Vanderbilt University Department of Economics Working Papers 15-00013

Improving the Finite Sample Performance of Autoregression Estimators in Dynamic Factor Models: A Bootstrap Approach

Mototsugu Shintani
University of Tokyo and Vanderbilt University

Zi-yi Guo
Vanderbilt University

Abstract

We investigate the finite sample properties of the estimator of a persistence parameter of an unobservable common factor when the factor is estimated by the principal components method. When the number of cross-sectional observations is not sufficiently large, relative to the number of time series observations, the autoregressive coefficient estimator of a positively autocorrelated factor is biased downward and the bias becomes larger for a more persistent factor. Based on theoretical and simulation analyses, we show that bootstrap procedures are effective in reducing the bias, and bootstrap confidence intervals outperform naive asymptotic confidence intervals in terms of the coverage probability.

Improving the Finite Sample Performance of Autoregression Estimators in Dynamic Factor Models: A Bootstrap Approach[*]

Mototsugu Shintani[†]and Zi-Yi Guo[‡]

This version: Novermber 2015

Abstract

We investigate the finite sample properties of the estimator of a persistence parameter of an unobservable common factor when the factor is estimated by the principal components method. When the number of cross-sectional observations is not sufficiently large, relative to the number of time series observations, the autoregressive coefficient estimator of a positively autocorrelated factor is biased downward and the bias becomes larger for a more persistent factor. Based on theoretical and simulation analyses, we show that bootstrap procedures are effective in reducing the bias, and bootstrap confidence intervals outperform naive asymptotic confidence intervals in terms of the coverage probability.

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

JEL classification: C15; C53

[†]RCAST, University of Tokyo and Department of Economics, Vanderbilt University; e-mail: `moto.shintani@gmail.com`.

[‡]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 since 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. In this paper, we investigate the finite sample properties of the two-step persistence estimator in dynamic factor models when an unobservable common factor is estimated by the principal components method in the first step. The first-step estimation is followed in the second step by the estimation of autoregressive models of the common factor. 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 and propose a simple bootstrap procedure that works well when N is relatively small.

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 in the implications of these models. For example, in a real business cycle model, there is a well-known trade-off between the persistence of the 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 largely depends on the persistence of real shocks in the economy (Woodford, 1999). In open economy models, the welfare gain from the introduction of in-

ternational 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 for measuring the persistence of the common components that may cause 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, the 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 consider only the case in which a single common factor is generated from a univariate autoregressive (AR) model of order one. This specification makes our problem simple and transparent 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 method is computationally convenient in estimating unobserved common factors with a large number of cross-sectional observations N. This 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 estimator of the common factor 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 large N asymptotic theory, an approximation

---

[1]Recently, Boivin and Giannoni (2006) proposed 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 models of higher order, however, there are several measures of persistence, including the sum of AR coefficients, the largest characteristic root and first-order autocorrelation.

[3]The principal 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).

may not work well when N is small relative to the time series observation T that is typically available in practice. Consistent with our theoretical prediction, the results from our Monte Carlo experiment using the positively autocorrelated factor suggest the downward bias in the AR coefficient estimator and significant under-coverage of the naive confidence interval when N is small. We show that a simple bootstrap procedure works well in correcting the bias and improves the performance of the confidence interval.

The bootstrap part of our analysis is closely related to recent studies by Gonçalves and Perron (2014) 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 (2014) employ a bootstrap procedure in factor-augmented forecasting regression models with multiple factors. The factor-augmented forecasting regression models are very useful in utilizing information from many predictors without including too many regressors. 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 (2014) provide the first order asymptotic validity of their bootstrap procedure for factor-augmented forecasting regression models, but not in the context of estimating the persistence parameter of the common factor. It should also be noted that, unlike their factor-augmented forecasting regression models with multiple factors, the bootstrap procedure for our univariate AR model of the common factor is not subject to scaling and rotation issues.[4] Yamamoto (2012) examines the performance of the bootstrap procedure applied to the 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.

---

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

The remainder of the paper is organized as follows. We first review the asymptotic theory of the two-step estimator, and discuss its small sample issues in Section 2. A bootstrap approach to reduce the bias is introduced in Section 3, and its usefulness is shown by the simulation in Section 4. An empirical illustration of our procedure is provided in Section 5. Some concluding remarks are made in Section 6. All the proofs of theoretical results are provided in the Appendix.

2 Two-Step Estimation of the Autoregressive Model of Latent Factor

We begin our discussion by reviewing the literature of finite sample bias correction of an 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 $t = 1, \ldots, T$. We consider a simple one-factor model given by

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

for $i = 1, \ldots, N$, 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. To introduce a dynamic structure in (1), we assume a zero-mean linear stationary AR(1) model of a common factor given by,

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

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

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

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

4

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

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

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 $SE(\widehat{\rho}) = (\widehat{\sigma}_\varepsilon^2 / \sum_{t=2}^{T+1} f_{t-1}^2)^{1/2}$, $\widehat{\sigma}_\varepsilon^2 = (T - 1)^{-1} \sum_{t=2}^{T} \widehat{\varepsilon}_t^2$ 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 $\widehat{\rho}_{KBC} = \widehat{\rho} + 2T^{-1}\widehat{\rho}_{KBC}$ obtained from the bias approximation formula $E(\widehat{\rho}) - \rho = -2T^{-1}\rho + O(T^{-2})$.[5] 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. A bootstrap method for AR models based on resampling residuals was later formalized by Bose (1988) and was extended to the multivariate case by Kilian (1998), among others. In particular, the bias-corrected estimator is given by $\widehat{\rho}_{BC} = \widehat{\rho} - \widehat{bias}$ where $\widehat{bias} = B^{-1} \sum_{b=1}^{B} \widehat{\rho}_b^* - \widehat{\rho}$ is the bootstrap bias estimator, $\widehat{\rho}_b^*$ is the b-th AR estimate from the bootstrap sample and B is the number of bootstrap replications. Both the Kendall-type bias correction and bootstrap bias

---

[5]This formula is valid if the intercept term is not included in the AR(1) model. With an intercept term, the bias-corrected estimator becomes $\widehat{\rho}_{KBC} = (T\widehat{\rho}+1)/(T-3)$ which is a solution to $\widehat{\rho}_{KBC} = \widehat{\rho}+T^{-1}(1+3\widehat{\rho}_{KBC})$ obtained from the bias approximation formula $E(\widehat{\rho}) - \rho = -T^{-1}(1 + 3\rho) + O(T^{-2})$.

correction reduce the small T bias by the order of T^{-1}.

To examine the finite sample properties of the OLS estimator $\widehat{\rho}$, we use the sample sizes $T = 100$ and 200, and generate the common factor f_t from (2) with the AR parameters, $\rho = 0.5$ and 0.9 combined with $\varepsilon_t \sim iidN(0, 1 - \rho^2)$. The initial value of f_t is drawn from the unconditional distribution of f_t, that is $N(0, 1)$. The mean values of $\widehat{\rho}$ along with the effective coverage rates of the nominal 90% conventional asymptotic confidence intervals (5) in 10,000 replications are reported in Table 1.[6] 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 $B = 199$. 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 the 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'.[7] 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

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

---

[6] Since our results are based on 10,000 replications, the standard error of 90% coverage rate in the simulation is about 0.003 ($\approx \sqrt{0.9 \times 0.1/10000}$).

[7] 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.

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 the factor, factor loadings and idiosyncratic errors. Below, we let M be some finite positive constant.

Assumption F (the factor): (i) $E|f_t|^4 \leq M$ and (ii) $F'F/T \xrightarrow{p} \sigma_f^2 = 1$ where $F = [f_1, \cdots, f_T]'$ as $T \to \infty$.

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

Assumption E (errors): (i) For all (i,t), $E(e_{it}) = 0$, $E|e_{it}|^8 \leq M$, (ii) $E(e_{is}e_{it}) = 0$ for all $t \neq s$, and $N^{-1}\sum_{i,j=1}^{N}|\tau_{ij}| \leq M$ where $\tau_{ij} = E(e_{it}e_{jt})$, (iii) $E|N^{-1/2}\sum_{i=1}^{N}[e_{it}e_{is} - E(e_{it}e_{is})]|^4 \leq M$ for all t and s and (iv) $(TN)^{-1}\sum_{t=1}^{T}\sum_{i,j=1}^{N}\lambda_i\lambda_j e_{it}e_{jt} \xrightarrow{p} \Gamma > 0$, as $N, T \to \infty$.

Since we focus on the AR(1) process of the factor, Assumption F is equivalent to the finite fourth moment condition of an i.i.d. error ε_t with variance $\sigma_\varepsilon^2 = 1 - \rho^2$ given the stationarity condition $|\rho| < 1$. Assumption FL can be replaced by the bounded deterministic sequence of factor loadings, but we only consider the case of random sequence in this paper. 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). Finally, 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 Assumptions F, FL, E and I hold. Then, as $T \to \infty$ and $N \to \infty$ such that $\sqrt{T}/N \to c$ where $0 \le c < \infty$,*

$$\sqrt{T}\,(\widetilde{\rho} - \rho) \xrightarrow{d} N(-c\rho\sigma_\lambda^{-4}\Gamma, 1 - \rho^2). \tag{7}$$

The proposition relies on the asymptotic framework employed by Bai (2003) and Gonçalves and Perron (2014) in their analysis of the 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 $T^{-1/2}$ is analogous to the bias term in the factor-augmented forecasting regression discussed by Ludvigson and Ng (2010) and Gonçalves and Perron (2014). 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 estimating 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

$$\sqrt{T}\,(\widetilde{\rho} - \rho) \xrightarrow{d} N(0, 1 - \rho^2) \tag{8}$$

as T tends to infinity, so that the limiting distribution of $\widetilde{\rho}$ in Proposition 1 is same as that of $\widehat{\rho}$ given by (4). In fact, we can further show the asymptotic equivalence of $\widetilde{\rho}$ and $\widehat{\rho}$ with their difference given by $\widetilde{\rho} - \widehat{\rho} = o_P(T^{-1/2})$.[8] 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

---

[8]See the proof of Proposition 1.

two-step estimator $\widetilde{\rho}$. For example, the 90% confidence interval can be constructed as

$$[\widetilde{\rho} - 1.645 \times SE(\widetilde{\rho}), \widetilde{\rho} + 1.645 \times SE(\widetilde{\rho})] \tag{9}$$

where $SE(\widetilde{\rho})$ is the standard error of $\widetilde{\rho}$ defined as $SE(\widetilde{\rho}) = (\widetilde{\sigma}_\varepsilon^2 / \sum_{t=2}^{T+1} \widetilde{f}_{t-1}^2)^{1/2}$, $\widetilde{\sigma}_\varepsilon^2 = (T - 1)^{-1} \sum_{t=2}^{T} \widetilde{\varepsilon}_t^2$ 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 (7) in Proposition 1, rather than by (8). In such a case, the presence of a bias term in (7) can result in bad coverage performance of a naive asymptotic confidence interval (9). Since the asymptotic bias term $-T^{-1/2} c \rho \sigma_\lambda^{-4} \Gamma$ can also be approximated by $-N^{-1} \rho \sigma_\lambda^{-4} \Gamma$, in what follows, we refer to this bias as the small N bias as opposed 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 (Γ).

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 $\lambda_i \sim iidN(0,1)$, the serially and cross-sectionally uncorrelated idiosyncratic error $e_{it} \sim iidN(0, \sigma_e^2)$, and the factor f_t from the same data generating process as before. The relative size of the common component and idiosyncratic error in x_{it} is expressed in terms 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 (2014) in considering 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 (9). The theoretical result for $c = 0$ implies that the coverage probability of (9) should be close to 0.90 only if N is sufficiently large relative to T, but we are interested in examining its finite sample 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 the two-step estimator $\widetilde{\rho}$ are clearly biased downward when N is small. 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 the 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 Γ increases). For the same parameter values for ρ, c and signal-to-noise ratio, the 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 Γ. However, when the signal-to-noise ratio is highest, the difference in point estimates between cross-sectionally uncorrelated and cross-sectionally correlated cases is 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 closest coverage to the nominal rate is obtained when $\rho = 0.5$ is combined with a small c (a large N) and a large signal-to-noise ratio. In this case, there is about a 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 $-c\rho\sigma_\lambda^{-4}\Gamma$ in (7) is not negligible, the confidence interval (9), which is based on approximation (8), cannot be expected to perform well. The presence of the small N bias results in under-coverage of the confidence interval (9) when N is small relative to T. 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 when N is small, by employing bootstrap procedures.

3 Bootstrapping the Autoregressive Model of the Latent Factor

In the previous section, we conjectured that the presence of the small N bias is likely the main source of poor coverage of the asymptotic confidence interval when N is small. 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 the bias analytically using the formula in Proposition 1, or estimate the bias using the bootstrap method for the purpose of correction. For example, Ludvigson and Ng (2010) consider the former approach in reducing bias in the context of the factor-augmented forecasting regression model. Here, we take the latter approach and employ the bootstrap procedure designed to work with cross-sectionally and serially uncorrelated errors. To be specific, we set $\tau_{ij} = 0$ for all $i \neq j$ in Assumption E(ii). However, in simulation, we also investigate its performance in the presence of cross-sectionally correlated errors ($\tau_{ij} \neq 0$). We first describe a simple bootstrap procedure for the bias correction.

Bootstrap I

1. Estimate the factor and factor loadings using the 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}^* to compute $\widetilde{f}_t^*$ and set $\widetilde{\rho}^* = \left(\sum_{t=2}^{T+1} \widetilde{f}_{t-1}^{*2} \right)^{-1}$ $\sum_{t=2}^{T} \widetilde{f}_{t-1}^* \widetilde{f}_t^*$ if $v_{NT}^* \geq \epsilon$ and $\widetilde{\rho}^* = \widetilde{\rho}$ otherwise. Here, ϵ is some small positive number, v_{NT}^* is the largest eigenvalue of $(1/TN) X^* X^{*\prime}$ where X^* is the $T \times N$ bootstrap data matrix with t-th row $X_t^{*\prime} = (x_{1t}^*, \ldots, x_{Nt}^*)$ and $\widetilde{\rho}$ is the AR estimate from $\widetilde{f}_t$.

4. Repeat steps 2 to 3 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. 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 the estimated factor model in step 2.

In the implementation of the bootstrap, theoretically, it is possible that the first principal components cannot be computed for some bootstrap sample if an associated eigenvalue is extremely small. In such a case, we just set $\widetilde{\rho}^* = \widetilde{\rho}$ for the corresponding bootstrap sample. This modification, however, does not affect the asymptotic property of the bootstrap estimator of bias.

It should be noted that the procedure above is designed to evaluate the small N bias rather than the small T bias. In order to incorporate both the small T bias and the small N bias simultaneously, we may combine the procedure above with bootstrapping AR models. This possibility is considered in the second bootstrap bias correction method described below.

Bootstrap II

1. Estimate the factor and factor loadings using the 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 the pseudo factor using $f_t^* = \widetilde{\rho}f_{t-1}^* + \varepsilon_t^*$.

4. Recenter $\widetilde{e}_{it}$ and $\widetilde{\lambda}_i$ 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 the principal components method to x_{it}^* to compute $\widetilde{f}_t^*$ and set $\widetilde{\rho}^* = \left(\sum_{t=2}^{T+1} \widetilde{f}_{t-1}^{*2} \right)^{-1} \sum_{t=2}^{T} \widetilde{f}_{t-1}^* \widetilde{f}_t^*$ if $v_{NT}^* \geq \epsilon$ and $\widetilde{\rho}^* = \widetilde{\rho}$ otherwise.

6. Repeat steps 2 to 5 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. 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 AR 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).

Let P^* denote the probability measure induced by the bootstrap conditional on the original sample, and let E^* denote expectation with respect to the distribution of the bootstrap sample conditional on the original sample. The following proposition provides the consistency of the bootstrap distribution.

Proposition 2. *Let all the assumptions of Proposition 1 hold with $\tau_{ij} = 0$ for all $i \neq j$, and the bootstrap data be generated as described in Bootstrap I or in Bootstrap II. 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 2 implies the first-order asymptotic validity of our bootstrap procedure in the sense that the limiting distribution of the bootstrap estimator $\tilde{\rho}^*$ is asymptotically equivalent to that of $\tilde{\rho}$.[9] Since the limiting distribution of $\tilde{\rho}$ is given by (7) in Proposition 1, the same distribution can be used to describe the limiting behavior of $\tilde{\rho}^*$. Thus, we conjecture that the small N bias term $-T^{-1/2} c \rho \sigma_\lambda^{-4} \Gamma$ can be corrected by using the bootstrap procedure. However, since the consistency of the bootstrap distribution does not necessarily imply the convergence of the bootstrap moment estimator, a bootstrap version of the uniform integrability condition is required to establish the consistency of the bootstrap bias estimator. While direct verification of the uniform integrability is typically complicated, Gonçalves and White (2005) utilized a convenient sufficient condition of the uniform integrability to prove the consistency of the bootstrap variance estimator in the context of regression models. In this paper, we focus on a similar sufficient condition $E^* \left(|\sqrt{T}(\tilde{\rho}^* - \tilde{\rho})|^{1+\delta} \right) = O_p(1)$ for some $\delta > 0$ in order to obtain the uniform integrability of the sequence $\{\sqrt{T}(\tilde{\rho}^* - \tilde{\rho})\}$. The asymptotic justification of using our bootstrap methods to correct the small N bias is established in the following proposition.

Proposition 3. *Let all the assumptions of Proposition 1 hold with $\tau_{ij} = 0$ for all $i \neq j$, $E|f_t|^{32} \leq M$, $E|\lambda_i|^{32} \leq M$, $E|e_{it}|^{64} \leq M$, and the bootstrap data be generated as described in Bootstrap I or in Bootstrap II. Then, as $T \to \infty$ and $N \to \infty$ such that $\sqrt{T}/N \to c$ where $0 \leq c < \infty$, $E^*(\tilde{\rho}^* - \tilde{\rho}) = -T^{-1/2} c \rho \sigma_\lambda^{-4} \Gamma + o_P(T^{-1/2})$.*

---

[9]In general, signs of the coefficients in the factor forecasting regression cannot be identified, and Gonçalves and Perron (2014) 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.

Proposition 3 implies the consistency of the bootstrap bias estimator $bias^*$ since $E^*(\widetilde{\rho}^* - \widetilde{\rho})$ can be accurately approximated by $bias^*$ with a suitably large value of B. The proposition also suggests that the bias-corrected estimator $\widetilde{\rho}_{BC} = \widetilde{\rho} - bias^*$ has the asymptotic bias of order smaller than $T^{-1/2}$. Since the same result holds for both *Bootstrap I* and *Bootstrap II,* whether or not bootstrapping AR models is included in the procedure does not matter asymptotically.

4   Monte Carlo Experiments

Let us now conduct the simulation to evaluate the performance of the bootstrap bias correction method. The results of the simulation under the same specification as in Table 2 are shown in Table 3. For each specification, the true bias is first evaluated by using the mean value of $\tilde{\rho} - \rho$ in 10,000 replications. The theoretical asymptotic bias $-T^{-1/2}c\rho\sigma_\lambda^{-4}\Gamma$ is also reported. The performance of bootstrap bias estimator based on *Bootstrap I* and *Bootstrap II* is evaluated by using the mean value of $bias^*$ in 10,000 replications. The number of the bootstrap replications is set at $B = 199$.

The results of the simulation can be summarized as follows. First, results turn out to be very similar between the cases of *Bootstrap I* and *Bootstrap II.* This finding suggests that the small T bias is almost negligible for the size of T we consider, which is consistent with the results in Table 1. Two bootstrap bias estimates match closely with the true bias for both cases of $\rho = 0.5$ and $\rho = 0.9$ unless the signal-to-noise ratio is too small. Second, while the direction of the changes in bias is consistent with the theoretical prediction, the asymptotic bias only accounts for a fraction of the actual bias. In many cases, bootstrap bias estimates are much closer to the actual bias than the asymptotic bias predicted by the theory. 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 the bootstrap correction method seems to be satisfactory.

Since the bootstrap bias correction method has been 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 a naive asymptotic confidence interval (9) from the nominal rate is proportional to the size of bias in Table 2. Thus, it is natural to expect that recentered asymptotic confidence intervals using the bootstrap bias-corrected estimates improve the coverage accuracy. For example, the 90% confidence interval can be constructed as

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

The asymptotic validity of the confidence interval (10) can easily be shown by combining the results in Propositions 1 to 3.

Instead of using a bias-corrected estimator, we can directly utilize the bootstrap distribution of the estimator to construct bootstrap confidence intervals. Here, we consider the percentile confidence interval based on the recentered bootstrap estimator $\widetilde{\rho}^* - \widetilde{\rho}$ as well as the percentile-t equal-tailed confidence interval based on the bootstrap t statistic defined as $t(\widetilde{\rho}^*) = (\widetilde{\rho}^* - \widetilde{\rho})/SE(\widetilde{\rho}^*)$ where $SE(\widetilde{\rho}^*)$ is the standard error of $\widetilde{\rho}^*$, which is asymptotically pivotal.[10] For example, the 90% percentile confidence interval and 90% percentile-t equal-tailed confidence interval can be constructed as

$$[\widetilde{\rho} - q_{0.95}(\widetilde{\rho}^* - \widetilde{\rho}), \widetilde{\rho} - q_{0.05}(\widetilde{\rho}^* - \widetilde{\rho})] \tag{11}$$

and

$$[\widetilde{\rho} - q_{0.95}(t(\widetilde{\rho}^*)) \times SE(\widetilde{\rho}), \widetilde{\rho} - q_{0.05}(t(\widetilde{\rho}^*)) \times SE(\widetilde{\rho})] \tag{12}$$

respectively, where $q_\alpha(x)$ denotes $100 \times \alpha$-th percentile of x. We now describe our procedure

---

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

of constructing the bootstrap confidence intervals.

Bootstrap Confidence Interval

1. Follow either steps 1 to 2 in *Bootstrap I* or steps 1 to 4 in *Bootstrap II*.

2. Compute the bootstrap AR coefficient estimate $\widetilde{\rho}^* = \left(\sum_{t=2}^{T+1} \widetilde{f}_{t-1}^{*2} \right)^{-1} \sum_{t=2}^{T} \widetilde{f}_{t-1}^* \widetilde{f}_t^*$ and $t(\widetilde{\rho}^*) = (\widetilde{\rho}^* - \widetilde{\rho})/SE(\widetilde{\rho}^*)$ if $v_{NT}^* \geq \epsilon$ and $\widetilde{\rho}^* = \widetilde{\rho}$ and $t(\widetilde{\rho}^*) = t(\widetilde{\rho}) = (\widetilde{\rho} - \rho)/SE(\widetilde{\rho})$ otherwise.

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

Note that, as in Kilian's (1998) argument on vector autoregression, $\widetilde{\rho}$ in step 3 in *Bootstrap II* can be replaced by the bias-corrected estimator $\widetilde{\rho}_{BC}$ without changing the limiting distribution of the bootstrap estimator. Proposition 2 implies that the coverage rate of the percentile bootstrap confidence interval approaches the nominal coverage rate in the limit. Similarly, we can modify Proposition 2 and replace $\tilde{\rho}^*$ and $\tilde{\rho}$ by their studentized statistics $t(\widetilde{\rho}^*)$ and $t(\widetilde{\rho})$ and show the bootstrap consistency of $t(\widetilde{\rho}^*)$ and the asymptotic validity of the percentile-t confidence interval.

Table 4 reports coverage of three confidence intervals based on the bootstrap applied to the two-step estimator $\widetilde{\rho}$ for the $\rho = 0.5$ and $\rho = 0.9$ cases. Here, for the bootstrap bias correction method required in the confidence interval (10), we use *Bootstrap II*. Similarly, we report percentile and percentile-t confidence intervals based on *Bootstrap Confidence Interval* combined with *Bootstrap II*. The table shows that all three confidence intervals significantly improve over the naive asymptotic interval (9) in Table 2. Especially, when $T = 200$, $c = 0.5$ and $\rho = 0.5$, the coverage rates of all three bootstrap intervals are very close to each other and are nearly the nominal rate, regardless of the signal-to-noise ratio. The percentile confidence

17

interval (11) seems to work relatively well when $T = 100$. The percentile-t confidence interval (12) seems to dominate the bias-corrected confidence interval (10) for all the cases.

As in the case of the bias correction result, the performance of confidence intervals tends to improve 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 when compared to the corresponding results for the naive asymptotic confidence interval. In summary, the percentile-t confidence interval works at least as well as the bias-corrected confidence interval, but does not uniformly dominates the percentile confidence interval. Therefore, we suggest using three methods 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 the 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 observations, 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, we 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 the 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 data sets with $T = 478$ and $N = 53$. Similarly, we can select variables 1, 6, 11 and so on to construct five balanced panels with $T = 478$ and $N = 31$. Since the number of the series in the full balanced panel and the two subsamples are $N = 159$, 53 and 31, corresponding $\sqrt{T}/N$ are 0.14, 0.41 and 0.71. 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 as the cumulative sums of the first principal components of panel data sets, 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 $N = 31$. As the asymptotic theory predicts, we observe that the variation among the indexes based on $N = 31$ is much larger than the variation among indexes based on $N = 53$.

In the next step, we estimate the persistence of three diffusion indexes using the AR(1) specification. Table 5 reports the point estimates $\widetilde{\rho}$, naive 90% confidence intervals (9), bias-corrected estimates $\widetilde{\rho}_{BC}$ and bootstrap-based 90% confidence intervals (10), (11) and (12). The bias-corrected estimates and bootstrap-based confidence intervals are computed with the number of bootstrap replication $B = 799$. One notable observation from this empirical exercise is that the magnitude of the bootstrap bias correction is substantial for all three cases. The estimated bias is largest in the case of $N = 31$ and is smallest in the case of $N = 159$. 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 findings from the Monte Carlo simulation.

6 Conclusion

In this paper, we examined the finite sample properties of the two-step estimator of a persistence parameter in dynamic factor models when an unobservable common factor is estimated by the principal components method in the first step. As a result of the simulation experiment with small N, we found that the AR coefficient estimator of a positively autocorrelated factor is biased downward, and the bias is larger for a more persistent factor. This finding is consistent with the theoretical prediction. The properties of the small N bias somewhat resemble those of the small T bias of the AR estimator. However, the bias caused by the small N is also present in the large T case. When there is a possibility of such a downward bias, a bootstrap procedure can be effective in correcting bias and controlling the coverage rate of the 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 $(1/TN)XX'$. Following the proof of Theorem 5 in Bai (2003), the estimation error of the factor $\widetilde{f}_t - H_{NT}f_t = O_P\left(N^{-1/2}\right)$, where $H_{NT} = (\widetilde{F}'F/T)(\Lambda'\Lambda/N)v_{NT}^{-1}$ and $\delta_{NT} = \min\{\sqrt{N}, \sqrt{T}\}$. From Bai's (2003) Lemma A.3, we have $\plim_{T,N\to\infty} v_{NT} = \sigma_\lambda^2\sigma_f^2 = v$ and $\plim_{T,N\to\infty} H_{NT}^2 = \plim_{T,N\to\infty} (\widetilde{F}'F/T)(\Lambda'\Lambda/N)^2(F'\widetilde{F}/T)v_{NT}^{-2} = v\sigma_\lambda^2 v^{-2} = \sigma_\lambda^2(\sigma_\lambda^2\sigma_f^2)^{-1} = \sigma_f^{-2} = 1$.

If f_t is observable,

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

since $T^{-1}\sum_{t=1}^T f_t^2 = 1 + o_P(1)$. If f_t is replaced with $\widetilde{f}_t$, we have

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

Under the assumptions, we can show (i) $T^{-1}\rho \sum_{t=2}^{T} \widetilde{f}_{t-1}(\widetilde{f}_{t-1} - H_{NT}f_{t-1}) = 2\rho v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$; (ii) $T^{-1}\sum_{t=2}^{T} \widetilde{f}_{t-1}(\widetilde{f}_t - H_{NT}f_t) = \rho v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$; and (iii) $T^{-1}H_{NT}\sum_{t=2}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})\varepsilon_t = o_P(\delta_{NT}^{-2})$. Since proofs of (i) and (iii) are almost same as those of Lemma A.2 (b) and Lemma A.1 in Gonçalves and Perron (2014), respectively, we only show (ii) below. Note that

$$T^{-1}\sum_{t=2}^{T} \widetilde{f}_{t-1}(\widetilde{f}_t - H_{NT}f_t)$$

$$= T^{-1}\sum_{t=2}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})(\widetilde{f}_t - H_{NT}f_t) + H_{NT}T^{-1}\sum_{t=2}^{T} f_{t-1}(\widetilde{f}_t - H_{NT}f_t).$$

For the first term, we have

$$T^{-1}\sum_{t=2}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})(\widetilde{f}_t - H_{NT}f_t)$$

$$= v_{NT}^{-2}H_{NT}^2 T^{-3}\sum_{t=2}^{T}(\sum_{s=1}^{T} f_s N^{-1}f_s \sum_{i=1}^{N}\lambda_i e_{it-1})(\sum_{s=1}^{T} f_s N^{-1}f_s \sum_{i=1}^{N}\lambda_i e_{it}) + o_P(\delta_{NT}^{-2})$$

$$= v_{NT}^{-2}H_{NT}^2 T^{-3}(\sum_{s=1}^{T} f_s^2)^2 \sum_{t=2}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i e_{it-1})(N^{-1}\sum_{i=1}^{N}\lambda_i e_{it}) + o_P(\delta_{NT}^{-2})$$

$$= o_P(\delta_{NT}^{-2}).$$

For the second term, we have

$$H_{NT}T^{-1}\sum_{t=2}^{T} f_{t-1}(\widetilde{f}_t - H_{NT}f_t)$$

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

$$= \rho T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} + o_P(\delta_{NT}^{-2})$$

$$= N^{-1}\rho H_{NT}\Gamma + o_P(\delta_{NT}^{-2}) = N^{-1}\rho v^{-2}\Gamma + o_P(\delta_{NT}^{-2}).$$

Combining the two results yields (ii). We can thus use (i), (ii), (iii), $H_{NT}^2 - 1 = o_P(1)$ and $T^{1/2}N^{-1} - c = o(1)$ to obtain

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

The desired result follows from the central limit theorem applied to the martingale difference sequence $f_{t-1}\varepsilon_t$ with $E(f_{t-1}^2\varepsilon_t^2) = 1 - \rho^2$ combined with Slutsky's theorem. ∎

Proof of Proposition 2.

In this proof, we only consider the case of *Bootstrap II* because the proof for *Bootstrap I* is similar but simpler. 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 $(1/TN)X^*X^{*\prime}\widetilde{F}^* = v_{NT}^*\widetilde{F}^*$ where v_{NT}^* is the largest eigenvalue of $(1/TN)X^*X^{*\prime}$. Let $\zeta_{st}^* = N^{-1}\sum_{i=1}^N e_{is}^* e_{it}^*$, $\eta_{st}^* = N^{-1}f_s^*\sum_{i=1}^N \lambda_i^* e_{it}^*$, and $\xi_{st}^* = N^{-1}f_t^*\sum_{i=1}^N \lambda_i^* e_{is}^* = \eta_{ts}^*$. The estimation error of the factor can be decomposed as

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

where $H_{NT}^* = (\widetilde{F}^{*\prime}F^*/T)(\Lambda^{*\prime}\Lambda^*/N)v_{NT}^{*-1}$. We denote $S_T^* = o_{P^*}(\alpha_T^{-1})$ if the bootstrap statistic S_T^* satisfies $P^*(\alpha_T|S_T^*| > \delta) = o_P(1)$ for any $\delta > 0$ as $\alpha_T \to \infty$. From Lemma B.1 in Gonçalves and Perron (2014), we have $v_{NT}^* = v + o_{P^*}(1)$, and $H_{NT}^{*2} - 1 = o_{P^*}(1)$.

The dominant term of the bootstrap estimation error can be decomposed as

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

$$= T^{-1/2}\sum_{t=2}^T \widetilde{f}_{t-1}^*\left\{\widetilde{f}_t^* - H_{NT}^* f_t^* - \widetilde{\rho}\left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right)\right\} + T^{-1/2}H_{NT}^*\sum_{t=2}^T \widetilde{f}_{t-1}^*\varepsilon_t^* + o_{P^*}(1)$$

$$= T^{-1/2}H_{NT}^{*2}\sum_{t=2}^T f_{t-1}^*\varepsilon_t^* - T^{-1/2}\widetilde{\rho}\sum_{t=2}^T \widetilde{f}_{t-1}^*\left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right)$$

$$+ T^{-1/2}\sum_{t=2}^T \widetilde{f}_{t-1}^*\left(\widetilde{f}_t^* - H_{NT}^* f_t^*\right) + T^{-1/2}H_{NT}^*\sum_{t=2}^T \left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right)\varepsilon_t^* + o_{P^*}(1).$$

The leading term can be written as

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

The last equality follows from the fact that $H_{NT}^{*2} - 1 = o_{P^*}(1)$. Analogous to the proofs of Proposition 1, we have (i) $T^{-1}\widetilde{\rho}\sum_{t=2}^T \widetilde{f}_{t-1}^*\left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right) =$

$2\rho v^{-2} N^{-1}\Gamma + o_{P*}(\delta_{NT}^{-2})$; (ii) $T^{-1}\tilde{\rho}\sum_{t=2}^{T}\widetilde{f}_{t-1}^{*}\left(\widetilde{f}_{t}^{*} - H_{NT}^{*}f_{t}^{*}\right) = \rho v^{-2} N^{-1}\Gamma + o_{P*}(\delta_{NT}^{-2})$; and (iii) $T^{-1}H_{NT}^{*}\sum_{t=2}^{T}(\widetilde{f}_{t-1}^{*} - H_{NT}^{*}f_{t-1}^{*})\varepsilon_{t}^{*} = o_{P*}(\delta_{NT}^{-2})$. Therefore,

$$\sqrt{T}\left(\widetilde{\rho}^{*} - \tilde{\rho}\right) = T^{-1/2}\sum_{t=2}^{T} f_{t-1}^{*}\varepsilon_{t}^{*} - c\rho\sigma_{\lambda}^{-4}\Gamma + o_{P*}(1).$$

We apply the bootstrap central limit theorem to the term $T^{-1/2}\sum_{t=2}^{T} f_{t-1}^{*}\varepsilon_{t}^{*}$. 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 $P^{*}(\sqrt{T}(\tilde{\rho}^{*} - \tilde{\rho}) \le x)$ approaches normal distribution function with mean $-c\rho\sigma_{\lambda}^{-4}\Gamma$ and variance $E^{*}(f_{t-1}^{*2}\varepsilon_{t}^{*2})$ under the bootstrap probability measure. In the residual bootstrap procedure for the AR(1) model, since f_{t}^{*} is generated by $f_{s}^{*} = \tilde{\rho}f_{s-1}^{*} + \varepsilon_{s}^{*}$ for $s = 0, \pm 1, \pm 2, ...$(see Bose, 1988, p. 1711), $E^{*}(f_{t-1}^{*2}) = \sum_{s=0}^{\infty}\tilde{\rho}^{2s}E^{*}(\varepsilon_{t}^{*2}) = (1 - \tilde{\rho}^{2})^{-1}E^{*}(\varepsilon_{t}^{*2})$. Because, $\tilde{\rho} \to^{P} \rho$ and $E^{*}(\varepsilon_{t}^{*2}) = T^{-1}\sum_{s=1}^{T}\widetilde{\varepsilon}_{t}^{2} \to^{P} \sigma_{\varepsilon}^{2} = 1 - \rho^{2}$, we have $E^{*}(f_{t-1}^{*2}\varepsilon_{t}^{*2}) = E^{*}(f_{t-1}^{*2})E^{*}(\varepsilon_{t}^{*2}) \to^{P} 1 - \rho^{2}$. Thus, we have $P^{*}(\sqrt{T}(\tilde{\rho}^{*} - \tilde{\rho}) \le x) - P(\sqrt{T}(\tilde{\rho} - \rho) \le x) \to^{P} 0$ for any x. By using Polya's theorem, we have the uniform convergence result. ∎

Proof of Proposition 3.

We show a sufficient condition $E^{*}[T\left(\widetilde{\rho}^{*} - \tilde{\rho}\right)^{2}] = O_{p}(1)$ for the uniform integrability of $\sqrt{T}\left(\widetilde{\rho}^{*} - \tilde{\rho}\right)$. From Lemma C.1 of Gonçalves and Perron (2014), with mutual independence of f_{t}, λ_{i}, and e_{it}, when $E|f_{t}|^{p} \le M$, $E|\lambda_{i}|^{p} \le M$, and $E|e_{it}|^{2p} \le M$ for some $p \ge 2$, we have (i) $T^{-1}\sum_{t=1}^{T}|\widetilde{f}_{t} - H_{NT}f_{t}|^{p} = O_{P}(1)$; (ii) $N^{-1}\sum_{i=1}^{N}|\tilde{\lambda}_{i} - H_{NT}^{-1}\lambda_{i}|^{p} = O_{P}(1)$; and (iii) $(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{e}_{it} = O_{P}(1)$. (i), (ii) and (iii) imply that $E^{*}(e_{it}^{*p}) = (NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{e}_{it}^{p} = O_{P}(1)$,

$$E^{*}|\lambda_{i}^{*}|^{p} = N^{-1}\sum_{i=1}^{N}|\tilde{\lambda}_{i}|^{p} \le 2^{p-1}N^{-1}(\sum_{i=1}^{N}|\tilde{\lambda}_{i} - H_{NT}^{-1}\lambda_{i}|^{p} + \sum_{i=1}^{N}|H_{NT}^{-1}\lambda_{i}|^{p}) = O_{P}(1),$$

and

$$
\begin{aligned}
E^{*}|\varepsilon_{t}^{*}|^{p} &= (T-1)^{-1}\sum_{t=2}^{T}|\widetilde{f}_{t} - \tilde{\rho}\widetilde{f}_{t-1}|^{p} \\
&= (T-1)^{-1}\sum_{t=2}^{T}|\widetilde{f}_{t} - H_{NT}f_{t} + H_{NT}f_{t} - \tilde{\rho}(\widetilde{f}_{t-1} - H_{NT}f_{t-1}) - \tilde{\rho}H_{NT}f_{t-1}|^{p} \\
&\le 4^{p-1}(T-1)^{-1} \\
&\quad \times \sum_{t=2}^{T}[|\widetilde{f}_{t} - H_{NT}f_{t}|^{p} + |H_{NT}f_{t}|^{p} + |\tilde{\rho}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})|^{p} + |\tilde{\rho}H_{NT}f_{t-1}|^{p}] \\
&= O_{P}(1),
\end{aligned}
$$

and

$$E^*|f_{t-1}^*|^p = E^*|\varepsilon_{t-1}^* + \tilde{\rho}\varepsilon_{-2}^* + ...|^p \leq (1 + |\tilde{\rho}|^p + |\tilde{\rho}|^{2p} + ...)E^*|\varepsilon_t^*|^p = O_P(1).$$

In addition, if $E|f_t|^{4p} \leq M$, $E|\lambda_i|^{4p} \leq M$, and $E|e_{it}|^{8p} \leq M$, we have

$$
\begin{aligned}
E^*|H_{NT}^*|^p &= E^* \left[|\tilde{F}^{*\prime}F^*/T|^p |\Lambda^{*\prime}\Lambda^*/N|^p |v_{NT}^*|^{-p} \right] \\
&\leq \epsilon^{-p} E^* \left[|\tilde{F}^{*\prime}F^*/T|^p |\Lambda^{*\prime}\Lambda^*/N|^p \right] \\
&\leq \epsilon^{-p} \left\{ E^* \left[(\tilde{F}^{*\prime}F^*/T)^{2p} \right] E^* \left[(\Lambda^{*\prime}\Lambda^*/N)^{2p} \right] \right\}^{1/2} = O_P(1)
\end{aligned}
$$

since

$$\tilde{F}^{*\prime}F^*/T = T^{-1}\sum_{t=2}^{T+1} \tilde{f}_{t-1}^* f_{t-1}^* \leq \left(T^{-1}\sum_{t=2}^{T+1} \tilde{f}_{t-1}^{*2} T^{-1}\sum_{t=2}^{T+1} f_{t-1}^{*2} \right)^{1/2} = \left(T^{-1}\sum_{t=2}^{T+1} f_{t-1}^{*2} \right)^{1/2},$$

$$E^* \left[(\tilde{F}^{*\prime}F^*/T)^{2p} \right] \leq E^* \left[\left(T^{-1}\sum_{t=2}^{T+1} f_{t-1}^{*2} \right)^p \right] \leq E^* \left[T^{-1}\sum_{t=2}^{T+1} f_{t-1}^{*2p} \right] = E^*(f_{t-1}^{*2p}) = O_P(1)$$

and

$$E^* \left[(\Lambda^{*\prime}\Lambda^*/N)^{2p} \right] = E^*[\left(N^{-1}\sum_{i=1}^{N} \lambda_i^{*2} \right)^{2p}] = E^*(\lambda_i^{*4p}) = O_P(1).$$

From the decomposition in the proof of Proposition 2, the second moment of the bootstrap estimator under the bootstrap measure is

$$E^*[T\left(\widetilde{\rho}^* - \widetilde{\rho}\right)^2]$$

$$= TE^*\{[\left(\sum_{t=2}^{T+1} \widetilde{f}_{t-1}^{*2}\right)^{-1} (\sum_{t=2}^{T} \widetilde{f}_{t-1}^* \widetilde{f}_t^* - \widetilde{\rho} \sum_{t=2}^{T+1} \widetilde{f}_{t-1}^{*2})]^2\}$$

$$= T^{-1}E^*\{[\sum_{t=2}^{T} \widetilde{f}_{t-1}^* \left(\widetilde{f}_t^* - \widetilde{\rho}\widetilde{f}_{t-1}^*\right)]^2\} + o_P(1)$$

$$= T^{-1}E^*\left\{\left[\sum_{t=2}^{T} \widetilde{f}_{t-1}^* \left\{\widetilde{f}_t^* - H_{NT}^* f_t^* - \widetilde{\rho}\left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right)\right\} + H_{NT}^* \widetilde{f}_{t-1}^* \varepsilon_t^*\right]^2\right\} + o_P(1)$$

$$= T^{-1}E^*\left\{\left[H_{NT}^{*2} \sum_{t=2}^{T} f_{t-1}^* \varepsilon_t^* - \widetilde{\rho} \sum_{t=2}^{T} \widetilde{f}_{t-1}^* \left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right)\right.\right.$$

$$\left.\left. + \sum_{t=2}^{T} \widetilde{f}_{t-1}^* \left(\widetilde{f}_t^* - H_{NT}^* f_t^*\right) + H_{NT}^* \sum_{t=2}^{T} \left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right) \varepsilon_t^*\right]^2\right\} + o_P(1)$$

$$\leq 4T^{-1}E^*\left[H_{NT}^{*4} \sum_{t=2}^{T} f_{t-1}^{*2} \varepsilon_t^{*2} - \widetilde{\rho}^2 \sum_{t=2}^{T} \widetilde{f}_{t-1}^{*2} \left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right)^2\right.$$

$$\left. + \sum_{t=2}^{T} \widetilde{f}_{t-1}^{*2} \left(\widetilde{f}_t^* - H_{NT}^* f_t^*\right)^2 + H_{NT}^{*2} \sum_{t=2}^{T} \left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right)^2 \varepsilon_t^{*2}\right] + o_P(1).$$

Combining the moment conditions introduced before, we can show that each term in this expansion is $O_P(1)$. For example, the leading term is bounded in probability because

$$T^{-1}E^* H_{NT}^{*4} \sum_{t=2}^{T} f_{t-1}^{*2} \varepsilon_t^{*2}$$

$$\leq \left[E^* H_{NT}^{*8} E^* \left(T^{-1} \sum_{t=2}^{T} f_{t-1}^{*2} \varepsilon_t^{*2}\right)^2\right]^{1/2}$$

$$\leq \left[E^* H_{NT}^{*8} E^* \left(f_{t-1}^{*4} \varepsilon_t^{*4}\right)\right]^{1/2} = O_p(1).$$

The last equality follows from $E^* \left(f_{t-1}^{*4} \varepsilon_t^{*4}\right) \leq \left[E^* \left(f_{t-1}^{*8}\right)E^*(\varepsilon_t^{*8})\right]^{1/2} = O_P(1)$ and $E^* H_{NT}^{*8} = O_P(1)$ under $E|f_t|^{32} \leq M$, $E|\lambda_i|^{32} \leq M$, and $E|e_{it}|^{64} \leq M$. ∎

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.

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.

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

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.

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 (2014). "Bootstrapping factor-augmented regression models." Journal of Econometrics 182(1), 156-173.

Gonçalves, S. and H. White (2005). "Bootstrap standard error estimates for linear regression." Journal of the American Statistical Association 100, 970-979.

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	0.90
	200	0.89	0.90	0.90	0.90

Note: Mean values of the OLS estimator ($\hat{\rho}$), the Kendall-type bias-corrected estimator ($\hat{\rho}_{KBC}$) and the bootstrap bias-corrected estimator ($\hat{\rho}_{BC}$) and coverage rates of the asymptotic confidence interval (5) in 10,000 replications.

Table 2: Two-Step AR Estimation

ρ	T	c	$\tilde{\rho}$ $1/\sigma_e^2=0.5$	0.75	1	1.5	2	Coverage Rate $1/\sigma_e^2=0.5$	0.75	1	1.5	2
			(A) No cross-sectional correlation									
0.5	100	0.5	0.42	0.43	0.44	0.45	0.46	0.77	0.82	0.84	0.86	0.86
		1	0.36	0.39	0.41	0.42	0.44	0.58	0.69	0.73	0.79	0.82
		1.5	0.32	0.36	0.38	0.40	0.41	0.44	0.56	0.64	0.73	0.76
	200	0.5	0.45	0.46	0.46	0.47	0.47	0.79	0.83	0.85	0.87	0.88
		1	0.41	0.43	0.44	0.45	0.46	0.60	0.70	0.75	0.81	0.83
		1.5	0.36	0.40	0.41	0.43	0.44	0.39	0.53	0.61	0.71	0.77
0.9	100	0.5	0.73	0.77	0.79	0.81	0.82	0.25	0.40	0.47	0.58	0.63
		1	0.62	0.68	0.71	0.75	0.77	0.07	0.13	0.22	0.32	0.41
		1.5	0.54	0.61	0.65	0.70	0.73	0.03	0.07	0.11	0.19	0.27
	200	0.5	0.80	0.82	0.83	0.85	0.85	0.27	0.42	0.51	0.62	0.70
		1	0.72	0.76	0.78	0.81	0.82	0.05	0.12	0.21	0.35	0.43
		1.5	0.65	0.70	0.73	0.77	0.79	0.01	0.04	0.08	0.16	0.25
			(B) Cross-sectional correlation									
0.5	100	0.5	0.40	0.42	0.44	0.45	0.45	0.71	0.79	0.81	0.84	0.86
		1	0.29	0.35	0.38	0.41	0.42	0.39	0.55	0.65	0.73	0.78
		1.5	0.21	0.28	0.32	0.37	0.39	0.24	0.39	0.49	0.61	0.69
	200	0.5	0.44	0.45	0.46	0.47	0.48	0.75	0.82	0.84	0.86	0.87
		1	0.37	0.41	0.43	0.45	0.46	0.44	0.61	0.69	0.78	0.82
		1.5	0.28	0.34	0.38	0.41	0.43	0.22	0.38	0.49	0.64	0.70
0.9	100	0.5	0.67	0.74	0.77	0.80	0.81	0.19	0.33	0.43	0.54	0.61
		1	0.44	0.56	0.63	0.70	0.74	0.05	0.11	0.17	0.29	0.38
		1.5	0.32	0.44	0.52	0.62	0.67	0.02	0.06	0.10	0.18	0.24
	200	0.5	0.78	0.81	0.83	0.84	0.85	0.22	0.37	0.48	0.61	0.66
		1	0.63	0.71	0.76	0.80	0.81	0.05	0.11	0.18	0.31	0.40
		1.5	0.46	0.59	0.65	0.73	0.76	0.01	0.05	0.08	0.16	0.23

Note: Mean values of the two-step estimator ($\tilde{\rho}$) and coverage rates of the asymptotic confidence interval (9) in 10,000 replications. $1/\sigma_e^2$ is the signal-to-noise ratio.

Table 3: Bootstrap Bias Corrections

ρ	c		$1/\sigma_e^2 =0.5$	0.75	1	1.5	2	$1/\sigma_e^2 =0.5$	0.75	1	1.5	2
					$T=100$					$T=200$		
			(A) No cross-sectional correlation									
0.5	0.5	bias	-0.08	-0.07	-0.06	-0.05	-0.05	-0.05	-0.04	-0.04	-0.03	-0.02
		asy bias	-0.05	-0.03	-0.03	-0.02	-0.01	-0.04	-0.02	-0.02	-0.01	-0.01
		bias I*	-0.05	-0.04	-0.03	-0.02	-0.02	-0.04	-0.03	-0.02	-0.01	-0.01
		bias II*	-0.07	-0.06	-0.05	-0.05	-0.04	-0.05	-0.04	-0.03	-0.03	-0.03
	1	bias	-0.13	-0.11	-0.10	-0.07	-0.07	-0.09	-0.08	-0.06	-0.04	-0.04
		asy bias	-0.10	-0.07	-0.05	-0.03	-0.03	-0.07	-0.05	-0.04	-0.02	-0.02
		bias I*	-0.08	-0.07	-0.06	-0.05	-0.04	-0.06	-0.05	-0.04	-0.04	-0.03
		bias II*	-0.09	-0.09	-0.08	-0.07	-0.07	-0.07	-0.06	-0.05	-0.05	-0.04
	1.5	bias	-0.18	-0.14	-0.12	-0.10	-0.08	-0.13	-0.10	-0.09	-0.07	-0.06
		asy bias	-0.15	-0.10	-0.07	-0.05	-0.04	-0.11	-0.07	-0.05	-0.04	-0.03
		bias I*	-0.09	-0.08	-0.08	-0.07	-0.06	-0.08	-0.07	-0.07	-0.06	-0.05
		bias II*	-0.10	-0.10	-0.10	-0.09	-0.08	-0.09	-0.08	-0.08	-0.07	-0.06
0.9	0.5	bias	-0.17	-0.13	-0.11	-0.10	-0.09	-0.10	-0.08	-0.07	-0.06	-0.05
		asy bias	-0.09	-0.06	-0.04	-0.03	-0.02	-0.06	-0.04	-0.03	-0.02	-0.02
		bias I*	-0.09	-0.08	-0.07	-0.05	-0.04	-0.07	-0.05	-0.05	-0.03	-0.03
		bias II*	-0.13	-0.12	-0.10	-0.09	-0.08	-0.09	-0.07	-0.07	-0.05	-0.05
	1	bias	-0.28	-0.22	-0.19	-0.15	-0.13	-0.18	-0.14	-0.12	-0.09	-0.08
		asy bias	-0.18	-0.12	-0.09	-0.06	-0.05	-0.13	-0.08	-0.06	-0.04	-0.03
		bias I*	-0.14	-0.13	-0.11	-0.10	-0.09	-0.12	-0.10	-0.09	-0.07	-0.06
		bias II*	-0.17	-0.16	-0.15	-0.14	-0.13	-0.13	-0.12	-0.10	-0.09	-0.08
	1.5	bias	-0.36	-0.29	-0.24	-0.20	-0.17	-0.26	-0.20	-0.17	-0.13	-0.11
		asy bias	-0.27	-0.18	-0.14	-0.09	-0.07	-0.19	-0.13	-0.10	-0.06	-0.05
		bias I*	-0.15	-0.15	-0.15	-0.14	-0.12	-0.15	-0.14	-0.12	-0.11	-0.09
		bias II*	-0.18	-0.18	-0.18	-0.17	-0.16	-0.16	-0.15	-0.14	-0.13	-0.11

Table 3 (continued)

ρ	c		$1/\sigma_e^2 =0.5$	0.75	1	1.5	2	$1/\sigma_e^2 =0.5$	0.75	1	1.5	2
					$T=100$					$T=200$		
					(B) Cross-sectional correlation							
0.5	0.5	bias	-0.10	-0.08	-0.07	-0.05	-0.05	-0.06	-0.04	-0.04	-0.03	-0.02
		asy bias	-0.05	-0.03	-0.03	-0.02	-0.01	-0.04	-0.02	-0.02	-0.01	-0.01
		bias I*	-0.05	-0.04	-0.03	-0.02	-0.02	-0.04	-0.03	-0.02	-0.02	-0.01
		bias II*	-0.07	-0.06	-0.05	-0.05	-0.04	-0.05	-0.04	-0.03	-0.03	-0.03
	1	bias	-0.21	-0.15	-0.12	-0.09	-0.08	-0.14	-0.09	-0.07	-0.05	-0.04
		asy bias	-0.10	-0.07	-0.05	-0.03	-0.03	-0.07	-0.05	-0.04	-0.02	-0.02
		bias I*	-0.06	-0.07	-0.06	-0.05	-0.05	-0.06	-0.05	-0.04	-0.04	-0.03
		bias II*	-0.08	-0.08	-0.08	-0.07	-0.07	-0.07	-0.06	-0.05	-0.05	-0.04
	1.5	bias	-0.29	-0.22	-0.18	-0.14	-0.10	-0.23	-0.16	-0.13	-0.09	-0.07
		asy bias	-0.15	-0.10	-0.07	-0.05	-0.04	-0.11	-0.07	-0.05	-0.04	-0.03
		bias I*	-0.07	-0.08	-0.08	-0.07	-0.07	-0.07	-0.07	-0.07	-0.06	-0.05
		bias II*	-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	-0.22	-0.17	-0.13	-0.11	-0.09	-0.12	-0.08	-0.07	-0.06	-0.05
		asy bias	-0.09	-0.06	-0.04	-0.03	-0.02	-0.06	-0.04	-0.03	-0.02	-0.02
		bias I*	-0.09	-0.08	-0.07	-0.05	-0.04	-0.07	-0.05	-0.05	-0.03	-0.03
		bias II*	-0.12	-0.11	-0.10	-0.09	-0.08	-0.09	-0.07	-0.07	-0.05	-0.05
	1	bias	-0.45	-0.35	-0.26	-0.18	-0.16	-0.27	0.18	-0.14	-0.10	-0.08
		asy bias	-0.18	-0.12	-0.09	-0.06	-0.04	-0.13	-0.08	-0.06	-0.04	-0.03
		bias I*	-0.11	-0.11	-0.11	-0.10	-0.09	-0.11	-0.10	-0.09	-0.07	-0.06
		bias II*	-0.13	-0.14	-0.14	-0.13	-0.13	-0.12	-0.11	-0.10	-0.09	-0.08
	1.5	bias	-0.57	-0.45	-0.37	-0.28	-0.23	-0.45	-0.31	-0.25	-0.18	-0.14
		asy bias	-0.27	-0.18	-0.14	-0.09	-0.07	-0.19	-0.13	-0.10	-0.06	-0.05
		bias I*	-0.12	-0.13	-0.14	-0.13	-0.12	-0.12	-0.13	-0.12	-0.11	-0.09
		bias II*	-0.13	-0.15	-0.16	-0.16	-0.16	-0.13	-0.14	-0.14	-0.13	-0.11

Note: The actual bias (bias), bootstrap bias estimator based on *Bootstrap I* (bias I*) and bootstrap bias estimator based on *Bootstrap II* (bias II*) are mean values in 10,000 replications. The asymptotic bias (asy bias) is $-T^{-1/2}c\rho\sigma_\lambda^{-4}\Gamma$. $1/\sigma_e^2$ is the signal-to-noise ratio.

Table 4: Coverage Rate of Bootstrap Confidence Intervals

ρ	c		$T=100$					$T=200$				
			$1/\sigma_e^2$ =0.5	0.75	1	1.5	2	$1/\sigma_e^2$ =0.5	0.75	1	1.5	2
			(A) No cross-sectional correlation									
0.5	0.5	Bc	0.85	0.86	0.86	0.86	0.87	0.86	0.87	0.87	0.87	0.88
		Per	0.87	0.87	0.88	0.87	0.88	0.88	0.88	0.88	0.89	0.89
		Per-t	0.86	0.87	0.87	0.87	0.88	0.87	0.88	0.87	0.88	0.89
	1	Bc	0.77	0.80	0.83	0.83	0.85	0.81	0.83	0.84	0.87	0.86
		Per	0.80	0.84	0.86	0.86	0.87	0.86	0.87	0.87	0.88	0.87
		Per-t	0.79	0.83	0.85	0.85	0.87	0.84	0.85	0.86	0.87	0.87
	1.5	Bc	0.68	0.75	0.77	0.80	0.82	0.72	0.79	0.81	0.82	0.85
		Per	0.73	0.80	0.81	0.84	0.85	0.78	0.83	0.85	0.85	0.87
		Per-t	0.72	0.79	0.80	0.84	0.84	0.75	0.82	0.84	0.84	0.87
0.9	0.5	Bc	0.78	0.82	0.83	0.84	0.84	0.84	0.87	0.88	0.89	0.89
		Per	0.90	0.93	0.93	0.93	0.93	0.95	0.95	0.95	0.94	0.93
		Per-t	0.80	0.86	0.87	0.88	0.88	0.86	0.90	0.90	0.90	0.89
	1	Bc	0.60	0.70	0.75	0.79	0.80	0.70	0.80	0.83	0.84	0.86
		Per	0.74	0.84	0.88	0.91	0.93	0.87	0.93	0.94	0.95	0.95
		Per-t	0.62	0.73	0.79	0.84	0.87	0.72	0.82	0.86	0.89	0.89
	1.5	Bc	0.45	0.60	0.66	0.73	0.76	0.50	0.64	0.71	0.76	0.80
		Per	0.60	0.74	0.79	0.87	0.88	0.70	0.84	0.90	0.92	0.93
		Per-t	0.48	0.63	0.70	0.79	0.82	0.53	0.70	0.79	0.84	0.88
			(B) Cross-sectional correlation									
0.5	0.5	Bc	0.81	0.84	0.86	0.86	0.87	0.85	0.86	0.87	0.88	0.88
		Per	0.83	0.86	0.87	0.87	0.88	0.87	0.88	0.88	0.89	0.88
		Per-t	0.82	0.85	0.87	0.87	0.88	0.86	0.86	0.87	0.89	0.89
	1	Bc	0.58	0.71	0.78	0.82	0.84	0.68	0.79	0.83	0.86	0.87
		Per	0.62	0.75	0.81	0.84	0.86	0.71	0.83	0.86	0.88	0.87
		Per-t	0.61	0.73	0.80	0.84	0.86	0.69	0.81	0.84	0.87	0.87
	1.5	Bc	0.45	0.60	0.67	0.75	0.78	0.48	0.64	0.73	0.78	0.82
		Per	0.48	0.64	0.70	0.78	0.81	0.52	0.69	0.78	0.83	0.85
		Per-t	0.47	0.63	0.69	0.77	0.81	0.51	0.66	0.76	0.81	0.84
0.9	0.5	Bc	0.62	0.73	0.78	0.81	0.83	0.75	0.83	0.85	0.87	0.88
		Per	0.73	0.85	0.89	0.91	0.92	0.87	0.93	0.93	0.93	0.93
		Per-t	0.62	0.76	0.81	0.85	0.86	0.73	0.82	0.87	0.88	0.89
	1	Bc	0.32	0.48	0.59	0.70	0.74	0.43	0.63	0.71	0.80	0.82
		Per	0.42	0.60	0.73	0.83	0.87	0.57	0.77	0.84	0.91	0.93
		Per-t	0.33	0.50	0.63	0.75	0.79	0.44	0.63	0.72	0.82	0.85
	1.5	Bc	0.21	0.36	0.46	0.60	0.65	0.24	0.42	0.56	0.65	0.71
		Per	0.29	0.46	0.58	0.72	0.78	0.36	0.57	0.71	0.81	0.87
		Per-t	0.23	0.39	0.50	0.65	0.71	0.26	0.44	0.60	0.71	0.78

Note: Coverage rates of three nominal 90% confidence intervals in 10,000 replications. Bc denotes the bootstrap bias-corrected asymptotic confidence interval (10), Per denotes the percentile bootstrap confidence interval (11) and Per-t denotes the percentile-t equal-tailed bootstrap confidence interval (12). $1/\sigma_e^2$ is the signal-to-noise ratio.

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

Series	$\tilde{\rho}$	Asymptotic Confidence interval	$\tilde{\rho}_{BC}$	Bootstrap confidence intervals		
				Bc	Per	Per-t
		(A) Full sample ($N = 159$)				
1	0.66	(0.60, 0.71)	0.69	(0.64, 0.75)	(0.64, 0.76)	(0.64, 0.75)
		(B) Long subsample ($N = 53$)				
1	0.65	(0.60, 0.71)	0.74	(0.69, 0.80)	(0.68, 0.82)	(0.68, 0.80)
2	0.58	(0.52, 0.64)	0.66	(0.60, 0.72)	(0.59, 0.74)	(0.59, 0.72)
3	0.68	(0.63, 0.73)	0.78	(0.72, 0.83)	(0.71, 0.86)	(0.71, 0.83)
average	0.64	(0.58, 0.69)	0.73	(0.67, 0.79)	(0.66, 0.80)	(066., 0.78)
		(C) Short subsample ($N = 31$)				
1	0.57	(0.51, 0.63)	0.75	(0.69, 0.81)	(0.66, 0.84)	(0.65, 0.80)
2	0.83	(0.79, 0.87)	0.95	(0.91, 1.00)	(0.88, 1.06)	(0.88, 0.99)
3	0.63	(0.58, 0.69)	0.75	(0.69, 0.80)	(0.67, 0.83)	(0.67, 0.80)
4	0.55	(0.49, 0.61)	0.65	(0.58, 0.71)	(0.57, 0.73)	(0.57, 0.71)
5	0.54	(0.48, 0.60)	0.67	(0.61, 0.74)	(0.59, 0.77)	(0.59, 0.75)
average	0.62	(0.57, 0.68)	0.75	(0.70, 0.81)	(0.67, 0.84)	(0.67, 0.81)

Note: The sample period is from 1959:3 to 1998:12 ($T = 478$). $c = \sqrt{T}/N$ is 0.14, 0.41 and 0.71, respectively, for series A, B and C. The first confidence interval next to $\tilde{\rho}$ is the 90% asymptotic confidence interval (9). For the boostrap confidence intervals, Bc denotes the 90% bootstrap bias-corrected asymptotic confidence interval (10), Per denotes the 90% percentile interval (11) and Per-t denotes the 90% percentile-t equal-tailed interval (12).
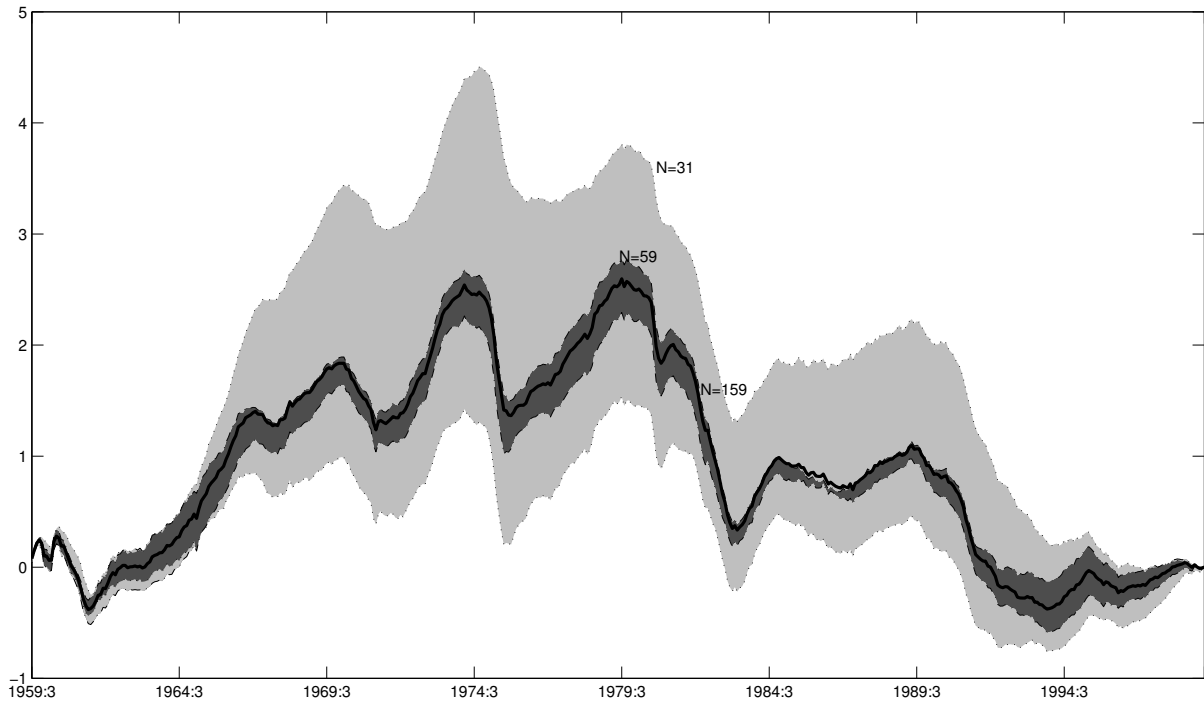
Figure 1: The US Diffusion Index