

HIAS-E-72

Identifying Factor-Augmented Vector Autoregression Models via Changes in Shock Variances

Yohei Yamamoto^(a)

(a) Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601 Japan

August, 2018



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.

Identifying Factor-Augmented Vector Autoregression Models via Changes in Shock Variances*

Yohei Yamamoto[†]

Hitotsubashi University

August 6, 2018

Abstract

We propose a new method for the structural identification of a dynamic causal relationship in factor-augmented vector autoregression models based on changes in the unconditional shock variances that occur on a historical date. The proposed method can incorporate both observed and unobserved factors in the structural vector autoregression system and it allows the contemporaneous matrix to be fully unrestricted. We derive the asymptotic distribution of the impulse response estimator and consider a bootstrap inference method. Monte Carlo experiments show that the proposed method is robust to the misspecification of the contemporaneous matrix unlike the existing methods. Both the asymptotic and bootstrap methods obtain a satisfactory coverage rate when the shock of an observed factor is studied, although the latter is more accurate when the shock of an unobserved factor is considered. An empirical example based on the same data employed by Bernanke et al. (2005) provides similar point estimates and somewhat wider confidence intervals, thereby supporting their identification strategy.

JEL Classification Numbers: C14, C22

Keywords: dynamic causal effect, factor-augmented vector autoregression, identification through heteroskedasticity, impulse response

*The author acknowledges financial support for this study under MEXT Grants-in-Aid for Scientific Research No.17H00985 and from Hitotsubashi Institute for Advanced Study, Hitotsubashi University. I also thank participants at the 2018 Kansai Econometrics Study Group Meeting and the 5th IAAE Annual Conference for useful comments and suggestions. All remaining errors are my own.

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

1 Introduction

One of the most challenging issues in structural VAR analysis¹ is that the model can only accommodate a small number of variables in the equation system. If the model is lacking important variables, the empirical results can suffer from a severe omitted variable bias. One solution to address this concern is to incorporate a small number of unobserved factors extracted from an information set spanned by a large panel of data. This approach is called factor-augmented VAR (FAVAR) and it was initiated by Bernanke et al. (2005), advanced in studies summarized by Stock and Watson (2016). It is well-known that the factors extracted using the popular principal component method are only identified up to a random rotation (Bai and Ng, 2013) and this often causes difficulties when making a particular interpretation of the estimated unobserved factors, which plays a critical role when identifying a dynamic causal relationship in the VAR system. Recently, Bai et al. (2016) addressed this concern by proposing exclusion restrictions for identifying the unobserved factors or, more specifically, to fix the random rotation matrix in the FAVAR model. However, their exclusion restrictions assume zero contemporaneous correlations between the sets of observed and unobserved factors, which may not be supported without an economic interpretation of the unobserved factors. This chicken-and-egg problem demands a method for identifying a structural FAVAR model where its validity can be investigated by a statistical procedure.²

Thus, in this study, we consider an approach for identifying a simultaneous equation system based on changes in the unconditional variances proposed by Rigobon (2003) and Rigobon and Sack (2003). Recently, this method was extended to the conventional small-scale structural VAR system by Lanne and Lütkepohl (2008), Lütkepohl (2013), and Lütkepohl and Netsunajev (2017).³ The fundamental idea of this method is to increase the number of equations in the system by considering changes in the variances⁴ to achieve the order con-

¹See Sims (1980) and the vast amount of work surveyed by Kilian and Lütkepohl (2017).

²It should also be noted that identification restrictions motivated by economic reasoning may be difficult even in the conventional structural VAR when its dimension becomes relatively large. For instance, Brunnermeier et al. (2017) used the structural VAR to investigate the effects of up to 10 orthogonal shocks on macroeconomic variables and showed that their contemporaneous ordering is not a trivial problem when the dimension is relatively large.

³See Section 4.5 of Stock and Watson (2016) and Chapter 14 of Kilian and Lütkepohl (2017) for a survey of a wider class of identification methods that use heteroskedasticity. These include discrete changes in the unconditional variances, variances that follow Markov switching and smooth transition dynamics, and conditionally heteroskedastic variances. Lewis (2018) discusses a method that is robust to the misspecification of heteroskedasticity process.

⁴Importantly, the coefficient parameters are typically assumed to be constant.

dition for identification. As a result, we can avoid the agnostic zero restrictions that are imposed by particularly ordering the VAR variables or imposed on the impulse responses based on economic meaning. Instead, we can use a statistical procedure to verify the identification conditions. See Qu and Perron (2007) and Casini and Perron (2017), for recent developments in structural break analysis that incorporates changes in the error variances.⁵

More specifically, in this study, we develop a new method for the structural identification of the impulse responses in FAVAR models based on the changes in shock variances that occur on a historical date. In particular, we identify the rotated version of the contemporaneous matrix and this problem is fundamentally the same as that addressed by Bai et al. (2016) but we allow the contemporaneous matrix to be fully unrestricted. We derive the asymptotic distributions of the identified structural coefficients and impulse response estimators under the condition that $\sqrt{T}/N \rightarrow 0$ as N and $T \rightarrow \infty$, where T and N are the time and cross-section dimensions, respectively. We also consider a bootstrap method for constructing the confidence interval, based on the seminal studies of Gonçalves and Perron (2014)⁶ and Yamamoto (2017).

Our Monte Carlo experiments show that the impulse responses are estimated more accurately by the proposed method as the magnitude of the variance change and the sample size increase. More importantly, the proposed method is robust to the misspecification of the contemporaneous matrix unlike the existing methods proposed by Bai et al. (2016). Even when it is correctly specified, the mean squared error (MSE) with the proposed method is comparable to that with the existing methods when the magnitude of the variance change and the time dimension are large. We also find that both the asymptotic and bootstrap confidence intervals yield satisfactory finite sample properties when we study the responses to the shock of an observed factor. However, the latter provides a more accurate coverage rate when the shock of an observed factor is considered.

⁵Many studies considered factor models. In particular, Han and Inoue (2015) proposed a test for changes in the factor loadings at a common time by investigating a change in the variances of factors. Chen et al. (2014) considered the relationships between the estimated first factor and the remaining estimated factors. Breitung and Eickmeier (2011) and Yamamoto and Tanaka (2015) investigated changes in the factor loadings for an individual response variable. Because a change in the variance of factors and some type of changes in the factor loadings at the same timing may not be identified, we need to “assume” that the factor loadings do not have such common breaks.

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

Finally, we investigate whether the recursive identification strategy of the well-known study of Bernanke et al. (2005) is valid by using the same data employed in the original study but with the proposed identification strategy. To ensure consistency with the empirical evidence of the Great Moderation, we consider changes in the shock variances in December 1984. We find that our identification method could obtain similar point estimates of the impulse responses to those in the original study but with somewhat wider confidence intervals. These findings support the validity of the identification strategy implemented by Bernanke et al. (2005).

The remainder of this paper is organized as follows. In Section 2, we introduce the model and assumptions. In Section 3, we propose a new identification method for FAVAR models and derive the asymptotic distributions of the estimators. We also design a bootstrap method for constructing the confidence interval in the same section. In Section 4, we examine the finite sample properties of the proposed estimator based on comparisons with alternative methods proposed by Bai et al. (2016), and an empirical illustration is provided in Section 5. In Section 6, we give our conclusions, and the appendices provide the technical derivations and details of estimators that are not explained in the main text.

Throughout this study, we use the following notation. The Euclidean norm of vector x is denoted by $\|x\|$. The vector-induced norm is used for matrices. The symbols “ \xrightarrow{p} ” and “ \xrightarrow{d} ” represent convergence in probability under the probability measure P and convergence in distribution, respectively. The symbols $O_p(\cdot)$ and $o_p(\cdot)$ denote the order of convergence under P . We use the symbol “ $A \approx B$ ” if, for two random matrices A and B , $\|A - B\| = o_p(1)$ as $N, T \rightarrow \infty$. Let L be the standard lag operator. The operator $vec(X)$ transforms an $m \times m$ matrix X into an $m^2 \times 1$ vector by stacking the columns, whereas $vech(X)$ stacks only the elements on and above the main diagonal elements of a square matrix X . We denote the Cholesky decomposition of a positive definite matrix X by $chol(X)$. $[\cdot]$ returns the integer part.

2 Model and assumptions

2.1 Structural FAVAR models

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

$$h_t^* = v^* + \sum_{j=1}^p A_j^* h_{t-j}^* + \varepsilon_t, \quad (1)$$

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

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

where Λ^* and Γ^* are $N \times r_1$ and $N \times r_2$ factor loading matrices of the unobserved and observed factors, respectively. Let $C^* = [\Lambda^*, \Gamma^*]$ denote the composite factor loadings. If we write c_i^* , we denote an $r \times 1$ vector of factor loadings of the i th response variable for $i = 1, \dots, N$. Similarly, λ_i^* and γ_i^* are $r_1 \times 1$ and $r_2 \times 1$ vectors of the factor loadings attached to the unobserved and the observed factors, respectively. The error term u_t^* is an $N \times 1$ vector referred to as the idiosyncratic errors. Again, the constant term ι^* can be omitted for simplicity.

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

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

or equivalently

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

To simplify the notation, we define the reduced-form VAR coefficient of the j th lag $A_j = B A_j^* B^{-1}$ and the reduced-form errors $e_t = B\varepsilon_t$ and we write the reduced-form VAR model

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

This is the same model as that described by Bai et al. (2016). If we simplify the notation by letting $W = [H_{(-1)}, H_{(-2)}, \dots, H_{(-p)}]$ be a $T \times rp$ matrix where $H_{(-j)} = [h_{1-j}, \dots, h_{T-j}]'$, and

⁷This is a simplified model of the formulation given by Stock and Watson (2016) where the number of shocks is the same as the number of factors.

$A = [A_1, \dots, A_p]'$ is an $rp \times r$ matrix, then (3) can be written equivalently as $H = WA + e$. The corresponding reduced-form factor model is

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

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

In this study, we develop a method for identifying the structural FAVAR model by estimating the matrix B but without imposing any exclusion restrictions on it. Our approach follows the method for identification based on the changes in shock variances proposed by Rigobon (2003) and Rigobon and Sack (2003) in a demand-and-supply system, which was extended to the conventional small-scaled structural VAR models by Lanne and Lütkepohl (2008), Lütkepohl (2013) and Lütkepohl and Netsunajev (2017).⁸ In particular, we define the structural impulse response of an observed variable $x_{i,t+h}$ to the structural shocks in ε_t at horizon h as

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

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

2.2 Assumptions

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

Assumption A (VARs)

(i) The structural VAR shocks ε_t are independent and identically distributed (i.i.d.) with mean zero, and the unconditional covariance matrix changes on a known date $T_b = [\kappa T]$ where $\kappa \in (0, 1)$, i.e.,

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

⁸The proposed method considers a discrete change in the unconditional variance of errors, where changes in the conditional variances and various types of volatility changes are studied (see chapter 14 of Kilian and Lütkepohl (2017)). Brunnermeier et al. (2017) apply this method to measure the effects of up to 10 orthogonal shocks on macroeconomic variables in financial markets.

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

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

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

(iv) The roots of $\det(I_r - A_1y - A_2y^2 - \dots - A_p y^p) = 0$ lie outside the unit circle.

Assumption B (idiosyncratic errors)

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

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

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

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

Assumption C (factor model)

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

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

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

Assumption D (normalization through structural impulse responses) The values are known for the main diagonal elements of the matrix of contemporaneous structural impulse responses of the first r variables.

The key assumptions of this paper are Assumptions A (i), (ii) and D. Assumptions A (i) and (ii) are indeed similar to that considered by Lütkepohl (2013) and Lütkepohl and Netsunajev (2017). They normalize the covariance matrix of one regime to be an identity matrix to fix the scale of the structural shocks, we also fix the scale of the eigenvectors based on Assumption D. Assumption A (iii) assumes a known break date for simplicity, but this could be relaxed by using the standard break date estimate developed by Qu and Perron (2007).⁹ The standard break date fraction estimator, i.e., the break date estimator divided by the sample size, would have a faster rate than \sqrt{T} so the estimation uncertainty associated

⁹Strictly speaking, as far as the author knows, this particular technique in the comon factor model is still an open question.

with the break date does not affect the asymptotic inference for the least square coefficient estimators. Assumptions A (iv) guarantees a stable VAR system and it also ensures that the probability limit of $F'M_G F/T$ becomes a fixed positive definite matrix which we use in Assumption C (ii). Assumption B is essentially the same as the standard literature of the factor model, e.g., see Bai (2003) and Bai and Ng (2006), and it allows idiosyncratic errors to have weak serial and cross-sectional dependence. Assumption C is the standard regularity condition and it guarantees the statistical identification of the unobserved factors up to a random rotation, as discussed in the next section. Finally, by Assumption D, we need to fix the scale of the unobserved economic shocks where this normalization depends on the researcher's discretion, discussed in the empirical illustration.¹⁰

3 Estimation and inference

3.1 Estimation of the reduced-form models

We use the two-step principal component method to estimate the reduced-form models (3) and (4). First, we obtain the regression residuals for x_{it} on the observed factors g_t . Then, the unobserved factors f_t are estimated as the r_1 principal components of the residuals. Thus, we obtain an estimate \tilde{F} as a $T \times r_1$ eigenvector matrix corresponding to the r_1 largest eigenvalues of $M_G X X' M_G / (TN)$ where $M_G \equiv I_T - G(G'G)^{-1}G'$, under the normalization of $\tilde{F}'\tilde{F}/T = I_{r_1}$. We then construct a $T \times r$ matrix of factor estimate $\tilde{H} \equiv [\tilde{F}, G]$ and the lagged variables up to p th order $\tilde{W} \equiv [\tilde{H}_{(-1)}, \dots, \tilde{H}_{(-p)}]$. The factor loadings are estimated using the least squares method $\tilde{C} = X'\tilde{H}(\tilde{H}'\tilde{H})^{-1}$ and the VAR coefficients are estimated by the feasible generalized least squares (FGLS) such that

$$vec(\tilde{A}') = [(\tilde{W}' \otimes I_r)\tilde{\Sigma}_e^{-1}(\tilde{W} \otimes I_r)]^{-1}(\tilde{W}' \otimes I_r)\tilde{\Sigma}_e^{-1}vec(\tilde{H}'),$$

where

$$\tilde{\Sigma}_e = \begin{bmatrix} \tilde{\Sigma}_1 & 0 & \cdots & 0 \\ 0 & \tilde{\Sigma}_2 & & \\ \vdots & & \ddots & \\ 0 & & & \tilde{\Sigma}_T \end{bmatrix},$$

¹⁰This is a common problem when normalizing the structural shocks in FAVARs, e.g., see Stock and Watson (2016).

is a $Tr \times Tr$ block diagonal matrix with $\tilde{\Sigma}_t = \frac{1}{T_b - p} \sum_{t=p+1}^{T_b} \hat{e}_t \hat{e}_t'$ for $t = 1, \dots, T_b$ and $\tilde{\Sigma}_t = \frac{1}{T - T_b} \sum_{t=T_b+1}^T \hat{e}_t \hat{e}_t'$ for $t = T_b + 1, \dots, T$ and \hat{e}_t comprises the ordinary least squares residuals of the regression of \tilde{H} on \tilde{W} . It should be noted that if we denote the random rotation matrix induced by the two-step principal component estimation by R , then it is known that the least squares estimators \tilde{c}_i and \tilde{A}_j are consistent estimators¹¹ for $c_i' R^{-1}$ and $RA_j R^{-1}$ and the following asymptotic normality results hold according to Bai and Ng (2006)¹²

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

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

Remark 1 *Bai et al. (2016) considers the same problem but without explicitly considering the structural identification, i.e., the contemporaneous correlation matrix B . Hence, their setup is the same as our reduced-form model (3) and (4), where they aim to identify the matrix R . We do not separately identify R and B , so our problem is essentially the same as theirs. We explain their approaches in Section 3.2 and Appendix B.*

3.2 Identification of the structural parameters

We propose a method for identifying the structural coefficients and impulse responses. The aim is to consistently estimate C^* , A^* , and $\Theta_{i,h}$ and the asymptotic distribution results are presented.

Algorithm:

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

¹¹In particular, $\|\tilde{c}_i - R'^{-1}c_i\| \xrightarrow{p} 0$ and $\|\tilde{A}_j - RA_j R^{-1}\| \xrightarrow{p} 0$ where R is a random matrix.

¹²In the Proof of Theorem 1 in Appendix A, we show that

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

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

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

$$\begin{aligned}\frac{1}{T_b} \sum_{t=p+1}^{T_b} \tilde{e}_t \tilde{e}_t' &= \tilde{\Omega}_1, \\ \frac{1}{T - T_b} \sum_{t=T_b+1}^T \tilde{e}_t \tilde{e}_t' &= \tilde{\Omega}_2.\end{aligned}$$

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

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

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

4. Estimate the contemporaneous matrix B by

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

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

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

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

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

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

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

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

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

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

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

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

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

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

The asymptotic variance of $\tilde{C}_{1:r}$ is denoted by V_Ξ and it is constructed based on the existing results given by Bai (2003) and Bai and Ng (2006). The asymptotic variance of the eigenvector matrix $\tilde{\Delta}$ denoted by V_Δ is described in Lemma A5 in Appendix A. We now derive the asymptotic distributions of the structural parameters and impulse responses estimators in the following theorem.

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

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

as $N, T \rightarrow \infty$ and $\sqrt{T}/N \rightarrow 0$ for any $i = 1, \dots, N$, where

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

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

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

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

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

with

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

(iii) For the impulse responses:

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

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

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

with

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

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

Several remarks are given pertaining to this result. First, all of the asymptotic variances given above can be consistently estimated by combining the reduced-form estimates, because all of the entities in the expressions are the rotated versions (some of them are cancelled out in the expressions given above). For example $R'^{-1}c_i'$ is consistently estimated by \tilde{c}_i' and $B'R'$ is consistently estimated by \tilde{B}' . Second, we only derive the asymptotic distribution under $\sqrt{T}/N \rightarrow 0$ in this study. Gonçalves and Perron (2014) showed that relaxing this condition in the form of $\sqrt{T}/N \rightarrow c$ where $0 \leq c < \infty$ generates asymptotic bias of the estimators and the inference that dealing explicitly with this asymptotic bias is important. However, we leave this issue for future research. Finally, when there are more than one break in the shock variances, utilizing all these changes yields an overidentifying situation and the proposed method is not directly applicable. A simple solution would be to use the sample covariances in only two regimes in step 3 at a cost of efficiency.

3.3 Strategies employed by Bai et al. (2016)

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

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

for all t . The following three sets of exclusion restrictions are considered.

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

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

It should be noted that Bai et al. (2016) considers the quasi-maximum likelihood method instead of the principal component method for estimating the unobserved factors. However, their identification strategies are valid when the principal component factor estimate is used as they point out, and the steps after factor extraction can be followed exactly. Hence, in the following, we compare our method with that of Bai et al. (2016) by including their identification methods in the two-step estimation using the principal component factor estimate.

3.4 Bootstrap confidence interval

The seminal studies of Gonçalves and Perron (2014) and Yamamoto (2017) show that asymptotic inference of the factor-augmented models may under-evaluate the sampling errors (asymptotic bias) of the coefficient estimators because it is assumed that the asymptotic results are free from factor estimation errors under $\sqrt{T}/N \rightarrow 0$. They develop bootstrap inference as an alternative method that is consistent even under $\sqrt{T}/N \rightarrow c$ where $0 \leq c < \infty$. We also consider the bootstrap method in a similar manner. Thus, we also assume independence¹³ but allow for heteroskedasticity in the idiosyncratic errors u_{it} . The procedure is described as follows.

¹³Gonçalves and Perron (2014) show that the serial correlations in the idiosyncratic errors are irrelevant when we focus on the coefficients of the factor-autoregression model in which the dependent variable is observed. However, this is not the case when some of the dependent variables are the estimated factors as in the FAVAR model.

Bootstrap algorithm:

1. Estimate the reduced-form models (3) and (4), and obtain the parameter estimate \tilde{C} , \tilde{A} , \tilde{B} , and the residuals \tilde{e}_t and \tilde{u}_{it} . Construct the structural impulse response estimator $\hat{\Theta}_{i,h}$ by using the algorithm proposed in Section 3.2.
2. Ensure that the pre- and post- break VAR residuals $\{\tilde{e}_t\}_{t=1}^{T_b}$ and $\{\tilde{e}_t\}_{t=T_b+1}^T$ are demeaned, respectively, in the time direction. Resample with replacements the pre-break residuals $\{\tilde{e}_t\}_{t=1}^{T_b}$ as $r \times 1$ vectors in an i.i.d. manner and label them $\{e_t^b\}_{t=1}^{T_b}$. Construct the same process for the post-break residuals and label them $\{e_t^b\}_{t=T_b+1}^T$. Generate the bootstrapped factors via $h_t^b = \sum_{j=1}^p \tilde{A}_j h_{t-j}^b + e_t^b$, for $t = p + 1, \dots, T$.¹⁴
3. Demean the idiosyncratic residuals \tilde{u}_{it} in both the time and cross-sectional directions. For each $i = 1, \dots, N$, if $\{u_{it}\}_{t=1}^T$ are considered homoskedastic, then we propose the i.i.d. resampling of $\{\tilde{u}_{it}\}_{t=1}^T$ to obtain $\{u_{it}^b\}_{t=1}^T$. If $\{u_{it}\}_{t=1}^T$ are considered heteroskedastic, then use the wild bootstrap $u_{it}^b = \tilde{u}_{it}\eta_{it}$, where $\eta_{it} \sim i.i.d.(0, 1)$ is an external random variable to obtain $\{u_{it}^b\}_{t=1}^T$. Generate the bootstrapped observations $x_{it}^b = \tilde{c}_i h_t^b + u_{it}^b$ for $t = 1, \dots, T$ and $i = 1, \dots, N$.
4. Using the bootstrap observations x_{it}^b , obtain the parameter estimate \tilde{C}^b , \tilde{A}^b , and \tilde{B}^b , by using the same estimation and identification methods proposed in Section 3.2, thereby obtaining the bootstrap estimate of the structural impulse response $\hat{\Theta}_{i,h}^b$.
5. Repeat steps 2 to 4 n_b times and store the recentered statistic $s_{i,h} \equiv \hat{\Theta}_{i,h}^b \zeta - \hat{\Theta}_{i,h} \zeta$, where ζ is an $r \times 1$ vector of shocks of interest.
6. Sort the statistics and select the $100\frac{\alpha}{2}^{th}$ and $100(1 - \frac{\alpha}{2})^{th}$ percentiles $[s_{i,h}^{(\alpha/2)}, s_{i,h}^{(1-\alpha/2)}]$. The resulting $100(1 - \alpha)\%$ confidence interval for $\Theta_{i,h}\zeta$ is $[\hat{\Theta}_{i,h}\zeta - s_{i,h}^{(1-\alpha/2)}, \hat{\Theta}_{i,h}\zeta + s_{i,h}^{(\alpha/2)}]$.

The bootstrap consistency of the method under $\sqrt{T}/N \rightarrow c$ ($0 \leq c < \infty$) would be similar to that shown by Yamamoto (2017) when explicitly deriving the asymptotic bias of the original and bootstrap estimates for the impulse responses, but this problem is rather trivial and it is not presented in this study.

¹⁴The bias-correction method discussed by Kilian (1998) can be applied. The bias is estimated by taking the average of $\hat{A}_j^* - R^* \hat{A}_j R^{*-1}$ in another bootstrap loop, where R^* is a bootstrap analogue of R and it can be constructed based on the original estimate and the bootstrap samples.

4 Monte Carlo experiments

We investigate the finite sample properties of the impulse response estimator by using Monte Carlo simulations. First, we compare the accuracy of estimator obtained using our identification method with that using the three methods based on the IR_a , IR_b and IR_c assumptions proposed by Bai et al. (2016). To ensure that the comparisons are fair and to simplify the implementation, we estimate the unobserved factors by the principal component method instead of the quasi-maximum likelihood method.¹⁵ We also investigate the empirical coverage properties of the asymptotic and bootstrap confidence intervals when the proposed identification method is used. We use the following reduced-form models as the data generating process (DGP):

$$h_t = Ah_{t-1} + e_t, \quad (5)$$

$$x_{it} = c'_i h_t + u_{it}, \quad (6)$$

with $h_t = [f'_t, g'_t]'$ where f_t is an $r_1 \times 1$ vector of unobserved factors and g_t is an $r_2 \times 1$ vector of observed factors. The number of unobserved factors is $r_1 = 2$ and that of the observed factors is $r_2 = 1$, so the total number of factors is $r = r_1 + r_2 = 3$. The order of the VAR model (5) is 1 with the VAR coefficient matrix

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

while the VAR errors are generated by $e_t = B\varepsilon_t$ with the structural shock $\varepsilon_t \sim i.i.d.N(0, \Pi_t)$ and the matrix B is arbitrarily set at

$$B = \begin{bmatrix} 1.0 & 0.4 & 0.8 \\ 0.2 & 1.0 & -0.3 \\ -0.6 & 0.4 & 1.0 \end{bmatrix}, \quad (7)$$

unless stated otherwise. The unconditional shock covariance matrix Π_t changes in the middle of the sample period such that

$$\Pi_t = \begin{cases} \text{diag}(2 + d, 1 + d/2, 0.25), & \text{for } t = 1, \dots, [T/2] \\ I_r & \text{for } t = [T/2] + 1, \dots, T \end{cases},$$

¹⁵As discussed in Section 3.3, this does not affect the qualitative results for the main conclusion.

where the parameter d controls the magnitude of the variance change, and we specifically consider $d = 0.5, 1$, and 2 in this experiment. For the factor model (6), the factor loadings c_i are generated as an $r \times 1$ vector where each element is independently drawn from the standard normal distribution. Furthermore, the idiosyncratic errors are generated by $u_{it} \sim i.i.d.N(0, 1)$. We consider the impulse responses of the N th variable $\Theta_{N,h}$.¹⁶ The number of Monte Carlo replications is 3,000.

4.1 MSE of the impulse response estimator

First, we examine the MSE of the impulse response estimator to one unit structural shock of the observed factor and an unobserved factor. In particular, we compare our method (in the column denoted as “ IR_v ”) with the methods proposed by Bai et al. (2016) (in the columns denoted as “ IR_a ”, “ IR_b ”, and “ IR_c ”). In this experiment, the matrix B for IR_a , IR_b , and IR_c is restricted as follows. Let $\Omega = BB'$ where B is specified by (7). Then, for IR_a , IR_b , and IR_c , we impose the off-diagonal blocks Ω_{12} and Ω_{21} to be zero matrices. Further, for IR_a and IR_b , we impose the upper left $r_1 \times r_1$ block Ω_{11} to be the identity matrix. This results in

$$\Omega = \begin{cases} \begin{bmatrix} I_{r_1} & 0 \\ 0 & \Omega_{22} \end{bmatrix} & \text{for } IR_a \text{ and } IR_b \\ \begin{bmatrix} \Omega_{11} & 0 \\ 0 & \Omega_{22} \end{bmatrix} & \text{for } IR_c \end{cases}, \quad (8)$$

where Ω_{22} is the lower right $r_2 \times r_2$ block of $\Omega = BB'$. After obtaining such Ω , we construct the matrix B by imposing a recursive assumption such that $B = chol(\Omega)$ for IR_a , IR_b , and IR_c , respectively. Hence, the true or correct B matrix can be different in different identification strategies.

The DGPs are different when diverse identification methods are compared, so it is difficult to determine whether the difference in MSEs is due to the DGP or the identification method. To address this concern, we consider the following four scenarios. In scenario A, the DGPs conform with the identification assumptions, and thus the four methods are applied to different DGPs. In scenario B, the DGP uses the unrestricted B matrix (7) for every

¹⁶The factor loadings of the first r response variables are used for identification and they can be different according to the identification method employed in this experiment. The remaining factor loadings are the same even when different identification methods are considered. The exception is IR_a where the factor loading matrix has a block structure (if Σ_u is diagonal). In this experiment the first factor loading of the N th response variable is zero.

method, and thus the method based on IR_v is regarded as a correct method, whereas IR_a , IR_b , and IR_c are not and the latter is affected by misspecification. In scenario C, the DGP uses the matrix B based on IR_a and IR_b (which are the same). In scenario D, the DGP uses the B matrix based on IR_c . As shown in (8), scenario C is the case with the most restrictions and all four methods satisfy their assumptions. Scenario D is somewhat less restrictive than scenario C because the Ω_{11} part is unrestricted, so the methods based on the IR_a and IR_b assumptions involve the misspecification, whereas the methods based on the IR_c and IR_v assumptions are correctly specified.

By contrast, the identification assumptions on the factor model are correctly imposed in every case as follows. When we use the method based on IR_a , the first $[0.7N]$ response variables have zero second factor loadings and the last $[0.3N]$ response variables have zero first factor loadings, so $\Lambda'\Sigma_u\Lambda/N$ becomes a diagonal matrix. When we use the method based on IR_b , we put three zeros in the factor loading matrix so that the first r_1 response variables $\Lambda_{1:r_1}$ becomes a lower triangular matrix. When we use the method based on IR_c , $\Lambda_{1:r_1}$ is the identity matrix. When we use the method based on IR_v , we follow Assumption D and sets $C_{1:r}$ such that the main diagonal elements of $C_{1:r}B$ are unity.¹⁷ We use the impulse response of the N th variable, so the objectives are the same regardless of the assumptions, except for IR_a where the first factor loading is restricted to be zero. However, that has very little effect on the MSEs.

Table 1 shows the MSEs of the impulse response estimator of the time horizon $h = 1$ for scenarios A to D. For each scenario, the left panel presents the MSEs of the impulse response estimator to a shock of the observed factor and the right panel shows the MSEs of the impulse response estimator to a shock of the first unobserved factor. We investigated the MSEs for different values of the horizon h , but the results were qualitatively similar, so we only present the case where $h = 1$.¹⁸ In each table, we consider the sample size pairs $(T, N) = (50, 50), (50, 150), (150, 50)$, and $(150, 150)$.

Throughout the four scenarios, the MSE of IR_v tends to decrease as d increases because the eigenvalues of $\Pi_1\Pi_2^{-1}$ are more distinct and the asymptotic variance of the eigenvectors will become smaller.¹⁹ By contrast, the MSEs of IR_a , IR_b , and IR_c tend to increase as

¹⁷First, we compute $C_{1:r}B$ and set its diagonal elements as one and get back $C_{1:r}$ by right-multiplying B^{-1} .

¹⁸When the sample size and the break size are not large, the MSE of IR_v may be driven by a very few large errors and the results may be unstable. Hence, we report the MSEs after truncating the 1% largest and 1% smallest squared errors. However, this does not affect qualitative results of the Monte Carlo simulation.

¹⁹This is consistent with Lemma A5, in which the asymptotic variance is shown to have $(\pi_{1/2}^{(k)} - \pi_{1/2}^{(j)})^2$ as the denominator.

d becomes large because the variance of the VAR errors increases in (the first half of) the sample period, thereby resulting in larger sampling errors for the impulse response estimator.

We then compare the MSEs of the four identification methods in each scenario. In scenario A, when the shock of the observed factor is considered, the MSE of IR_v is larger than that with any other method for every sample size pair. However, it becomes more comparable to the others, especially with IR_b , under a large break $d = 2$ with a large time dimension $T = 150$. A larger N also reduces the MSE except in a few cases, but its effect is more blurred than that of T because the estimator is \sqrt{T} consistent. When the shock of the first unobserved factor is considered, the MSEs of IR_a and IR_b increase whereas those of IR_c decrease compared with the shock of the observed factor case. The MSE of IR_v is much smaller than those of IR_a and IR_b and it outweighs that of IR_c when T is large.

Scenario B is the most interesting experiment because the matrix B is fully unrestricted and it is the same for all of the identification methods. When the shock of the observed factor is considered, the MSEs of IR_a , IR_b , and IR_c are considerably larger than those in scenario A, whereas the MSEs of IR_v does not increase. This is because the methods based on IR_a , IR_b , and IR_c suffer from misspecification whereas the method based on IR_v does not. When we consider the shock of the first unobserved factor, IR_v still tends to give the smallest MSE when $T = 150$.

In scenario C, the matrix B has the most restrictions, and thus the relative merit of IR_v is expected to be the smallest among the four scenarios. However, we find that IR_v gives MSEs that are comparable with those of IR_a , IR_b , and IR_c when d and T are large and the shock of the observed factor is considered. The results remained qualitatively similar even when the shock of the first unobserved factor is considered. Finally, the results obtained under scenario D agree with those under scenario C even when matrix B is slightly more restricted.

In summary, the Monte Carlo results confirm that the impulse responses are estimated more accurately using the proposed method as the magnitude of the variance change increases and as the sample size grows. More importantly, the proposed method is robust to the misspecification of the matrix B unlike the existing methods proposed by Bai et al. (2016). Even when the matrix B is correctly specified, the MSEs with the proposed method are comparable with those as d and T become large.

4.2 Coverage property of the confidence interval

Next, we investigate the coverage property of the confidence interval for the impulse responses identified using the proposed method. Table 2 presents the empirical coverage rate and the median value of the length, i.e., the upper confidence limit minus the lower confidence limit, for the 95% asymptotic confidence interval based on 3,000 Monte Carlo replications. Given that the order of VAR is 1, it is sufficient to investigate the horizon h up to 5. The same set of sample size pairs described in the previous subsection, i.e., $(T, N) = (50, 50), (50, 150), (150, 50),$ and $(150, 150)$ are used to study the sample size effect. The upper table shows the results for the impulse responses to a unit shock of the observed factor and the lower table shows the impulse responses to a unit shock of the first unobserved factor. When we consider the shock of the observed factor, the coverage rate is close to the nominal level of 95% in every case, validating Theorem 2. Importantly, when we consider the responses to an unobserved factor shock as shown in the lower table, the asymptotic confidence interval tends to fall short of the nominal coverage rate. This under-coverage property is exacerbated when the sample size is small, especially when N is smaller than T ($T = 150, N = 50$) because the condition that $\sqrt{T}/N \rightarrow 0$ is considered less appropriate in this empirical situation. Unsurprisingly, the coverage rate deteriorates as h increases, because the degree of nonlinearity of impulse responses increases and this undermines the asymptotic approximation. The median length decreases as d becomes larger because the asymptotic variance decreases for the same reason explained in the previous subsection. We also observe that the median length decreases as T increases, which is consistent with the \sqrt{T} rate of the estimator, whereas the effect of a large N is more blurred.

In the aforementioned studies of bootstrapping factor models, it is suggested that bootstrap methods can improve the coverage property. To examine this suggestion, Table 3 shows the coverage rate and median length of the 95% bootstrap confidence interval, as proposed in Section 3.3. We consider the i.i.d. bootstrap based on 499 repetitions. The upper table again shows the results for a unit shock of the observed factor, where the coverage rate is very close to the nominal level of 95% and the median length is somewhat longer than the asymptotic interval, thereby reflecting the factor estimation errors. The lower table shows the results with respect to a unit shock of the first unobserved factor. Clearly, the bootstrap interval has a coverage rate that is much closer to the nominal level than the asymptotic interval. As suggested in the seminal study of Gonçalves and Perron (2014), the bootstrap accounts for the effects of factor estimation errors that are not accounted for by the asymptotic interval under the assumption of $\sqrt{T}/N \rightarrow 0$. A well-known caveat that the coverage

rate can deteriorate as h becomes larger even with the bootstrap method is also relevant in this case. This could potentially be fixed by introducing the double bootstrap bias correction procedure proposed by Kilian (1998).

5 Empirical illustration

We apply the proposed identification method to the well-known empirical study by Bernanke et al. (2005), who investigated the effect of a contractionary monetary shock on macroeconomic variables using the same FAVAR model as (3), with the federal funds rate (FFR) as the observable factor g_t and the unobservable factors f_t extracted from 120 macroeconomic data series for the United States. The aim of this study is to examine whether we can obtain impulse response estimates similar to those in the original study by using the proposed identification method.

Thus, we first explain their identification strategy. Bernanke et al. (2005) assumed a recursive structure among the unobserved and observed factors. They first extract the unobserved factors from a subset of “slow-moving variables” among the 120 macroeconomic variables. The remaining unobserved factors are identified as the residuals from the regression of the standard principal component estimate on the slow-moving factors and g_t . Clearly, the identified unobserved factors are orthogonal to the observed factor and the slow-moving factors. Our method does not impose these restrictions but a change in the unconditional shock variances is accounted for in December 1984. The timing of this change in the variance is consistent with the evidence for the so called Great Moderation²⁰ and a reasonable variation in the timing does not qualitatively affect our results.²¹ To fix the scale of the shocks, we use the point estimate of the reduced-form factor loading estimate for the main diagonal elements of the contemporaneous impulse responses. Our specific choice is based on the variables that take a large value for each factor loading. Our specific choices comprise the industrial production (final products) for the first factor, the NYSE common stock price index (composite) for the second factor, the producer price index (finished goods) for the third factor, and the spread of FYGT1 from FYFF for the observed factor. As discussed in Section 2, this choice may affect the empirical results and making this step more robust

²⁰Stock and Watson (2002) document that about 40% of 168 U.S. macroeconomic variables experienced a change of the variance during 1983 and 1985.

²¹Han and Inoue (2015) pointed out that common breaks in the factor loadings can increase the number of factors and Yamamoto (2016) investigated a consequence of this fact in a forecasting exercise. Any criticism of the time-invariant factor loading assumption may apply to this example, so to the original study of Bernanke et al. (2005).

should be addressed in future research.

Following the original study, we use the VAR lag order $p = 13$ and the number of unobserved factors $r = 3$. We use exactly the same data employed by Bernanke et al. (2005), which ranges from January 1959 to August 2001. A 25-basis point increase in the FFR is investigated, and the impulse responses and associated bootstrap confidence intervals are retransformed to level. The confidence level is 90% and the number of bootstrap repetitions is 1,000. We could compute the impulse responses of the 120 variables, but given space limitations, the following six selected variables are presented: FFR, industrial production index (total, IP), consumer price index (all item; CPI), monetary base, exchange rate yen, commodity price index up to 48 periods ahead as shown in Figure 1.

First, we find that the directions of the point estimates of the impulse responses are largely consistent with those in the original study.²² This suggests that imposing the identification assumption of Bernanke et al. (2005) has little effect on their main results. Second, not surprisingly, the confidence intervals with our identification method are somewhat wider than those in the original study in most cases. We recall that the MSE with our method tends to become larger when the DGP is restricted correctly (in Scenarios C and D) in our Monte Carlo investigations, so this finding also supports the identification strategy implemented by Bernanke et al. (2005).

6 Conclusion

In this study, we develop a new structural identification method for FAVAR models. The main idea is based on previous studies that use changes in the unconditional shock variances on a historical date. Similar to the methods developed by Bai et al. (2016), the proposed method can identify the rotated version of the contemporaneous matrix, but it is advantageous because it allows the contemporaneous matrix to be fully unrestricted. Our method can utilize statistical methods to verify the identification assumptions rather than relying on economically motivated assumptions. We derive the asymptotic distributions of the structural parameter and impulse response estimators under $\sqrt{T}/N \rightarrow 0$ as $N, T \rightarrow \infty$ and propose a bootstrap method for constructing the confidence interval. Monte Carlo experiments suggest that the MSEs with the proposed method are larger than the existing methods if the identifying assumptions of the latter is indeed correct. However, the proposed method obtains much smaller MSEs when the contemporaneous matrix is indeed unrestricted. This

²²We can directly compare the results with Figure II (page 408) of Bernanke et al. (2005).

feature becomes more distinct when the magnitude of the change in the variance is large and as T increases. Both the asymptotic and bootstrap confidence intervals provide a satisfactory coverage rate when the impulse response to a shock of the observed factor is investigated. The bootstrap interval gives more accurate coverage rate when a shock of an unobserved factor is considered, especially when N is relatively smaller than T , as suggested in the previous studies. Finally, our empirical investigation of the same model employed by Bernanke et al. (2005) provides similar point estimates and wider confidence intervals in general. These findings support the validity of the identification strategy used by Bernanke et al. (2005) in their context. Several important issues need to be addressed in future research. The asymptotic theory under $\sqrt{T}/N \rightarrow c$ ($0 \leq c < \infty$) must be developed to justify the bootstrap methods. More importantly, the new identification strategy can potentially be applied to empirical data given the large number of studies of structural VARs reviewed by Stock and Watson (2016) and Kilian and Lütkepohl (2017), e.g., especially in a data-rich environment.

Appendix A: Proof of Theorems

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

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

where $\theta_{T,i} \equiv [(G'G)^{-1}G'F\lambda_i + \gamma_i]$. Because M_GF and G are orthogonal, by using Theorem 1 of Bai and Ng (2006), the reduced-form factor loading estimator has the following property

$$\tilde{c}_i \approx \begin{bmatrix} Q^{-1} & 0_{r_1 \times r_2} \\ 0_{r_2 \times r_1} & I_{r_2} \end{bmatrix} \begin{bmatrix} \lambda_i \\ \theta_{T,i} \end{bmatrix},$$

so that

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

where $Q \equiv V^{-1}(\tilde{F}'M_GF/T)(\Lambda'\Lambda/N)$ and V is an $r_1 \times r_1$ diagonal matrix where the main diagonal elements are the r_1 largest eigenvalues of $M_GXX'M_G/(TN)$. Thus, we have

$$\tilde{\Xi} \equiv \tilde{C}_{1:r} \xrightarrow{p} C_{1:r}\bar{R}^{-1} \equiv \Xi\bar{R}^{-1}, \quad (\text{A.1})$$

where

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

and $\bar{R} = p \lim(R)$. This implies that the inverse has the probability limit

$$\tilde{\Xi}^{-1} \xrightarrow{p} \bar{R}\Xi^{-1}. \quad (\text{A.2})$$

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

$$\tilde{\Omega}_j \xrightarrow{p} \bar{R}B\Pi_j B'\bar{R}', \quad (\text{A.3})$$

for $j = 1$ and 2 because $e_t = B\varepsilon_t$ and Assumption A (i).

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

$$\begin{aligned}\tilde{S} &\equiv \tilde{\Xi}\tilde{\Omega}_1\tilde{\Omega}_2^{-1}\tilde{\Xi}^{-1} \xrightarrow{p} (\Xi\bar{R}^{-1})(\bar{R}B\Pi_1B'\bar{R}')(\bar{R}B\Pi_2B'\bar{R}')^{-1}(\bar{R}\Xi^{-1}), \\ &= \Xi B\Pi_{1/2}B^{-1}\Xi^{-1} \equiv S,\end{aligned}$$

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

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

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

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

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

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

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

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

and

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

The results follow. ■

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

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

(i)

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

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

(ii)

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

where

$$V_A \equiv p \lim T[\{(I_p \otimes R')W'\} \otimes I_r]^{-1}[(I_T \otimes R)\Sigma_e(I_T \otimes R')][\{W(I_p \otimes R)\} \otimes I_r]^{-1}.$$

Proof of Lemma A1: (i) This is a direct result of Theorem 1 of Bai and Ng (2006).

(ii) First, we note that

$$\begin{aligned}\tilde{H} &\equiv HR + (\tilde{H} - HR), \\ &= WAR + eR + WAR + (H - HR), \\ &= \tilde{W}(I_p \otimes R^{-1})AR + eR + [W - \tilde{W}(I_p \otimes R^{-1})]AR + (\tilde{H} - HR).\end{aligned}$$

Consider the FGLS estimator. To simplify the notation, we let $D \equiv [(\tilde{W}' \otimes I_r)\tilde{\Sigma}_e^{-1}(\tilde{W} \otimes I_r)]$. Then,

$$\begin{aligned}vec(\tilde{A}') &= D^{-1}(\tilde{W}' \otimes I_r)\tilde{\Sigma}_e^{-1}vec(\tilde{H}'), \\ &= D^{-1}(\tilde{W}' \otimes I_r)\tilde{\Sigma}_e^{-1}vec[R'A'(I_p \otimes R^{-1})\tilde{W}'], \\ &\quad + D^{-1}(\tilde{W}' \otimes I_r)\tilde{\Sigma}_e^{-1}vec(R'e'), \\ &\quad + D^{-1}(\tilde{W}' \otimes I_r)\tilde{\Sigma}_e^{-1}vec[R'A'\{W' - (I_p \otimes R^{-1})\tilde{W}'\}], \\ &\quad + D^{-1}(\tilde{W}' \otimes I_r)\tilde{\Sigma}_e^{-1}vec(\tilde{H}' - R'H), \\ &= I + II + III + IV.\end{aligned}$$

For I , because

$$vec[R'A'(I_p \otimes R^{-1})\tilde{W}'] = (\tilde{W} \otimes I_r)vec[R'A'(I_p \otimes R^{-1})],$$

we obtain

$$I = vec[R'A'(I_p \otimes R^{-1})].$$

Hence,

$$\sqrt{T}vec[R'A'(I_p \otimes R^{-1})\tilde{W}'] = \sqrt{T}(II + III + IV).$$

Terms III and IV are factor estimation errors and they diminish as $N, T \rightarrow \infty$ under $\sqrt{T}/N \rightarrow 0$. This yields

$$\begin{aligned}AVar(\sqrt{T}vec[R'A'(I_p \otimes R^{-1})\tilde{W}']) \\ = p \lim(TD^{-1})(\tilde{W}' \otimes I_r)\tilde{\Sigma}_e^{-1}[T^{-1}vec(R'e')vec(R'e')']\tilde{\Sigma}_e^{-1}(\tilde{W} \otimes I_r)(TD^{-1}).\end{aligned}$$

However, because

$$E(\tilde{\Sigma}_e) = E[vec(R'e')vec(R'e')'],$$

we have

$$p \lim(\tilde{W}' \otimes I_r)\tilde{\Sigma}_e^{-1}[vec(R'e')vec(R'e')'/T]\tilde{\Sigma}_e^{-1}(\tilde{W} \otimes I_r) = p \lim T^{-1}D,$$

or

$$AVar(\sqrt{T}vec[R'A'(I_p \otimes R^{-1})\tilde{W}']) = p \lim(TD^{-1}).$$

We note that

$$D \approx [\{(I_p \otimes R)W'\} \otimes I_r][(I_T \otimes R)\Sigma_e(I_T \otimes R)]^{-1}[\{W(I_p \otimes R)\} \otimes I_r].$$

The result follows. ■

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

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

where

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

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

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

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

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

where

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

with

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

where D_r is a duplication matrix of dimension r .

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

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

where

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

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

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

where

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

Furthermore,

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

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

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

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

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

where

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

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

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

$$\begin{aligned} \tilde{Q} &\approx \bar{\Xi} \bar{\Omega}_{1/2} \bar{\Xi}^{-1} \\ &= (\Xi \bar{R}^{-1}) (\bar{R} \bar{\Omega}_1 \bar{\Omega}_2^{-1} \bar{R}^{-1}) (\bar{R} \Xi^{-1}) \\ &= \Xi \bar{\Omega}_1 \bar{\Omega}_2^{-1} \Xi^{-1} \equiv Q. \end{aligned}$$

Using Lemmas A2 and A3, the delta method yields

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

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

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

Furthermore,

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

and

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

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

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

$$\sqrt{T} \text{vec}(\tilde{\Delta} - C_{1:r}^*) \xrightarrow{d} N(0, V_\Delta),$$

where

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

with

$$\begin{aligned} V_{\delta_{kk}} &\equiv \sum_{j=1, j \neq k}^r \frac{A \text{Var}(v_{ii})}{(\pi_{1/2}^{(k)} - \pi_{1/2}^{(j)})^2} c_{\bullet j}^* c_{\bullet j}^{*'}, \\ V_{\delta_{kl}} &\equiv -\frac{A \text{Var}(v_{kl})}{(\pi_{1/2}^{(k)} - \pi_{1/2}^{(j)})^2} c_{\bullet k}^* c_{\bullet l}^{*'}, \quad \text{for } k \neq l, \end{aligned}$$

v_{ij} is the (i, j) th element of $\sqrt{T}(\tilde{Q} - Q)$ and $c_{\bullet j}^*$ is the j th column of $C_{1:r}^*$.

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

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

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

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

so that

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

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

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

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

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

Because $\tilde{\delta}_k$ is the k th column of $C_{1,r}^*E$,

$$\sqrt{T}(\tilde{\delta}_k - c_{\bullet j}^*) \xrightarrow{d} N \left(0, \sum_{j=1, j \neq k}^r \frac{AVar(v_{ij})}{(\pi_{1/2}^{(k)} - \pi_{1/2}^{(j)})^2} c_{\bullet j}^* c_{\bullet j}^{*'} \right),$$

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

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

The result follows. ■

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

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

but

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

and

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

yield the results.

(ii) The delta method yields

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

but

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

and

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

yield the result.

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

Appendix B: Estimation under the IR_a , IR_b , and IR_c assumptions

We obtain \tilde{B} under the three sets of identification assumptions given by Bai et al. (2016) as follows.

1. Estimate the reduced-form models to obtain \tilde{A} , \tilde{C} , and \tilde{e}_t . Let $\tilde{C}_{1:r}$ be the first r rows of C .
2. Construct the sample covariance matrices of the VAR residuals for the whole sample by

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

and obtain its partitioned components $\tilde{\Omega}_{11}$, $\tilde{\Omega}_{12}$, $\tilde{\Omega}_{21}$, and $\tilde{\Omega}_{22}$. Let

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

IRa Let V_a be the eigenvector matrix of

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

in descending order of their associated eigenvalues. Calculate

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

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

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

IRc Let

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

References

- Anderson, T.W., 2003, An introduction to multivariate statistical analysis: third edition, Wiley.
- Bai, J., 2003, Inferential theory for factor models of large dimensions, *Econometrica* 71(1), 135-171.
- Bai, J., K. Li, and L. Lu, 2016, Estimation and inference on FAVAR models, *Journal of Business and Economic Statistics* 34(4), 620-641.
- Bai, J., and S. Ng, 2006, Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions, *Econometrica* 74(4), 1133-1150.
- Bai, J., and S. Ng, 2013, Principal components estimation and identification of the factors, *Journal of Econometrics* 176, 18-29.
- Bernanke, B.S., J. Boivin, and P. Eliasch, 2005, Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach, *Quarterly Journal of Economics* 120(1), 387-422.
- Breitung, J. and S. Eickmeier, 2011, Testing for structural breaks in dynamic factor models, *Journal of Econometrics* 163, 71-84.
- Brunnermeier, M., D. Palia, K.A. Sastry, and C.A. Sims, 2017, Feedbacks: Financial markets and economic activity, Working Paper, Princeton University.
- Casini A. and P. Perron, 2017, Structural breaks in time series, Unpublished manuscript, Boston University.
- Chen, L., J.J.Dolado. and J. Gonzalo, 2014, Detecting big structural breaks in large factor models, *Journal of Econometrics* 180, 30-48.
- Djogbenou, A., S. Gonçalves, and B. Perron, 2015, Bootstrap inference in regressions with estimated factors and serial correlation, *Journal of Time Series Analysis* 36, 481-502.
- Gonçalves, S., and B. Perron, 2014, Bootstrapping factor-augmented regression models, *Journal of Econometrics* 182, 156-173.
- Gonçalves, S., and B. Perron, 2016, Bootstrapping factor models with cross-sectional dependence, mimeo.
- Gonçalves, S., B. Perron, and A. Djogbenou, 2017, Bootstrap prediction intervals for factor models, *Journal of Business and Economic Statistics* 35(1), 53-69.

- Han, X. and A. Inoue, 2015, Tests for parameter instability in dynamic factor models, *Econometric Theory* 31, 1117-1152.
- Kilian, L., 1998, Small sample confidence intervals for impulse response functions, *Review of Economics and Statistics* 80(2), 218-230.
- Kilian, L., and H. Lütkepohl, 2017, Structural vector autoregressive analysis, Cambridge University Press (forthcoming), <http://www-personal.umich.edu/~lkilian/book.html>
- Lanne, M. and H. Lütkepohl, 2008, Identifying monetary policy shocks via changes in volatility, *Journal of Money, Credit and Banking* 40, 1131-1149.
- Lewis D.J., 2018, Identifying shocks via time-varying volatility, Harvard University.
- Lütkepohl, H., 2005, New introduction to multiple time series analysis, Springer.
- Lütkepohl, H., 2013, Identifying structural vector autoregressions via changes in volatility, *Advances in Econometrics* 32, 169-203.
- Lütkepohl, H. and A. Netsunajev, 2017, Structural vector autoregressions with heteroskedasticity: A review of different volatility models, *Economics and Statistics* 1, 2-18.
- Qu, Z. and P. Perron, 2007, Estimating and testing structural changes in multivariate regressions, *Econometrica* 75(2), 459-502.
- Rigobon, R., 2003, Identification through heteroskedasticity, *The Review of Economics and Statistics* 85(4), 777-792.
- Rigobon, R. and Sack, B., 2003, Measuring the effect of monetary policy to the stock market, *Quarterly Journal of Economics* 118, 639-669.
- Sims, C.A., 1980, Macroeconomic and reality, *Econometrica* 48(1), 1-48.
- Stock, J.H., and M.W. Watson, 2002, Has the business cycle changed and why? NBER Macroeconomic Annual 17, 159-218.
- Stock, J.H., and M.W. Watson, 2016, Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. Handbook of Macroeconomics, J.B. Taylor and H. Uhlig Eds. Elsevier.
- Yamamoto, Y., 2016, Forecasting with non-spurious factors in U.S. macroeconomic time series, *Journal of Business and Economic Statistics* 34(1), 81-106.
- Yamamoto, Y., 2017, Bootstrap inference for impulse response functions in factor-augmented vector autoregressions, HIAS Discussion Paper series E-4.

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

Table 1. Mean squared errors of the impulse response estimator

Scenario A: Correct DGP

| Responses to a unit shock of the observed factor | | | | | |
|--|-----|-------|-------|-------|-------|
| | d | IRa | IRb | IRc | IRv |
| T=50, N=50 | 0.5 | 0.068 | 0.126 | 0.116 | 0.425 |
| | 1.0 | 0.067 | 0.132 | 0.114 | 0.278 |
| | 2.0 | 0.066 | 0.128 | 0.113 | 0.262 |
| T=50, N=150 | 0.5 | 0.065 | 0.125 | 0.104 | 0.274 |
| | 1.0 | 0.063 | 0.122 | 0.104 | 0.264 |
| | 2.0 | 0.068 | 0.132 | 0.112 | 0.260 |
| T=150, N=50 | 0.5 | 0.021 | 0.041 | 0.034 | 0.054 |
| | 1.0 | 0.022 | 0.040 | 0.034 | 0.050 |
| | 2.0 | 0.022 | 0.045 | 0.036 | 0.051 |
| T=150, N=50 | 0.5 | 0.022 | 0.043 | 0.035 | 0.054 |
| | 1.0 | 0.021 | 0.043 | 0.035 | 0.054 |
| | 2.0 | 0.022 | 0.043 | 0.034 | 0.048 |

| Responses to a unit shock of the first unobserved factor | | | | | |
|--|-----|-------|-------|-------|-------|
| | d | IRa | IRb | IRc | IRv |
| T=50, N=50 | 0.5 | 0.521 | 0.705 | 0.084 | 0.289 |
| | 1.0 | 0.619 | 0.751 | 0.077 | 0.235 |
| | 2.0 | 0.589 | 0.815 | 0.074 | 0.304 |
| T=50, N=150 | 0.5 | 0.496 | 0.657 | 0.084 | 0.318 |
| | 1.0 | 0.528 | 0.759 | 0.080 | 0.304 |
| | 2.0 | 0.641 | 0.895 | 0.076 | 0.241 |
| T=150, N=50 | 0.5 | 0.505 | 0.730 | 0.040 | 0.033 |
| | 1.0 | 0.567 | 0.914 | 0.042 | 0.030 |
| | 2.0 | 0.588 | 1.048 | 0.039 | 0.026 |
| T=150, N=50 | 0.5 | 0.451 | 0.770 | 0.037 | 0.030 |
| | 1.0 | 0.532 | 0.835 | 0.037 | 0.029 |
| | 2.0 | 0.583 | 1.031 | 0.038 | 0.026 |

Scenario B: Unrestricted DGP

| Responses to a unit shock of the observed factor | | | | | |
|--|-----|-------|-------|-------|-------|
| | d | IRa | IRb | IRc | IRv |
| T=50, N=50 | 0.5 | 0.414 | 0.543 | 0.449 | 0.425 |
| | 1.0 | 0.455 | 0.614 | 0.494 | 0.278 |
| | 2.0 | 0.472 | 0.662 | 0.527 | 0.262 |
| T=50, N=150 | 0.5 | 0.405 | 0.550 | 0.453 | 0.274 |
| | 1.0 | 0.431 | 0.584 | 0.469 | 0.264 |
| | 2.0 | 0.487 | 0.648 | 0.521 | 0.261 |
| T=150, N=50 | 0.5 | 0.384 | 0.481 | 0.413 | 0.054 |
| | 1.0 | 0.418 | 0.532 | 0.445 | 0.049 |
| | 2.0 | 0.512 | 0.637 | 0.542 | 0.052 |
| T=150, N=50 | 0.5 | 0.389 | 0.473 | 0.411 | 0.054 |
| | 1.0 | 0.421 | 0.530 | 0.457 | 0.052 |
| | 2.0 | 0.507 | 0.618 | 0.530 | 0.049 |

| Responses to a unit shock of the first unobserved factor | | | | | |
|--|-----|-------|-------|-------|-------|
| | d | IRa | IRb | IRc | IRv |
| T=50, N=50 | 0.5 | 0.678 | 0.868 | 0.070 | 0.280 |
| | 1.0 | 0.693 | 0.849 | 0.068 | 0.295 |
| | 2.0 | 0.756 | 1.036 | 0.068 | 0.220 |
| T=50, N=150 | 0.5 | 0.644 | 0.893 | 0.068 | 0.276 |
| | 1.0 | 0.685 | 0.862 | 0.070 | 0.335 |
| | 2.0 | 0.786 | 0.939 | 0.066 | 0.318 |
| T=150, N=50 | 0.5 | 0.684 | 1.023 | 0.031 | 0.034 |
| | 1.0 | 0.788 | 1.103 | 0.033 | 0.040 |
| | 2.0 | 0.847 | 1.125 | 0.032 | 0.030 |
| T=150, N=50 | 0.5 | 0.699 | 0.944 | 0.032 | 0.034 |
| | 1.0 | 0.759 | 1.050 | 0.032 | 0.032 |
| | 2.0 | 0.763 | 1.103 | 0.030 | 0.025 |

Table 1. Mean squared errors of the impulse response estimator (continued)

Scenario C: DGP with the IR_a and IR_b assumptions

| Responses to a unit shock of the observed factor | | | | | | Responses to a unit shock of the first unobserved factor | | | | | |
|--|-----|-------|-------|-------|-------|--|-----|-------|-------|-------|-------|
| | d | IRa | IRb | IRc | IRv | | d | IRa | IRb | IRc | IRv |
| T=50, N=50 | 0.5 | 0.068 | 0.126 | 0.094 | 0.373 | T=50, N=50 | 0.5 | 0.536 | 0.718 | 0.046 | 0.300 |
| | 1.0 | 0.067 | 0.132 | 0.093 | 0.266 | | 1.0 | 0.563 | 0.738 | 0.040 | 0.287 |
| | 2.0 | 0.066 | 0.126 | 0.092 | 0.219 | | 2.0 | 0.656 | 0.850 | 0.038 | 0.296 |
| T=50, N=150 | 0.5 | 0.061 | 0.116 | 0.085 | 0.251 | T=50, N=150 | 0.5 | 0.507 | 0.661 | 0.044 | 0.272 |
| | 1.0 | 0.069 | 0.126 | 0.094 | 0.246 | | 1.0 | 0.542 | 0.749 | 0.045 | 0.347 |
| | 2.0 | 0.071 | 0.141 | 0.099 | 0.242 | | 2.0 | 0.568 | 0.762 | 0.038 | 0.343 |
| T=150, N=50 | 0.5 | 0.022 | 0.042 | 0.029 | 0.050 | T=150, N=50 | 0.5 | 0.547 | 0.858 | 0.011 | 0.034 |
| | 1.0 | 0.022 | 0.041 | 0.029 | 0.050 | | 1.0 | 0.532 | 0.833 | 0.011 | 0.035 |
| | 2.0 | 0.023 | 0.043 | 0.029 | 0.049 | | 2.0 | 0.639 | 1.003 | 0.011 | 0.032 |
| T=150, N=50 | 0.5 | 0.019 | 0.039 | 0.027 | 0.045 | T=150, N=50 | 0.5 | 0.516 | 0.831 | 0.012 | 0.035 |
| | 1.0 | 0.021 | 0.042 | 0.029 | 0.048 | | 1.0 | 0.513 | 0.835 | 0.011 | 0.028 |
| | 2.0 | 0.021 | 0.045 | 0.029 | 0.048 | | 2.0 | 0.587 | 1.079 | 0.011 | 0.031 |

Scenario D: DGP with the IR_c assumption

| Responses to a unit shock of the observed factor | | | | | | Responses to a unit shock of the first unobserved factor | | | | | |
|--|-----|-------|-------|-------|-------|--|-----|-------|-------|-------|-------|
| | d | IRa | IRb | IRc | IRv | | d | IRa | IRb | IRc | IRv |
| T=50, N=50 | 0.5 | 0.080 | 0.162 | 0.116 | 0.471 | T=50, N=50 | 0.5 | 0.821 | 1.010 | 0.082 | 0.385 |
| | 1.0 | 0.079 | 0.169 | 0.114 | 0.322 | | 1.0 | 0.795 | 1.071 | 0.077 | 0.358 |
| | 2.0 | 0.079 | 0.161 | 0.113 | 0.273 | | 2.0 | 0.898 | 1.270 | 0.078 | 0.351 |
| T=50, N=150 | 0.5 | 0.077 | 0.156 | 0.104 | 0.308 | T=50, N=150 | 0.5 | 0.774 | 1.067 | 0.084 | 0.482 |
| | 1.0 | 0.075 | 0.153 | 0.103 | 0.312 | | 1.0 | 0.805 | 1.175 | 0.080 | 0.325 |
| | 2.0 | 0.081 | 0.171 | 0.114 | 0.318 | | 2.0 | 0.910 | 1.275 | 0.074 | 0.322 |
| T=150, N=50 | 0.5 | 0.025 | 0.052 | 0.034 | 0.063 | T=150, N=50 | 0.5 | 0.744 | 1.080 | 0.041 | 0.040 |
| | 1.0 | 0.024 | 0.051 | 0.033 | 0.056 | | 1.0 | 0.933 | 1.299 | 0.042 | 0.035 |
| | 2.0 | 0.026 | 0.056 | 0.036 | 0.061 | | 2.0 | 1.025 | 1.455 | 0.041 | 0.035 |
| T=150, N=50 | 0.5 | 0.025 | 0.051 | 0.034 | 0.059 | T=150, N=50 | 0.5 | 0.777 | 1.076 | 0.038 | 0.038 |
| | 1.0 | 0.025 | 0.052 | 0.035 | 0.061 | | 1.0 | 0.837 | 1.308 | 0.041 | 0.038 |
| | 2.0 | 0.025 | 0.053 | 0.033 | 0.055 | | 2.0 | 0.924 | 1.438 | 0.040 | 0.036 |

Table 2. Coverage rate and median length of the asymptotic confidence interval

Responses to a shock of the observed factor

| | | Coverage rate | | | | | Median length | | | | |
|--------------|-----|---------------|------|------|------|------|---------------|------|------|------|------|
| | d | h=1 | 2 | 3 | 4 | 5 | h=1 | 2 | 3 | 4 | 5 |
| T=50, N=50 | 0.5 | 97.7 | 94.6 | 94.6 | 93.5 | 91.5 | 2.63 | 1.69 | 1.06 | 0.70 | 0.45 |
| | 1.0 | 98.5 | 95.1 | 94.8 | 93.8 | 91.8 | 2.54 | 1.66 | 1.08 | 0.71 | 0.45 |
| | 2.0 | 98.6 | 96.1 | 96.2 | 94.6 | 92.7 | 2.46 | 1.66 | 1.09 | 0.72 | 0.47 |
| T=50, N=150 | 0.5 | 97.5 | 95.2 | 95.2 | 93.8 | 92.2 | 2.56 | 1.68 | 1.07 | 0.69 | 0.45 |
| | 1.0 | 98.4 | 95.6 | 95.5 | 93.0 | 91.9 | 2.54 | 1.69 | 1.09 | 0.72 | 0.47 |
| | 2.0 | 98.5 | 95.8 | 95.3 | 94.2 | 92.5 | 2.44 | 1.68 | 1.10 | 0.72 | 0.48 |
| T=150, N=50 | 0.5 | 97.3 | 94.9 | 94.8 | 94.1 | 93.0 | 1.17 | 0.81 | 0.54 | 0.36 | 0.23 |
| | 1.0 | 97.6 | 95.7 | 95.5 | 94.9 | 94.0 | 1.15 | 0.82 | 0.55 | 0.36 | 0.24 |
| | 2.0 | 97.5 | 95.2 | 95.4 | 94.8 | 93.7 | 1.15 | 0.83 | 0.58 | 0.39 | 0.26 |
| T=150, N=150 | 0.5 | 97.9 | 94.6 | 95.0 | 94.2 | 93.2 | 1.18 | 0.81 | 0.55 | 0.36 | 0.24 |
| | 1.0 | 97.9 | 95.1 | 94.9 | 94.7 | 93.7 | 1.17 | 0.82 | 0.56 | 0.37 | 0.24 |
| | 2.0 | 97.4 | 94.3 | 94.5 | 94.4 | 93.3 | 1.14 | 0.82 | 0.56 | 0.38 | 0.25 |

Responses to a shock of the first unobserved factor

| | | Coverage rate | | | | | Median length | | | | |
|--------------|-----|---------------|------|------|------|------|---------------|------|------|------|------|
| | d | h=1 | 2 | 3 | 4 | 5 | h=1 | 2 | 3 | 4 | 5 |
| T=50, N=50 | 0.5 | 90.3 | 94.0 | 93.4 | 90.3 | 87.3 | 1.51 | 1.12 | 0.77 | 0.51 | 0.33 |
| | 1.0 | 89.7 | 94.5 | 93.5 | 90.6 | 87.0 | 1.45 | 1.08 | 0.75 | 0.51 | 0.33 |
| | 2.0 | 89.1 | 93.5 | 92.0 | 88.9 | 86.3 | 1.35 | 1.04 | 0.72 | 0.48 | 0.32 |
| T=50, N=150 | 0.5 | 90.7 | 94.9 | 94.4 | 91.9 | 89.1 | 1.52 | 1.12 | 0.78 | 0.52 | 0.35 |
| | 1.0 | 89.2 | 93.4 | 92.4 | 90.2 | 87.6 | 1.43 | 1.07 | 0.74 | 0.51 | 0.33 |
| | 2.0 | 89.0 | 93.4 | 91.9 | 90.3 | 87.5 | 1.33 | 1.03 | 0.71 | 0.48 | 0.32 |
| T=150, N=50 | 0.5 | 91.7 | 94.5 | 94.6 | 92.9 | 89.6 | 0.77 | 0.55 | 0.39 | 0.26 | 0.17 |
| | 1.0 | 91.5 | 93.9 | 93.8 | 91.5 | 89.4 | 0.76 | 0.56 | 0.39 | 0.27 | 0.18 |
| | 2.0 | 90.4 | 94.1 | 94.1 | 91.3 | 87.7 | 0.69 | 0.53 | 0.37 | 0.25 | 0.17 |
| T=150, N=150 | 0.5 | 91.9 | 94.2 | 94.1 | 91.2 | 89.2 | 0.79 | 0.56 | 0.39 | 0.27 | 0.17 |
| | 1.0 | 92.5 | 95.1 | 94.6 | 91.7 | 89.0 | 0.76 | 0.55 | 0.40 | 0.27 | 0.18 |
| | 2.0 | 90.8 | 94.1 | 94.5 | 92.1 | 89.3 | 0.69 | 0.52 | 0.37 | 0.25 | 0.17 |

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

Responses to a unit shock of the observed factor

| | d | Coverage rate | | | | | Median length | | | | |
|--------------|-----|---------------|------|------|------|------|---------------|------|------|------|------|
| | | h=1 | 2 | 3 | 4 | 5 | h=1 | 2 | 3 | 4 | 5 |
| T=50, N=50 | 0.5 | 98.3 | 98.7 | 96.6 | 96.1 | 94.6 | 2.80 | 1.96 | 1.12 | 0.72 | 0.47 |
| | 1.0 | 97.9 | 98.1 | 96.0 | 95.3 | 93.9 | 2.59 | 1.89 | 1.09 | 0.69 | 0.45 |
| | 2.0 | 97.6 | 98.4 | 96.1 | 95.6 | 93.7 | 2.38 | 1.78 | 1.07 | 0.67 | 0.42 |
| T=50, N=150 | 0.5 | 97.9 | 98.8 | 96.5 | 96.4 | 95.3 | 2.77 | 1.93 | 1.14 | 0.73 | 0.47 |
| | 1.0 | 97.9 | 98.5 | 97.1 | 96.2 | 94.4 | 2.59 | 1.88 | 1.13 | 0.71 | 0.46 |
| | 2.0 | 97.7 | 98.3 | 96.8 | 96.2 | 94.7 | 2.36 | 1.81 | 1.07 | 0.67 | 0.44 |
| T=150, N=50 | 0.5 | 96.1 | 96.2 | 94.3 | 92.7 | 90.3 | 1.10 | 0.83 | 0.53 | 0.33 | 0.21 |
| | 1.0 | 96.5 | 96.7 | 94.5 | 93.6 | 91.4 | 1.06 | 0.83 | 0.52 | 0.33 | 0.21 |
| | 2.0 | 96.2 | 95.8 | 94.1 | 93.0 | 91.0 | 1.03 | 0.82 | 0.54 | 0.34 | 0.22 |
| T=150, N=150 | 0.5 | 96.7 | 96.1 | 94.4 | 93.4 | 91.7 | 1.09 | 0.83 | 0.53 | 0.34 | 0.22 |
| | 1.0 | 96.3 | 96.1 | 94.1 | 93.0 | 91.6 | 1.07 | 0.83 | 0.54 | 0.34 | 0.22 |
| | 2.0 | 95.6 | 95.4 | 92.8 | 92.1 | 90.4 | 1.02 | 0.81 | 0.52 | 0.33 | 0.21 |

Responses to a unit shock of the first unobserved factor

| | d | Coverage rate | | | | | Median length | | | | |
|--------------|-----|---------------|------|------|------|------|---------------|------|------|------|------|
| | | h=1 | 2 | 3 | 4 | 5 | h=1 | 2 | 3 | 4 | 5 |
| T=50, N=50 | 0.5 | 98.6 | 98.3 | 95.2 | 93.0 | 89.2 | 2.91 | 1.73 | 1.01 | 0.63 | 0.40 |
| | 1.0 | 98.5 | 98.3 | 94.7 | 91.6 | 89.3 | 2.98 | 1.75 | 1.02 | 0.64 | 0.40 |
| | 2.0 | 98.3 | 97.5 | 93.6 | 90.6 | 87.5 | 2.82 | 1.68 | 0.97 | 0.60 | 0.39 |
| T=50, N=150 | 0.5 | 98.3 | 98.3 | 95.2 | 93.3 | 91.0 | 3.00 | 1.73 | 1.03 | 0.66 | 0.43 |
| | 1.0 | 98.5 | 97.7 | 94.3 | 91.8 | 88.8 | 2.93 | 1.72 | 0.99 | 0.62 | 0.41 |
| | 2.0 | 98.6 | 97.8 | 94.5 | 92.1 | 89.7 | 2.91 | 1.69 | 0.99 | 0.62 | 0.41 |
| T=150, N=50 | 0.5 | 98.0 | 97.7 | 95.0 | 92.6 | 89.9 | 0.99 | 0.64 | 0.43 | 0.29 | 0.19 |
| | 1.0 | 98.9 | 97.2 | 95.1 | 92.2 | 90.0 | 1.03 | 0.66 | 0.44 | 0.29 | 0.20 |
| | 2.0 | 98.3 | 97.4 | 95.3 | 92.9 | 89.7 | 0.99 | 0.64 | 0.43 | 0.29 | 0.19 |
| T=150, N=150 | 0.5 | 98.1 | 97.3 | 94.4 | 91.9 | 89.7 | 1.02 | 0.65 | 0.43 | 0.30 | 0.20 |
| | 1.0 | 98.8 | 97.6 | 95.2 | 92.8 | 90.3 | 1.02 | 0.66 | 0.43 | 0.30 | 0.20 |
| | 2.0 | 98.5 | 97.9 | 96.0 | 93.4 | 90.5 | 0.98 | 0.64 | 0.42 | 0.28 | 0.19 |

Figure 1. Point estimates and 90% bootstrap confidence intervals using the same data employed by Bernanke et al. (2005)

