Shintani, Mototsugu; Guo, Zi-Yi

Preprint

Finite Sample Performance of Principal Components Estimators for Dynamic Factor Models: Asymptotic vs. Bootstrap Approximations

Finite Sample Performance of Principal Components Estimators for Dynamic Factor Models: Asymptotic and Bootstrap Approximations*

Mototsugu Shintani† and Zi-Yi Guo‡

This version: January 2011

Abstract

This paper investigates the finite sample properties of the two-step estimators of dynamic factor models when unobservable common factors are estimated by the principal components methods in the first step. Effects of the number of individual series on the estimation of an autoregressive model of a common factor are investigated both by theoretical analysis and by a Monte Carlo simulation. When the number of the series is not sufficiently large relative to the number of time series observations, the autoregressive coefficient estimator of positively autocorrelated factor is biased downward and the bias is larger for a more persistent factor. In such a case, bootstrap procedures are effective in reducing the bias and bootstrap confidence intervals outperform naive asymptotic confidence intervals in terms of controlling the coverage probability.

Keywords: Bias Correction; Bootstrap; Dynamic Factor Model; Principal Components

JEL classification: C15; C53

*The earlier version of the paper was circulated under the title "Finite sample performance of principal components estimators for dynamic factor models." The authors would like to thank the Editor, two anonymous referees, Todd Clark, Mario Crucini, Silvia Gonçalves, James MacKinnon, Serena Ng, Benoit Perron and seminar and conference participants at Indiana University, University of Montreal and 2012 Meetings of the Midwest Econometrics Group for helpful comments and discussions.

†Department of Economics, Vanderbilt University; e-mail: mototsugu.shintani@vanderbilt.edu.

‡Department of Economics, Vanderbilt University; e-mail: ziyi.guo@vanderbilt.edu.

1 Introduction

The estimation of dynamic factor models has become popular in macroeconomic analysis after influential works by Sargent and Sims (1977), Geweke (1977) and Stock and Watson (1989). Later studies by Stock and Watson (1998, 2002), Bai and Ng (2002) and Bai (2003) emphasize the consistency of the principal components estimator of unobservable common factors under the asymptotic framework with a large number of cross-sectional observations. This paper investigates the finite sample properties of two-step persistence estimators in dynamic factor models when unobservable common factors are estimated by the principal components method in the first step. The first-step estimation is followed by the estimation of autoregressive models of common factors in the second step. Using analytical results and simulation experiments, we evaluate the effect of the number of the series (N) relative to the time series observations (T) on the performance of the two-step estimator of a persistence parameter. Furthermore, we propose a simple bootstrap procedure that works well in the case of relatively small N.

In this paper, we focus on the persistence parameter of the common factor because of its empirical relevance in macroeconomic analysis. In modern macroeconomics literature, dynamic stochastic general equilibrium (DSGE) models predict that a small set of driving forces is responsible for covariation in macroeconomic variables. Theoretically, the persistence of the common factor often plays a key role on implications of these models. For example, in the real business cycle model, there is a well-known trade-offs between the persistence of technology shock and the performance of the model. When the shock becomes more persistent, the performance improves along some dimensions but deteriorates along other dimensions (King et al., 1988, Hansen, 1997, Ireland, 2001). In DSGE models with a monetary sector, the optimal monetary policy highly depends on the persistence of real shocks in the economy (Woodford, 1999). In open economy models, the welfare gain from the introduction of international risk sharing becomes larger when the technology shock becomes more persistent (Baxter and Crucini, 1995). Since these common shocks are not directly observable, a dynamic factor model offers a simple robust statistical framework to measure the persistence of

such common components that are responsible in explaining macroeconomic fluctuations.[1]

Dynamic factor models have also been used to construct a business cycle index (e.g., Stock and Watson, 1989, Kim and Nelson, 1993) and to extract a measure of underlying, or core, inflation (e.g., Bryan and Cecchetti, 1993). In such applications, persistence of a single factor can often be of main interest. For example, Clark (2006) examines the possibility of a structural shift in the persistence of a single common factor estimated using the first principal component of disaggregate inflation series. In this paper, we only consider the case where a single common factor is generated from a univariate autoregressive (AR) model of order one. This keeps our problem simple since the persistence measure corresponds to the AR coefficient. However, in principle, the main idea of our approach can be applicable to AR models of higher order.[2]

The principal components estimation of the unobserved common factors is computationally simple and feasible with a large number of cross-sectional observations N. The method also allows for an approximate factor structure with possible cross-sectional correlations of idiosyncratic errors.[3] The large N asymptotic results obtained by Connor and Korajczyk (1986) and Bai (2003) imply $\sqrt{N}$-consistency of the principal components estimators of common factors up to a scaling constant. Therefore, if N is sufficiently large, we can treat the estimated common factor as if we directly observe the true common factor when conducting inference. However, since this argument is based on the asymptotic theory, an approximation may not work with small N relative to the time series observation T typically available in practice. Consistent with our theoretical prediction, results from our Monte Carlo experiment using positively autocorrelated factors suggest the downward bias in the AR coefficient estimator and significant under-coverage of the naive confidence interval when N is small. The simulation results also suggest that a simple bootstrap procedure works well in correcting the bias and improves the coverage rate of the confidence interval.

[1] Recently, Boivin and Giannoni (2006) propose estimating a dynamic factor model in which they impose the full structure of the DSGE model on the transition equation of the latent factors.

[2] In the case of AR(p) models, however, there are several measures of persistence, including the sum of AR coefficients, largest characteristic root, and first-order autocorrelation.

[3] The principle components estimator of the common factor with large N can also be used to estimate nonlinear models (Connor, Korajczyk and Linton, 2006, Diebold, 1998, Shintani, 2005, 2008) or to test the hypothesis of a unit root (Bai and Ng, 2004, and Moon and Perron, 2004).

The bootstrap part of our analysis is closely related to recent studies by Gonçalves and Perron (2012) and Yamamoto (2012). Both papers also employ bootstrap procedures for the purpose of improving the finite sample performance of estimators of dynamic factor models. Gonçalves and Perron (2012) employ a bootstrap procedure in factor-augmented forecasting regression models with multiple factors. The factor-augmented forecasting regression models offer a very useful framework in forecasting macroeconomic time series using information extracted from many variables. This aspect is emphasized in Stock and Watson (1998, 2002), Marcellino, Stock and Watson (2003) and Bai and Ng (2006) among others. Gonçalves and Perron (2012) provide the first order asymptotic validity of their bootstrap procedure for factor-augmented forecasting regression models, but not in the context of estimation of persistence parameter of the common factor. It should also be noted that, unlike their factor-augmented forecasting regression models with multiple factors, bootstrapping common factor in our univariate AR model is not subject to the scaling and rotation issues.[4] Yamamoto (2012) examines the performance of the bootstrap procedure applied to factor-augmented vector autoregressive (FAVAR) models of Bernanke, Boivin and Eliasz (2005). While his multiple factor structure is more general than our single factor structure, his main focus is the identification of structural parameters in the FAVAR analysis using various identifying assumptions. In contrast, we are more interested in the role of parameters in the model in explaining the deviation from the large N asymptotics when N is small.

There are several simulation results available in the literature on the principal components estimator of dynamic factor models. Stock and Watson (1998) report the finite sample simulation results on the magnitude of the first-step estimation error of the common factor as well as the performance of an out-of-sample forecast based on the estimated factor relative to that of an infeasible forecast with a true factor. Boivin and Ng (2006) report similar performance measures in investigating the marginal effect of increasing N when there is a strong cross-sectional correlation of the errors. In addition, Stock and Watson (1998) and Bai and Ng (2002) find that information criteria designed to determine the number of the factors

[4] To be more specific, under our normalizing assumption, the factor is estimated up to sign but autoregressive coefficient can be identifed as the sign cancels out from both side of the regression.

perform well in a finite sample. However, none of these studies directly investigate the effect of N on the estimation of dynamic structure of the common factors.

The remainder of the paper is organized as follows: Section 2 reviews the asymptotic theory of the two-step estimator and investigate the finite sample performance of the estimator in simulation. Section 3 considers a bootstrap approach to reduce the bias. Section 4 considers a bootstrap approach to improve the coverage performance of the confidence interval. Section 5 provides an empirical illustration of our procedures. Some concluding remarks are made in Section 6. All the proofs of theoretical results are provided in the Appendix.

2 Two-Step Estimation of Autoregressive Model of Latent Factor

We begin our discussion by reviewing the literature of finite sample bias correction of infeasible estimator of an AR(1) model and then provide asymptotic properties of a two-step estimator of dynamic factor structure. Let x_{it} be an i-th component of N-dimensional multiple time series $X_t = (x_{1t}, \ldots, x_{Nt})'$ and $t = 1, ..., T$. A natural way to explain the comovement of x_{it}'s caused by a single factor, such as productivity shocks, is to use a simple one-factor model

$$x_{it} = \lambda_i f_t + e_{it} \tag{1}$$

for $i = 1, ..., N$, where λ_i's are factor loadings with respect to i-th series, f_t is a scalar common factor, and e_{it}'s are possibly cross-sectionally correlated idiosyncratic shocks. If a dynamic structure is introduced by incorporating (i) a dynamic data generating process for f_t, (ii) lags of f_t in (1) or (iii) serial correlation in e_{it}'s, then the model becomes a dynamic factor model. In this paper, we limit our attention to a simple case with a single factor generated from a zero-mean linear stationary AR(1) model,

$$f_t = \rho f_{t-1} + \varepsilon_t \tag{2}$$

where $|\rho| < 1$, and ε_t is i.i.d. with $E(\varepsilon_t) = 0, E({\varepsilon_t}^2) = \sigma^2$ and a finite fourth moment.

When f_t is directly observable, the AR parameter ρ can be estimated by ordinary least squares (OLS),

$$\widehat{\rho} = \left(\sum_{t=1}^{T} f_{t-1}^2 \right)^{-1} \sum_{t=1}^{T} f_{t-1} f_t. \tag{3}$$

Under the assumption described above, the limiting distribution of the OLS estimator (3) is given by

$$\sqrt{T}\,(\widehat{\rho} - \rho) \xrightarrow{d} N(0, 1 - \rho^2), \tag{4}$$

as T tends to infinity, which justifies the use of the asymptotic confidence intervals for ρ. For example, the 90% confidence interval is typically constructed as

$$[\widehat{\rho} - 1.645 \times SE(\widehat{\rho}), \widehat{\rho} + 1.645 \times SE(\widehat{\rho})] \tag{5}$$

where $SE(\widehat{\rho})$ is the standard error of $\widehat{\rho}$ defined as $SE(\widehat{\rho}) = (\widehat{\sigma}^2 / \sum_{t=1}^{T} f_{t-1}^2)^{1/2}$, $\widehat{\sigma}^2 = (T-1)^{-1} \sum_{t=1}^{T} \widehat{\varepsilon}_t^2$ and $\widehat{\varepsilon}_t = f_t - \widehat{\rho} f_{t-1}$.

When T is small, the presence of bias of the OLS estimator (3) is well-known and several procedures have been proposed to reduce the bias in the literature. Using the approximation formula of the bias obtained in early studies by Hurwicz (1950), Marriott and Pope (1954) and Kendall (1954), one can construct a simple bias-corrected estimator. For example, in the current setting with a zero-mean restriction, the bias-corrected estimator is given by $\widehat{\rho}_{KBC} = T(T-2)^{-1}\widehat{\rho}$ which is a solution to the bias approximation formula $E(\widehat{\rho}) - \rho = -2T^{-1}\rho + O(T^{-2})$ for ρ with $E(\widehat{\rho})$ replaced by $\widehat{\rho}$. Alternatively, one can use the bootstrap method for the bias correction. A similar methodology was first employed by Quenouille (1949) who proposed a subsampling procedure to correct the bias. The bootstrapping AR models from resampling residuals was later formalized by Bose (1988) and extended to the multivariate case by Kilian (1998). In particular, the bias-corrected estimator is given by $\widehat{\rho}_{BC} = \widehat{\rho} - \widehat{bias}$ where the bootstrap bias estimate is $\widehat{bias} = B^{-1} \sum_{b=1}^{B} \widehat{\rho}_b^* - \widehat{\rho}$ and $\widehat{\rho}_b^*$ is the b-th AR estimate from the bootstrap sample and B is the number of bootstrap replications. By using either the Kendall-type bias correction or bootstrap bias correction procedures, the small T bias is reduced by the order of T^{-1}.

Table 1 reports the mean values of the OLS estimator $\widehat{\rho}$ along with the effective coverage rates of the nominal 90% conventional asymptotic confidence intervals (5) in 1,000 replications using f_t generated from (2) with the AR parameter, $\rho = 0.5$ and 0.9 combined with $\varepsilon_t \sim iidN(0, 1-\rho^2)$.[5] The sample sizes are $T = 100$ and 200. The initial value f_t is drawn from the unconditional distribution of f_t, that is $N(0,1)$. In addition to the OLS estimator $\widehat{\rho}$, the mean values of the Kendall-type bias-corrected estimator $\widehat{\rho}_{KBC}$ and the bootstrap bias-corrected estimator $\widehat{\rho}_{BC}$ are also reported. For the bootstrap bias correction, we use $B = 499$. The results suggest that the coverage of conventional asymptotic confidence intervals seems very accurate for sample sizes $T = 100$ and 200. In addition, comparisons between two bias correction methods suggest that the small T bias of the OLS estimator ($\widehat{\rho}$) can be corrected reasonably well either by Kendall-type correction ($\widehat{\rho}_{KBC}$) or the bootstrap-type correction ($\widehat{\rho}_{BC}$). In what follows, we use the results in Table 1 as a benchmark to evaluate the performance of the two-step estimator when the factor f_t is not known.

Let us now review the asymptotic property of the two-step estimator for the persistence parameter ρ when only x_{it} from (1) is observable. Under very general conditions, f_t can still be consistently estimated (up to scale) by using the first principal component of the $N \times N$ matrix $X'X$ where X is the $T \times N$ data matrix with t-th row X_t', or by using the first eigenvector of the $T \times T$ matrix XX'.[6] We denote this common factor estimator by $\widetilde{f}_t$ with a normalization $T^{-1}\sum_{t=1}^{T} \widetilde{f}_t^2 = 1$. Once $\widetilde{f}_t$ is obtained, we can replace f_t in (3) by $\widetilde{f}_t$ and the feasible estimator of ρ is

$$\widetilde{\rho} = \left(\sum_{t=1}^{T} \widetilde{f}_{t-1}^2 \right)^{-1} \sum_{t=1}^{T} \widetilde{f}_{t-1} \widetilde{f}_t. \tag{6}$$

Below, we first show the asymptotic validity of this two-step estimator, followed by the examination of its finite sample performance using a simulation. To this end, we employ the following assumptions on the moment conditions for factors, factor loadings and idiosyncratic errors. Below, we let M be some finite positive constant.

[5] Since our results are based on 1,000 replications, the standard error of 90% coverage rate in the simulation is about 0.01 ($\approx \sqrt{0.9 \times 0.1/1000}$).

[6] Since principal components are not scale-invariant, it is common practice to standardized all x_{it}'s to have zero sample mean and unit sample variance before applying the principal components method.

Assumption F (factors): (i) $E(f_t) = 0$, $E({f_t}^2) = \sigma_f^2 = 1$, $E({f_t}^4) < M$, and (ii) $F'F/T - \sigma_f^2 = o_P(1)$ where $F = [f_1, \cdots, f_T]'$ as $T \to \infty$.

Assumption FL (factor loadings): (i) $E(\lambda_i) = 0$, $E({\lambda_i}^2) = \sigma_\lambda^2$, $E({\lambda_i}^4) < M$, and (ii) $\Lambda'\Lambda/N - \sigma_\lambda^2 = o_P(1)$ where $\Lambda = [\lambda_1, \cdots, \lambda_N]'$ as $N \to \infty$.

Assumption E (errors): (i) $E(e_{it}) = 0$, $E({e_{it}}^2) = \sigma_{ei}^2 \leq M$, $E|{e_{it}}^8| \leq M$, (ii) $E(e_{is}e_{it}) = 0$ for all $t \neq s$, (iii) $E(e_{it}e_{jt}) = \tau_{ij} \leq M$ for all t, i and j, (iv) $E[N^{-1/2}\sum_{i=1}^{N}(e_{it}e_{is} - E(e_{it}e_{is}))]^2 \leq M$ and (v) $\Xi \equiv \lim_{N,T\to\infty} T^{-1}\sum_{t=1}^{T}\Xi_t > 0$, where $\Xi_t \equiv Var(N^{-1/2}\sum_{i=1}^{N}\lambda_i e_{it})$.

Assumption E allows cross-sectional correlation and heteroskedasticity but not serial correlation of idiosyncratic error terms. It should be noted that Assumption E can be replaced by a weaker assumption that allows serial correlations of idiosyncratic errors (see Bai, 2003, and Bai and Ng, 2002). In addition, we employ the following assumption on the relation among three random variables.

Assumption I (independence): The variables $\{f_t\}$, $\{\lambda_i\}$ and $\{e_{it}\}$ are three mutually independent groups. Dependence within each group is allowed.

The following proposition provides the asymptotic properties of the two-step estimator of the autoregressive coefficient.

Proposition 1. *Let x_{it} and f_t be generated from (1) and (2), respectively, and suppose that assumptions F, FL, E and I are satisfied. Then, as $T \to \infty$* and *$N \to \infty$ such that $\sqrt{T}/N \to c$ where $0 \leq c < \infty$,*

$$E(\tilde{\rho} - \rho) = -T^{-1/2}c\rho\sigma_\lambda^{-4}\Xi + o(T^{-1/2}) \tag{7}$$

and

$$\sqrt{T}(\widetilde{\rho} - \rho) \xrightarrow{d} N(-c\rho\sigma_\lambda^{-4}\Xi, 1 - \rho^2). \tag{8}$$

The proposition is derived using the asymptotic framework employed by Bai (2003) and Gonçalves and Perron (2012) in their analysis of factor-augmented forecasting regression model. In particular, it relies on the simultaneous limit theory where both N and T are allowed to grow simultaneously with a rate of N being at least as fast as $\sqrt{T}$. The bias term of order $T^{-1/2}$ in (7) is analogues to the bias term in the factor-augmented forecasting regression discussed by Ludvigson and Ng (2010) and Gonçalves and Perron (2012). Bai (2003) emphasizes that the factor estimation error has no effect on the estimation of the factor-augmented forecasting regression model if $\sqrt{T}/N$ is sufficiently small in the limit ($c = 0$). Similarly, in the context of estimation of the autoregressive model of the common factor, the factor estimation error can be negligible for small $\sqrt{T}/N$. A special case of Proposition 1 with $c = 0$ implies

$$\sqrt{T}\,(\widetilde{\rho} - \rho) \xrightarrow{d} N(0, 1 - \rho^2) \tag{9}$$

as T tends to infinity, so that the limiting distribution of $\widetilde{\rho}$ in Theorem 1 is same as that of $\widehat{\rho}$ given by (4). In fact, we can also show the asymptotic equivalence of $\widetilde{\rho}$ and $\widehat{\rho}$ with their difference given by $\widetilde{\rho} - \widehat{\rho} = o_P(T^{-1/2})$.[7] Therefore, when the number of the series (N) is sufficiently large relative to the time series observations (T), the estimated factor $\widetilde{f}_t$ can be treated in exactly the same way as in the case of observable f_t. Combined with the consistency of the standard error, asymptotic confidence intervals analogues to (4) can be used for the two-step estimator $\widetilde{\rho}$. For example, the 90% confidence interval can be constructed as

$$[\widetilde{\rho} - 1.645 \times SE(\widetilde{\rho}), \widetilde{\rho} + 1.645 \times SE(\widetilde{\rho})] \tag{10}$$

where $SE(\widetilde{\rho})$ is the standard error of $\widetilde{\rho}$ defined as $SE(\widetilde{\rho}) = (\widetilde{\sigma}^2 / \sum_{t=1}^{T} \widetilde{f}_t^2)^{1/2}$, $\widetilde{\sigma}^2 = (T - 1)^{-1} \sum_{t=1}^{T} \widetilde{\varepsilon}_t^2$ and $\widetilde{\varepsilon}_t = \widetilde{f}_t - \widetilde{\rho}\widetilde{f}_{t-1}$.

When N is small (relative to T), however, the distribution of $\widetilde{\rho}$ may better be approximated by (8) in Proposition 1, rather than by (9). In such a case, the presence of bias term in (8) can result in bad coverage performance of a naive asymptotic confidence interval (10). Since $-T^{-1/2} c\rho\sigma_\lambda^{-4}\Xi$ in (7) can also be rewritten as $-N^{-1}\rho\sigma_\lambda^{-4}\Xi$, in what follows, we refer this bias

[7] See the proof of Proposition 1.

as the small N bias as oppose to the small T bias, $-2T^{-1}\rho$, discussed above. Within our asymptotic framework, the small N bias dominates the small T bias since the former is of order $T^{-1/2}$ and the latter is of order T^{-1}. However, it is interesting to note some similarity between the small N bias and the small T bias. For positive values of ρ, both types of bias are downward and increasing in ρ. However, the small N bias also depends on the dispersion of the factor loadings (σ_λ^2) and covariance structure of the factor loadings and idiosyncratic errors (Ξ).

To examine the finite sample performance of the two-step estimator $\widetilde{\rho}$ in a simulation, we now generate x_{it} from (1) with the factor loading $\lambda_i \sim N(0,1)$, serially and cross-sectionally uncorrelated idiosyncratic error $e_{it} \sim N(0,\sigma_e^2)$, and the factor f_t from the same DGP as before. Relative size of common component and idiosyncratic error in x_{it} is expressed in term of the signal-to-noise ratio defined by $Var(\lambda_i f_t)/Var(e_{it}) = 1/\sigma_e^2$ which is controlled by changing σ_e^2. The set of values of the signal-to-noise ratio we consider is $\{0.5, 0.75, 1.0, 1.5, 2.0\}$. We also follow Bai and Ng (2006) and Gonçalves and Perron (2012) and consider the performance in the presence of cross-sectionally correlated errors where the correlation between e_{it} and e_{jt} is given by $0.5^{|i-j|}$ if $|i-j| \leq 5$. For a given value of T, the relative sample size N is set according to $N = [\sqrt{T}/c]$ for $c = \{0.5, 1.0, 1.5\}$ where $[x]$ is integer part of x. Therefore, sets of Ns under consideration are $\{7, 10, 20\}$ for $T = 100$ and $\{9, 14, 28\}$ for $T = 200$.

Table 2 reports the mean values of the two-step estimator $\widetilde{\rho}$, along with the effective coverage rates of the nominal 90% asymptotic confidence intervals (10). While the theory predicts that the coverage probability should be 0.90 only if N is sufficiently large relative to T, we are interested in examining its finite performance when N is small. The upper panel of the table shows the results with cross-sectionally uncorrelated errors while the lower panel shows those with cross-sectionally correlated errors.

Overall, the point estimates of two-step estimator $\widetilde{\rho}$ are clearly biased downward for small N. Compared to the results for the infeasible estimator $\widehat{\rho}$ in Table 1, the magnitude of bias is much larger with $\widetilde{\rho}$ reflecting the fact that the theoretical order of small N bias dominates that of the small T bias. In addition, consistent with the theoretical prediction in Proposition 1, the magnitude with bias increases when (i) ρ increases, (ii) c increases (or N decreases) and

(iii) the signal-to-noise ratio decreases (or Ξ increases). For the same parameter values for ρ, c and signal-to-noise ratio, introduction of the cross-sectional correlation seems to increase the bias of $\widetilde{\rho}$. This effect does not show up in the first order asymptotics in Proposition 1 since it does not change the value of Ξ. However, when the signal-to-noise ratio is the highest, the difference in point estimates between cross-sectionally uncorrelated and cross-sectionally correlated cases is the smallest.

The coverage performance of the standard asymptotic confidence intervals also becomes worse compared to the results in Table 1. For all the cases, the actual coverage frequency is much less than the nominal coverage rate of 90%. The most closest coverage to the nominal rate is obtained when $\rho = 0.5$ combined with a small c (a large N) and a large signal-to-noise ratio. In this case, there are about 2 to 4% under-coverage. The deviation from the nominal rate becomes larger when ρ increases, c increases, the signal-to-noise ratio decreases and the cross-sectional correlation is introduced. The fact that the degree of under-coverage is in parallel relationship to the magnitude of the small N bias can also be explained by Proposition 1. When $-c\rho\sigma_{\lambda}^{-4}\Xi$ in (8) is not negligible, confidence interval (10), which is based on approximation (9), cannot be expected to perform well. In summary, the asymptotic confidence interval (10) may work well in terms of the coverage rate when N is as large as a half of T and when the AR parameter is not close to unity. Otherwise, the presence of the small N bias results in a poor coverage of the naive confidence interval. The effect of this downward bias becomes more severe as the AR parameter approaches to unity. In the next section, we consider the possibility of improving the performance of the two-step estimator for small N by approximating the true distribution by bootstrap procedures.

3 Bootstrap Approach to Bias Correction

In the previous section, we find that the presence of small sample bias associated with finite N is likely the main source of poor coverage of the asymptotic confidence interval. Recall that in the case of correcting the small T bias, an analytical bias formula is utilized to obtain $\widehat{\rho}_{KBC}$ while the bootstrap estimate of bias is used to construct $\widehat{\rho}_{BC}$. Similarly, we

can either utilize the explicit bias function and correct bias analytically using the formula in Proposition 1, or estimate bias using the bootstrap method for the purpose of correction. Here we take the latter approach and employ the bootstrap procedure designed to work with cross-sectionally uncorrelated errors ($\tau_{ij} = 0$ for all $i \neq j$). However, in simulation, we also investigate its performance in the presence of cross-sectionally correlated errors. We first describe a simple bootstrap procedure for the bias correction.

Bootstrap Bias Correction I

1. Estimate factors and factor loadings using principal components method and obtain residuals $\widetilde{e}_{it} = x_{it} - \widetilde{\lambda}_i \widetilde{f}_t$.

2. Recenter $\widetilde{e}_{it}$, $\widetilde{\lambda}_i$ and $\widetilde{f}_t$ around zero. Generate $x^*_{1t} = \lambda^*_1 \widetilde{f}_t + e^*_{1t}$ for $t = 1, ..., T$ by first drawing λ^*_1 from $\widetilde{\lambda}_i$ and then drawing e^*_{1t} for $t = 1, ..., T$ from $\widetilde{e}_{jt}$ given $\lambda^*_1 = \widetilde{\lambda}_j$. Repeat the same procedure N times to generate all x^*_{it}'s for $i = 1, ..., N$.

3. Apply the principal components method to x^*_{it} and estimate $\widetilde{f}^*_t$.

4. Compute bootstrap AR coefficient estimate $\widetilde{\rho}^*$ from $\widetilde{f}^*_t$.

5. Repeat steps 2 to 4 B times to obtain the bootstrap bias estimator $bias^* = B^{-1} \sum_{b=1}^{B} \widetilde{\rho}^*_b - \widetilde{\rho}$ where $\widetilde{\rho}^*_b$ is the b-th bootstrap AR estimate and $\widetilde{\rho}$ is the AR estimate from $\widetilde{f}_t$. The bias-corrected estimator of ρ is given by $\widetilde{\rho}_{BC} = \widetilde{\rho} - bias^*$.

Beran and Srivastava (1985) have established the validity of applying the bootstrap procedure to the principal components analysis. Our procedure slightly differs from theirs in that we resample x^*_{it} using estimated factor model in step 2. It should be noted that the procedure above is designed to evaluate the bias from small N in the principal components method rather than the bias from small T in the autoregression. In order to incorporate the small T bias and the small N bias simultaneously, we may combine the procedure with bootstrapping autoregressive models. This possibility is considered in the second bootstrap bias correction method described below.

Bootstrap Bias Correction II

1. Estimate factors and factor loadings using principal components method and obtain residuals $\widetilde{e}_{it} = x_{it} - \widetilde{\lambda}_i \widetilde{f}_t$.

2. Compute the AR coefficient estimate $\widetilde{\rho}$ from $\widetilde{f}_t$ and obtain residuals $\widetilde{\varepsilon}_t = \widetilde{f}_t - \widetilde{\rho}\widetilde{f}_{t-1}$.

3. Recenter $\widetilde{\varepsilon}_t$ around zero if necessary and generate ε_t^* by resampling from $\widetilde{\varepsilon}_t$. Then generate pseudo factors using $f_t^* = \widetilde{\rho} f_{t-1}^* + \varepsilon_t^*$.

4. Recenter $\widetilde{e}_{it}$, $\widetilde{\lambda}_i$ and $\widetilde{f}_t$ around zero. Generate $x_{1t}^* = \lambda_1^* f_t^* + e_{1t}^*$ for $t = 1, ..., T$ by first drawing λ_1^* from $\widetilde{\lambda}_i$ and then drawing e_{1t}^* for $t = 1, ..., T$ from $\widetilde{e}_{jt}$ given $\lambda_1^* = \widetilde{\lambda}_j$. Repeat the same procedure N times to generate all x_{it}^*'s for $i = 1, ..., N$.

5. Apply principal components method to x_{it}^* and estimate $\widetilde{f}_t^*$.

6. Compute bootstrap AR coefficient estimate $\widetilde{\rho}^*$ from $\widetilde{f}_t^*$.

7. Repeat steps 2 to 6 B times to obtain the bootstrap bias estimator $bias^* = B^{-1} \sum_{b=1}^{B} \widetilde{\rho}_b^* - \widetilde{\rho}$ where $\widetilde{\rho}_b^*$ is the b-th bootstrap AR estimate and $\widetilde{\rho}$ is the AR estimate from $\widetilde{f}_t$. The bias-corrected estimator of ρ is given by $\widetilde{\rho}_{BC} = \widetilde{\rho} - bias^*$.

The second procedure for the bias correction involves a combination of bootstrapping principal components and bootstrapping the residuals in autoregressive models (Freedman, 1984, and Bose, 1988). Note that our procedures employ the bootstrap bias correction based on a constant bias function. While this form of bias correction seems to be the one most frequently used in practice (e.g., Kilian, 1998), the performance of the bias-corrected estimator may be improved by introducing linear or nonlinear bias functions in the correction (see MacKinnon and Smith, 1998).

The asymptotic justification of using our bootstrap methods to correct the small N bias is established in the following proposition.

Proposition 2. *Let all the assumptions of Proposition 1 and of cross-sectionally uncorrelated errors are satisfied and the bootstrap data* $\{X^*\}$ *are generated as described in Bootstrap Bias*

Correction I or in Bootstrap Bias Correction II. Then, as $T \to \infty$ and $N \to \infty$ *such that* $\sqrt{T}/N \to c$ *where* $0 \leq c < \infty$, $bias^* = E^*(\widetilde{\rho}^* - \widetilde{\rho}) = E(\tilde{\rho} - \rho) + o_P(T^{-1/2})$.

Since the first order bias term is of order $T^{-1/2}$ in (7) unless $c = 0$, Proposition 2 implies the consistency of the bootstrap bias estimator. This also suggest that the bias-corrected estimator by $\widetilde{\rho}_{BC} = \widetilde{\rho} - bias^*$ has bias of order smaller than $T^{-1/2}$. Since the consistency holds for both *Bootstrap Bias Correction I* and *Bootstrap Bias Correction II,* whether or not bootstrapping autoregressive models does not matter asymptotically.

Let us now conduct the simulation to evaluate the performance the bootstrap bias correction method. The results of simulation under the same specification as in Table 2 are shown in Table 3. For each specification, the true bias $E(\tilde{\rho} - \rho)$ is first evaluated by using the mean value of $\tilde{\rho} - \rho$ in 1,000 replications. The first-order theoretical bias term $-T^{-1/2} c\rho\sigma_\lambda^{-4}\Xi$ in (7) is also reported. The performance of bootstrap bias estimator is evaluated by using the mean value of $bias^*$ in 1,000 replications. Since the results turns out to be very similar between the cases of *Bootstrap Bias Correction I* and *Bootstrap Bias Correction II,* we only report the results from the latter method. The similar performance between the two can be expected since the results in Table 1 suggest that the small T bias is almost negligible for the size of T we consider. Number of the bootstrap replications is set at $B = 499$.

The results of the simulation can be summarized as follows. First, the bootstrap bias estimate matches closely with the true bias for both $\rho = 0.5$ and $\rho = 0.9$ cases unless the signal-to-noise ratio is too small. Second, while the direction of the changes of theoretical first-order bias term is consistent with that of true bias, it only accounts for the fraction of the actual bias. In many case, the bootstrap bias estimates are much closer to the actual bias than the first-order theoretical bias term. Third, the bootstrap bias estimate does not seem to capture the effect of increased bias in the presence of the cross-sectional correlation. However, this is not surprising because our bootstrap procedure is designed for the case of cross-sectionally uncorrelated errors. Overall, the performance of bootstrap correction method seem to be satisfactory.

4 Bootstrap Approach to Confidence Intervals

Since the bootstrap bias correction method is proven to be effective in simulation, we now turn to the issue of improving the performance of confidence intervals using a bootstrap approach. Recall that the deviation of the actual coverage rate of naive asymptotic confidence interval (10) from the nominal rate is proportional to the size of bias in Table 2. Thus, it is natural to expect that a recentered asymptotic confidence interval using the bootstrap bias estimates improves the coverage accuracy. For example, the 90% confidence interval can be constructed as

$$[\widetilde{\rho}_{BC} - 1.645 \times SE(\widetilde{\rho}), \widetilde{\rho}_{BC} + 1.645 \times SE(\widetilde{\rho})]. \tag{11}$$

The asymptotic validity of the confidence interval (11) can be easily shown using the consistency result of bootstrap bias estimator provided in Proposition 2.

Alternatively, we can directly utilize the bootstrap distribution of the estimator to construct bootstrap confidence intervals. Here we consider Efron's equal-tailed percentile bootstrap confidence interval based on the original estimator $\widetilde{\rho}^*$ as well as the equal-tailed percentile-t bootstrap confidence interval based on the t statistic of $\widetilde{\rho}^*$,defined as $t(\widetilde{\rho}^*) = (\widetilde{\rho}^* - \widetilde{\rho}_{BC})/SE(\widetilde{\rho}^*)$ where $SE(\widetilde{\rho}^*)$ is the standard error of $\widetilde{\rho}^*$, which is asymptotically pivotal.[8] We now describe our procedure of constructing the bootstrap confidence intervals.

Bootstrap Confidence Interval

1. Follow either steps 1 to 3 in *Bootstrap Bias Correction I* or steps 1 to 5 in *Bootstrap Bias Correction II*.

2. Compute bootstrap AR coefficient estimate $\widetilde{\rho}^*$ or $t(\widetilde{\rho}^*)$ from $\widetilde{f}_t^*$.

3. Repeat steps 1 to 2 B times to obtain the empirical distribution of $\widetilde{\rho}^* - bias^*$ to construct the percentile confidence interval and of $t(\widetilde{\rho}^*)$ to construct the percentile-t confidence interval.

[8] See Hall (1992) on the importance of using asymptically pivotal statistics in achieving the higher order accuracy of the bootstrap confidence interval.

Note that, as in Kilian's (1998) argument on vector autoregression, $\widetilde{\rho}$ in step 3 in *Bootstrap Bias Correction II* can be replaced by bias corrected estimates $\widetilde{\rho}_{BC}$ without changing the limiting distribution of the bootstrap estimator. The following proposition provides the asymptotic validity of the bootstrap confidence intervals.

Proposition 3. *Let all the assumptions of Proposition 1 and of cross-sectionally uncorrelated errors are satisfied and the bootstrap data* $\{X^*\}$ *are generated as described in Bootstrap Confidence Interval. Then, as* $T \to \infty$ and $N \to \infty$ *such that* $\sqrt{T}/N \to c$ *where* $0 \leq c < \infty$, $\sup_{x \in \Re} |P^*(\sqrt{T}(\tilde{\rho}^* - \tilde{\rho}) \leq x) - P(\sqrt{T}(\tilde{\rho} - \rho) \leq x)| \xrightarrow{P} 0$.

Proposition 3 implies the consistency of our bootstrap procedure in the sense that the limiting distribution of the bootstrap estimator $\tilde{\rho}^*$ is asymptotically equivalent to that of $\widetilde{\rho}$.[9] Since the limiting distribution of $\widetilde{\rho}$ is given by (8) in Proposition 1, the same distribution can be used to describe the limiting behavior of $\tilde{\rho}^*$. Since the coverage rate of the asymptotic confidence interval around the bias-corrected estimate, given by (11), approaches the nominal coverage rate in the limit, the same is true for the percentile bootstrap confidence interval. Similarly, we can modify Proposition 3 and replace $\tilde{\rho}^*$ and $\tilde{\rho}$ by their studentized statistics $t(\widetilde{\rho}^*)$ and $(\widetilde{\rho} - E(\tilde{\rho}))/SE(\widetilde{\rho})$ and show the bootstrap consistency of $t(\widetilde{\rho}^*)$ and the asymptotic validity of the percentile-t confidence interval.[10]

Table 4 reports coverage of three confidence intervals based on the bootstrap applied to the two-step estimator $\widetilde{\rho}$ for $\rho = 0.5$ and $\rho = 0.9$ cases. For the bootstrap bias correction method required in the confidence interval (11), we use *Bootstrap Bias Correction I* mainly because it involves less computation. The table shows that such a simple bias corrected asymptotic interval (11) significantly improves over the conventional asymptotic interval without bias correction (10) in Table 2. Especially, when $T = 200$ and $c = 0.5$, the coverage rates are nearly the nominal rate regardless of the signal-to-noise ratio. For per-

[9] In general, signs of the coefficients in the factor forecasting regression cannot be identified and Gonçalves and Perron (2012) argue the consistency of their bootstrap procedure in renormalized parameter space. In contrast, our result is not subject to the sign identification problem since slope coefficients in univariate AR models can still be identified.

[10] Note that here we are not claiming the higher order refinement of the percentile-t bootstrap confidence interval.

centile and percentile-t confidence intervals, we report the case when we use *Bootstrap Bias Correction II* combined with *Bootstrap Confidence Interval.* Both percentile and percentile-t confidence intervals also improve over the naive asymptotic intervals (10). The percentile confidence interval seem to dominate other intervals when c is small (N is large) or $\rho = 0.5$. The percentile-t confidence interval seem to work relatively well when c is small and the signal-to-noise ratio is small. As in the case of the bias correction result, the performance of all confidence intervals generally improves when the signal-to-noise ratio increases. Likewise, the performance deteriorates when errors are cross-sectionally correlated. Yet, their coverage is much closer to the nominal rate compared to the corresponding results for the naive asymptotic confidence interval. In summary, none of the three bootstrap confidence intervals uniformly dominate others so that three methods may be used complementarily in practice.

5 Empirical Application to US Diffusion Index

In this section, we apply our bootstrap procedure to the analysis of a diffusion index based on a dynamic factor model. Stock and Watson (1998, 2002) extract common factors from 215 U.S. monthly macroeconomic time series and report that the forecasts based on such diffusion indexes outperform the conventional forecasts.[11] We use the same data source (and transformations) as Stock and Watson and sample period is from 1959:3 to 1998:12 giving a maximum number of time series observation $T = 478$. By excluding the series with missing observation, we first construct a balance panel with $N = 159$.[12] For the purpose of visualizing the effect of small N on the estimation of persistence parameter of the single common factor, we then generate multiple subsamples using the following procedure. Based on the full balanced panel, we select variables 1, 4, 7 and so on to construct a balanced panel subsample. Next, construct another subsample by selecting variables 2, 5, 8 and so on. By

[11] The list provided in Appendix B of Stock and Watson (2002) shows that individual series are from 14 categories that consist of (1) real output and income; (2) employment and hours; (3) real retail, manufacturing and trade sales; (4) consumption; (5) housing starts and sales; (6) real inventories and inventory-sales ratios; (7) orders and unfilled orders; (8)stock prices; (9) exchange rates; (10) interest rates; (11) money and credit quantity aggregates; (12) price indexes; (13) average hourly earnings; and (14) miscellaneous.

[12] The number of the series in the full balanced panel differs from that of Stock and Watson (2002) due to the different treatment of outliers.

repeating such a selection three times, we can construct three balanced panel with $T = 478$ and $N = 53$. Similarly, we can select variables 1, 6, 11 and so on to construct five balanced panel with $T = 478$ and $N = 32$. Since numbers of the series in the full balanced panel and the two subsamples are $N = 159$, 53 and 32, corresponding $\sqrt{T}/N$ are 0.14, 0.41 and 0.68. Since the values of $\sqrt{T}/N$ are not close to zero, the bootstrap method is likely more appropriate than the naive asymptotic approximation in the two-step estimation. Diffusion indexes obtained in the first step by applying the principal components method are shown in Figure 1. The bold line shows the estimated common factor using the full balanced panel with $N = 159$. The darker shaded area represents the range of common factor estimates among three subsamples with $N = 53$ while the lighter shaded area represents the range of common factor estimates among five subsamples with $N = 32$. As the asymptotic theory predicts, we observe that the variation among the indexes based on $N = 32$ is much larger than the variation among indexes based on $N = 53$.

In the next step, we estimate the dynamic structure of three diffusion indexes using the AR(1) specification. Table 5 reports the point estimates $\widetilde{\rho}$, naive 90% confidence intervals (10), bias-corrected estimates $\widetilde{\rho}_{BC}$, and 90% confidence intervals (11) which are based on the bootstrap bias-corrected estimates. The bias-corrected estimates are computed with the number of bootstrap replication $B = 799$. One notable observation from this empirical exercise is that the size of the bootstrap bias correction is substantial for all three cases with the size largest for the $N = 32$ case and smallest for the $N = 159$ case. In addition, the non-overlapping range between the naive and bootstrap intervals seems to be wider when N is smaller. These observations are consistent with our finding in the Monte Carlo section.

6 Conclusion

In this paper, we examined the finite sample properties of the two-step estimator of the persistence parameter in dynamic factor models when unobservable common factors are estimated by the principal components methods in the first step. As a result of the simulation experiment with small N, we found that the AR coefficient estimator of positively autocor-

related factor is biased downward and the bias is larger for a more persistent factor. This finding is consistent with the theoretical prediction. The property of small N bias somewhat resembles the bias problem of AR estimator for small T. However, the bias caused by small N is also present in the large T case. When there is a possibility of such a downward bias, a bootstrap procedure proposed in the paper is effective in correcting bias and controlling the coverage rate of confidence interval.

Using a large number of series in the dynamic factor analysis has become a very popular approach in applications. The finding of this paper suggests that practitioners need to pay attention to the relative size of N and T before relying too much on a naive asymptotic approximation. Finally, it would be interesting to extend the experiments to allow for higher order AR models and nonlinear factor dynamics.

Appendix: Proofs

Proof of Proposition 1.

The principal components estimator $\widetilde{F} = \left[\widetilde{f}_1, \cdots, \widetilde{f}_T\right]'$ is the first eigenvector of the $T \times T$ matrix XX' with normalization $T^{-1}\sum_{t=1}^{T}\widetilde{f}_t^2 = 1$, where

$$X = \begin{bmatrix} X_1' \\ \vdots \\ X_T' \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{N1} \\ \vdots & \ddots & \vdots \\ x_{1T} & \cdots & x_{NT} \end{bmatrix}.$$

By definition, $(1/TN)XX'\widetilde{F} = \widetilde{F}v_{NT}$ where v_{NT} is the largest eigenvalue of $(1/TN)XX'$. Let $\zeta_{st} = N^{-1}\sum_{i=1}^{N} e_{is}e_{it} - E(e_{is}e_{it})$, $\eta_{st} = N^{-1}f_s\sum_{i=1}^{N}\lambda_i e_{it}$, and $\xi_{st} = N^{-1}f_t\sum_{i=1}^{N}\lambda_i e_{is}$. Following the proof of Theorem 5 in Bai (2003), the estimation error of the factor can be decomposed as

$$\begin{aligned} \widetilde{f}_t - H_{NT}f_t &= J_{NT}T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st} + J_{NT}T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\eta_{st} + J_{NT}T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\xi_{st} \\ &= O_P\left(N^{-1/2}\delta_{NT}^{-1}\right) + O_P\left(N^{-1/2}\right) + O_P\left(N^{-1/2}\delta_{NT}^{-1}\right) = O_P\left(N^{-1/2}\right) \end{aligned}$$

where $H_{NT} = (\widetilde{F}'F/T)(\Lambda'\Lambda/N)J_{NT}$, $J_{NT} = \left(v_{NT} - T^{-1}\overline{\sigma}_e^2\right)^{-1}$, $\overline{\sigma}_e^2 = N^{-1}\sum_{i=1}^{N}\sigma_{ei}^2$ and $\delta_{NT} = \min\{\sqrt{N}, \sqrt{T}\}$. From Bai's (2003) Lemma A.3, we have $\operatorname*{plim}_{T,N\to\infty} v_{NT} = \Sigma_\Lambda\Sigma_F = v$ and $\operatorname*{plim}_{T,N\to\infty} H_{NT}^2 = \operatorname*{plim}_{T,N\to\infty}(\widetilde{F}'F/T)(\Lambda'\Lambda/N)^2(F'\widetilde{F}/T)J_{NT}^2 = v\Sigma_\Lambda v^{-2} = \Sigma_\Lambda(\Sigma_\Lambda\Sigma_F)^{-1} = \Sigma_F^{-1} = 1$.

If f_{t-1} is observable,

$$\sqrt{T}\left(\widehat{\rho} - \rho\right) = \sqrt{T}\left(\sum_{t=1}^{T} f_{t-1}^2\right)^{-1}\sum_{t=1}^{T} f_{t-1}\varepsilon_t = T^{-1/2}\sum_{t=1}^{T} f_{t-1}\varepsilon_t + o_P(1)$$

since $T^{-1}\sum_{t=1}^{T} f_{t-1}^2 - 1 = o_P(1)$. If f_{t-1} is replaced with the factor estimator,

$$\begin{aligned} \sqrt{T}\left(\widetilde{\rho} - \rho\right) &= \sqrt{T}\left(\sum_{t=1}^{T}\widetilde{f}_{t-1}^2\right)^{-1}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t - \rho\widetilde{f}_{t-1}\right) \\ &= T^{-1/2}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t - \rho\widetilde{f}_{t-1}\right) + o_P(1) \end{aligned}$$

since $T^{-1}\sum_{t=1}^{T}\widetilde{f}_{t-1}^2 - 1 = o_P(1)$. By decomposing the dominant term, we have

$$\begin{aligned}
& T^{-1/2}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t - \rho\widetilde{f}_{t-1}\right) \\
= \;& T^{-1/2}H_{NT}\sum_{t=1}^{T}\widetilde{f}_{t-1}\varepsilon_t + T^{-1/2}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left\{\widetilde{f}_t - H_{NT}f_t - \rho\left(\widetilde{f}_{t-1} - H_{NT}f_{t-1}\right)\right\} \\
= \;& T^{-1/2}H_{NT}^2\sum_{t=1}^{T}f_{t-1}\varepsilon_t - T^{-1/2}\rho\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_{t-1} - H_{NT}f_{t-1}\right) \\
& + T^{-1/2}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t - H_{NT}f_t\right) + T^{-1/2}H_{NT}\sum_{t=1}^{T}\left(\widetilde{f}_{t-1} - H_{NT}f_{t-1}\right)\varepsilon_t.
\end{aligned}$$

We next show (i) $T^{-1}\rho\sum_{t=1}^{T}\widetilde{f}_{t-1}(\widetilde{f}_{t-1} - H_{NT}f_{t-1}) = 2\rho v^{-2}N^{-1}\Xi + o_P(\delta_{NT}^{-2})$; (ii) $T^{-1}\sum_{t=1}^{T}\widetilde{f}_{t-1}(\widetilde{f}_t - H_{NT}f_t) = \rho v^{-2}N^{-1}\Xi + o_P(\delta_{NT}^{-2})$; (iii) $T^{-1}H_{NT}\sum_{t=1}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})\varepsilon_t = o_P(\delta_{NT}^{-2})$.

We decompose (i) as,

$$\begin{aligned}
T^{-1}\rho\sum_{t=1}^{T}\widetilde{f}_{t-1}(\widetilde{f}_{t-1} - H_{NT}f_{t-1}) &= T^{-1}\rho\sum_{t=1}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})^2 + T^{-1}\rho\sum_{t=1}^{T}H_{NT}f_{t-1}(\widetilde{f}_{t-1} - H_{NT}f_{t-1}) \\
&= \rho(A + B).
\end{aligned}$$

For A, we have,

$$\begin{aligned}
A &= T^{-1}\sum_{t=1}^{T}T^{-2}J_{NT}^2[\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st-1} + \sum_{s=1}^{T}\widetilde{f}_s\eta_{st-1} + \sum_{s=1}^{T}\widetilde{f}_s\xi_{st-1}]^2 \\
&= J_{NT}^2T^{-1}\sum_{t=1}^{T}(A_{1t-1} + A_{2t-1} + A_{3t-1})^2.
\end{aligned}$$

First,

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T}A_{1t-1}^2 &= T^{-1}\sum_{t=1}^{T}[T^{-1}(\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st-1})]^2 = T^{-1}\sum_{t=1}^{T}[T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s + H_{NT}f_s)\zeta_{st-1}]^2 \\
&\leq T^{-1}\sum_{t=1}^{T}[T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)\zeta_{st-1}]^2 + T^{-1}\sum_{t=1}^{T}[T^{-1}\sum_{s=1}^{T}H_{NT}f_s\zeta_{st-1}]^2 \\
&= O_P(\delta_{NT}^{-2}N^{-1}) = o_P(\delta_{NT}^{-2}),
\end{aligned}$$

since

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T}[T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)\zeta_{st-1}]^2 &\leq [T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)^2][T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\zeta_{st-1}^2] \\
&= O_P(\delta_{NT}^{-2}N^{-1}),
\end{aligned}$$

where the last equality follows from Assumption E(iv), and

$$\begin{aligned} T^{-1}E\sum_{t=1}^{T}[T^{-1}\sum_{s=1}^{T}f_s\zeta_{st-1}]^2 &= T^{-1}E\sum_{t=1}^{T}[T^{-1}\sum_{s=1}^{T}f_sN^{-1}\sum_{i=1}^{N}(e_{it-1}e_{is}-E(e_{it-1}e_{is}))]^2 \\ &= \sigma_f^2T^{-1}E\sum_{t=1}^{T}T^{-2}\sum_{s=1}^{T}[N^{-1}\sum_{i=1}^{N}(e_{it-1}e_{is}-E(e_{it-1}e_{is}))]^2 \\ &= O_P((NT)^{-1}), \end{aligned}$$

provided $\sigma_f^2 = 1$. Second,

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}A_{2t-1}^2 &= T^{-3}\sum_{t=1}^{T}[\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)\eta_{st-1} + \sum_{s=1}^{T}H_{NT}f_s\eta_{st-1}]^2, \\ &= T^{-1}\sum_{t=1}^{T}(A_{2t-1.1}+A_{2t-1.2}+A_{2t-1.2}+A_{2t-1.4}). \end{aligned}$$

We have

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}A_{2t-1.1} &\leq [T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s^0)^2][T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\eta_{st-1}^2] \\ &= O_P(\delta_{NT}^{-2}N^{-1}), \end{aligned}$$

where the last equality follows from

$$\begin{aligned} T^{-2}E\sum_{t=1}^{T}\sum_{s=1}^{T}\eta_{st-1}^2 &= T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}E(N^{-1}f_s\sum_{i=1}^{N}\lambda_ie_{it-1})^2 = T^{-1}\sigma_f^2\sum_{t=1}^{T}E(N^{-1}\sum_{i=1}^{N}\lambda_ie_{it})^2 \\ &= O_P(N^{-1}); \end{aligned}$$

and

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}A_{2t-1.4} &= H_{NT}^2T^{-3}\sum_{t=1}^{T}(\sum_{s=1}^{T}f_sN^{-1}f_s\sum_{i=1}^{N}\lambda_ie_{it-1})(\sum_{s=1}^{T}f_sN^{-1}f_s\sum_{i=1}^{N}\lambda_ie_{it-1}) \\ &= H_{NT}^2(T^{-1}\sum_{s=1}^{T}f_s^2)^2T^{-1}\sum_{t=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_ie_{it-1})^2 \\ &= N^{-1}\Xi + o_P(\delta_{NT}^{-2}) = O_P(\delta_{NT}^{-2}); \end{aligned}$$

and

$$T^{-1}\sum_{t=1}^{T}A_{2t-1.2} \leq [T^{-1}\sum_{t=1}^{T}A_{2t-1.1}]^{1/2}[T^{-1}\sum_{t=1}^{T}A_{2t-1.4}]^{1/2} = O_P(\delta_{NT}^{-2}N^{-1/2}).$$

Therefore, $T^{-1}\sum_{t=1}^{T}A_{2t-1}^2 = N^{-1}\Xi + o_P(\delta_{NT}^{-2}) = O_P(\delta_{NT}^{-2})$. Third,

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}A_{3t-1}^2 &= T^{-1}\sum_{t=1}^{T}T^{-2}[\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)\xi_{st-1} + \sum_{s=1}^{T}H_{NT}f_s\xi_{st-1}]^2, \\ &= T^{-1}\sum_{t=1}^{T}(A_{3t-1.1}+A_{3t-1.2}+A_{3t-1.2}+A_{3t-1.4}) = o_P(\delta_{NT}^{-2}). \end{aligned}$$

The proof of $T^{-1}\sum_{t=1}^{T} A_{3t-1.1} = o_P(\delta_{NT}^{-2})$ and $T^{-1}\sum_{t=1}^{T} A_{3t-1.2} = o_P(\delta_{NT}^{-2})$ is similar as the proof of $T^{-1}\sum_{t=1}^{T} A_{2t-1.1} = o_P(\delta_{NT}^{-2})$ and $T^{-1}\sum_{t=1}^{T} A_{2t-1.2} = o_P(\delta_{NT}^{-2})$.

$$T^{-1}\sum_{t=1}^{T} A_{3t-1.4} = H_{NT}^2 (T^{-1}\sum_{t=1}^{T} f_{t-1}^2)^2 (T^{-1}\sum_{s=1}^{T} N^{-1}\sum_{i=1}^{N} f_s \lambda_i e_{is})^2 = O_P((NT)^{-1}),$$

since

$$\begin{aligned} E(T^{-1}N^{-1}\sum_{s=1}^{T}\sum_{i=1}^{N} f_s\lambda_i e_{is})^2 &= E[T^{-2}N^{-2}\sum_{s=1}^{T} f_s^2 (\sum_{i=1}^{N}\lambda_i e_{is})^2] \\ &= \sigma_f^2 E[T^{-2}\sum_{s=1}^{T} N^{-2}(\sum_{i=1}^{N}\lambda_i e_{is})^2] = O_P((NT)^{-1}). \end{aligned}$$

By the Cauchy-Schwartz inequality, we can show that $T^{-1}\sum_{t=1}^{T} A_{1t-1}A_{2t-1} = o_P(\delta_{NT}^{-2})$, $T^{-1}\sum_{t=1}^{T} A_{1t-1}A_{3t-1} = o_P(\delta_{NT}^{-2})$, and $T^{-1}\sum_{t=1}^{T} A_{2t-1}A_{3t-1} = o_P(\delta_{NT}^{-2})$. By combining all the results, we have $T^{-1}\sum_{t=1}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})^2 = J_{NT}^2 N^{-1}\Xi + o_P(\delta_{NT}^{-2}) = v^{-2}N^{-1}\Xi + o_P(\delta_{NT}^{-2})$. For B, we have

$$\begin{aligned} B &= H_{NT}J_{NT}T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}[f_{t-1}\widetilde{f}_s\zeta_{st-1} + f_{t-1}\widetilde{f}_s\eta_{st-1} + f_{t-1}\widetilde{f}_s\xi_{st-1}] \\ &= H_{NT}J_{NT}(B_1 + B_2 + B_3). \end{aligned}$$

First,

$$\begin{aligned} B_1 &= T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T} f_{t-1}\widetilde{f}_s\zeta_{st-1} \le [T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2]^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=1}^{T} f_{t-1}\zeta_{st-1})^2]^{1/2} \\ &= O_P((NT)^{-1/2}). \end{aligned}$$

Second,

$$\begin{aligned} B_2 &= T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T} f_{t-1}\widetilde{f}_s\eta_{st-1} \le [T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2]^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=1}^{T} f_{t-1}\eta_{st-1})^2]^{1/2} \\ &= [T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2]^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=1}^{T} f_{t-1}N^{-1}\sum_{i=1}^{N} f_s\lambda_i e_{it-1})^2]^{1/2} \\ &= [T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2]^{1/2}[(T^{-1}\sum_{s=1}^{T} f_s^2)(T^{-1}N^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N} f_{t-1}\lambda_i e_{it-1})^2]^{1/2} \\ &= O_P((NT)^{-1/2}). \end{aligned}$$

Third,

$$\begin{aligned} B_3 &= T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T} f_{t-1}\widetilde{f}_s\xi_{st-1} = T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T} f_{t-1}(\widetilde{f}_s - H_{NT}f_s + H_{NT}f_s)\xi_{st-1} \\ &= B_{31} + B_{32}. \end{aligned}$$

For B_{31},

$$\begin{aligned} B_{31} &= [T^{-1}\sum_{t=1}^{T} f_{t-1}^2][T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)N^{-1}\sum_{i=1}^{N}\lambda_i e_{is}] \\ &= T^{-2}\sum_{s=1}^{T} J_{NT}\sum_{t=1}^{T}[\widetilde{f}_t\zeta_{ts} + \widetilde{f}_t\eta_{ts} + \widetilde{f}_t\xi_{ts}]N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} + o_P(\delta_{NT}^{-2}) \\ &= J_{NT}(B_{31.1} + B_{31.2} + B_{31.3}) + o_P(\delta_{NT}^{-2}), \end{aligned}$$

where

$$\begin{aligned} B_{31.1} &= T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}\widetilde{f}_t\zeta_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} \\ &\leq [T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=1}^{T}\widetilde{f}_t\zeta_{ts})^2]^{1/2}[T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i e_{is})^2]^{1/2} = O_P(\delta_{NT}^{-1}N^{-1}), \end{aligned}$$

and

$$\begin{aligned} B_{31.2} &= T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}\widetilde{f}_t\eta_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} = (T^{-1}\sum_{t=1}^{T}\widetilde{f}_t f_t)[T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i e_{is})^2] \\ &= H_{NT}N^{-1}\Xi + o_P(\delta_{NT}^{-2}). \end{aligned}$$

and

$$B_{31.3} = T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}(\widetilde{f}_t - H_{NT}f_t + H_{NT}f_t)\xi_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} = O_P(\delta_{NT}^{-1}N^{-1}),$$

since

$$\begin{aligned} & T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}(\widetilde{f}_t - H_{NT}f_t)\xi_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} \\ \leq\ & [T^{-1}\sum_{t=1}^{T}(\widetilde{f}_t - H_{NT}f_t)^2]^{1/2}[T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{s=1}^{T}\xi_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is})^2]^{1/2} \\ \leq\ & [T^{-1}\sum_{t=1}^{T}(\widetilde{f}_t - H_{NT}f_t)^2]^{1/2}[(T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\xi_{ts}^2)T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i e_{is})^2]^{1/2} \\ =\ & O_P(\delta_{NT}^{-1}N^{-1}T^{-1/2}), \end{aligned}$$

and

$$\begin{aligned} T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}H_{NT}f_t\xi_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} &= H_{NT}T^{-2}[\sum_{s=1}^{T}N^{-1}\sum_{i=1}^{N}f_s\lambda_i e_{is}]^2. \\ &= O_P((NT)^{-1}). \end{aligned}$$

Thus, $B_{31} = H_{NT}N^{-1}\Xi + o_P(\delta_{NT}^{-2})$. For B_{32},

$$\begin{aligned} B_{32} &= T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}f_{t-1}H_{NT}f_s\xi_{st-1} = T^{-2}H_{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}f_{t-1}f_sN^{-1}\sum_{i=1}^{N}f_{t-1}\lambda_i e_{is} \\ &= (T^{-1}H_{NT}\sum_{t=1}^{T}f_{t-1}^2)(T^{-1}N^{-1}\sum_{s=1}^{T}\sum_{i=1}^{N}f_s\lambda_i e_{is}) = O_P((NT)^{-1/2}). \end{aligned}$$

Therefore, $B_3 = H_{NT}N^{-1}\Xi + o_P(\delta_{NT}^{-2})$. In total, we have $T^{-1}\rho\sum_{t=1}^{T} H_{NT}f_{t-1}(\tilde{f}_{t-1} - H_{NT}f_{t-1}) = \rho J_{NT}^2 H_{NT}^2 N^{-1}\Xi + o_P(\delta_{NT}^{-2}) = \rho v^{-2}N^{-1}\Xi + o_P(\delta_{NT}^{-2})$. Thus, $T^{-1}\rho\sum_{t=1}^{T}\widetilde{f}_{t-1}(\widetilde{f}_{t-1} - H_{NT}f_{t-1}) = 2\rho v^{-2}N^{-1}\Xi + o_P(\delta_{NT}^{-2})$.

Using an argument similar to the one used in (i), we can decompose (ii) as

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}\widetilde{f}_{t-1}(\widetilde{f}_t - H_{NT}f_t) &= T^{-1}\sum_{t=1}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})(\widetilde{f}_t - H_{NT}f_t) + H_{NT}T^{-1}\sum_{t=1}^{T} f_{t-1}(\widetilde{f}_t - H_{NT}f_t) \\ &= \rho v^{-2}N^{-1}\Xi + o_P(\delta_{NT}^{-2}). \end{aligned}$$

The proof is almost the same as the proof of (i). We only mention the difference. To show $T^{-1}\sum_{t=1}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})(\widetilde{f}_t - H_{NT}f_t) = o_P(\delta_{NT}^{-2})$, we need use

$$\begin{aligned} & T^{-1}\sum_{t=1}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})(\widetilde{f}_t - H_{NT}f_t) \\ = \; & J_{NT}^2 H_{NT}^2 T^{-3}\sum_{t=1}^{T}(\sum_{s=1}^{T} f_s N^{-1} f_s \sum_{i=1}^{N}\lambda_i e_{it-1})(\sum_{s=1}^{T} f_s N^{-1} f_s \sum_{i=1}^{N}\lambda_i e_{it}) + o_P(\delta_{NT}^{-2}) \\ = \; & J_{NT}^2 H_{NT}^2 T^{-3}(\sum_{s=1}^{T} f_s^2)^2 \sum_{t=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i e_{it-1})(N^{-1}\sum_{i=1}^{N}\lambda_i e_{it}) + o_P(\delta_{NT}^{-2}) \\ = \; & o_P(\delta_{NT}^{-2}). \end{aligned}$$

To show $H_{NT}T^{-1}\sum_{t=1}^{T} f_{t-1}(\widetilde{f}_t - H_{NT}f_t) = \rho v^{-2}N^{-1}\Xi + o_P(\delta_{NT}^{-2})$, we need use

$$\begin{aligned} H_{NT}T^{-1}\sum_{t=1}^{T} f_{t-1}(\widetilde{f}_t - H_{NT}f_t) &= [T^{-1}\sum_{t=1}^{T} f_{t-1}f_t][T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)N^{-1}\sum_{i=1}^{N}\lambda_i e_{is}] + o_P(\delta_{NT}^{-2}) \\ &= \rho T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} + o_P(\delta_{NT}^{-2}) \\ &= \rho H_{NT}N^{-1}\Xi + o_P(\delta_{NT}^{-2}). \end{aligned}$$

For (iii)

$$\begin{aligned} T^{-1}H_{NT}\sum_{t=1}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})\varepsilon_t &= H_{NT}J_{NT}T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}[\widetilde{f}_s\zeta_{st-1}\varepsilon_t + \widetilde{f}_s\eta_{st-1}\varepsilon_t + \widetilde{f}_s\xi_{st-1}\varepsilon_t] \\ &= H_{NT}J_{NT}(C_1 + C_2 + C_3). \end{aligned}$$

For C_1,

$$\begin{aligned} C_1 &= T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st-1}\varepsilon_t \le (T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2)^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=1}^{T}\zeta_{st-1}\varepsilon_t)^2]^{1/2} \\ &= O_P((NT)^{-1/2}), \end{aligned}$$

where the last equality follows from

$$\begin{aligned} T^{-1}\sum_{s=1}^{T} E(T^{-1}\sum_{t=1}^{T}\zeta_{st-1}\varepsilon_t)^2 &= T^{-1}\sum_{s=1}^{T} E(T^{-1}\sum_{t=1}^{T}[N^{-1}\sum_{i=1}^{N}(e_{it-1}e_{is}-E(e_{it-1}e_{is}))]\varepsilon_t)^2 \\ &= \sigma^2 T^{-1}\sum_{s=1}^{T} T^{-2}E\sum_{t=1}^{T}[N^{-1}\sum_{i=1}^{N}(e_{it-1}e_{is}-E(e_{it-1}e_{is}))]^2 \\ &= O_P((NT)^{-1}). \end{aligned}$$

For C_2,

$$C_2 = T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{f}_s\eta_{st-1}\varepsilon_t \leq (T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2)^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=1}^{T}\eta_{st-1}\varepsilon_t)^2]^{1/2} = O_P((NT)^{-1/2}),$$

where the last equality follows from

$$\begin{aligned} T^{-1}\sum_{s=1}^{T} E(T^{-1}\sum_{t=1}^{T}\eta_{st-1}\varepsilon_t)^2 &= \sigma^2 T^{-1}\sum_{s=1}^{T} E(T^{-2}\sum_{t=1}^{T}\eta_{st-1}^2) \\ &= \sigma^2 T^{-1}\sum_{s=1}^{T} E[T^{-2}\sum_{t=1}^{T}(N^{-1}\sum_{i=1}^{N} f_s^0\lambda_i^0 e_{it-1})^2] \\ &= \sigma^2\sigma_f^2 T^{-2}E[\sum_{t=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^0 e_{it-1})^2] = O_P((NT)^{-1}). \end{aligned}$$

Similarly, we can show $C_3 = O_P((NT)^{-1/2})$. In total, we have $T^{-1}H_{NT}\sum_{t=1}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})\varepsilon_t = O_P((NT)^{-1/2})$. Finally,

$$\sqrt{T}\,(\widetilde{\rho}-\rho) = T^{-1/2}\sum_{t=1}^{T} f_{t-1}\varepsilon_t - c\rho v^{-2}\Xi + o_P(1).$$

■

Proof of Proposition 2.

The bootstrap principal components estimator $\widetilde{F}^* = \left[\widetilde{f}_1^*, \cdots, \widetilde{f}_T^*\right]'$ is the first eigenvector of the $T \times T$ matrix $X^* X^{*\prime}$ with normalization $T^{-1}\sum_{t=1}^{T} \widetilde{f}_t^{*2} = 1$, where the bootstrap sample is given by

$$X^* = \begin{bmatrix} X_1^{*\prime} \\ \vdots \\ X_T^{*\prime} \end{bmatrix} = \begin{bmatrix} x_{11}^* & \cdots & x_{N1}^* \\ \vdots & \ddots & \vdots \\ x_{1T}^* & \cdots & x_{NT}^* \end{bmatrix}.$$

Analogous to the original version, we have $(1/TN)X^* X^{*\prime}\widetilde{F}^* = \widetilde{F}^* v_{NT}^*$ where v_{NT}^* is the largest eigenvalue of $(1/TN)X^* X^{*\prime}$. Let $\zeta_{st}^* = N^{-1}\sum_{i=1}^{N} e_{is}^* e_{it}^* - E^*(e_{is}^* e_{it}^*)$, $\eta_{st}^* = N^{-1} f_s^* \sum_{i=1}^{N} \lambda_i^* e_{it}^*$, and $\xi_{st}^* = N^{-1} f_t^* \sum_{i=1}^{N} \lambda_i^* e_{is}^* = \eta_{ts}^*$. As the proof of proposition 1, the estimation error of the factor can be decomposed as

$$\widetilde{f}_t^* - H_{NT}^* f_t^* = J_{NT}^* T^{-1} \sum_{s=1}^{T} \widetilde{f}_s^* \zeta_{st}^* + J_{NT}^* T^{-1} \sum_{s=1}^{T} \widetilde{f}_s^* \eta_{st}^* + J_{NT}^* T^{-1} \sum_{s=1}^{T} \widetilde{f}_s^* \xi_{st}^*$$

where $H_{NT}^* = (\widetilde{F}^{*\prime} F^*/T)(\Lambda^{*\prime}\Lambda^*/N) J_{NT}^*$, $J_{NT}^* = \left(v_{NT}^* - T^{-1}\bar{\sigma}_e^{*2}\right)^{-1}$, $\bar{\sigma}_e^{*2} = N^{-1}\sum_{i=1}^{N} \sigma_{ei}^{*2}$, and $\sigma_{ei}^{*2} = E^*(e_{is}^{*2})$. From Lemma C.1 of Gonçalves and Perron (2012), under our assumption F, FL and E, we have (i) $T^{-1}\sum_{t=1}^{T} |\widetilde{f}_t - H_{NT} f_t|^4 = O_P(1)$, (ii) $N^{-1}\sum_{i=1}^{N} |\tilde{\lambda}_i - H_{NT}^{-1}\lambda_i|^4 = O_P(1)$, (iii) $(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} \tilde{e}_{it}^4 = O_P(1)$. This result implies that $E^*(e_{it}^{*4}) = O_P(1)$,

$$E^* \lambda_i^{*4} = N^{-1}\sum_{i=1}^{N} \tilde{\lambda}_i^4 \leq 8N^{-1}\left(\sum_{i=1}^{N} |\tilde{\lambda}_i - H_{NT}^{-1}\lambda_i|^4 + \sum_{i=1}^{N} |H_{NT}^{-1}\lambda_i|^4\right) = O_P(1),$$

and

$$\begin{aligned} E^* \varepsilon_t^{*4} &= T^{-1}\sum_{t=1}^{T} (\widetilde{f}_t - \tilde{\rho}\widetilde{f}_{t-1})^4 = T^{-1}\sum_{t=1}^{T} [\widetilde{f}_t - H_{NT} f_t + H_{NT} f_t - \tilde{\rho}(\widetilde{f}_{t-1} - H_{NT} f_{t-1}) - \tilde{\rho} H_{NT} f_{t-1}]^4 \\ &\leq 4^3 T^{-1}\sum_{t=1}^{T} [(\widetilde{f}_t - H_{NT} f_t)^4 + (H_{NT} f_t)^4 + \tilde{\rho}^4 (\widetilde{f}_{t-1} - H_{NT} f_{t-1})^4 + (\tilde{\rho} H_{NT} f_{t-1})^4] = O_P(1). \end{aligned}$$

The bootstrap estimation error can be approximated as

$$\begin{aligned} \sqrt{T}\,(\tilde{\rho}^* - \tilde{\rho}) &= \sqrt{T}\left(\sum_{t=1}^{T} \left(\widetilde{f}_{t-1}^*\right)^2\right)^{-1} \sum_{t=1}^{T} \widetilde{f}_{t-1}^* \left(\widetilde{f}_t^* - \tilde{\rho}\widetilde{f}_{t-1}^*\right) \\ &= T^{-1/2}\sum_{t=1}^{T} \widetilde{f}_{t-1}^* \left(\widetilde{f}_t^* - \tilde{\rho}\widetilde{f}_{t-1}^*\right) + o_{P^*}(1) \end{aligned}$$

since $T^{-1}\sum_{t=1}^{T}\left(\widetilde{f}_{t-1}^{*}\right)^{2}-1=o_{P^*}(1)$. By decomposing the dominant term, we have

$$\begin{aligned}
& T^{-1/2}\sum_{t=1}^{T}\widetilde{f}_{t-1}^{*}\left(\widetilde{f}_{t}^{*}-\tilde{\rho}\widetilde{f}_{t-1}^{*}\right) \\
= \; & T^{-1/2}\sum_{t=1}^{T}\widetilde{f}_{t-1}^{*}\left\{\widetilde{f}_{t}^{*}-H_{NT}^{*}f_{t}^{*}-\tilde{\rho}\left(\widetilde{f}_{t-1}^{*}-H_{NT}^{*}f_{t-1}^{*}\right)\right\}+T^{-1/2}H_{NT}^{*}\sum_{t=1}^{T}\widetilde{f}_{t-1}^{*}\varepsilon_{t}^{*} \\
= \; & T^{-1/2}H_{NT}^{*2}\sum_{t=1}^{T}f_{t-1}^{*}\varepsilon_{t}^{*}-T^{-1/2}\tilde{\rho}\sum_{t=1}^{T}\widetilde{f}_{t-1}^{*}\left(\widetilde{f}_{t-1}^{*}-H_{NT}^{*}f_{t-1}^{*}\right) \\
& +T^{-1/2}\sum_{t=1}^{T}\widetilde{f}_{t-1}^{*}\left(\widetilde{f}_{t}^{*}-H_{NT}^{*}f_{t}^{*}\right)+T^{-1/2}H_{NT}^{*}\sum_{t=1}^{T}\left(\widetilde{f}_{t-1}^{*}-H_{NT}^{*}f_{t-1}^{*}\right)\varepsilon_{t}^{*}.
\end{aligned}$$

The leading term can be written as

$$T^{-1/2}\left(H_{NT}^{*2}-1\right)\sum_{t=1}^{T}f_{t-1}^{*}\varepsilon_{t}^{*}+T^{-1/2}\sum_{t=1}^{T}f_{t-1}^{*}\varepsilon_{t}^{*}=T^{-1/2}\sum_{t=1}^{T}f_{t-1}^{*}\varepsilon_{t}^{*}+o_{P^*}(1).$$

The last equality follows from the fact that $v_{NT}^{*}=v^{*}+o_{P^*}(1)$ where $v^{*}=\Sigma_{\Lambda}^{*}\Sigma_{F}^{*}$, $\Sigma_{\Lambda}^{*}=\widetilde{\Lambda}'\widetilde{\Lambda}/N\rightarrow^{P}v$ and $\Sigma_{F}^{*}=\widetilde{F}'\widetilde{F}/T=1$, and $H_{NT}^{*2}-1=o_{P^*}(1)$ because

$$H_{NT}^{*2}=(\widetilde{F}^{*\prime}F^{*}/T)(\Lambda^{*\prime}\Lambda^{*}/N)(F^{*\prime}\widetilde{F}^{*\prime}/T)J_{NT}^{*^{2}}=\Sigma_{F}^{*-1}+o_{p^*}(1).$$

In what follows, we show that (i) $T^{-1/2}\tilde{\rho}\sum_{t=1}^{T}\widetilde{f}_{t-1}^{*}\left(\widetilde{f}_{t-1}^{*}-H_{NT}^{*}f_{t-1}^{*}\right)=2\rho v^{2}N^{-1}\Xi+o_{P}^{*}(\delta_{NT}^{-2})$; (ii) $T^{-1/2}\tilde{\rho}\sum_{t=1}^{T}\widetilde{f}_{t-1}^{*}\left(\widetilde{f}_{t}^{*}-H_{NT}^{*}f_{t}^{*}\right)=\rho v^{2}N^{-1}\Xi+o_{P}^{*}(\delta_{NT}^{-2})$; (iii) $T^{-1}H_{NT}^{*}\sum_{t=1}^{T}(\widetilde{f}_{t-1}^{*}-H_{NT}^{*}f_{t-1}^{*})\varepsilon_{t}^{*}=o_{P}^{*}(\delta_{NT}^{-2})$

The proof of the above three equations is quite similar as those in Proposition 1. Here, we only focus on the differences. For (i), we decompose it as

$$\begin{aligned}
T^{-1}\tilde{\rho}\sum_{t=1}^{T}\widetilde{f}_{t-1}^{*}(\widetilde{f}_{t-1}^{*}-H_{NT}^{*}f_{t-1}^{*}) &= T^{-1}\tilde{\rho}\sum_{t=1}^{T}(\widetilde{f}_{t-1}^{*}-H_{NT}^{*}f_{t-1}^{*})^{2}+T^{-1}H_{NT}^{*}\tilde{\rho}\sum_{t=1}^{T}f_{t-1}^{*}(\widetilde{f}_{t-1}^{*}-H_{NT}^{*}f_{t-1}^{*}) \\
&= \tilde{\rho}(A^{*}+B^{*}).
\end{aligned}$$

We first show $T^{-1}\sum_{t=1}^{T}(\widetilde{f}_{t-1}^{*}-H_{NT}^{*}f_{t-1}^{*})^{2}=O_{P^*}(\delta_{NT}^{-2})$,

$$\begin{aligned}
A^{*} &= J_{NT}^{*2}\left[T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{s=1}^{T}\tilde{f}_{s}^{*}\zeta_{st}^{*}+T^{-1}\sum_{s=1}^{T}\tilde{f}_{s}^{*}\eta_{st}^{*}+T^{-1}\sum_{s=1}^{T}\tilde{f}_{s}^{*}\xi_{st}^{*})^{2}\right] \\
&\leq 3J_{NT}^{*2}\left[T^{-1}\sum_{t=1}^{T}T^{-2}(\sum_{s=1}^{T}\tilde{f}_{s}^{*}\zeta_{st}^{*})^{2}+T^{-1}\sum_{t=1}^{T}T^{-2}(\sum_{s=1}^{T}\tilde{f}_{s}^{*}\eta_{st}^{*})^{2}+T^{-1}\sum_{t=1}^{T}T^{-2}(\sum_{s=1}^{T}\tilde{f}_{s}^{*}\xi_{st}^{*})^{2}\right] \\
&= 3J_{NT}^{*2}\left[A_{1}^{*}+A_{2}^{*}+A_{3}^{*}\right].
\end{aligned}$$

First,

$$A_{1}^{*}\leq(T^{-1}\sum_{s=1}^{T}\tilde{f}_{s}^{*2})(T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\zeta_{st}^{*2})=O_{P^*}(\delta_{NT}^{-2})$$

where the last equality follows from

$$\begin{aligned} T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}E^*\zeta_{st}^{*2} &= T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}E^*(N^{-1}\sum_{i=1}^{N}[e_{is}^*e_{it}^* - E^*(e_{is}^*e_{it}^*)])^2 \\ &= (NT)^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{i=1}^{N}Var^*(e_{is}^*e_{it}^*) = O_P(N^{-1}) \end{aligned}$$

provided $E^*(e_{it}^{*4}) = O_P(1)$. Second,

$$\begin{aligned} A_2^* &= T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{s=1}^{T}\tilde{f}_s^*N^{-1}f_s^*\sum_{i=1}^{N}\lambda_i^*e_{it}^*)^2 = T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{s=1}^{T}\tilde{f}_s^*f_s^*)^2(N^{-1}\sum_{i=1}^{N}\lambda_i^*e_{it}^*)^2 \\ &\le T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{s=1}^{T}\tilde{f}_s^{*2})(T^{-1}\sum_{s=1}^{T}f_s^{*2})(N^{-1}\sum_{i=1}^{N}\lambda_i^*e_{it}^*)^2 \\ &= (T^{-1}\sum_{s=1}^{T}\tilde{f}_s^{*2})(T^{-1}\sum_{s=1}^{T}f_s^{*2})[T^{-1}\sum_{t=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^*e_{it}^*)^2] = O_{P^*}(\delta_{NT}^{-2}) \end{aligned}$$

which follows from $T^{-1}\sum_{s=1}^{T}f_s^{*2} = O_{P^*}(1)$ using Theorem 4.1 of Freedman (1984) and from

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}E^*(N^{-1}\sum_{i=1}^{N}\lambda_i^*e_{it}^*)^2 &= N^{-2}T^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}E^*(\lambda_i^{*2}e_{it}^{*2}) = N^{-2}T^{-1}\sum_{i=1}^{N}E^*(\lambda_i^{*2})(\sum_{t=1}^{T}e_{it}^{*2}) \\ &\le N^{-2}T^{-1}(\sum_{i=1}^{N}E^*\lambda_i^{*4})^{1/2}(\sum_{i=1}^{N}E^*(\sum_{t=1}^{T}e_{it}^{*2})^2)^{1/2} \\ &\le N^{-1}(N^{-1}\sum_{i=1}^{N}E^*\lambda_i^{*4})^{1/2}(N^{-1}\sum_{i=1}^{N}T^{-1}E^*\sum_{t=1}^{T}e_{it}^{*4})^{1/2} = O_P(N^{-1}), \end{aligned}$$

provided $E^*(\lambda_i^{*4}) = O_P(1)$. Third,

$$\begin{aligned} A_3^* &= (T^{-1}\sum_{t=1}^{T}f_t^{*2})(T^{-1}\sum_{s=1}^{T}\tilde{f}_s^*N^{-1}\sum_{i=1}^{N}\lambda_i^*e_{is}^*)^2 \\ &\le (T^{-1}\sum_{t=1}^{T}f_t^{*2})(T^{-1}\sum_{s=1}^{T}\tilde{f}_s^{*2})[T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^*e_{is}^*)^2] = O_{P^*}(\delta_{NT}^{-2})\ . \end{aligned}$$

Therefore, $A^* = O_{P^*}(\delta_{NT}^{-2})$. By using one additional condition that $N^{-1}T^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}f_t^*\lambda_i^*e_{it}^* =$

$O_P^*((NT)^{-1/2})$,

$$\begin{aligned}
E^*(N^{-1}T^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N} f_t^*\lambda_i^* e_{it}^*)^2 &= (NT)^{-2}[E^*\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} f_t^* f_s^* \lambda_i^*\lambda_j^* e_{it}^* e_{js}^*] \\
&= (NT)^{-2}E^*\sum_{t=1}^{T}\sum_{i=1}^{N} f_t^{*2}\lambda_i^{*2} e_{it}^{*2} \\
&\leq (NT)^{-2}(T^{-1}E^*\sum_{t=1}^{T} f_t^{*4})[T^{-1}E^*\sum_{t=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^{*2}e_{it}^{*2})^2] \\
&\leq (NT)^{-1}(T^{-1}E^*\sum_{t=1}^{T} f_t^{*4})[(NT)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N} E^*\lambda_i^{*4}E^* e_{it}^{*4}] \\
&= O_P(\delta_{NT}^{-4}),
\end{aligned}$$

then we show

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T}(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)^2 &= T^{-1}J_{NT}^{*2}\sum_{t=1}^{T}[T^{-1}\sum_{s=1}^{T} H_{NT}^* f_s^* \eta_{st-1}]^2 + o_P^*(\delta_{NT}^{-2}) \\
&= J_{NT}^{*2}H_{NT}^{*2}(T^{-1}\sum_{s=1}^{T} f_s^{*2})^2 T^{-1}\sum_{t=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{it-1}^*)^2 + o_P^*(\delta_{NT}^{-2}) \\
&= J_{NT}^{*2}H_{NT}^{*2}T^{-1}\sum_{t=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{it-1}^*)^2 + o_P^*(\delta_{NT}^{-2}) \\
&= v^{-2}N^{-1}\Xi + o_P^*(\delta_{NT}^{-2}),
\end{aligned}$$

where the last equality follows from

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T}(N^{-1/2}\sum_{i=1}^{N}\lambda_i^* e_{it}^*)^2 &= T^{-1}\sum_{t=1}^{T} Var^*(N^{-1/2}\sum_{i=1}^{N}\lambda_i^* e_{it}^*) + o_P^*(1) \\
&= T^{-1}N^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}\widetilde{\lambda}_i^2\widetilde{e}_{it}^2 + o_P^*(1) = \Xi + o_P^*(1),
\end{aligned}$$

provided we assume cross-sectionally uncorrelated errors ($\tau_{ij} = 0$) in the Bootstrap Bias Correction and the Bootstrap Confidence Interval. Therefore, we have $T^{-1}\tilde{\rho}\sum_{t=1}^{T}(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)^2 = \tilde{\rho}v^{-2}N^{-1}\Xi + o_P^*(\delta_{NT}^{-2}) = \rho v^{-2}N^{-1}\Xi + o_P^*(\delta_{NT}^{-2})$. To show $T^{-1}H_{NT}^*\tilde{\rho}\sum_{t=1}^{T} f_{t-1}^*(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*) = \rho v^{-2}N^{-1}\Xi + o_P^*(\delta_{NT}^{-2})$, we need

$$\begin{aligned}
T^{-1}H_{NT}^*\sum_{t=1}^{T} f_{t-1}^*(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*) &= J_{NT}^{*2}H_{NT}^* T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}\widetilde{f}_t^*\eta_{ts}^* N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{is}^* + o_P^*(\delta_{NT}^{-2}) \\
&= J_{NT}^{*2}H_{NT}^*(T^{-1}\sum_{t=1}^{T}\widetilde{f}_t^* f_t^*)[T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{is}^*)^2] + o_P^*(\delta_{NT}^{-2}) \\
&= v^{-2}N^{-1}\Xi + o_P^*(\delta_{NT}^{-2}).
\end{aligned}$$

Therefore, $T^{-1}H_{NT}^*\tilde{\rho}\sum_{t=1}^{T} f_{t-1}^*(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*) = \rho v^{-2}N^{-1}\Xi + o_P^*(\delta_{NT}^{-2})$. In total, we show $T^{-1/2}\tilde{\rho}\sum_{t=1}^{T}\widetilde{f}_{t-1}^*\left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right) = 2\rho v^2 N^{-1}\Xi + o_P^*(\delta_{NT}^{-2})$. Similarly, we can

show $T^{-1/2}\tilde{\rho}\sum_{t=1}^{T}\widetilde{f}_{t-1}^{*}\left(\widetilde{f}_{t}^{*}-H_{NT}^{*}f_{t}^{*}\right)=\rho v^{2}N^{-1}\Xi+o_{P}^{*}(\delta_{NT}^{-2})$ by using one additional condition that

$$\begin{aligned}
& T^{-1/2}\sum_{t=1}^{T}(\widetilde{f}_{t-1}^{*}-H_{NT}^{*}f_{t-1}^{*})(\widetilde{f}_{t}^{*}-H_{NT}^{*}f_{t}^{*}) \\
= \; & J_{NT}^{*2}H_{NT}^{*2}(T^{-1}\sum_{s=1}^{T}f_{s}^{*2})^{2}T^{-1}\sum_{t=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_{i}^{*}e_{it-1}^{*})(N^{-1}\sum_{i=1}^{N}\lambda_{i}^{*}e_{it}^{*})+o_{P}^{*}(\delta_{NT}^{-2}) \\
= \; & o_{P}^{*}(\delta_{NT}^{-2}).
\end{aligned}$$

Finally, we prove that

$$\sqrt{T}\,(\tilde{\rho}^{*}-\tilde{\rho})=T^{-1/2}\sum_{t=1}^{T}f_{t-1}^{*}\varepsilon_{t}^{*}-c\rho v^{-2}\Xi+o_{P}^{*}(1).$$

Since $E^{*}(T^{-1/2}\sum_{t=1}^{T}f_{t-1}^{*}\varepsilon_{t}^{*})=0$, $B^{-1}\sum_{b=1}^{B}\widetilde{\rho}_{b}^{*}-\widetilde{\rho}$ is a consistent estimator of the dominant term of $E(\tilde{\rho})-\rho$. ■

Proof of Proposition 3.

From proposition 2, we can apply the bootstrap central limit theorem to the term $T^{-1/2}\sum_{t=1}^{T}f_{t-1}^{*}\varepsilon_{t}^{*}$. Since $E^{*}[f_{t-1}^{*}\varepsilon_{t}^{*}|f_{t-2}^{*}\varepsilon_{t-1}^{*},...]=0$, we can use the central limit theorem for the martingale difference sequence under the bootstrap probability measure and thus $P^{*}(\sqrt{T}(\tilde{\rho}^{*}-\tilde{\rho})\leq x)$ approaches normal distribution function with mean $-c\rho v^{-2}\Xi$ and variance $E^{*}(f_{t-1}^{*2}\varepsilon_{t}^{*2})=T^{-1}\sum_{t=1}^{T}\tilde{f}_{t-1}^{2}\tilde{\varepsilon}_{t}^{2}$ under the bootstrap probability measure. Combining it with $T^{-1}\sum_{t=1}^{T}\tilde{f}_{t-1}^{2}\tilde{\varepsilon}_{t}^{2}\rightarrow^{P}E(f_{t-1}^{2}\varepsilon_{t}^{2})=\Gamma^{-1}$, we have $P^{*}(\sqrt{T}(\tilde{\rho}^{*}-\tilde{\rho})\leq x)-P(\sqrt{T}(\tilde{\rho}-\rho)\leq x)\rightarrow^{P}0$ for any x. By using Polya's theorem, we have the uniform convergence result given by

$$\sup_{x\in R}|P^{*}(\sqrt{T}(\tilde{\rho}^{*}-\tilde{\rho})\leq x)-P(\sqrt{T}(\tilde{\rho}-\rho)\leq x)|\rightarrow^{P}0.$$

■

References

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

Bai, J. and S. Ng (2002). "Determining the number of factors in approximate factor models." Econometrica 70(1), 191-221.

Bai, J. and S. Ng (2004). "A PANIC attack on unit roots and cointegration." Econometrica 72(4),1127-1177.

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

Baxter, M. and M. J. Crucini (1995). "Business cycles and the asset structure of foreign trade." International Economic Review 36, 821-854.

Beran, R. and M. S. Srivastava (1985). "Bootstrap tests and confidence regions for functions of a covariance matrix." Annals of Statistics 13, 95-115.

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

Boivin, J. and M. Giannoni (2006). "DSGE models in a data-rich environment." NBER Working Paper no. 12772.

Boivin, J. and S. Ng (2006). "Are more data always better for factor analysis?" Journal of Econometrics 132, 169-194.

Bose, A. (1988). "Edgeworth correction by bootstrap in autoregressions." Annals of Statistics 16, 1709-1722.

Bryan, M. F. and S. G. Cecchetti (1993). "The consumer price index as a measure of inflation," Federal Reserve Bank of Cleveland Economic Review, 15–24.

Connor, G., R. A. Korajczyk (1986). "Performance measurement with the Arbitrage Pricing Theory: A new framework for analysis." Journal of Financial Economics 15, 373-394.

Connor, G., R. A. Korajczyk and O. Linton (2006). "The common and specific components of dynamic volatility." Journal of Econometrics 132(1), 231-255.

Clark, T. E. (2006). "Disaggregate evidence on the persistence of consumer price inflation." Journal of Applied Econometrics 21, 563–587.

Diebold, F.X. (2000). ""Big Data" dynamic factor models for macroeconomic measurement and forecasting." manuscript presented at World Congress of the Econometric Society 2000.

Freedman, D. (1984). "On bootstrapping two-stage least-squares estimates in stationary linear models," Annals of Statistics 12, 827-842.

Geweke, J. (1977). "The dynamic factor analysis of economic time-series models." in D. J. Aigner and A.S. Goldberger (eds.), Latent Variable in Socioeconomic Models, Amsterdam, North-Holland, 365-387.

Gonçalves, S. and B. Perron (2012). "Bootstrapping factor-augmented regression models," working paper, University of Montreal.

Hall, P. (1992). The Bootstrap and Edgeworth Expansion. New York: Springer-Verlag.

Hansen, G.D. (1997). "Technical progress and aggregate fluctuations." Journal of Economic Dynamics and Control 21, 1005-1023.

Hurwicz, L. (1950). "Least squares bias in time series," in T. Koopmans (ed.) Statistical Inference in Dynamic Economic Models. Wiley, New York, 365-383.

Ireland, P. N. (2001). "Technology shocks and the business cycle: On empirical investigation." Journal of Economic Dynamics and Control 25, 703–719.

Kendall, M. G. (1954). "Note on bias in the estimation of autocorrelation." Biometrika 41, 403-403.

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

Kim, C.-J. and C. R. Nelson (1998). "Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime-switching." Review of Economics and Statistics 80, 188-201.

King, R. G., C. I. Plosser and S. T. Rebelo (1988). "Production, growth and business cycles: I. the basic neoclassical model." Journal of Monetary Economics 21, 195-232.

Ludvigson, S. and S. Ng (2010). "A factor analysis of bond risk premia." In A. Ullah and D. E. A. Giles (eds.), Handbook of Empirical Economics and Finance, 313-372.

MacKinnon, J. G. and A. A. Smith (1998). "Approximate bias correction in econometrics." Journal of Econometrics 85, 205-230.

Marcellino, M., J. H. Stock and M. W. Watson (2003). "Macroeconomic forecasting in the Euro area: Country specific versus area-wide information." European Economic Review 47, 1-18.

Marriott, F. H. C., and J. A. Pope (1954). "Bias in the estimation of autocorrelations." Biometrika, 41, 390–402.

Moon, H. R. and B. Perron (2004). "Testing for a unit root in panels with dynamic factors." Journal of Econometrics 122, 81-126.

Quenouille, M. H., (1949). "Approximate tests of correlation in time series." Journal of the Royal Statistical Society, Series B 11, 68–83.

Sargent, T. J. and C. Sims (1977). "Business cycle modeling without pretending to have too much a priori theory." In C. Sims (ed.), New Methods of Business Cycle Research. Minneapolis: Federal Reserve Bank of Minneapolis.

Shintani, M. (2005). "Nonlinear forecasting analysis using diffusion indexes: An application to Japan." Journal of Money, Credit, and Banking 37(3), 517-538.

Shintani, M. (2008). "A dynamic factor approach to nonlinear stability analysis." Journal of Economic Dynamics and Control 32(9), 2788-2808.

Stock, J. H. and M. W. Watson (1989). "New indexes of coincident and leading economic indicators." In O. Blanchard and S. Fischer (eds.), NBER Macroeconomics Annual (Cambridge, Mass. MIT Press).

Stock, J. H. and M. W. Watson (1998). "Diffusion indexes." NBER Working Paper no. 6702.

Stock, J. H. and M. W. Watson (2002). "Macroeconomic forecasting using diffusion indexes." Journal of Business and Economic Statistics 20, 147-162.

Woodford, M. (1999). "Optimal monetary policy inertia," NBER Working Paper no. 7261.

Yamamoto, Y. (2012). "Bootstrap inference for impulse response functions in factor-augmented vector autoregressions." working paper, University of Alberta.

Table 1: AR Estimation

		Estimator			
ρ	T	$\hat{\rho}$	$\hat{\rho}_{KBC}$	$\hat{\rho}_{BC}$	Coverage Rate
0.5	100	0.49	0.50	0.50	0.90
	200	0.50	0.50	0.50	0.90
0.9	100	0.88	0.90	0.90	0.91
	200	0.89	0.90	0.90	0.89

Note: Mean values of the OLS estimator ($\hat{\rho}$), Kendall-type bias-corrected estimator ($\hat{\rho}_{KBC}$) and bootstrap bias-corrected estimator ($\hat{\rho}_{BC}$) and coverage rates of the asymptotic confidence interval (5) in 1,000 replications.

Table 2: Two-step AR Estimation

			$\tilde{\rho}$					Coverage Rate				
ρ	T	c	S/N =0.5	0.75	1	1.5	2	S/N=0.5	0.75	1	1.5	2
(A) No cross-sectional correlation												
0.5	100	0.5	0.42	0.43	0.44	0.45	0.45	0.78	0.81	0.81	0.88	0.86
		1	0.36	0.39	0.41	0.42	0.43	0.58	0.68	0.76	0.78	0.80
		1.5	0.32	0.36	0.37	0.40	0.41	0.45	0.56	0.62	0.73	0.76
	200	0.5	0.45	0.46	0.46	0.47	0.47	0.79	0.83	0.85	0.88	0.86
		1	0.41	0.42	0.44	0.45	0.46	0.59	0.69	0.76	0.81	0.85
		1.5	0.37	0.39	0.41	0.43	0.44	0.41	0.54	0.60	0.72	0.76
0.9	100	0.5	0.73	0.76	0.78	0.81	0.81	0.26	0.38	0.47	0.58	0.61
		1	0.62	0.68	0.72	0.75	0.77	0.07	0.16	0.24	0.30	0.41
		1.5	0.54	0.60	0.65	0.70	0.72	0.04	0.06	0.12	0.18	0.25
	200	0.5	0.80	0.82	0.83	0.85	0.85	0.27	0.42	0.51	0.62	0.68
		1	0.71	0.76	0.79	0.81	0.82	0.05	0.14	0.22	0.35	0.44
		1.5	0.64	0.70	0.73	0.77	0.79	0.02	0.05	0.06	0.18	0.27
(B) Cross-sectional correlation												
0.5	100	0.5	0.40	0.42	0.43	0.45	0.45	0.72	0.78	0.81	0.86	0.85
		1	0.28	0.35	0.38	0.41	0.42	0.38	0.55	0.68	0.72	0.77
		1.5	0.22	0.28	0.32	0.37	0.39	0.25	0.36	0.50	0.62	0.69
	200	0.5	0.44	0.45	0.46	0.47	0.47	0.75	0.81	0.85	0.87	0.85
		1	0.37	0.41	0.43	0.45	0.45	0.41	0.59	0.70	0.79	0.83
		1.5	0.27	0.34	0.37	0.41	0.43	0.22	0.37	0.46	0.63	0.70
0.9	100	0.5	0.66	0.74	0.77	0.80	0.81	0.18	0.33	0.41	0.54	0.59
		1	0.43	0.57	0.64	0.71	0.73	0.05	0.13	0.19	0.28	0.36
		1.5	0.33	0.43	0.52	0.61	0.66	0.04	0.05	0.09	0.17	0.23
	200	0.5	0.78	0.81	0.83	0.84	0.85	0.22	0.38	0.48	0.59	0.66
		1	0.62	0.72	0.76	0.80	0.81	0.04	0.11	0.19	0.32	0.41
		1.5	0.47	0.60	0.65	0.74	0.77	0.02	0.05	0.06	0.17	0.24

Note: Mean values of the two-step estimator ($\tilde{\rho}$) and coverage rates of the asymptotic confidence interval (10) in 1,000 replications. S/N denotes signal-to-noise ratio.

Table 3: Bootstrap Bias Corrections

			$T = 100$					$T = 200$				
ρ	c		S/N =0.5	0.75	1	1.5	2	S/N =0.5	0.75	1	1.5	2
(A) No cross-sectional correlation												
0.5	0.5	bias	-0.09	-0.06	-0.06	-0.05	-0.04	-0.05	-0.04	-0.04	-0.03	-0.03
		1st bias	-0.05	-0.03	-0.03	-0.02	-0.01	-0.04	-0.02	-0.02	-0.01	-0.01
		bias*	-0.07	-0.06	-0.05	-0.05	-0.04	-0.05	-0.04	-0.03	-0.03	-0.03
	1	bias	-0.14	-0.11	-0.09	-0.08	-0.07	-0.10	-0.08	-0.07	-0.05	-0.04
		1st bias	-0.10	-0.07	-0.05	-0.03	-0.03	-0.07	-0.05	-0.04	-0.02	-0.02
		bias*	-0.09	-0.09	-0.08	-0.07	-0.07	-0.07	-0.06	-0.06	-0.05	-0.04
	1.5	bias	-0.18	-0.14	-0.12	-0.10	-0.09	-0.13	-0.10	-0.09	-0.07	-0.06
		1st bias	-0.15	-0.10	-0.08	-0.05	-0.04	-0.11	-0.07	-0.05	-0.04	-0.03
		bias*	-0.10	-0.10	-0.10	-0.09	-0.09	-0.09	-0.08	-0.08	-0.07	-0.06
0.9	0.5	bias	-0.17	-0.13	-0.12	-0.09	-0.08	-0.10	-0.08	-0.07	-0.06	-0.05
		1st bias	-0.09	-0.06	-0.05	-0.03	-0.02	-0.06	-0.04	-0.03	-0.02	-0.02
		bias*	-0.13	-0.12	-0.10	-0.09	-0.08	-0.09	-0.08	-0.07	-0.06	-0.05
	1	bias	-0.28	-0.22	-0.19	-0.15	-0.13	-0.19	-0.14	-0.12	-0.09	-0.08
		1st bias	-0.18	-0.12	-0.09	-0.06	-0.05	-0.13	-0.08	-0.06	-0.04	-0.03
		bias*	-0.17	-0.16	-0.15	-0.13	-0.12	-0.14	-0.12	-0.11	-0.09	-0.08
	1.5	bias	-0.36	-0.29	-0.24	-0.20	-0.17	-0.26	-0.20	-0.16	-0.13	-0.11
		1st bias	-0.27	-0.18	-0.14	-0.09	-0.07	-0.19	-0.13	-0.10	-0.06	-0.05
		bias*	-0.18	-0.18	-0.17	-0.16	-0.15	-0.16	-0.15	-0.14	-0.12	-0.11
(B) Cross-sectional correlation												
0.5	0.5	bias	-0.10	-0.07	-0.06	-0.05	-0.04	-0.06	-0.05	-0.04	-0.03	-0.03
		1st bias	-0.05	-0.03	-0.03	-0.02	-0.01	-0.04	-0.02	-0.02	-0.01	-0.01
		bias*	-0.07	-0.06	-0.05	-0.05	-0.04	-0.05	-0.04	-0.03	-0.03	-0.03
	1	bias	-0.21	-0.16	-0.12	-0.09	-0.08	-0.14	-0.10	-0.08	-0.05	-0.05
		1st bias	-0.10	-0.07	-0.05	-0.03	-0.03	-0.07	-0.05	-0.04	-0.02	-0.02
		bias*	-0.08	-0.08	-0.08	-0.07	-0.07	-0.07	-0.06	-0.06	-0.05	-0.04
	1.5	bias	-0.30	-0.22	-0.18	-0.14	-0.11	-0.22	-0.16	-0.12	-0.08	-0.07
		1st bias	-0.15	-0.10	-0.08	-0.05	-0.04	-0.11	-0.07	-0.05	-0.04	-0.03
		bias*	-0.08	-0.09	-0.09	-0.09	-0.09	-0.08	-0.08	-0.08	-0.07	-0.06
0.9	0.5	bias	-0.23	-0.16	-0.13	-0.10	-0.09	-0.12	-0.09	-0.08	-0.06	-0.05
		1st bias	-0.09	-0.06	-0.05	-0.03	-0.02	-0.06	-0.04	-0.03	-0.02	-0.02
		bias*	-0.12	-0.11	-0.10	-0.09	-0.08	-0.09	-0.07	-0.07	-0.05	-0.05
	1	bias	-0.45	-0.34	-0.27	-0.19	-0.16	-0.29	0.19	-0.15	-0.10	-0.09
		1st bias	-0.18	-0.12	-0.09	-0.06	-0.05	-0.13	-0.08	-0.06	-0.04	-0.03
		bias*	-0.14	-0.14	-0.14	-0.13	-0.12	-0.12	-0.11	-0.10	-0.09	-0.08
	1.5	bias	-0.58	-0.46	-0.37	-0.29	-0.23	-0.44	-0.31	-0.24	-0.17	-0.14
		1st bias	-0.27	-0.18	-0.14	-0.09	-0.07	-0.19	-0.13	-0.10	-0.06	-0.05
		bias*	-0.13	-0.15	-0.16	-0.16	-0.16	-0.13	-0.14	-0.14	-0.12	-0.11

Note: Actual bias (bias) and bootstrap bias estimator (bias*) are mean values in 1,000 replications. The first-order theoretical bias term (1st bias) is the first term in the right hand side of (7). S/N denotes signal-to-noise ratio.

Table 4: Coverage Rate of Bootstrap Confidence Intervals

			$T = 100$					$T = 200$				
ρ	c		S/N=0.5	0.75	1	1.5	2	S/N=0.5	0.75	1	1.5	2
(A) No cross-sectional correlation												
0.5	0.5	Bc	0.85	0.84	0.85	0.86	0.86	0.86	0.89	0.87	0.90	0.87
		Per	0.87	0.87	0.87	0.88	0.88	0.87	0.89	0.87	0.90	0.87
		Per-t	0.85	0.85	0.85	0.88	0.86	0.86	0.89	0.87	0.90	0.87
	1	Bc	0.77	0.81	0.83	0.85	0.84	0.83	0.83	0.84	0.87	0.87
		Per	0.78	0.83	0.87	0.87	0.88	0.82	0.86	0.87	0.89	0.89
		Per-t	0.81	0.81	0.83	0.85	0.83	0.86	0.84	0.84	0.87	0.87
	1.5	Bc	0.69	0.74	0.79	0.81	0.81	0.72	0.81	0.81	0.82	0.86
		Per	0.68	0.76	0.81	0.85	0.86	0.73	0.84	0.86	0.87	0.91
		Per-t	0.73	0.77	0.79	0.81	0.81	0.78	0.85	0.84	0.83	0.87
0.9	0.5	Bc	0.77	0.81	0.83	0.83	0.83	0.85	0.90	0.87	0.89	0.87
		Per	0.78	0.83	0.85	0.88	0.89	0.89	0.94	0.94	0.95	0.94
		Per-t	0.76	0.75	0.75	0.77	0.75	0.86	0.87	0.85	0.87	0.87
	1	Bc	0.61	0.72	0.74	0.80	0.80	0.71	0.80	0.82	0.86	0.84
		Per	0.54	0.68	0.71	0.80	0.81	0.72	0.83	0.87	0.93	0.92
		Per-t	0.70	0.73	0.72	0.76	0.72	0.83	0.84	0.85	0.86	0.82
	1.5	Bc	0.46	0.60	0.69	0.73	0.77	0.51	0.65	0.74	0.76	0.80
		Per	0.38	0.52	0.61	0.66	0.72	0.51	0.67	0.78	0.82	0.88
		Per-t	0.60	0.67	0.68	0.68	0.71	0.71	0.80	0.80	0.80	0.80
(B) Cross-sectional correlation												
0.5	0.5	Bc	0.83	0.85	0.85	0.87	0.86	0.84	0.88	0.87	0.89	0.87
		Per	0.82	0.86	0.86	0.88	0.88	0.85	0.89	0.87	0.90	0.88
		Per-t	0.84	0.85	0.85	0.87	0.87	0.85	0.89	0.87	0.90	0.88
	1	Bc	0.59	0.70	0.77	0.83	0.83	0.66	0.78	0.82	0.86	0.87
		Per	0.57	0.69	0.78	0.84	0.85	0.65	0.77	0.83	0.87	0.88
		Per-t	0.65	0.74	0.79	0.84	0.83	0.73	0.81	0.83	0.87	0.87
	1.5	Bc	0.43	0.60	0.67	0.75	0.78	0.48	0.67	0.74	0.80	0.83
		Per	0.38	0.56	0.66	0.75	0.80	0.45	0.64	0.74	0.83	0.85
		Per-t	0.51	0.64	0.71	0.77	0.79	0.57	0.71	0.78	0.83	0.84
0.9	0.5	Bc	0.62	0.74	0.77	0.81	0.81	0.75	0.85	0.86	0.87	0.88
		Per	0.60	0.75	0.79	0.83	0.87	0.79	0.89	0.90	0.93	0.94
		Per-t	0.63	0.70	0.73	0.76	0.75	0.78	0.86	0.84	0.85	0.87
	1	Bc	0.32	0.46	0.57	0.72	0.75	0.40	0.62	0.70	0.80	0.84
		Per	0.29	0.44	0.54	0.71	0.75	0.41	0.64	0.75	0.86	0.90
		Per-t	0.40	0.54	0.58	0.70	0.69	0.50	0.70	0.74	0.82	0.83
	1.5	Bc	0.23	0.36	0.48	0.60	0.68	0.24	0.40	0.54	0.64	0.70
		Per	0.17	0.30	0.41	0.53	0.63	0.24	0.42	0.59	0.71	0.81
		Per-t	0.31	0.42	0.51	0.60	0.66	0.38	0.54	0.66	0.72	0.73

Note: Coverage rates of three nominal 90% confidence intervals in 1,000 replications. Bc denotes bootstrap bias corrected asymptotic confidence interval (11), Per denotes equal-tailed percentile bootstrap confidence interval and Per-t denotes equal-tailed percentile-t bootstrap confidence interval. S/N denotes signal-to-noise ratio.

Table 5: AR(1) Estimates of US diffusion index

Series	$\tilde{\rho}$	Confidence interval	$\tilde{\rho}_{BC}$	Confidence interval
		(A) Full sample ($N = 159$)		
1	0.66	(0.60, 0.71)	0.69	(0.64, 0.75)
		(B) Long subsample ($N = 53$)		
1	0.65	(0.60, 0.71)	0.74	(0.69, 0.80)
2	0.58	(0.52, 0.64)	0.66	(0.60, 0.72)
3	0.68	(0.63, 0.73)	0.78	(0.72, 0.83)
average	0.64	(0.58, 0.69)	0.73	(0.67, 0.79)
		(C) Short subsample ($N = 32$)		
1	0.57	(0.51, 0.63)	0.75	(0.69, 0.81)
2	0.83	(0.79, 0.87)	0.95	(0.91, 0.99)
3	0.63	(0.58, 0.69)	0.75	(0.70, 0.81)
4	0.55	(0.49, 0.61)	0.65	(0.58, 0.71)
5	0.54	(0.48, 0.60)	0.67	(0.61, 0.73)
average	0.62	(0.57, 0.68)	0.75	(0.70, 0.81)

Note: Sample period is from 1959:3 to 1998:12 ($T = 478$). $c = \sqrt{T}/N$ is 0.14, 0.41 and 0.68, respectively, for series A, B and C. The first confidence interval next to $\tilde{\rho}$ is the 90% asymptotic confidence interval (10). The second confidence interval next to $\tilde{\rho}_{BC}$ is the 90% bootstrap bias-corrected confidence interval (11).

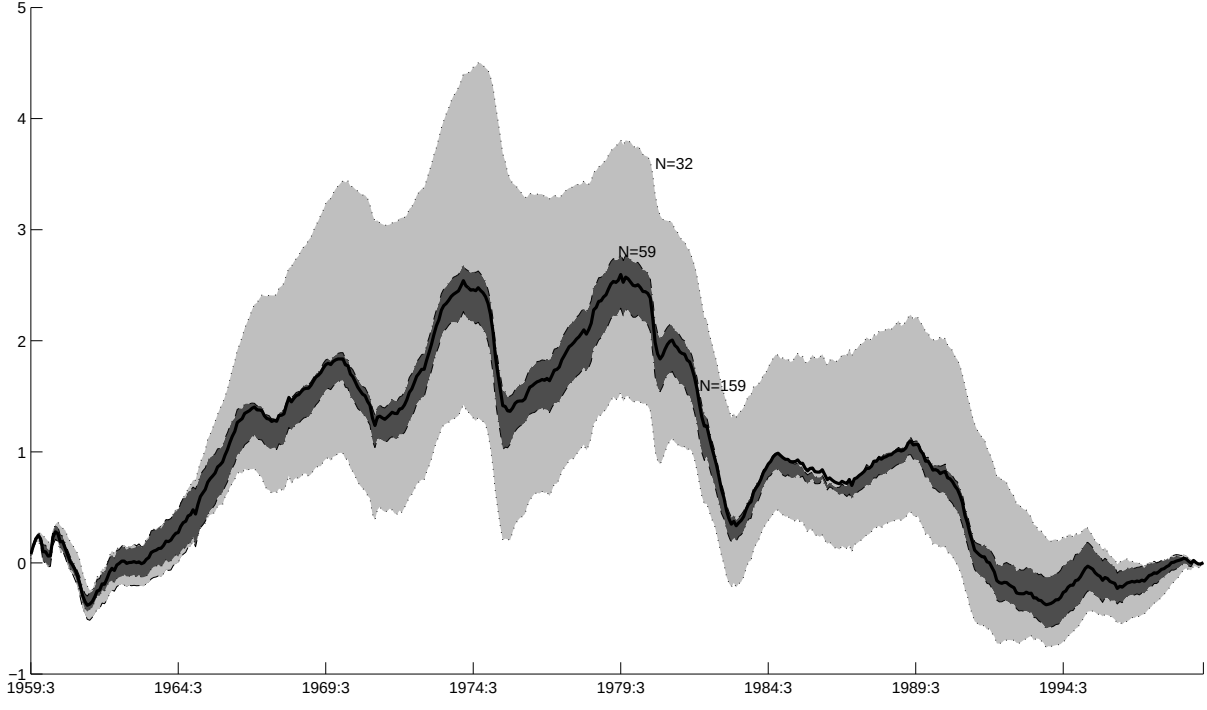

Figure 1: US Diffusion Index