Resampling uncertainty of Principal Components Factors

Javier de Vicente and Esther Ruiz *

Department of Statistics, Universidad Carlos III de Madrid

June 2017

Abstract

In the context of Dynamic Factor Models (DFM), one of the most popular procedures for factor extraction is Principal Components (PC). Measuring the uncertainty associated to PC factor estimates should be part of interpreting these estimates. Although, the asymptotic distribution of factors is known, for the sample sizes and cross-sectional dimensions usually encountered in practice, it is not an appropriate approximation to the finite sample one. The main problem is that it does not take into account parameter uncertainty. Alternatively, several bootstrap procedures have been proposed in DFM with goals related to inference. We show that these procedures are not appropriate to measure the uncertainty of PC factor estimates and propose a resampling procedure designed with this purpose. The asymptotic and finite sample properties of the proposed procedure are analyzed and compared with those of the asymptotic and alternative extant bootstrap procedures. The results are empirically illustrated obtaining confidence intervals of the underlying factor in a system of Spanish macroeconomic variables and in a system of house prices of advanced and emerging markets.

Keywords: Resampling Procedures, Dynamic Factor Models, Unobserved Components.

*Financial Support from the Spanish Government contract grant ECO2015-70331-C2-2-R (MINECO/FEDER) is gratefully acknowledged. We are also grateful to participants at Computational and Financial Econometrics meeting (CFE 2016, Seville) and at the Time Series Workshop meeting (2017, Zaragoza) and at seminars at the Joint Research Center (JRC) of the European Commision at Ispra and at the Joint Statistics and Econometrics Workshop at London School of Economics. Any remaining errors are obviously our responsability.
1 Introduction

Currently, large systems of macroeconomic variables are easily accessible, and the consequent extraction of the underlying common factors is an important issue for econometricians and policy decision makers as these factors often have interpretations of interest. The latent factors are useful instruments for a wide range of applications: i) to represent economic cycles, trends and structural shocks; see Arouba et al. (2012), Camacho et al. (2015) and Breitung and Eickmeier (2016) for some recent references; ii) to serve as instrumental variables; see Favero et al. (2005), Bai and Ng (2010) and Kapetanios and Marcellino (2010); iii) as regressors for the construction of Factor-Augmented Vector Autorregresive models (FAVAR) or Factor-augmented Error Correction models (FECM); see, for example, Bernanke et al. (2005), Banerjee et al. (2014), Abbate et al. (2016) and Bai et al. (2016) or iv) in the context of factor-augmented predictive regressions; see, for example, Stock and Watson (2006), Ludvigson and Ng (2007, 2009), Ando and Tsay (2014), Bräuning and Koopman (2014) and Neely et al. (2014)1.

In this context, the Dynamic Factor Models (DFMs), originally introduced by Geweke (1977) and Sargent and Sims (1977), have received a lot of attention; see Breitung and Eickmeier (2006), Bai and Ng (2008), Stock and Watson (2011), Breitung and Choi (2013) and Bai and Wang (2016) for excellent surveys on DFMs. The main goal of DFMs is to explain the dynamics of the system using a reduced number of unobservable common factors, which determine the dynamics of the macroeconomic variables. Several methods have been proposed the literature for factor extraction in DFM. The most popular procedures for large data sets are still based on Principal Components (PC) techniques; see, for example, Ludvigson and Ng (2007, 2009, 2010), Wang (2009), Foerster et al. (2011), Ando and Tsay (2014), Gonçalves and Perron (2014), Neely et al. (2014), Djogbenov et al. (2015), Fossati (2016) and Jackson et al. (2016) for recent references. The factors correspond to the first few principal components (arranged in decreasing order by their eigenvalues) of the entire system of variables; see, for example, Stock and Watson (2002) for an excellent discussion on PC factor extraction. The popularity of PC factor extraction relies on the fact that it is computationally simple and allows dealing with very large systems. However, it is crucial to obtain not only accurate point estimates of the latent factors, but also

---

1 Several authors have look at the relationship between DFM and the economic-based Dynamic Stochastic General Equilibrium (DSGE) models. For example, Foerster et al. (2011) construct a multi-sector DSGE model where sectoral productivity shocks have a factor model structure and show that the model solution has an approximate factor model representation. Alternatively, Onatski and Ruge-Murcia (2013) use a fully specified DSGE model as a laboratory to understand the practical benefits and limitations of factor analysis techniques on economic data; see also Giannone et al. (2006) who examine the relative performance of VAR and DFM approximation to DSGE model and Wang (2009) who compares both models in an out-of-sample forecasting exercise.
of their associated uncertainty. For example, Bai (2003) remarks the importance of constructing confidence intervals of the extracted factors in empirical applications in which they represent economic indices. Boivin and Ng (2006) also pay attention to the uncertainty of factor estimates in the context of predictive regressions while Bai and Ng (2006) argue about the importance of measuring correctly the uncertainty of the factor extraction in FAVAR models. More recently, Jackson et al. (2016) argue that measures of factor uncertainty should always accompany applied work in order to establish the statistical legitimacy of the results.

The asymptotic distribution of the factors extracted using PC is derived by Bai (2003) assuming weak dependence in the idiosyncratic term while Bai and Ng (2006) propose three different estimators of the asymptotic covariance matrix of the factors depending on the structure of the errors. More recently, Bai and Ng (2013) derive the limiting distribution of the factors and its corresponding covariance matrix estimation for different identification restrictions. However, results on the performance of the asymptotic distribution to approximate the finite sample distribution of the estimated factors are scarce. As far as we are concerned, only Poncela and Ruiz (2016) show that PC intervals based on the asymptotic distribution could underestimate the uncertainty of the extracted factors. The poor performance of the asymptotic distribution could be attributed to the fact that parameter uncertainty is not considered. Alternatively, the finite sample distribution of the estimated factors can be obtained using resampling procedures that incorporate parameter uncertainty. Several authors propose using bootstrap in the context of DFM with other objectives than obtaining the distribution of the underlying factors. For example, Yamamoto (2016) obtains bootstrap bands for impulse response functions (IRF) in the context of FAVAR models. In empirical applications, Barigozzi et al. (2016) and Foroni et al. (2014) also implement resampling methods to construct confidence bands for IRF in the context of FAVAR models. Ludvigson and Ng (2007, 2009 and 2010), Gospodinov and Ng (2013), Gonçalves and Perron (2014), Djogbenou et al. (2015), Jackson et al. (2016) and Gonçalves et al. (2017) implement bootstrap procedures in the context of the parameters of factor-augmented predictive regression models. Shin and Guo (2015) also propose using bootstrap to test about the autoregressive parameter governing the dependence of the latent factor. Furthermore, Alonso et al. (2008) and Alonso et al. (2011) use bootstrap procedures for constructing forecasting intervals for population projections and electricity prices respectively. However, none of these papers analyze the performance of the bootstrap procedures implemented when they are used to obtain confidence bands for the extracted factors.

This paper has three main contributions. First, we provide extensive Monte Carlo experiments
in order to assess the conditions under which the asymptotic distribution of the factors extracted using PC is a good approximation of the finite sample distribution. In concordance with the results in Poncela and Ruiz (2016), we show that, in a wide range of scenarios, the asymptotic confidence intervals of the estimated factors are unrealistically tiny. The second contribution is to analyze the performance of the available bootstrap methods mentioned above when implemented to obtain confidence bands of the PC extracted factors. We show that these methods either obtain the marginal distribution of the factors and, consequently, the corresponding intervals are too wide as to be informative or they do not incorporate parameter uncertainty. The third and main contribution of this paper is to propose and theoretically justify a new resampling procedure designed to construct conditional confidence bands for PC that takes into account the dynamic dependence in the system. The asymptotic and finite sample performance of the proposed procedure is analyzed.

The rest of the paper is organized as follows. Section 2 describes the PC factor extraction procedure and its asymptotic distribution. Monte Carlo experiments are carried out to assess the adequacy of the asymptotic distribution to approximate the finite sample distribution of the factors. Section 3 describes available bootstrap procedures proposed for DFM and analyzes their finite sample performance. In Section 4, the new resampling procedure is proposed and its asymptotic validity is established. We also carry out Monte Carlos experiments to analyze its finite sample performance. Section 5 illustrates the results with an empirical application. Finally, Section 6 concludes.

2 Factor extraction

In this section, we describe the DFM considered in this paper and introduce notation. We also describe the asymptotic properties of PC factor estimates. Finally, we carry out Monte Carlo experiments to assess the adequacy of the asymptotic distribution to approximate the finite sample distribution and to construct confidence intervals of the extracted factors.

2.1 The Dynamic Factor Model

We consider the following stationary DFM in which the latent factors and the idiosyncratic components are VAR(1) processes

\[ Y_t = PF_t + \varepsilon_t, \quad (1) \]
\[ F_t = \Phi F_{t-1} + \eta_t, \quad (2) \]
\[ \varepsilon_t = \Gamma \varepsilon_{t-1} + a_t \quad (3) \]
where $Y_t = (y_{1t}, ..., y_{Nt})'$ is the $N \times 1$ vector of observed variables at time $t$ for $t = 1, ..., T$, $P$ is the $N \times r$ matrix of factor loadings, $F_t = (f_{1t}, ..., f_{rt})'$ is the $r \times 1$ matrix of unobservable factors and $\varepsilon_t = (\varepsilon_{1t}, ..., \varepsilon_{Nt})'$ is the $N \times 1$ vector of idiosyncratic noises. The disturbances $\eta_t = (\eta_{1t}, ..., \eta_{rt})'$ and $a_t = (a_{1t}, ..., a_{Nt})'$ are mutually independent Gaussian white noise vectors with finite covariance matrices $\Sigma_\eta$ and $\Sigma_a$, respectively. The matrices $\Phi$ and $\Gamma$ are diagonal with their parameters restricted so that $Y_t$ is stationary. The number of factors, $r$, is assumed to be known and fixed as the cross-sectional and temporal dimensions, $N$ and $T$, respectively, grow.

The DFM in equations (1) to (3) has been frequently considered in the related literature; see, for example, Jungbacker and Koopman (2015), Alvarez et al. (2016) and Jackson et al. (2016) for some recent references.

To uniquely fix the $T \times r$ matrix of factors, $F = (F_1, ..., F_T)'$, and $P$ (up to a column sign change), we assume that $\frac{1}{T}F'F = I_r$ and $P'P$ is a diagonal matrix with its main diagonal values ordered in decreasing order; see Bai and Ng (2013) for an extensive discussion on identification issues.

The matrix representation of model (1) is

$$Y = FP' + \varepsilon,$$

where $Y$ and $\varepsilon$ are $T \times N$ matrices.

Note that, according to (2) and assuming that $E(F_tF_t') = I_r$, the marginal distribution of the factors is given by

$$F_t \sim N(0, I_r),$$

and, consequently, one can always construct confidence intervals for the unobserved factors using this distribution. However, the corresponding confidence intervals will be uninformative. Confidence intervals with less uncertainty can be constructed conditional on the available information $\{Y_t\}_{t=1}^T$. Also, note that the marginal Mean Square Error (MSE) is not appropriate when the intervals are not centered at the marginal mean (zero) but in a estimation that uses information contained in $\{Y_t\}_{t=1}^T$.

### 2.2 Principal Components Factor Extraction

Although it is well known that PC is not efficient, it is still among the most popular factor extraction procedures due to its simplicity and low computational burden when dealing with very large systems of macroeconomic or financial variables. The method of PC minimizes:

$$V(r) = \min_{P,F} (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T (Y_{it} - P_i'F_t)^2.$$
Mechanically speaking, the factor estimates can be obtained in one of two ways. The first solution is obtained concentrating out $P$. Using the normalization, $F'F/T = I_r$, the estimated factors, $\tilde{f}_i$, is $\sqrt{T}$ times the eigenvectors corresponding to the $r$ largest eigenvalues of $YY'$ and $\tilde{P}' = \frac{1}{N} \tilde{f}'Y$, with $\tilde{P}'\tilde{P}$ being diagonal. The second solution is obtained after concentrating out $F$. Then, $\tilde{F}$ is $\sqrt{N}$ times the eigenvectors corresponding to the $r$ largest eigenvalues of $Y'Y$. Using the normalization $\frac{1}{N} P'P = I_r$, yields

$$\tilde{f} = \frac{1}{N} YP.$$  \hfill (7)

Note that the matrices $YY'$ and $Y'Y$ have identical nonzero eigenvalues and, consequently,

$$\frac{1}{T} \tilde{f}'\tilde{f} = \frac{1}{N} \tilde{P}'\tilde{P} = \tilde{V},$$  \hfill (8)

where $\tilde{V}$ is the $r \times r$ diagonal matrix consisting of the first $r$ eigenvalues of the matrix $\frac{1}{N} Y Y'$ arranged in decreasing order. Then, $\tilde{f} = \tilde{f}V^{1/2}$ and $\tilde{P} = \tilde{P}V^{1/2}$; see Bai and Ng (2008). Let $\hat{f} = f\left(\frac{1}{T} \tilde{f}'\tilde{f}\right)^{1/2} = \tilde{f}V^{1/2} = \tilde{f}\hat{V}$. Note that $\hat{f}_t$ and $\tilde{f}_t$ are equivalent in the sense that knowing one will lead to the other. From the results above, we can see that

$$\hat{f} = \tilde{f}V^{1/2} = \frac{1}{N} YP\hat{V}^{1/2} = \frac{1}{N} Y\hat{P},$$  \hfill (9)

and, consequently,

$$\hat{f}_t = \frac{1}{N} \hat{P}'Y_t.$$  \hfill (10)

The interest in expression (10) lies in the fact that the factor estimates are expressed as a linear filter of the original observations as in (7) while, simultaneously, they satisfy the restriction $\frac{1}{T} \tilde{f}'\tilde{f} = I_r$.

It is well known that the extracted factors, $\hat{f}_t$, estimate only a rotation of the true factors, $HF_t$, where $H = \left(\frac{P'P}{N}\right)$. As discussed above, given that the filter used to estimate the factors is based on $\{Y_t\}_{t=1}^T$, the MSE should also be computed conditional on this information. The MSE of the estimated factors can be obtained as follows:

$$E_T \left[ \left( \hat{f}_t - HF_t \right) \left( \hat{f}_t - HF_t \right)' \right] =$$

$$E_T \left[ \left( \hat{f}_t - f_t \right) \left( \hat{f}_t - f_t \right)' \right] + E_T \left[ \left( f_t - HF_t \right) \left( f_t - HF_t \right)' \right] + 2E_T \left[ \left( \hat{f}_t - f_t \right) \left( f_t - HF_t \right)' \right],$$  \hfill (11)

where $f_t$ is the factor extracted if the loadings were known, i.e.

$$f_t = \frac{1}{N} P'Y_t,$$  \hfill (12)

and the $T$ below the expectation means that it is conditional on $\{Y_t\}_{t=1}^T$. Note that the MSE of $\hat{f}_t$ in expression (11) decomposes the total MSE in the uncertainty due to parameter estimation,
the disturbance uncertainty and the covariance between them. First, using (10) and (12), we can obtain the following expression of the MSE attributed to parameter uncertainty

$$E_T\left[ (\tilde{f}_t - f_t)(\tilde{f}_t - f_t)' \right] = \frac{1}{N^2 T} E_T\left[ (\tilde{P} - P)' \right] Y_t Y_t' \left( \tilde{P} - P \right).$$

(13)

On the other hand, from equation (1) we can obtain the following expression for the rotated true factors

$$HF_t = \frac{1}{NP} Y_t' - \frac{1}{NP} \epsilon_t$$

(14)

and, consequently, the disturbance uncertainty is given by

$$E_T\left[ (f_t - HF_t)(f_t - HF_t)' \right] = \frac{1}{N^2 T} E_T\left[ P' \epsilon_t \epsilon_t' \right] P.$$  

(15)

Finally, the covariance is given by

$$E_T\left[ (\tilde{f}_t - f_t)(f_t - HF_t)' \right] = \frac{1}{N^2 T} E_T\left[ (\tilde{P} - P)' \right] Y_t Y_t' \left( \tilde{P} - P \right).$$

(16)

The required MSE is given by

$$E_T\left[ (\tilde{f}_t - HF_t)(\tilde{f}_t - HF_t)' \right] = \frac{1}{N^2 T} E_T\left[ (\tilde{P} - P)' \right] Y_t Y_t' \left( \tilde{P} - P \right) + \frac{1}{N^2 T} E_T\left[ (2\tilde{P} - P)' \epsilon_t \epsilon_t' \right] P.$$  

(17)

### 2.3 Asymptotic distribution of PC factors

The first asymptotic result on PC factor estimates, in the context of strict DFM, is due to Connor and Korajczyk (1986) who prove consistency of PC factors when $N$ goes to infinity and $T$ is fixed. Bai (2003) shows that, in this case, consistency requires to assume asymptotic orthogonality and homoscedasticity of the idiosyncratic components. Only under large $N$ and $T$, Bai (2003) establishes consistency in the presence of serial correlation and heteroscedasticity; see also Stock and Watson (2002) who show that the space spanned by the estimated factors is consistent when both $N$ and $T$ tend simultaneously to infinity if the serial and cross-sectional correlations of the idiosyncratic noises are weak and the factors are pervasive. Furthermore, if $\sqrt{N} \rightarrow 0$, Bai (2003) derives the limiting distribution of the factors. Under the restrictions $\frac{1}{T} F' F = I_r$ and the diagonal elements of $P' P$ being distinct and positive and arranged in decreasing order, Bai and Ng (2013) show that

$$\sqrt{N} \left( \tilde{f}_t - F_t \right) \overset{d}{\rightarrow} N\left( 0, \Sigma_p^{-1} \Gamma I \Sigma_p^{-1} \right),$$

(18)

where $\Sigma_p = \lim_{N \rightarrow \infty} \frac{1}{N} P' P$ and $\frac{1}{N} \sum_{i=1}^{N} \epsilon_{it} \overset{d}{\rightarrow} N\left( 0, \Gamma_i \right)$. This result shows that $\tilde{f}_t$ is asymptotically equivalent to the least squares estimator for $F_t$ in a cross-section regression with $P$ as the regressor, as if $P$ were observable. Furthermore, Bai (2003) shows that, if the idiosyncratic noises
are serially uncorrelated, the limiting distributions are asymptotically independent across \( t \) i. Given that \( \hat{f} = fV \) and \( \frac{1}{N} \tilde{P}'\tilde{P} = \tilde{V} \), it is straightforward to see that

\[
\sqrt{N} \left( \hat{f}_t - \frac{P'P}{N} F_t \right) \xrightarrow{d} N(0, \Gamma_t).
\]

(19)

where, according to Bai and Ng (2006) \( \Gamma \) can be estimated assuming that the idiosyncratic errors are cross-sectionally uncorrelated, as follows \(^3\),

\[
\tilde{\Gamma}_t = \frac{1}{N} \sum_{i=1}^{N} \tilde{p}_i\tilde{p}_i'\tilde{\varepsilon}_{it}^2
\]

(20)

where, \( \tilde{p}_i \) is the \( i-th \) row of the factor loading matrix \( \tilde{P} \) and the residuals are given by \( \tilde{\varepsilon}_{it} = Y_{it} - \tilde{p}'_i\tilde{F}_t \).

Based on the asymptotic distribution in (19), \( (1 - \alpha) \% \) confidence bands for the factors can be constructed as follows

\[
\left( \frac{\tilde{P}'\tilde{P}}{N} \right)^{-1} \hat{f}_t \pm Z_{\alpha/2} \left[ \left( \frac{\tilde{P}'\tilde{P}}{N} \right)^{-1} \tilde{\Gamma}_t \left( \frac{\tilde{P}'\tilde{P}}{N} \right)^{-1} \right]^{1/2}
\]

(21)

with \( Z_{\alpha/2} \) being the \( \alpha/2 \) quantile of the standard normal distribution. Note that the estimated finite sample approximation of the asymptotic covariance matrix of \( \hat{f}_t \) (and, consequently, \( \hat{f}_t \)) is asymptotically equivalent to a least square estimator in which \( P \) is treated as if it were known, this asymptotic approximation underestimates the covariance of \( f_t \) as it does not take into account neither the MSE attributed to parameter uncertainty in (13) nor the covariance in (16). Consequently, unless \( T \) is very large, the asymptotic MSE will underestimate the finite sample MSE and the corresponding confidence bands for \( F_t \) will show undercoverage. As an illustration, the first row, first column of Figure 1 plots a factor and its corresponding asymptotic confidence bands.

### 2.4 Finite sample performance

We carry out Monte Carlo experiments in order to assess the finite sample adequacy of the asymptotic distribution when constructing confidence bands for the unobserved factors. The Monte Carlo experiments are performed using DFM of increasing complexity. The first model considered is the ubiquitous single factor model with temporal and cross-sectionally independent idiosyncratic errors. Then, we consider the same model in which the idiosyncratic term is either cross-correlated, temporally dependent or heteroscedastic. Jackson et al (2016) show that the

\(^3\)Bai and Ng (2006) propose this estimator of the asymptotic covariance matrix arguing that if the cross correlation in the errors is small, assuming that they are zero could be convenient because the sampling variability from their estimation could cause an efficiency loss.
conclusions for \( r = 1 \) could not always be generalized to cases in with \( r > 1 \). Consequently, we also perform Monte Carlo experiments in a DFM which \( r = 2 \). The data generating process (DGP) is given by the DFM in equations (1)-(3) with \( N, T = 20, 50 \) and \( 100 \). The number of Monte Carlo replicates is \( R = 1000 \). The number of factors is \( r = 1 \), and the idiosyncratic noises are homoscedastic and cross-sectionally uncorrelated white noises. The matrix of factor loadings, \( P \), is generated once from a uniform distribution in \([0,1]\). In order to analyze the effect of the temporal dependence of the factor on its accuracy, we consider several values of the autoregressive parameter of the factor, \( \Phi = 0.3, 0.5 \) and \( 0.7 \). In each case, the noise in equation (2), \( \eta_t \), has variance such that \( \text{Var}(F_t) = 1 \). Finally, the covariance matrix of the idiosyncratic noises is given by \( \Sigma_a = q^{-1}I \). Given that \( \text{Var}(F_t) = 1 \), the signal to noise ratio is given by \( q = \frac{\text{Var}(F_t)}{\text{Var}(\varepsilon_t)} \).

We consider several values of \( q \), namely \( q = 2, 1 \) and \( 0.5 \); see, Breitung and Eickmeier (2016) point out that the accuracy of factor estimates can depend on the signal to noise ratio.

For each replicate, \( i = 1, ..., R \), and moment of time \( t = 1, ..., T \) we construct point wise intervals as in (21) with nominal coverages 70\% and 95\% \(^4\). Then, at each moment of time, the empirical coverage is computed by counting how many true factors \( F_t \), lie inside the corresponding interval through the Monte Carlo replicates as \( C_{t}^{(i)} = \frac{1}{R} \sum_{i=1}^{R} I \left( F_t^{(i)} \in CI_t^{(i)} \right) \) where \( I(\cdot) \) is the indicator function. We should mention that the coverages are rather constant over time. Figure 2 illustrates the evolution of \( C_{t}^{(i)} \) for \( t = 1, ..., T \) for \( T = N = 50 \). It can be shown that the coverages remains almost constant across \( t \) \(^5\). Finally, we also compute the length of each interval at each moment of time and for each replicate. Table 1 reports the average coverage across time and the average length across time and Monte Carlo replicates for different temporal and cross-sectional dimensions when \( \phi = 0.7 \) and \( \text{SNR} = 1 \). We should point out that, although the coverages of the asymptotic intervals are slightly smaller when \( \text{SNR} \) decreases, conclusions are roughly the same as those reported in Table 1 \(^6\).

Observe that, regardless the sample, \( N \) and \( T \), Table 1 shows that the coverages of the asymptotic bands are always bellow the nominal coverages. Furthermore, we can observe that if \( T = 20 \), the undercoverage is larger as \( N \) increases. On the other hand, if \( N = 20 \), increasing \( T \) reduces the undercoverage.

In order to have a better understanding of the finite sample properties of the PC estimator when data has a more realistic structure, we also generate the idiosyncratic errors by equation (3) with \( \Gamma = \gamma I_N \) and \( \gamma = 0.5 \) and \( 0.7 \). When \( \gamma = 0 \), cross-sectionally correlated errors and cross-section

\(^{4}\)Forni et al. (2014) and Barigozzi et al. (2016) construct 64\% confidence bands for IRFS. Forni et al. (2014) consider a nominal coverage of 90\% while Bai (2003) considers 95\%.

\(^{5}\)Results for other temporal and cross-section are similar. They are available by request.

\(^{6}\)Conclusions for other values of \( \phi \), \( \text{SNR} \) and different samples sizes are similar. They are available upon request.
heteroscedastic errors are also generated such that $\Sigma_a$ is a Toeplitz matrix with parameter 0.5 and $\Sigma_a = \text{diag} \left[ \sigma_a^2 U(a, b) \right]$, respectively. When the errors are heteroscedastic, $\sigma_a^2 = 0.1, 1, 2$ and 10, $a = 0.1, 0.5$ and 0.9 and $b = 2, 1.5$ and 1.1. Table 2 provides the Monte Carlo averages of the coverages when there is serial or cross-sectional dependence in the idiosyncratic term ($\gamma = 0.7$) and also in the presence of cross-sectional heteroscedasticity. Introducing temporal dependence in the idiosyncratic noise produce extremely low sample coverages. This is consistent with Bai (2003) who point out that for a fixed $T$, it is not possible to obtain a consistent estimation in presence of serial correlation. Results when the idiosyncratic terms are heteroscedastic or cross-correlated are quite similar to the independency scenario.

3 Extant bootstrap procedures for PC factors

Several authors propose implementing resampling techniques in the context of PC for iid observations; see, for example, Beran and Srivastava (1985), Stauffer et al. (1985), Timmerman et al. (2007), Babamoradi et al. (2013), Van Aelst et al. (2013) and Fisher et al. (2015). However, these procedures are not appropriate for DFM. In recent years, several resampling methods have been proposed in the context of DFMs. Although these procedures have other objectives than constructing confidence bands for the extracted factors they allow for the construction of bootstrap bands of the factors as a subproduct. In this section, we describe these extant resampling algorithms and carry out Monte Carlo experiments to assess their adequacy to construct confidence bands for the extracted PC factors. The extant algorithms can be classified into two main groups: i) Block bootstrap algorithms and ii) residual algorithms.

3.1 Block bootstrap

Gospodinov and Ng (2013) propose a moving block bootstrap of the original vector of observations. Denoting by $B_{t,m} = (Y_t, Y_{t+1}, \ldots, Y_{t+m-1})$ a block of $m$ ($1 \leq m < T$) consecutive observations of $Y_t$, bootstrap replicates $\tilde{Y}_t^{*(b)}$ are obtained by drawing with replacement $K = T/m$ blocks from $(B_{1,m}, B_{2,m}, \ldots, B_{T-m+1,m})$ for $b = 1, \ldots, B$. PC estimates $\tilde{F}_t^{*(b)}$ are obtained as $\sqrt{T}$ times the eigenvectors correspont to the $r$ largest eigenvalues of the $T \times T$ matrix $Y_t^{*(b)} Y_t^{*(b)'}$. The block size $m$ should grow at a slower rate than $T$. Gospodinov and Ng (2013) consider $m = 4$ for forecasting purposes and obtain similar results for other block sizes for $m \in [4, 24]$. Denote by $\tilde{G}_t^*(x)$ the empirical distribution of $\tilde{F}_t^{*(b)}$ given by

$$\tilde{G}_t^*(x) = \#(\tilde{F}_t^{*(b)} \leq x)/B.$$  \hspace{1cm} (22)

7 Results for different sample sizes and different idiosyncratic structures are available from the authors from request.
Then, for each \( t = 1, \ldots, T \), \((1 - \alpha)\%\) confidence bands for the extracted factors can be constructed as

\[
\left( q_{(\alpha/2)t}, q_{(1-\alpha/2)t} \right),
\]

where \( q_{it} \) is the \( \alpha/2 \) and \( i \) empirical quantile of \( \hat{G}^*_t(x) \). Alternatively, it is possible to compute the bootstrap RMSE, as follows

\[
RMSE_{Bt} = \sqrt{\frac{1}{B} \sum_{b=1}^{B} \left( \tilde{F}^*_t(b) - \frac{1}{B} \sum_{b=1}^{B} \tilde{F}^*_t(b) \right)^2}
\]

(24)

Then, assuming normality of the factors, \((1 - \alpha)\%\) confidence intervals are given by

\[
\tilde{f}_t \pm Z_{\alpha/2} \text{RMSE}_{Bt}.
\]

(25)

It is important to note that, when bootstrapping \( Y^*_t(b) \) as proposed by Gospodinov and Ng (2013), one obtains replicates of the marginal distribution of \( \{Y_t\} \) and, consequently, of the marginal distribution of \( F_t \) in (5). If confidence bands are constructed as in (23), they will be centered at zero with RMSE given by (24). They will have the correct coverages but they are uninformative. On the other hand, when the intervals are computed as in (25), the RMSE is marginal while the interval is centered at \( \tilde{f}_t \). Note that these intervals will be too wide with coverages well above the nominal. As an illustration, the second row of Figure 1 plots a factor generated by the DFM described in the previous section with \( r = 1 \) and white noise and cross-sectionally uncorrelated and homoscedastic idiosyncratic errors. The bands are constructed as in expressions (23) and (25) with \( B = 1000 \) bootstrap replicates; see Gonçalves and Perron (2014) who consider \( B = 399 \) while Ludvigson and Ng (2007, 2009, 2010) consider \( B = 1000 \). Table 3 reports the empirical coverages and lengths for the 70% and 95% confidence intervals constructed in the same way. As we have already pointed out, the coverages are around the nominal but the lengths are extremely large.

### 3.2 Residual bootstrap

Gonçalves and Perron (2014), Djogbénou et al. (2015) and Gonçalves et al. (2015) propose a residual-based bootstrap consisting in obtaining bootstrap replicates of \( Y_t \) as follows

\[
Y^*_t(b) = \tilde{P} \tilde{F}^*_t + \tilde{\epsilon}^*_t(b)
\]

(26)

where \( \tilde{\epsilon}^*_t(b) \) are random extractions with replacement from \( \hat{G}^*_\epsilon \), the empirical distribution of \( \hat{\epsilon}_t = Y_t - \hat{P} \hat{F}_t \). Then, using \( Y^*_t(b) \), obtain PC estimates of the factors, \( \hat{F}^*_t(b) \) as before. The

---

\(^8\)Gonçalves and Perron (2014) propose a wild bootstrap algorithm to obtain replicates of \( \tilde{\epsilon}^*_t(b) \) that take into account potential heteroscedasticity while Breitung and Eckmeier (2016) propose a block bootstrap scheme to account for the serial correlation of the idiosyncratic noises. This procedure has also been proposed by Ludvigson and Ng (2007, 2009, 2010).
confidence intervals can be constructed as in (23) or as in (25). The bootstrap replicates of $Y_t$ obtained in equation (26) are centered in the estimated common factor $\hat{P}_t F_t$ and incorporate uncertainty about the idiosyncratic noise. However, they do not add the uncertainty associated with the estimation of the common factor. As a consequence, the confidence bands are conditional but narrower than they should be. The first row, second column of Figure 1 illustrates this procedure with the same factor as used for the block bootstrap algorithm described above.

Gonçalves and Perron (2014) proof the asymptotic validity of their proposed procedure for the OLS estimator of the factor-augmented regression under the assumption of stationarity and when the idiosyncratic errors are weakly dependent across individuals and overt time. Their asymptotics require that $\sqrt{T}/N \to c$, thus allowing for the possibility that the factor estimation error enters into the limiting distribution of the OLS estimator as an asymptotic bias term. The bias problem has been discussed by Ludvigson and Ng (2011) who propose an analytical bias correction. They suggest that the form of $\Sigma_\varepsilon$ dictates the type of resampling that should be used. They show that a wild bootstrap algorithm outperforms the Bai and Ng (2006) asymptotic distribution. Djogebeou et al. (2015) generalize the results in Gonçalves and Perron (2014) and make the resampling procedure applicable to forecasting more than one step-ahead. Gonçalves and Perron (2014) show that just as $\tilde{f}_t$ estimates a rotation of $F_t$, $\tilde{f}_t^*$ estimates $H^* \tilde{f}_t$ where $H^*$ is the resampling analogue of the rotation matrix

$$
\frac{1}{T} \sum_{t=1}^T \| \tilde{f}_t^* - H^* \tilde{f}_t^* \|^2 = O_p \left( \frac{\delta^{2}_{NT}}{NT} \right)
$$

where $\delta_{NT} = \min \left( \sqrt{N}, \sqrt{T} \right)$. The sign of the factors is not identified. The validity of the residual wild bootstrap depends on moments of order 12 of factors and idiosyncratic noises.

Yamamoto (2016) considers two residual-based bootstrap procedures. The first one is based on factor estimation in the bootstrap replicates as recommended by Gonçalves and Perron (2014) while the second one treats the original factor as if it were the observed process $F_t^{(b)} = \hat{\Phi} F_{t-1}^{(b)} + u_t^{(b)}$. Although Yamamoto (2016) shows that both procedures are asymptotically valid to construct confidence bands for IRFs when $\sqrt{T}/N \to 0$, in finite samples, the first procedure has good coverage while the second tends to undercover. When dealing with confidence for the factors, the first procedure of Yamamoto (2016) is the same as that Gonçalves and Perron (2014) and, therefore, we do not consider it further in this paper.

The bands constructed using the second procedure are either marginal or based on the marginal RMSE as, at each moment of time, they are based on bootstrap replicates of the factors which are not based on the available information set. Therefore, we expect a similar behaviour as that of the bands constructed using Gospodinov and Ng (2013) algorithm; see the second row of

---

9 Alonzo et al. (2008 and 2011) propose a bootstrap with the same structure for forecasting purposes.
Figure 1 and the table 1.

Shintani and Guo (2015) propose a slight modification of the above algorithm. Obtaining bootstrap replicates of the factor \( F_t^{\ast(b)} = \Phi F_{t-1}^{\ast(b)} + \eta_t^{\ast(b)} \) where \( \eta_t^{\ast} \) are random extractions with replacement from \( \hat{G}_t \) to subsequently obtain bootstrap replicates of \( Y_t \) with \( \hat{y}_t^{\ast(b)} = \hat{p}_{tr}^{\ast(b)} \hat{f}_t^{\ast(b)} + \epsilon_t^{\ast(b)} \) by drawing pairs \( (\hat{p}_{tr}^{\ast(b)}, \epsilon_t^{\ast(b)}) \) from \( (\hat{p}_{tr}, \epsilon_t) \). Using \( Y_t^{\ast(b)} \) extract the factors \( \hat{F}_t^{\ast(b)} \) same as before. As explained before, the confidence bands of the extracted factors can also be constructed using the corresponding percentiles of \( \hat{G}_t^{\ast}(x) \) or with Gaussian approximations. It is important to point out that the confidence bands are also marginal as they are based on bootstrap replicates of the factors generated as \( F_t^{\ast(b)} = \Phi F_{t-1}^{\ast(b)} + \eta_t^{\ast(b)} \). As a consequence, we expect these bands to be wider than if the conditional bands were obtained\(^{10} \). Table 1 reports the finite sample performance of the second bootstrap procedure proposed by Shintani and Guo (2015). The estimation of the uncertainty of the factor extraction is above the real one leading to coverages well above the nominal.

Therefore, we can conclude that none of the bootstrap procedures already proposed in the context of PC factor extraction in DFM, are valid to measure the uncertainty of the factor, however, it is important to mention that none of these methods was designed with this purpose.

4 Conditional Resampling

Given that the bootstrap procedures described in the previous section are not adequate when the aim is to construct confidence intervals of the underlying factors, in this section, we propose a new resampling strategy designed with this purpose. Its asymptotic validity is established and its finite sample performance is analyzed through extensive Monte Carlo experiments.

4.1 Resampling Procedure

The new proposed resampling algorithm is based on computing first the MSE of \( \hat{f}_t \) given in equation (17) conditional on the parameter estimates and then its average over the distribution of the parameter estimator; see, Pfeffermann and Tiller (2005) and Rodriguez and Ruiz (2012) for the same strategy. Conditioning on the parameters estimates, the MSE in (17) becomes

\[
E_t \left[ \left( \hat{f}_t - HF_t \right) \left( \hat{f}_t - HF_t \right)^\prime | \hat{P} \right] = \frac{1}{N^2} \left[ \left( \hat{P} - P \right)^\prime Y_t Y_t^\prime \left( \hat{P} - P \right) + \frac{1}{N^2} \left( 2\hat{P} - P \right)^\prime \epsilon_t \epsilon_t^\prime P \right]
\]

(28)

The resampling algorithm is designed as follows:

\(^{10}\)The first algorithm proposed by Shintani and Guo (2015) is valid under the assumption of independence of \( Y_t \) given that the cross-sectional dimension is being bootstrapped.
1. Denote by $Y^{*(b)}$ a random sample of size $p \ (1 \leq p \leq N)$ from $Y$.

2. Using $Y^{*(b)}$, obtain PC estimates $\tilde{P}^{*(b)}$.

3. Compute $\hat{f}^{*(b)}_t = \frac{1}{N} \tilde{P}^{*(b)} Y_t$. Note that $\hat{f}^{*(b)}_t$ is conditional on $Y_t$ and incorporates the parameter uncertainty through $\tilde{P}^{*(b)}$.

4. Repeat steps 2, 3 and 4, $B$ times.

In order to have a measure of the uncertainty associated with PC factor extraction is necessary to compute a resampling analog of the real MSE.

The resampling analog of $E \left[ \left( \hat{f}_t - f_t \right) \left( \hat{f}_t - f_t \right)' \right]$ is given by

$$\frac{1}{B} \sum_{b=1}^{B} \left( \hat{f}^{*(b)}_t - \hat{f}_t \right) \left( \hat{f}^{*(b)}_t - \hat{f}_t \right)' = \frac{1}{BN^2} \sum_{b=1}^{B} \left( \tilde{P}^{*(b)} - \tilde{P} \right)' Y_t Y_t' \left( \tilde{P}^{*(b)} - \tilde{P} \right)$$

(29)

Note that to approximate the MSE when the loadings are known we can use the following relation:

$$E \left( \hat{MSE} - MSE \right) = E^* \left( MSE^* - \hat{MSE} \right)$$

(30)

The resampling analog of the second term of the MSE is given by

$$\frac{2}{N^2} E \left[ \tilde{P}' \varepsilon_t \varepsilon_t' P \right] = \frac{2}{BN^2} \sum_{b=1}^{B} \tilde{P}^{*(b)} \varepsilon_t \varepsilon_t' \tilde{P}$$

(31)

Finally, the third term of the MSE can be approximated as follows

$$\frac{1}{N^2} E \left[ P' \varepsilon_t \varepsilon_t' P \right] = \frac{1}{BN} - \frac{1}{BN} \sum_{b=1}^{B} \tilde{f}^{*(b)}$$

(32)

Alternatively, we can do it without the bias bootstrap correction and computing it as $\frac{1}{N} \tilde{\Gamma}$. Finally, the bootstrap analog of the MSE of $\hat{f}_t$ is given by

$$RMSE^* = \frac{1}{BN^2} \sum_{b=1}^{B} \left( \tilde{P}^{*(b)} - \tilde{P} \right)' Y_t Y_t' \left( \tilde{P}^{*(b)} - \tilde{P} \right) + \frac{1}{N} \tilde{\Gamma}$$

Assuming normality, the corresponding resampling $(1 - \alpha)\%$ confidence interval for the true factors, $F_t$, is given by

$$\left( \frac{\tilde{P}' \tilde{P}}{N} \right)^{-1} \hat{f}_t \pm z_{\alpha/2} \left( \frac{\tilde{P}' \tilde{P}}{N} \right)^{-1} \left( \frac{1}{B} \sum_{b=1}^{B} \left( \tilde{P}^{*(b)} - \tilde{P} \right)' Y_t Y_t' \left( \tilde{P}^{*(b)} - \tilde{P} \right) + \frac{1}{N} \tilde{\Gamma} \right) \left( \frac{\tilde{P}' \tilde{P}}{N} \right)^{-1}$$

(33)

where $z_{\alpha/2}$ is the $\alpha/2$ quantile of the normal distribution.
4.2 Finite Sample Performance

We carry out Monte Carlo experiments in order to assess the adequacy of the new resampling procedure to approximate the finite sample distribution of the PC factors and, consequently, to construct confidence bands. The Monte Carlo experiments are performed using the data generating process previously explained. The number of Resampling replicates is $B = 1000$. It should be noted that the sample size $p$ varies in the same direction as the ratio $T/N$.


\[\text{Nota, la relación es } p = 0.67 + 0.09 \ln(T/N). \quad \text{No sé si es muy cutre poner una relación de ese tipo. Estoy sacando mas puntos para aproximar mejor pero esa ecuación lo clava. } R^2=0.98.\]

For each Monte Carlo replicate, $i = 1, \ldots, R$, we construct the point wise intervals as in equation (32). Table 1 reports, for nominal coverages of 70% and 95%, the average coverage across time and the average length across time and Monte Carlo replicates for different temporal and cross-sectional dimensions when $\phi = 0.7$ and $SNR = 1$. We should point out that the coverages of the resampling intervals for different values of $\phi$ and $SNR$ are almost the same as those reported in Table 1. Figure 2 illustrates the evolution of $C_t^{(i)}$ for $t = 1, \ldots, T$ for $T = N = 50$. It can be shown that the coverages remains almost constant across $t$ (results for other samples sizes are similar. They are also available by request).

Observe that, regardless the sample dimensions, $N$ and $T$, the new resampling procedure estimate correctly the uncertainty of PC factor extraction with coverages always around the nominal. Table 2 provides the Monte Carlo averages of the coverages when there is serial or cross-sectional dependence in the idiosyncratic term ($\gamma = 0.7$) and also in the presence of cross-sectional heteroscedasticity. It can be also observed that, changing the structure of the idiosyncratic term does not affect the finite sample performance of the new procedure and it is still performing well.

5 Empirical Analysis

In this section, we exemplify the importance of a proper measurement of the uncertainty associated with the factors extracted by PC. For this purpose we analyze the quarterly series belonging to the database of the Ministry of Treasure and Public Administration, "Base de datos trimestrales de la economía española", which consists of a panel of 75 Spanish macroeconomic variables observed quarterly from the first quarter of 1980 to the last of 2014. The variables have been seasonally adjusted and converted to stationary. Moreover, they have been standardized to have zero mean and unite variance. Therefore, the total panel of data consists of $N=75$ variables and $T=140$ observations. We start the analysis estimating the number of factors to extract.

\footnote{Results for other values of $\phi$, $SNR$ are available upon request.}
More specially, we consider the information criteria of Bai and Ng (2002), the edge distribution of Onatski (2010) and the ratios of eigenvalues proposed by Ahn and Horenstein (2013). The number of factors to compute is one. The factors are extracted by PC and the confidence intervals are constructed following the procedures described above. The sum of the weights is $\sum_{i=1}^{N} p_i = 12.21$ with estimated weights larger than 0.8 in absolute value corresponding to: Gross capital formation, capital stock, imports, unemployment rate, rest of the world clients’ GDP, total resources of public administrations. \( \hat{\Phi} = 0.7, \hat{\sigma}_a^2 = [-0.93, 0.86] \) with the mode around 0, and serial dependence \( \hat{\gamma} = [-0.74, 0.97] \) distributed uniformly in this interval. Figure 3 plots the results of this analysis. It can be observed that the asymptotic confidence intervals are considerably narrower than the ones constructed following the procedure proposed in the previous section. The 95% asymptotic confidence intervals are almost equivalent to the 75% confidence intervals constructed using our procedure which, through an intense Monte Carlo experiment, has proven to be a better approach for measuring the uncertainty of factor extraction with PC in this kind of scenario. In other words, it is important to point out the importance of measuring the uncertainty of the factors correctly. If practitioners and policy decision makers use the asymptotic covariance matrix for measuring the uncertainty or for constructing confidence bands for the latent factors, could lead to a wrong interpretation of the economic reality -cycles and recessions- and to an incorrect density forecast in the context of diffusion indexes.

6 Conclusions

This paper explores different methods for improving the computation of the uncertainty associated to factor extraction using PC in DFM. By means of an intense Monte Carlo exercise, the finite sample performance of the asymptotic covariance matrix proposed by Bai and Ng (2006) is investigated, we see that this estimation underestimates the uncertainty of the PC factors, causing narrower gaussian confidence bands than desired. Moreover, it has been shown that the existing resampling procedures in the context of PC in DFM are not capable of measuring the uncertainty associated to the factor extraction correctly. Partly because some of them compute the marginal variance instead of the conditional one, or because they do not take into account the parameter uncertainty of the DFM. Finally, a resampling algorithm to compute the uncertainty of PC factors and to construct confidence intervals is presented. This algorithm has proven to have a better finite sample performance than the existing methods for a wide range of scenarios of very different nature. Many topics remain to be developed. First of all, it is desirable to expand the algorithm for the cases in which more than one factor is extracted and also, it is necessary to improve the performance of the procedure when strong serial dependence exists both in the factors and in the idiosyncratic term. A second interesting area of research would be
to study the importance of a correct measurement of the factor extraction uncertainty in density forecast using diffusion indexes and, moreover, to study the effect of a non-Gaussian idiosyncratic term. Another important extension would be to apply the procedure to empirical cases with the objective of constructing a stress indicator and warning signals for economic recessions.

We can extend the resampling methods proposed in this paper to analyse their performance in the context of constructing forecast intervals in the context of factor-augmented regressions, also known as diffusion index forecasts, and compare them with the extant procedures proposed by Goncalves et al. (2017) which do not require assuming normality of the forecast errors.

References


**Table 1**: Monte Carlo results on the asymptotic approximation and the new resampling procedure when the idiosyncratic component is homoscedastic and serial and cross-sectionally uncorrelated. $\phi = 0.7$ and $SNR = 1$ (Conclusions for other values of $\phi$, $SNR$ and different sample sizes are similar. They are available upon request).

<table>
<thead>
<tr>
<th></th>
<th>$T=20; N=20$</th>
<th>$T=20; N=50$</th>
<th>$T=20; N=100$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coverage</td>
<td>Length</td>
<td>Coverage</td>
</tr>
<tr>
<td>Asymptotic</td>
<td>0.78</td>
<td>1.39</td>
<td>0.77</td>
</tr>
<tr>
<td>(0.95)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asymptotic</td>
<td>0.50</td>
<td>0.74</td>
<td>0.48</td>
</tr>
<tr>
<td>(0.70)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Procedure</td>
<td>0.72</td>
<td>0.90</td>
<td>0.69</td>
</tr>
<tr>
<td>(0.70)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Procedure</td>
<td>0.94</td>
<td>1.71</td>
<td>0.93</td>
</tr>
<tr>
<td>(0.95)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2**: Monte Carlo results on the asymptotic approximation for different idiosyncratic structures with $T=50$ and $N=50$ (Conclusions for other values of $\phi$, $\gamma$, $(a, b)$ and $SNR$ are similar. They are available upon request).

<table>
<thead>
<tr>
<th></th>
<th>Indepenendency</th>
<th>Serial Dependency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coverage</td>
<td>Length</td>
</tr>
<tr>
<td>Asymptotic</td>
<td>0.86</td>
<td>0.96</td>
</tr>
<tr>
<td>(0.95)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asymptotic</td>
<td>0.58</td>
<td>0.51</td>
</tr>
<tr>
<td>(0.70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Procedure</td>
<td>0.70</td>
<td>0.65</td>
</tr>
<tr>
<td>(0.70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Procedure</td>
<td>0.93</td>
<td>1.24</td>
</tr>
<tr>
<td>(0.95)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Cross-Section Dependency</th>
<th>Heteroscedasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coverage</td>
<td>Length</td>
</tr>
<tr>
<td>Asymptotic</td>
<td>0.85</td>
<td>0.95</td>
</tr>
<tr>
<td>(0.95)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asymptotic</td>
<td>0.56</td>
<td>0.50</td>
</tr>
<tr>
<td>(0.70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Procedure</td>
<td>0.70</td>
<td>0.66</td>
</tr>
<tr>
<td>(0.70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Procedure</td>
<td>0.93</td>
<td>1.25</td>
</tr>
<tr>
<td>(0.95)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Finite sample performance of the extant Bootstrap procedures when the idiosyncratic component is homoscedastic and serial and cross-sectionally uncorrelated. $T = N = 50$, $\phi = 0.7$ and $SNR = 1$(Conclusions for other values of $\phi$, $SNR$ and different sample sizes are similar. They are available upon request).

<table>
<thead>
<tr>
<th>Bootstrap Method</th>
<th>Coverage</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block Bootstrap (0.7)</td>
<td>0.76</td>
<td>2.03</td>
</tr>
<tr>
<td>Goncalves and Perron (0.7)</td>
<td>0.51</td>
<td>0.41</td>
</tr>
<tr>
<td>Yamamoto (0.7)</td>
<td>0.76</td>
<td>1.98</td>
</tr>
<tr>
<td>Shintani and Guo (0.7)</td>
<td>0.83</td>
<td>1.11</td>
</tr>
<tr>
<td>Block Bootstrap (0.95)</td>
<td>0.95</td>
<td>3.89</td>
</tr>
<tr>
<td>Goncalves and Perron (0.95)</td>
<td>0.72</td>
<td>0.63</td>
</tr>
<tr>
<td>Yamamoto (0.95)</td>
<td>0.96</td>
<td>3.73</td>
</tr>
<tr>
<td>Shintani and Guo (0.95)</td>
<td>0.97</td>
<td>2.14</td>
</tr>
</tbody>
</table>
Figure 1: Illustration of 95% confidence bands constructed using different methods. Asymptotic approximation (first row, first column), idiosyncratic residual bootstrap (first row, second column) and block bootstrap (second row). The first column of the second row is based on Gaussian bands with bootstrap RMSEs while the second column plots the bands constructed from the bootstrap densities.
Figure 2: Asymptotic (blue lines) and Resampling coverages (red lines) from $t = 1, \ldots, T$. Nominals: 70%, 80% and 95%.

Figure 3: Asymptotic (blue lines) and Resampling intervals (red lines) for the economic cycle in Spain (black line).