Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

A Parametric Estimation Method for Dynamic Factor Models of Large Dimensions

George Kapetanios and Massimiliano Marcellino

Working Paper n. 305

January 2006

IGIER – Università Bocconi, Via Salasco 5, 20136 Milano –Italy
http://www.igier.uni-bocconi.it

A Parametric Estimation Method for Dynamic Factor Models of Large Dimensions*

George Kapetanios†
Queen Mary University of London

Massimiliano Marcellino‡
IEP-Bocconi University, IGIER and CEPR

Revised Version: January 2006

Abstract

The estimation of dynamic factor models for large sets of variables has attracted considerable attention recently, due to the increased availability of large datasets. In this paper we propose a new parametric methodology for estimating factors from large datasets based on state space models and discuss its theoretical properties. In particular, we show that it is possible to estimate consistently the factor space. We also develop a consistent information criterion for the determination of the number of factors to be included in the model. Finally, we conduct a set of simulation experiments that show that our approach compares well with existing alternatives.

J.E.L. Classification: C32, C51, E52

Keywords: Factor models, Principal components, Subspace algorithms

*We would like to thank an anonymous Referee for helpful comments on a previous draft.

†Corresponding author. Department of Economics, Queen Mary, University of London, Mile End Rd., London E1 4NS. Email: G.Kapetanios@qmul.ac.uk

‡IGIER - Università Bocconi, Via Salasco 5, 20136, Milano, Italy. Phone: +39-02-5836-3327. Fax: +39-02-5836-3302. E-mail: massimiliano.marcellino@uni-bocconi.it

1 Introduction

Recent work in the econometric literature considers the problem of summarising efficiently a large set of variables and using this summary for a variety of purposes including forecasting. Work in this field has been carried out in a series of recent papers by Stock and Watson (2001, 2002) (SW) and Forni, Lippi, Hallin and Reichlin (1999,2000) (FHLR). Factor analysis has been the main tool used in summarising the large datasets.

The static version of the factor model was analyzed, among others, by Chamberlain and Rothschild (1983), Connor and Korajczyk (1986, 1993). Geweke (1977) and Sargent and Sims (1977) studied a dynamic factor model for a limited number of series. Further developments were due to Stock and Watson (1989, 1991), Quah and Sargent (1993) and Camba-Mendez et al (2001), but all these methods are not suited when the number of variables is very large due to the computational cost, even when a sophisticated EM algorithm is used for optimization, as in Quah and Sargent (1993).

For this reason, SW have suggested a non-parametric principal component based estimation approach in the time domain, and shown that principal components can estimate consistently the factor space asymptotically. FHLR have developed an alternative non-parametric procedure in the frequency domain, based on dynamic principal components (see Chapter 9 of Brillinger (1981)), that incorporates an explicitly dynamic element in the construction of the factors.

In this paper we suggest a third approach for factor estimation that retains the attractive framework of a parametric state space model but is computationally feasible for very large datasets because it does not use maximum likelihood but linear algebra methods, based on subspace algorithms used extensively in engineering, to estimate the state. To the best of our knowledge, this is the first time that these algorithms are used for factor estimation.

We analyze the asymptotic properties of the new estimators, first for a fixed number of series, N, and then allowing N to diverge. We show that as long as N grows less than $T^{1/3}$, where T is the number of observations, the subspace algorithm yields consistent estimators for the space spanned by the factors. We have also developed an information criterion that leads to consistent selection of the number of factors to be included in the model, along the lines of Bai and Ng (2002) for the static principal component approach. Finally, we have compared the finite sample performance of our estimator and of those by SW and FHLR by means of simulation experiments, finding that our proposal performs quite well.

The paper is organised as follows. Section 2 presents the state space model approach and derives the properties of the estimators for the fixed N case. Section 3 deals with the diverging N case, with correlation of the idiosyncratic components, and with a modified algorithm to analyze datasets with $N > T$. Section 4 discusses the Monte Carlo experiments. Finally, Section 5 summarizes and concludes.

2 The state space factor estimator

In this section we present and discuss the basic state space representation for the factor model, discuss the subspace estimators, and derive their asymptotic properties when T diverges and N is fixed. In the following section we extend the framework to deal with the N going to infinity case, with the analysis of datasets with a larger cross-section than time-series dimension, and with cross-sectionally or serially correlated idiosyncratic errors.

2.1 The basic state space model

Following Deistler and Hannan (1988), we consider the following state space model.

$$\begin{aligned} x_{Nt} &= Cf_t + \epsilon_t, \quad t = 1, \ldots, T \\ f_t &= Af_{t-1} + B^* v_{t-1}, \end{aligned} \tag{1}$$

where x_{Nt} is an N-dimensional vector of stationary zero-mean variables observed at time t, f_t is a k-dimensional vector of unobserved states (factors) at time t, and ϵ_t and v_t are multivariate, mutually uncorrelated, standard orthogonal white noise sequences of dimension, respectively, N and k. B^* is assumed to be nonsingular.[1] The aim of the analysis is to obtain estimates of the states f_t, for $t = 1, \ldots, T$. We make the following assumption

Assumption 1 *(a) $|\lambda_{\max}(A)| < 1$ and $|\lambda_{\min}(A)| > 0$ where $|\lambda_{\max}(.)|$ and $|\lambda_{\min}(.)|$ denote, respectively, the maximum and minimum eigenvalue of a matrix in absolute value.*

(b) The elements of C are bounded

The first part of assumption 1-(a), combined with assumption 1-(b) ensures that x_{Nt} is stationary. The second part of assumption 1-(a) implies that each factor is correlated over

[1] In general, in the dynamic factor model the idiosyncratic errors are allowed to be correlated over time and across variables, see e.g. Stock and Watson (2002a, 2002b) for precise conditions. We consider this extension in Section 3.2 from a theoretical point of view and in Section 4 in a set of Monte Carlo experiments.

time, which is important to distinguish it from the idiosyncratic white noise error terms. Notice also that the factors are driven by lagged errors. Without this assumption, it would not be possible to obtain consistent estimators for the factors when N is finite, as we will discuss in more detail below.

This model is quite general. Its aim is to use the states as a summary of the information available from the past on the future evolution of the system. To illustrate its generality we give an example where a factor model with factor lags in the measurement equation can be recast in the above form indicating the ability of the model to model dynamic relationships between x_{Nt} and f_t. Define the original model to be

$$\begin{aligned} x_{Nt} &= C_1 f_t + C_2 f_{t-1} + \epsilon_t, \quad t = 1, \dots, T \\ f_t &= A f_{t-1} + B^* v_{t-1}, \end{aligned} \tag{2}$$

This model can be written as

$$\begin{aligned} x_{Nt} &= (C_1, C_2)\tilde{f}_t + \epsilon_t, \quad t = 1, \dots, T \\ \tilde{f}_t &= \begin{pmatrix} f_t \\ f_{t-1} \end{pmatrix} = \begin{pmatrix} A & 0 \\ I & 0 \end{pmatrix} \begin{pmatrix} f_{t-1} \\ f_{t-2} \end{pmatrix} + \begin{pmatrix} B^* & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_{t-1} \\ 0 \end{pmatrix}, \end{aligned} \tag{3}$$

which is a special case of the specification in (1), even though by not taking into account the particular structure of the A matrix and the reduced rank of the error process we are losing in terms of efficiency.

A large literature exists on the identification issues related with the state space representation given in (1). An extensive discussion may be found in Deistler and Hannan (1988). In particular, they show in Chapter 1 that (1) is equivalent to the prediction error representation of the state space model given by

$$\begin{aligned} x_{Nt} &= C f_t + D u_t, \quad t = 1, \dots, T \\ f_t &= A f_{t-1} + B u_{t-1}. \end{aligned} \tag{4}$$

where D is nonsingular and u_t is an orthogonal white noise process.

While (1) and (4) are equivalent, the latter is more convenient for the derivation of our estimation algorithm. Note that as at this stage the number of series, N, is large but fixed we need to impose no conditions on the structure of C. Conditions on this matrix will be discussed later when we consider the case of N tending to infinity and possibly correlated idiosyncratic errors.

2.2 Subspace Estimators

As we have mentioned in the introduction, maximum likelihood techniques, possibly using the Kalman filter, may be used to estimate the parameters of the model under some identification scheme. Yet, for large datasets this is very computationally intensive. Quah and Sargent (1993) developed an EM algorithm that allows to consider up to 50-60 variables, but it is still so time-consuming that it is not feasible to evaluate its performance in a simulation experiment.

To address this issue, we propose to exploit subspace algorithms, which avoid expensive iterative techniques by relying on matrix algebraic methods, and can be used to provide estimates for the factors as well as the parameters of the state space representation. To the best of our knowledge, our paper is the first application of subspace algorithms for factor estimation.

There are many subspace algorithms, and vary in many respects, but a unifying characteristic is their view of the state as the interface between the past and the future in the sense that the best linear prediction of the future of the observed series is a linear function of the state. A review of existing subspace algorithms is given by Bauer (1998) in an econometric context. Another review with an engineering perspective may be found in Van Overschee and De Moor (1996).

The starting point of most subspace algorithms is the following representation of the system which follows from the state space representation in (4) and the assumed nonsingularity of D.

$$X_t^f = \mathcal{O}\mathcal{K}X_t^p + \mathcal{E}E_t^f \tag{5}$$

where $X_t^f = (x'_{Nt}, x'_{Nt+1}, x'_{Nt+2}, \ldots)'$, $X_t^p = (x'_{Nt-1}, x'_{Nt-2}, \ldots)'$, $E_t^f = (u'_t, u'_{t+1}, \ldots)'$, $\mathcal{O} = [C', A'C', (A^2)'C', \ldots]'$, $\mathcal{K} = [\bar{B}, (A-\bar{B}C)\bar{B}, (A-\bar{B}C)^2\bar{B}, \ldots]$, $\bar{B} = BD^{-1}$ and

$$\mathcal{E} = \begin{pmatrix} D & 0 & \ldots & 0 \\ CB & D & \ddots & \vdots \\ CAB & \ddots & \ddots & 0 \\ \vdots & & CB & D \end{pmatrix}.$$

The derivation of this representation is simple once we note that (i) $X_t^f = \mathcal{O}f_t + \mathcal{E}E_t^f$ and (ii) $f_t = \mathcal{K}X_t^p$. The best linear predictor of the future of the series at time t is given by $\mathcal{O}\mathcal{K}X_t^p$. The state is given in this context by $\mathcal{K}X_t^p$ at time t. The task is therefore to provide an estimate for $\mathcal{K}$.[2]

The above representation involves infinite dimensional vectors. In practice, truncation is used to end up with finite sample approximations given by $X^f_{s,t} = (x'_{Nt}, x'_{Nt+1}, x'_{Nt+2}, \ldots, x'_{Nt+s-1})'$ and $X^p_{p,t} = (x'_{Nt-1}, x'_{Nt-2}, \ldots, x'_{Nt-p})'$. Then an estimate of $\mathcal{F} = \mathcal{OK}$ may be obtained by regressing $X^f_{s,t}$ on $X^p_{p,t}$. Following that, the most popular subspace algorithms use a singular value decomposition (SVD) of an appropriately weighted version of the least squares estimate of $\mathcal{F}$, denoted by $\hat{\mathcal{F}}$. In particular the algorithm we will use, due to Larimore (1983), applies an SVD to $\hat{\Gamma}^f \hat{\mathcal{F}} \hat{\Gamma}^p$, where $\hat{\Gamma}^f$ and $\hat{\Gamma}^p$ are the sample covariances of $X^f_{s,t}$ and $X^p_{p,t}$ respectively. These weights are used to determine the importance of certain directions in $\hat{\mathcal{F}}$. Then, the estimate of $\mathcal{K}$ is given by

$$\hat{\mathcal{K}} = \hat{S}_k^{1/2} \hat{V}_k' \hat{\Gamma}^{p^{-1/2}}$$

where $\hat{U}\hat{S}\hat{V}'$ represents the SVD of $\hat{\Gamma}^{f^{-1/2}} \hat{\mathcal{F}} \hat{\Gamma}^{p^{1/2}}$, $\hat{V}_k$ denotes the matrix containing the first k columns of $\hat{V}$ and $\hat{S}_k$ denotes the heading $k \times k$ submatrix of $\hat{S}$. $\hat{S}$ contains the singular values of $\hat{\Gamma}^{f^{-1/2}} \hat{\mathcal{F}} \hat{\Gamma}^{p^{1/2}}$ in decreasing order. Then, the factor estimates are given by $\hat{\mathcal{K}} X^p_t$. We refer to this method as SSS.

For what follows it is important to note that the choice of the weighting matrices $\hat{\Gamma}^f$ and $\hat{\Gamma}^p$ is important but not crucial for the asymptotic properties of the estimation method. This is because the choice does not affect neither the consistency nor the rate of convergence of the factor estimator. For these properties, the weighting matrices are only required to be nonsingular. Therefore, for the sake of simplicity, in the theoretical analysis and in the Monte Carlo study,

Assumption 2 *We set* $\hat{\Gamma}^f = I_{sN}$ *and* $\hat{\Gamma}^p = I_{pN}$

A second point to note is that consistent estimation of the factor space requires the "lag" truncation parameter p to increase at a rate greater than $ln(T)^{\alpha}$, for some $\alpha > 1$ that depends on the maximum eigenvalue of A, but at a rate lower than $T^{1/3}$. A simplified condition for p is to set it to $T^{1/r}$ for any $r > 3$.

For consistency, the "lead" truncation parameter s is also required to be set so as to satisfy $sN \geq k$. As N is usually going to be very large for the applications we have in

[2]This "interface property" of the state was already emphasized in the '60s by Kalman and Akaike in related contexts.

mind, this restriction is not binding and we can use $s = 1$. This is relevant in particular in a forecasting context because with $s = 1$ only contemporaneous and lagged values of the variables are used for factor estimation.

Once estimates of the factors have been obtained, if estimates of the parameters of the model (including the factor loadings) are subsequently required, least squares methods may be used with the estimated factors instead of the true ones. The resulting estimates have been proved to be $\sqrt{T}$-consistent and asymptotically normal in Bauer (1998) in a related context. We note that the identification scheme underlying the above estimators of the parameters is implicit, and depends on the normalisation used in the computation of the SVD.

It is worth pointing out that the estimated parameters can be used with the Kalman filter on the state space model to obtain both filtered and smoothed estimates of the factors. Since the SSS method produces factor estimates at time t conditional on data available at time $t-1$, it may be possible that smoothed estimates from the Kalman filter are superior to those obtained by the SSS method. However, the parameter estimates are conditional on the factor estimates obtained in the first step by the SSS method. Limited experimentation using the Monte Carlo setup reported in Kapetanios and Marcellino (2004) suggests that the loss in performance of the smoothed Kalman filter factor estimate because of the use of estimated factors from the SSS method, is roughly similar to the benefit of using all the data. Moreover, in general, factors estimated using the SSS method outperform filtered Kalman filter factor estimates.

Finally, we must note that the SSS method is also applicable in the case of unbalanced panels. In analogy to the work of SW, use of the EM algorithm, described there, can be made to provide estimates both of the factors and of the missing elements in the dataset.

2.3 Asymptotic properties

Let us denote the true number of factors by k^0 and investigate in more detail OLS estimation of the multivariate regression model

$$X^f_{s,t} = \mathcal{F}X^p_{p,t} + \mathcal{E}E^f_{s,t} \tag{6}$$

where $E^f_t = (u'_t, u'_{t+1}, \ldots, u'_{t+s})'$. Estimation of the above is equivalent to estimation of each equation separately. We make the following assumptions

Assumption 3 *u_t is an i.i.d. $(0, \Sigma_u)$ sequence with finite fourth moments.*

Assumption 4 $p = O\left((\ln(T))^{\alpha}\right)$, $1 < \alpha < \infty$.

Denote $X^p = (X^p_{p,1}, ..., X^p_{p,T})'$. Then we have the following theorem:

Theorem 1 ***(Consistency and Rate of Convergence).*** *If we define $\hat{f}_t = \hat{\mathcal{K}} X^p_{p,t}$, then, under assumptions 1-4, $vec(\hat{f}) - vec(H^k f^0) = o_p(p^2 T^{-a})$ for all $a < 1/2$, where H^k is a square matrix of full rank.*

Proof. By (4) and (5) we can see that $\mathcal{K} X^p_{p,t}$ spans the space of the true factors and is therefore an estimator for $H^k f$. So we need to concentrate on the properties of $\hat{\mathcal{K}}$ as an estimator of $\mathcal{K}$. We want to show that

$$vec(\hat{\mathcal{K}}') - vec(\mathcal{K}') = o_p(pT^{-a}), \text{ for all } a < 1/2 \tag{7}$$

We now define formally the function $g(.)$ such that

$$vec(\hat{\mathcal{K}}') = g\left(vec(\hat{\mathcal{F}})\right)$$

Therefore, $g(.)$ defines the singular value decomposition operator. By theorems 5.6 and 5.8 of Chatelin (1983) $g(.)$ is continuous, differentiable and therefore admits a first order Taylor expansion. Therefore,

$$vec(\hat{\mathcal{K}}') - vec(\mathcal{K}') = \frac{\partial g'}{\partial \mathcal{F}}\left(vec(\hat{\mathcal{F}}) - vec(\mathcal{F})\right)\frac{\partial g'}{\partial \mathcal{F}} + o_p(T^{-1/2}) \tag{8}$$

As a result if

$$vec(\hat{\mathcal{F}}) - vec(\mathcal{F}) = o_p(T^{-a}), \text{ for all } a < 1/2 \tag{9}$$

then (7) follows. Note that we have $o_p(pT^{-a})$ and not $o_p(T^{-a})$ in (7) since the dimension of $\frac{\partial g'}{\partial \mathcal{F}}$ is of order p. We need to show (9). We first note that by assumption 1, the absolute value of the maximum eigenvalue of $\mathcal{F} = \mathcal{O}\mathcal{K}$, denoted $|\lambda_{\max}(\mathcal{F})|$, is less than one implying exponentially declining coefficients with respect to p. This implies that

$$\sum_{i=1}^{\infty} i^{1/2} ||\mathcal{F}(i)|| < \infty \tag{10}$$

where $\mathcal{F}(i)$ denotes the matrix coefficient of x_{Nt-i} in the regression of $X^f_{s,t}$ on $X^p_{p,t}$. Then, by Theorem 2.1 and (4.3) of Hannan and Kavalieris (1986) it follows that $vec(\hat{\mathcal{F}}) - vec(\mathcal{F}) = O_p\left((T/\ln(T))^{1/2}\right)$ which immediately implies (9). The final step involves obtaining the rate of convergence for $(vec(\hat{f}) - vec(H^k f))$. The factor estimates are linear combinations of the elements of $\hat{\mathcal{K}}$. Since both T and p grow, by assumption 4 $vec(\hat{f}) - vec(H^k f) = o_p(p^2 T^{-a})$ for all $a < 1/2$ since the factor estimates are made up of a linear combination of the elements of $\hat{\mathcal{K}}$ whose number is of order p. ■

Remark 1 *Assumption 4 is more restrictive than strictly necessary. Theorem 2.1 of Hannan and Kavalieris (1986) requires that* $p \to \infty$ *as* $T \to \infty$ *and* $p = o\left((T/\ln T)^{1/2}\right)$. *However, we use 4 as our aim is to provide a good rate of convergence of the factor estimate to the space spanned by the true factors. Note that under assumption 4* $vec(\hat{f}) - vec(H^k f) = o_p(T^{-a})$ *for all* $a < 1/2$ *since* $(\ln T)^\alpha = o(T^a)$ *for all* $a > 0$. *In practical terms this is not binding since by setting* α *appropriately a large range of values for* p *is possible in small samples.*

Remark 2 *The derivation of the rate of convergence hinges crucially on two steps: Firstly the fact that* $vec(\hat{\mathcal{K}}')$ *is asymptotically just a linear combination of* $vec(\hat{\mathcal{F}})$ *via (8) and secondly that* $vec(\hat{f})$ *is a linear combination of* $vec(\hat{\mathcal{K}}')$. *We expect that the derived rate of convergence is conservative in terms of* p *since the above two steps introduce the term* p^2 *which could be made smaller via, e.g., further analysis of* $\frac{\partial g'}{\partial \mathcal{F}}$. *However, given assumption 4, the gain from such an analysis will be minimal in terms of* T.

Remark 3 *It is important to mention that consistency is possible because in the model (1) the factors depend on lagged errors. Without this assumption, i.e., if* f_t *depends on* v_t *rather than on* v_{t-1}, *the SSS estimator would be consistent for* Af_{t-1} *but not for the space spanned by* f_t.

Remark 4 *At this point it is worth considering briefly the novelty of the results presented in Theorem 1. The main novelty lies in the application of the results of Hannan and Kavalieris (1986) which relate to infinite order AR processes to the context of factor analysis. In particular, the main insight lies in noting that the factor estimates can be analysed by analysing the behaviour of the OLS estimates of the coefficients of (6) and using the existing theory summarised in Hannan and Kavalieris (1986) to provide such an analysis.*

3 The case $N \to \infty$

In this section we firstly investigate the conditions for consistency of the SSS method when N diverges. Second, we discuss correlation of the idiosyncratic errors. Finally, we derive an information criterion for the selection of the number of factors.

3.1 Consistency of the SSS estimator

To prove consistency of the SSS estimator, we need to introduce a further assumption in addition to those given in the previous Section. In particular, we require

Assumption 5 $N = o\left(T^{1/6}\right)$

Then we have

Theorem 2 ***(Consistency and Rate of Convergence when*** $N \to \infty$***).*** *As* N *and* T *diverge, and under assumptions 1-5 ,* $vec(\hat{f}) - vec(H^k f) = o_p(N^3 p^2 T^{-a})$ *for all* $a < 1/2$*, where* H^k *is a square matrix of full rank.*

Proof. Denote the i-th row of $\hat{\mathcal{F}}$ and $\mathcal{F}$ by $\hat{\mathcal{F}}_i = \left(\hat{\mathcal{F}}_{i,1}, ..., \hat{\mathcal{F}}_{i,Np}\right)'$ and $\mathcal{F}_i = (\mathcal{F}_{i,1}, ..., \mathcal{F}_{i,Np})'$ respectively. For simplicity assume that $s = 1$. The proof of this Theorem follows along the lines considered in Theorem 1 once we prove that for each of the N equations in (6),

$$\hat{\mathcal{F}}_i - \mathcal{F}_i = o_p(T^{-a}) \text{ for all } a < 1/2 \tag{11}$$

. In particular note that now as each of the N equations in (6) contains Np regressors the order of $\frac{\partial g'}{\partial \mathcal{F}}$ is $N^2 p$ and that $\hat{f}$ is a linear combination of the elements of $\hat{\mathcal{K}}$ whose number is of order Np. These considerations lead to the $o_p(N^3 p^2 T^{-a})$ result in the statement of the theorem. To prove (11) we mirror the analysis of Theorems 4 and 5 of An et al (1982). For simplicity we consider Yule-Walker estimation of $\hat{\mathcal{F}}_i$ which is asymptotically equivalent to OLS estimation. Let γ_i^{fp} denote the covariance of $x_{i,Nt}$ and $X_{p,t}^p$ and $\hat{\gamma}_i^{fp}$ its sample counterpart. Then, by (25) of An et al (1982)

$$\Gamma^p \left(\hat{\mathcal{F}}_i - \mathcal{F}_i\right) = -\left(\hat{\Gamma}^p - \Gamma^p\right)\left(\hat{\mathcal{F}}_i - \mathcal{F}_i\right) - \left(\hat{\gamma}_i^{fp} - \gamma_i^{fp}\right) - \left(\hat{\Gamma}^p - \Gamma^p\right)\mathcal{F}_i$$

Since each $x_{i,Nt}$ is a linear combination of the factors and an i.i.d. process it satisfies the assumptions of Theorem 5 of An et al (1982). Then we have by that Theorem

$$\left\|\left(\hat{\Gamma}^p - \Gamma^p\right)\left(\hat{\mathcal{F}}_i - \mathcal{F}_i\right)\right\|^2 = o_p(1)\sum_{j=1}^{Np}\left(\hat{\mathcal{F}_{i,j}} - \mathcal{F}_{i,j}\right)^2$$

$$\left\|\hat{\gamma}_i^{fp} - \gamma_i^{fp}\right\|^2 = o_p\left((\ln T/T)^{1/2}\right)$$

and

$$\left\|\left(\hat{\Gamma}^p - \Gamma^p\right)\mathcal{F}_i\right\|^2 = o_p\left((\ln T/T)^{1/2}\right)$$

Hence,

$$(1 + o_p(1)\left\|\hat{\mathcal{F}}_i - \mathcal{F}_i\right\|^2 = o_p\left((\ln T/T)^{1/2}\right)$$

which implies (11) and completes the proof of the theorem. ■

Thus, divergence of N requires to be accompanied by a faster divergence of T for the SSS factor estimators to remain consistent.

3.2 Correlation in the idiosyncratic errors

In this subsection we discuss the case of cross-sectional and/or serial correlation of the idiosyncratic errors. This extension can be rather simply handled within the state space method. Basically, the idiosyncratic errors can be treated as additional pseudo-factors that enter only a few of the variables via restrictions on the matrix of loadings C. These pseudo-factors can be serially correlated processes or not depending on the matrix A in equation (1).

The problem becomes one of distinguishing common factors and pseudo-factors, i.e., cross-sectionally correlated idiosyncratic errors. This is virtually impossible for finite N, while when N diverges a common factor is one which enters an infinite number of series, i.e, the column of the, now infinite dimensional, matrix C associated with a common factor will have an infinity of non-zero entries, and likewise a pseudo-factor will only have a finite number of non-zero entries in the respective column of C. Let k_1 denote the number of common factors thus defined and k_2 the number of pseudo-factors. Note that k_2 may tend to infinity but not faster than N. Then, following Forni et al. (2000), we make the following assumption.

Assumption 6 *The matrix $\mathcal{OK}$ in (5) has k_1 singular values tending to infinity as N tends to infinity and k_2 non-zero finite singular values.*

For example, the condition in the assumption is satisfied if k_1 common factors enter a non zero fraction, bN, $0 < b < 1$, of the series x_{Nt}, in the state space model given by (1), while $k_2(N)$ pseudo-factors enter a vanishing proportion of the series x_{Nt}, i.e. each such factors enter $c(N)N$ of the series x_{Nt} where $lim_{N\to\infty}c(N)N = 0$ and $k_2(N)$ is at most equal to N.. In particular, the assumed setup from now on becomes one where the pseudo factors enter only a finite number of $x_{i,Nt}$. For simplicity, we further assume the following

Assumption 7 *For each i, $x_{i,Nt} = c_i' f_t + \epsilon_t$ where $\epsilon_t = c_i^{k_2\prime} f_t^{2,k_2} + \varepsilon_t$, $c_i^{k_2} = (c_{i,1}^{k_2}, ..., c_{i,k_2}^{k_2})'$, $f_t^{2,k_2} = (f_{1,t}^{2,k_2}, ..., f_{k_2,t}^{2,k_2})'$. f_t^{2,k_2} is the set of pseudo factors. For each i, $A_i(L) f_{i,t}^{2,k_2} = \zeta_{i,t}$ where $A_i(L)$ is a lag polynomial of, at most, order s. $\zeta_{i,t}$ and ε_t are i.i.d. processes with finite eighth moments. At most a finite number of the elements of $c_i^{k_2}$ are different from zero for all i.*

Theorem 3 ***(Consistency and Rate of Convergence when $N \to \infty$ with serially correlated idiosyncratic errors)*** *As N and T diverge, and under assumptions 1-5 and 7, $vec(\hat{f}) - vec(H^k f) = o_p(N^3 p^2 T^{-a})$ for all $a < 1/2$, where H^k is a square matrix of full rank.*

Proof. Assumption 6 is implied by assumption 7. The theorem follows if we prove that for each of the N equations in (6), (11) holds under the assumptions of the theorem. This result will follow immediately if we show that the assumptions of theorem 5 of An et al (1982) are satisfied. But this follows immediately if we note that each $x_{i,Nt}$ is a linear combination of a finite number of AR processes of finite order.and is therefore a finite AR process. ■

Remark 5 *The results of the theorem can easily be extended to accommodate stationary infinite AR process of the sort considered in An et al (1982). However, we choose not to do this as the state space representation then becomes one where the number of pseudo factors is of order larger than N and requires that s tends to infinity. However, we note that the subspace estimation as specified in this paper is still valid for such processes. Therefore, the finite order AR assumption for the idiosyncratic errors is not binding.*

3.3 Choice of the number of factors

The choice of the number of factors to be included in the model is a relevant issue, see e.g. Bai and Ng (2002). We will show that it is possible to obtain a consistent estimator of the number of factors even when N diverges or the idiosyncratic errors are correlated using an information criterion of the form

$$IC(k_1) = V(k_1, \hat{f}^{k_1}) + k_1 g(N,T) \tag{12}$$

where

$$V(k_1, \hat{f}^{k_1}) = (NT)^{-1} \sum_{t=1}^{T} tr[(x_{Nt} - \hat{C}\hat{f}_t^{k_1})(x_{Nt} - \hat{C}\hat{f}_t^{k_1})'], \tag{13}$$

$\hat{f}^{k_1} = (\hat{f}_1^{k_1}, ..., \hat{f}_T^{k_1})'$, $\hat{f}_t^{k_1}$ are the factor estimates for the k_1 first common factors (according to the singular values), $\hat{C}$ is the OLS estimate of C based on $\hat{f}_t^{k_1}$ and $g(N,T)$ is a penalty term.

Before examining the properties of this criterion, note that, since the factors are orthogonal, any set of up to k_1^0 factor estimators are consistent for the respective set of true factors up to a nonsingular transformation determined by the normalisation used in the SVD carried out during the estimation and the identification of the state space model, see SW for a similar point . Thus, denoting the $T \times k_1$ matrix of the k_1 first true factors by f^{k_1}, we have that

$$\left(T^{1/2}/N^3p^2\right) ||\hat{f}_t^{k_1} - H^{k_1'} f_t^{k_1}|| = O_p(1)$$

for some nonsingular matrix H^{k_1}. This follows from Theorem 2. Then, strengthening assumptions 3 and 5 with

Assumption 8 *u_t is an i.i.d. $(0, \Sigma_u)$ sequence with finite eighth moments.*

and

Assumption 9 $N = o\left(T^{1/7}\right)$

the following theorem holds

Theorem 4 *Let the factors be estimated by the SSS method and denote the true number of common factors k_1^0. Let $\hat{k}_1 = argmin_{1\leq k\leq kmax} IC(k_1)$. Then, under assumptions 1,2,4,7-9, $\lim_{T\to\infty} Pr(\hat{k}_1 = k_1^0) = 1$ if i) $g(N,T) \to 0$ and ii) $Ng(N,T) \to \infty$ as $N, T \to \infty$.*

Proof. The proof builds upon a set of results by Bai and Ng (2002). Therefore, to start with, we examine whether our parametric setting in terms of the representation 1 satisfies their assumptions. Assumption A of Bai and Ng (2002) is satisfied if $|\lambda_{max}(A)| < 1$, where $|\lambda_{max}(A)|$ denotes the maximum eigenvalue of A in absolute value and the fourth moments of u_t exist. These conditions are satisfied by our assumptions 1 and 3. Their Assumption B on factor loadings is straightforwardly satisfied by assuming boundedness of the elements of the C matrix. Their assumption C is satisfied by assuming that the eighth moments of u_t exist combined with our cross correlation structure in Assumption 6. Finally, their Assumption D is trivially satisfied because we assume that factors and idiosyncratic errors are uncorrelated.

We must now prove that $lim_{N(T),T\to\infty} Pr(IC(k_1) < IC(k_1^0)) = 0$ for all $k_1 \neq k_1^0$, $k_1 < k^{max}$. Denoting the $T \times k_2$ matrix of the first k_2 true idiosyncratic pseudo factors by f^{2,k_2}, we examine

$$V(k_1, (f^{k_1}, f^{2,k_2})) - V(k_1, (f^{k_1}))$$

for any finite k_2. We know that, for all elements of x_{Nt} in which f^{2,k_2} does not enter, it is

$$1/T \sum_{t=1}^{T} (x_{i,Nt} - \hat{C}'_{i,1,2}(f_t^{k_1'}, f_t^{2,k_2'})')^2 - 1/T \sum_{t=1}^{T} (x_{i,Nt} - \hat{C}'_{i,1} f_t^{k_1})^2 = O_p(T^{-1})$$

For a finite number of elements of x_{Nt}

$$1/T \sum_{t=1}^{T} (x_{i,Nt} - \hat{C}'_{i,1,2}(f_t^{k_1'}, f_t^{2,k_2'})')^2 - 1/T \sum_{t=1}^{T} (x_{i,Nt} - \hat{C}'_{i,1} f_t^{k_1})^2 = O_p(1)$$

Therefore, overall

$$V(k_1, (f^{k_1}, f^{2,k_2})) - V(k_1, (f^{k_1})) = O_p(N^{-1}) \tag{14}$$

First consider $k_1 < k_1^0$. Then

$$IC(k_1) - IC(k_1^0) = V(k_1, \hat{f}^{k_1}) - V(k_1^0, \hat{f}^{k_1^0}) - (k_1^0 - k_1)g(N,T)$$

and the required condition for the result is

$$Pr[V(k_1, \hat{f}^{k_1}) - V(k_1^0, \hat{f}^{k_1^0}) < (k_1^0 - k_1)g(N,T)] = 0$$

as $N(T), T \to \infty$. Now

$$V(k_1, \hat{f}^{k_1}) - V(k_1^0, \hat{f}^{k_1^0}) = [V(k_1, \hat{f}^{k_1}) - V(k_1, f^{k_1}H^{k_1})] + [V(k_1, f^{k_1}H^{k_1}) - V(k_1^0, f^{k_1^0}H^{k_1^0})] + \\ [V(k_1^0, f^{k_1^0}H^{k_1^0}) - V(k_1^0, \hat{f}^{k_1^0})]$$

By the rate of convergence of the factor estimators and Lemma 2 of Bai and Ng (2002) we have

$$V(k_1, \hat{f}^{k_1}) - V(k_1, f^{k_1}H^{k_1}) = O_p((T/Np)^{-1})$$

and

$$V(k_1^0, \hat{f}^{k_1^0}) - V(k_1^0, f^{k_1^0}H^{k_1^0}) = O_p((T/Np)^{-1})$$

Note that Lemma 2 of Bai and Ng (2002) stands independently from the factor estimation method discussed in that paper and only uses the rate of convergence of the factor estimators derived in their Theorem 1. Then $V(k_1, f^{k_1}H^{k_1}) - V(k_1^0, f^{k_1^0}H^{k_1^0})$ can be written as $V(k_1, f^{k_1}H^{k_1}) - V(k_1^0, f^{k_1^0})$ which has positive limit by Lemma 3 of Bai and Ng (2002). Thus, as long as $g(N,T) \to 0$, $Pr(IC(k_1) < IC(k_1^0)) = 0$ for all $k_1 < k_1^0$.

Then, to prove $Pr(IC(k_1) < IC(k_1^0)) = 0$ for all $k_1 > k_1^0$ we have to prove that

$$Pr[V(k_1^0, \hat{f}^{k_1^0}) - V(k_1, \hat{f}^{k_1}) < (k_1 - k_1^0)g(N,T)] \to 0$$

By (14) we know that asymptotically the analysis of the state space model will be equivalent to the case of a model where there are no idiosyncratic pseudo factors up to an order of probability of N^{-1}. Then

$$|V(k_1^0, \hat{f}^{k_1^0}) - V(k_1, \hat{f}^{k_1})| \leq 2max_{k_1^0 < k_1 \leq kmax}|V(k_1, \hat{f}^{k_1}) - V(k_1, f^{k_1^0})|.$$

By following the analysis of Lemma 4 of Bai and Ng (2002) we know that

$$max_{k_1^0 < k_1 \leq kmax}|V(k_1, \hat{f}^{k_1}) - V(k_1, f^{k_1^0})| = O_p((T^{1/2}/N^3p^2)^{-1})$$

Combining this expression with (14), gives the required result since, under assumption 9 then $(T^{1/2}/N^3p^2)^{-1} < N^{-1}$. Note again that Lemma 4 of Bai and Ng (2002) stands independently from the factor estimation method discussed in that paper and only uses the rate of convergence of the factor estimators derived in their Theorem 1. ■

4 A comparison of the estimation methods

In this section we summarize the results of an extensive set of simulation experiments to investigate the small sample properties of our factor estimation method and compare it with existing alternative approaches, i.e. static principal components (PCA, SW) as in Stock and Watson (2002a, 2002b) and dynamic principal components (DPCA, FHLR) as in Forni et al. (2000). The first subsection describes the simulation set-up; the second one the results.

4.1 Monte Carlo experiments, set-up

The basic data generating process (DGP) we use is:

$$\begin{aligned} x_{Nt} &= Cf_t + \epsilon_t, \quad t = 1, \ldots, T \\ A(L)f_t &= B(L)u_t \end{aligned} \tag{15}$$

where $A(L) = I - A_1(L) - \ldots - A_p(L)$, $B(L) = I + B_1(L) + \ldots + B_q(L)$.

An important comment is in order for this model. We have developed our theory for predetermined factors, i.e. factors that are determined at time $t-1$. This is reflected by (1) where the error term of the factor equation is dated at time $t-1$. As mentioned, this assumption is not considered restrictive in the state space model literature, see e.g. Deistler and Hannan (1988). Yet, the specification we use for the simulations allows for factors that are determined at time t. This brings us in line with the nonparametric context of SW and FHLR. However, as the simulations will show, this choice still leaves the new estimation method performing comparably and, in a majority of cases, better than either PCA or DPCA. The rationale underlying this results is that the SSS estimator, when contemporaneous errors drive the factors, is consistent for the expected value of the factors conditional on information up to period $t-1$. Of course, the performance of the SSS estimator further improves when u_{t-1} is used in (15) rather than u_t.

For the SSS method, the "lag" truncation parameter is set at $p = \ln(T)^\alpha$. We have found that a range of α between 1.05 and 1.5 provides a satisfactory performance, and we have used the value $\alpha = 1.25$ in the reported results.

The "lead" truncation parameter s is set equal to the assumed number of factors for SSS, which typically coincides with the true number of factors, i.e. $s = k$. For robustness, and since it is relevant for forecasting, we will present selected result for the case $s = 1$ as well.[3] For the DPCA method we use 3 leads and 3 lags.

[3] We have also experimented with other values of s but $s = 1$ or $s = k$ appear to be the preferable choices. To select the value of s we can either include this parameter as a variable in the information criterion search or, perhaps more straightforwardly, we can choose the value that maximises the proportion of the variance of each series explained by the factors, averaged over all series.

With the exceptions noted below, the C matrix is generated using standard normal variates as elements and the error terms are generated as uncorrelated standard normal pseudo-random variables. We have considered several combinations of N, T and report results for the following N, T pairs: $(50, 50)$, $(50, 100)$, $(100, 50)$, $(100, 100)$, $(50, 500)$, $(100, 500)$ and $(200, 50)$.

To provide a comprehensive evaluation of the relative performance of the three factor estimation methods, we consider several types of experiments. They differ for the number of factors (one or several), the choice of s ($s = k$ or $s = 1$), the factor loadings (static or dynamic), the choice of the number of factors (true number or misspecified), the properties of the idiosyncratic errors (uncorrelated or serially correlated), and the way the C matrix is generated (standard normal or uniform with non-zero mean). Each experiment is replicated 500 times. Depending on these characteristics, the experiments can be divided into five groups.

In the first group, we assume that we have a single VARMA factor with 8 specifications that differ for the extent of serial correlation and the AR and MA order:

(1) $a_1 = 0.2, b_1 = 0.4$;

(2) $a_1 = 0.7, b_1 = 0.2$;

(3) $a_1 = 0.3, a_2 = 0.1, b_1 = 0.15, b_2 = 0.15$;

(4) $a_1 = 0.5, a_2 = 0.3, b_1 = 0.2, b_2 = 0.2$;

(5) $a_1 = 0.2, b_1 = -0.4$;

(6) $a_1 = 0.7, b_1 = -0.2$;

(7) $a_1 = 0.3, a_2 = 0.1, b_1 = -0.15, b_2 = -0.15$;

(8) $a_1 = 0.5, a_2 = 0.3, b_1 = -0.2, b_2 = -0.2$.

Experiment 9 is as experiment 1 but both the ARMA factor and its lag enter the measurement equation, i.e., the C matrix is $C(L) = C_0 + C_1 L$ where L is the lag operator. We fix a priori the number of factors to $p + q$, which is the true number in the state space representation. It is larger than the true number in the FHLR setup, and it should provide a reasonable approximation for SW too. As a robustness check, we consider the case where the factor is generated as in Experiment 1 but only one factor is assumed to exist rather than $p + q$. We refer to this experiment as Experiment 10. In the case of experiments 9 and 10, qualitatively similar results are obtained when the mentioned modifications are applied to the parameter specifications 2-8 (results available upon request).

In the second group of experiments, we investigate the case of serially correlated idiosyncratic errors. The DGP for that is specified as in experiments 1-10 but with each idiosyncratic error being an AR(1) process with coefficient 0.2 rather than an i.i.d. process. These experiments are labelled 11-20. The results are rather robust to higher values of serial correlation but 0.2 is a reasonable value in practice since usually the common component captures most of the persistence of the series. We have also investigated the case of cross-correlated errors by assuming that the contemporaneous covariance matrix of the idiosyncratic errors is tridiagonal with diagonal elements equal to 1 and off-diagonal elements equal to 0.2. These experiments produced the same ranking of methods as in the case of serial correlation and virtually no deterioration of performance with respect to the idiosyncratic errors case (results available upon request).

In the third group of experiments, we use a 3 dimensional VAR(1) as the data generation process for the factors as opposed to an ARMA process. We report results for the case where the A matrix is diagonal with elements equal to 0.5. This is labelled experiment 21.

In the fourth group of experiments, we consider the DGPs in experiments 1-21 but generate the C matrix using standard uniform variates, thereby allowing for the factor loadings to have a non zero mean. To save space, we only report results for $(N, T) =$ $(50, 50)$ for this case.

Finally, we consider again experiments 1-21 but using $s = 1$ instead of $s = k$. We present results for the (N, T) pairs $(50, 50)$ and $(100, 100)$.

We concentrate on the relationship between the true and estimated common components (Cf_t and $\widehat{C}\widehat{f_t}$), measured by their correlation, and on the properties of the estimated idiosyncratic components ($\widehat{\epsilon}_t$), using an LM(4) test to evaluate whether they are white noise as in the DGP, and presenting the rejection probabilities of the test. These are the most common evaluation criteria used in the literature. Throughout, we report the average values of the different evaluation criteria (averaging over all variables for each replication and then over all replications), and the standard errors of the averages over replications.

4.2 Monte Carlo experiments, results

The results are summarized in Tables 1 to 7 for different combinations of N and T, while Table 8 presents the outcome for the uniform factor loadings C and $(N, T) = (50, 50)$. Finally, Tables 9-11 present results for the case $s = 1$.

Starting with the $(N, T) = (50, 50)$ case in Table 1, and the single ARMA factor experiments (1-8), the SSS method clearly outperforms the other two. The gains with respect to PCA are rather limited, in the range 5-10%, but systematic across experiments. The gains are larger with respect to DPCA, about 20%, and again systematic across experiments. For all the three methods the correlation is higher the higher the persistence of the factor. There is little evidence that the idiosyncratic component is serially correlated on the basis of the

LM(4) test for any of the methods, but the DPCA yields systematically larger rejection probabilities.

The presence of serially correlated idiosyncratic errors (experiments 11-18) does not affect significantly the results. The values for each method, the ranking of the methods and the relative gains are virtually the same as in the basic case. Non correlation of the errors is rejected more often, but still in a very low number of cases. This is related to the low power of the LM test in small (T) samples, for larger values of T the rejection rate increases substantially, see Tables 2 and 3.

Allowing for a lagged effect of the factor on the variables, instead, leads to a serious deterioration of the SSS performance, with a drop of about 25% in the correlation values, compare experiments 1 and 9, and 11 and 19. The performance of DPCA, which is particularly suited for this generating process from a theoretical point of view, does improve, but it is still beaten by PCA even though the difference shrinks. The choice of a lower value for s improves substantially the performance of SSS in this case, making it comparable with PCA, compare the relevant lines of Table 9 for $s = 1$. This finding, combined with the fact that DPCA is still beaten by PCA, suggests that the use of leads of the variables for factor estimation is complicated when the factors can have a dynamic impact on the variables.

When a lower number of factors than true is assumed for SSS, one instead of two in experiments 10 and 20, the performance does not deteriorate. Actually, comparing experiments 1 and 10, and 11 and 20, there is a slight increase in correlation. A similar improvement can be observed for PCA and DPCA, and it is likely due to the fact that a single factor can do most of the work of capturing the true common component, while estimation uncertainty is reduced.

The presence of three autoregressive factors, experiment 21, reduces the gap PCA-DPCA. The correlation values are higher than in the single factor case, reflecting in general the higher persistence of the factors. Yet, the performance of SSS deteriorates substantially. The latter improves and becomes comparable to PCA with $s = 1$, see table 11.

The next three issues we consider are the effects of larger temporal dimension, cross-sectional dimension, and uniform rather than standard normal loading matrix.

Tables 2 and 3 report results for $N = 50$ and, respectively, $T = 100$ and $T = 500$. The correlation between the true and estimated common component increases monotonically for all the three methods, but neither the ranking of methods nor the performance across experiments are affected. The performance of the LM tests in detecting serial correlation in the error process gets also closer and closer to the theoretical one.

When N increase to 100 while T remains equal to 50 (Table 4), the figures for SSS are basically unchanged in all experiments, while the performance of PCA and DPCA improves systematically. Yet, the gains are not sufficient to match the SSS approach, which still yields the highest correlation in all cases, except with a dynamic effect of the factors of the variables (experiments 9 and 19), and with three autoregressive factors (experiment 21). This pattern continues if we further increase N to 200 (Table 7).

When both N and T increase, $N = 100, T = 100$ in Table 5 while $N = 100, T = 500$ in Table 6, the performance of all methods improves with respect to Table 1, proportionally more so for PCA and DPCA that benefit more for the larger value of N, as mentioned before. But also in these cases SSS is in general the best in terms of correlation.

The final issue we consider is the choice of s. This is examined through Tables 9-11 where we set $s = 1$. For this case PCA and SSS perform very similarly. The advantage SSS had for the ARMA experiments shrinks substantially, SSS is still better but only marginally so. On the other hand, the large disadvantage SSS had for VAR experiments and experiments with factor lags disappears, as mentioned above, with SSS and PCA performing equally well.

In summary, the DPCA method shows consistently lower correlation between true and estimated common components than SSS and PCA. It shows, in general, more evidence of serial correlation, although not to any significant extent. Additionally, from results we are not presenting here the DPCA method has the lowest variance for the idiosyncratic component or, in other words, has the highest explanatory power of the series in terms of the common components. These results seem to indicate that i) part of the idiosyncratic component seems to leak into the estimated common component in the DPCA case, thus reducing the correlation between true and estimated common components and the variance of the idiosyncratic component and ii) some (smaller in terms of variance) part of the common component leaks into the estimated idiosyncratic component thus increasing the serial correlation of the idiosyncratic component. The conclusion from these results is that if one cares about isolating common components as summaries of underlying common features of the data, then a high R^2 may not always be the appropriate guide. When instead the factors have a dynamic effect on the variables, the performance of DPCA improves, but it is still beaten by PCA. This experiment and the one with three autoregressive factors are the only cases where PCA beats SSS, but the difference can be annihilated by means of a proper choice of the s parameter. In all other experiments SSS leads to gains in terms of higher correlation in the range 5-10%.

5 Conclusion

In this paper we have developed a parametric estimation method for dynamic factor models of large dimension based on a subspace algorithm applied to the state space representation of the model. We have proved consistency of the estimators, also in the case $N \to \infty$ and with correlated idiosyncratic errors. We have also proposed information criteria for a consistent selection of the number of factors. Finally, we have shown that the method performs well compared with existing nonparametric estimators.

References

[1] An, H. Z., Z. Chen and E. J. Hannan (1982), "Autocorrelation, autoregression and autoregressive approximation", *Annals of Statistics*, 10, 926-936.

[2] Bai, J. and S. Ng (2002), "Determining the number of factors in approximate factor models", *Econometrica*, 70, 191-223.

[3] Bauer, D. (1998), "Some Asymptotic Theory for the Estimation of Linear Systems Using Maximum Likelihood Methods or Subspace Algorithms", Ph.d. Thesis.

[4] Berk K. N. (1974), "Consistent Autoregressive Spectral Estimates" , *Annals of Statistics*, 2(3) , 489–502.

[5] Brillinger D.R. (1981), Time Series: Data Analysis and Theory, *McGraw-Hill*.

[6] Camba-Mendez G., G. Kapetanios, R. J. Smith and M. R. Weale (2001), "An Automatic Leading Indicator of Economic Activity: Forecasting GDP Growth for European Countries", *Econometrics Journal,* 4(1), S56-90.

[7] Chatelin (1983), Spectral Approximation of Linear Operators, *Academic Press*

[8] Chamberlain, G. and M. Rothschild (1983), "Arbitrage factor structure, and mean variance analysis of large asset markets", *Econometrica*, 51, 1281-1304.

[9] Connor, G. and R.A. Korajczyk (1986), "Performance measurement with the arbitrage pricing theory", *Journal of Financial Economics*, 15, 373-394.

[10] Connor, G. and R.A. Korajczyk (1993), "A test for the number of factors in an approximate factor model", *Journal of Finance*, 48, 1263-1291.

[11] Deistler M. and E.J. Hannan (1988), "The Statistical Analysis of Linear Systems", John Wiley.

[12] Favero, C., Marcellino, M. and Neglia, F. (2002), "Principal components at work: the empirical analysis of monetary policy with large datasets", IGIER WP 223

[13] Forni, M. and L. Reichlin (1996), "Dynamic common factors in large cross-sections", *Empirical Economics*, 21, 27-42.

[14] Forni, M. and L. Reichlin (1997), "National policies and local economies: Europe and the United States", CEPR WP 1632.

[15] Forni, M. and L. Reichlin (1998), "Let's get real: A dynamic factor analytical approach to disaggregated business cycle", *Review of Economic Studies*, 65, 453-474.

[16] Forni, M., Hallin, M., Lippi, M. and L. Reichlin (1999), "Reference cycles: The NBER methodology revisited", manuscript.

[17] Forni, M., Hallin, M., Lippi, M. and L. Reichlin (2000), "The generalised factor model: identification and estimation", *The Review of Economic and Statistics*, 82, 540-554.

[18] Forni, M., Hallin, M., Lippi, M. and L. Reichlin (2003), "The Generalized Dynamic Factor Model one-sided estimation and forecasting", mimeo.

[19] Geweke, J. (1977), "The dynamic factor analysis of economic time series", ch. 19 in Aigner, D.J. and A.S. Goldberger (eds.), *Latent variables in socio-economic models*, Amsterdam: North Holland.

[20] Hannan, E. J. and L. Kavalieris (1986), "Regression, autoregression models", *Journal of Time Series Analysis*, 7, 27-49.

[21] Kapetanios, G. and M. Marcellino (2004), "A comparison of estimation methods for dynamic factor models of large dimension", Queen Mary Working Paper No. 489.

[22] Larimore, W.E. (1983), "System Identification, Reduced Order Filters and Modelling via Canonical Variate Analysis", Proc. 1983 Amer. Control Conference, pp. 445-451.

[23] Magnus, J. and H. Neudecker (1988), "Matrix Differential Calculus with Applications to Statistics and Econometrics", John Wiley.

[24] Van Overschee, P. and B. De Moor (1996), "Subspace Identification for Linear Systems", Kluwer Academic Publishers.

[25] Quah, D. and T.J. Sargent (1993), "A dynamic index model for large cross-sections", ch. 7 in Stock, J.H. and M.W. Watson (eds.), *Business cycles, indicators and forecasting*, University of Chicago Press for the NBER.

[26] Sargent, T.J. and C.A. Sims (1977), "Business cycle modelling without pretending to have too much a-priori economic theory", in Sims, C.A. (ed.), *New methods in business cycle research*, Minneapolis: Federal Reserve Bank of Minneapolis.

[27] Sin, C. Y. and H. White (1996), "Information criteria for selecting possibly misspecied parametric models", *Journal of Econometrics,* 71, 207-225.

[28] Stock, J.H. and M.W. Watson (1989), "New indexes of coincident and leading economic indicators", *NBER Macroeconomic Annual*, 351-393.

[29] Stock, J.H. and M.W. Watson (1991), "A probability model of the coincident economic indicators", ch.4 in Lahiri, K. and G.H. Moore (eds.), *Leading economic indicators: New approaches and forecasting records*, New York: Cambridge University Press.

[30] Stock, J.H. and M.W. Watson (2001), "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business and Economic Statistics*, 20, 147-62

[31] Stock, J.H. and M.W. Watson (2002), "Forecasting Using Principal Components from a Large Number of Predictors", *Journal of the American Statistical Association*, 97, 1167–1179.

Table 1: Results for case: N=50, T=50

Exp. [a]	Corr. with True[b]			Serial Correlation[c]		
	PCA	SSS	DPCA	PCA	SSS	DPCA
Exp 1	$0.821_{(0.052)}$	$0.860_{(0.054)}$	$0.727_{(0.053)}$	$0.067_{(0.033)}$	$0.066_{(0.035)}$	$0.097_{(0.042)}$
Exp 2	$0.859_{(0.049)}$	$0.890_{(0.050)}$	$0.780_{(0.056)}$	$0.072_{(0.040)}$	$0.075_{(0.039)}$	$0.103_{(0.045)}$
Exp 3	$0.740_{(0.054)}$	$0.805_{(0.054)}$	$0.634_{(0.056)}$	$0.073_{(0.036)}$	$0.081_{(0.040)}$	$0.137_{(0.052)}$
Exp 4	$0.803_{(0.058)}$	$0.855_{(0.054)}$	$0.713_{(0.068)}$	$0.076_{(0.040)}$	$0.086_{(0.038)}$	$0.143_{(0.054)}$
Exp 5	$0.806_{(0.053)}$	$0.848_{(0.055)}$	$0.703_{(0.052)}$	$0.067_{(0.034)}$	$0.066_{(0.034)}$	$0.094_{(0.042)}$
Exp 6	$0.823_{(0.053)}$	$0.861_{(0.053)}$	$0.731_{(0.055)}$	$0.068_{(0.035)}$	$0.070_{(0.038)}$	$0.103_{(0.042)}$
Exp 7	$0.717_{(0.053)}$	$0.787_{(0.054)}$	$0.604_{(0.052)}$	$0.064_{(0.034)}$	$0.076_{(0.038)}$	$0.135_{(0.049)}$
Exp 8	$0.724_{(0.057)}$	$0.791_{(0.058)}$	$0.616_{(0.057)}$	$0.067_{(0.035)}$	$0.080_{(0.038)}$	$0.137_{(0.053)}$
Exp 9	$0.898_{(0.028)}$	$0.693_{(0.061)}$	$0.823_{(0.036)}$	$0.071_{(0.036)}$	$0.039_{(0.030)}$	$0.123_{(0.049)}$
Exp 10	$0.904_{(0.061)}$	$0.904_{(0.060)}$	$0.848_{(0.050)}$	$0.068_{(0.037)}$	$0.068_{(0.036)}$	$0.079_{(0.039)}$
Exp 11	$0.813_{(0.055)}$	$0.855_{(0.055)}$	$0.721_{(0.052)}$	$0.102_{(0.043)}$	$0.116_{(0.045)}$	$0.132_{(0.050)}$
Exp 12	$0.848_{(0.051)}$	$0.881_{(0.052)}$	$0.772_{(0.056)}$	$0.100_{(0.042)}$	$0.112_{(0.045)}$	$0.132_{(0.050)}$
Exp 13	$0.722_{(0.058)}$	$0.789_{(0.058)}$	$0.620_{(0.059)}$	$0.084_{(0.037)}$	$0.123_{(0.045)}$	$0.155_{(0.053)}$
Exp 14	$0.791_{(0.060)}$	$0.846_{(0.055)}$	$0.704_{(0.068)}$	$0.089_{(0.040)}$	$0.123_{(0.049)}$	$0.162_{(0.056)}$
Exp 15	$0.798_{(0.055)}$	$0.845_{(0.057)}$	$0.697_{(0.053)}$	$0.113_{(0.045)}$	$0.130_{(0.049)}$	$0.150_{(0.051)}$
Exp 16	$0.813_{(0.055)}$	$0.854_{(0.056)}$	$0.724_{(0.055)}$	$0.105_{(0.043)}$	$0.118_{(0.046)}$	$0.143_{(0.050)}$
Exp 17	$0.703_{(0.055)}$	$0.776_{(0.058)}$	$0.596_{(0.053)}$	$0.082_{(0.039)}$	$0.125_{(0.047)}$	$0.157_{(0.056)}$
Exp 18	$0.715_{(0.057)}$	$0.785_{(0.059)}$	$0.610_{(0.057)}$	$0.082_{(0.039)}$	$0.127_{(0.048)}$	$0.165_{(0.058)}$
Exp 19	$0.889_{(0.031)}$	$0.685_{(0.063)}$	$0.814_{(0.037)}$	$0.086_{(0.039)}$	$0.052_{(0.032)}$	$0.138_{(0.049)}$
Exp 20	$0.892_{(0.064)}$	$0.893_{(0.063)}$	$0.840_{(0.053)}$	$0.119_{(0.047)}$	$0.120_{(0.047)}$	$0.128_{(0.050)}$
Exp 21	$0.974_{(0.009)}$	$0.692_{(0.051)}$	$0.947_{(0.014)}$	$0.078_{(0.038)}$	$0.111_{(0.068)}$	$0.125_{(0.046)}$

[a] PCA: Principal Component Estimation Method; DPCA: Dynamic Principal Component Estimation Method; SSS: Subspace algorithm on state space form. Exp. 1-8 : one factor, different ARMA DGP, no correlation among idiosyncratic components; Exp 9: as Exp. 1 but dynamic impact on variables; Exp 10: as Exp. 1 but one factor imposed in estimation rather than p+q; Exp. 11-20: as 1-10 but temporal correlation among idiosyncratic components; Exp. 21: three AR factors (non correlated), no correlation among idiosyncratic components.

[b] Mean Correlation between true and estimated common component, with MC st.dev. in ().

[c] Mean rejection rate of LM serial correlation test of idiosyncratic component, with MC st.dev. in ().

Table 2: Results for case: N=50, T=100

Exp. [a]	Corr. with True[b]			Serial Correlation[c]		
	PCA	SSS	DPCA	PCA	SSS	DPCA
Exp 1	$0.856_{(0.044)}$	$0.903_{(0.045)}$	$0.781_{(0.044)}$	$0.057_{(0.033)}$	$0.057_{(0.032)}$	$0.068_{(0.036)}$
Exp 2	$0.890_{(0.041)}$	$0.928_{(0.039)}$	$0.830_{(0.045)}$	$0.060_{(0.034)}$	$0.061_{(0.033)}$	$0.073_{(0.036)}$
Exp 3	$0.777_{(0.044)}$	$0.862_{(0.042)}$	$0.689_{(0.045)}$	$0.057_{(0.034)}$	$0.064_{(0.036)}$	$0.086_{(0.040)}$
Exp 4	$0.844_{(0.044)}$	$0.906_{(0.038)}$	$0.776_{(0.052)}$	$0.061_{(0.034)}$	$0.068_{(0.035)}$	$0.086_{(0.040)}$
Exp 5	$0.839_{(0.043)}$	$0.891_{(0.045)}$	$0.754_{(0.043)}$	$0.056_{(0.034)}$	$0.056_{(0.033)}$	$0.069_{(0.038)}$
Exp 6	$0.859_{(0.043)}$	$0.904_{(0.044)}$	$0.785_{(0.044)}$	$0.057_{(0.033)}$	$0.058_{(0.035)}$	$0.070_{(0.036)}$
Exp 7	$0.752_{(0.044)}$	$0.847_{(0.044)}$	$0.658_{(0.045)}$	$0.056_{(0.032)}$	$0.061_{(0.033)}$	$0.084_{(0.039)}$
Exp 8	$0.767_{(0.046)}$	$0.855_{(0.045)}$	$0.677_{(0.049)}$	$0.057_{(0.032)}$	$0.064_{(0.034)}$	$0.088_{(0.041)}$
Exp 9	$0.923_{(0.021)}$	$0.703_{(0.055)}$	$0.869_{(0.026)}$	$0.061_{(0.034)}$	$0.028_{(0.025)}$	$0.081_{(0.039)}$
Exp 10	$0.935_{(0.047)}$	$0.935_{(0.047)}$	$0.894_{(0.040)}$	$0.056_{(0.032)}$	$0.057_{(0.032)}$	$0.061_{(0.033)}$
Exp 11	$0.849_{(0.043)}$	$0.898_{(0.043)}$	$0.776_{(0.043)}$	$0.212_{(0.060)}$	$0.242_{(0.061)}$	$0.235_{(0.061)}$
Exp 12	$0.888_{(0.039)}$	$0.926_{(0.038)}$	$0.830_{(0.041)}$	$0.204_{(0.057)}$	$0.229_{(0.058)}$	$0.226_{(0.059)}$
Exp 13	$0.770_{(0.045)}$	$0.859_{(0.043)}$	$0.686_{(0.048)}$	$0.157_{(0.051)}$	$0.240_{(0.062)}$	$0.228_{(0.059)}$
Exp 14	$0.836_{(0.042)}$	$0.902_{(0.037)}$	$0.771_{(0.050)}$	$0.157_{(0.050)}$	$0.233_{(0.060)}$	$0.221_{(0.058)}$
Exp 15	$0.836_{(0.041)}$	$0.890_{(0.042)}$	$0.753_{(0.041)}$	$0.232_{(0.061)}$	$0.263_{(0.064)}$	$0.263_{(0.062)}$
Exp 16	$0.853_{(0.043)}$	$0.900_{(0.045)}$	$0.782_{(0.044)}$	$0.208_{(0.060)}$	$0.239_{(0.064)}$	$0.239_{(0.064)}$
Exp 17	$0.743_{(0.043)}$	$0.840_{(0.042)}$	$0.652_{(0.044)}$	$0.167_{(0.053)}$	$0.245_{(0.062)}$	$0.229_{(0.064)}$
Exp 18	$0.764_{(0.046)}$	$0.853_{(0.045)}$	$0.677_{(0.049)}$	$0.162_{(0.054)}$	$0.246_{(0.062)}$	$0.230_{(0.061)}$
Exp 19	$0.916_{(0.022)}$	$0.695_{(0.050)}$	$0.862_{(0.027)}$	$0.183_{(0.055)}$	$0.097_{(0.042)}$	$0.220_{(0.058)}$
Exp 20	$0.931_{(0.049)}$	$0.932_{(0.049)}$	$0.889_{(0.041)}$	$0.244_{(0.062)}$	$0.245_{(0.061)}$	$0.250_{(0.062)}$
Exp 21	$0.984_{(0.005)}$	$0.686_{(0.040)}$	$0.970_{(0.007)}$	$0.062_{(0.033)}$	$0.205_{(0.100)}$	$0.083_{(0.038)}$

[a] PCA: Principal Component Estimation Method; DPCA: Dynamic Principal Component Estimation Method; SSS: Subspace algorithm on state space form. Exp. 1-8 : one factor, different ARMA DGP, no correlation among idiosyncratic components; Exp 9: as Exp. 1 but dynamic impact on variables; Exp 10: as Exp. 1 but one factor imposed in estimation rather than p+q; Exp. 11-20: as 1-10 but temporal correlation among idiosyncratic components; Exp. 21: three AR factors (non correlated), no correlation among idiosyncratic components.

[b] Mean Correlation between true and estimated common component, with MC st.dev. in ().

[c] Mean rejection rate of LM serial correlation test of idiosyncratic component, with MC st.dev. in ().

Table 3: Results for case: N=50, T=500

Exp. [a]	Corr. with True[b]			Serial Correlation[c]		
	PCA	SSS	DPCA	PCA	SSS	DPCA
Exp 1	$0.899_{(0.028)}$	$0.939_{(0.042)}$	$0.855_{(0.031)}$	$0.052_{(0.030)}$	$0.062_{(0.043)}$	$0.058_{(0.032)}$
Exp 2	$0.922_{(0.027)}$	$0.951_{(0.036)}$	$0.889_{(0.030)}$	$0.050_{(0.031)}$	$0.064_{(0.044)}$	$0.056_{(0.034)}$
Exp 3	$0.822_{(0.033)}$	$0.907_{(0.049)}$	$0.773_{(0.036)}$	$0.056_{(0.033)}$	$0.088_{(0.075)}$	$0.066_{(0.033)}$
Exp 4	$0.885_{(0.030)}$	$0.946_{(0.026)}$	$0.851_{(0.034)}$	$0.051_{(0.031)}$	$0.083_{(0.059)}$	$0.064_{(0.036)}$
Exp 5	$0.881_{(0.033)}$	$0.937_{(0.039)}$	$0.830_{(0.035)}$	$0.050_{(0.031)}$	$0.055_{(0.036)}$	$0.056_{(0.032)}$
Exp 6	$0.900_{(0.030)}$	$0.943_{(0.039)}$	$0.857_{(0.033)}$	$0.052_{(0.031)}$	$0.059_{(0.043)}$	$0.056_{(0.031)}$
Exp 7	$0.803_{(0.036)}$	$0.904_{(0.055)}$	$0.749_{(0.039)}$	$0.051_{(0.029)}$	$0.071_{(0.067)}$	$0.062_{(0.035)}$
Exp 8	$0.822_{(0.037)}$	$0.914_{(0.049)}$	$0.773_{(0.039)}$	$0.052_{(0.033)}$	$0.077_{(0.070)}$	$0.065_{(0.035)}$
Exp 9	$0.946_{(0.014)}$	$0.718_{(0.055)}$	$0.924_{(0.017)}$	$0.050_{(0.031)}$	$0.122_{(0.143)}$	$0.058_{(0.033)}$
Exp 10	$0.967_{(0.031)}$	$0.966_{(0.032)}$	$0.948_{(0.026)}$	$0.052_{(0.031)}$	$0.052_{(0.031)}$	$0.053_{(0.031)}$
Exp 11	$0.893_{(0.030)}$	$0.941_{(0.044)}$	$0.851_{(0.033)}$	$0.945_{(0.032)}$	$0.945_{(0.040)}$	$0.950_{(0.030)}$
Exp 12	$0.920_{(0.026)}$	$0.954_{(0.032)}$	$0.889_{(0.028)}$	$0.944_{(0.032)}$	$0.937_{(0.043)}$	$0.949_{(0.030)}$
Exp 13	$0.820_{(0.037)}$	$0.914_{(0.043)}$	$0.772_{(0.040)}$	$0.924_{(0.038)}$	$0.933_{(0.054)}$	$0.941_{(0.034)}$
Exp 14	$0.879_{(0.031)}$	$0.944_{(0.030)}$	$0.846_{(0.034)}$	$0.922_{(0.038)}$	$0.920_{(0.062)}$	$0.940_{(0.036)}$
Exp 15	$0.883_{(0.031)}$	$0.937_{(0.042)}$	$0.834_{(0.034)}$	$0.950_{(0.031)}$	$0.954_{(0.031)}$	$0.956_{(0.030)}$
Exp 16	$0.897_{(0.029)}$	$0.943_{(0.048)}$	$0.856_{(0.031)}$	$0.942_{(0.034)}$	$0.940_{(0.052)}$	$0.950_{(0.031)}$
Exp 17	$0.793_{(0.036)}$	$0.901_{(0.051)}$	$0.740_{(0.038)}$	$0.925_{(0.038)}$	$0.943_{(0.046)}$	$0.943_{(0.033)}$
Exp 18	$0.817_{(0.035)}$	$0.911_{(0.049)}$	$0.769_{(0.038)}$	$0.926_{(0.037)}$	$0.940_{(0.053)}$	$0.942_{(0.032)}$
Exp 19	$0.945_{(0.015)}$	$0.721_{(0.052)}$	$0.922_{(0.018)}$	$0.932_{(0.036)}$	$0.662_{(0.176)}$	$0.940_{(0.034)}$
Exp 20	$0.965_{(0.036)}$	$0.961_{(0.051)}$	$0.945_{(0.030)}$	$0.956_{(0.029)}$	$0.956_{(0.029)}$	$0.955_{(0.029)}$
Exp 21	$0.991_{(0.001)}$	$0.609_{(0.030)}$	$0.988_{(0.002)}$	$0.053_{(0.031)}$	$0.569_{(0.117)}$	$0.058_{(0.033)}$

[a] PCA: Principal Component Estimation Method; DPCA: Dynamic Principal Component Estimation Method; SSS: Subspace algorithm on state space form. Exp. 1-8 : one factor, different ARMA DGP, no correlation among idiosyncratic components; Exp 9: as Exp. 1 but dynamic impact on variables; Exp 10: as Exp. 1 but one factor imposed in estimation rather than p+q; Exp. 11-20: as 1-10 but temporal correlation among idiosyncratic components; Exp. 21: three AR factors (non correlated), no correlation among idiosyncratic components.

[b] Mean Correlation between true and estimated common component, with MC st.dev. in ().

[c] Mean rejection rate of LM serial correlation test of idiosyncratic component, with MC st.dev. in ().

Table 4: Results for case: N=100, T=50

Exp. [a]	Corr. with True[b]			Serial Correlation[c]		
	PCA	SSS	DPCA	PCA	SSS	DPCA
Exp 1	$0.841_{(0.038)}$	$0.868_{(0.038)}$	$0.740_{(0.040)}$	$0.069_{(0.026)}$	$0.069_{(0.026)}$	$0.102_{(0.032)}$
Exp 2	$0.871_{(0.036)}$	$0.895_{(0.034)}$	$0.790_{(0.041)}$	$0.072_{(0.026)}$	$0.073_{(0.027)}$	$0.108_{(0.031)}$
Exp 3	$0.758_{(0.044)}$	$0.806_{(0.042)}$	$0.639_{(0.048)}$	$0.070_{(0.027)}$	$0.079_{(0.027)}$	$0.156_{(0.041)}$
Exp 4	$0.818_{(0.052)}$	$0.856_{(0.047)}$	$0.721_{(0.063)}$	$0.078_{(0.029)}$	$0.088_{(0.027)}$	$0.163_{(0.042)}$
Exp 5	$0.821_{(0.038)}$	$0.852_{(0.039)}$	$0.713_{(0.039)}$	$0.063_{(0.025)}$	$0.068_{(0.025)}$	$0.096_{(0.030)}$
Exp 6	$0.836_{(0.041)}$	$0.863_{(0.040)}$	$0.736_{(0.044)}$	$0.072_{(0.026)}$	$0.073_{(0.026)}$	$0.108_{(0.032)}$
Exp 7	$0.734_{(0.040)}$	$0.786_{(0.040)}$	$0.609_{(0.041)}$	$0.068_{(0.025)}$	$0.077_{(0.029)}$	$0.149_{(0.039)}$
Exp 8	$0.749_{(0.042)}$	$0.798_{(0.041)}$	$0.629_{(0.045)}$	$0.069_{(0.025)}$	$0.081_{(0.028)}$	$0.156_{(0.040)}$
Exp 9	$0.912_{(0.022)}$	$0.696_{(0.058)}$	$0.833_{(0.032)}$	$0.071_{(0.026)}$	$0.036_{(0.021)}$	$0.130_{(0.036)}$
Exp 10	$0.904_{(0.043)}$	$0.904_{(0.043)}$	$0.852_{(0.037)}$	$0.065_{(0.026)}$	$0.065_{(0.026)}$	$0.075_{(0.027)}$
Exp 11	$0.829_{(0.039)}$	$0.859_{(0.039)}$	$0.736_{(0.041)}$	$0.102_{(0.031)}$	$0.115_{(0.034)}$	$0.135_{(0.035)}$
Exp 12	$0.855_{(0.042)}$	$0.880_{(0.041)}$	$0.776_{(0.047)}$	$0.104_{(0.030)}$	$0.112_{(0.033)}$	$0.137_{(0.035)}$
Exp 13	$0.746_{(0.044)}$	$0.800_{(0.042)}$	$0.634_{(0.046)}$	$0.084_{(0.028)}$	$0.119_{(0.034)}$	$0.172_{(0.044)}$
Exp 14	$0.805_{(0.049)}$	$0.847_{(0.044)}$	$0.712_{(0.060)}$	$0.093_{(0.029)}$	$0.124_{(0.034)}$	$0.179_{(0.043)}$
Exp 15	$0.817_{(0.039)}$	$0.853_{(0.040)}$	$0.713_{(0.039)}$	$0.109_{(0.032)}$	$0.128_{(0.034)}$	$0.152_{(0.038)}$
Exp 16	$0.825_{(0.043)}$	$0.857_{(0.043)}$	$0.731_{(0.046)}$	$0.101_{(0.031)}$	$0.118_{(0.032)}$	$0.146_{(0.037)}$
Exp 17	$0.721_{(0.043)}$	$0.780_{(0.043)}$	$0.602_{(0.044)}$	$0.085_{(0.029)}$	$0.122_{(0.034)}$	$0.171_{(0.043)}$
Exp 18	$0.735_{(0.045)}$	$0.790_{(0.045)}$	$0.620_{(0.048)}$	$0.088_{(0.030)}$	$0.124_{(0.032)}$	$0.176_{(0.044)}$
Exp 19	$0.904_{(0.024)}$	$0.686_{(0.055)}$	$0.826_{(0.032)}$	$0.088_{(0.030)}$	$0.050_{(0.023)}$	$0.148_{(0.039)}$
Exp 20	$0.902_{(0.046)}$	$0.902_{(0.047)}$	$0.847_{(0.039)}$	$0.117_{(0.034)}$	$0.117_{(0.034)}$	$0.125_{(0.036)}$
Exp 21	$0.979_{(0.006)}$	$0.696_{(0.048)}$	$0.952_{(0.010)}$	$0.076_{(0.028)}$	$0.109_{(0.063)}$	$0.123_{(0.037)}$

[a]PCA: Principal Component Estimation Method; DPCA: Dynamic Principal Component Estimation Method; SSS: Subspace algorithm on state space form. Exp. 1-8 : one factor, different ARMA DGP, no correlation among idiosyncratic components; Exp 9: as Exp. 1 but dynamic impact on variables; Exp 10: as Exp. 1 but one factor imposed in estimation rather than p+q; Exp. 11-20: as 1-10 but temporal correlation among idiosyncratic components; Exp. 21: three AR factors (non correlated), no correlation among idiosyncratic components.

[b]Mean Correlation between true and estimated common component, with MC st.dev. in ().

[c]Mean rejection rate of LM serial correlation test of idiosyncratic component, with MC st.dev. in ().

Table 5: Results for case: N=100, T=100

Exp. [a]	Corr. with True[b]			Serial Correlation[c]		
	PCA	SSS	DPCA	PCA	SSS	DPCA
Exp 1	$0.875_{(0.029)}$	$0.910_{(0.029)}$	$0.798_{(0.030)}$	$0.058_{(0.022)}$	$0.058_{(0.023)}$	$0.070_{(0.025)}$
Exp 2	$0.904_{(0.028)}$	$0.931_{(0.028)}$	$0.843_{(0.032)}$	$0.061_{(0.024)}$	$0.061_{(0.024)}$	$0.075_{(0.026)}$
Exp 3	$0.807_{(0.033)}$	$0.870_{(0.031)}$	$0.711_{(0.036)}$	$0.058_{(0.024)}$	$0.062_{(0.024)}$	$0.091_{(0.028)}$
Exp 4	$0.865_{(0.033)}$	$0.910_{(0.029)}$	$0.793_{(0.041)}$	$0.062_{(0.024)}$	$0.066_{(0.025)}$	$0.093_{(0.030)}$
Exp 5	$0.860_{(0.032)}$	$0.897_{(0.032)}$	$0.773_{(0.033)}$	$0.058_{(0.023)}$	$0.059_{(0.023)}$	$0.072_{(0.026)}$
Exp 6	$0.876_{(0.030)}$	$0.910_{(0.030)}$	$0.798_{(0.032)}$	$0.060_{(0.024)}$	$0.060_{(0.025)}$	$0.072_{(0.027)}$
Exp 7	$0.783_{(0.032)}$	$0.852_{(0.031)}$	$0.679_{(0.034)}$	$0.055_{(0.024)}$	$0.060_{(0.025)}$	$0.090_{(0.029)}$
Exp 8	$0.796_{(0.035)}$	$0.860_{(0.033)}$	$0.696_{(0.037)}$	$0.061_{(0.026)}$	$0.063_{(0.025)}$	$0.093_{(0.030)}$
Exp 9	$0.938_{(0.015)}$	$0.702_{(0.042)}$	$0.883_{(0.021)}$	$0.058_{(0.024)}$	$0.024_{(0.016)}$	$0.081_{(0.028)}$
Exp 10	$0.938_{(0.034)}$	$0.938_{(0.034)}$	$0.898_{(0.028)}$	$0.057_{(0.023)}$	$0.057_{(0.022)}$	$0.063_{(0.024)}$
Exp 11	$0.867_{(0.030)}$	$0.902_{(0.030)}$	$0.792_{(0.031)}$	$0.213_{(0.040)}$	$0.238_{(0.042)}$	$0.236_{(0.044)}$
Exp 12	$0.896_{(0.031)}$	$0.923_{(0.030)}$	$0.837_{(0.034)}$	$0.210_{(0.040)}$	$0.233_{(0.043)}$	$0.229_{(0.045)}$
Exp 13	$0.797_{(0.034)}$	$0.864_{(0.032)}$	$0.704_{(0.037)}$	$0.161_{(0.036)}$	$0.236_{(0.045)}$	$0.230_{(0.044)}$
Exp 14	$0.857_{(0.034)}$	$0.905_{(0.029)}$	$0.786_{(0.040)}$	$0.161_{(0.036)}$	$0.230_{(0.044)}$	$0.228_{(0.044)}$
Exp 15	$0.858_{(0.030)}$	$0.899_{(0.029)}$	$0.772_{(0.032)}$	$0.227_{(0.043)}$	$0.260_{(0.044)}$	$0.264_{(0.045)}$
Exp 16	$0.870_{(0.032)}$	$0.905_{(0.033)}$	$0.798_{(0.033)}$	$0.210_{(0.041)}$	$0.241_{(0.042)}$	$0.245_{(0.044)}$
Exp 17	$0.773_{(0.033)}$	$0.848_{(0.032)}$	$0.672_{(0.035)}$	$0.167_{(0.037)}$	$0.245_{(0.042)}$	$0.235_{(0.044)}$
Exp 18	$0.790_{(0.034)}$	$0.859_{(0.032)}$	$0.694_{(0.038)}$	$0.164_{(0.037)}$	$0.242_{(0.044)}$	$0.238_{(0.041)}$
Exp 19	$0.934_{(0.015)}$	$0.694_{(0.040)}$	$0.879_{(0.020)}$	$0.179_{(0.039)}$	$0.091_{(0.030)}$	$0.228_{(0.043)}$
Exp 20	$0.933_{(0.036)}$	$0.933_{(0.036)}$	$0.891_{(0.032)}$	$0.247_{(0.043)}$	$0.247_{(0.043)}$	$0.251_{(0.043)}$
Exp 21	$0.988_{(0.003)}$	$0.688_{(0.037)}$	$0.974_{(0.005)}$	$0.062_{(0.023)}$	$0.215_{(0.104)}$	$0.082_{(0.026)}$

[a] PCA: Principal Component Estimation Method; DPCA: Dynamic Principal Component Estimation Method; SSS: Subspace algorithm on state space form. Exp. 1-8 : one factor, different ARMA DGP, no correlation among idiosyncratic components; Exp 9: as Exp. 1 but dynamic impact on variables; Exp 10: as Exp. 1 but one factor imposed in estimation rather than p+q; Exp. 11-20: as 1-10 but temporal correlation among idiosyncratic components; Exp. 21: three AR factors (non correlated), no correlation among idiosyncratic components.

[b] Mean Correlation between true and estimated common component, with MC st.dev. in ().

[c] Mean rejection rate of LM serial correlation test of idiosyncratic component, with MC st.dev. in ().

Table 6: Results for case: N=100, T=500

Exp. [a]	Corr. with True[b]			Serial Correlation[c]		
	PCA	SSS	DPCA	PCA	SSS	DPCA
Exp 1	$0.918_{(0.020)}$	$0.958_{(0.019)}$	$0.874_{(0.021)}$	$0.051_{(0.022)}$	$0.052_{(0.022)}$	$0.054_{(0.022)}$
Exp 2	$0.939_{(0.017)}$	$0.970_{(0.016)}$	$0.908_{(0.020)}$	$0.050_{(0.022)}$	$0.051_{(0.022)}$	$0.054_{(0.023)}$
Exp 3	$0.859_{(0.024)}$	$0.939_{(0.019)}$	$0.806_{(0.026)}$	$0.052_{(0.022)}$	$0.053_{(0.023)}$	$0.056_{(0.024)}$
Exp 4	$0.910_{(0.019)}$	$0.963_{(0.015)}$	$0.876_{(0.022)}$	$0.054_{(0.023)}$	$0.053_{(0.022)}$	$0.058_{(0.023)}$
Exp 5	$0.906_{(0.022)}$	$0.951_{(0.021)}$	$0.857_{(0.023)}$	$0.052_{(0.022)}$	$0.051_{(0.022)}$	$0.054_{(0.024)}$
Exp 6	$0.920_{(0.021)}$	$0.960_{(0.019)}$	$0.878_{(0.024)}$	$0.052_{(0.021)}$	$0.052_{(0.021)}$	$0.053_{(0.021)}$
Exp 7	$0.841_{(0.024)}$	$0.931_{(0.021)}$	$0.782_{(0.026)}$	$0.051_{(0.021)}$	$0.052_{(0.022)}$	$0.055_{(0.022)}$
Exp 8	$0.856_{(0.023)}$	$0.939_{(0.019)}$	$0.802_{(0.026)}$	$0.051_{(0.023)}$	$0.051_{(0.022)}$	$0.057_{(0.022)}$
Exp 9	$0.963_{(0.008)}$	$0.709_{(0.035)}$	$0.941_{(0.010)}$	$0.053_{(0.021)}$	$0.021_{(0.016)}$	$0.055_{(0.023)}$
Exp 10	$0.971_{(0.022)}$	$0.971_{(0.022)}$	$0.952_{(0.019)}$	$0.051_{(0.022)}$	$0.052_{(0.022)}$	$0.052_{(0.022)}$
Exp 11	$0.913_{(0.021)}$	$0.954_{(0.021)}$	$0.871_{(0.022)}$	$0.945_{(0.022)}$	$0.952_{(0.022)}$	$0.948_{(0.022)}$
Exp 12	$0.934_{(0.019)}$	$0.965_{(0.018)}$	$0.903_{(0.021)}$	$0.944_{(0.023)}$	$0.949_{(0.021)}$	$0.946_{(0.022)}$
Exp 13	$0.854_{(0.024)}$	$0.937_{(0.019)}$	$0.803_{(0.026)}$	$0.929_{(0.025)}$	$0.950_{(0.022)}$	$0.943_{(0.023)}$
Exp 14	$0.907_{(0.020)}$	$0.962_{(0.016)}$	$0.872_{(0.023)}$	$0.927_{(0.027)}$	$0.950_{(0.023)}$	$0.941_{(0.024)}$
Exp 15	$0.905_{(0.021)}$	$0.953_{(0.020)}$	$0.856_{(0.023)}$	$0.950_{(0.022)}$	$0.956_{(0.021)}$	$0.954_{(0.021)}$
Exp 16	$0.916_{(0.022)}$	$0.957_{(0.021)}$	$0.875_{(0.023)}$	$0.944_{(0.022)}$	$0.952_{(0.021)}$	$0.949_{(0.021)}$
Exp 17	$0.834_{(0.024)}$	$0.929_{(0.020)}$	$0.777_{(0.026)}$	$0.933_{(0.024)}$	$0.954_{(0.020)}$	$0.945_{(0.023)}$
Exp 18	$0.852_{(0.024)}$	$0.937_{(0.020)}$	$0.799_{(0.027)}$	$0.929_{(0.025)}$	$0.952_{(0.022)}$	$0.945_{(0.023)}$
Exp 19	$0.963_{(0.008)}$	$0.712_{(0.034)}$	$0.940_{(0.011)}$	$0.935_{(0.025)}$	$0.533_{(0.088)}$	$0.943_{(0.025)}$
Exp 20	$0.968_{(0.025)}$	$0.968_{(0.025)}$	$0.947_{(0.021)}$	$0.952_{(0.020)}$	$0.952_{(0.021)}$	$0.951_{(0.020)}$
Exp 21	$0.995_{(0.001)}$	$0.675_{(0.021)}$	$0.992_{(0.002)}$	$0.053_{(0.022)}$	$0.810_{(0.076)}$	$0.057_{(0.023)}$

[a] PCA: Principal Component Estimation Method; DPCA: Dynamic Principal Component Estimation Method; SSS: Subspace algorithm on state space form. Exp. 1-8 : one factor, different ARMA DGP, no correlation among idiosyncratic components; Exp 9: as Exp. 1 but dynamic impact on variables; Exp 10: as Exp. 1 but one factor imposed in estimation rather than p+q; Exp. 11-20: as 1-10 but temporal correlation among idiosyncratic components; Exp. 21: three AR factors (non correlated), no correlation among idiosyncratic components.

[b] Mean Correlation between true and estimated common component, with MC st.dev. in ().

[c] Mean rejection rate of LM serial correlation test of idiosyncratic component, with MC st.dev. in ().

Table 7: Results for case: N=200, T=50

Exp. [a]	Corr. with True[b]			Serial Correlation[c]		
	PCA	SSS	DPCA	PCA	SSS	DPCA
Exp 1	$0.849_{(0.030)}$	$0.869_{(0.029)}$	$0.748_{(0.035)}$	$0.067_{(0.018)}$	$0.069_{(0.018)}$	$0.108_{(0.024)}$
Exp 2	$0.881_{(0.029)}$	$0.897_{(0.028)}$	$0.797_{(0.038)}$	$0.074_{(0.019)}$	$0.075_{(0.020)}$	$0.112_{(0.027)}$
Exp 3	$0.775_{(0.035)}$	$0.810_{(0.033)}$	$0.648_{(0.040)}$	$0.069_{(0.018)}$	$0.078_{(0.020)}$	$0.179_{(0.033)}$
Exp 4	$0.830_{(0.041)}$	$0.857_{(0.038)}$	$0.726_{(0.054)}$	$0.077_{(0.020)}$	$0.088_{(0.022)}$	$0.181_{(0.035)}$
Exp 5	$0.833_{(0.031)}$	$0.855_{(0.031)}$	$0.721_{(0.032)}$	$0.066_{(0.018)}$	$0.066_{(0.018)}$	$0.103_{(0.024)}$
Exp 6	$0.849_{(0.031)}$	$0.869_{(0.031)}$	$0.748_{(0.037)}$	$0.070_{(0.017)}$	$0.071_{(0.018)}$	$0.112_{(0.025)}$
Exp 7	$0.753_{(0.031)}$	$0.791_{(0.031)}$	$0.618_{(0.034)}$	$0.067_{(0.018)}$	$0.077_{(0.019)}$	$0.169_{(0.033)}$
Exp 8	$0.765_{(0.036)}$	$0.801_{(0.034)}$	$0.635_{(0.040)}$	$0.071_{(0.019)}$	$0.080_{(0.020)}$	$0.176_{(0.035)}$
Exp 9	$0.921_{(0.017)}$	$0.689_{(0.053)}$	$0.838_{(0.027)}$	$0.069_{(0.018)}$	$0.035_{(0.014)}$	$0.144_{(0.030)}$
Exp 10	$0.912_{(0.030)}$	$0.912_{(0.030)}$	$0.857_{(0.028)}$	$0.067_{(0.018)}$	$0.067_{(0.018)}$	$0.079_{(0.021)}$
Exp 11	$0.840_{(0.031)}$	$0.862_{(0.030)}$	$0.743_{(0.035)}$	$0.102_{(0.021)}$	$0.114_{(0.024)}$	$0.139_{(0.029)}$
Exp 12	$0.866_{(0.032)}$	$0.885_{(0.030)}$	$0.788_{(0.038)}$	$0.105_{(0.022)}$	$0.110_{(0.024)}$	$0.141_{(0.029)}$
Exp 13	$0.764_{(0.034)}$	$0.805_{(0.033)}$	$0.645_{(0.039)}$	$0.092_{(0.022)}$	$0.119_{(0.024)}$	$0.195_{(0.041)}$
Exp 14	$0.814_{(0.045)}$	$0.848_{(0.040)}$	$0.714_{(0.057)}$	$0.098_{(0.023)}$	$0.125_{(0.026)}$	$0.201_{(0.039)}$
Exp 15	$0.831_{(0.031)}$	$0.858_{(0.031)}$	$0.722_{(0.033)}$	$0.111_{(0.022)}$	$0.130_{(0.024)}$	$0.160_{(0.027)}$
Exp 16	$0.839_{(0.031)}$	$0.863_{(0.030)}$	$0.743_{(0.037)}$	$0.105_{(0.022)}$	$0.118_{(0.023)}$	$0.152_{(0.028)}$
Exp 17	$0.742_{(0.032)}$	$0.787_{(0.031)}$	$0.614_{(0.034)}$	$0.089_{(0.021)}$	$0.123_{(0.023)}$	$0.190_{(0.038)}$
Exp 18	$0.752_{(0.037)}$	$0.795_{(0.035)}$	$0.629_{(0.041)}$	$0.091_{(0.023)}$	$0.124_{(0.027)}$	$0.200_{(0.041)}$
Exp 19	$0.913_{(0.019)}$	$0.687_{(0.050)}$	$0.833_{(0.028)}$	$0.089_{(0.022)}$	$0.049_{(0.017)}$	$0.161_{(0.032)}$
Exp 20	$0.902_{(0.033)}$	$0.902_{(0.033)}$	$0.848_{(0.030)}$	$0.118_{(0.023)}$	$0.118_{(0.022)}$	$0.126_{(0.024)}$
Exp 21	$0.981_{(0.005)}$	$0.694_{(0.046)}$	$0.954_{(0.009)}$	$0.077_{(0.019)}$	$0.111_{(0.057)}$	$0.126_{(0.029)}$

[a] PCA: Principal Component Estimation Method; DPCA: Dynamic Principal Component Estimation Method; SSS: Subspace algorithm on state space form. Exp. 1-8 : one factor, different ARMA DGP, no correlation among idiosyncratic components; Exp 9: as Exp. 1 but dynamic impact on variables; Exp 10: as Exp. 1 but one factor imposed in estimation rather than p+q; Exp. 11-20: as 1-10 but temporal correlation among idiosyncratic components; Exp. 21: three AR factors (non correlated), no correlation among idiosyncratic components.

[b] Mean Correlation between true and estimated common component, with MC st.dev. in ().

[c] Mean rejection rate of LM serial correlation test of idiosyncratic component, with MC st.dev. in ().

Table 8: Results for case: N=50, T=50 and non zero mean factor loadings C

Exp. [a]	Corr. with True[b]			Serial Correlation[c]		
	PCA	SSS	DPCA	PCA	SSS	DPCA
Exp 1	$0.881_{(0.039)}$	$0.916_{(0.039)}$	$0.815_{(0.040)}$	$0.070_{(0.037)}$	$0.071_{(0.038)}$	$0.101_{(0.045)}$
Exp 2	$0.904_{(0.036)}$	$0.932_{(0.035)}$	$0.852_{(0.039)}$	$0.073_{(0.037)}$	$0.074_{(0.038)}$	$0.105_{(0.047)}$
Exp 3	$0.817_{(0.045)}$	$0.873_{(0.042)}$	$0.734_{(0.049)}$	$0.068_{(0.035)}$	$0.081_{(0.040)}$	$0.135_{(0.051)}$
Exp 4	$0.865_{(0.046)}$	$0.908_{(0.040)}$	$0.799_{(0.055)}$	$0.075_{(0.038)}$	$0.090_{(0.041)}$	$0.143_{(0.053)}$
Exp 5	$0.867_{(0.042)}$	$0.905_{(0.042)}$	$0.794_{(0.042)}$	$0.064_{(0.033)}$	$0.068_{(0.035)}$	$0.091_{(0.040)}$
Exp 6	$0.881_{(0.042)}$	$0.915_{(0.040)}$	$0.817_{(0.045)}$	$0.070_{(0.036)}$	$0.070_{(0.037)}$	$0.101_{(0.045)}$
Exp 7	$0.798_{(0.046)}$	$0.860_{(0.043)}$	$0.712_{(0.048)}$	$0.065_{(0.035)}$	$0.074_{(0.037)}$	$0.131_{(0.049)}$
Exp 8	$0.807_{(0.047)}$	$0.867_{(0.044)}$	$0.722_{(0.051)}$	$0.070_{(0.036)}$	$0.082_{(0.038)}$	$0.143_{(0.052)}$
Exp 9	$0.921_{(0.023)}$	$0.757_{(0.048)}$	$0.863_{(0.031)}$	$0.071_{(0.036)}$	$0.034_{(0.026)}$	$0.126_{(0.048)}$
Exp 10	$0.938_{(0.048)}$	$0.945_{(0.049)}$	$0.907_{(0.041)}$	$0.071_{(0.036)}$	$0.071_{(0.036)}$	$0.081_{(0.040)}$
Exp 11	$0.878_{(0.042)}$	$0.913_{(0.040)}$	$0.815_{(0.043)}$	$0.101_{(0.043)}$	$0.114_{(0.044)}$	$0.133_{(0.048)}$
Exp 12	$0.900_{(0.041)}$	$0.927_{(0.040)}$	$0.849_{(0.043)}$	$0.103_{(0.042)}$	$0.111_{(0.043)}$	$0.129_{(0.050)}$
Exp 13	$0.808_{(0.046)}$	$0.870_{(0.042)}$	$0.728_{(0.049)}$	$0.085_{(0.042)}$	$0.122_{(0.048)}$	$0.161_{(0.056)}$
Exp 14	$0.860_{(0.048)}$	$0.903_{(0.043)}$	$0.796_{(0.055)}$	$0.091_{(0.043)}$	$0.125_{(0.049)}$	$0.162_{(0.057)}$
Exp 15	$0.863_{(0.041)}$	$0.905_{(0.041)}$	$0.792_{(0.043)}$	$0.105_{(0.043)}$	$0.129_{(0.046)}$	$0.147_{(0.051)}$
Exp 16	$0.875_{(0.045)}$	$0.910_{(0.043)}$	$0.813_{(0.046)}$	$0.104_{(0.046)}$	$0.121_{(0.046)}$	$0.144_{(0.050)}$
Exp 17	$0.790_{(0.044)}$	$0.857_{(0.040)}$	$0.704_{(0.047)}$	$0.083_{(0.039)}$	$0.126_{(0.046)}$	$0.153_{(0.054)}$
Exp 18	$0.797_{(0.046)}$	$0.861_{(0.043)}$	$0.714_{(0.050)}$	$0.085_{(0.040)}$	$0.127_{(0.047)}$	$0.159_{(0.051)}$
Exp 19	$0.919_{(0.025)}$	$0.755_{(0.049)}$	$0.864_{(0.032)}$	$0.086_{(0.040)}$	$0.048_{(0.032)}$	$0.140_{(0.052)}$
Exp 20	$0.932_{(0.048)}$	$0.937_{(0.047)}$	$0.900_{(0.040)}$	$0.121_{(0.049)}$	$0.121_{(0.047)}$	$0.129_{(0.047)}$
Exp 21	$0.983_{(0.006)}$	$0.777_{(0.052)}$	$0.975_{(0.007)}$	$0.075_{(0.036)}$	$0.154_{(0.096)}$	$0.122_{(0.049)}$

[a]PCA: Principal Component Estimation Method; DPCA: Dynamic Principal Component Estimation Method; SSS: Subspace algorithm on state space form. Exp. 1-8 : one factor, different ARMA DGP, no correlation among idiosyncratic components; Exp 9: as Exp. 1 but dynamic impact on variables; Exp 10: as Exp. 1 but one factor imposed in estimation rather than p+q; Exp. 11-20: as 1-10 but temporal correlation among idiosyncratic components; Exp. 21: three AR factors (non correlated), no correlation among idiosyncratic components.

[b]Mean Correlation between true and estimated common component, with MC st.dev. in ().

[c]Mean rejection rate of LM serial correlation test of idiosyncratic component, with MC st.dev. in ().

Table 9: Results for case: N=50, T=50, $s = 1$

Exp.[a]	Corr. with True[b]			Serial Correlation[c]		
	PCA	SSS	DPCA	PCA	SSS	DPCA
Exp 1	$0.827_{(0.051)}$	$0.829_{(0.050)}$	$0.733_{(0.049)}$	$0.066_{(0.035)}$	$0.066_{(0.035)}$	$0.096_{(0.039)}$
Exp 2	$0.858_{(0.047)}$	$0.860_{(0.048)}$	$0.779_{(0.052)}$	$0.069_{(0.035)}$	$0.073_{(0.036)}$	$0.103_{(0.046)}$
Exp 3	$0.737_{(0.052)}$	$0.741_{(0.052)}$	$0.631_{(0.054)}$	$0.067_{(0.035)}$	$0.071_{(0.038)}$	$0.147_{(0.051)}$
Exp 4	$0.803_{(0.057)}$	$0.806_{(0.057)}$	$0.713_{(0.067)}$	$0.074_{(0.039)}$	$0.079_{(0.039)}$	$0.149_{(0.053)}$
Exp 5	$0.810_{(0.052)}$	$0.814_{(0.052)}$	$0.708_{(0.050)}$	$0.064_{(0.037)}$	$0.069_{(0.037)}$	$0.094_{(0.041)}$
Exp 6	$0.823_{(0.055)}$	$0.825_{(0.055)}$	$0.728_{(0.056)}$	$0.068_{(0.036)}$	$0.070_{(0.035)}$	$0.099_{(0.041)}$
Exp 7	$0.713_{(0.053)}$	$0.717_{(0.053)}$	$0.602_{(0.050)}$	$0.066_{(0.035)}$	$0.070_{(0.037)}$	$0.134_{(0.048)}$
Exp 8	$0.725_{(0.055)}$	$0.728_{(0.055)}$	$0.617_{(0.056)}$	$0.072_{(0.037)}$	$0.072_{(0.039)}$	$0.147_{(0.051)}$
Exp 9	$0.897_{(0.027)}$	$0.897_{(0.028)}$	$0.822_{(0.037)}$	$0.066_{(0.037)}$	$0.071_{(0.037)}$	$0.123_{(0.050)}$
Exp 10	$0.907_{(0.060)}$	$0.908_{(0.060)}$	$0.853_{(0.049)}$	$0.068_{(0.036)}$	$0.069_{(0.036)}$	$0.078_{(0.037)}$
Exp 11	$0.815_{(0.054)}$	$0.820_{(0.055)}$	$0.724_{(0.053)}$	$0.101_{(0.043)}$	$0.111_{(0.044)}$	$0.129_{(0.047)}$
Exp 12	$0.852_{(0.051)}$	$0.856_{(0.051)}$	$0.777_{(0.055)}$	$0.103_{(0.044)}$	$0.114_{(0.045)}$	$0.136_{(0.047)}$
Exp 13	$0.727_{(0.058)}$	$0.733_{(0.056)}$	$0.625_{(0.059)}$	$0.084_{(0.042)}$	$0.105_{(0.044)}$	$0.170_{(0.058)}$
Exp 14	$0.795_{(0.055)}$	$0.800_{(0.056)}$	$0.709_{(0.064)}$	$0.093_{(0.043)}$	$0.113_{(0.044)}$	$0.173_{(0.057)}$
Exp 15	$0.801_{(0.056)}$	$0.805_{(0.056)}$	$0.701_{(0.053)}$	$0.110_{(0.042)}$	$0.124_{(0.045)}$	$0.149_{(0.052)}$
Exp 16	$0.813_{(0.056)}$	$0.818_{(0.055)}$	$0.726_{(0.055)}$	$0.104_{(0.045)}$	$0.116_{(0.048)}$	$0.143_{(0.052)}$
Exp 17	$0.707_{(0.050)}$	$0.713_{(0.050)}$	$0.598_{(0.048)}$	$0.087_{(0.039)}$	$0.109_{(0.044)}$	$0.168_{(0.059)}$
Exp 18	$0.723_{(0.055)}$	$0.729_{(0.055)}$	$0.617_{(0.056)}$	$0.083_{(0.038)}$	$0.106_{(0.043)}$	$0.171_{(0.059)}$
Exp 19	$0.895_{(0.028)}$	$0.896_{(0.028)}$	$0.821_{(0.034)}$	$0.087_{(0.039)}$	$0.107_{(0.041)}$	$0.148_{(0.053)}$
Exp 20	$0.893_{(0.063)}$	$0.894_{(0.062)}$	$0.839_{(0.052)}$	$0.120_{(0.047)}$	$0.119_{(0.047)}$	$0.129_{(0.046)}$

[a]PCA: Principal Component Estimation Method; DPCA: Dynamic Principal Component Estimation Method; SSS: Subspace algorithm on state space form. Exp. 1-8 : one factor, different ARMA DGP, no correlation among idiosyncratic components; Exp 9: as Exp. 1 but dynamic impact on variables; Exp 10: as Exp. 1 but one factor imposed in estimation rather than p+q; Exp. 11-20: as 1-10 but temporal correlation among idiosyncratic components.

[b]Mean Correlation between true and estimated common component, with MC st.dev. in ().

[c]Mean rejection rate of LM serial correlation test of idiosyncratic component, with MC st.dev. in ().

Table 10: Results for case: N=100, T=100, $s = 1$

Exp. [a]	Corr. with True[b]			Serial Correlation[c]		
	PCA	SSS	DPCA	PCA	SSS	DPCA
Exp 1	$0.874_{(0.030)}$	$0.877_{(0.030)}$	$0.794_{(0.032)}$	$0.058_{(0.022)}$	$0.059_{(0.023)}$	$0.073_{(0.026)}$
Exp 2	$0.905_{(0.028)}$	$0.906_{(0.028)}$	$0.844_{(0.031)}$	$0.061_{(0.025)}$	$0.062_{(0.024)}$	$0.075_{(0.026)}$
Exp 3	$0.806_{(0.032)}$	$0.810_{(0.032)}$	$0.709_{(0.035)}$	$0.058_{(0.024)}$	$0.059_{(0.025)}$	$0.092_{(0.028)}$
Exp 4	$0.865_{(0.032)}$	$0.868_{(0.032)}$	$0.792_{(0.041)}$	$0.061_{(0.024)}$	$0.064_{(0.024)}$	$0.095_{(0.029)}$
Exp 5	$0.859_{(0.031)}$	$0.861_{(0.030)}$	$0.771_{(0.032)}$	$0.056_{(0.023)}$	$0.055_{(0.024)}$	$0.067_{(0.025)}$
Exp 6	$0.877_{(0.031)}$	$0.880_{(0.031)}$	$0.800_{(0.033)}$	$0.059_{(0.025)}$	$0.060_{(0.024)}$	$0.074_{(0.026)}$
Exp 7	$0.784_{(0.033)}$	$0.789_{(0.033)}$	$0.680_{(0.034)}$	$0.056_{(0.023)}$	$0.058_{(0.023)}$	$0.089_{(0.028)}$
Exp 8	$0.800_{(0.033)}$	$0.804_{(0.033)}$	$0.701_{(0.037)}$	$0.058_{(0.023)}$	$0.059_{(0.023)}$	$0.094_{(0.029)}$
Exp 9	$0.939_{(0.014)}$	$0.940_{(0.013)}$	$0.884_{(0.019)}$	$0.058_{(0.023)}$	$0.059_{(0.024)}$	$0.085_{(0.029)}$
Exp 10	$0.938_{(0.036)}$	$0.938_{(0.035)}$	$0.896_{(0.029)}$	$0.057_{(0.022)}$	$0.057_{(0.023)}$	$0.062_{(0.025)}$
Exp 11	$0.868_{(0.031)}$	$0.872_{(0.032)}$	$0.792_{(0.032)}$	$0.217_{(0.043)}$	$0.238_{(0.044)}$	$0.244_{(0.044)}$
Exp 12	$0.897_{(0.029)}$	$0.901_{(0.029)}$	$0.839_{(0.032)}$	$0.209_{(0.040)}$	$0.228_{(0.043)}$	$0.231_{(0.044)}$
Exp 13	$0.796_{(0.033)}$	$0.802_{(0.033)}$	$0.703_{(0.036)}$	$0.171_{(0.037)}$	$0.218_{(0.044)}$	$0.238_{(0.044)}$
Exp 14	$0.859_{(0.034)}$	$0.864_{(0.033)}$	$0.790_{(0.041)}$	$0.167_{(0.038)}$	$0.213_{(0.044)}$	$0.232_{(0.043)}$
Exp 15	$0.859_{(0.031)}$	$0.863_{(0.031)}$	$0.773_{(0.033)}$	$0.232_{(0.044)}$	$0.255_{(0.046)}$	$0.261_{(0.044)}$
Exp 16	$0.872_{(0.032)}$	$0.876_{(0.032)}$	$0.798_{(0.034)}$	$0.215_{(0.040)}$	$0.234_{(0.042)}$	$0.245_{(0.046)}$
Exp 17	$0.775_{(0.032)}$	$0.783_{(0.032)}$	$0.673_{(0.034)}$	$0.174_{(0.037)}$	$0.223_{(0.045)}$	$0.243_{(0.044)}$
Exp 18	$0.794_{(0.033)}$	$0.801_{(0.032)}$	$0.698_{(0.036)}$	$0.171_{(0.040)}$	$0.218_{(0.043)}$	$0.247_{(0.046)}$
Exp 19	$0.935_{(0.014)}$	$0.937_{(0.013)}$	$0.880_{(0.019)}$	$0.187_{(0.039)}$	$0.226_{(0.044)}$	$0.235_{(0.041)}$
Exp 20	$0.934_{(0.038)}$	$0.934_{(0.037)}$	$0.893_{(0.032)}$	$0.241_{(0.045)}$	$0.241_{(0.045)}$	$0.245_{(0.045)}$

[a]PCA: Principal Component Estimation Method; DPCA: Dynamic Principal Component Estimation Method; SSS: Subspace algorithm on state space form. Exp. 1-8 : one factor, different ARMA DGP, no correlation among idiosyncratic components; Exp 9: as Exp. 1 but dynamic impact on variables; Exp 10: as Exp. 1 but one factor imposed in estimation rather than p+q; Exp. 11-20: as 1-10 but temporal correlation among idiosyncratic components.

[b]Mean Correlation between true and estimated common component, with MC st.dev. in ().

[c]Mean rejection rate of LM serial correlation test of idiosyncratic component, with MC st.dev. in ().

Table 11: Results for Experiment 21 (3 AR factors (non correlated), no correlation among idiosyncratic components) and $s = 1$

N/T	Corr. with True[a]			Serial Correlation[b]		
	PCA	SSS	DPCA	PCA	SSS	DPCA
$N = 50, T = 50$	$0.9754_{(0.008)}$	$0.9751_{(0.008)}$	$0.9478_{(0.013)}$	$0.076_{(0.040)}$	$0.074_{(0.038)}$	$0.125_{(0.048)}$
$N = 50, T = 100$	$0.9844_{(0.004)}$	$0.9843_{(0.004)}$	$0.9703_{(0.007)}$	$0.062_{(0.033)}$	$0.060_{(0.033)}$	$0.082_{(0.038)}$
$N = 100, T = 50$	$0.9792_{(0.006)}$	$0.9789_{(0.006)}$	$0.9520_{(0.011)}$	$0.076_{(0.028)}$	$0.076_{(0.027)}$	$0.124_{(0.037)}$
$N = 100, T = 100$	$0.9880_{(0.004)}$	$0.9879_{(0.004)}$	$0.9745_{(0.006)}$	$0.063_{(0.025)}$	$0.063_{(0.025)}$	$0.084_{(0.028)}$
$N = 500, T = 50$	$0.9827_{(0.003)}$	$0.9825_{(0.003)}$	$0.9554_{(0.007)}$	$0.076_{(0.013)}$	$0.075_{(0.012)}$	$0.126_{(0.021)}$
$N = 100, T = 500$	$0.9914_{(0.002)}$	$0.9913_{(0.002)}$	$0.9777_{(0.003)}$	$0.061_{(0.010)}$	$0.061_{(0.010)}$	$0.082_{(0.012)}$
$N = 200, T = 50$	$0.9835_{(0.006)}$	$0.9878_{(0.005)}$	$0.9741_{(0.008)}$	$0.074_{(0.039)}$	$0.074_{(0.038)}$	$0.127_{(0.050)}$

[a]Mean Correlation between true and estimated common component, with MC st.dev. in ().

[b]Mean rejection rate of LM serial correlation test of idiosyncratic component, with MC st.dev. in ().