \xrightarrow{p^*} 0$$

in probability, using the facts  $E^*(e^{*'} e^* / T) = \hat{B} \hat{B}'$ ,  $Q^{*-1'} \hat{B}$  is triangular and the sign restrictions hold.

**Step 2 (Procedure B):** We first show that

$$\|\hat{e}^{**'} \hat{e}^{**} / T - e^{*'} e^* / T\| = o_p(1),$$

by following standard OLS estimates (because no factor estimation is involved in Procedure B). Hence,

$$\left\| \hat{B}^{**} - \hat{B} \right\| = \left\| \underbrace{\text{Chol}(\underbrace{\hat{e}^{**'} \hat{e}^{**} / T}_{\rightarrow \hat{B} \hat{B}'})}_{\hat{B}} - \hat{B} \right\| \xrightarrow{D^*} 0$$

in probability, by using  $E^*(e' e / T) = \hat{B} \hat{B}'$ .

**Lemma BT3:** Suppose Assumptions A, B, C, and D hold. a. Let  $\hat{\sigma}^* = \text{vech} \left( H^{*-1} \frac{\hat{e}' \hat{e}}{T} H^{*-1} \right)$  and  $\tilde{\sigma} = \text{vech} \left( \hat{B} \hat{B}' \right)$ . Then,

$$\sqrt{T} \Omega_B^{*-1/2} (\hat{\sigma}^* - \tilde{\sigma}) \xrightarrow{d} N(0, I_{r(r+1)/2}),$$

in probability as  $N, T \rightarrow \infty$  and  $\sqrt{T}/N \rightarrow 0$ . Moreover,  $p \lim(\Sigma_\sigma^*) = \Sigma_\sigma$ . b. Let  $\hat{\sigma}^{**} = \text{vech} \left( \frac{\hat{e}^{**'} \hat{e}^{**}}{T} \right)$  and  $\tilde{\sigma} = \text{vech} \left( \hat{B} \hat{B}' \right)$ . Then,

$$\sqrt{T} \Omega_B^{**1/2} (\hat{\sigma}^{**} - \tilde{\sigma}) \xrightarrow{d} N(0, I_{r(r+1)/2}),$$

in probability as  $N, T \rightarrow \infty$ . Moreover,  $p \lim(\Omega_B^{**}) = \Omega_B$ .

**Proof of Lemma BT3:** Part (a) is a bootstrap analog of Lemma 4 and the proof directly follows its proof by replacing the original entities with the bootstrap counterparts. Part (b) straightforwardly follows Lütkepohl (2005) because under Procedure B the VAR residuals do not include the factor estimation errors. From Lemma 2,  $p \lim(\hat{B} \hat{B}') = Q^{-1} B B' Q^{-1}$  so that  $p \lim(\Omega_B^*) = \Omega_B$  and  $p \lim(\Omega_B^{**}) = \Omega_B$ .

**Lemma BT4:** Suppose Assumptions A, B, C, and D hold. Then,  $H^{*'} H^* \xrightarrow{p^*} I_r$  in probability as  $N, T \rightarrow \infty$ .

**Proof of Lemma BT4:** First,

$$\hat{F}^{*'} F^* / T = (\hat{F}^* - F^* H^{*'})' F^* / T + H^* F^{*'} F^* / T,$$

then,

$$H^{*'} = H^* - H^* F^{*'} F^* / T + \hat{F}^{*'} F^* / T - (\hat{F}^* - F^* H^{*'})' F^* / T. \quad (\text{A.7})$$

Right multiplying (A.7) by  $H^*$  yields

$$\begin{aligned} H^{*'} H^* &= [H^* H^* - H^* (F^{*'} F^* / T) H^*] + \hat{F}^{*'} \hat{F}^* / T \\ &\quad + \hat{F}^{*'} (F^* H^{*'} - \hat{F}^*) / T - (\hat{F}^* - F^* H^{*'})' F^* H^* / T, \\ &= I + II + III + IV. \end{aligned}$$

Since  $\frac{F'^*F^*}{T} \xrightarrow{p^*} \frac{\hat{F}'\hat{F}}{T} = I_r$  in probability,  $I = o_{p^*}(1)$ ,  $II = I_r$  by construction, and  $III, IV = O_{p^*}(\delta^{-2})$  by Lemma BT1.

**Proof of Theorem 3:** We equivalently show that  $\sqrt{T}(\hat{\varphi}_{ih} - \varphi_{ih}) \xrightarrow{d} N(0, \Omega_{\varphi_{ih}})$  and  $\sqrt{T}(\hat{\varphi}_{ih}^* - \hat{\varphi}_{ih}) \xrightarrow{d^*} N(0, \Omega_{\varphi_{ih}})$  in probability as  $N, T \rightarrow \infty$  and  $\sqrt{T}/N \rightarrow 0$ . In the first step, we confirm that  $\hat{\varphi}_{ih}^* - \hat{\varphi}_{ih} = o_{p^*}(1)$  in probability by showing that it holds for the individual parameters  $\lambda_i^s$  and  $\Phi^s$ . In the second step, we work on the limit distributions. First, let  $\varepsilon^* = \hat{B}^* - H^*\hat{B} = o_{p^*}(1)$  in probability by Lemma BT2(a). This implies  $\xi^* = o_{p^*}(1)$  in probability. For  $\lambda_i^s$ ,

$$\begin{aligned} \hat{\lambda}_i^{s*} - \hat{\lambda}_i^s &= T^{-1}\hat{B}^{*'}H^*H^*F'^*u_i^* + \varepsilon^{*'}H^{*'-1}\hat{\lambda}_i + T^{-1}\varepsilon^{*'}H^*F'^*u_i^* \\ &\quad + T^{-1}\hat{B}^{*'}\hat{F}'(F^* - \hat{F}^*H^{*'-1})\hat{\lambda}_i + T^{-1}\hat{B}^{*'}(\hat{F}^* - F^*H^{*'})'u_i^*, \quad (\text{A.8}) \\ &= I + II + III + IV + V. \end{aligned}$$

This is the bootstrap counterpart of (A.4). Hence,  $I = O_{p^*}(T^{-1/2})$  from Lemma BT-B(d),  $II = o_{p^*}(1)$  from Lemma BT2(a),  $III = o_{p^*}(T^{-1/2})$  from Lemma BT-B(d) and Lemma BT2(a), and  $IV, V = O_{p^*}(\delta^{-2})$  from Lemma BT1 in probability. Hence, the RHS of (A.8) is  $o_{p^*}(1)$  in probability as  $N, T \rightarrow \infty$ . For  $\Phi^s$ , it follows that

$$\begin{aligned} \hat{\Phi}^{s*} - \hat{\Phi}^s &= (I_p \otimes \hat{B}'H^{*'})(T^{-1}\hat{Z}^{*'}\hat{Z}^*)^{-1}[T^{-1}(I_p \otimes H^*Z'^*e^*\hat{B}'^{-1}) \\ &\quad + (I_p \otimes \hat{B}')\left[\hat{\Phi}H^{*'} + (T^{-1}\hat{Z}^{*'}\hat{Z}^*)^{-1}(I_p \otimes H^*)Z'^*e^*H^{*'}\right]\xi^* \\ &\quad + (I_p \otimes \varepsilon^{*'})[(I_p \otimes H^{*'-1})\hat{\Phi} + (T^{-1}\hat{Z}^{*'}\hat{Z}^*)^{-1}T^{-1}(I_p \otimes H^*)Z'^*e^*]\hat{B}'^{-1} \\ &\quad + (I_p \otimes \varepsilon^{*'})\left[(I_p \otimes H^{*'-1})\hat{\Phi} + (T^{-1}\hat{Z}^{*'}\hat{Z}^*)^{-1}T^{-1}(I_p \otimes H^*)Z'^*e^*H^{*'}\right]\xi^* \\ &\quad + (I_p \otimes \hat{B}')\left[T^{-1}\hat{Z}^{*'}(\hat{F}^* - F^*H^{*'})\right]\hat{B}'^{-1}, \quad (\text{A.9}) \\ &= I + II + III + IV + V, \end{aligned}$$

in probability under Lemma BT1. Now, (A.9) is analogous to (A.6). It is similarly shown that  $I = o_{p^*}(T^{-1/2})$  and  $II, III$ , and  $IV$  are all  $o_{p^*}(1)$  in probability under Lemma BT2.  $V = O_{p^*}(\delta^{-2})$  according to Lemma BT1. Hence,  $\hat{\Phi}^{s*} - \hat{\Phi}^s = o_{p^*}(1)$  in probability. These two facts imply  $\hat{\varphi}_{ih}^* - \hat{\varphi}_{ih} \xrightarrow{p^*} 0$  in probability as  $N, T \rightarrow \infty$ . The second step involves the limit distributions. I hypothetically consider the “rotation-adjusted” version of the reduced-form parameter estimates induced by  $H^*$  in the bootstrap space. Specifically, the rotation-adjusted estimates of  $\hat{\lambda}_i^*$ ,  $\hat{\Phi}^*$ , and  $\hat{B}^*$  are  $H^{*'}\hat{\lambda}_i^*$ ,  $(I_p \otimes H^{*'})\hat{\Phi}^*H^{*'-1}$ , and  $H^{*-1}\hat{B}^*$ , respectively. Given the fact that the structural IRF is identified, all the  $H^*$ s are eventually cancelled out in the structural IRF estimate. For  $\lambda_i$ , I construct a bootstrap analogue of (A.3) left-multiplied by  $H^{*'}$  and scaled by  $\sqrt{T}$ :

$$\begin{aligned} &\sqrt{T}(H^{*'}\hat{\lambda}_i^* - H^{*'}H^{*'-1}\hat{\lambda}_i) \\ &= T^{-1/2}H^{*'}H^*F'^*u_i^* + T^{-1/2}H^{*'}\hat{F}'(F^* - \hat{F}^*H^{*'-1})\hat{\lambda}_i + T^{-1/2}H^{*'}(\hat{F}^* - F^*H^{*'})'u_i^*, \\ &= T^{-1/2}H^{*'}H^*F'^*u_i^* + O_{p^*}(T^{1/2}\delta^{-2}) + O_{p^*}(T^{1/2}\delta^{-2}) \quad (\text{A.10}) \end{aligned}$$

in probability from Lemma BT1, so that

$$\sqrt{T}(H^{*'}\hat{\lambda}_i^* - \hat{\lambda}_i) \xrightarrow{d^*} N(0, Q^{-1'}\Theta_i Q^{-1})$$

in probability as  $N, T \rightarrow \infty$  and  $\sqrt{T}/N \rightarrow 0$  under Lemmas BT-B (i) and (j) and Lemma BT4. For  $\hat{\Phi}$ , constructing a bootstrap analogue of (A.5) left-multiplied by  $(I_p \otimes H^{*'})$ , right-multiplied by  $H^{*'-1}$ , and scaled by  $\sqrt{T}$  gives

$$\begin{aligned} & \sqrt{T}[(I_p \otimes H^{*'})\hat{\Phi}^* H^{*'-1} - (I_p \otimes H^{*'} H^{*'-1})\hat{\Phi} H^{*'} H^{*'-1}] \\ = & T^{-1/2}(I_p \otimes H^{*'}) \left( T^{-1}\hat{Z}^{*'}\hat{Z}^* \right)^{-1} [Z^*(I_p \otimes H^{*'})]' e^* H^{*'} H^{*'-1} \\ & + T^{-1/2}(I_p \otimes H^{*'}) \left( T^{-1}\hat{Z}^{*'}\hat{Z}^* \right)^{-1} \left[ \hat{Z}^* - Z^*(I_p \otimes H^{*'}) \right]' e^* H^{*'} H^{*'-1} \\ & + T^{-1/2}(I_p \otimes H^{*'}) \left( T^{-1}\hat{Z}^{*'}\hat{Z}^* \right)^{-1} \hat{Z}^{*'} (Z^*(I_p \otimes H^{*'}) - \hat{Z}^*) (I_p \otimes H^{*'-1}) \hat{\Phi} H^{*'} H^{*'-1} \\ & + T^{-1/2}(I_p \otimes H^{*'}) \left( T^{-1}\hat{Z}^{*'}\hat{Z}^* \right)^{-1} \hat{Z}^{*'} (\hat{F}^{*'} - F^* H^{*'}) H^{*'-1}, \\ = & T^{-1/2}(I_p \otimes H^{*'}) \left( T^{-1}\hat{Z}^{*'}\hat{Z}^* \right)^{-1} [Z^*(I_p \otimes H^{*'})]' e^* \\ & + O_{p^*}(T^{1/2}\delta^{-2}) + O_{p^*}(T^{1/2}\delta^{-2}) + O_{p^*}(T^{1/2}\delta^{-2}) \end{aligned} \tag{A.11}$$

in probability under Lemma BT1. Since

$$\begin{aligned} T^{-1}\hat{Z}^{*'}\hat{Z}^* &= T^{-1}[Z^*(I_p \otimes H^{*'})]' Z^*(I_p \otimes H^{*'}) \\ & \quad + T^{-1} \left[ \hat{Z}^* - Z^*(I_p \otimes H^{*'}) \right]' Z^*(I_p \otimes H^{*'}) \\ & \quad + T^{-1} [Z^*(I_p \otimes H^{*'})]' \left[ \hat{Z}^* - (I_p \otimes H^{*'}) Z^* \right], \\ &= T^{-1}(I_p \otimes H^{*'}) Z^{*'} Z^*(I_p \otimes H^{*'}) + O_{p^*}(T^{1/2}\delta^{-2}) \end{aligned}$$

in probability, we obtain

$$\begin{aligned} & \sqrt{T} \left[ (I_p \otimes H^{*'})\hat{\Phi}^* H^{*'-1} - \hat{\Phi} \right], \\ = & T^{-1/2}(I_p \otimes H^{*'})(I_p \otimes H^{*'})(Z^{*'} Z^*/T)(I_p \otimes H^{*'})(I_p \otimes H^{*'}) Z^{*'} e^* + O_{p^*}(T^{1/2}\delta^{-2}), \\ = & (T^{-1} Z^{*'} Z^*)^{-1} T^{-1/2} Z^{*'} e^* + O_{p^*}(T^{1/2}\delta^{-2}), \end{aligned}$$

in probability by using Lemma BT4. Hence,

$$\sqrt{T} \text{vec} \left[ (I_p \otimes H^{*'})\hat{\Phi}^* H^{*'-1} - \hat{\Phi} \right] \xrightarrow{d^*} N(0, [I_r \otimes \Sigma_{Z^*}^{-1}]' \Sigma^* [I_r \otimes \Sigma_{Z^*}^{-1}]),$$

with  $\Sigma_{Z^*} = p \lim Z^{*'} Z^*/T$  as  $N, T \rightarrow \infty$  and  $\sqrt{T}/N \rightarrow 0$  under Lemma BT-C(b). Since

$$\Sigma_{Z^*} \xrightarrow{p} (I_p \otimes Q^{-1}) \Sigma_Z (I_p \otimes Q^{-1})$$

in probability, and the probability limit of  $\Sigma^*$  is given in Lemma BT-C(b),

$$\begin{aligned} & \sqrt{T} \text{vec} \left[ (I_p \otimes H^{*'}) \hat{\Phi}^* H^{*'-1} - \hat{\Phi} \right] \\ & \xrightarrow{d^*} N(0, [(I_p \otimes Q^{-1'}) \Sigma_Z (I_p \otimes Q^{-1})]^{-1} \\ & \quad \times [(I_p \otimes Q^{-1'}) \Sigma_Z (I_p \otimes Q^{-1})] \otimes [Q^{-1'} \Sigma_e Q^{-1}] \\ & \quad \times [(I_p \otimes Q^{-1'}) \Sigma_Z (I_p \otimes Q^{-1})]^{-1'}) \end{aligned}$$

in probability. Finally, for  $B$ , the rotation-adjusted version of  $\hat{B}^*$  is  $H^{*-1} \hat{B}^*$  and  $\sqrt{T} \text{vech}(H^{*-1} \hat{B}^* - \hat{B}) \xrightarrow{d^*} N(0, \Omega_B)$  in probability from Lemma BT3(a). For the original estimate, it is straightforward to show from (A.3), (A.5), and Lemma 3 that

$$\begin{aligned} & \sqrt{T}(\hat{\lambda}_i - H'^{-1} \lambda_i) \xrightarrow{d} N(0, Q^{-1'} \Theta_i Q^{-1}), \\ & \sqrt{T} \text{vec} \left[ \hat{\Phi} - (I_p \otimes H'^{-1}) \hat{\Phi} H' \right] \xrightarrow{d} N(0, [(I_p \otimes Q^{-1'}) \Sigma_Z (I_p \otimes Q^{-1})]^{-1} \\ & \quad \times [(I_p \otimes Q^{-1'}) \Sigma_Z (I_p \otimes Q^{-1})] \otimes [Q^{-1'} \Sigma_e Q^{-1}] \\ & \quad \times [(I_p \otimes Q^{-1'}) \Sigma_Z (I_p \otimes Q^{-1})]^{-1'}), \\ & \sqrt{T} \text{vech}(\hat{B} - HB) \xrightarrow{d} N(0, \Omega_B), \end{aligned}$$

and the result follows.

**Proof of Theorem 4:** With Procedure B,  $H^*$  is not generated. Hence, we do not need to introduce the rotation-adjusted parameter estimates. In addition, the expansions of the bootstrap parameter estimates will have fewer terms in the absence of the factor estimation errors. For  $\lambda_i^s$ ,

$$\hat{\lambda}_i^{s**} - \hat{\lambda}_i^s = T^{-1} \hat{S}' F^{*'} u_i^* + \varepsilon^{**'} \hat{\lambda}_i + T^{-1} \varepsilon^{**'} F^{*'} u_i^* = I + II + III. \quad (\text{A.12})$$

$I = O_{p^*}(T^{-1/2})$  under Lemma BT-B(d) and  $II, III = o_{p^*}(1)$  in probability. For  $\Phi^s$ ,

$$\begin{aligned} \hat{\Phi}^{s**} - \hat{\Phi}^s &= (I_p \otimes \hat{B})(Z^{*'} Z^*)^{-1} (Z^{*'} e^* \hat{B}^{-1}) \\ & \quad + (I_p \otimes \hat{B}) \left[ \hat{\Phi} + (Z^{*'} Z^*)^{-1} Z^{*'} e^* \right] \xi^* \\ & \quad + (I_p \otimes \varepsilon^*) \left[ \hat{\Phi} + (Z^{*'} Z^*)^{-1} Z^{*'} e^* \right] \hat{B}^{-1} \\ & \quad + (I_p \otimes \varepsilon^*) \left[ \hat{\Phi} + (T^{-1} Z^{*'} Z^*)^{-1} T^{-1} Z^{*'} e^* \right] \xi^* \\ & \xrightarrow{p^*} 0 \end{aligned}$$

in probability. Next, consider the asymptotic distributions. For  $\lambda_i$ ,

$$\sqrt{T}(\hat{\lambda}_i^{**} - \hat{\lambda}_i) = T^{-1/2} F^* u_i^* \xrightarrow{d^*} N(0, Q^{-1'} \Theta_i Q^{-1})$$

in probability as  $N, T \rightarrow \infty$  from Lemma BT-B(d). For  $\Phi$ , since

$$\hat{\Phi}^{**} - \hat{\Phi} = (Z^{*'} Z^* / T)^{-1} (Z^{*'} e^* / T),$$

we obtain

$$\begin{aligned} \sqrt{T}vec(\hat{\Phi}^{**} - \hat{\Phi}) &\xrightarrow{d^*} N(0, [(I_p \otimes Q^{-1'})\Sigma_Z(I_p \otimes Q^{-1})]^{-1} \\ &\quad \times [(I_p \otimes Q^{-1'})\Sigma_Z(I_p \otimes Q^{-1})] \otimes [Q^{-1'}\Sigma_e Q^{-1}] \\ &\quad \times [(I_p \otimes Q^{-1'})\Sigma_Z(I_p \otimes Q^{-1})]^{-1'}) \end{aligned}$$

in probability as  $N, T \rightarrow \infty$ , following the proof of Theorem 3. Finally, Lemma BT3(b) guarantees  $\sqrt{T}vec(\hat{B}^{**} - \hat{B}) \xrightarrow{d^*} N(0, \Omega_B)$  in probability as  $N, T \rightarrow \infty$ .

## Appendix B: Proof of Lemmas BT-B and BT-C

This appendix provides a proof that Lemmas BT-B and BT-C hold under Assumptions A, B, and C and the bootstrapped residuals  $\{u_t^*\}$  and  $\{e_t^*\}$  generated by the proposed designs. When we say that the cross-sectional dependence of  $\{\hat{u}_t\}$  is preserved in  $\{u_t^*\}$ , it means that  $N^{-1/2} \sum_{i=1}^N u_{it} = O_p(1)$  implies  $N^{-1/2} \sum_{i=1}^N u_{it}^* = O_p(1)$  in probability.

**Proof of Lemma BT-B:** We closely follow the proof of Conditions A\*(a), A\*(b), A\*(c), and D of Gonçalves and Perron (2014), but under the i.i.d. bootstrap scheme with  $\lambda_i$ ,  $F_t$ , and  $u_{js}$  considered to be mutually independent. For part (a),  $E^*(u_{it}^*) = 0$  holds since the procedure requires  $\{\hat{u}_{it}\}$  to be demeaned in the time and cross-sectional directions before resampling. Part (b) is shown by the following. Since  $\{u_t^*\}$  is an independent sequence over time,  $\gamma_{st}^* = 0$  for  $s \neq t$  so that

$$T^{-1} \sum_{t=1}^T \sum_{s=1}^T |\gamma_{st}^*|^2 = T^{-1} \sum_{t=1}^T |\gamma_{tt}^*|^2 = O_p(1),$$

since  $|\gamma_{tt}^*| = \left| N^{-1} \sum_{i=1}^N \hat{u}_{it}^2 \right| = O_p(1)$ . The result for  $\tau_{ij}^*$  is trivial since the cross-sectional dependence in  $u_{it}$  is preserved in  $u_{it}^*$ . For part (c), by using the fact that  $\{u_{it}^*\}$  is independent from  $\{u_{is}^*\}$  for  $t \neq s$

$$\begin{aligned} E^* \left| N^{-1/2} \sum_{i=1}^N [u_{is}^* u_{it}^* - E^*(u_{is}^* u_{it}^*)] \right|^2 &= N^{-1} \sum_{i=1}^N \sum_{j=1}^N Cov^*(u_{is}^* u_{it}^*, u_{js}^* u_{jt}^*), \\ &= N^{-1} \sum_{i=1}^N \sum_{j=1}^N E^*(u_{is}^* u_{it}^* u_{js}^* u_{jt}^*), \\ &= N^{-1} \sum_{i=1}^N \sum_{j=1}^N E^*(u_{is}^* u_{js}^*) E^*(u_{it}^* u_{jt}^*), \\ &\leq N^{-1} \sum_{i=1}^N \sum_{j=1}^N |\tau_{ij}^*|^2 = O_p(1), \end{aligned}$$

by using part (b) and the result follows. Finally, we show part (d). Since  $\{u_{it}^*\}$  is independent over time with mean zero and  $E(u_{it}^{*2}) = \sigma_i^2$  and  $\{F_t^*\}$  and  $\{u_{it}^*\}$  are independent, the central limit theorem holds with the covariance matrix

$$\begin{aligned} \Theta_i^* &= E^* \left[ T^{-1} \sum_{t=1}^T F_t^* F_t^{*'} u_{it}^{*2} \right] = \hat{\Phi}(L) E^* \left[ T^{-1} \sum_{t=1}^T e_t^* e_t^{*'} u_{it}^{*2} \right] \hat{\Phi}(L)', \\ &= H \Phi(L) H^{-1} E^* \left[ T^{-1} \sum_{t=1}^T e_t^* e_t^{*'} u_{it}^{*2} \right] H^{-1'} \Phi(L)' H' + o_p^*(1). \end{aligned}$$

Since

$$T^{-1} \sum_{t=1}^T e_t^* e_t^{*'} u_{it}^{*2} \xrightarrow{p} p \lim \left[ T^{-1} \sum_{t=1}^T \hat{e}_t \hat{e}_t' \hat{u}_{it}^2 \right] = \sigma_i^2 Q^{-1'} \Sigma_e Q^{-1} \text{ in probability,}$$

and it is shown that

$$\Theta_i^* = \sigma_i^2 Q^{-1'} \Phi(L) \Sigma_e \Phi(L)' Q^{-1},$$

where  $\sigma_i^2 \Phi(L) \Sigma_e \Phi(L)' = \Theta_i$ .

**Proof of Lemma BT-C:** For part (a),  $E^*(e_t^*) = 0$  is trivial since the proposed algorithms ensure that  $\{\hat{e}_t\}$  is demeaned. It also holds that  $E^* \|e_t^*\|^2 = O_p(1)$  and independence between  $e_t^*$  and  $e_s^*$  for  $s \neq t$  holds because  $\{e_t^*\}$  is an i.i.d. resampling of  $\{\hat{e}_t\}$  by construction. For part (b), the central limit theorem holds since  $\{e_t^*\}$  is an i.i.d. sequence. The covariance is given by

$$\begin{aligned} & T^{-1} \sum_{t=1}^T Z_t^* e_t^{*'} \xrightarrow{p} p \lim T^{-1} \sum_{t=1}^T \hat{Z}_t \hat{e}_t' \text{ in probability} \\ & = \left[ (I_p \otimes Q^{-1'}) \Sigma_Z (I_p \otimes Q^{-1}) \right] \otimes (Q^{-1'} \Sigma_e Q^{-1}). \end{aligned}$$

For part (c), by Assumption C(e), the roots of  $\det(I_r - \Phi_1 z - \Phi_2 z^2 - \dots - \Phi_p z^p) = 0$  are outside the unit circle. These roots are equivalent to the roots of

$$\det(H'^{-1}) \det(I_r - \Phi_1 z - \Phi_2 z^2 - \dots - \Phi_p z^p) \det(H') = 0,$$

because  $\det(H')$  and  $\det(H'^{-1})$  are nonzero. Hence, the roots are equivalent to those of

$$\det[H'^{-1}(I_r - \Phi_1 z - \Phi_2 z^2 - \dots - \Phi_p z^p)H'] = 0,$$

or those of

$$\det(I_r - H'^{-1} \Phi_1 H' z - H'^{-1} \Phi_2 H' z^2 - \dots - H'^{-1} \Phi_p H' z^p) = 0.$$

However, these are asymptotically equivalent to the roots of

$$\det(I_r - \hat{\Phi}_1 z - \hat{\Phi}_2 z^2 - \dots - \hat{\Phi}_p z^p) = 0,$$

as  $N, T \rightarrow \infty$  by (A.5).

**Acknowledgements:** The author acknowledges financial support for this work from Hitotsubashi Institute for Advanced Study, Hitotsubashi University. I am grateful to Silvia Ginçalves, Atsushi Inoue, Benoit Perron, Zhongjun Qu, Moto Shintani, and the participants of the 27th Annual Meeting of the Canadian Econometrics Study Group, the fifth CIREQ Time Series Conference, the 2011 North American Summer Meeting of the Econometric Society, and numerous seminars for their helpful comments. All remaining errors are mine.

## References

- Acconcia, A. and S. Simonelli, 2008, Interpreting aggregate fluctuations looking at sectors, *Journal of Economic Dynamics and Control* 32(9), 3009-3031.
- Ang, A. and M. Piazzesi, 2003, A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, *Journal of Monetary Economics* 50(4), 745-787.
- Bai, J., 2003, Inferential theory for factor models of large dimensions, *Econometrica* 71(1), 135-171.
- 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. Elias, 2005, Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach, *Quarterly Journal of Economics* 120(1), 387-422.
- Boivin, J. and M.P. Giannoni, 2006, DGSE models in a data-rich environment, NBER Working Paper No. 12772.
- Boivin, J., M.P. Giannoni, and I. Mihov, 2009, Sticky prices and monetary policy: evidence from disaggregated U.S. data, *American Economic Review* 99(1), 350-384.
- Boivin, J., M.P. Giannoni, and D. Stevanović, 2013, Dynamic effects of credit shocks in a data-rich environment, FRB of New York Staff Report No. 615.
- Efron, B. and R.J. Tibshirani, 1994, An introduction to the bootstrap. Chapman & Hall/CRC.
- Giannone, D., L. Reichlin, and L. Sala, 2005, Monetary policy in real time. NBER Macroeconomics Annual pp. 161-200, MIT Press.
- Gilchrist, S., V. Yankov, and E. Zakrajšek, 2009, Credit market shocks and economic fluctuations: evidence from corporate bond and stock markets, *Journal of Monetary Economics* 56(4), 471-493.
- Gonçalves, S. and B. Perron, 2014, Bootstrapping factor-augmented regression models, *Journal of Econometrics* 182, 156-173.

- Hall, P., 1992, The bootstrap and Edgeworth expansion, Springer-Verlag.
- Kilian, L., 1998, Small sample confidence intervals for impulse response functions, *Review of Economics and Statistics* 80(2), 218-230.
- Kilian, L., 1999, Finite-sample properties of percentile and percentile-t bootstrap confidence intervals for impulse responses, *Review of Economics and Statistics* 81(4), 652-660.
- Ludvigson, S.C. and S. Ng, 2009a, Macro factors in bond risk premia, *Review of Financial Studies* 22(12), 5027-5067.
- Ludvigson, S.C. and S. Ng, 2009b, A factor analysis of bond risk premia. Handbook of Empirical Economics and Finance, A. Ulah and D. Giles Eds. pp. 313-372, Chapman and Hall.
- 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.
- McCracken, M.W. and S. Ng, 2015, FRED-MD: a monthly database for macroeconomic research, Federal Reserve Bank of St. Louis Working Paper Series 2015-012B.
- Moench, E., 2008, Forecasting the yield curve in a data-rich environment: a no-arbitrage factor-augmented VAR approach, *Journal of Econometrics* 146, 26-43.
- Sargent, T.J. and C.A. Sims, 1977, Business cycle modeling without pretending to have too much a priori economic theory, FRB Working Paper 55.
- Shintani, M. and Z-Y. Guo, 2015, Improving the finite sample performance of autoregression estimators in dynamic factor models: a bootstrap approach, *Econometric Reviews*.
- Stock, J.H., M.W. Watson, 2005, Implications of dynamic factor models for VAR analysis. NBER Working Paper 11467.
- Stock, J.H., M.W. Watson, 2010, Dynamic factor models. Oxford Handbook of Economic Forecasting, M.P. Clements and D.F. Hendry Eds. Oxford University Press.
- Yamamoto, Y., 2015, Asymptotic inference for common factor models in the presence of jumps, HIAS Discussion Paper series E-4.

**Table 1. Coverage properties of IRFs (Gaussian, 95% level)**

ID1: recursive restriction

		Coverage Ratio (%)						Length of C.I. (Median)						
	$T$	$N$	$h=0$	1	2	3	4	5	$h=0$	1	2	3	4	5
Proc. A	40	50	92.6	94.1	90.5	94.1	94.4	95.3	1.18	1.41	1.02	0.75	0.52	0.37
	40	200	93.2	93.4	90.6	93.5	93.9	94.8	1.17	1.40	1.02	0.75	0.54	0.38
	120	50	91.0	95.8	94.0	92.7	92.6	93.3	0.72	0.84	0.61	0.38	0.23	0.14
	120	200	91.5	95.9	93.7	93.0	93.1	93.7	0.72	0.85	0.63	0.40	0.24	0.15
Proc. B	40	50	90.8	89.6	86.7	92.5	94.8	95.8	1.09	1.30	0.92	0.66	0.47	0.33
	40	200	92.0	88.3	86.4	92.8	95.1	95.9	1.09	1.29	0.91	0.66	0.47	0.34
	120	50	88.5	92.7	89.9	89.1	90.5	92.8	0.67	0.78	0.56	0.35	0.21	0.12
	120	200	89.8	92.1	89.0	89.6	91.0	92.4	0.68	0.78	0.57	0.36	0.22	0.13
Normal	40	50	93.8	92.0	86.8	84.6	85.0	83.9	1.24	1.36	0.88	0.49	0.26	0.13
	40	200	94.3	91.4	87.4	85.1	84.9	84.4	1.24	1.35	0.87	0.49	0.26	0.14
	120	50	90.6	93.7	92.0	90.0	88.7	87.9	0.71	0.80	0.55	0.31	0.16	0.08
	120	200	91.3	93.4	91.8	90.0	88.9	88.4	0.71	0.79	0.57	0.33	0.17	0.09

ID2: short-run restriction

		Coverage Ratio (%)						Length of C.I. (Median)						
	$T$	$N$	$h=0$	1	2	3	4	5	$h=0$	1	2	3	4	5
Proc. A	40	50	85.7	97.2	96.7	97.1	97.0	97.5	0.68	2.44	1.25	0.70	0.40	0.24
	40	200	88.3	97.8	97.3	97.7	97.6	98.0	0.64	2.28	1.30	0.79	0.47	0.29
	120	50	90.2	98.0	97.0	96.4	95.7	95.5	0.41	0.72	0.44	0.24	0.12	0.06
	120	200	92.8	98.8	98.6	98.3	98.1	98.1	0.37	0.64	0.46	0.28	0.16	0.09
Proc. B	40	50	83.8	96.8	95.4	96.3	95.7	96.7	0.54	1.16	0.70	0.44	0.27	0.17
	40	200	85.7	96.8	96.2	97.2	97.0	97.2	0.53	1.22	0.84	0.57	0.37	0.25
	120	50	85.9	95.3	92.1	91.2	92.4	92.7	0.34	0.49	0.32	0.18	0.10	0.05
	120	200	91.2	97.0	96.2	95.9	96.2	96.5	0.34	0.50	0.37	0.23	0.13	0.08
Normal	40	50	98.5	89.7	85.6	81.0	78.4	76.1	1.54	0.71	0.35	0.17	0.07	0.03
	40	200	99.4	92.1	87.7	85.0	83.1	82.8	1.68	0.76	0.40	0.20	0.10	0.05
	120	50	98.6	89.5	81.5	77.6	76.1	74.9	0.99	0.42	0.21	0.10	0.04	0.02
	120	200	99.5	93.8	87.8	84.7	82.9	82.2	1.06	0.45	0.23	0.12	0.06	0.03

ID3: long-run restriction

		Coverage Ratio (%)						Length of C.I. (Median)						
	$T$	$N$	$h=0$	1	2	3	4	5	$h=0$	1	2	3	4	5
Proc. A	40	50	86.3	96.2	96.1	96.7	96.8	97.4	2.94	2.95	1.54	0.87	0.48	0.28
	40	200	84.9	96.5	96.6	97.1	97.2	98.0	2.84	2.62	1.57	0.96	0.57	0.34
	120	50	91.2	97.2	96.4	95.8	95.3	95.3	0.95	0.72	0.47	0.26	0.14	0.07
	120	200	88.1	97.3	97.7	97.4	97.2	97.4	0.84	0.59	0.47	0.29	0.16	0.09
Proc. B	40	50	87.8	96.1	92.8	95.8	95.5	97.6	1.88	1.15	0.71	0.47	0.28	0.17
	40	200	87.1	95.4	92.9	95.4	96.3	97.5	2.00	1.13	0.84	0.58	0.38	0.25
	120	50	91.7	95.2	88.8	88.2	89.7	91.3	0.79	0.42	0.29	0.17	0.09	0.05
	120	200	91.1	95.2	92.8	92.9	93.9	95.2	0.81	0.40	0.33	0.21	0.13	0.07
Normal	40	50	97.2	91.8	87.7	82.5	80.0	77.2	1.91	0.81	0.38	0.18	0.08	0.04
	40	200	98.0	94.6	90.3	87.0	85.6	84.0	1.80	0.79	0.41	0.21	0.10	0.05
	120	50	95.2	93.5	84.7	80.2	78.0	76.2	1.11	0.45	0.22	0.10	0.05	0.02
	120	200	97.9	96.6	91.4	87.1	84.7	83.7	1.09	0.46	0.23	0.12	0.06	0.03

**Table 2. Coverage properties of IRFs (Chi, 95% level)**

ID1: recursive restriction

		Coverage Ratio (%)						Length of C.I. (Median)					
		$h=0$	1	2	3	4	5	$h=0$	1	2	3	4	5
Proc. A	$T=40, N=50$	90.7	93.8	91.8	93.5	94.0	94.3	1.69	1.45	1.02	0.72	0.50	0.34
	$T=40, N=200$	90.2	94.9	92.5	94.5	94.8	95.4	1.70	1.43	1.02	0.74	0.51	0.36
	$T=120, N=50$	94.1	96.0	93.9	93.0	93.0	93.5	1.17	0.90	0.62	0.39	0.23	0.13
	$T=120, N=200$	94.7	96.4	94.6	93.4	93.7	94.1	1.17	0.91	0.63	0.39	0.23	0.14
Proc. B	$T=40, N=50$	88.3	89.5	87.7	92.5	94.8	95.5	1.50	1.31	0.92	0.65	0.45	0.32
	$T=40, N=200$	87.5	90.5	87.9	93.4	95.7	96.7	1.53	1.31	0.92	0.66	0.46	0.33
	$T=120, N=50$	91.4	92.9	90.2	89.3	90.3	92.1	1.09	0.84	0.57	0.35	0.21	0.12
	$T=120, N=200$	92.0	92.4	89.9	89.9	91.1	93.1	1.09	0.84	0.57	0.36	0.21	0.13
Normal	$T=40, N=50$	85.6	91.0	88.0	86.4	85.9	84.6	1.34	1.35	0.87	0.48	0.25	0.12
	$T=40, N=200$	84.8	91.6	88.2	87.3	87.2	86.1	1.33	1.34	0.87	0.48	0.26	0.13
	$T=120, N=50$	80.8	92.1	92.0	90.7	89.1	87.7	0.77	0.80	0.56	0.31	0.16	0.08
	$T=120, N=200$	80.3	91.7	91.8	90.9	89.4	88.9	0.78	0.80	0.56	0.32	0.16	0.08

ID2: short-run restriction

		Coverage Ratio (%)						Length of C.I. (Median)					
		$h=0$	1	2	3	4	5	$h=0$	1	2	3	4	5
Proc. A	$T=40, N=50$	78.9	95.3	95.0	95.9	95.5	96.7	0.99	3.45	1.39	0.74	0.38	0.22
	$T=40, N=200$	82.2	96.8	96.1	96.7	96.9	97.9	0.97	3.75	1.78	0.99	0.54	0.33
	$T=120, N=50$	84.8	95.6	94.4	93.9	93.3	93.2	0.72	2.18	0.95	0.42	0.19	0.09
	$T=120, N=200$	90.9	97.1	96.4	96.2	95.9	96.0	0.65	1.67	0.91	0.46	0.23	0.11
Proc. B	$T=40, N=50$	77.4	95.8	94.7	96.2	94.3	95.9	0.74	1.88	0.80	0.47	0.24	0.14
	$T=40, N=200$	82.8	96.5	96.0	97.3	96.7	97.5	0.73	1.74	0.99	0.60	0.34	0.21
	$T=120, N=50$	79.9	92.2	86.7	86.5	86.4	87.1	0.50	0.70	0.35	0.18	0.09	0.04
	$T=120, N=200$	90.6	95.9	94.7	94.7	94.8	95.0	0.50	0.63	0.43	0.25	0.14	0.07
Normal	$T=40, N=50$	95.9	86.5	84.5	77.8	74.2	70.2	2.01	0.91	0.37	0.15	0.06	0.02
	$T=40, N=200$	97.6	91.3	86.9	82.4	79.9	77.3	1.93	0.84	0.39	0.18	0.08	0.03
	$T=120, N=50$	94.3	83.5	76.2	72.2	69.9	68.1	1.19	0.49	0.22	0.09	0.04	0.02
	$T=120, N=200$	99.0	93.3	85.9	82.2	80.7	79.4	1.18	0.49	0.24	0.11	0.05	0.02

ID3: long-run restriction

		Coverage Ratio (%)						Length of C.I. (Median)					
		$h=0$	1	2	3	4	5	$h=0$	1	2	3	4	5
Proc. A	$T=40, N=50$	81.9	94.5	94.8	96.0	96.1	97.9	3.77	4.47	1.74	0.95	0.48	0.27
	$T=40, N=200$	84.1	96.2	96.1	96.9	97.2	98.1	4.31	4.81	2.22	1.29	0.71	0.42
	$T=120, N=50$	89.2	95.4	94.8	94.7	94.3	94.5	2.61	2.73	1.20	0.54	0.24	0.11
	$T=120, N=200$	88.7	96.5	96.3	96.2	96.0	96.3	2.11	1.96	1.09	0.56	0.28	0.14
Proc. B	$T=40, N=50$	83.6	95.3	93.7	96.7	94.9	97.6	2.97	2.15	0.88	0.55	0.26	0.15
	$T=40, N=200$	86.0	96.1	94.2	96.7	97.0	98.2	2.73	1.66	1.04	0.65	0.36	0.22
	$T=120, N=50$	87.5	92.1	84.5	85.2	85.7	86.4	1.20	0.69	0.36	0.18	0.09	0.04
	$T=120, N=200$	90.3	94.5	92.3	92.2	93.0	94.0	1.07	0.52	0.41	0.25	0.14	0.08
Normal	$T=40, N=50$	94.9	89.6	86.6	79.8	76.2	71.6	2.77	1.13	0.44	0.18	0.07	0.03
	$T=40, N=200$	96.8	94.2	89.5	85.2	82.0	78.7	2.35	0.96	0.44	0.20	0.09	0.04
	$T=120, N=50$	92.3	88.7	80.2	75.1	72.3	70.3	1.54	0.58	0.24	0.10	0.04	0.02
	$T=120, N=200$	97.5	96.2	89.2	84.9	82.3	80.7	1.31	0.53	0.25	0.12	0.05	0.02

**Table 3. Coverage properties of IRFs (Hetero, 95% level)**

ID1: recursive restriction

		Coverage Ratio (%)						Length of C.I. (Median)					
<i>T</i>	<i>N</i>	<i>h=0</i>	1	2	3	4	5	<i>h=0</i>	1	2	3	4	5
Proc. A	40 50	93.6	94.3	91.1	93.5	94.5	95.1	1.20	1.42	1.05	0.75	0.54	0.39
	40 200	93.5	93.8	91.3	94.0	94.2	94.5	1.18	1.40	1.02	0.74	0.53	0.38
	120 50	91.7	96.1	93.8	92.5	92.1	92.6	0.72	0.84	0.61	0.38	0.23	0.14
	120 200	92.3	96.3	94.1	92.6	92.5	93.1	0.73	0.85	0.62	0.39	0.24	0.14
Proc. B	40 50	91.6	90.4	86.3	92.2	94.6	95.6	1.11	1.31	0.93	0.68	0.50	0.36
	40 200	91.2	90.5	86.7	92.7	94.3	95.2	1.09	1.30	0.93	0.67	0.48	0.34
	120 50	88.6	92.6	90.1	90.0	90.6	92.3	0.67	0.77	0.56	0.35	0.21	0.13
	120 200	89.7	92.7	90.7	89.9	91.5	93.0	0.68	0.79	0.57	0.36	0.22	0.13
Normal	40 50	94.8	92.9	86.9	85.0	84.8	84.4	1.25	1.37	0.89	0.51	0.28	0.15
	40 200	94.8	92.6	87.4	85.6	85.5	84.9	1.25	1.35	0.88	0.50	0.27	0.14
	120 50	91.1	93.5	92.2	90.8	89.0	88.5	0.71	0.79	0.56	0.31	0.17	0.08
	120 200	92.0	94.1	92.9	91.4	90.4	89.6	0.72	0.80	0.57	0.32	0.17	0.09

ID2: short-run restriction

		Coverage Ratio (%)						Length of C.I. (Median)					
<i>T</i>	<i>N</i>	<i>h=0</i>	1	2	3	4	5	<i>h=0</i>	1	2	3	4	5
Proc. A	40 50	86.0	97.8	96.7	97.2	96.1	97.2	0.63	1.73	0.94	0.55	0.33	0.20
	40 200	87.0	98.5	98.0	98.2	97.7	98.0	0.57	1.73	1.15	0.74	0.46	0.31
	120 50	92.0	98.3	96.1	94.9	94.2	94.1	0.39	0.59	0.35	0.18	0.09	0.05
	120 200	91.6	98.5	98.0	97.8	97.6	97.3	0.36	0.60	0.42	0.25	0.14	0.08
Proc. B	40 50	82.5	96.2	94.7	96.0	94.8	96.4	0.56	1.22	0.72	0.44	0.27	0.17
	40 200	86.7	97.4	96.6	97.2	97.2	97.8	0.55	1.28	0.87	0.58	0.37	0.25
	120 50	84.6	95.1	92.0	91.9	92.5	93.0	0.35	0.48	0.31	0.18	0.10	0.05
	120 200	90.0	97.2	96.3	96.2	96.4	96.3	0.35	0.51	0.37	0.23	0.13	0.07
Normal	40 50	98.8	88.5	83.9	79.0	77.1	75.2	1.61	0.72	0.34	0.16	0.07	0.03
	40 200	99.6	91.6	88.1	84.8	83.4	82.6	1.75	0.77	0.40	0.20	0.10	0.05
	120 50	98.1	89.1	81.3	76.9	74.2	73.3	0.98	0.41	0.20	0.09	0.04	0.02
	120 200	99.5	93.1	86.4	83.4	81.5	80.8	1.06	0.44	0.22	0.11	0.06	0.03

ID3: long-run restriction

		Coverage Ratio (%)						Length of C.I. (Median)					
<i>T</i>	<i>N</i>	<i>h=0</i>	1	2	3	4	5	<i>h=0</i>	1	2	3	4	5
Proc. A	40 50	84.9	97.3	96.2	97.2	96.1	97.6	2.18	2.17	1.22	0.73	0.41	0.25
	40 200	82.9	97.7	96.8	97.3	97.4	98.1	2.25	1.98	1.41	0.92	0.58	0.37
	120 50	90.2	97.9	94.9	93.8	93.1	92.2	0.79	0.58	0.35	0.19	0.10	0.05
	120 200	87.5	97.9	97.6	97.3	97.5	97.3	0.78	0.54	0.41	0.25	0.14	0.08
Proc. B	40 50	88.3	96.1	92.0	95.6	95.1	97.4	1.95	1.22	0.73	0.48	0.28	0.17
	40 200	86.1	95.9	93.6	95.4	96.7	97.8	2.16	1.23	0.91	0.63	0.41	0.27
	120 50	91.7	95.6	89.6	88.5	89.5	90.9	0.77	0.41	0.28	0.16	0.09	0.05
	120 200	91.0	95.6	93.6	93.4	94.5	95.7	0.80	0.39	0.33	0.21	0.13	0.07
Normal	40 50	97.2	92.0	87.1	82.2	79.1	76.5	1.90	0.80	0.38	0.18	0.08	0.04
	40 200	98.0	94.5	91.0	87.8	86.1	84.3	1.85	0.81	0.42	0.21	0.11	0.05
	120 50	95.4	93.8	84.8	80.3	77.2	76.0	1.08	0.43	0.21	0.10	0.04	0.02
	120 200	98.1	96.8	90.4	87.0	84.9	82.8	1.12	0.47	0.24	0.12	0.06	0.03

**Table 4. Coverage properties of IRFs under small  $N$  (Gaussian, 95% level)**

ID1: recursive restriction

		Coverage Ratio (%)						Length of C.I. (Median)					
$T$	$N$	$h=0$	1	2	3	4	5	$h=0$	1	2	3	4	5
Proc. A	120 10	91.3	96.2	93.0	91.2	90.5	90.8	0.73	0.86	0.61	0.38	0.22	0.13
	120 30	91.8	94.9	93.0	92.1	92.0	92.6	0.72	0.85	0.61	0.38	0.23	0.14
	240 10	84.0	96.3	94.7	92.2	90.7	89.7	0.52	0.60	0.44	0.26	0.14	0.08
	240 30	85.6	96.2	94.7	93.0	91.7	91.4	0.53	0.60	0.44	0.26	0.15	0.08
Proc. B	120 10	87.6	92.1	88.6	87.6	88.5	90.2	0.67	0.79	0.57	0.35	0.20	0.12
	120 30	88.9	91.2	88.6	88.8	90.2	91.7	0.67	0.78	0.56	0.35	0.21	0.12
	240 10	79.8	92.3	91.1	89.3	87.7	87.4	0.48	0.56	0.40	0.24	0.13	0.07
	240 30	81.2	91.9	91.1	90.0	88.6	89.3	0.49	0.56	0.41	0.24	0.14	0.07
Normal	120 10	91.0	93.3	90.9	88.5	87.0	86.1	0.71	0.80	0.56	0.31	0.16	0.08
	120 30	92.0	92.9	91.4	89.8	88.4	87.6	0.71	0.79	0.56	0.32	0.17	0.08
	240 10	82.5	93.3	92.7	91.1	89.8	88.5	0.50	0.57	0.40	0.23	0.12	0.06
	240 30	83.1	92.6	92.8	91.6	90.6	89.4	0.51	0.57	0.41	0.23	0.12	0.06

ID2: short-run restriction

		Coverage Ratio (%)						Length of C.I. (Median)					
$T$	$N$	$h=0$	1	2	3	4	5	$h=0$	1	2	3	4	5
Proc. A	120 10	67.3	95.2	89.4	86.9	84.3	83.6	0.56	0.87	0.38	0.17	0.08	0.03
	120 30	87.1	97.4	95.4	94.3	93.6	92.9	0.45	0.76	0.43	0.22	0.11	0.05
	240 10	57.9	93.3	80.7	75.5	72.3	70.8	0.43	0.46	0.21	0.09	0.04	0.02
	240 30	83.6	96.0	91.7	88.7	86.6	85.9	0.32	0.39	0.22	0.11	0.06	0.03
Proc. B	120 10	50.4	88.4	74.3	74.7	74.3	75.8	0.33	0.45	0.21	0.10	0.05	0.02
	120 30	79.5	92.8	87.0	86.7	87.6	88.5	0.34	0.47	0.28	0.15	0.08	0.04
	240 10	37.5	79.1	60.0	56.2	56.3	57.8	0.24	0.29	0.13	0.06	0.03	0.01
	240 30	73.6	90.0	80.5	77.7	78.3	78.7	0.24	0.29	0.18	0.09	0.04	0.02
Normal	120 10	84.2	78.0	68.1	62.7	59.8	57.9	0.82	0.39	0.17	0.07	0.03	0.01
	120 30	97.0	86.3	75.9	72.3	70.4	69.6	0.94	0.40	0.19	0.09	0.04	0.02
	240 10	76.7	69.7	55.0	48.4	46.7	45.7	0.58	0.27	0.11	0.05	0.02	0.01
	240 30	96.0	84.7	71.2	66.2	63.7	61.4	0.68	0.29	0.13	0.06	0.03	0.01

ID3: long-run restriction

		Coverage Ratio (%)						Length of C.I. (Median)					
$T$	$N$	$h=0$	1	2	3	4	5	$h=0$	1	2	3	4	5
Proc. A	120 10	80.9	95.7	90.1	88.3	86.0	85.9	1.15	0.99	0.43	0.19	0.08	0.04
	120 30	90.7	96.8	94.6	94.0	93.4	93.1	1.03	0.80	0.47	0.24	0.12	0.06
	240 10	69.9	93.5	79.7	74.8	72.5	71.5	0.72	0.49	0.22	0.10	0.04	0.02
	240 30	86.5	96.4	91.1	87.8	86.1	85.7	0.55	0.37	0.23	0.12	0.06	0.03
Proc. B	120 10	68.8	89.9	68.8	69.2	69.2	71.7	0.70	0.45	0.20	0.10	0.04	0.02
	120 30	89.6	93.5	83.4	82.7	83.7	86.3	0.77	0.43	0.27	0.14	0.08	0.04
	240 10	51.1	82.2	55.3	51.3	51.5	53.4	0.44	0.29	0.12	0.06	0.02	0.01
	240 30	84.6	91.3	78.2	74.4	74.3	75.5	0.47	0.26	0.16	0.08	0.04	0.02
Normal	120 10	77.0	82.8	70.6	65.2	61.7	59.3	1.08	0.43	0.18	0.07	0.03	0.01
	120 30	92.9	90.4	79.8	74.9	71.9	70.1	1.10	0.43	0.20	0.09	0.04	0.02
	240 10	64.3	75.1	57.6	50.0	47.4	46.5	0.72	0.28	0.12	0.05	0.02	0.01
	240 30	89.3	91.0	76.4	68.3	64.8	62.9	0.81	0.31	0.14	0.06	0.03	0.01

**Table 5-a. Coverage rate of confidence intervals for the calibrated US macroeconomic model (95% level)**

	<i>Proc</i>	<i>h=0</i>	1	2	3	4	5	6	7
Inflation	A	90.8	92.0	91.7	92.1	92.2	91.6	90.9	90.3
	B	89.6	91.6	90.9	91.3	91.6	91.4	91.2	90.9
Short rate	A	90.6	90.4	90.1	91.1	92.3	92.2	92.1	91.9
	B	71.4	80.2	76.3	82.5	85.4	87.8	89.6	90.9
Production	A	97.0	95.8	97.4	96.4	96.0	95.0	94.0	92.7
	B	79.7	91.2	84.2	88.5	87.2	88.0	87.7	87.9
Unemployment rate	A	96.5	95.6	97.6	96.6	96.4	95.3	94.2	92.9
	B	86.3	92.2	88.4	91.0	89.9	90.5	89.4	89.4

**Table 5-b. Length of confidence intervals for the calibrated US macroeconomic model (95% level)**

	<i>Proc</i>	<i>h=0</i>	1	2	3	4	5	6	7
Inflation	A	7.71	4.78	4.90	4.26	3.97	3.67	3.43	3.22
	B	7.34	4.54	4.70	4.11	3.86	3.62	3.41	3.23
Short rate	A	8.87	6.50	6.28	5.89	5.72	5.53	5.39	5.21
	B	5.39	4.56	4.18	4.38	4.58	4.79	4.93	5.01
Production	A	16.55	13.84	14.00	13.63	14.72	14.75	15.01	14.86
	B	10.80	12.59	10.38	11.79	12.42	12.98	13.32	13.38
Unemployment rate	A	13.69	11.55	11.25	11.10	11.81	11.86	12.11	11.94
	B	10.29	10.75	9.12	10.07	10.49	10.91	11.12	11.12