Received: 27 May 2016 | Revised: 12 June 2018
DOI: 10.1002/jae.2659

RESEARCH ARTICLE

WILEY Journal of APPLIED ECONOMETRICS

Bootstrap inference for impulse response functions in factor-augmented vector autoregressions

Yohei Yamamoto

Department of Economics, Hitotsubashi University, Tokyo, Japan

Correspondence
Yohei Yamamoto, Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan.
Email: yohei.yamamoto@econ.hit-u.ac.jp

Summary

In this study, we consider residual-based bootstrap methods 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). Both procedures are asymptotically valid under the condition $\sqrt{T}/N \to 0$, where N and T are the cross-sectional dimension and the time dimension, respectively. However, Procedure A is also valid even when $\sqrt{T}/N \to c$ with $0 \leq c < \infty$ because it accounts for the effect of the factor estimation errors on the impulse response function estimator. Our simulation results suggest that Procedure A achieves more accurate coverage rates than those of Procedure B, especially when N is much smaller than T. In the monetary policy analysis of Bernanke et al. (*Quarterly Journal of Economics*, 2005, *120*(1), 387–422), the proposed methods can produce statistically different results.

1 | INTRODUCTION

Factor-augmented vector autoregressions (FAVARs), introduced by Bernanke, Boivin, and Eliasz (2005) and further explored by Stock and Watson (2005, 2011), 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, FAVARs provide a suitable specification if certain concepts in economic models, such as productivity and inflation, are better captured by unobservable factors measured using multiple indicators, rather than by a single specific series (e.g., Boivin & Giannoni, 2006; Sargent & Sims, 1977). As a result, the literature on the empirical applications of FAVARs is rapidly growing in various fields. A current but noncomprehensive list includes the works of Ang and Piazzesi (2003), Giannone, Reichlin, and Sala (2005), Boivin, Giannoni, and Mihov (2009), Boivin, Giannoni, and Stevanović (2013), Acconcia and Simonelli (2008), Moench (2008), Ludvigson and Ng (2009a, 2009b), Gilchrist, Yankov, and Zakrajšek (2009), and Forni and Gambetti (2010), among others. For further discussion, see Kilian and Lütkepohl (2017, chapter 16).

In this study, we discuss methods to construct the confidence interval for structural impulse response functions (IRFs) in FAVARs when the unobservable factors are estimated using the two-step principal component (PC) method. The seminal work by Bai and Ng (2006) developed asymptotic inferential methods for the coefficients in factor-augmented regression models. They showed that under certain conditions, including $\sqrt{T}/N \to 0$ as $N, T \to \infty$, where N and T are the cross-sectional dimension and the time dimension, respectively, the effect of the factor estimation errors on the coefficient estimators was negligible. This result implies that the PC factor estimate can replace the true factors, and the asymptotic confidence interval is constructed in the usual manner. Furthermore, this approach is straightforward to apply to IRFs because they are a function of the model coefficients. For example, Bai, Li, and Lu (2016) proposed identification assumptions to consistently estimate the parameters and impulse responses in FAVARs—that is, to consistently estimate the random rotation and provide asymptotic distributions under $\sqrt{T}/N \to 0$.

In contrast, the effect of factor estimation errors on the coefficient estimators can be relevant when the data set involves N that is much smaller than T, such that the condition $\sqrt{T}/N \to 0$ is inappropriate. Bootstrap methods are an alternative to the asymptotic intervals of Bai and Ng (2006). In this context, the seminal works of Gonçalves and Perron (2014) and Shintani and Guo (2015) studied the theoretical properties of residual-based bootstrap intervals for factor-augmented regression models, showing their asymptotic validity under the more general framework of $\sqrt{T}/N \to c$, with $0 \leq c < \infty$. However, their models were reduced form. This study extends their analysis to structural VAR models. In particular, we consider recursively identified models as well as nonrecursively identified models and models identified by long-run restrictions (see Kilian & Lütkepohl, 2017, chapters 8–11). In doing so, we compare the confidence intervals produced by the method, à la Gonçalves and Perron (2014), and an alternative method that can also be used in empirical studies. The former method estimates the factors in the bootstrap replications and is referred to as Procedure A. The alternative method estimates the factors in the original data, but not in the bootstrap space, treating the original factor estimate as the observed processes. This method is referred to as Procedure B and can be implemented in a straightforward manner from the standard structural VAR. In doing so, we restrict the errors in the vector autoregressions to be independent and identically distributed (i.i.d.), and the idiosyncratic errors in the factor model to be cross-sectionally and serially independent.[1]

Our findings are summarized as follows. In theory, the original estimator of the structural impulse responses has the asymptotic bias $c\Delta$, say, with a nonzero constant Δ. Procedure A mimics the factor estimation errors and reproduces the bias $c\Delta^*$ in the bootstrap estimator. We show that Δ^* is asymptotically equivalent to Δ. In contrast, Procedure B provides a bootstrap estimator that has no asymptotic bias. Hence the original and the bootstrap distributions match only when $\sqrt{T}/N \to 0$.[2] In conclusion, both procedures are asymptotically valid under the condition $\sqrt{T}/N \to 0$. However, Procedure A is also valid when $\sqrt{T}/N \to c$ with $0 \leq c < \infty$ because it accounts for the effect of the factor estimation errors. In finite samples, Procedure A achieves better coverage rates than those of Procedure B, especially when N is much smaller than T. The median length of the interval in Procedure B is shorter than that of Procedure A.

Finally, we provide an empirical example using the monetary policy analysis of Bernanke et al. (2005) to illustrate how their results are affected by considering the current findings. The bootstrap method employed by Bernanke et al. (2005) is essentially similar to Procedure B in this study, but has some methodological differences—as we elaborate on—that are questionable or minor. We observe that their results are largely consistent with the results from using the proposed methods; however, in some cases, the proposed methods provide a difference in the statistical significance and policy implications. For example, our methods provide upper confidence limits of the price response that do not cross the zero axis at longer horizons, implying that the price does not significantly decrease after a contractionary monetary policy shock. This caveat is strengthened for an empirical analysis with fewer cross-sections.

One of the shortcomings of this study is that we circumvent a new insight into identification strategies of the FAVAR models that include observed factors. We do not exclude this possibility; however, by following Bernanke et al. (2005), the observed and unobserved factors are restricted orthogonal, whereas the unobserved factors may be contemporaneously correlated. The statistical properties of the identified FAVAR models are more comprehensively discussed by Bai et al. (2016); however, their goal was to identify the reduced-form coefficients and impulse responses and not the structural ones. Further work in this direction would offer an interesting agenda.

The remainder 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 methods for the structural identification and asymptotic properties of the impulse response estimator. In Section 4, we propose the bootstrap inference procedures and investigate their asymptotic properties. Section 5 assesses the finite-sample properties of the suggested procedures using Monte Carlo simulations. Section 6 serves as an empirical application of our methods to the monetary policy analysis of Bernanke et al. (2005). Section 7 concludes the paper. The Supporting Information Appendices include all technical derivations.

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

[2] This is the same idea as the bootstrap test of Corradi and Swanson (2014) for factor loading stability.

Throughout the paper, we use the following notation. The Euclidean norm of vector x is denoted by $\|x\|$. For matrices, the vector-induced norm is used. The symbols $\xrightarrow{p}$ and $\xrightarrow{d}$ represent convergence in probability under the probability measure P and convergence in distribution, respectively. Symbols $O_p(\cdot)$ and $o_p(\cdot)$ are the order of convergence under probability 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 $\lim_{N,T\to\infty} P[P^*(|T^*| > M(\epsilon)) > \epsilon] = 0$. 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 $\delta = \min\left\{\sqrt{N}, \sqrt{T}\right\}$, L be the standard lag operator and Chol(X) denote 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, whereas 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:

$$X_t = \mu^s + \Lambda^s F_t^s + u_t^s, \quad t = 1, \cdots, T, \tag{1}$$

where X_t is an $N \times 1$ vector of the response variables and N is the number of equations. The large set of response variables X_t is assumed to be driven by much lower dimensional unobservable factors F_t^s $(r \times 1)$ with time-invariant unobservable factor loadings $\Lambda^s = [\lambda_1^s, \cdots, \lambda_N^s]'$ $(N \times r)$. Then, $u_t^s = [u_{1t}^s, \cdots, u_{Nt}^s]'$ is an $N \times 1$ vector of idiosyncratic errors and μ^s is an $N \times 1$ vector of constants. Of particular interest, the model allows the unobservable 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 A_j^s $(j = 1, \ldots, p)$, an $r \times 1$ constant vector v^s, and structural shocks e_t^s $(q \times 1)$ with a $r \times q$ $(r \geq q)$ matrix G, such that

$$F_t^s = v^s + \sum_{j=1}^{p} A_j^s F_{t-j}^s + G e_t^s, \tag{2}$$

where $E(e_t^s e_t^{s\prime}) = I_q$. This is called the static factor representation of the dynamic factor model, meaning that the factors F_t^s may include the lagged dynamic factors at time t and, if so, the number of shocks on the dynamic factors q can be strictly smaller than r. For simplicity, we focus on the case $q = r$ and $G = I_r$. When the variables are written without their associated t subscript, they denote the entire matrix of the data, for example, $X = [X_1, \ldots, X_T]'$ is a $T \times N$ matrix and $F^s = [F_1^s, \cdots, F_T^s]'$ is a $T \times r$ matrix. The goal is to establish a valid inference procedure on the structural impulse responses generated from these models. Given that invertibility is assured, the two models (Equations 1 and 2) can be rewritten in vector moving-average form, such that

$$X_t = \mu^s + \Lambda^s \Psi^s(L) e_t^s + u_t^s, \tag{3}$$

where $\Psi^s(L) \equiv \sum_{j=0}^{\infty} \Psi_j^s L^j$, with $\Psi_0^s = I_r$ and $\Psi^s(L) = \left[I_r - \sum_{j=1}^{p} A_j^s L^j\right]^{-1}$. Let the structural impulse responses of the variable X_t at time horizon h $(h = 0, 1, 2, \ldots)$ be Θ_h.[3] Then:

$$\Theta_h \equiv \frac{\partial X_{t+h}}{\partial e_t^{s\prime}} = \Lambda^s \Psi_h^s.$$

This class of models is called FAVARs, introduced by Bernanke et al. (2005) and explored further by Stock and Watson (2005). Bernanke et al. (2005) included an observable structural factor as well as unobservable factors in F_t^s. Under the assumption that the observable and unobservable structural factors do not contemporaneously affect each other, this extension does not add complications to the analysis and the results presented here still hold.[4] Second, in many empirical applications, accounting for individual serial correlation in X_{it} is desirable, such that X_t can be replaced by $\alpha(L)X_t$ in the previous models, where $\alpha(L) = [\alpha_1(L), \ldots, \alpha_N(L)]'$ and $\alpha_i(L) = 1 + \alpha_{i,1}L + \cdots + \alpha_{i,p_i}L^{p_i}$. However, consistency and

[3] In the following, the IRFs of the ith response variable are denoted by $\Theta_{i,h}$, a $1 \times r$ vector.
[4] Bai et al. (2016) assumed that the observable and unobservable factors were contemporaneously uncorrelated by imposing $\Omega_{\epsilon v} = 0$ in all of the three sets of their identification assumptions.

asymptotic normality of the coefficients $\alpha_{i,j}$ ($i = 1, \dots, N$, and $j = 1, \dots, p_i$) are obtained 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 two simple models (Equations (1) and 2).

Although these models can be estimated using the maximum likelihood principle, with respect to the identification assumptions, the maximum likelihood estimation for models of such a large dimension can be computationally demanding and needs stronger assumptions. Thus we consider instead the less computationally demanding PC estimation proposed in the seminal work of Bai and Ng (2006). To this end, we state the estimable reduced-form representations of Equations 1 and 2, following the conventional strategy in the VAR literature. We rewrite the models using an $r \times r$ invertible matrix B that captures the contemporaneous correlations among F_t^s, such that $F_t = BF_t^s$. It then follows that $\Lambda = \Lambda^s B^{-1}$, $A_j = BA_j^s B^{-1}$, $e_t = Be_t^s$, and $\upsilon = B\upsilon^s$. Because the constant and idiosyncratic errors in Equation 1 are not affected by the factor structure, $\mu = \mu^s$ and $u_t = u_t^s$. The reduced-form models are

$$X_t = \mu + \Lambda F_t + u_t, \tag{4}$$

$$F_t = \upsilon + \sum_{j=1}^{p} A_j F_{t-j} + e_t. \tag{5}$$

To simplify the notation, let $Z = [\iota, F_{(-1)}, F_{(-2)}, \dots, F_{(-p)}]$ be a $T \times (rp+1)$ matrix, where ι is a $T \times 1$ vector of ones and $F_{(-j)} = [F_{1-j}, \dots, F_{T-j}]'$, and let $A = [\upsilon, A_1, \dots, A_p]'$ be an $(rp+1) \times r$ matrix such that Equation 5 can equivalently be written as $F = ZA + e$. The constant terms in the models can be omitted if the data are demeaned. Hence we use the models without constant terms in the following analysis for notational brevity.[5]

Let the reduced-form impulse responses of X_t at horizon h ($h = 0, 1, 2, \dots$) be Φ_h. Given invertibility, Equations (4) and 5 can be rewritten as

$$X_t = \Lambda \Psi(L) e_t + u_t, \tag{6}$$

where $\Psi(L) \equiv \sum_{j=0}^{\infty} \Psi_j L^j$ with $\Psi_0 = I_r$ and $\Psi(L) = \left[I_r - \sum_{j=1}^{p} A_j L^j \right]^{-1}$. Then:

$$\Phi_h \equiv \frac{\partial X_{t+h}}{\partial e_t'} = \Lambda \Psi_h.$$

As in the standard structural VAR analysis, the structural impulse responses can be represented by

$$\Theta_h = \Phi_h B = \Lambda \Psi_h B, \tag{7}$$

using the reduced-form parameters and B.

The two-step PC estimation procedure is as follows. First, the factors are extracted from Equation 4 using the PC method to find the solution of

$$(\hat{F}, \hat{\Lambda}) = \arg\min_{F,\Lambda} \sum_{t=1}^{T} (X_t - \Lambda F_t)'(X_t - \Lambda F_t), \tag{8}$$

with normalization $\hat{F}'\hat{F}/T = I_r$. Second, we estimate the VAR equation (5) using $\hat{F}_t$ and the least squares method. When estimating models 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 as

$$H_{NT} = V_{NT}^{-1}(\hat{F}'F/T)(\Lambda'\Lambda/N), \tag{9}$$

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

2.2 | Assumptions

We introduce the standard regularity assumptions for the remainder of the analysis. Let the data-generating processes introduced in the previous subsection be defined on a probability space $(\Omega, \mathcal{F}, P)$ and let $M < \infty$ be a generic constant.

[5]The theoretical derivations presented 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 ensure that the residuals are demeaned in the bootstrap procedures. See Section 4.1.

Assumption 1.

(a) The loadings λ_i are deterministic and $\Lambda'\Lambda/N \to \Sigma_\Lambda$, where Σ_Λ is an $r \times r$ positive definite matrix.
(b) The factors specified by Equation 5 satisfy $F'F/T \xrightarrow{p} \Sigma_F$, where Σ_F is an $r \times r$ positive definite matrix.
(c) The eigenvalues of the $r \times r$ matrix $\Sigma_\Lambda \Sigma_F$ are distinct.
(d) u_{is} and e_t are mutually independent for all (i, s, t).

Assumption 2.

(a) $E(u_{it}) = 0$ and $E|u_{it}|^8 \leq M$, for all (i, t).
(b) $N^{-1}\sum_{i=1}^{N}\sum_{k=1}^{N}|\tau_{ik}| \leq M$, where $\tau_{ik} = E(u_{it}u_{kt})$. $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}|\gamma_{st}| \leq M$, where $\gamma_{st} = E(N^{-1}\sum_{i=1}^{N}u_{is}u_{it})$.
(c) For every (s, t), $E\left|N^{-1/2}\sum_{t=1}^{N}[u_{is}u_{it} - E(u_{is}u_{it})]\right|^4 \leq M$.
(d) For any i, $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}F_s u_{it}\left(N^{-1}\sum_{k=1}^{N}u_{ks}u_{kt}\right) = O_p(1)$.
(e) For any i, $T^{-1/2}\sum_{t=1}^{T}F_t u_{it} \xrightarrow{d} N(0, \Omega_{Fui})$, where $\Omega_{Fui} \equiv \mathrm{var}\left(T^{-1/2}\sum_{t=1}^{T}F_t u_{it}\right)$.
(f) For $j = 0, 1, \cdots, p$, $\frac{1}{NT}\sum_{t=j+1}^{T}\sum_{i=1}^{N}\sum_{k=1}^{N}\lambda_i\lambda_k' u_{it}u_{k,t-j} - \Gamma_j \xrightarrow{p} 0$, where

$$\Gamma_j \equiv \lim_{N,T\to\infty} T^{-1}\sum_{t=1}^{T}\Gamma_{t,j},$$

and

$$\Gamma_{t,j} \equiv \mathrm{cov}\left(N^{-1/2}\sum_{i=1}^{N}\lambda_i u_{it}, N^{-1/2}\sum_{k=1}^{N}\lambda_k u_{k,t-j}\right).$$

Assumption 3.

(a) e_t is an $r \times 1$ vector of i.i.d. random variables, with $E(e_t) = 0$, and $E(e_t e_t') = \Sigma_e$ is an $r \times r$ positive definite matrix.
(b) $E|e_{lt}|^4 \leq M$, for all $l = 1, \ldots, r$ and t.
(c) $T^{-1/2}\mathrm{vec}\left(Z'e\right) \xrightarrow{d} N(0, \Sigma_e \otimes \Sigma_Z)$, with $\Sigma_Z \equiv \mathrm{plim}_{T\to\infty} Z'Z/T$.
(d) $\sqrt{T}\mathrm{vech}\left(e'e/T - \Sigma_e\right) \xrightarrow{d} N(0, \Omega_{e0})$, where $\Omega_{e0} \equiv 2(D_r'D_r)^{-1}D_r'[\Sigma_e \otimes \Sigma_e]D_r(D_r'D_r)^{-1})$, with D_r being the duplication matrix of size r defined such that, for any $r \times r$ matrix X, $\mathrm{vec}(X) = D_r\mathrm{vech}(X)$.
(e) The roots of $\det(I_r - A_1 z - A_2 z^2 - \cdots - A_p z^p) = 0$ lie outside the unit circle.
(f) The $r \times r$ matrix B has full rank.

Assumptions 1 and 2 follow the standard literature on factor models, such as Bai and Ng (2006) and Gonçalves and Perron (2014). In particular, Assumption 1(c) guarantees the uniqueness of the limit of the factor rotation matrix and Assumption 1(d) ensures the independence of the factors and idiosyncratic errors, as in Bai and Ng (2006), which greatly simplifies the moment conditions without losing much substance in our context. Assumptions 2(b) and 2(c) allow for the idiosyncratic errors to have weak serial and cross-sectional correlations. Assumption 2(d) is a unique, high-level assumption because it facilitates the bias derivation in our theorem. We can show that this assumption holds if u_{it} has no cross-sectional and serial dependence. Assumption 3 consists of the basic assumptions in the VAR literature. In particular, we impose an i.i.d. property with a fixed covariance matrix on e_t in Assumption 3(a). Assumption 3(e) ensures a stable system and invertibility of the VAR model into the moving average form (Equation 6) via the Wold representation theorem. Finally, these assumptions are on the parameters and variables $\{F_t, \Lambda, A, u_t, e_t\}$ in reduced-form representation rather than on the structural entities $\left\{F_t^s, \Lambda^s, A^s, u_t^s, e_t^s\right\}$. This approach was adopted mainly for notational simplicity in the proofs. However, researchers may need to validate these assumptions for the structural counterparts. For example, Assumption 3(a) can be replaced by the same assumption on e_t^s, with a small notational change from Σ_e to $B^{-1}\Sigma_e B'^{-1}$. This assumption holds because $e_t^s = B^{-1}e_t$ and B is a full-rank matrix.

3 | STRUCTURAL INFERENCE

3.1 | Structural identification

Once the reduced-form models are estimated, structural parameter estimates can be obtained using the contemporaneous coefficient matrix B. Because of the well-known random rotation of the PC estimate, the conventional structural

VAR identification schemes do not simply go through with FAVARs. Therefore, we propose identification schemes by imposing restrictions on the structural parameters that account for the rotation. Indeed, these identification schemes are conceptually common in many existing structural VAR studies. To this effect, we introduce the following identification assumptions.

Assumption 4.

(a) The lag order p and the number of factors r are known.
(b) $E(e_t^s e_t^{s\prime}) = I_r$.

Assumption 5. We have either

(a) (recursive restriction) $Q'^{-1}B$ is an upper or lower triangular matrix and the signs of its diagonal elements are known; or
(b) (short-run restriction) The short-run impulse responses $\Theta_0 \equiv \Lambda^s = \begin{bmatrix} \Lambda^s_{1:r} \\ \Lambda^s_{r+1:N} \end{bmatrix}$ have $\Lambda^s_{1:r} = [\lambda_1^s, \cdots, \lambda_r^s]'$ as an $r \times r$ upper or lower triangular matrix with positive diagonal elements; or
(c) (long-run restriction) The long-run impulse responses $\Theta_{1:r,\infty} \equiv \Lambda^s_{1:r}\left(\sum_{j=0}^{\infty} \Psi_h^s\right)$ from the first to the rth observations are an $r \times r$ upper or lower triangular matrix with positive diagonal elements.

Assumption 4(a) excludes model uncertainty from the analysis and simplifies the identification and inference problems. Any attempt to relax this assumption should be practically relevant and of great interest. However, this problem is beyond the scope of this study.[6] Assumption 4(b) imposes the orthogonality of the structural shocks, which is common in the structural VAR literature. Note that Assumption 4(a) fixes the total number of parameters in the model and Assumption 4(b) imposes restrictions on the $\frac{r^2+r}{2}$ parameters.

Any of the three conditions in Assumption 5 imposes $\frac{r^2-r}{2}$ zeros on the model parameters. Hence they achieve the necessary order condition of $\frac{r^2-r}{2} + \frac{r^2+r}{2} = r^2$ parameter restrictions on the structural models. They also satisfy the sufficient condition, as we show in Theorem 1. Assumption 5(a) is similar to the popular recursive restriction in structural VARs because it imposes a recursive structure on an invertible matrix $Q'^{-1}B$. However, in FAVARs we do not restrict the matrix B itself, but consider its rotation $Q'^{-1}B$ because it is the estimate of $\hat{B}$ obtained by the Cholesky decomposition. As suggested in the literature, such as Lütkepohl (2005), if the sign of a main diagonal of $Q'^{-1}B$ is negative, then the signs of all elements in the same column are flipped.[7] In contrast, Assumption 5(b) provides a set of restrictions on the short-run structural impulse responses. This requires at least $r-1$ observable variables where the kth ($k \le r-1$) element is contemporaneously affected only by the first k factors. Assumption 5(c) works similarly but restricts the long-run impulse responses. The implication of the long-run impulse response restriction follows from, for example, Blanchard and Quah (1989). These two assumptions formalize the exact identification methods suggested by Stock and Watson (2005), and they are consistent with IRb of Bai et al. (2016), in which the shocks of the unobservable factors are uncorrelated and in which the submatrix of the factor loadings is triangular.

Note that the recursive restriction Assumption 5(a) includes the limit of the rotation matrix Q, and it is not straightforward to interpret restrictions on such a nonstructural entity. Hence, we further break down Assumption 5(a) into the following set of sufficient conditions to ascertain its feasibility.[8]

Assumption 6. (a′)
The following three restrictions imply

(i) Σ_F is diagonal;
(ii) Σ_A is diagonal; and
(iii) B is an upper or lower triangular matrix and the signs of $\lambda_i' B = \Theta_{i,0}$ are known.

[6]For example, Dufour and Stevanović (2013) discussed that when the factors are a linear combination of observables, their dynamics are represented, in general, by VARMA processes, rather than by finite-order VARs.
[7]This is because $B = \mathrm{Chol}(BB')$ only if B is an upper triangular matrix with positive diagonal elements.
[8]Lemma 5 in Supporting Information Appendix A provides the proof that Assumption 5(a′) implies Assumption 5(a). Note that the former is a set of sufficient conditions.

The first two parts of Assumption 5(a$'$) imply a statistical normalization that the reduced-form model involves orthogonal factors and loadings. Therefore, we are able to impose a recursive structure directly on B, as in a conventional structural VAR. Finally, the signs of the elements of $\lambda_i' B$ (i.e., the instantaneous structural impulse responses) can be deduced from the underlying economic model.

3.2 | Estimation of identified structural models

The following three schemes are simple to implement to obtain $\hat{B}$. Note that each estimation achieves a consistent estimator for $Q'^{-1}B$.

ID1 (short-run restriction: recursive model):

1. Obtain $\hat{B}$, such that

$$\hat{B} = \text{Chol}\left[\hat{e}'\hat{e}/T\right]; \tag{10}$$

2. Adjust the signs of $\hat{\Theta}_{i,h}$ $(h = 0, 1, \ldots)$ if $sign(\hat{\Theta}_{i,0})$ is not what was expected.

ID2 (short-run restriction: nonrecursive model):

1. Construct a short-run impulse response estimate for observations from 1 to r:

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

2. Obtain $\hat{B}$, such that

$$\hat{B} = \hat{\Phi}^{-1}_{1:r,0}\hat{\Theta}_{1:r,0}, \tag{11}$$

where $\hat{\Theta}_{1:r,0}$ is obtained in the previous step.

ID3 (long-run restriction):

1. Construct a long-run impulse response estimate for the observations from 1 to r:

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

with $\hat{\Phi}_{1:r,\infty} = \hat{\Lambda}_{1:r}\left[I_r - \sum_{j=1}^{p}\hat{A}_j\right]^{-1}$, where $\hat{A}_j$ for $j = 1, \ldots, p$ is a reduced-form VAR coefficient estimate.

2. Obtain $\hat{B}$, such that

$$\hat{B} = \hat{\Phi}^{-1}_{1:r,\infty}\hat{\Theta}_{1:r,\infty}, \tag{12}$$

where $\hat{\Theta}_{1:r,\infty}$ is obtained in the previous step.

The ID1 scheme achieves a consistent estimate for $Q'^{-1}B$ under Assumption 5(a$'$). Note that the second step is to normalize the signs of $Q'^{-1}B$, which are not directly known. However, they are deduced through the sign of the structural impulse responses $\Theta_{i,0}$ in practice.[9] In contrast, the ID2 and ID3 schemes achieve a triangular matrix in the Cholesky decomposition, where the diagonal elements are restricted to be positive in the first step, making sign normalization unnecessary. In the second step, the resulting $\hat{B}$ achieves a consistent estimate for $Q'^{-1}B$.

Under any of these identification assumptions and the two-step PC estimation method, we obtain the following theorems on the consistency and asymptotic normality of the structural parameters and impulse response estimators. We let $\hat{\lambda}_i^s = \hat{B}'\hat{\lambda}_i$, $\hat{A}^s = (I_p \otimes \hat{B}')\hat{A}\hat{B}'^{-1}$ and $\hat{\Theta}_{i,h}$ be a function of them.

Theorem 1. (Consistency of the structural parameters). *Suppose Assumptions 1–5 hold. If the reduced-form parameters are estimated by the two-step PC method and matrix B is estimated by ID1, ID2, or ID3, then* $\hat{\lambda}_i^s - \lambda_i^s \xrightarrow{p} 0$, $\hat{A}^s - A^s \xrightarrow{p} 0$, *and* $\hat{\Theta}_{i,h} - \Theta_{i,h} \xrightarrow{p} 0$, *uniformly in h, for any i as* $N, T \to \infty$.

[9]When we consider impulse responses of the same structural shock, we only fix the sign of a certain cross-section i at a certain horizon h (e.g., GDP at four periods after the shock) to determine whether the obtained signs of the entire cross-sections and the entire sequence are flipped. To see this, consider a shock vector $\xi^{(k)}$, an $r \times 1$ vector with one in the kth element and zero otherwise. Then, the structural impulse response to this shock is $\lambda_i'\Psi_h B\xi^{(k)}$. We now let $B^{(k)}$ be a B matrix with the kth column having the opposite signs. Then, it is shown that

$$\lambda_i'\Psi_h B\xi^{(k)} = -\lambda_i'\Psi_h B^{(k)}\xi^{(k)}.$$

i and for any h if a certain estimate has a fixed sign.

Theorem 2. (Asymptotic distribution of the structural impulse response estimator). *Suppose Assumptions 1–5 hold. If the reduced-form parameters are estimated by the two-step PC method and matrix B is estimated by ID1, ID2, or ID3, then*

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

for any i and uniformly in h as $T, N \to \infty$ and $\sqrt{T/N} \to c$ $(0 \leq c < \infty)$, provided $\partial\Theta_{i,h}/\partial\lambda_i \neq 0$, $\partial\Theta_{i,h}/\partial vec(A) \neq 0$, and $\partial\Theta_{i,h}/\partial vec(B) \neq 0$. The asymptotic bias is given by

$$\Delta_{\Theta ih} = \frac{\partial\Theta'_{i,h}}{\partial\lambda'_i}\Delta_{\lambda i} + \frac{\partial\Theta'_{i,h}}{\partial \mathrm{vec}(A)'}\Delta_A + \frac{\partial\Theta'_{i,h}}{\partial \mathrm{vec}(B)'}\Delta_B,$$

where

$$\Delta_{\lambda i} \equiv (Q'^{-1}\Sigma_\Lambda^{-1}\Gamma_0\Sigma_\Lambda^{-1}Q^{-1} + Q'^{-1}\Sigma_F\Gamma_0\Sigma_\Lambda^{-1}\Sigma_F^{-1}\Sigma_\Lambda^{-1}Q^{-1})\lambda_i,$$

$$\Delta_A \equiv vec\{[(I_p \otimes Q'^{-1})\Sigma_Z(I_p \otimes Q^{-1})]^{-1}[\Pi_3(I_p \otimes Q)AQ^{-1} + \Delta_3]\},$$

with

$$\Pi_3 \equiv \begin{bmatrix} \Delta_{3,0} & \Delta_{3,1} & \cdots & \Delta_{3,p-1} \\ \Delta'_{3,1} & \Delta_{3,0} & & \\ \vdots & & \ddots & \\ \Delta'_{3,p-1} & \Delta'_{3,p-2} & & \Delta_{3,0} \end{bmatrix},$$
$$\Delta_3 \equiv [\Delta'_{3,1}\Delta'_{3,2}\cdots\Delta'_{3,p}]',$$

and

$$\Delta_B \equiv \bar{C}(\Delta_{1,0} + Q'^{-1}A'\Pi_1 AQ^{-1} + \Delta'_1(I_p \otimes Q)AQ^{-1} + Q'^{-1}A'(I_p \otimes Q')\Delta_1),$$

with

$$\bar{C} \equiv \frac{\partial \mathrm{vec}(Q'^{-1}B)}{\partial \mathrm{vec}(Q'^{-1}\Sigma_e Q^{-1})'} = [L_r(I_{r^2} + K_{rr})(Q'^{-1}B \otimes I_r)L'_r]^{-1},$$
$$\Pi_1 \equiv \begin{bmatrix} \Delta_{1,0} & \Delta_{1,1} & \cdots & \Delta_{1,p-1} \\ \Delta'_{1,1} & \Delta_{1,0} & & \\ \vdots & & \ddots & \\ \Delta'_{1,p-1} & \Delta'_{1,p-2} & & \Delta_{1,0} \end{bmatrix},$$
$$\Delta_1 \equiv [\Delta'_{1,1}\Delta'_{1,2}\cdots\Delta'_{1,p}]',$$

where $\Delta_{1,j}$ and $\Delta_{3,j}$ are defined in Lemma 1. The asymptotic variance is given by

$$\Omega_{\Theta ih} = \frac{\partial\Theta'_{i,h}}{\partial\lambda'_i}\Omega_{\lambda i}\frac{\partial\Theta_{i,h}}{\partial\lambda_i} + \frac{\partial\Theta'_{i,h}}{\partial \mathrm{vec}(A)'}\Omega_A\frac{\partial\Theta_{i,h}}{\partial \mathrm{vec}(A)} + \frac{\partial\Theta'_{i,h}}{\partial \mathrm{vec}(B)'}\Omega_B\frac{\partial\Theta_{i,h}}{\partial \mathrm{vec}(B)},$$

where $\Omega_{\lambda i} \equiv Q'^{-1}\Omega_{Fui}Q^{-1}$,

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

and $\Omega_B \equiv \bar{C}\Omega_e\bar{C}'$, with

$$\Omega_e \equiv 2(D'_r D_r)^{-1}D'_r[Q'^{-1}\Sigma_e Q^{-1} \otimes Q'^{-1}\Sigma_e Q^{-1}]D_r(D'_r D_r)^{-1}.$$

Note that L_r is the elimination matrix of size r, and K_{rr} is the commutation matrix for the $r \times r$ matrix.

The following should be noted about the results. Theorem 1 implies that the structural parameters and impulse responses are consistently estimated to ensure they are free from the rotation. Second, despite the implication of Theorem 1, we conventionally use the expression in terms of the reduced-form parameters in Theorem 2, which facilitates constructing the covariance estimate in practice.

Finally, the structural impulse response estimator has an asymptotic bias $c\Delta_{\Theta ih}$ if N is much smaller than T, such that $\sqrt{T}/N \to c$ with $0 < c < \infty$, and the estimator has no bias if $\sqrt{T}/N \to 0$. This is consistent with the seminal works of Gonçalves and Perron (2014) and Shintani and Guo (2015). We newly find that the asymptotic bias of the impulse response estimator depends on the serial correlation in the idiosyncratic errors through Γ_j, defined in Assumption 2(f). This contrasts with the findings of Gonçalves and Perron (2014), who consider the coefficient estimator in the factor-augmented regression model outside the factor model to ensure that the bias depends on Γ_j only for $j = 0$. This is because the impulse response estimator includes VAR coefficients of the estimated factors. Intuitively, in the FAVAR model, both the response variable and the regressors involve factor estimation errors, and their product consists of an asymptotic bias because of the serial correlation in the idiosyncratic errors. In contrast, in the factor-augmented regression model of Gonçalves and Perron (2014), the response variable is observed and has no factor estimation errors, and no bias is generated. Shintani and Guo (2015) have a setting similar to ours, in which they consider an autoregressive coefficient of an estimated factor, but rule out the serial correlation in the idiosyncratic errors.

4 | BOOTSTRAP INFERENCE

4.1 | Algorithms

This section outlines the residual-based bootstrap algorithms to construct the confidence intervals for the structural impulse responses. To this end, we propose the i.i.d. bootstrap for the VAR residuals as assumed in Assumption 3(a). We need to impose a restriction that the idiosyncratic errors in the factor model are serially and cross-sectionally independent, although heteroskedasticity is allowed. This contrasts with the setting of Gonçalves and Perron (2014), in which the serial correlation in the idiosyncratic errors is irrelevant in the bootstrap. At the end of the next section, we provide remarks on the case of serially correlated idiosyncratic errors.

Assumption 7.

(a) u_{it} and u_{ks} are independent for all $i \neq k$ and for any (t, s).
(b) u_{it} and u_{is} are independent for all $t \neq s$ and for any i.
(c) $E\|F_t\|^{12} \leq M$, $E|u_{it}|^{12} \leq M$, and $E^*|\eta_{it}|^{4q} \leq M$, for some $q > 1$, and for all i and t.

Assumptions 7(a) and 7(b) impose no cross-sectional and no serial independence in the idiosyncratic errors, respectively. Assumption 7(c) strengthens the moment conditions to ensure the wild bootstrap procedures. In particular, we compare two alternative algorithms that are implemented in a straightforward manner in empirical studies. The first method estimates the unobservable factors in the bootstrap replications, as recommended in the literature. This is called Procedure A. The second method does not do so in the bootstrap loop, treating the original factor estimate as the observed processes. Therefore, the implementation is no more involved than standard small-scale structural VARs. This is called Procedure B. We first outline Procedure A as follows.

Procedure A: Bootstrap with factor estimation

1. Estimate the reduced-form models (Equations 4 and 5) using the two-step PC procedure, obtaining the parameter estimates $\hat{\Lambda}$, $\hat{A}$, and $\hat{B}$, as well as the residuals $\hat{u}_t$ and $\hat{e}_t$. Construct the structural impulse response estimate $\hat{\Theta}_{i,h}$.
2. Make sure that the VAR residuals $\{\hat{e}_t\}_{t=1}^T$ are demeaned in the time direction. Resample with replacements the residuals $\{\hat{e}_t\}_{t=1}^T$ as $r \times 1$ vectors in an i.i.d. fashion and label them $\left\{e_t^*\right\}_{t=1}^T$. Generate the bootstrapped sample F_t^* using $F_t^* = \sum_{j=1}^{p} \hat{A}_j F_{t-j}^* + e_t^*$, for $t = 1, \dots, T$.[10]
3. Make sure that the idiosyncratic residuals $\{\{\hat{u}_{it}\}_{t=1}^T\}_{i=1}^N$ are demeaned in both time and cross-sectional directions. For each $i = 1, \dots, N$:

   (a) if $\{u_{it}\}_{t=1}^T$ are homoskedastic, we propose the i.i.d. resampling of $\{\hat{u}_{kt}\}_{t=1}^T$ to obtain $\left\{u_{kt}^*\right\}_{t=1}^T$;

[10]The bias correction method discussed by Kilian (1998) is applied. The bias is estimated by taking the average of $\hat{A}_j^* - H^*\hat{A}_j H^{*-1}$ in another bootstrap loop in Procedure A, and by $\hat{A}_j^{**} - \hat{A}_j$ in Procedure B.

(b) if $\{u_{it}\}_{t=1}^{T}$ are heteroskedastic, we propose the wild bootstrap $u_{it}^{*} = \hat{u}_{it}\eta_{it}$, where $\eta_{it} \sim$ i.i.d.$(0, 1)$ is an external random variable to obtain $\left\{u_{it}^{*}\right\}_{t=1}^{T}$.

Generate the bootstrapped sample X^{*} from $X_t^{*} = \hat{\Lambda}F_t^{*} + u_t^{*}$, for $t = 1, \ldots, T$.

4. Using the bootstrapped sample X^{*}, estimate $(\hat{F}^{*}, \hat{\Lambda}^{*})$ following the first step of the PC procedure. Then, estimate the VAR equation (5) using $\hat{F}^{*}$ to obtain the bootstrapped estimates $\hat{A}^{*}$ and $\hat{B}^{*}$ with the same identification scheme as the original estimate. This yields the bootstrap estimate of the structural impulse response $\hat{\Theta}_{i,h}^{*}$.
5. Repeat steps 2–4 R times.
6. With a $r \times 1$ vector of shocks of interest ζ, store the recentered statistic $s_{i,h} \equiv \hat{\Theta}_{i,h}^{*}\zeta - \hat{\Theta}_{i,h}\zeta$. Sort the statistics and select the $100 \cdot \alpha/2$th and $100 \cdot (1 - \alpha/2)$th percentiles $[s_{i,h}^{(\alpha/2)}, s_{i,h}^{(1-\alpha/2)}]$. The resulting $100 \cdot (1 - \alpha)\%$ confidence interval for $\Theta_{i,h}\xi$ is $[\hat{\Theta}_{i,h}\zeta - s^{(1-\alpha/2)}, \hat{\Theta}_{i,h}\zeta - s^{(\alpha/2)}]$.

In Procedure A, the bootstrap sample X_t^{*} shares the same data-generating process as the original data X_t in steps 2 and 3. In step 4, the bootstrap impulse response estimator involves the same identification method as the original estimator, suggesting that the dispersions of the bootstrap estimator can mimic the sampling distribution of the original estimator.

Procedure B: Bootstrap without factor estimation. Procedure B is the same as Procedure A except for a change in step 4, which is formalized as follows:

4 Using the bootstrapped sample X^{*} and factors F^{*}, estimate $\hat{\Lambda}^{**}$, $\hat{A}^{**}$, and $\hat{B}^{**}$ using the same identification method as the original estimate. This yields the bootstrap estimate of the structural impulse response $\hat{\Theta}_{i,h}^{**}$.

Procedure B is a natural and simple extension of the method conducted in standard structural VAR because it generates the VAR variables as if they were observable. In both procedures, it is essential for the bootstrap residuals to mimic the property of the true error terms. In step 2, we consider the i.i.d. resampling for the VAR errors as assumed in the original ones $\{e_t\}_{t=1}^{T}$.[11] The cross-sectional dependence in e_t is preserved because we resample the residuals as $r \times 1$ vectors. In contrast, in step 3, heteroskedasticity in the idiosyncratic errors of the factor model can be replicated by the wild bootstrap procedure, as suggested by Gonçalves and Perron (2014). When the idiosyncratic errors are serially correlated, those of every cross-section must be replicated in principle. This is because the sampling distribution of the impulse response estimator contains the variance of $\hat{\lambda}_i$, which is affected by the serial correlation in u_{it} even when $c = 0$. When $c > 0$, its asymptotic bias is also affected by the serial correlation in u_{kt} for $k = 1, \ldots, N$ through $\hat{A}$ and $\hat{B}$.[12] Finally, the percentile interval in step 6 of Hall (1992) can be replaced by Efron's percentile method by storing $s_{i,h} \equiv \hat{\Theta}_{i,h}^{*}\zeta$ and constructing $[s_{i,h}^{(\alpha/2)}, s_{i,h}^{(1-\alpha/2)}]$ or the percentile-t interval in theory, although their finite-sample properties are to be further investigated.[13]

4.2 | Asymptotic validity

This section studies the asymptotic validity of the proposed procedures. For Procedure A, we obtain the following theorem.

Theorem 3. (Asymptotic validity of Procedure A)
Suppose Assumptions 1–7 hold. If the models (Equations 4 and 5) are estimated using the two-step PC method and matrix B is estimated by ID1, ID2, or ID3, and Procedure A is implemented with the i.i.d. or the wild bootstrap, then for any $r \times 1$ shock vector ζ

$$\sup_{x \in \boldsymbol{R}} \left| P^{*}\left[(\hat{\Theta}_{i,h}^{*}\zeta - \hat{\Theta}_{i,h}\zeta) \leq x\right] - P\left[(\hat{\Theta}_{i,h}\zeta - \Theta_{i,h}\zeta) \leq x\right] \right| \xrightarrow{p} 0,$$

for any i as $N, T \to \infty$ and $\sqrt{T}/N \to c\ (0 \leq c < \infty)$, uniformly in h.

This theorem implies that, under Procedure A, the bootstrap asymptotic distribution is equivalent to that of the original estimator. The result is free from the random rotations H and H^{*} because the impulse responses are identified in the structural sense in both the original and the bootstrap estimators. Second, the result is straightforward when $c = 0$ because we obtain the standard asymptotic distributions. However, the method is also valid when $c > 0$ to enable Procedure A to accommodate N being significantly smaller than T, because the bias of the bootstrap distribution converges in probability to the original bias. We now prove the validity of Procedure B.

[11]We rule out heteroskedasticity in the VAR errors, although an extension to the heteroskedastic case is an interesting topic for future research.
[12]Note that in the setup of Gonçalves and Perron (2014) the estimated factors are used only as the regressors outside the factor model.
[13]See Kilian (1999) for the standard VAR analysis. In our unreported limited Monte Carlo study, they give similar coverage properties.

Theorem 4. (Asymptotic validity of Procedure B)
Suppose Assumptions 1–7 hold. If models 4 and 5 are estimated using the two-step PC method and matrix B is estimated by ID1, ID2, or ID3, and Procedure B is implemented with the i.i.d. or the wild bootstrap, then for any $r \times 1$ *shock vector* ζ

$$\sup_{x \in \boldsymbol{R}} \left| P^* \left[(\hat{\Theta}^{**}_{i,h}\zeta - \hat{\Theta}_{i,h}\zeta) \le x \right] - P \left[(\hat{\Theta}_{i,h}\zeta - \Theta_{i,h}\zeta) \le x \right] \right| \xrightarrow{p} 0,$$

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

This theorem shows that Procedure B is also valid, but only if $\sqrt{T}/N \to 0$. Therefore, we predict that the bootstrap provides a good coverage ratio when N is not significantly smaller than T. However, if it is, it may not replicate the original distribution because this algorithm does not estimate the factors in the bootstrap space and does not account for the effect of the factor estimation errors.

Remark 1. In Theorems 3 and 4, serial correlation in the idiosyncratic errors is ruled out, although this is an issue in most empirical applications. Suppose we are interested in the structural IRF of the ith cross-section $\Theta_{i,h}$. One direction to address this concern under $\sqrt{T}/N \to 0$ is to replicate the serial correlation in the idiosyncratic errors only of the ith cross-section. This is because Theorem 2 suggests that all we need to replicate is the asymptotic variance ($\Omega_{\lambda i}$, Ω_A, and Ω_B), but no asymptotic bias under $\sqrt{T}/N \to 0$. It also shows that Ω_A and Ω_B do not depend on the serial correlation in the idiosyncratic errors. Thus mimicking the serial correlation only in the ith idiosyncratic errors yields the valid asymptotic variance of $\hat{\lambda}_i$ (i.e., $\Omega_{\lambda i}$) and hence that of $\hat{\Theta}_{i,h}$.

Remark 2. In contrast, under $\sqrt{T}/N \to c$ with $0 \le c < \infty$, the asymptotic bias ($\Delta_{\lambda i}$, Δ_A, and Δ_B) matters to the limiting distributions as well. Theorem 2 suggests that the asymptotic bias depends on the autocovariances of the idiosyncratic errors of all cross-sections (Γ_j) through $\Delta_{1,j}$ and $\Delta_{3,j}$. Therefore, in principle, one needs to mimic the serial correlation in the idiosyncratic errors of all cross-sections. We leave this analysis for a future research topic.

5 | FINITE-SAMPLE PROPERTIES

5.1 | Monte Carlo simulations with artificial data

This section reports Monte Carlo simulation results to assess the finite-sample properties of the bootstrap procedures. We consider a simple VAR(1) model, where the $N \times 1$ response variables X_t are generated by (4) with the factors F_t evolving by Equation (5) with $p = 1$ for $t = 1, \ldots, T$. The VAR errors are generated by $e_t = Be^s_t$ with $e^s_t \sim$ i.i.d.$N(0, I_r)$ for $t = 0, 1, \ldots, T$. We set $r = 2$, and the choice of the parameter values is

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

unless otherwise stated. The factor loadings $\lambda_{i,j}$ ($j = 1, 2$) are generated from i.i.d. uniform distributions $U[0, 1]$ and are transformed to ensure that the identification restrictions hold. For ID1, the generated $N \times r$ matrix of Λ is divided by Chol($\Lambda' \Lambda / N$). For ID2 and ID3, we use the second and third rows of Λ for the identification, such that

$$\Lambda_{2:3} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

This makes the short-run structural impulse responses $\Lambda_{2:3}B$ and the long-run structural impulse responses $\Lambda_{2:3}(I_r - A)^{-1}B$ lower triangular matrices.

For $u_{i,t}$, we consider the following two patterns: (1) i.i.d. standard normal distributions, which we call "i.i.d."; and (2) the independent normal distribution of random variance generated by $U[0.5, 1.5]$, which we call "Hetero."

We report the results for the impulse responses of the first variable to a unit structural shock of the first factor. Because the sample size and the estimation errors in $\hat{B}$ can affect the results, we study the four (T, N) combinations $(T, N) = (40, 50)$, $(40, 200)$, $(120, 50)$, and $(120, 200)$ using ID1, ID2, and ID3. The experiment is based on 3,000 replications with 399 bootstrap repetitions. The confidence interval is based on the equal-tailed Hall's percentile interval at the 95% nominal level. By default, the bias correction of Kilian (1998) is applied, where the bias of $\hat{A}$ is estimated using another 300-bootstrap loop. The coverage rate and the median length of the confidence interval for the structural

TABLE 1 Coverage properties of impulse responses (i.i.d., 95% level)

		Coverage rate (%)						Length of CI (median)					
T	N	$h = 0$	1	2	3	4	5	$h = 0$	1	2	3	4	5
ID1													
Procedure A													
40	50	94.1	94.2	91.2	94.0	94.0	94.5	1.19	1.44	1.07	0.77	0.55	0.39
40	200	95.3	93.5	91.0	93.3	94.4	95.6	1.21	1.42	1.03	0.75	0.53	0.38
120	50	92.7	95.2	94.3	93.3	92.9	93.4	0.74	0.85	0.62	0.39	0.24	0.14
120	200	93.7	96.3	93.0	91.7	91.8	92.3	0.72	0.85	0.62	0.39	0.24	0.14
Procedure B													
40	50	92.2	89.6	84.4	92.7	95.0	95.8	1.10	1.31	0.93	0.68	0.48	0.35
40	200	92.3	90.1	85.7	92.0	94.1	96.2	1.11	1.33	0.93	0.67	0.48	0.35
120	50	89.8	91.5	90.0	89.7	90.4	92.7	0.69	0.80	0.57	0.36	0.22	0.13
120	200	90.4	93.1	89.3	89.7	91.4	92.9	0.68	0.78	0.56	0.35	0.21	0.13
Normal													
40	50	94.1	92.4	85.7	84.0	85.1	84.3	1.24	1.37	0.90	0.53	0.29	0.15
40	200	94.3	92.1	84.1	83.0	84.3	84.7	1.24	1.36	0.88	0.50	0.28	0.14
120	50	92.0	93.8	92.5	90.2	88.5	88.2	0.72	0.81	0.57	0.33	0.17	0.09
120	200	92.2	93.6	91.4	89.9	89.3	88.1	0.72	0.81	0.56	0.32	0.17	0.09
ID2													
Procedure A													
40	50	94.2	96.0	93.9	94.3	92.4	93.1	0.80	0.72	0.45	0.29	0.19	0.12
40	200	94.1	97.6	95.5	95.7	95.1	95.3	0.78	0.71	0.48	0.34	0.23	0.16
120	50	95.9	97.5	94.9	90.7	90.1	89.6	0.49	0.42	0.27	0.15	0.08	0.05
120	200	95.7	98.5	95.7	94.4	94.0	93.3	0.47	0.41	0.27	0.16	0.10	0.05
Procedure B													
40	50	90.7	91.7	89.3	92.7	92.5	93.7	0.75	0.63	0.40	0.27	0.18	0.12
40	200	93.5	92.6	91.8	95.3	95.9	95.9	0.75	0.64	0.44	0.30	0.21	0.15
120	50	91.0	94.2	89.7	88.3	87.4	87.0	0.45	0.37	0.25	0.15	0.08	0.05
120	200	94.8	95.1	91.7	90.7	92.2	93.2	0.45	0.37	0.25	0.15	0.09	0.05
Normal													
40	50	92.9	92.8	87.5	84.7	82.0	80.4	0.79	0.64	0.36	0.19	0.09	0.04
40	200	94.8	93.3	90.3	87.3	86.2	84.8	0.78	0.65	0.40	0.22	0.11	0.06
120	50	91.5	93.1	91.1	89.3	87.6	85.5	0.45	0.38	0.24	0.13	0.06	0.03
120	200	94.4	94.0	92.5	90.2	88.9	88.0	0.45	0.37	0.24	0.13	0.07	0.03
ID3													
Procedure A													
40	50	87.6	95.1	91.1	91.7	90.0	90.5	0.86	0.73	0.46	0.31	0.20	0.13
40	200	89.7	94.3	92.4	94.9	94.2	94.6	0.86	0.70	0.48	0.34	0.23	0.16
120	50	90.9	95.4	93.3	90.7	90.8	88.0	0.53	0.44	0.28	0.16	0.09	0.05
120	200	93.5	95.6	95.8	94.6	93.7	93.2	0.52	0.42	0.29	0.18	0.10	0.06
Procedure B													
40	50	87.2	91.6	88.8	91.1	90.9	91.0	0.83	0.65	0.44	0.30	0.21	0.15
40	200	85.9	92.7	89.6	93.1	94.2	94.5	0.84	0.65	0.45	0.32	0.23	0.17
120	50	89.8	92.4	89.2	87.0	86.2	86.4	0.50	0.39	0.26	0.16	0.09	0.05
120	200	90.3	94.4	91.6	89.9	90.6	91.6	0.51	0.38	0.26	0.16	0.10	0.06
Normal													
40	50	90.4	92.9	87.9	84.8	82.3	79.6	0.77	0.64	0.38	0.21	0.10	0.05
40	200	92.5	93.2	90.1	87.3	85.7	83.8	0.77	0.63	0.39	0.22	0.11	0.06
120	50	88.6	92.1	91.5	89.3	86.5	85.5	0.45	0.37	0.24	0.13	0.07	0.03
120	200	91.0	93.6	93.4	91.6	90.5	89.0	0.44	0.36	0.25	0.14	0.07	0.04

TABLE 2 Coverage properties of impulse responses (hetero, 95% level)

		Coverage rate (%)						Length of CI (median)					
T	N	$h = 0$	1	2	3	4	5	$h = 0$	1	2	3	4	5
ID1													
Procedure A													
40	50	95.0	95.0	91.1	92.7	93.2	93.7	1.20	1.43	1.05	0.76	0.55	0.39
40	200	94.6	95.2	92.0	93.9	95.1	95.6	1.22	1.45	1.05	0.76	0.55	0.38
120	50	91.5	95.6	93.0	92.8	92.7	93.0	0.73	0.84	0.60	0.38	0.23	0.14
120	200	90.8	95.4	93.3	93.0	94.1	95.0	0.75	0.87	0.63	0.39	0.24	0.14
Procedure B													
40	50	92.4	90.5	88.5	93.8	94.9	95.5	1.11	1.33	0.95	0.70	0.50	0.36
40	200	91.8	91.2	86.6	93.3	95.4	96.5	1.12	1.32	0.94	0.68	0.49	0.36
120	50	89.3	91.8	89.1	88.7	90.6	92.8	0.67	0.77	0.56	0.35	0.21	0.13
120	200	89.1	90.7	89.0	88.3	89.6	92.7	0.70	0.81	0.56	0.35	0.21	0.12
Normal													
40	50	94.7	93.1	88.4	86.9	86.2	85.4	1.25	1.36	0.90	0.52	0.29	0.15
40	200	94.2	91.5	87.2	86.4	86.4	86.7	1.26	1.37	0.88	0.51	0.27	0.14
120	50	91.3	92.9	90.4	89.2	88.4	87.4	0.71	0.79	0.56	0.32	0.17	0.08
120	200	91.5	91.9	90.7	89.2	87.2	86.8	0.73	0.81	0.56	0.31	0.16	0.08
ID2													
Procedure A													
40	50	93.1	97.3	93.5	92.8	89.3	90.0	0.86	0.74	0.45	0.29	0.18	0.12
40	200	94.3	97.5	94.2	95.5	94.8	95.2	0.79	0.70	0.48	0.33	0.23	0.15
120	50	94.1	97.2	93.7	91.9	90.5	88.4	0.51	0.43	0.26	0.14	0.08	0.04
120	200	96.2	97.2	95.9	95.2	94.6	94.3	0.47	0.40	0.27	0.17	0.10	0.06
Procedure B													
40	50	89.3	91.3	88.2	90.2	89.5	89.7	0.78	0.64	0.41	0.27	0.18	0.12
40	200	93.1	92.7	89.2	93.0	93.7	94.7	0.76	0.61	0.42	0.29	0.21	0.14
120	50	89.5	90.3	88.4	86.5	85.6	85.7	0.46	0.37	0.24	0.14	0.08	0.04
120	200	95.1	93.9	90.8	90.4	92.1	93.8	0.46	0.36	0.24	0.15	0.09	0.05
Normal													
40	50	91.7	91.9	87.6	82.3	77.8	76.1	0.81	0.66	0.37	0.19	0.09	0.04
40	200	94.8	92.1	87.5	85.7	84.8	83.6	0.79	0.62	0.38	0.21	0.11	0.05
120	50	88.8	91.2	89.7	87.7	86.3	84.6	0.46	0.37	0.23	0.12	0.06	0.03
120	200	94.1	93.4	92.0	91.1	89.6	88.3	0.45	0.36	0.24	0.13	0.07	0.03
ID3													
Procedure A													
40	50	89.5	95.7	92.2	92.4	89.5	90.1	0.89	0.74	0.48	0.31	0.21	0.14
40	200	90.0	95.9	93.7	94.4	94.6	94.9	0.87	0.72	0.50	0.35	0.24	0.17
120	50	93.8	96.6	92.3	90.8	88.6	88.0	0.55	0.43	0.27	0.15	0.08	0.05
120	200	93.4	96.0	95.0	94.0	92.8	92.1	0.53	0.42	0.28	0.17	0.10	0.06
Procedure B													
40	50	89.3	90.8	86.6	90.0	90.1	90.9	0.84	0.66	0.44	0.30	0.21	0.15
40	200	87.1	93.1	90.8	93.6	94.9	95.5	0.85	0.66	0.46	0.32	0.23	0.16
120	50	92.1	94.3	88.7	87.2	85.8	85.2	0.53	0.38	0.25	0.15	0.09	0.05
120	200	89.8	94.2	93.0	91.4	91.5	93.0	0.52	0.37	0.26	0.16	0.10	0.06
Normal													
40	50	92.8	92.4	87.8	83.7	81.1	78.5	0.79	0.64	0.38	0.20	0.10	0.05
40	200	92.8	93.5	90.3	87.8	86.1	84.2	0.80	0.62	0.39	0.22	0.11	0.06
120	50	91.3	92.7	91.3	88.4	85.2	82.9	0.46	0.37	0.24	0.13	0.06	0.03
120	200	90.2	93.4	94.2	92.4	90.1	88.9	0.45	0.36	0.24	0.14	0.07	0.04

TABLE 3 Coverage properties of impulse responses under small N (i.i.d., 95% level)

		Coverage rate (%)						Length of CI (median)					
T	N	$h=0$	1	2	3	4	5	$h=0$	1	2	3	4	5
ID1													
Procedure A													
120	10	92.9	96.6	92.7	90.5	89.9	89.7	0.75	0.88	0.61	0.36	0.21	0.12
120	30	92.9	96.3	94.4	93.7	93.5	93.9	0.71	0.84	0.61	0.39	0.23	0.14
240	10	86.5	96.4	94.6	92.6	90.7	89.7	0.54	0.62	0.44	0.26	0.14	0.07
240	30	87.0	96.0	94.2	92.8	91.6	91.5	0.52	0.61	0.44	0.26	0.14	0.08
Procedure B													
120	10	88.5	93.2	88.9	87.2	87.9	89.9	0.68	0.79	0.56	0.34	0.20	0.12
120	30	88.8	92.7	89.9	89.3	90.2	92.0	0.66	0.77	0.56	0.35	0.21	0.13
240	10	80.4	92.8	89.8	88.0	87.2	87.9	0.49	0.57	0.41	0.24	0.13	0.07
240	30	83.0	93.1	91.0	89.5	87.8	88.2	0.49	0.57	0.41	0.24	0.14	0.07
Normal													
120	10	92.1	94.6	90.9	88.4	86.7	84.8	0.71	0.80	0.55	0.30	0.16	0.08
120	30	92.6	93.7	92.1	89.8	88.2	87.9	0.71	0.78	0.56	0.31	0.17	0.08
240	10	83.3	93.1	90.9	90.0	88.6	87.5	0.51	0.58	0.41	0.23	0.12	0.06
240	30	83.8	93.5	92.9	91.4	89.6	87.9	0.50	0.57	0.41	0.23	0.12	0.06
ID2													
Procedure A													
120	10	84.5	98.6	89.7	82.0	75.4	72.8	0.52	0.46	0.24	0.12	0.06	0.03
120	30	93.1	98.1	93.8	86.7	83.8	81.8	0.49	0.42	0.25	0.14	0.07	0.04
240	10	85.3	95.9	86.3	79.6	75.1	75.8	0.37	0.33	0.17	0.08	0.04	0.02
240	30	94.7	98.2	94.4	88.5	83.5	85.6	0.36	0.31	0.18	0.10	0.05	0.02
Procedure B													
120	10	70.4	89.8	82.8	78.3	74.6	70.4	0.45	0.38	0.22	0.12	0.06	0.03
120	30	88.3	91.8	88.0	84.9	82.7	83.2	0.44	0.36	0.23	0.14	0.08	0.04
240	10	58.1	84.2	78.6	74.5	68.5	63.4	0.32	0.27	0.16	0.08	0.04	0.02
240	30	85.5	92.9	88.3	85.7	82.0	79.8	0.33	0.27	0.17	0.09	0.05	0.02
Normal													
120	10	72.5	89.7	85.1	80.4	75.8	72.3	0.47	0.39	0.21	0.10	0.05	0.02
120	30	86.8	91.6	89.3	85.7	83.5	81.0	0.45	0.36	0.22	0.12	0.06	0.03
240	10	60.6	83.3	81.0	78.4	75.2	71.4	0.33	0.28	0.16	0.08	0.03	0.01
240	30	82.8	92.6	90.1	89.1	86.5	83.7	0.32	0.27	0.17	0.09	0.04	0.02
ID3													
Procedure A													
120	10	86.0	96.3	88.7	83.0	83.8	78.3	0.56	0.46	0.25	0.13	0.06	0.03
120	30	90.4	96.8	93.2	88.9	84.9	80.9	0.53	0.42	0.26	0.14	0.08	0.04
240	10	88.1	96.1	90.1	82.6	85.6	80.7	0.41	0.33	0.18	0.09	0.04	0.02
240	30	91.2	96.3	93.8	88.4	84.0	84.3	0.38	0.31	0.18	0.10	0.05	0.02
Procedure B													
120	10	80.5	91.7	84.8	79.3	73.8	70.2	0.50	0.39	0.24	0.13	0.07	0.04
120	30	89.7	93.6	89.3	85.9	84.4	82.9	0.49	0.38	0.24	0.14	0.08	0.05
240	10	79.3	90.6	86.4	80.8	75.4	70.3	0.36	0.28	0.17	0.09	0.05	0.02
240	30	88.4	93.7	89.6	86.0	82.8	80.0	0.36	0.27	0.17	0.10	0.05	0.03
Normal													
120	10	77.1	91.7	88.4	83.0	77.6	73.3	0.44	0.38	0.22	0.11	0.05	0.02
120	30	88.0	93.9	92.3	89.7	86.6	83.7	0.43	0.36	0.23	0.12	0.06	0.03
240	10	74.1	89.9	86.6	83.9	81.2	77.6	0.31	0.27	0.16	0.08	0.04	0.02
240	30	84.9	92.7	90.4	88.9	87.0	84.7	0.31	0.26	0.17	0.09	0.04	0.02

impulse responses up to five periods are reported. As a comparison, the results obtained from the asymptotic normal approximations are also reported (denoted by "Normal.")

Table 1 shows the results of Procedures A, B, and Normal under the "i.i.d." case. Here, the i.i.d. bootstrap is used in step 3 of the algorithm. We observe that the coverage rate of Procedure A is close to the nominal level of 95% in most cases using any identification method. Procedure B also achieves a satisfactory coverage rate, but tends to undercover especially when $(T, N) = (120, 50)$. This is consistent with the fact that Procedure B does not account for the effect of the factor estimation errors in the bootstrap estimator, and the gap is wider when N is smaller than T. In contrast, the median length of Procedure B is almost always shorter than that of Procedure A. The asymptotic approximation becomes erratic as the horizon becomes longer because the impulse responses become highly nonlinear and the delta method does not work well. Table 2 reports the results under "Hetero." Here, we use the wild bootstrap in which η_{it} is an independent draw from the standard normal distribution for all i and t. The coverage rates in Table 2 are very similar to those in Table 1, which confirms that the wild bootstrap works well in finite samples. The median lengths are also similar to Table 1; thus Procedure B provides shorter median intervals. The asymptotic approximation has similar properties where the coverage rate deteriorates as h becomes larger. In Table 3, we investigate the cases of small N to assess how differently Procedures A and B behave. In particular, we set $(T, N) = (120, 10), (120, 30), (240, 10)$, and $(240, 30)$, and present the i.i.d. case because the other cases are qualitatively the same. Procedure B is reported to suffer from severe undercoverage, which often amounts to less than 90% with ID1 and less than 80% with ID2 and ID3, whereas Procedure A can mitigate the problem. The difference in the coverage rates between the two procedures is larger than in the previous cases, as the theory predicts.

5.2 | Monte Carlo simulation using empirical data

We next consider an experiment using a data-generating process modeled after actual economic data. To this end, we use a model calibrated by 128 monthly US macroeconomic time series as investigated by McCracken and Ng (2015).[14] Although the data span the period from January 1960 to December 2014,[15] we use the data up to December 2007 to exclude the financial crisis period. After transforming the data, the sample size is $T = 564$ and $N = 128$, and outliers are removed using the method proposed by Yamamoto (2015).[16] In addition, the data are demeaned and standardized to have unit standard deviations. We use a model with two factors and lag order four, although slight variations in the number of factors and lag order do not affect the qualitative results. We further 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 and financial variables, which is consistent with Stock and Watson (2005). Hence we use the ID1 method for structural identification assuming that the first factor is not contemporaneously affected by structural shocks of the second factor; however, using other identification methods provides very similar results.

The aim is to evaluate the coverage properties of Procedures A and B. However, the coverage rate of the confidence interval constructed from the actual data cannot be calculated. Hence we use the following calibration experiment to approximate the actual data-generating process.

1. Estimate Equations 4 and 5 and obtain the coefficient estimates and residuals.
2. Generate the quasi-observations from the calibrated model, with the error terms resampled 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 sample covariance matrix of $\hat{e}_t$. This allows e_t^s to be interpreted as structural innovations.
3. Using each generated data set, construct 95% confidence intervals for the impulse responses using the proposed bootstrap procedure with 399 repetitions and determine whether the true (calibrated) impulse responses 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 impulse responses of the CPI (consumer price index: all items), short rate (3-month Treasury bill rate), IP (industrial production index: total), and unemployment rate (civilian unemployment rate) to the first factor shock. Table 4 provides the coverage rates for the impulse responses for eight periods using Procedures A and B, and Table 5

[14]These data can be downloaded from the author's website. The original data consist of 133 series. However, five series ("real manufacturing and trade industries sales," "new orders for consumer goods," "new orders for nondefense capital goods," "consumer sentiment index," and "trade-weighted US dollar index") are not available for the sample period and thus are omitted from our data set.
[15]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.
[16]This method corrects outliers that exceed five times the standard error estimate in the filtered individual series.

TABLE 4 Coverage rate of confidence intervals for the calibrated US macroeconomic model (95% level)

	Proc.	$h = 0$	1	2	3	4	5	6	7
CPI	A	89.7	92.1	92.1	92.4	92.4	91.9	92.1	91.5
	B	88.3	91.1	91.0	91.7	91.8	91.8	91.4	91.1
	N	92.0	91.4	92.7	90.3	90.5	89.0	88.8	88.1
Short rate	A	90.5	90.8	90.7	91.1	91.8	92.3	92.4	92.1
	B	70.1	79.7	76.6	81.5	85.4	87.9	89.7	90.5
	N	69.2	80.4	78.2	82.9	86.1	88.5	90.0	90.7
IP	A	97.3	96.1	98.4	97.1	97.1	95.3	93.6	92.1
	B	79.1	92.5	84.2	88.8	87.6	88.9	88.2	88.6
	N	77.2	92.7	80.0	85.7	84.3	85.5	86.2	87.5
Unemployment rate	A	95.3	95.6	97.8	97.2	96.5	95.6	94.4	93.1
	B	85.3	92.5	88.2	91.4	90.4	90.5	90.0	90.1
	N	84.9	92.8	85.0	87.7	86.6	87.7	88.0	88.8

TABLE 5 Median length of confidence intervals for the calibrated US macroeconomic model (95% level)

	Proc.	$h = 0$	1	2	3	4	5	6	7
CPI	A	7.74	4.78	4.93	4.29	4.00	3.70	3.45	3.23
	B	7.31	4.52	4.68	4.10	3.86	3.60	3.40	3.22
	N	7.69	4.64	4.92	4.06	3.85	3.43	3.21	2.96
Short rate	A	8.85	6.44	6.24	5.81	5.65	5.45	5.30	5.12
	B	5.39	4.50	4.15	4.34	4.53	4.74	4.87	4.94
	N	5.26	4.50	4.11	4.24	4.46	4.66	4.83	4.93
IP	A	16.29	13.67	13.86	13.48	14.48	14.54	14.87	14.76
	B	10.69	12.48	10.28	11.69	12.30	12.87	13.18	13.26
	N	10.42	12.66	10.02	11.57	12.19	12.94	13.38	13.61
Unemployment rate	A	13.12	11.42	10.88	10.79	11.53	11.63	11.87	11.80
	B	10.18	10.64	9.00	9.97	10.36	10.79	11.00	11.03
	N	10.18	10.86	8.91	9.88	10.32	10.83	11.14	11.26

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 lower than the nominal level, as the theory suggests.

6 | EMPIRICAL ILLUSTRATION

This section provides an empirical example of the proposed bootstrap methods on the basis of the monetary policy analysis of Bernanke et al. (2005). We describe their FAVAR model in the reduced form such that

$$X_t = \Lambda F_t + CY_t + u_t,$$
$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \sum_{j=1}^{p} A_j \begin{bmatrix} F_{t-j} \\ Y_{t-j} \end{bmatrix} + e_t,$$

with $e_t \equiv Be_t^s$, where F_t consists of r unobservable factors extracted from the US macroeconomic data set and Y_t is the federal funds rate (FFR). We are interested in the impulse of various macroeconomic variables in X_t in response to a structural shock of the FFR. To this end, we assume i.i.d. VAR errors with $E(e_t^s e_t^{s\prime}) = I_{r+1}$. As we mentioned in Section 2, we impose a simple orthogonality restriction between the observable and unobservable factors; hence we first obtain $M_Y X$ with $M_Y = I_T - Y(Y'Y)^{-1}Y'$, where Y is a $T \times 1$ vector of the observed factor, and we estimate the unobserved factors as the PCs of $M_Y X$. The matrix of contemporaneous correlations is estimated through ID1. More recently, Bai et al. (2016)

discuss identification and estimation methods in such models, and we leave incorporating their method in this example to a future agenda.

Methodological differences between Bernanke et al. (2005) and our procedure arise from several sources. First, their method does not involve factor estimation in the bootstrap loops and is essentially similar to Procedure B. Second, they identify unobservable factors as residuals of the regression of the standard PCs on the PCs from "slow-moving variables" and Y_t, and this step is performed only in the estimation but not in the bootstrap loops. Third, their method involves a specification difference between the original and bootstrap data generations that should have minor effects in this example.

We use the same 120 monthly US macroeconomic variables for the sample period from January 1959 to August 2001 and the same data transformation. The model specification also follows Bernanke et al. (2005): We include three unobservable factors, together with the federal funds rate (FFR) as an observable factor, and set the lag order as $p = 13$. A 25-basis-point innovation in the FFR (i.e., a contractionary monetary policy shock) is investigated and the obtained impulse responses and associated confidence intervals are retransformed to level. The confidence level is 90% and the number of replications is 3,000, as per the original study. Although we could compute the impulse responses of 120 variables, given space limitations, the following six selected variables are presented: the FFR, the industrial production index (total; IP), the consumer price index (all items; CPI), the exchange rate yen (Japan/US foreign exchange rate), the unemployment rate (civilian unemployment rate), and new orders (new orders for durable goods).

Figure 1 is an exact replication of Figure 2 of Bernanke et al. (2005), whereas Figure 2 shows the results of six variables using our procedures with the i.i.d. bootstrap for the idiosyncratic errors. In any case of Figure 2, Procedure A (thick dotted lines) and Procedure B (thin dotted lines) are not very different, for the following two reasons. First, we have relatively

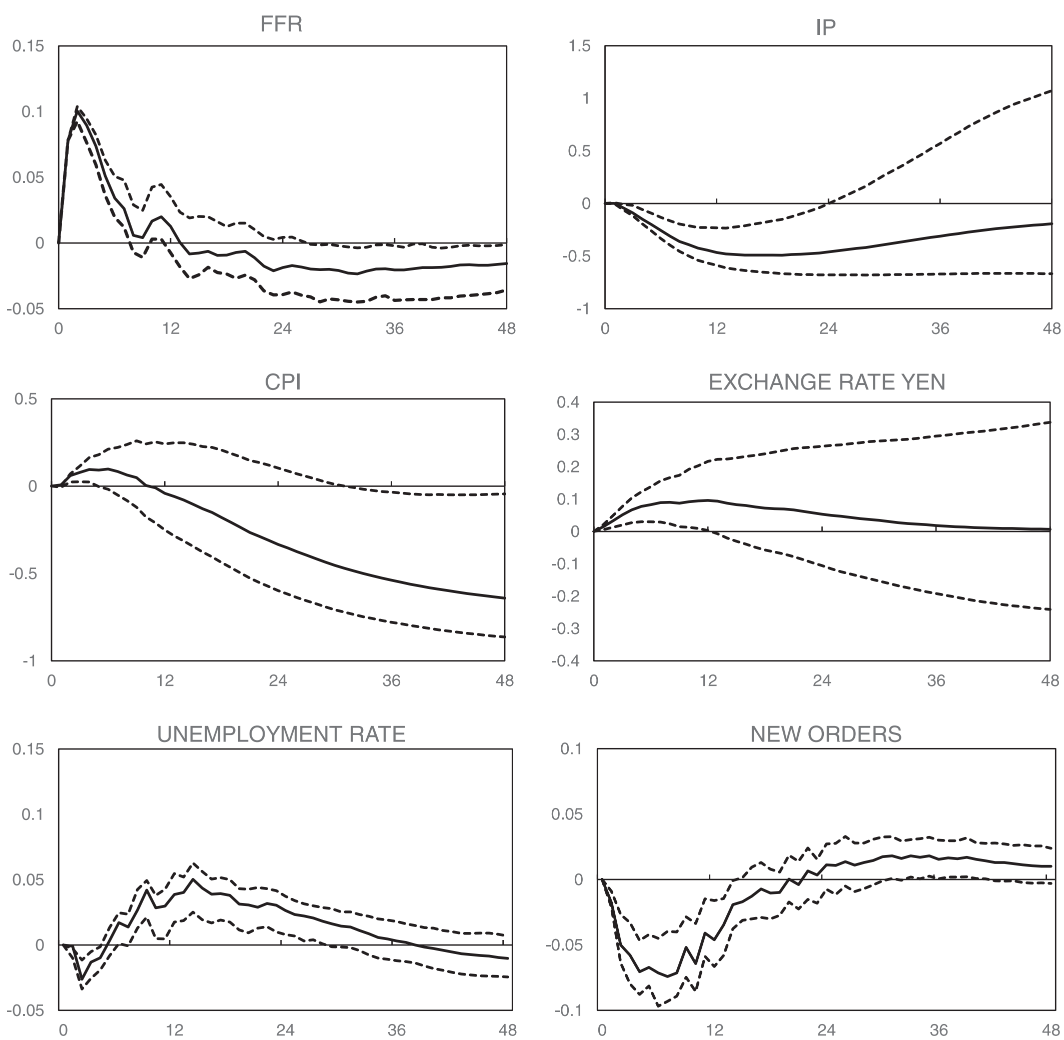

FIGURE 1 Impulse responses to a contractionary monetary policy shock (Bernanke et al., 2005)

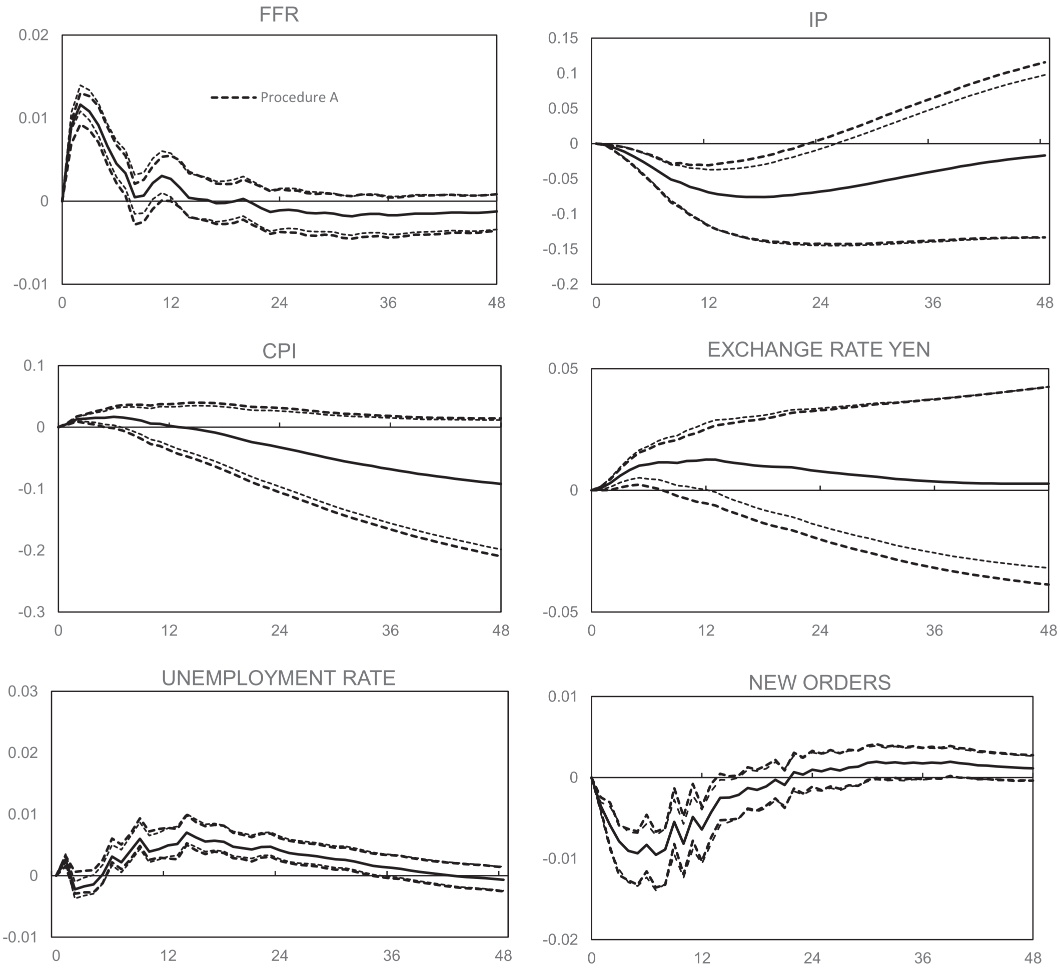

FIGURE 2 Impulse responses to a contractionary monetary policy shock using the i.i.d. bootstrap

large cross-sectional units. This is consistent with our theoretical result that, under a relatively large N, Procedures A and B yield similar properties. Second, we consider impulse responses to a shock to the observable factor, and we impose an orthogonal restriction between the observed and unobserved factor estimates. Hence the impulse response estimate has a smaller effect of factor estimation errors compared with the impulse responses to an unobservable factor shock.[17] We also observe that, in most cases, a shift in the central location occurs, because Procedure A replicates the asymptotic bias whereas Procedure B does not. Regarding the interval length, Procedure A produces a wider interval for the IP, the CPI, and the exchange rate, which is consistent with the theory. The lengths for the FFR, the unemployment rate, and the new orders are approximately the same, which is a reasonable result under a large N.

We now compare our results in Figure 2 with those of Bernanke et al. (2005). We first observe that the shapes of the point estimates and confidence intervals are similar in all six impulse responses in the study. Note that the scale difference of the impulse responses between Figures 1 and 2 is caused by the particular unobservable factor estimation of Bernanke et al. (2005), as previously described. Interestingly, we obtain a difference in statistical significance, particularly in the CPI response for which the upper confidence limits do not cross the zero axis at longer horizons in Figure 2. Therefore, in contrast to Bernanke et al. (2005), the proposed methods suggest that the price level does not significantly decrease after a contractionary monetary policy shock. This caveat will be strengthened in an empirical analysis with fewer cross-sections. Third, the intervals are more symmetric around the point estimate in our results than in theirs, possibly because the bootstrap data-generating process is inconsistent with the original data-generating process in Bernanke et al. (2005).

Finally, we consider the effect of the heteroskedasticity in the idiosyncratic errors by using the wild bootstrap procedure with the same specification as in Section 5.1 and the results are reported in Figure 3. Because Figures 2 and 3 are very

[17]This effect must be more accurately investigated in a further study.

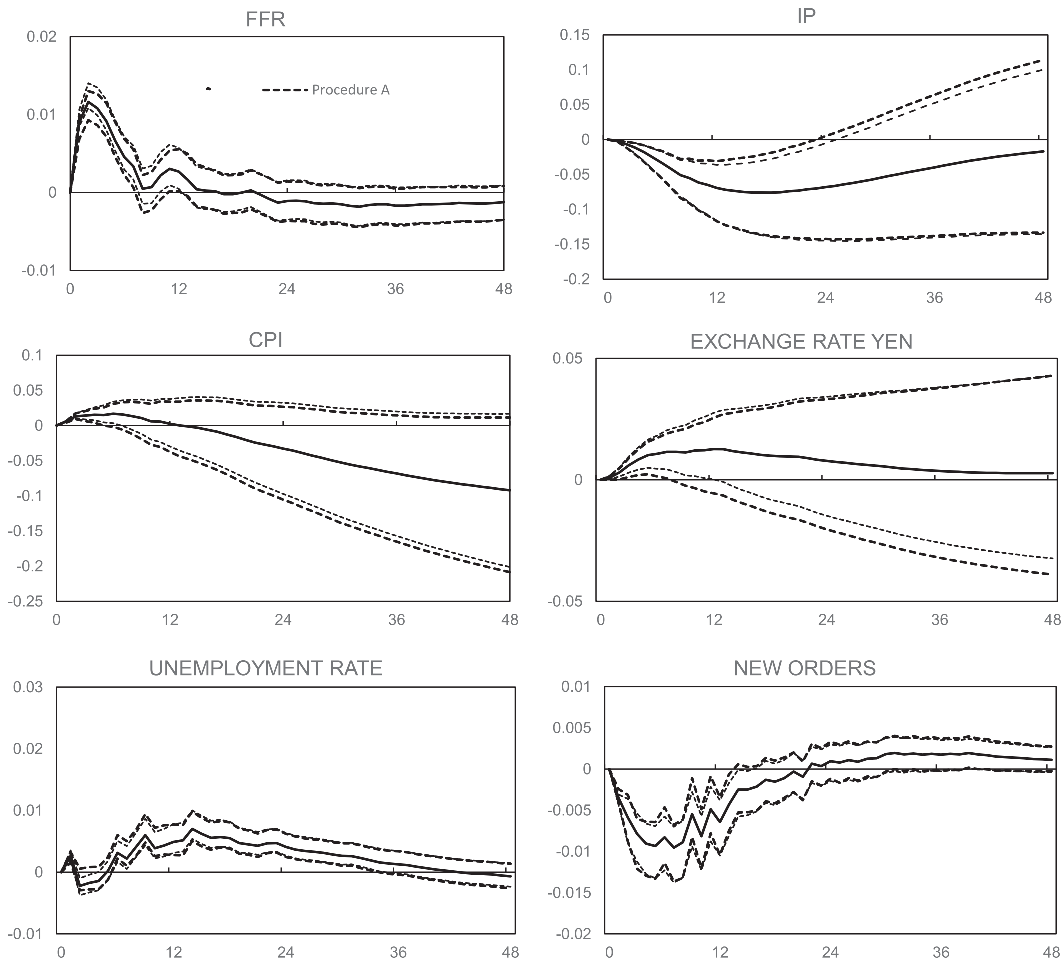

FIGURE 3 Impulse responses to a contractionary monetary policy shock using the wild bootstrap

similar, the heteroskedasticity of the idiosyncratic errors does not have a significant impact on the confidence intervals in this example.

7 | CONCLUSIONS

This study investigated residual-based bootstrap methods for the confidence interval of structural impulse responses in FAVARs. We are particularly interested in a comparison between the method with factor estimation (Procedure A) and that without factor estimation (Procedure B) in the bootstrap replications. In doing so, we restrict the VAR errors to be i.i.d. and the idiosyncratic errors in the factor model to be cross-sectionally and serially independent, although heteroskedasticity in the latter is allowed. Our theoretical results suggest that both procedures are asymptotically valid under $\sqrt{T}/N \to 0$. However, the former is also valid even when $\sqrt{T}/N \to c$ with $0 \leq c < \infty$. This is because Procedure A accounts for the effect of the factor estimation errors, whereas Procedure B does not. Our Monte Carlo simulations find that the former achieves better coverage rates than those of the latter, especially when N is significantly smaller than T. We also show that in the monetary policy analysis of Bernanke et al. (2005) the proposed methods result in a different statistical significance from the original study, and this caveat will be strengthened in an empirical analysis with fewer cross-sections. Finally, there are a number of ways to extend this study to meet requirements in a wide range of empirical applications. The list includes dealing with the serial correlation in the idiosyncratic errors and non-i.i.d. VAR errors.[18] Another agenda pertains to developing structural identification schemes because the ones in this study are limited.

[18]Brüggemann, Jentsch, and Trenkler (2016) show that the wild bootstrap procedure does not work in the standard structural VAR when the VAR errors are serially uncorrelated but also dependent, and the impulse response estimate involves an estimate for the covariance matrix.

ACKNOWLEDGMENTS

The author acknowledges the financial support from MEXT Grants-in-Aid for Scientific Research No. 16K03593 and Hitotsubashi Institute for Advanced Study, Hitotsubashi University. I am grateful to the editor Barbara Rossi and three anonymous referees for useful suggestions to revise the early draft. I also thank Silvia Gonçalves, Atsushi Inoue, Benoit Perron, Zhongjun Qu, Moto Shintani, and the participants of the 27th Annual Meeting of the CESG, the fifth CIREQ Time Series Conference, the 2011 North American Summer Meeting of the Econometric Society, the second Hitotsubashi Summer Institute Workshop, the 10th CFE Conference, the fourth IAAE Conference, and numerous seminars for their helpful comments. All remaining errors are mine.

OPEN RESEARCH BADGES

This article has earned an Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at [http://qed.econ.queensu.ca/jae/2019-v34.2/yamamoto/].

REFERENCES

Acconcia, A., & Simonelli, S. (2008). Interpreting aggregate fluctuations looking at sectors. *Journal of Economic Dynamics and Control*, *32*(9), 3009–3031.

Ang, A., & Piazzesi, M. (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., Li, K., & Lu, L. (2016). Estimation and inference on FAVAR models. *Journal of Business and Economic Statistics*, *34*(4), 620–641.

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

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

Blanchard, O. J., & Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, *79*(4), 655–673.

Boivin, J., & Giannoni, M. P. (2006). DGSE models in a data-rich environment. (*NBER Working Paper No. 12772*). Cambridge, MA: National Bureau of Economic Research.

Boivin, J., Giannoni, M. P., & Mihov, I. (2009). Sticky prices and monetary policy: Evidence from disaggregated U.S. data. *American Economic Review*, *99*(1), 350–384.

Boivin, J., Giannoni, M. P., & Stevanović, D. (2013). Dynamic effects of credit shocks in a data-rich environment. (*FRB of New York Staff Report No. 615*). New York, NY: Federal Reserve Bank.

Brüggemann, R., Jentsch, C., & Trenkler, C. (2016). Inference in VARs with conditional heteroskedasticity of unknown form. *Journal of Econometrics*, *191*, 69–85.

Corradi, V., & Swanson, N. R. (2014). Testing for structural stability of factor augmented forecasting models. *Journal of Econometrics*, *182*, 100–118.

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

Dufour, J.-M., & Stevanović, D. (2013). Factor-augmented VARMA models with macroeconomic applications. *Journal of Business and Economic Statistics*, *31*(4), 491–506.

Forni, M., & Gambetti, L. (2010). The dynamic effects of monetary policy: A structural factor model approach. *Journal of Monetary Economics*, *57*, 203–216.

Giannone, D., Reichlin, L., & Sala, L. (2005). Monetary policy in real time, *NBER Macroeconomics Annual*. Cambridge, MA: MIT Press, pp. 161–200.

Gilchrist, S., Yankov, V., & Zakrajšek, E. (2009). Credit market shocks and economic fluctuations: Evidence from corporate bond and stock markets. *Journal of Monetary Economics*, *56*(4), 471–493.

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

Gonçalves, S., & Perron, B. (2016). Bootstrapping factor models with cross-sectional dependence. (*Working paper*). Montreal, Canada: University of Montreal Department of Economic Sciences.

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

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. Berlin, Germany: Springer.

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.

Kilian, L., & Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis*. Cambridge, UK: Cambridge University Press.

Ludvigson, S. C., & Ng, S. (2009a). Macro factors in bond risk premia. *Review of Financial Studies*, *22*(12), 5027–5067.

Ludvigson, S. C., & Ng, S. (2009b). A factor analysis of bond risk premia. In Ulah, A., & Giles, D. (Eds.), *Handbook of empirical economics and finance* . (pp. 313-372), London, UK: Chapman & Hall, pp. 313–372.

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

McCracken, M. W., & Ng, S. (2015). FRED-MD: A monthly database for macroeconomic research. (*Working Paper Series 2015-012B*). St Louis, MO: Federal Reserve Bank of St. Louis.

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., & Sims, C. A. (1977). Business cycle modeling without pretending to have too much a priori economic theory. (*FRB Working Paper 55*). Minneapolis, MN: Federal Reserve Bank.

Shintani, M., & Guo, Z.-Y. (2015). Improving the finite sample performance of autoregression estimators in dynamic factor models: A bootstrap approach. *Econometric Reviews*, *37*(4), 360–379.

Stock, J. H., & Watson, M. W. (2005). Implications of dynamic factor models for VAR analysis. (*NBER Working Paper 11467*). Cambridge, MA: National Bureau of Economic Research.

Stock, J. H., & Watson, M. W. (2011). Dynamic factor models. In Clements, M. P., & Hendry, D. F. (Eds.), *Oxford Handbook of Economic Forecasting* (pp. 35–60). Oxford, UK: Oxford University Press.

Yamamoto, Y. (2015). Asymptotic inference for common factor models in the presence of jumps. (*HIAS Discussion Paper series E-4*). Tokyo, Japan: Hitotsubashi Institute for Advanced Study, Hitotsubashi University.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Yamamoto Y. Bootstrap inference for impulse response functions in factor-augmented vector autoregressions. *J Appl Econ*. 2019;34:247–267. https://doi.org/10.1002/jae.2659