

# **Consistent Estimation Of The Number Of Dynamic Factors In A Large *N* And *T* Panel**

July 2005  
(This Draft: July 3, 2006)

Dante Amengual  
Department of Economics  
Princeton University  
Princeton, NJ 08544  
amengual@princeton.edu

and

Mark W. Watson\*  
Woodrow Wilson  
Princeton University  
Princeton, NJ 08544  
mwatson@princeton.edu

Dante Amengual is graduate student, Department of Economics, Princeton University, Princeton, NJ 08544 ([amengual@princeton.edu](mailto:amengual@princeton.edu)); and Mark W. Watson is Professor, Department of Economics and Woodrow Wilson School, Princeton University, Princeton, NJ 08544 ([mwatson@princeton.edu](mailto:mwatson@princeton.edu)). This research is an outgrowth of joint work with Jim Stock, who we thank for comments and suggestions. Thanks also to Jushan Bai, Serena Ng, two referees and the Associate Editor for useful comments. The work was funded in part by NSF grant SBR-0214131.

## Abstract

Bai and Ng (2002) propose a consistent estimator for the number of *static* factors in a large  $N$  and  $T$  approximate factor model. This paper shows how the Bai-Ng estimator can be modified to consistently estimate the number of *dynamic* factors in a restricted dynamic factor model. The modification is straightforward: the standard Bai-Ng estimator is applied to residuals obtained by projecting the observed data onto lagged values of principal component estimates of the static factors.

Key Words: approximate factor model, dynamic factor model, Bai-Ng estimator

# 1. Introduction

Panel datasets with large time series dimension ( $T$ ) and cross section dimension ( $N$ ) are being increasingly used in macroeconomics for both forecasting and structural analysis. Often, these data are analyzed in the context of an assumed latent factor structure of the form

$$X_t = \Lambda F_t + e_t, \quad (1.1)$$

for  $t = 1, \dots, T$ , where  $X_t$  denotes an  $N \times 1$  vector of observed variables,  $F_t$  is an  $r \times 1$  vector of latent factors,  $\Lambda$  is a matrix of coefficients, and  $e_t$  is a vector of errors. When the elements of  $e_t$  have weak cross sectional and serial correlation, the factors  $F_t$  summarize the important cross covariance properties of the variables.

A question of fundamental interest is the number of latent factors,  $r$ , that are required in (1.1). Significant progress on addressing this problem was made in Bai and Ng (2002) who proposed consistent estimators of  $r$  based on a penalized least squares objective function associated with the classic principal components estimator. However, in dynamic models it is important to differentiate between the number of “static” factors (necessary to fit the covariance matrix of  $X$ ) and the number of “dynamic” factors (necessary to fit the spectral density matrix of  $X$ ). While the Bai-Ng estimator was developed to estimate the number of static factors, this paper shows that it can be easily modified to consistently estimate the number of dynamic factors.

Dynamics can be incorporated in the model by assuming that  $F_t$  evolves as a VAR:

$$F_t = \sum_{i=1}^p \Phi_i F_{t-i} + \varepsilon_t \quad (1.2)$$

with innovations  $\varepsilon_t$  that can be represented as  $\varepsilon_t = G\eta_t$  where  $G$  is  $r \times q$  with full column rank and  $\eta_t$  is sequence of shocks with mean zero and covariance matrix  $\Sigma_{\eta\eta} = I_q$ ;  $\eta_t$  is the vector of dynamic factor shocks. Several papers show how (1.1) and (1.2) can be derived from a restricted version of a general dynamic factor model driven by  $q$  dynamic factors; in this case  $F_t$  contains linear combinations of current and lagged values of the dynamic factors and (1.1)-(1.2) is analogous to the companion form representation of the dynamic factor model. (See Bai and Ng (2005), Forni, Hallin, Lippi, and Reichlin (2005), Giannone, Reichlin, and Sala (2004), and Stock and Watson (2004, 2005).)

To see how the Bai-Ng estimator might be used to estimate the number of dynamic factors,  $q$ , substitute (1.2) into (1.1) to obtain

$$Y_t = \Gamma \eta_t + e_t \quad (1.3)$$

where  $Y_t = X_t - \sum_{i=1}^p \Lambda \Phi_i F_{t-i}$  and  $\Gamma = \Lambda G$ . Thus  $Y_t$  can be represented as a factor model with  $q$  factors that correspond to the common shocks  $\eta_t$ . Were  $Y_t$  observed data,  $q$  could be consistently estimated by applying the Bai-Ng estimator to  $Y_t$ . This is infeasible because  $Y_t$  depends on unknown parameters and lags of the unobserved factors.

This paper studies the consistency properties of the Bai-Ng estimator applied to  $\hat{Y}_t = X_t - \sum_{i=1}^p \hat{\Pi}_i \hat{F}_{t-i}$  where  $\hat{\Pi}_i$  is an estimator of  $\Lambda \Phi_i$  and  $\hat{F}_{t-i}$  is an estimator of  $F_{t-i}$ . The analysis proceeds in two steps. In the first step, the Bai-Ng estimator is shown to remain consistent if the estimation error  $\hat{Y}_t - Y_t$  is sufficiently small (specifically  $\sum_{t=1}^T \sum_{i=1}^n (\hat{Y}_{it} - Y_{it})^2 = O_p[\max(N, T)]$ ). The second step shows that the principal components estimator of  $F$  and feasible estimators of  $\Pi$  yield estimators  $\hat{Y}_{it}$  that achieve this degree of accuracy. Together these results yield a feasible consistent estimator of the number of dynamic factors.

The estimator studied in this paper was proposed in Stock and Watson (2005) and applied to the problem of estimating the number of dynamic factors in a large panel of U.S. macroeconomic time series. Stock and Watson (2005) did not study the consistency properties of the estimator, and that is the purpose of the present paper. Other estimators have also been proposed and used in applied work. Notably, Forni, Hallin, Lippi and Reichlin (2000) suggest informal methods based on the relative size of eigenvalues from the estimated spectral density matrix for  $X$ , related methods have been proposed and applied in the empirical analysis of Forni, Lippi and Reichlin (2003), Giannone, Reichlin and Sala (2004) and elsewhere, and Hallin and Liška (2005) show how a consistent estimator of  $q$  can be constructed from the estimated spectrum. Bai and Ng (2005a) propose an estimator for  $q$  based on the residual covariance matrix of the VAR in (1.3) estimated using the principal components estimator of  $F_t$ , and show that the estimator is consistent. Section 3 studies the relative performance of various consistent estimators using a simulation study.

More generally, the plan of this paper is as follows. Section 2 briefly summarizes the Bai-Ng estimator, shows the estimator remains consistent when applied to data contaminated with a small amount of measurement error, and uses this result to show that the Bai-Ng estimator applied to  $\hat{Y}$  is a consistent estimator of the number of dynamic factors. A Monte Carlo study is presented in Section 3 to gauge the performance of the estimator, Section 4 contains some concluding remarks, and the Appendix includes the proofs to the results given in Section 2.

## 2. Assumptions and Asymptotic Results

### 2.1 Review of Existing Work with a Small Extension

We begin by reviewing results for the model (1.1) under a standard set of assumptions. Transposing (1.1) and stacking the  $T$  equations yields

$$X = F\Lambda' + e \tag{2.1}$$

where  $X$  is  $T \times N$ ,  $F$  is  $T \times r$ ,  $\Lambda$  is  $N \times r$ , and  $e$  is  $T \times N$ . The  $t$ 'th rows of  $X$ ,  $F$  and  $e$  are  $X_t'$ ,  $F_t'$ , and  $e_t'$ ; the  $i$ 'th row of  $\Lambda$  is  $\lambda_i'$ ; the  $i$ 'th element of  $X_t$  is  $X_{it}$  and similarly for  $e_{it}$ , so that  $X_{it} = \lambda_i'F_t + e_{it}$ .

Asymptotic properties of various statistics generated by this model have been studied in Stock and Watson (2002), Bai and Ng (2002), Bai (2003), and Bai and Ng (2005a, 2005b) under a similar set assumptions. The focus is on datasets in which both  $N$  and  $T$  are large, so that the asymptotics assume that  $N, T \rightarrow \infty$  jointly (equivalently that  $N = N(T)$  with  $\lim_{T \rightarrow \infty} N(T) = \infty$ ). The minimum value of  $N$  and  $T$  plays an important

role in the analysis and this value is denoted by  $s_{NT} = \min(N, T)$ . The remaining assumptions concern moments and dependence properties of the variables; for the purposes of this paper, the following assumptions suffice:

$$(A.1) \quad E(F_t F_t') = I_r.$$

$$(A.2) \quad E(\lambda_i \lambda_i') = \Sigma_{\Lambda\Lambda}, \text{ where } \Sigma_{\Lambda\Lambda} \text{ is a diagonal matrix with elements } \sigma_{ii} > \sigma_{jj} > 0 \text{ for } i < j.$$

(When  $\Lambda$  is deterministic,  $\Sigma_{\Lambda\Lambda}$  is interpreted as the limiting empirical average.)

$$(A.3) \quad T^{-1} \sum_{t=1}^T F_t F_t' \xrightarrow{p} I_r.$$

$$(A.4) \quad N^{-1} \sum_{i=1}^N \lambda_i \lambda_i' \xrightarrow{p} \Sigma_{\Lambda\Lambda}.$$

$$(A.5) \quad (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T e_{it}^2 \xrightarrow{p} \sigma_e^2 > 0.$$

$$(A.6) \quad \text{For some integer } m \geq 2 \text{ and for all integers } j \leq m, \text{Etrace}[(ee')^j] =$$

$$O(NT \times \max[N, T]^{j-1}).$$

$$(A.7) \quad E \sum_{t=1}^T \sum_{s=1}^T \left( \sum_{i=1}^N \lambda_i' F_t e_{is} \right)^2 = O(NT^2).$$

$$(A.8) \quad E \sum_{t=1}^T \sum_{i=1}^N \lambda_i' \lambda_i e_{it}^2 = O(NT).$$



$$(A.9) \quad E \sum_{i=1}^N \left\| \sum_{t=1}^T F_t e_{it} \right\|^2 = O(NT).$$

Assumptions (A.1)-(A.5) rule out explosive or trending behavior in both the time series and cross section dimensions; the particular values of  $E(F_t F_t')$  and  $E(\lambda_i \lambda_i')$  listed in (A.1) and (A.2) are normalizations (because  $\Lambda F_t = \Lambda H H^{-1} F_t$  for arbitrary  $H$ ), and assumption (A.5) rules out degenerate cases in which the factors explain all of the variance of the  $X_{it}$ 's. Assumption (A.6) limits the variability and dependence in the errors  $e_{it}$ . For  $j = 1$ , it implies that  $\sum_{i=1}^N \sum_{t=1}^T E(e_{it}^2) = O(NT)$ ; for  $j = 2$ , it implies that  $\sum_{i=1}^N \sum_{j=1}^N \left( \sum_{t=1}^T e_{it} e_{jt} \right)^2 = \sum_{t=1}^T \sum_{\tau=1}^T \left( \sum_{i=1}^N e_{it} e_{i\tau} \right)^2 = O_p(NT \times \max[N, T])$ , and so forth for larger values of  $j$ . Assumptions (A7)-(A9) limit the dependence across elements of  $\Lambda$ ,  $F$  and  $e$ . Importantly, all of these assumptions hold for sequences of i.i.d. random variables with the appropriate number of moments, and assumptions (A6)-(A9) can be interpreted as relaxing the i.i.d. assumption to allow weak dependence.

The Bai-Ng estimators of  $r$  are based on penalized least squares objective functions. The penalty function depends on a deterministic function  $g(N, T)$  that satisfies  $g(N, T) \rightarrow 0$  and  $s_{NT}^\delta g(N, T) \rightarrow \infty$  for  $\delta = (m - 1)/m$ , where  $m$  is given in assumption (A6). The least squares objective function is conveniently written in terms of the eigenvalues of the  $XX'$  moment matrix. Let  $\omega_i$  denote the  $i$ 'th largest eigenvalue of  $(NT)^{-1}XX'$ , and consider the least squares problem:  $\min_{\{\lambda_i^k\} \{F_t^k\}} (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T (X_{it} - \lambda_i^{k'} F_t^k)^2$ , where  $\lambda_i^k$  and  $F_t^k$  are arbitrary  $k \times 1$  vectors. The usual principal components calculations imply

that the average predicted sum of squares associated with the least squares solution is given by  $R(k, X) = \sum_{i=1}^k \omega_i$ . Letting  $\hat{\sigma}_X^2 = (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T X_{it}^2$  denote the average total sum of squares, the penalized average sum of squared residuals is  $PC(k, X) = \hat{\sigma}_X^2 - R(k, X) + kg(N, T)$ , and the Bai-Ng “PC” estimator is

$$\widehat{BN}^{PC}(X) = \arg \min_{0 \leq k \leq r^{\max}} PC(k, X). \quad (2.2)$$

Letting  $ICP(k, X) = \ln[\hat{\sigma}_X^2 - R(k, X)] + kg(N, T)$ , the Bai-Ng “ICP” estimator is

$$\widehat{BN}^{ICP}(X) = \arg \min_{0 \leq k \leq r^{\max}} ICP(k, X), \quad (2.3)$$

where  $r^{\max}$  is a finite constant that satisfies  $r \leq r^{\max}$ .

Consistency of the Bai-Ng estimator is given in the following lemma:

**Lemma 1 (Bai-Ng):** Under assumptions (A1)-(A9),  $\widehat{BN}^{PC}(X) \xrightarrow{P} r$  and  $\widehat{BN}^{ICP}(X) \xrightarrow{P} r$ .

As discussed on the last section, we will study consistency of the Bai-Ng estimators applied to variables measured with error ( $\hat{Y}_t$  in the notation of the last section). The following result shows that the Bai-Ng estimators remain consistent in the presence of sufficiently small measurement error.

**Lemma 2:** Suppose (A1)-(A9) are satisfied and  $\tilde{X} = X + b$  where

$$T^{-1}N^{-1} \sum_{i=1}^N \sum_{t=1}^T b_{it}^2 = O_p(s_{NT}^{-1}), \text{ then } \widehat{BN}^{PC}(\tilde{X}) \xrightarrow{p} r \text{ and } \widehat{BN}^{ICP}(\tilde{X}) \xrightarrow{p} r.$$

Bai and Ng (2002) showed consistency of  $\widehat{BN}^{PC}(X)$  and  $\widehat{BN}^{ICP}(X)$  for  $\delta = 1$  using assumptions like those in A1-A9, but without A6. However, there was an error in their proof. As shown in their errata, their proof is valid using a stronger condition on  $e$ . In particular, for  $e = R\xi H$  where  $R$  and  $H$  are  $N \times N$  and  $T \times T$  matrices with bounded eigenvalues,  $\xi$  is required to be a  $T \times N$  matrix of independently distributed random variables with mean zero and bounded 7<sup>th</sup> moments.

## 2.2 Consistent Estimation of the Number of Dynamic Factors

The results from Lemma 2 suggest that the estimators  $\widehat{BN}^{PC}(\hat{Y})$  and  $\widehat{BN}^{ICP}(\hat{Y})$  will be consistent for the number of dynamic factors if the error  $\hat{Y} - Y$  is small. We consider two versions of  $\hat{Y}$  that are sufficiently accurate for this purpose. Both rely on a first-stage estimate of  $F$ . Thus, let  $\hat{F}$  and  $\hat{\Lambda}$  denote the principal components estimators of  $F$  and  $\Lambda$  constructed from (2.1) using a consistent estimator of  $r$ . Let  $(\hat{\Phi}_1, \hat{\Phi}_2, \dots, \hat{\Phi}_p)$  denote the OLS estimators from the regression of  $\hat{F}_t$  onto  $(\hat{F}_{t-1}, \dots, \hat{F}_{t-p})$ .

The first version of  $\hat{Y}$  is

$$\hat{Y}_t^A = X_t - \sum_{i=1}^p \hat{\Lambda} \hat{\Phi}_i \hat{F}_{t-i}. \quad (2.4)$$

The second version of  $\hat{Y}$  uses direct estimates of the regression of  $X_t$  onto lags of  $F_t$ . Let  $(\hat{\Pi}_1^{OLS}, \hat{\Pi}_2^{OLS}, \dots, \hat{\Pi}_p^{OLS})$  denote the OLS estimators from the regression of  $X_t$  onto  $(\hat{F}_{t-1}, \dots, \hat{F}_{t-p})$ . The second version of  $\hat{Y}$  is

$$\hat{Y}_t^B = X_t - \sum_{i=1}^p \hat{\Pi}_i^{OLS} \hat{F}_{t-i} \quad (2.5)$$

which does not impose the cross-equation constraint  $\Pi_i = \Lambda \Phi_i$ .

Consistency of the Bai-Ng estimator for  $q$  is then readily shown if (i) the factor model (1.3) for  $Y$  satisfies the analogues of conditions (A.1)-(A.9) above, and (ii) the estimators  $\hat{F}$ ,  $\hat{\Lambda}$ ,  $\hat{\Phi}$  and  $\hat{\Pi}^{OLS}$  are sufficiently accurate. Thus to begin, assume that (A.1)-(A.9) hold with  $\eta$  replacing  $F$  and  $\Gamma$  replacing  $\Lambda$ . (Note that the normalization in (A.1)-(A.2) can be achieved by appropriate choice of  $G$ .) Stock and Watson (2002) and Bai (2003) discuss the accuracy of the estimators  $\hat{F}$  and  $\hat{\Lambda}$  under assumptions like those listed as (A.1)-(A.9) above. As in Bai and Ng (2005),  $\hat{\Phi}$  will be  $T^{1/2}$ -consistent under a standard set of assumptions for the VAR for  $F_t$ :

(A10) Let  $\mathbf{F}_t = (F'_{t-1}, \dots, F'_{t-p})'$ , then

- (i) the stochastic process  $\{F_t\}$  is stationary and ergodic;
- (ii)  $E(\mathbf{F}_t \mathbf{F}_t')$  is non-singular;
- (iii)  $\text{vec}(\mathbf{F}_t \eta_t')$  is a martingale difference sequence with finite second moments.

Finally, accuracy of  $\hat{\Pi}^{OLS}$  requires the additional assumption:

$$(A.11) \quad E \sum_{i=1}^N \left\| \sum_{t=1}^T \mathbf{F}_t e_{it} \right\|^2 = O(NT).$$

We then have:

**Theorem:** Consider the model (1.1)-(1.3). Suppose that (1.1) satisfies (A.1)-(A.9), that the analogous assumptions are satisfied for (1.3) and that (A.10) is satisfied. Then

(a)  $\widehat{BN}^{PC}(\hat{Y}^A) \xrightarrow{p} q$  and  $\widehat{BN}^{ICP}(\hat{Y}^A) \xrightarrow{p} q$ .

(b) In addition, suppose that (A.11) is satisfied. Then  $\widehat{BN}^{PC}(\hat{Y}^B) \xrightarrow{p} q$  and  $\widehat{BN}^{ICP}(\hat{Y}^B) \xrightarrow{p} q$ .

The remaining ingredient in the testing problem is  $p$ , the number of lags in the VAR. It is straightforward to show that, under the usual VAR assumptions,  $p$  can be estimated consistently by BIC.

In some models, innovations in a subset of the  $X_t$  variables may depend on only a subset of the dynamic shocks  $\eta_t$ . For example in Bernanke, Boivin and Elias (2005) and Stock and Watson (2005),  $X_t$  is partitioned into a set of “slow moving” variables and other variables,  $X_t = (X_t^{Slow}, X_t^{Other})'$  where innovations in  $X_t^{Slow}$  depend on only a subset of the  $\eta_t$ . It is straightforward to show that the size of this subset can be consistently estimated ( $N_{Slow}, T \rightarrow \infty$ ) using  $\widehat{BN}^{ICP}$  applied to the relevant subset of elements of  $\hat{Y}$ .

### 3. Comparing the Estimators Using Simulated Data

#### 3.1 Experimental Design

The experimental design is taken from Bai and Ng (2005a) where four data generating processes (DGPs) are considered.

**DGPs.** In the first design (DGP1),  $X_{it} = \lambda_i' F_t + e_{it}$  and  $F_t = \Phi F_{t-1} + G \eta_t$ , where  $F_t$  is  $5 \times 1$  and  $\eta_t$  is  $3 \times 1$ , so that  $r = 5$  and  $q = 3$ ;  $\{\lambda_i\}$ ,  $\{e_{it}\}$ , and  $\{\eta_t\}$  are mutually independent, with  $\{\lambda_i\}$  and  $\{\eta_t\}$  i.i.d. standard normal random variables/vectors;  $\Phi$  is a diagonal matrix with elements (0.2, 0.375, 0.55, 0.725, 0.90), and the columns of  $G$  are randomly chosen from the unit sphere and are independent of the other random variables. To allow cross sectional dependence in the idiosyncratic errors,  $e_t$  is  $N(0, \Omega)$ , where  $\Omega_{ij} = \rho^{|i-j|}$ . Results are presented for  $\rho = 0$  and  $\rho = 0.5$ .

DGP2 is the same as DGP1, but with  $r = 3$  and  $\Phi = 0.5 \times I_3$ .

In the final two designs,  $X_t$  is a moving average of factors  $f_t$  that follow an AR (DGP3) or MA (DGP4) process. In DGP3,  $X_{it} = (\lambda_{i0} + \lambda_{i1}L)f_t + e_{it}$  and  $f_t = \phi f_{t-1} + \eta_t$ , where  $f_t$  is  $2 \times 1$ , so that  $r = 4$  and  $q = 2$ . This model can be written as (1.1) and (1.2) with  $F_t = (f_t' f_{t-1}')'$ ,  $\lambda_i' = (\lambda_{i0}' \lambda_{i1}')$ ,  $\Phi = \begin{bmatrix} \phi & 0 \\ I_2 & 0 \end{bmatrix}$ , and  $G = [I_2 \ 0_{2 \times 2}]'$ . The factor loadings and errors are generated as in DGP1, and  $\phi = 0.5 \times I_2$ . In DGP4,  $X_{it} = (\lambda_{i0} + \lambda_{i1}L + \lambda_{i2}L^2)f_t + e_{it}$  and  $f_t = (I_2 + \Theta L)\eta_t$ , where  $f_t$  is  $2 \times 1$ , so that  $r = 6$  and  $q = 2$ . In this design  $F_t = (f_t' f_{t-1}' f_{t-2}')$ , but now  $F_t$  follows an MA process, so that the VAR in (1.2) serves as an approximation. The MA coefficient matrix is diagonal with elements 0.2 and 0.9.

**Estimators.** The  $\widehat{BN}^{ICP}$  estimators are implemented using the penalty factor  $g(N, T) = \ln(s_{NT})/A$ , where  $A = NT/(N+T)$ . (This is the ‘‘ICP2’’ penalty factor in Bai and

Ng (2002).)  $r$  is estimated using  $\widehat{BN}^{ICP}(X)$ , where  $X$  is the standardized version of the data generated by DGP1-DGP4, and where  $r^{max} = 10$ .  $q$  is estimated using  $\widehat{BN}^{ICP}(\hat{Y}^A)$  and  $\widehat{BN}^{ICP}(\hat{Y}^B)$  constructed using an estimated VAR(2) for  $\hat{F}_t$ .

Two alternative estimators,  $\hat{q}_3$  and  $\hat{q}_4$  from Bai and Ng(2005a), were also constructed. These estimators use the eigenvalues of the residual covariance matrix of the VAR for  $\hat{F}_t$  to estimate  $q$ . Specifically, let  $\hat{\varepsilon}_t = \hat{F}_t - \sum_{i=1}^p \hat{\Phi}_i \hat{F}_{t-i}$  where  $\hat{F}_t$  is the  $\hat{r} \times 1$  vector of factors estimated by principal components using the normalization  $N^{-1} \hat{\Lambda}' \hat{\Lambda} = I_r$  and  $T^{-1} \hat{F}' \hat{F} = \text{diag}(\sigma_{ii})$ ,  $\hat{\Sigma}_{\hat{\varepsilon}\hat{\varepsilon}} = T^{-1} \sum_{t=p+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_t'$  denote the estimated covariance matrix, and denote the ordered eigenvalues of  $\hat{\Sigma}_{\hat{\varepsilon}\hat{\varepsilon}}$  by  $c_1 \geq c_2 \geq \dots \geq c_{\hat{r}}$ . Let  $D_{1,k} = [c_{k+1}^2 / \sum_{i=1}^{\hat{r}} c_i^2]^{1/2}$  and  $D_{2,k} = [\sum_{i=k+1}^{\hat{r}} c_i^2 / \sum_{i=1}^{\hat{r}} c_i^2]^{1/2}$ ; the estimators are  $\hat{q}_3 = \min_k [k : D_{1,k} < m / s_{NT}^{2/5}]$ , and  $\hat{q}_4 = \min_k [k : D_{2,k} < m / s_{NT}^{2/5}]$ , where  $m$  is a positive constant. Following Bai and Ng(2005a) we implement these estimators using  $m = 1.0$ .

### 3.2 Results

Results are shown in Table 1 for each DGP and various values of  $N$  and  $T=100$ . Panel A shows results with  $\rho = 0$  (so that the  $e_{it}$  errors are mutually uncorrelated) and panel B shows results with  $\rho = 0.5$  (so that  $e_{it}$  are correlated in the cross section).

Looking first at panel A, five results stand out. First, the estimators are quite accurate for  $N$  as small as 50, at least for the simple designs considered. All of the estimators produce the correct answer in more that 98% of the simulations when  $N = 50$ ,

and perform nearly as well when  $N = 40$ . Second, the constraint  $\Pi_i = \Lambda\Phi_i$  used by  $\widehat{BN}^{ICP}(\hat{Y}^A)$  but ignored by  $\widehat{BN}^{ICP}(\hat{Y}^B)$  is useful:  $\widehat{BN}^{ICP}(\hat{Y}^A)$  has a smaller root mean squared error than  $\widehat{BN}^{ICP}(\hat{Y}^B)$  in all of the cases considered in the table. Third, for DGP1 and DGP2,  $\widehat{BN}^{ICP}(\hat{Y}^A)$  achieves a higher proportion of correct values of  $q$  than the other estimators; for DGP3 and DGP4,  $\hat{q}_3$  achieves the highest proportion of correct values. Fourth, in DGP1, while  $\Sigma_{FF}$  has rank 5, 2 of its eigenvalues are small and  $\widehat{BN}^{IPC}(X)$  tends to underestimate the number of static factors when  $N$  is large. In spite of this,  $\widehat{BN}^{IPC}(\hat{Y}^a)$  and  $\widehat{BN}^{IPC}(\hat{Y}^b)$  accurately estimate the number of dynamic factors. Finally, comparing the results from DGP3 and DGP4, the AR approximation for DGP4 does not appear to lead to a serious deterioration of performance in any of the estimators.

Panel B shows that the performance of the  $\widehat{BN}^{IPC}$  deteriorates when there is cross sectional correlation in the errors:  $\widehat{BN}^{IPC}(X)$  tends to overestimate  $r$ , the number of static factors, and, while not as severe, this upward bias is also evident in  $\widehat{BN}^{IPC}(\hat{Y}^a)$  and  $\widehat{BN}^{IPC}(\hat{Y}^b)$ .  $\hat{q}_3$  and  $\hat{q}_4$  suffer only a small deterioration in accuracy. Both  $\widehat{BN}^{IPC}(\hat{Y})$  and  $\hat{q}$  provide accurate estimates of the number of dynamic factors when  $N = 100$ .



#### 4. Summary and Concluding Remarks

This paper has proposed a modification of the Bai-Ng (2002) estimator and shown that the modification provides a consistent estimator for the number of dynamic factors in an approximate dynamic factor model. The modification uses a result (Lemma 2) that shows that the Bai-Ng estimator remains consistent even when the data are contaminated with a suitably small amount of error. This result may prove useful in other settings, for example in models in which the equation for  $X_{it}$  has the form  $X_{it} = \lambda_i' F_t + \beta_i' Z_{it} + e_{it}$ , where  $Z_{it}$  are observed regressors and  $\beta_i$  must be estimated. We leave these calculations for future work.

## Appendix

This appendix summarizes key details of proofs to the results given in the text. A complete set of proofs are available in the detailed appendix (D-Appendix hereafter) available at <http://www.wws.princeton.edu/~mwatson>.

Proof of Lemma 1: This is a version of Theorem 1 and Corollary 1 in Bai and Ng (2002) under slightly different assumptions. See D-Appendix for a detailed proof using the assumptions listed above.

Proof of Lemma 2: Let  $\tilde{\omega}_k$  denote the  $k$ 'th ordered eigenvalue of  $(NT)^{-1} \tilde{X}\tilde{X}'$ . As shown in D-Appendix, Lemma 2 is implied by (i)  $\tilde{\omega}_k - \omega_k = o_p(1)$  for  $k \leq r$  and (ii)  $\tilde{\omega}_k - \omega_k = O_p(s_{NT}^{-\delta})$  for  $k > r$ . To verify (i) and (ii), let  $\mu$  denote the largest eigenvalue of  $(NT)^{-1}bb'$ , then

$$\omega_k + \mu - 2(\omega_k \mu)^{1/2} \leq \tilde{\omega}_k \leq \omega_k + \mu + 2(\omega_k \mu)^{1/2} \quad (2.6)$$

follows from Horn and Johnson 3.3.16 (1991). By the assumption of the lemma,

$\text{trace}(bb') = O_p(s_{NT}^{-1})$ , so that  $\mu = O_p(s_{NT}^{-1})$ . For  $k \leq r$   $\omega_k \xrightarrow{p} \sigma_{kk}$  (D-Appendix R11), so

that  $\tilde{\omega}_k - \omega_k = o_p(1)$  for  $k = 1, \dots, r$  follows from (2.6), and this shows (i). For  $k > r$ ,

$\omega_k = O_p(s_{NT}^{-\delta})$  (D-Appendix R28), thus (2.6) implies  $\tilde{\omega}_k - \omega_k = O_p(s_{NT}^{-1}) + O_p(s_{NT}^{-(1+\delta)/2})$ ,

and this shows (ii).

Proof of Theorem 1: Let  $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_p)$ ,  $\Pi = \Lambda\Phi$ , and  $\mathbf{F}_t = (F'_{t-1}, \dots, F'_{t-p})'$ , so that  $F_t = \Phi\mathbf{F}_t + G\eta_t$  and  $Y_t = X_t - \Pi\mathbf{F}_t$ . Let  $\pi_i'$  denote the  $i$ 'th row of  $\Pi$  and  $\gamma_i'$  denote the  $i$ 'th row of  $\Gamma$ , then  $X_{it} = \eta_i'\gamma_i + \mathbf{F}_t'\pi_i + e_{it}$ . The following results are versions of Theorem 1 in Bai and Ng (2002) (see D-Appendix):

$$T^{-1} \sum_{t=1}^T \left\| \hat{F}_t - J_{NT} F_t \right\|^2 = O_p(s_{NT}^{-1}), \text{ where } J_{NT} \text{ is an } r \times r \text{ matrix that satisfies } J_{NT} \xrightarrow{p} J \text{ a non-singular matrix,} \quad (2.7)$$

$$T^{-1} \sum_{t=1}^T \left\| \hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t \right\|^2 = O_p(s_{NT}^{-1}), \text{ where } \mathbf{J}_{NT} \text{ is an } (pr) \times (pr) \text{ matrix that satisfies } \mathbf{J}_{NT} \xrightarrow{p} \mathbf{J} \text{ a non-singular matrix,} \quad (2.8)$$

$$N^{-1} \sum_{i=1}^N \left\| \hat{\lambda}_i - J_{NT}^{-1'} \lambda_i \right\|^2 = O_p(s_{NT}^{-1}). \quad (2.9)$$

The following lemma is useful:

**Lemma 3:** Let  $\hat{\pi}_i$  denote an estimator of  $\pi_i$  and  $b_{it} = \hat{\mathbf{F}}_t' \hat{\pi}_i - \mathbf{F}_t' \pi_i$ . If

$$N^{-1} \sum_{i=1}^N \left\| \hat{\pi}_i - \mathbf{J}_{NT}^{-1'} \pi_i \right\|^2 = O_p(s_{NT}^{-1}), \text{ then } T^{-1} N^{-1} \sum_{t=1}^T \sum_{i=1}^N b_{it}^2 = O_p(s_{NT}^{-1}).$$

Proof: Write  $\hat{\mathbf{F}}_t = \mathbf{J}_{NT} \mathbf{F}_t + (\hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t)$  and  $\hat{\pi}_i = \mathbf{J}_{NT}^{-1'} \pi_i + (\hat{\pi}_i - \mathbf{J}_{NT}^{-1'} \pi_i)$ , so that

$$b_{it} = \mathbf{F}_t' \mathbf{J}_{NT}' (\hat{\pi}_i - \mathbf{J}_{NT}^{-1'} \pi_i) + (\hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t)' \mathbf{J}_{NT}^{-1'} \pi_i + (\hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t)' (\hat{\pi}_i - \mathbf{J}_{NT}^{-1'} \pi_i). \text{ Thus}$$

$$\begin{aligned}
T^{-1}N^{-1} \sum_{i=1}^T \sum_{i=1}^N b_{it}^2 &\leq \left[ T^{-1} \sum_{i=1}^T \|\mathbf{F}_t\|^2 \right] \|\mathbf{J}_{NT}\|^2 \left[ N^{-1} \sum_{i=1}^N \|\hat{\pi}_i - \mathbf{J}_{NT}^{-1'} \pi_i\|^2 \right] \\
&+ \left[ T^{-1} \sum_{i=1}^T \|\hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t\|^2 \right] \|\mathbf{J}_{NT}^{-1}\|^2 \left[ N^{-1} \sum_{i=1}^N \|\pi_i\|^2 \right] \\
&+ \left[ T^{-1} \sum_{i=1}^T \|\hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t\|^2 \right] \left[ N^{-1} \sum_{i=1}^N \|\hat{\pi}_i - \mathbf{J}_{NT}^{-1'} \pi_i\|^2 \right]
\end{aligned}$$

and the result follows from  $T^{-1} \sum_{i=1}^T \|\mathbf{F}_t\|^2 = O_p(1)$  (A.10),  $\|\mathbf{J}_{NT}\|^2 \xrightarrow{p} \|\mathbf{J}\|^2 < \infty$ ,

$$N^{-1} \sum_{i=1}^N \|\hat{\pi}_i - \mathbf{J}_{NT}^{-1'} \pi_i\|^2 = O_p(s_{NT}^{-1}) \text{ (assumption of the lemma), and } T^{-1} \sum_{i=1}^T \|\hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t\|^2 =$$

$O_p(s_{NT}^{-1})$  (from (2.8)).

Part (a) of Theorem 1: The feasible OLS estimator of  $\Phi$  is

$$\hat{\Phi} = \left[ T^{-1} \sum_{t=p+1}^T \hat{F}_t \hat{\mathbf{F}}_t' \right] \left[ T^{-1} \sum_{t=p+1}^T \hat{\mathbf{F}}_t \hat{\mathbf{F}}_t' \right]^{-1}. \text{ Using } F_t = \Phi \mathbf{F}_t + G \eta_t, \hat{F}_t \text{ can be written as}$$

$$\hat{F}_t = J_{NT} \Phi \mathbf{J}_{NT}^{-1} \hat{\mathbf{F}}_t + J_{NT} G \eta_t + (\hat{F}_t - J_{NT} F_t) - J_{NT} \Phi \mathbf{J}_{NT}^{-1} (\hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t). \text{ Thus}$$

$$\begin{aligned}
\hat{\Phi} - J_{NT} \Phi \mathbf{J}_{NT}^{-1} &= \left[ J_{NT} G T^{-1} \sum_{t=p+1}^T \eta_t \hat{\mathbf{F}}_t' + T^{-1} \sum_{t=p+1}^T (\hat{F}_t - J_{NT} F_t) \hat{\mathbf{F}}_t' - J_{NT} \Phi \mathbf{J}_{NT}^{-1} T^{-1} \sum_{t=p+1}^T (\hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t) \hat{\mathbf{F}}_t' \right] \\
&\times \left[ T^{-1} \sum_{t=p+1}^T \hat{\mathbf{F}}_t \hat{\mathbf{F}}_t' \right]^{-1}
\end{aligned}$$

Straightforward calculations (see D-Appendix) show that each of the terms  $T^{-1} \sum_{t=p+1}^T \eta_t \hat{\mathbf{F}}_t'$ ,

$T^{-1} \sum_{t=p+1}^T (\hat{F}_t - J_{NT} F_t) \hat{\mathbf{F}}_t'$ , and  $T^{-1} \sum_{t=p+1}^T (\hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t) \hat{\mathbf{F}}_t'$  are  $O_p(s_{NT}^{-1/2})$ , and that

$T^{-1} \sum_{t=p+1}^T \hat{\mathbf{F}}_t \hat{\mathbf{F}}_t' \xrightarrow{p} \mathbf{J} E(\mathbf{F}_t \mathbf{F}_t') \mathbf{J}$  which is nonsingular. Thus  $\hat{\Phi} - J_{NT} \Phi \mathbf{J}_{NT}^{-1} = O_p(s_{NT}^{-1/2})$ .

To complete the proof, let  $\hat{\pi}_i = \hat{\Phi}' \hat{\lambda}_i$ , write  $\hat{\lambda}_i = J_{NT}^{-1} \lambda_i + (\hat{\lambda}_i - J_{NT}^{-1} \lambda_i)$  and

$\hat{\Phi} = J_{NT} \Phi \mathbf{J}_{NT}^{-1} + (\hat{\Phi} - J_{NT} \Phi \mathbf{J}_{NT}^{-1})$ , so that  $\hat{\pi}_i - \mathbf{J}_{NT}^{-1} \pi_i = \mathbf{J}_{NT}^{-1} \Phi J_{NT}' (\hat{\lambda}_i - J_{NT}^{-1} \lambda_i)$

+  $(\hat{\Phi} - J_{NT} \Phi \mathbf{J}_{NT}^{-1})' J_{NT}^{-1} \lambda_i + (\hat{\Phi} - J_{NT} \Phi \mathbf{J}_{NT}^{-1})' (\hat{\lambda}_i - J_{NT}^{-1} \lambda_i)$ . Thus  $N^{-1} \sum_{i=1}^N \left\| \hat{\pi}_i - \mathbf{J}_{NT}^{-1} \pi_i \right\|^2 =$

$O_p(s_{NT}^{-1})$  follows from  $\hat{\Phi} - J_{NT} \Phi \mathbf{J}_{NT}^{-1} = O_p(s_{NT}^{-1/2})$  and  $N^{-1} \sum_{i=1}^N \left\| \hat{\lambda}_i - J_{NT}^{-1} \lambda_i \right\|^2 = O_p(s_{NT}^{-1})$ . Part

(a) then follows from Lemma 3.

Part (b) of Theorem 1: The feasible OLS estimator of  $\pi_i$  is

$$\hat{\pi}_i^{OLS} = \left[ T^{-1} \sum_{t=p+1}^T \hat{\mathbf{F}}_t \hat{\mathbf{F}}_t' \right]^{-1} \left[ T^{-1} \sum_{t=p+1}^T \hat{\mathbf{F}}_t X_{it} \right].$$

Using  $X_{it} = \mathbf{F}_t' \pi_i + \eta_t' \gamma_i + e_{it} = \hat{\mathbf{F}}_t' \mathbf{J}_{NT}^{-1} \pi_i - (\hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t)' \mathbf{J}_{NT}^{-1} \pi_i + \eta_t' \gamma_i + e_{it}$ ,

$$\begin{aligned} \hat{\pi}_i^{OLS} - \mathbf{J}_{NT}^{-1} \pi_i &= \left[ T^{-1} \sum_{t=p+1}^T \hat{\mathbf{F}}_t \hat{\mathbf{F}}_t' \right]^{-1} \times \\ &\quad \left[ T^{-1} \sum_{t=p+1}^T \hat{\mathbf{F}}_t (\hat{\mathbf{F}}_t - \mathbf{J}_{NT} \mathbf{F}_t)' \mathbf{J}_{NT}^{-1} \pi_i + T^{-1} \sum_{t=p+1}^T \hat{\mathbf{F}}_t \eta_t' \gamma_i + T^{-1} \sum_{t=p+1}^T \hat{\mathbf{F}}_t e_{it} \right]. \end{aligned}$$

Straightforward calculations (see D-Appendix) show that each of the terms

$$N^{-1} \sum_{i=1}^N \left\| T^{-1} \sum_{t=p+1}^T \hat{\mathbf{F}}_t (\hat{\mathbf{F}}_t - \mathbf{J}_{\text{NT}} \mathbf{F}_t)' \mathbf{J}_{\text{NT}}^{-1} \boldsymbol{\pi}_i \right\|^2, \quad N^{-1} \sum_{i=1}^N \left\| T^{-1} \sum_{t=1}^T \hat{\mathbf{F}}_t \eta_t' \gamma_i \right\|^2, \text{ and}$$

$$N^{-1} \sum_{i=1}^N \left\| T^{-1} \sum_{t=1}^T \hat{\mathbf{F}}_t \eta_t' \gamma_i \right\|^2 \text{ are } O_p(s_{NT}^{-1}), \text{ and that } T^{-1} \sum_{t=p+1}^T \hat{\mathbf{F}}_t \hat{\mathbf{F}}_t' \xrightarrow{p} \mathbf{J} E(\mathbf{F}_t \mathbf{F}_t') \mathbf{J} \text{ which is}$$

nonsingular. Thus  $N^{-1} \sum_{i=1}^N \left\| \hat{\boldsymbol{\pi}}_i^{\text{OLS}} - \mathbf{J}_{\text{NT}}^{-1} \boldsymbol{\pi}_i \right\|^2 = O_p(s_{NT}^{-1})$ , and the result follows from Lemma

3.

## References

- Bai, J. (2003), “Inferential theory for factor models of large dimensions”, *Econometrica* 71:135-171.
- Bai, J., and S. Ng (2002), “Determining the number of factors in approximate factor models”, *Econometrica* 70:191-221.
- Bai, J. and S. Ng (2005a), “Determining the number of primitive shocks in factor models,” *Journal of Business and Economic Statistics*, forthcoming.
- Bai, J. and S. Ng (2005b), “Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions,” manuscript, University of Michigan.
- Bai, J. and S. Ng (2006), “Determining the number of factors in approximate factor models, Errata,” <http://www.umich.edu/~ngse/papers/correctionEcta2.pdf>
- Bernanke, B.S., J. Boivin and P. Elias (2005), “Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach”, *Quarterly Journal of Economics* 120:387-422.
- Forni, M., M. Hallin, M. Lippi and L. Reichlin (2000), “The generalized factor model: identification and estimation”, *The Review of Economics and Statistics* 82:540–554.
- Forni, M., M. Hallin, M. Lippi and L. Reichlin (2005), “The generalized dynamic factor model: One-sided estimation and forecasting”, *Journal of the American Statistical Association* 100: 830-840.
- Giannone, D., L. Reichlin, and L. Sala (2003), “Opening the Black Box: Structural Factor Models versus Structural VARs,” CEPR Discussion Paper No. 4133.

- Giannone, D., L. Reichlin and L. Sala (2004), “Monetary Policy in Real Time”, *NBER Macroeconomics Annual*, 2004.
- Hallin, M. and R. Liška (2005), “The generalized dynamic factor model: determining the number of factor,” mimeo, E.C.A.R.E.S. Université Libre d Bruxelles.
- Horn, R.A. and C.R. Johnson (1991), *Topics in Matrix Analysis*, Cambridge University Press, Cambridge.
- 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: Simulation Results

$$\text{Cov}(e_{it}e_{jt}) = \rho^{|i-j|}$$

A.  $\rho = 0$ 

N	T	$\widehat{BN}^{ICP}(\hat{Y}^A)$				$\widehat{BN}^{ICP}(\hat{Y}^B)$				$\hat{q}_3$				$\hat{q}_4$				$\widehat{BN}^{ICP}(X)$		
		< q	= q	> q	rmse	< q	= q	> q	rmse	< q	= q	> q	rmse	< q	= q	> q	rmse	< r	= r	> r
A. DGP1 $X_{it} = \lambda_i' F_t + e_{it}; F_t = \Phi F_{t-1} + G \eta_t; r = 5, q = 3$																				
20	100	0.00	0.84	0.16	0.84	0.00	0.63	0.37	2.50	0.31	0.69	0.00	0.56	0.15	0.83	0.01	0.41	0.27	0.17	0.56
30	100	0.00	1.00	0.00	0.07	0.00	0.98	0.02	0.17	0.05	0.95	0.00	0.23	0.04	0.96	0.00	0.20	0.64	0.24	0.12
40	100	0.00	1.00	0.00	0.01	0.00	1.00	0.00	0.03	0.01	0.99	0.00	0.09	0.01	0.99	0.00	0.08	0.79	0.18	0.03
50	100	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.05	0.00	1.00	0.00	0.05	0.85	0.14	0.01
100	100	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.86	0.14	0.00
B. DGP2 $X_{it} = \lambda_i' F_t + e_{it}; F_t = \Phi F_{t-1} + G \eta_t; r = 3, q = 3$																				
20	100	0.00	0.85	0.15	0.76	0.00	0.67	0.33	2.31	0.23	0.77	0.00	0.48	0.16	0.84	0.00	0.40	0.00	0.52	0.48
30	100	0.00	1.00	0.00	0.05	0.00	0.98	0.02	0.16	0.03	0.97	0.00	0.16	0.03	0.97	0.00	0.16	0.00	0.92	0.08
40	100	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.03	0.00	1.00	0.00	0.04	0.00	1.00	0.00	0.04	0.00	0.99	0.01
50	100	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00
100	100	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00
C. DGP3 $X_{it} = (\lambda_{i0} + \lambda_{i1}L)' f_t + e_{it}; f_t = \Phi f_{t-1} + \eta_t; r = 4, q = 2$																				
20	100	0.00	0.74	0.26	0.79	0.00	0.51	0.49	2.20	0.03	0.95	0.01	0.22	0.00	0.73	0.26	0.66	0.00	0.09	0.91
30	100	0.00	0.92	0.08	0.28	0.00	0.85	0.15	0.40	0.00	0.99	0.00	0.08	0.00	0.90	0.10	0.41	0.00	0.40	0.60
40	100	0.00	0.98	0.02	0.15	0.00	0.95	0.05	0.23	0.00	1.00	0.00	0.04	0.00	0.96	0.04	0.24	0.00	0.61	0.39
50	100	0.00	1.00	0.00	0.06	0.00	0.99	0.01	0.11	0.00	1.00	0.00	0.02	0.00	0.99	0.01	0.12	0.00	0.76	0.24
100	100	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.99	0.01
D. DGP4 $X_{it} = (\lambda_{i0} + \lambda_{i1}L + \lambda_{i2}L^2)' f_t + e_{it}; f_t = (I + \Theta L) \eta_t; r = 6, q = 2$																				
20	100	0.00	0.73	0.26	0.81	0.00	0.51	0.49	2.09	0.02	0.97	0.00	0.16	0.01	0.92	0.07	0.32	0.00	0.25	0.75
30	100	0.00	0.95	0.05	0.22	0.00	0.88	0.12	0.35	0.00	1.00	0.00	0.04	0.00	0.99	0.01	0.12	0.00	0.69	0.31
40	100	0.00	0.99	0.01	0.08	0.00	0.98	0.02	0.13	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.04	0.00	0.84	0.16
50	100	0.00	1.00	0.00	0.03	0.00	1.00	0.00	0.06	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.93	0.07
100	100	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00

Table 1 (Continued)

B.  $\rho = 0.5$ 

N	T	$\widehat{BN}^{ICP}(\hat{Y}^A)$				$\widehat{BN}^{ICP}(\hat{Y}^B)$				$\hat{q}_3$				$\hat{q}_4$				$\widehat{BN}^{ICP}(X)$		
		< q	= q	> q	rmse	< q	= q	> q	rmse	< q	= q	> q	rmse	< q	= q	> q	rmse	< r	= r	> r
A. DGP1 $X_{it} = \lambda_i' F_t + e_{it}; F_t = \Phi F_{t-1} + G \eta_t; r = 5, q = 3$																				
20	100	0.00	0.00	1.00	6.49	0.00	0.00	1.00	6.97	0.32	0.68	0.00	0.58	0.07	0.87	0.06	0.37	0.00	0.00	1.00
30	100	0.00	0.30	0.70	2.15	0.00	0.02	0.98	5.30	0.06	0.94	0.00	0.25	0.01	0.93	0.07	0.28	0.01	0.03	0.96
40	100	0.00	0.73	0.27	0.67	0.00	0.32	0.68	2.32	0.01	0.99	0.00	0.09	0.00	0.97	0.03	0.17	0.14	0.19	0.67
50	100	0.00	0.92	0.08	0.30	0.00	0.70	0.30	0.86	0.00	1.00	0.00	0.05	0.00	0.99	0.01	0.08	0.35	0.31	0.34
100	100	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.01	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.82	0.18	0.00
B. DGP2 $X_{it} = \lambda_i' F_t + e_{it}; F_t = \Phi F_{t-1} + G \eta_t; r = 3, q = 3$																				
20	100	0.00	0.00	1.00	6.28	0.00	0.00	1.00	6.96	0.27	0.73	0.00	0.52	0.04	0.90	0.06	0.33	0.00	0.00	1.00
30	100	0.00	0.25	0.75	2.14	0.00	0.03	0.97	5.18	0.03	0.97	0.00	0.18	0.00	0.95	0.05	0.24	0.00	0.02	0.98
40	100	0.00	0.71	0.29	0.73	0.00	0.36	0.64	2.15	0.00	1.00	0.00	0.05	0.00	0.99	0.01	0.09	0.00	0.25	0.75
50	100	0.00	0.92	0.08	0.30	0.00	0.73	0.27	0.75	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.60	0.40
100	100	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00
C. DGP3 $X_{it} = (\lambda_{i0} + \lambda_{i1}L)' f_t + e_{it}; f_t = \Phi f_{t-1} + \eta_t; r = 4, q = 2$																				
20	100	0.00	0.00	1.00	7.48	0.00	0.00	1.00	7.96	0.03	0.92	0.04	0.28	0.00	0.58	0.42	0.80	0.00	0.00	1.00
30	100	0.00	0.47	0.53	1.76	0.00	0.08	0.92	5.32	0.00	0.98	0.02	0.15	0.00	0.55	0.45	0.88	0.00	0.00	1.00
40	100	0.00	0.83	0.17	0.50	0.00	0.48	0.52	1.63	0.00	0.99	0.01	0.10	0.00	0.63	0.37	0.81	0.00	0.03	0.97
50	100	0.00	0.95	0.05	0.22	0.00	0.80	0.20	0.58	0.00	1.00	0.00	0.06	0.00	0.79	0.21	0.61	0.00	0.15	0.85
100	100	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.08	0.00	0.90	0.10
D. DGP4 $X_{it} = (\lambda_{i0} + \lambda_{i1}L + \lambda_{i2}L^2)' f_t + e_{it}; f_t = (I + \Theta L) \eta_t; r = 6, q = 2$																				
20	100	0.00	0.01	0.99	7.51	0.00	0.00	1.00	7.96	0.03	0.97	0.01	0.18	0.00	0.81	0.19	0.49	0.00	0.00	1.00
30	100	0.00	0.47	0.53	1.75	0.00	0.10	0.90	5.03	0.00	1.00	0.00	0.05	0.00	0.87	0.13	0.41	0.00	0.01	0.99
40	100	0.00	0.84	0.16	0.48	0.00	0.51	0.49	1.46	0.00	1.00	0.00	0.01	0.00	0.94	0.06	0.28	0.00	0.13	0.87
50	100	0.00	0.96	0.04	0.21	0.00	0.82	0.18	0.53	0.00	1.00	0.00	0.00	0.00	0.98	0.02	0.16	0.00	0.39	0.61
100	100	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.98	0.01

Notes: The first two columns show the values of  $N$  and  $T$  used in the simulations. The next four columns summarize the results for the estimator  $\widehat{BN}^{ICP}(\hat{Y}^A)$ ; the columns labeled <  $q$ , =  $q$  and >  $q$  shows the fraction of estimates that were less than, equal to, and greater than  $q$ ; the column labeled rmse is the root mean square error of the estimates. The same entries are provided for the other estimators of  $q$ . The final three columns summarize the results for the estimates of  $r$ . Results are based on 5,000 simulations.