A Comparison of the Finite Sample Properties of Selection Rules of Factor Numbers in Large Datasets

GUO-FITOUSSI, Liang

ADIS Paris sud

September 2013

A Comparison of the Finite Sample Properties of Selection Rules of Factor Numbers in Large Datasets

Liang GUO-FITOUSSI[a*]

Very Preliminary version, comments welcome.

Abstract

In this paper, we compare the properties of the main criteria proposed for selecting the number of factors in dynamic factor model in a small sample. Both static and dynamic factor numbers' selection rules are studied. Simulations show that the GR ratio proposed by Ahn and Horenstein (2013) and the criterion proposed by Onatski (2010) outperform the others. Furthermore, the two criteria can select accurately the number of static factors in a dynamic factors design. Also, the criteria proposed by Hallin and Liska (2007) and Breitung and Pigorsch (2009) correctly select the number of dynamic factors in most cases. However, empirical application show most criteria select only one factor in presence of one strong factor.

Key words: dynamic factor model, factor numbers, small sample properties

JEL Classification: C13, C52

1 Introduction

The improvements in computer technology, and collection and storage of data, and development of powerful mathematical and statistical software is allowing researchers and professionals in economics and finance to benefit from increasingly rich and increasingly disaggregated data. It is in this context that factor models of large dimensional dataset have been proposed and achieved popularity. The factor models of large datasets are widely applied because they constitute a good compromise between exploiting large amounts of information and parsimonious parameter estimation. For review of recent factor model developments, see Reichlin (2003), Breitung and Eickmeier (2006), Eickmeier and

*Université Paris-Sud, ADIS, 54, BD DESGRANGES, 92331, Sceaux.

Ziegler (2008), Boivin and Ng (2005), Bai and Ng (2008), Guo (2010) and Stock and Watson (2010).

In macroeconomics, factor models are used for nowcasting (Altissimo et al. (2006)), forecasting (Stock and Watson (1998)), construction of indexes, structural analysis (see e.g. Stock and Watson (2005), Bernanke and al. (2005)), and monetary policy (see e.g. Bernake and Boivin (2003)). In finance, they are used to study arbitrage pricing theory (APT) (see e.g. Chamberlain and Rothschild 1983), performance evaluation (Chaps 5 and 6 in Campell et al. 1997), factors in interest term structures (see e.g. Koopman and Van der Wel 2013), asset management strategies such as "momentum trading" (see e.g. Tong (2000)), and credit default correlation (see e.g. Cipollini and Missaglia (2007), Guo (2010)). In practice, the approach of Stock and Watson (1989) for constructing economic indicators is regularly used by NBER economists and the Federal Reserve Bank of Chicago. In Europe, economic and financial institutions use a coincident indicator of economic cycles in the euro zone (EuroCOIN, Forni et al., 2000), published monthly by the London-Based Centre for Economic Policy Research and Banca d'Italia for economic activity analysis and forecasting.

A critical step in the estimation of factor models is selecting the number of latent factors. In classical factor models, one of the most widely used methods is the Kaiser-Guttman criterion (Guttman (1954), Kaiser (1960)), in which only factors with eigenvalues greater than 1 are retained. The underlying idea is that "a factor must account for at least as much variance as an individual variable" (Nunnally and Bernstein (1994)). Another is the Scree test, a graphical tool proposed by Cattell (1966). However, these informal methods are subject to criticisms of vulnerability, subjectivity, and lack of statistical theory (Wislon and Cooper (2008)). Moreover, in presence of cross-sectional and temporal dependence of errors, typical features of macroeconomic and financial data, these methods cannot cleanly reveal the true number of factors (Ahn and Horenstein (2013)). In some economic theories, the number of factors and the factors themselves are imposed rather than being specified by the data, a well-known example is the CAPM. Under the assumption of cross-sectional and temporal dependence of errors, Connor and Korajcyk (1993), Chamberlain and Rothschild (1983), Cragg and Donald (1997), Lewbel (1991) and Donald (1997) propose criteria for selecting the number of factors. However, all of these criteria require one dimension (N or T) of dataset fixed.

For factor models with both N and T approaching infinity, early work on the issue of selection of number of factors includes Stock and Watson (1998), Forni and Reichlin (1998) and Forni et al. (2000). However, the pioneering formal statistical procedure is the information criteria developed by Bai and Ng (2002). Since then, a few researchers have proposed alternative consistent estimators. These estimators can be classified into four types. The first is information criteria, e.g., Bai and Ng (2002), Amengual and Watson (2007), Stock and Watson (2005) and Alessi et al. (2010). The second is based on the theory of random matrices and linked specifically to the proprieties of the largest eigenvalues of the matrix. Representative works are Hallin and Liska (2007, 2010). The third type is based on the rank of a matrix, such as Bai and Ng (2007). The fourth employs canonical correlation analysis. The representative papers are Jacobs and Otter (2008) and Breitung and Pigorsch (2009). However, these estimators are related to each other. For instance, Onatski (2007) shows the relation between the information criteria estimators and the eigenvalue estimators by pointing out that the information type estimator equals the number of eigenvalues greater than a threshold value specified by a penalty function. The criteria proposed by Onatski (2007) and Ahn and Horenstein (2009) exactly exploit this relation.

All these selection criteria deliver a consistent estimator of the number of factors; however, estimated results in finite samples often diverge. Furthermore, although the assumptions are more or less restrictive for different selection rules, most authors argue that their approach can be extended to the more general case. The purpose of this paper is to compare the properties of the main criteria proposed in a small sample and thus help the choice of criteria using different data. To the best of our knowledge, there is no existing complete comparison of criteria, apart from the article by Barhoumi et al. (2013). Compared to Barhoumi et al. (2013), which focuses on forecasting performance, our work focuses on performance of criteria under different assumptions.

The paper is organized as follows. Section 2 summarizes the existing factor models. Section 3 presents the different types of estimators. Section 4 reports Monte Carlo experiments. Section 5 provides two empirical applications, using respectively macroeconomic data and stock return data. Section 6 presents the conclusions.

2 Factor Models

Usually, the factor model is written in a general form as follows:

$$\mathbf{x}_t = \mathbf{\Lambda F}_t + \mathbf{e}_t \tag{1}$$

$\mathbf{x}_t$ are N-dimensional observable variables. When $\mathbf{x}_t$ admit a factorial representation, they can be decomposed into a small number of factors and N idiosyncratic errors. $\mathbf{F}_t$ is an $r-$dimensional vector of common factors, where r denotes the number of factors, $r \ll N$. $\mathbf{\Lambda}$ is an $N \times r$ dimensional matrix containing the factor loadings. We use χ_t to denote the common component, $\chi_t = \mathbf{\Lambda F}_t$. $\mathbf{e}_t$ is $N \times 1$ dimensional idiosyncratic errors. $\mathbf{\Lambda}$, $\mathbf{F}_t$ and $\mathbf{e}_t$ are unobservable.

Specifically, the representation (1) is a static factor model and F_t are termed static factors because the relationship between factors and factor loadings is static. Nonetheless, even in a static model, factors F_t can be "dynamic" in the sense that they can evolve following a dynamic process such as,

Table 1: Comparisons of Assumptions of Different Specification Criteria
Insert table 1

$$\Phi(L)F_t = B(L)\upsilon_t \tag{2}$$

The idiosyncratic errors might also be autocorrelated:

$$\Psi(L)e_{it} = D_i(L)\zeta_{it} \tag{3}$$

where υ_t and ζ_{it} are i.i.d. white noise with $E\|\upsilon_t\|^{4+\delta} < M < \infty$ and $E\|\zeta_{it}\|^{4+\delta} < M < \infty$ for some $\delta > 0$. $\Phi(L)$, $B(L)$, $\Psi(L)$ and $D_i(L)$ are lagged polynomials with roots which all lie outside the unit circle.

The dynamic factor model can be written as follows,

$$\mathbf{x}_t = \lambda_{\mathbf{i}}'(\mathbf{L})\mathbf{f}_t + \mathbf{e}_t \tag{4}$$

$\mathbf{f}_t$ are q-dimensional dynamic factors, where q is the number of dynamic factors. $\lambda_{\mathbf{i}}(\mathbf{L})$ are lagged polynomials with roots outside the unit circle. Factors and idiosyncratic errors follow dynamic processes similar to those in equations (2) and (3). In (1) and (4), both dependence and heteroskedasticity of idiosyncratic errors and dependence between factors and errors are allowed. The assumptions proposed by various researchers differ mainly in relation to the tradeoff between moment constraints and dependence properties of the factors and idiosyncratic errors. We do not report the detailed technical assumptions here[1]; we provide a brief summary in Table 1 to show the differences. We would point out for simplicity, stationarity is assumed, although it is not necessary for some criteria[2].

When $\lambda_i(L)$ are lagged polynomials of limited orders, we call (4) restricted dynamic factor model, which is in contrast to generalized dynamic factor model with $\lambda_i(L)$ of infinite orders. Bai and Ng (2007) show that the restricted dynamic model and the approximated static model can be deduced by mathematical identities. However, notice that only the contemporaneous effects of the factors on the variables are considered in the static model, while lagged dependencies are also allowed in the dynamic model. In addition, they imply different estimation methods. Asymptotic principal components analysis (APCA) could be applied to the sample covariance matrix for estimating factors of static factor models (see Stock and Watson (2002) and others.) However, one could use

[1] We refer the reader to Bai and Ng (2002) Forni et al. (2000) and others for the assumptions.

[2] For non stationary data, Bai and Ng (2004) suggest that the number of factors can be estimated with differenced data.

dynamic principal components analysis (DPCA) in the frequency domain for dynamic factor models (Brillinger 1981, Forni et al. (2000, 2004)). Alternatively, Doz et al. (2006) propose a quasi maximum likelihood approach. Kapetanios and Marcellino (2004) also proposed a parametric method for estimating large approximate factor models. For reviews and comparisons of these estimation methods, see Stock and Watson (2010), Boivin and Ng (2005), Marcellino et al. (2005) and D'agostino and Giannone (2006).

3 Criteria of selection of number of factors

In this section, we discuss various criteria. They are classified in four groups: information criteria type, criteria based on properties of eigenvalues or singular value, criteria exploiting the rank of matrix, and criteria using canonical correlation analysis.

3.1 Information criteria type

As is well-known, the general rule for information criteria is selecting the number of factors which minimizes the variance explained by the idiosyncratic component. A penalty function is introduced in order to avoid overparameterization. The choice of penalty function is often related to the rate of convergence of the estimators. Standard criteria AIC and BIC are good examples. However, these criteria are not applicable in large factor models because the factors are unobservable and do not take account of the double dimensions (T and N).

3.1.1 Estimation of static factors

I. Bai and Ng (2002) Bai and Ng (2002) modify AIC and BIC by taking account of both dimensions n and T of the dataset and suggest criteria PC_p to specify the number of static factors r:

$$PC(k) = V(k) + kp(n,T) \tag{5}$$

where $V(k) = (nT)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} (X_{it} - \hat{\lambda}_i^{k\prime} \hat{F}_t^k)^2$, $\hat{\lambda}_i^k$ and $\hat{F}_t^k$ are the APCA estimators of the factor loadings and factors, the superscript k signifies k static factors are used.

The selected number of factors should minimize $PC(k)$, i.e.,

$\widehat{k} = \arg\min_{0 \leq k \leq r\max} PC(k)$, where $rmax$ is a predetermined bounded integer.

As for AIC et BIC, $V(k)$ should be small if $k > r$. To avoid underestimation and overestimation, the penalty function must satisfy the conditions (i) $p(n,T) \to 0$ and (ii) $C_{N,T} \cdot p(n,T) \to \infty$ when $n, T \to \infty$, where

$C_{n,T} = min\left[\sqrt{n}, \sqrt{T}\right]$ (See Theorem 2 of Bai and Ng (2002)). The intuition behind these conditions is that the penalty function $p(n,T)$ converges to zero but less quickly than the convergence rate of estimator of factors, which is proven to be $C_{n,T}^{-1}$ by Bai and Ng (2002). Therefore, the penalty function approaches zero but it "dominates the difference in the sum of squared residuals between the true and the overparameterized model" (Bai and Ng (2002)). Another class of criteria allowing a consistent estimator of r is proposed by Bai and Ng (2002), and is the logarithmic version of $PC(k)$. For each classe of criteria, Bai and Ng (2002) propose three specific formulations . Since IC_{p1} and PC_{p1} are shown to be more robust than the others by the Monte Carlo Simulation in Bai and Ng (2002)[3], we consider only these two criteria in this paper,

$$PC_{p1}^{BN02}(k) = V(k, \widehat{F^k}) + k\widehat{\sigma}^2 \left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right) \tag{6}$$

$$IC_{p1}^{BN02}(k) = \ln(V(k, \widehat{F^k})) + k \left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right) \tag{7}$$

where $\widehat{\sigma}^2$ is a consistent estimate of $(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} E(e_{it})^2$. Bai and Ng (2002) suggest that $\widehat{\sigma}^2$ can be replaced by $V(rmax, \widehat{F^{rmax}})$ in reality. However, this implies that PC depend directly on the choice of $rmax$ (Alessi et al. (2008) and Forni et al. (2007)).

The criteria of Bai and Ng (2002) have attracted two criticisms. One is that the estimators need to pre-specify a maximum possible number of factors, $rmax$ (Ahn and Horenstein (2013)). Although Schwert (1989) suggests using $8int\left[(T/100)^{1/4}\right]$as a rule to set $rmax$ for time series analysis, no guide is available for panel analysis. Bai and Ng (2002) suggest an arbitrary choice, $8int\left[(c_{N,T}^2/100)^{1/4}\right]$, for large dimensional factor models without proofs. Another problem is that the threshold can be arbitrarily scaled. Namely, if $p(N,T)$ leads to consistent estimation of r, so does $\alpha p(N,T)$, where $\alpha \in R^+$. As pointed out by Hallin and Liska (2007) and Alessi et al. (2008), although multiplying the penalty function by an arbitrary constant has no influence on the asymptotic performance of the criteria, the result can be affected in a finite sample. Finally, in the applications in D'Agostino and Giannone (2013), Ahn and Horenstein (2013), Forni et al (2009) and Alessi et al. (2008), Bai and Ng (2002)'s criteria lead to underestimation and/or overestimation of number of factors in practice.

[3]Basically, the difference between IC_{p1}, PC_{p1} and the other criteria resides in use of the term $\frac{n+T}{NT}$ or use of the convergence rate $C_{N,T}$. Note the convergence rate fails to take account of both dimensions. For example, we obtain the same $C_{N,T}$ for $N = 50, T = 50$ and $N = 200, T = 50$, while the estimation error is smaller in the latter case. According to Bai and Ng (2002), the term $\frac{N+T}{NT}$ provides a small correction to the convergence rate and the authors' simulations show that it has a desirable upwards penalty adjustment effect.

II. Alessi et al. (2008) One of the criticisms of Bai and Ng's (2002) criteria, related to the degree of freedom in penalty function, is exploited by Alessi et al. (2008), who propose a refinement of the criteria in Bai and Ng (2002). The idea was inspired by Hallin and Liska (2007), who proposed selection criteria for dynamic factors (c.f section 3.3). Instead of using one specific penalty function, Alessi et al. (2008) evaluate a whole family of penalty functions. In particular, they propose the following information criteria based on $IC_{p1}(k)$ of Bai and Ng (2002)[4]:

$$IC_a^{ABC}(k) = \ln(V(k, \widehat{F^k})) + \alpha k \left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right) \tag{8}$$

The arbitrary positive real number α is called a tuning parameter, and tunes the penalizing power of the function. The estimated number of factors is $\widehat{k}_\alpha = \arg\min_{0\leq k \leq k\max} IC_a^{ABC}(k)$, which depends on the choice of α. The calibration of α is carried out in the following steps: First, the author set un upper bound for the constant α, $\alpha \in [0, \alpha max]$. Next, J subsamples of size (n_j, T_j) are considered, with $j = 0, \ldots, J$, $0 < n_1 < \ldots < n_J = n$ and $0 \leq T_1 \leq \ldots < T_J = T$. For each j, the number of the factors, denoted by $\hat{k}_{\alpha,n_j}^{T_j}$, is computed. If there exists an interval $[\underline{\alpha}, \bar{\alpha}]$ of α which has a stable behavior, i.e., the number of factors $\hat{k}_{\alpha,n_j}^{T_j}$ is constant across subsamples of different sizes, this means that the choice of α has not been affected by the size of the sample. This number $\hat{k}_{\alpha,n_j}^{T_j}$ is then the estimated number of factors.

Following the notation in Hallin and Liska (2007), the stability is measured by the empirical variance of $\hat{k}_{\alpha,n_j}^{T_j}$:

$$S_\alpha = \frac{1}{J}\sum_{j=1}^{J}\left[\hat{k}_{\alpha,n_j}^{T_j} - \frac{1}{J}\sum_{j=1}^{J}\hat{k}_{\alpha,n_j}^{T_j}\right] \tag{9}$$

This procedure is termed tuning-stability checkup procedure in Ahn and Horenstein (2013). The estimator has the same asymptotic properties as the original criteria, while it conveys a more robust estimation of the number of factors than it would were the penalty fixed.

3.1.2 Estimation of dynamic factors

Stock and Watson (2005) and Amengual and Watson (2007) To estimate the number of dynamic factors in a restricted dynamic model, Stock and Watson (2005) propose a modification of Bai and Ng's (2002) estimator. The proof of the consistency properties of the estimator is given by Amengual and Watson (2007). The modification is straightforward. Precisely, they assume

[4] Another criterion proposed by Alessi et al. (2008) is based on $IC_{p2}(k)$ of Bai and Ng (2002), for the reason given in fn 2, it is not reported here.

Table 2: Summary of eigenvalues criteria

insert table 2

that F_t is a VAR(p) process, i.e. (2) becomes $\Phi(L)F_t = \upsilon_t$, with $\Phi(L) = I - A_1 L - \cdots - A_p L^p$, and the innovations can be represented as $\upsilon_t = G\eta_t$,where G is $r \times q$ dimensional full column rank matrix and η_t is i.i.d. shocks. It follows that the number of common shocks is identical to the number of dynamic factors q. To estimate q, a two-step procedure is proposed. In the first step, the static factors are estimated from x_t using the APCA estimator and the number of static factors is determined by applying Bai and Ng's (2002) information criteria. In the second step, the number of dynamic factors is estimated by applying again Bai and Ng's (2002) information criteria to the sample covariance matrix of estimated innovations, which is obtained as the residual of a regression of x_t on lags of x_t and $\hat{F}_t$.

3.2 Application of theory of random matrix and eigenvalue properties

The second type of selection rules is based on some results developed according to the theory of random matrix and especially the eigenvalues' properties. The basic idea is that if the variables admit an r factor structure, the r largest eigenvalues in the sample covariance matrix should explode, while the rest should tend to 0. Thus, the number of eigenvalues diverging as N, T diverge is equal to the number of factors. The first exploration of properties of eigenvalues goes back to the Scree test introduced by Cattell (1966) in psychology. Cattell (1966) states that if one plots the decreasing eigenvalues in the sample covariance matrix of the data against their respective order numbers, the plot shows a sharp break when the true number of factors ends, which is the so-called "scree" corresponding to the beginning of idiosyncratic effects. However, the Scree test remains a visual inspection. Another heuristic eye-inspection rule based on the relative size of the eigenvalues is proposed by Forni et al. (2000) in frequency domain. More formal tests were developed by Kapetanios (2004, 2010) and Onatski (2009, 2010). Table 3 presents a summary of the eigenvalues criteria.

3.2.1 Estimation of static factors

III. Onatski (2009) Onatski (2009) develops a sequential procedure by applying the asymptotic distribution of the eigenvalues, namely, a few scaled and centered largest eigenvalues of the covariance matrix of a particular Hermitian random matrix, which asymptotically distribute as a Tracy-Widom of type 2 (TW_2, Tracy and Widom (1994)) as T grows noticeably faster than n [5]. More-

[5] The assumption that T grows faster than n is obviously not realistic in the macroeconomic application. While the Monte Carlo simulations in Onatski (2009) show the test developed works well even when n is much larger than T.

over, Onatski (2009) constructs a statistic by taking the ratio of the difference in adjacent eigenvalues, which gets rid of both the centering and scaling parameters of the eigenvalues. The selection rule in Onatski (2009) is developed for a generalized dynamic factor model, while it is also applicable for approximate factors. In the case of approximate factors, the selection procedure consists of:

1. Divide the sample to two subsamples of equal length, multiplying the second half by imaginary unit i, $\hat{X}_j = X_j + iX_{j+\frac{T}{2}}$ compute the discrete Fourier transformation $\hat{X}_j = \sum_{t=1}^{T} X_t \cdot e^{-iw_j t}$ of the data at frequencies $\omega_j = \frac{2\pi s_j}{T}$.

2. Compute i-th largest eigenvalue of the smoothed periodogram estimate $\frac{2}{T}\sum_{j=1}^{T/2} \hat{X}_j \hat{X}_j'$[6], μ_i, and construct the statistic.

$$R^{O09} \equiv max_{k_0<i<k_i} \frac{\mu_i - \mu_{i+1}}{\mu_{i+1} - \mu_{i+2}} \tag{10}$$

Under the null, statistic R converges in the distribution to $max_{k_0<i<k_i} \frac{\lambda_i - \lambda_{i+1}}{\lambda_{i+1} - \lambda_{i+2}}$, where λ_i are random variables with joint multivariate TW_2 distribution[7]. Under the alternative, R explodes since μ_k explodes while μ_{i+1} and μ_{i+2} are bounded. A table of critical values of test statistic is given in Onatski (2009). The null is rejected if and only if R is larger than or equal to the critical value.

VI. Onatski (2010) Another selection procedure is developed by Onatski (2010), based on the structure of the idiosyncratic component in the data. He imposes a structure on the idiosyncratic components in the data: $e = A\varepsilon B$, where A and B are two unrestricted deterministic matrices, and ε is an $N \times T$ matrix with i.i.d. gaussian entries[8]. Thus, both the cross-sectional and temporal correlation of the idiosyncratic components are allowed. Besides, comparing the assumptions about proportional growth to n of the cumulative effect of factors of Bai and Ng (2002, 2007), Onatski (2010) assumes only the cumulative effect of the "least influential factors" diverges to infinity in probability as $n \to \infty$. This assumption allows the existence of some "weak" factors whose explanatory power does not proportionally increase with N. However, instead of a closed form expression of the upper bound on the idiosyncratic eigenvalues, Onatski (2010) derives an implicit function for the upper bound. As proved by Zhang (2006), when the idiosyncratic components are non-trivially correlated both cross-sectionally and temporally, the eigenvalue distribution of $ee'/T(n)$ converges a.s. to non random cdf $F^{\kappa,A,B}$ (a sample size of n and$T(n)$ is assumed with $n/T(n) \to \kappa > 0$ as $n \to \infty$) (Zhang 2006, Theorem 1.2.1). However, $F^{\kappa,A,B}$ is a complicated function without explicit form. Onatski (2010) shows

[6] In the case of the estimation in frequency domain, the "prime" denotes the conjugate-complex transpose of the matrix.

[7] λ_i is the i-th largest eigenvalue of a complex Wishart $W_n^C(m, S_n^e(\omega_0)$ of dimension n and degrees of freedom m, $S_n^e(\omega_0)$ is the spectral density matrices of $e_t(n)$ at frequency ω_0.

[8] For non-Gaussian ε, either A or B is required to be diagonal the other remaining unrestricted.

that any finite number of the largest of the bounded eigenvalues in the sample covariance matrix cluster around a single point, $u(F^{\kappa,A,B})$, where $u(\cdot)$ denotes the upper bound of the support of the distribution $F^{\kappa,A,B}$. Thus, for any $k > r$, the difference between the two adjacent eigenvalues $\mu_k - \mu_{k+1}$ converges to zero, while $\mu_r - \mu_{r+1}$ diverges to infinity. Onatski (2010) defines a family of estimators:

$$\hat{r}^{O10}(\delta) = max\,\{k \leq rmax^n : \mu_k - \mu_{k+1} \geq \delta\} \tag{11}$$

where δ is a positive number.

The procedure to estimate the number of the factors is:

1. Compute the eigenvalue in the sample covariance matrix for the normalized data.

2. Set $j = rmax + 1$, run OLS regression of $\mu_j, \cdots, \mu_{j+4}$ on the constant and $(j-1)^{2/3}, \cdots, (j+3)^{2/3}$and denote the slope coefficient $\hat{\beta}$[9]. $\delta = 2 \mid \hat{\beta} \mid$.

3. If $\lambda_k - \lambda_{k+1} < \delta$ for all $k < rmax$, $\hat{r}(\delta) = 0$; otherwise, a factor structure exists, compute $\hat{r}(\delta) = max\,\{k \leq r^n_{max} : \mu_k - \mu_{k+1} \geq \delta\}$.

4. Set $j = \hat{r}(\delta) + 1$, repeat the step 2 and 3 until convergence.

V. Ahn and Horenstein (2013a) Similar to Onatski (2010), assumptions about cross-section and serial correlations on idiosyncratic component are imposed in Ahn and Horenstein (2013): $e = R_T^{1/2} \varepsilon G_N^{1/2}$. where R_T and G_N are positive semi-definite matrices, and ε is an $T \times N$ matrix with i.i.d. entries. Thus, both cross-sectional and temporal correlation of the idiosyncratic components are allowed. Furthermore, the smallest eigenvalue of R_T is bounded below by a positive number. That is to say, none of $\mathbf{e}_{it}$ and their linear functions can be perfectly predicted by their past values. The smallest eigenvalue of G_N is allowed to be zero, as long as an asymptotically no negligible number of eigenvalue of G_N are bounded below by a positive number. These assumptions are suitable for macroeconomic and financial data, where the variables are highly (perhaps perfectly) correlated, thus the smallest of eigenvalue of G_N could be zero.

The statistic proposed by Ahn and Horenstein (2013), "Eigenvalue Ratio" (ER) estimator, is obtained simply by maximizing the ratio of the two adjacent eigenvalues arranged in descending order:

$ER^{AH}(k) \equiv \frac{\tilde{\mu}_k}{\tilde{\mu}_{k+1}}$, $k = 1, 2, \ldots, kmax$

The idea is the ratio of the $r-th$ and $r+1-th$ eigenvalues of (XX'/TN) diverges to infinity, while all other ratios are asymptotically bounded. The estimators of r is the solution to the problem of maximization of $ER(k)$: $k = max_{1 \leq k \leq rmax} ER(k)$.

[9]The OLS regression is justified by the fact that $F^{\kappa,A,B}$ can be approximated by $1 - a((u-x)^+)^{3/2}$ for some positive a in the neighborhood of u, $(u-x)^+$ stands for the positive part of u-x. The choice of five regressors is suggested by the Monte Carlo simulations results in Onatski (2010). See Onatski (2010) for more details of the calibration of δ.

3.2.2 Estimation of dynamic factors

Onatski (2009) The estimation procedure for static factor number of Onatski (2009), $R \equiv max_{k_0<i<k_i} \frac{\mu_i - \mu_{i+1}}{\mu_{i+1} - \mu_{i+2}}$ (cf section 3.1.2) is also applicable to the number of dynamic factors. In this case, μ_i in the step 2 of this procedure is the i-th largest eigenvalue of the smoothed periodogram estimate $\frac{1}{2\pi m}\sum_{j=1}^{m} \hat{X}_j \hat{X}_j'$ of the spectral density of the data at frequency ω_0. The rest of the procedure is identical.

3.3 Information criteria based on properties of eigenvalues

Before introducing Hallin and Liska (2007), a brief discussion about the relation between the information criteria and eigenvalue is needed. In the information criteria, the PCA estimator can be considered as a solution to the problem of minimization of $V(k)$. While regression of X_{it} on the first k principal components is based on the eigenvalue. Thus, the information criteria estimator and eigenvalue estimator are tightly related. As pointed out by Onatski (2010), $V(k)$ and $\hat{\sigma}^2$ in (6) are respectively equal to $(nT)^{-1}\sum_{j=k+1}^{n}\mu_j$ and $(nT)^{-1}\sum_{j=rmax+1}^{n}\mu_j$, which means the information criteria are also based on the empirical distribution of eigenvalues.

Hallin and Liska (2007) Hallin and Liska (2007) develop information criteria in frequency domain to estimate the number of dynamic factors. The basic idea is similar to Bai and Ng (2002). Due to the complexity of the spectral technique, rather than using the expected mean of squared residuals as in (5), Hallin and Liska (2007) employ the average contribution of the bounded eigenvalue of the spectral density matrix. With the assumption that the divergence rate of the smallest diverging eigenvalue is n, the information criterion is of following form:

$$IC_n^{HL}(k) = \frac{1}{n}\sum_{i=k+1}^{n}\int_{-\pi}^{\pi}\mu_{ni}(\theta)d\theta + \alpha k p(n) \qquad (12)$$

where $\mu_{ni}(\theta)$ is the i-th eigenvalue , $\sum_n(\theta)$. As in Alessi et al. (2008), α, which is an arbitrary positive real number, is the tuning parameter.

For a finite sample, lag window estimation method is suggested by Hallin and Liska (2007) and the information criteria are:

$$IC_{1;n}^{T,HL}(k) = \frac{1}{n}\sum_{i=k+1}^{n}\frac{1}{2M_T+1}\sum_{l=-M_T}^{M_T}\mu_{ni}^{T}(\theta_l) + \alpha k p(n,T) \qquad (13)$$

$$IC_{2;n}^{T,HL}(k) = log\left[\frac{1}{n}\sum_{i=k+1}^{n}\frac{1}{2M_T+1}\sum_{l=-M_T}^{M_T}\mu_{ni}^{T}(\theta_l)\right] + \alpha k p(n,T) \tag{14}$$

with $\theta_l := \frac{\pi l}{M_T+1/2}$ for $l = -M_T, \ldots, M_T$, $M_T > 0$ a truncation parameter, $0 \leq k \leq qmax$. $qmax$ is a predetermined upper bound. $\mu_{ni}^{T}(\theta_l)$ are the eigenvalues of the lag window estimator of sample spectral density matrix. The penalty function satisfies two conditions (i) $p(N,T) \to 0$ and (ii) $min\left[n, M_T^2, M_T^{-1/2}T^{1/2}\right] \cdot p(T,N) \to \infty$ when $N,T \to \infty$ (see proposition 2 of Hallin and Liska (2007)). Three forms of penalty are proposed in Hallin and Liska (2007):

$$p_1(n,T) = (M_T^2 + M_T^{-1/2}T^{1/2} + n^{-1})log\left(min\left[n, M_T^2, M_T^{-1/2}T^{1/2}\right]\right) \tag{15}$$

$$p_2(n,T) = \left(min\left[n, M_T^2, M_T^{-1/2}T^{1/2}\right]\right)^{-1/2} \tag{16}$$

$$p_3(n,T) = \left(min\left[n, M_T^2, M_T^{-1/2}T^{1/2}\right]\right)^{-1} log\left(min\left[n, M_T^2, M_T^{-1/2}T^{1/2}\right]\right) \tag{17}$$

The calibration of α is the same as described for the criteria of Alessi et al. (2008) (cf section 3.1.1.II).

Ahn and Horenstein (2013b) Ahn and Horenstein (2013) proposed another related statistics: "Growth Ratio" (GR) estimator, which is the ratio of growth rates of residual variances as one fewer principal component is used in the time series regressions:

$$GR^{AH}(k) \equiv \frac{ln\left[V(k-1)/V(k)\right]}{ln\left[V(k)/V(k+1)\right]} \tag{18}$$

where $V(k)$=$\sum_{j=k+1}^{m}\widetilde{\mu}_{NT,j}$, and $k = argmax_{1\leq k\leq rmax}GR(k)$

3.4 The rank of the matrix

Based on the rank of the spectral density matrix, Bai and Ng (2007) propose an alternative criteria for selecting the number of dynamic factors. The factors are assumed to evolve as a VAR as in Stock and Watson (2005). Then, the r static factors can be dynamically related, and the spectrum of the static factors has reduced rank, which is actually the number of dynamic factors (or "primitive shocks" according to authors). In other words, the rank of the covariance matrix of υ_t, $\Sigma_\upsilon = E(\upsilon_t \upsilon_t')$, is equal to the number of dynamic factors.

Specifically, Bai and Ng (2005) define two statistics:

$$D_{a,k} = \left(\frac{\beta_{k+1}^2}{\sum_{j=1}^{r} \beta_j^2} \right)^{1/2} \tag{19}$$

and

$$D_{b,k} = \left(\frac{\sum_{j=k+1}^{r} \beta_j^2}{\sum_{j=1}^{r} \beta_j^2} \right)^{1/2} \tag{20}$$

where $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_r$ are the ordered eigenvalues of $\widehat{\Sigma}_u$. With a matrix of rank $q \leq r$, the $r - q$ smallest eigenvalues are zero. Thus, $D_{a,k} = D_{b,k} = 0$ if $k \geq q$.

The estimation of the number of the dynamic factors is carried in several steps. First, the principal component estimators of the static factors, $\hat{F}_t$, are obtained. Next the residuals $\widehat{u}$ are obtained from estimation of a VAR in $\hat{F}_t$ and construct $\widehat{\Sigma}_u = \frac{1}{T} \sum_{t=1}^{T} \widehat{u}_t' \widehat{u}_t$. Then, $\widehat{D}_{a,k}$ and $\widehat{D}_{b,k}$ are calculated from $\widehat{\Sigma}_\upsilon$. Bai and Ng (2007) suggest two selection rules (Proposition 2 of Bai and Ng (2007)):

$$\kappa_a = \left\{ k : \hat{D}_{a,k} < g/min \left[n^{1/2-\delta}, T^{1/2-\delta} \right] \right\} \tag{21}$$

$$\kappa_b = \left\{ k : \hat{D}_{b,k} < g/min \left[n^{1/2-\delta}, T^{1/2-\delta} \right] \right\} \tag{22}$$

for some $0 < g < \infty$ and $0 < \delta < 1/2$, and $\widehat{q}_a = min \{k \in \kappa_a\}$, $\widehat{q}_b = min \{k \in \kappa_b\}$.

In other words, q is the smallest k such as $\widehat{D}_{a,k}$ and $\widehat{D}_{b,k}$ are asymptotically zero. Since we know $\left\| \widehat{\Sigma}_u - H^* \Sigma_u H^{*\prime} \right\| = O_p(1/\delta_{n,T})$[10]. By continuity of eigenvalue, we have $\widehat{D}_{a,k} - D_{a,k} = O_p(\delta_{n,T}^{-1})$ and $\widehat{D}_{b,k} - D_{b,k} = O_p(\delta_{n,T}^{-1})$. For $k \geq q$, since $D_{a,k} = D_{b,k} = 0$, then $\widehat{D}_{a,k} = O_p(\delta_{n,T}^{-1})$. Thus, when n, $T \to \infty$, $\widehat{D}_{a,k} < M/min \left[N^{1/2-\epsilon}, T^{1/2-\epsilon} \right]$ with probability tending to 1 , which implies $q \in \kappa_a$ for large N, T. Whereas $q - 1 \notin \kappa_a$ since $\widehat{D}_{a,k} \geq g > 0$ if $k \prec q$, then $\widehat{D}_{a,k} > M/min \left[n^{1/2-\epsilon}, T^{1/2-\epsilon} \right]$ with probability tending to 1. $M/min \left[n^{1/2-\epsilon}, T^{1/2-\epsilon} \right]$ is the tolerated error induced by the estimation. Based on Monte Carlo simulations, Bai and Ng (2007) suggest $\epsilon = 0.1$ and the choice of M depends on whether the estimation is based on covariance matrix or correlation matrix of VAR residuals.

[10] See Bai and Ng (2002) theorem 1 and Bai and Ng (2007) proposition 1 and lemma 2

3.5 Canonical correlation analysis

Breitung and Pigorsch (2009) Instead of using PCA, Breitung and Pigorsch (2009) develop a selection procedure based on canonical correlation analysis (CCA). Compared to PCA, CCA is invariant to any rotation of the factor space. In particular, the procedure is based on,

$$| \hat{\mu}_j^* \hat{S}_{00} - \hat{S}_{01} \hat{S}_{11}^{-1} \hat{S}_{01}' |= 0 \tag{23}$$

where $\hat{S}_{00} = \sum_{t=s+1}^{T} \hat{F}_t \hat{F}_t{}'$, $\hat{S}_{01} = \sum_{t=s+1}^{T} \hat{F}_t \hat{G}_{t-1}'$,$\hat{S}_{11} = \sum_{t=s+1}^{T} \hat{G}_{t-1} \hat{G}_{t-1}$ and $\hat{G}_{t-1} = \left[\hat{F}_{t-1}', \ldots, \hat{F}_{t-s}'\right]'$, s is the lag order.

$\hat{\mu}_j^*$, the generalized eigenvalues resulting from (13), represent the canonical correlations between the current and past values of the r static factors, respectively denoted by F_t and G_{t-1}. We use * to distinguish them from the eigenvalues of PCA estimators. The generalized eigenvalues are equivalent to the R^2 measure of a regression of the associated linear combination of F_t on G_{t-1} (Anderson 1984). Hence, they are scale invariant and $0 \leq \hat{\mu}_j^* \leq 1$. The motivation for using $\hat{\mu}_j^*$ is that if some lags of the factors enter the static representation F_t, these lags are perfectly predictable from linear combinations of G_{t-1}. Thus, the corresponding canonical correlations (eigenvalues) converge to unity as the sample size tends to infinity, whereas the remaining eigenvalues converge to values smaller than 1. Furthermore, the convergence rate is given by Breitung and Pigorsch (2009, Theorem 1), i.e., for $j = 1, \cdots, r-k$, $\left(1 - \tilde{\mu}_j^*\right) = O_p(C_{N,T}^{-2})$ while $p\left(1 - \tilde{\mu}_j^* > M\right) \to 1$ for some constant $M > 0$ if $j > r - k$. It follows for $j = 1, \cdots, r-k$, $C_{NT}^{2-\delta}\left(1 - \tilde{\mu}_j^*\right)$ converges to zero with $0 < \delta < 2$, while $C_{NT}^{2-\delta}\left(1 - \tilde{\mu}_j^*\right)$ tends to infinity if $j > r - k$. The statistic constructed by Breitung and Pigorsch (2009) is:

$$\xi(k^*) = \widetilde{C}_{NT}^{2-\delta} \sum_{j=1}^{r-k^*} (1 - \tilde{\mu}_j^*) \tag{24}$$

As in BN 02, the convergence rate is replaced by $\widetilde{C}_{NT}^{-2} = \frac{N+T}{NT}$ to take account of the two sampling dimensions.

The number of dynamic factors is determined by choosing the smallest number k in the sequence $k^* = r, r-1, \cdots, 1$, where $\xi(k^*)$ is smaller than some fixed threshold τ: $\widehat{q} = \min\{k^* : \xi(k^*) < \tau\}$. Finally, based on the Monte Carlo simulations, Breitung and Pigorsch (2009) suggest $\delta = 0.5$ and $\tau = 4.5$ for a fraction of explained variance of $[0.4, 0.9]$.

3.6 Singular value

Otter, Jacobs and den Reijer (2011) Otter et al. (2011) propose an alternative criterion based on singular values instead of frequency domain eigenvalues. Furthermore, different from other studies, no explicit factor model is assumed. Instead, they are interested in the simple fact that if a factor structure is appropriate. In particular, applying the singular value decomposition to $n \times T$ stationary normalized random matrix X (with covariance matrix $\sum_x$, $tr(\sum_x) = \sqrt{nT}$), one has $X = USC'$, with $S = diag(\sigma_1, \sigma_2, \cdots, \sigma_m)$, $\sigma_1 > \sigma_2 \cdots > \sigma_m$.

$$E\left(\|X\|^2\right) = tr(X'X) = \sum_{j=1}^{k} \sigma_j^2$$

X has the factor structure if there exists a $r < min(n, T)$ such that for $j \leq r$, $\sigma_j^2 = O(\sqrt{nT})$ and for $j > r$, $\sigma_j^2 = o(\sqrt{nT})$.

Then, for the estimated covariance matrix,

$$E\left(\left\|\frac{X}{\sqrt{nT}}\right\|^2\right) = tr(C\frac{S}{\sqrt{nT}}U'U\frac{S}{\sqrt{nT}}C') = \sum_{j=1}^{k} \frac{\sigma_j^2}{nT} = 1$$

The singular values of $\frac{X}{\sqrt{nT}}$ is thus $\frac{\sigma_j^2}{nT}$, which can be denoted by s_j.

Furthermore, the scaled data matrix $\frac{X}{\sqrt{nT}}$ can be decomposed as

$$\frac{X}{\sqrt{nT}} = U_1 S_1 C_1 + U_2 S_2 C_2 = F + \varepsilon$$

with $S_1 = diag(s_1, s_2, \cdots, s_r)$, $S_2 = diag(s_{r+1}, s_{r+2}, \cdots, s_{min(n,T)})$,

While the Euclidean norm of x_t can be decomposed as

$$E\left(\left\|\frac{X}{\sqrt{NT}}\right\|^2\right) = E\left(\left\|\frac{F}{\sqrt{NT}}\right\|^2\right) + E\left(\left\|\frac{\varepsilon}{\sqrt{NT}}\right\|^2\right) = \sum_{j=1}^{r} s_j^2 + \sum_{j=r+1}^{m} s_j^2 = 1$$

and $E\left(\left\|\frac{F}{\sqrt{NT}}\right\|^2\right) = \sum_{j=1}^{k} s_j^2 = ks_j^2 + \sum_{j=1}^{k} \delta_j(k)$, with $\delta_j(k) = s_j - s_k > 0$, $j = 1, 2, \ldots, k-1$. Thus, $E\left(\|F_t\|^2\right)$ has a lower bound $J(k) = ks_j^2$, which can be viewed as a tradeoff between k and σ_k^2/NT. Denote the difference of $J(k)$ by $DJ(k)$ [11]. Since for $j \leq r$, $DJ(k)$ will be positive, and for $j > r$, $DJ(k)$ will be zero as $N, T \to \infty$. Otter et al. (2013) suggest to use $DJ(k)$ to determine the number of factors, i.e.,$k = argmin(DJ(k))$.

[11] $DJ(k) = \frac{\triangle J(k)}{k} = \frac{J(K+1)-J(k)}{K+1-K} = J(K+1) - J(k)$, which could be written as $\frac{k}{NT}(\sigma_k^2 - \sigma_{k+1}^2) + \frac{1}{NT}\sigma_{k+1}^2$

4 Simulation

In this section, we provide a detailed Monte Carlo study to evaluate the performance of each selection procedure in finite samples. First, we assess the performance of PC and IC in Bai nd Ng (2002) (hereafter BN02a, BN02b respectively), Alessi et al. (2007, hereafter ABC), Onatski (2009, 2010, Ona09 and Ona10 respectively), the ER and GR ratio of Ahn and Horenstein (2013, AH_ER and AH_GR respectively) to determine the number of static factors. Then, we assess the performance of Onatski (2009, hereafter Ona09), Hallin and Liska (2007, HL), Breitung and Pigorsch (2007, BP) and Otter et al. (2013, OJR) to determine the number of dynamic factors. All computations are performed with MATLAB R2013a.[12]

4.1 Static factors

To investigate the properties of different criteria to determine the number of static factors, two part simulations are conducted. In the first part (DGP 1-5), we are interested in the effects of the error covariance structure. In the second part (DGPs 6-8), we investigate the effect of the presence of weak and strong factors. The strongly correlated idiosyncratic errors and presence of strong factors are both meaningful in macroeconomic and financial applications. The simulation experiment design is adopted from Bai and Ng (2002) and Ahn and Horenstein (2013).

The base model is

$$x_{it} = \sum_{j=1}^{r} \lambda_{i,j} F_{jt} + \sqrt{\theta} e_{i,t}$$

with $e_{i,t} = \rho e_{i,t-1} + \upsilon_{it} + \sum_{j\neq 0,j=-J}^{J} \beta \upsilon_{i-j,t}$, $\upsilon_{i,t} \sim N(0,1)$.

$F_{j,t}$ and $\lambda_{i,j}$ are $N(0,1)$ variables, r is set to be 3. Considering that the number of variables (N) sufficient for Monter Carlo simulations are shown to be about 40 (Boivin and Ng (2006) and Inklaar et al. (2005)), we began with 40 variables up to 300. For each N, the number of observation is set at 40 to 300, i.e., N = 40, 50, 100, 150, 200, 300; T = 40, 50, 100, 150, 200, 300. The choice of N and T is quite plausible since it reflects the most frequent size of the empirical datasets used in dynamic factor model.

4.1.1 Simulation: Part 1

DGP1: Homeskedastic idiosyncratic component, and idiosyncratic component have the same variance as the common component:

[12] We are grateful that to Otter, Jacobs, den Reijer, Breitung and Pigorsh for providing their code. The codes of Bai and Ng (2002, 2007), Alessi et al. (2007), Hallin and Liska (2007), Onatski (2009, 2010), Anh and Horenstein (2013) are to be found on their personal homepages.

$e_{i,t} \sim N(0,1)$, $\theta = r$, $\rho = \beta = J = 0$.

DGP2: Heteroskedastic idiosyncratic component:

$e_{i,t} = \begin{matrix} e^1_{i,t} & if\ t\ odd \\ e^1_{i,t} + e^2_{i,t} & if\ t\ even \end{matrix}$, with $e^1_{i,t}$,$e^2_{i,t}$ i.i.d. $\sim N(0,1)$, $\rho = \beta = J = 0$.

DGP3: Only serial correlation is allowed for the idiosyncratic component:

$\beta = 0$, $e_{i,t} = \rho e_{i,t-1} + \upsilon_{it}$.

Instead of assuming $\rho = 0.7$ as in Bai and Ng (2002), we follow Onatski (2009), i.e., ρ are i.i.d $U[-0.8, 0.8]$, which fits the range of the first order autocorrelation of the idiosyncratic error of the data in Stock and Watson (2002).

DGP4: Only cross-section correlation is allowed for the idiosyncratic component: $\rho = 0$, then,

$e_{i,t} = \upsilon_{it} + \sum_{j\neq 0, j=-J}^{J} \beta \upsilon_{i-j,t}$, $\beta = 0.5$, $J = max(10, N/20)$

DGP5: Both serial and cross-section correlation are allowed for the idiosyncratic component:

$e_{i,t} = \rho e_{i,t-1} + \upsilon_{it} + \sum_{j\neq 0, j=-J}^{J} \beta \upsilon_{i-j,t}$, $\rho \sim U[-0.8, 0.8]$, $\beta = 0.2$, $J = max(10, N/20)$

To investigate the effect of the level of cross-section correlation, we run two versions of DGP5 with different magnititudes of correlation, one of $\beta = 0.2$ and the other of $\beta = 0.5$.

4.1.2 Simulation: Part 2

DGP6: common component has smaller variance than the idiosyncratic component:

$e_{i,t} \sim N(0,1)$, $\theta = [2, 4, 6, 8, 10]r$

By allowing θ to be a sequence of numbers, we investigate the effects of the varing explantory power of factors. To isolate the effect of the explanatory power of factors, three versions of DGP6 are conducted. One without serial and cross-section correlation, and the other two introducing serial and cross-section correlations for the idiosyncratic component, however, with different magnititudes of correlation, one with $\beta = 0.2$ and the other with $\beta = 0.5$, i.e.,

DGP6a: $e_{i,t} \sim N(0,1)$, $\theta = r$, $\rho = \beta = J = 0$

DGP6b: $e_{i,t} = \rho e_{i,t-1} + \upsilon_{it} + \sum_{j\neq 0, j=-J}^{J} \beta \upsilon_{i-j,t}$, $\rho \sim U[-0.8, 0.8]$, $\beta = 0.2$, $J = max(10, N/20)$

DGP6c: $e_{i,t} = \rho e_{i,t-1} + \upsilon_{it} + \sum_{j \neq 0, j=-J}^{J} \beta \upsilon_{i-j,t}$, $\rho \sim U[-0.8, 0.8]$, $\beta = 0.5$, $J = max(10, N/20)$.

DGP7: Homescedastic idiosyncratic component, and the idiosyncratic component has smaller variance than the common componentvariance (presence of one weak factor).

$F_{1,t}, F_{2,t} \sim N(0,1)$, $F_{3,t} \sim N(0, \sigma_{F3}^2)$, $\sigma_{F3}^2 = [0.45, 0.4, 0.35, 0.3, 0.25, 0.2, 0.15, 0.1]$

DGP8: One factor has dominately strong explanatory power (presence of a dominantly strong factor)

$\theta = 1$, $F_{1,t} \sim N(0, \sigma_{F1}^2)$, $F_{2,t}, F_{3,t} \sim N(0,1)$, σ_{F1}^2 takes the value of all pair numbers between 2 and 20.

Three versions similar to as DGPs 6a-6c are conducted for DGP7 and DGP8 respectively.

Finally, each series is standardized to have a mean of 0 and unit varaince. For all DGPs, rmax is set to be twice the real number of factors, i.e., 6[13]. The values of the tuning parameters are chosen the following: for the criteria of ABC and HL, the parmeter α in (10) and (14) is set up to 5 with step size of 0.01; for the criterion of BP, we follow their suggestion to set $\delta = 0.5$ and $\tau = 4.5$ (see (26)).

The results for the average selected number of factors over 500 replications for DGP1-5 are summarized in tables 1-4, 5a and 5b. The MSE of the factor number estimators are reported in parentheses below. The results for criteria BN02a, BN02b, ABC, Ona09, Ona10, AH_ER and AH_GR are displayed in the columns.

Monte Carlo simulations show that all methods perform well for DGP1. For DGP 2, most of the criteria give accurate estimates in the presence of heteroskedasticity in the idiosyncratic errors. Ona09 underestimate the number of factor only if N and T are both small (40,50). Futhermore, notice that MSEs are quite small in general, which suggests the estimator is consistent. Similar results are found for DGP 3. Almost all criteria remain robust to the presence of heteroskedasticity and serial correlation.

However, when cross-section correlation is allowed for the idiosyncratic component (DGP 4), the results are less accurate. The criteria BN02a and BN02b tend to overestimate the number of static factors and select the predetermined maximum number of factors. The increased sample size does not improve the results. ABC shows a slight improvement over Bai and Ng (2002), especially at size N=200. However, we should point out that ABC is much more time-consuming than the other method. Next, Onatski (2009, 2010) tend to underestimate the number of factors. Finally, AH_ER and AH_GR dominate the other criteria. ER overestimates the number of factors for small samples and gives an estimator very close to the real number when N and T increase.

[13] We also follow the choice of Bai and Ng (2002) and set rmax to be $8int\left[(c_{N,T}^2/100)^{1/4}\right]$, the results are similar and are not reported

Both serial and cross-section correlation are allowed for the idiosyncratic component in DGP 5. When cross-correlation is strong, $\beta = 0.5$, the results are similar to DGP4. However, if we allow only for low cross-section correlation,$\beta = 0.2$, the performance of Ona09 and Ona10 improve as N and T increase. The number of factors selected is close to the real number for large N (N$\geq$ 100 for Ona09 and N$\geq$ 50 for Ona10). Again, AH_ER and AH_GR outperform the other criteria.

For DGP 6-8, only the results for equaling N and T are reported here for simplicity[14]. The results for DGP 6 (common component has smaller variance than the idiosyncratic component) are displayed in figures 1-4. DGP 6a (absence of serial and cross-section correlation) are shown on the left, DGP 6b (presence of serial and cross-section correlation, $\beta = 0.2$) are shown in the middle, and DGP 6c ($\beta = 0.5$) is shown on the right. We can see that, in the absence of cross-section correlation, pure weak factors have little negative influence on the estimation. Most criteria yield satisfactory results, even for very small samples such as N=T=40. Next, introducing small cross-section correlation worsens the results (DGP6b). BN02a and BN02b always overestimate the number of factors, whereas Ona09 and Ona10 tend to produce underestimation. As the degree of cross-section correlation increases, the estimation results deteriorate further. However, AH_ER and AH_GR continue to suggest a number of factors quite close to the real numbers.

The results for DGP 7a-c (presence of one weak factor) are similarly displayed in figures 5-8. In the absence of serial and cross-section correlation (left column), the number of factors selected by ABC and Ona10 are closest to the real number most of time, while the ER and GR ratios tend to neglect the weak factor as the sample size increases. When serial and cross-section correlations are allowed, the results worsen (middle and right column). BN02a and BN02b always overestimate the number of factors all of the time. ABC outperforms BN02a and BN02b but produces less reliable results with an increase in the sample size. The other criteria neglect the weak factor as the sample increases. We would point out that increasing the sample size does not necessarily improve the results since the degree of cross-section correlation also increases according the experiment design ($e_{i,t} = \rho e_{i,t-1} + \upsilon_{it} + \sum_{j\neq 0, j=-J}^{J} \beta \upsilon_{i-j,t}$, $J = max(10, N/20)$).

The results for DGP 8a-c (presence of a dominant strong factor) are displayed in figure 9-12. In the absence of serial and cross-section correlation (left column), the numbers selected by BN02a, BN02b, ABC and Ona10 are almost always accurate. In contrast, the GR ratio gives precise estimations only if the dominant factor is not very "strong". Finally, ER tends to give only two factors as the sample increases. When small serial and cross-section correlations are allowed (middle column), only Ona10 continues to indicate a fairly accurate number. When large serial and cross-section correlations are allowed (right column), the results worsen. BN02a, BN02b and ABC overestimate the number of

[14]The results of all the combinations of N=40, 50, 100, 150, 200, 300; T=40, 50, 100, 150, 200, 300 for each different value of θ are available on request

factors and the remaining criteria tend to suggest only one factor as the sample size increases.

To summarize, in the case of i.i.d., heteroskedasticity or pure serial correlation in the idiosyncratic component, all the criteria perform well. Most criteria are more sensitive to cross-section correlation. ABC, Ona10, AH_ER and Ah_GR outperform the others in the presence of serial and mild cross-section correlation. In particular, Ah_GR does well even in presence of strong cross-correlation. However, it tends to underestimate the number of factors in the presence of one weak or one dominant factor. In the presence of both cross-correlation (moderate) and a dominant factor, Ona10 seems to be more reliable.

4.2 Dynamic factors

In relation to the criteria for selecting the number of dynamic factors, we generate samples with the same DGPs as Bai and Ng (2007), with slight modification. The base model is

$$\mathbf{x}_t = A_0 f_t + A_1 f_{t-1} + A_2 f_{t-2} + \mathbf{e}_t$$

A_0,A_1 and A_2 are drawn from the standard normally distributed random variables. The number of dynamic factors is assumed to be two in all DGPs. The number of static factor models is hence $r = q(s+1) = 6$. Following DGPs are considered:

DGP9: f_t is a Moving Average MA(1) process: $f_t = v_t + \Theta v_{t-1}$, with $\Theta = diag(0.2, 0.9)$ and $v_t \sim N(0,1)$,

DGP10: f_t is an Autoregressive AR(1) process: $f_t = \Gamma_1 f_{t-1} + v_t$, with $\Gamma = diag(0.2, 0.9)$ and $v_t \sim N(0,1)$.

In contrast of Bai and Ng (2007), where $\mathbf{e}_t$ are i.i.d. standard normal, both serial and cross-section correlation are allowed for idiosyncratic component in DGP9 and DGP10, which follows the process suggested by Onatski (2009):

$e_{i,t} = \rho e_{i,t-1} + v_{it}$, with $\rho \sim U[-0.8, 0.8]$, $\beta = 0.2$[15].

We also consider the presence of one dominantly strong factor in both cases, i.e. $f_{1,t} \sim N(0, \sigma^2_{f1})$, $f_{2,t} \sim N(0,1)$, σ^2_{f1} takes the value of all pair numbers between 2 and 20. Let MA with one dominant factor be DGP11 and AR with one dominant factor be DGP12.

For the procedures where $qmax$ is required, $qmax$ is set to be r, the number of static factors. The first step is thus estimation of the number of static factors. In light of the results obtained in the previous section, we rely on the results in Ona10, AH_ER and AH_GR primarily. The results for number of static factors

[15]The results with $\beta = 0.5$ and s=1 are similar and will be not repoted.

selected by the previous criteria for DGP9 (MA) are given in table 7[16]. BN02a and BN02b overestimate the number of static factors. Notice the number of static factors estimated approches the real number as the N or/and T increase since the cross-section correlation does not increase with sample size. And Ona09 underestimates the number of static factors. However, ABC, Ona10, AH_ER and AH_GR criteria give estimation close to the real number of static factors. Therefore, we set r to be the real number of static factor, i.e., 6.

The results for the average selected number of factors over 500 replications for DGP 9 are given in table 8. BN07a and BN07b invariably point to one dynamic factor, while Ona09, JOR, SWa and SWb always overestimate the number of factors. HL and BP accurately estimate the number of factors. However, as in ABC, HL is much more time consuming. Another problem related to ABC and HL is the authors suggest choosing the value of the second flat interval as the number of factors. However, they are not precise about the length of the flat interval. Therefore, very short flat intervals can lead to unstable results.

The results for DGP 10 (AR) are given in table 8. For the static factors, BN02a and BN02b overestimate the number for small T (T=50). As sample size increases, BN02a and BN02b underestimate the number of factors. The underestimation of Ona09 is more severe. Next, AH_ER and AH_GR criteria give less accurate estimations than in DGP9. They underestimate the number of factors for small samples and the problem of AH_ER is more severe. However, ABC and Ona10 continue to give an accurate estimation. For the number of dynamic factors, the results are similar to DGP9, BN07a and BN07b continue to underestimate the number of factors and SWa and SWb continue to overestimate the number of factors. In fact, SWa and SWb always take the value of $qmax$. Next, Ona09 slightly underestimates the number of factors. It is not surprising that the number of static factors and dynamic factors suggested by Ona09 are close since the approach is similar. However, our experiments shows that the criteria developped by Onatski (2009) is more suitable for selecting the number of dynamic factors. Concerning JOR, it overestimates the number of factors. Nontheless, as sample size increases, the estimation approaches the real number. Finally, HL chooses two (most of time) or three factors and BP accurately selects number of factors.

For DGP 11 (MA with one dominant factor, figures 13-16), the results are similar to DGP9 (MA). Despite the presence of one dominant strong factor, Ona10, AH_ER and AH_GR still correctly select the number of static factors. The estimation of ABC is less accurate than in DGP9 and still close to the real number. BP continues to choose the correct number of dynamic factors. In contrast, if we allow one dominant factor in AR (DGP12, figures 17-20), no criteria can accurately estimate point out the real number of static factors. To estimate the number of dynamic factors, if we set qmax to be the real number of factors, Ona09 and JOR give an accurate estimation and outperforms the others, while BP overestimates the number of dynamic factors. If we set qmax

[16] Similar results are found for DGP 10 and DGP11. The results are not reported here.

to be 2, as suggested by Ona10, then AH_ER and AH_GR, BP chooses 1 for small samples and 2 as the sample gets large, which is not surprising[17].

To conclude, when the dynamic factors follow an MA process, ABC, Ona10, AH_ER and AH_GR will accurately estimate the number of static factors. However, the performance of AH_ER and AH_GR deteriorate if the dynamic factors follow an AR process. The situation becomes worse if there is one dominant strong factor present, and no criteria can give the correct specification. In the specification for the number of dynamic factors, BP gives an accurate estimation in both the AR and MA cases. However, if there is one dominantly strong factor in AR, it fails. Another problem is that BP is quite sensitive to $qmax$.

5 Empirical application

Finally, we evaluate the performance of the criteria in one empirical study. Our data consist of weekly stock returns of the components of the S&P 500 between 01/2000-12/2013 (705 observations). Among the stock components in the S&P 500, we chose only those stocks where the returns are availabe for the entier observation periods. Thus, 397 stocks are included. The data source are Bloomberg[18]. All the series are transformed to be stationary and standardized. BN02a, BN02b provide estimations of six static factors. ABC and Ona10 estimate five static factors, and the others suggest one static factor. Considering our simulation results, there is a strong evidence of the presence of one strong factor.

The estimation results for the number of dynamic factors are given in table 9[19] The economic interpretations are strongly recommended. Six dynamic factors are suggested by BP and five dynamic factors are suggested by Ona 09. The others estimate one dynamic factor. There is strong support for one market factor.

6 Conclusion

This paper compared the small sample performance of selection rules for factors numbers in large datasets. We find that the GR ratio proposed by Ahn and Horenstein (2013) is robust to the presence of serial and (strong) cross-section correlation. However, it tends to underestimate the number of factors in the presence of one weak or one dominant factor. In the case of both presence of cross-correlation (moderate) and a dominant factor, the criterion proposed by

[17] In fact, we also set $qmax$ to be 3, 4 and 5, the results were similar, except BP is more sensitive to the number of static factors chosen. Similar results are found for DGP9-11.

[18] We are grateful to Jean-Etienne Carlotti for providing the data.

[19] We also estimate the number of factors of the dataset of Stock and Watson (2005). The results are quite similar.

Onatski (2010) seems more reliable. Furthermore, the two criteria can select accurately the number of static factors in a dynamic factors design. Also, the criteria proposed by Breitung and Pigorsch (2009) correctly select the number of dynamic factors in most cases. We hope these results help in the choice of the number determining criteria with different types of data, for future research.

References

[1] Ahn, S.C. and A.R. Horenstein (2013), Eigenvalue Ratio Test for the Number of Factors. Econometrica, 81(3), 1203-27

[2] Alessi L., Barigozzi M. and Capasso M. (2008), A robust criterion for determining the number of static factors in approximate factor models. Working paper No 903, European Central Bank.

[3] Amengual D. and Watson M.W. (2007), Consistent estimation of the number of dynamic factors in a large N and T panel. Journal of Business and Economic Statistics, 25, 91-96.

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

[5] Bai, J., and S. Ng (2006), Determining the Number of Factors in Approximate Factor Models, errata. manuscript, Columbia University.

[6] Bai J. and Ng S. (2007), Determining the number of primitive shocks in factor models. Journal of Business and Economic Statistics, 25, 52-60.

[7] Barhoumi K., Darné O. and Ferrara L. (2013), Testing the number of factors: An empirical assessment for forecasting purposes, Oxford Bulletin of Economics and Statistics, 75, 1, 64-79

[8] Bernanke, B. S., Boivin, J., Eliasz, P. (2005). Measuring the effects of monetary policy: A factoraugmented vector autoregressive (FAVAR) approach. Quarterly Journal of Economics 120 387-422.

[9] Boivin, J. and S. Ng (2005). Understanding and comparing factor-based forecasts. International Journal of Central Banking, 1, 3, 117-151.

[10] Breitung J. and Eickmeier S. (2006). Dynamic factor models. In Hübler O. and Frohn J. (eds.), Modern Econometric Analysis, Chap. 3, Springer.

[11] Breitung J. and Pigorsch U. (2012), A canonical correlation approach for selecting the number of dynamic factors, forthcoming in: Oxford Bulletin of Economics and Statistics.

[12] Brillinger D. R.(1981), Time series: Data analysis and theory. Holden-Day, Inc., San Francisco.

[13] Campbell, J.Y., Lo, A. and MacKinlay, A. C. (1997). The Econometrics of Financial Markets. Princeton University Press, Princeton, N.J.

[14] Cattell, R. B. (1966) "The Scree Test for the Number of Factors", Multivariate Behavioral Research, vol. 1, 245-76.

[15] Cipollini, A. and Missaglia, G. (2007), "Dynamic Factor analysis of industry sector default rates and implication for Portfolio Credit Risk Modelling," MPRA Paper 3582, University Library of Munich, Germany

[16] Connor, G. and Korajcyk, R. (1993), A Test for the Number of Factors in an Approximate Factor Model, Journal of Finance XLVIII:4 , 1263–1291.

[17] Cragg, J. and Donald, S. (1997), Inferring the Rank of a Matrix, Journal of Econometrics 76 , 223– 250.

[18] Antonello D' Agostino & Domenico Giannone (2013). "Comparing Alternative Predictors Based on Large-Panel Factor Models," Oxford Bulletin of Economics and Statistics, Department of Economics, University of Oxford, vol. 74(2), pages 306-326, 04.

[19] Donald, S. (1997), Inference Concerning the Number of Factors in a Multivariate Nonparameteric Relationship, Econometrica 65:1 , 103–132.

[20] Eickmeier, S., Ziegler C. (2008), How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach, Journal of Forecasting, 27(3), 237-265.

[21] Forni, M. and Reichlin, L. (1998), Let's Get Real: a Factor-Analytic Approach to Disaggregated Business Cycle Dynamics, Review of Economic Studies 65 , 453–473.

[22] Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000), The Generalized Dynamic Factor Model: Identification and Estimation, The Review of Economics and Statistics Vol. 82, No. 4 (Nov., 2000), pp. 540-554.

[23] Forni, M. & Giannone, D. & Lippi, M. & Reichlin, L, (2009). "Opening The Black Box: Structural Factor Models With Large Cross Sections," Econometric Theory, Cambridge University Press, vol. 25(05), pages 1319-1347, October.

[24] Guttman, L. (1954), Some necessary conditions for common-factor analysis. Psychometrika, 19, 149-161.

[25] Hallin M. and Liska R. (2007). Determining the number of factors in the general dynamic factor model. Journal of the American Statistical Association, 102, 603-617.

[26] Koopman S. J. and Van der Wel M. (2013), Forecasting the U.S. Term Structure of Interest Rates Using a Macroeconomic Smooth Dynamic Factor Model, International Journal of Forecasting, Forthcoming.

[27] Lewbel, A. (1991), The Rank of Demand Systems: Theory and Nonparameteric Estimation, Econometrica 59 , 711–730.

[28] Johnstone, I.M. (2001), "On the Distribution of the Largest Eigenvalue in Principal Component Analysis," Annals of Statistics, 29, 295-327.

[29] Kapetanios G. (2004). A new method for determining the number of factors in factor models with large datasets. Working paper No 525, Department of Economics, Queen Mary University of London.

[30] Kapetanios G. (2010) . A Testing Procedure for Determining the Number of Factors in Approximate Factor Models With Large Datasets. Journal of Business and Economic Statistics, 28(3), pp. 397–409.

[31] Kapetanios G. and Marcellino M. (2009). A Parametric Estimation Method for Dynamic Factor Models of Large Dimension. Journal of Time Series Analysis, 30(2), pp. 208-238.

[32] Jum C. Nunnally, Ira H. Bernstein (1994), Psychometric theory (3rd edition), New York: McGraw-Hill.

[33] Onatski A., (2009). "Testing Hypotheses About the Number of Factors in Large Factor Models," Econometrica, Econometric Society, vol. 77(5), pages 1447-1479, 09.

[34] Onatski A., (2010). "Determining the Number of Factors from Empirical Distribution of Eigenvalues," The Review of Economics and Statistics, MIT Press, vol. 92(4), pages 1004-1016, November.

[35] Schumacher C. (2007). Forecasting German GDP using alternative factor models based on large datasets, mimeo.

[36] Schwert G.W (1989) "Tests for Unit Roots: A Monte Carlo Investigation," Journal of Business and Economic Statistics, 7 (April), 147-159.

[37] Tracy, C. A.; Widom, H. (1994), "Level-spacing distributions and the Airy kernel", Communications in Mathematical Physics 159 (1): 151–174.

[38] Bai, Z. D. and Silverstein, J. W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices. Annals of Probability, 26 , 316-345.

[39] Wilson P. and Cooper C. (2008), Finding the magic number, The Psychologist, Volume 21, Part 10 and October 2008.

Appendix

Table 3: DGP1: Estimtion of number of factors

N	T	BN02a	BN02b	ABC	Ona09	Ona10	AH_ER	AH_GR
40	40	3.07 (0.07)	3 (0)	2.97 (0.03)	2.07 (3.21)	2.99 (0.01)	2.82 (0.25)	2.87 (0.16)
40	50	3.01 (0.01)	3 (0)	2.99 (0.01)	2.22 (2.52)	3.01 (0.01)	2.9 (0.14)	2.94 (0.07)
40	100	3 (0)	3 (0)	3 (0)	2.78 (0.57)	3.01 (0.02)	2.99 (0)	2.99 (0)
40	150	3 (0)	3 (0)	3 (0)	2.88 (0.24)	3 (0)	2.99 (0)	3 (0)
40	200	3 (0)	3 (0)	3 (0)	2.92 (0.16)	3 (0)	2.99 (0)	3 (0)
40	300	3 (0)	3 (0)	3 (0)	2.87 (0.25)	3 (0.01)	3 (0)	3 (0)
50	40	3.01 (0.01)	3 (0)	3 (0)	2.33 (2.17)	3.01 (0.01)	2.9 (0.15)	2.93 (0.07)
50	50	3 (0)	3 (0)	3 (0)	2.6 (1.31)	3.01 (0.03)	2.96 (0.05)	2.98 (0.01)
50	100	3 (0)	3 (0)	3 (0)	2.89 (0.22)	3.02 (0.04)	2.99 (0)	3 (0)
50	150	3 (0)	3 (0)	3 (0)	2.89 (0.21)	3.01 (0.04)	3 (0)	3 (0)
50	200	3 (0)	3 (0)	3 (0)	2.89 (0.2)	3 (0)	3 (0)	3 (0)
50	300	3 (0)	3 (0)	3 (0)	2.91 (0.17)	3 (0)	3 (0)	3 (0)
100	40	3 (0)	3 (0)	3 (0)	2.83 (0.38)	3.01 (0.03)	2.99 (0)	2.99 (0)
100	50	3 (0)	3 (0)	3 (0)	2.9 (0.19)	3 (0.01)	2.99 (0)	2.99 (0)
100	100	3 (0)	3 (0)	3 (0)	2.9 (0.18)	3 (0.01)	3 (0)	3 (0)
100	150	3 (0)	3 (0)	3 (0)	2.94 (0.12)	3 (0)	3 (0)	3 (0)
100	200	3 (0)	3 (0)	3 (0)	2.9 (0.18)	3 (0)	3 (0)	3 (0)
100	300	3 (0)	3 (0)	3 (0)	2.88 (0.24)	3 (0)	3 (0)	3 (0)
150	40	3 (0)	3 (0)	3 (0)	2.91 (0.18)	3 (0)	2.99 (0)	3 (0)
150	50	3 (0)	3 (0)	3 (0)	2.92 (0.16)	3.01 (0.01)	3 (0)	3 (0)
150	100	3 (0)	3 (0)	3 (0)	2.89 (0.2)	3.01 (0.02)	3 (0)	3 (0)
150	150	3 (0)	3 (0)	3 (0)	2.93 (0.12)	3 (0.02)	3 (0)	3 (0)
150	200	3 (0)	3 (0)	3 (0)	2.92 (0.15)	3 (0)	3 (0)	3 (0)
150	300	3 (0)	3 (0)	3.03 (0)	2.91 (0.17)	3 (0)	3 (0)	3 (0)
200	40	3 (0)	3 (0)	3 (0)	2.9 (0.18)	3 (0)	3 (0)	3 (0)
200	50	3 (0)	3 (0)	3 (0)	2.92 (0.15)	3.01 (0.02)	3 (0)	3 (0)
200	100	3 (0)	3 (0)	3 (0)	2.93 (0.13)	3 (0)	3 (0)	3 (0)
200	150	3 (0)	3 (0)	3 (0)	2.94 (0.12)	3 (0.01)	3 (0)	3 (0)
200	200	3 (0)	3 (0)	3 (0.27)	2.94 (0.11)	3 (0)	3 (0)	3 (0)
200	300	3 (0)	3 (0)	3 (0)	2.89 (0.21)	3 (0)	3 (0)	3 (0)
300	40	3 (0)	3 (0)	3 (0)	2.86 (0.26)	3 (0)	3 (0)	3 (0)
300	50	3 (0)	3 (0)	3 (0)	2.88 (0.24)	3 (0)	3 (0)	3 (0)
300	100	3 (0)	3 (0)	3 (0)	2.9 (0.19)	3 (0)	3 (0)	3 (0)
300	150	3 (0)	3 (0)	3 (0)	2.92 (0.16)	3 (0)	3 (0)	3 (0)
300	200	3 (0)	3 (0)	3 (0)	2.94 (0.12)	3 (0)	3 (0)	3 (0)
300	300	3 (0)	3 (0)	3 (0)	2.93 (0.12)	3 (0)	3 (0)	3 (0)

Table 4: DGP2: Estimtion of number of factors

N	T	BN02a	BN02b	ABC	Ona09	Ona10	AH_ER	AH_GR
40	40	3.09	2.99	2.97	2.05	3.02	2.82	2.89
		(0.09)	(0)	(0.03)	(3.28)	(0.04)	(0.27)	(0.15)
40	50	3	2.93	2.63	1.47	2.91	2.56	2.69
		(0)	(0.06)	(0.97)	(5.17)	(0.16)	(0.69)	(0.48)
40	100	3	3	3	2.88	3	2.99	2.99
		(0)	(0)	(0)	(0.32)	(0.02)	(0)	(0)
40	150	2.99	2.98	2.99	2.7	3.01	2.96	2.97
		(0)	(0.01)	(0.01)	(0.88)	(0.02)	(0.04)	(0.03)
40	200	3	3	3	2.9	3.01	2.99	3
		(0)	(0)	(0)	(0.19)	(0.02)	(0)	(0)
40	300	3	3	3	2.86	3	2.98	2.99
		(0)	(0)	(0)	(0.27)	(0.01)	(0.01)	(0)
50	40	3	3	3	2.31	3.01	2.9	2.94
		(0)	(0)	(0)	(2.29)	(0.01)	(0.14)	(0.07)
50	50	2.99	2.9	2.78	1.68	2.98	2.64	2.74
		(0)	(0.09)	(0.76)	(4.66)	(0.04)	(0.57)	(0.38)
50	100	3	3	3	2.89	3	3	3
		(0)	(0)	(0)	(0.23)	(0)	(0)	(0)
50	150	3	2.99	3	2.83	3.01	2.98	2.99
		(0)	(0)	(0)	(0.35)	(0.01)	(0.01)	(0.01)
50	200	3	3	3	2.88	3	3	3
		(0)	(0)	(0)	(0.23)	(0)	(0)	(0)
50	300	3	3	3	2.9	3.01	3	3
		(0)	(0)	(0)	(0.19)	(0.03)	(0)	(0)
100	40	3	3	3	2.85	3	2.99	2.99
		(0)	(0)	(0)	(0.41)	(0)	(0)	(0)
100	50	2.99	2.99	2.99	2.55	3.01	2.94	2.96
		(0)	(0)	(0.01)	(1.44)	(0.02)	(0.08)	(0.05)
100	100	3	3	3	2.92	3.01	3	3
		(0)	(0)	(0)	(0.14)	(0.02)	(0)	(0)
100	150	3	3	3	2.92	3.01	3	3
		(0)	(0)	(0)	(0.16)	(0.02)	(0)	(0)
100	200	3	3	3	2.88	3.01	3	3
		(0)	(0)	(0)	(0.22)	(0.01)	(0)	(0)
100	300	3	3	3	2.93	3.01	3	3
		(0)	(0)	(0)	(0.13)	(0.02)	(0)	(0)
150	40	3	3	3	2.93	3.01	3	3
		(0)	(0)	(0)	(0.16)	(0.01)	(0)	(0)
150	50	3	3	3	2.84	3.01	2.99	2.99
		(0)	(0)	(0)	(0.37)	(0.04)	(0)	(0)
150	100	3	3	3	2.88	3	3	3
		(0)	(0)	(0)	(0.23)	(0)	(0)	(0)
150	150	3	3	3	2.91	3.01	3	3
		(0)	(0)	(0)	(0.16)	(0.02)	(0)	(0)
150	200	3	3	3	2.9	3.01	3	3
		(0)	(0)	(0)	(0.18)	(0.02)	(0)	(0)
150	300	3	3	3	2.94	3	3	3
		(0)	(0)	(0)	(0.11)	(0)	(0)	(0)
200	40	3	3	3	2.93	3	2.99	3
		(0)	(0)	(0)	(0.13)	(0)	(0)	(0)
200	50	3	3	3	2.88	3.01	3	3
		(0)	(0)	(0)	(0.23)	(0.02)	(0)	(0)
200	100	3	3	3	2.92	3.01	3	3
		(0)	(0)	(0)	(0.15)	(0.03)	(0)	(0)
200	150	3	3	3	2.92	3.01	3	3
		(0)	(0)	(0)	(0.14)	(0.02)	(0)	(0)
200	200	3	3	3	2.92	3	3	3
		(0)	(0)	(0)	(0.14)	(0)	(0)	(0)
200	300	3	3	3	2.88	3	3	3
		(0)	(0)	(0)	(0.23)	(0)	(0)	(0)
300	40	3	3	3	2.9	3	3	3
		(0)	(0)	(0)	(0.2)	(0)	(0)	(0)
300	50	3	3	3	2.91	3	2.99	3
		(0)	(0)	(0)	(0.17)	(0)	(0)	(0)
300	100	3	3	3	2.86	3.01	3	3
		(0)	(0)	(0)	(0.26)	(0.03)	(0)	(0)
300	150	3	3	3	2.92	3.02	3	3
		(0)	(0)	(0)	(0.14)	(0.05)	(0)	(0)
300	200	3	3	3	2.92	3.01	3	3
		(0)	(0)	(0)	(0.15)	(0.02)	(0)	(0)
300	300	3	3	3	2.94	3	3	3
		(0)	(0)	(0)	(0.11)	(0.02)	(0)	(0)

Table 5: DGP3: Estimtion of number of factors

N	T	BN02a	BN02b	ABC	Ona09	Ona10	AH_ER	AH_GR
40	40	3.87 (1.14)	3.06 (0.06)	3.03 (0.03)	2.78 (0.61)	3.01 (0.02)	2.99 (0)	2.99 (0)
40	50	3.4 (0.42)	3.01 (0.01)	3.02 (0.02)	2.81 (0.47)	3.03 (0.05)	3 (0)	3 (0)
40	100	3 (0)	3 (0)	3 (0)	2.88 (0.24)	3 (0.01)	3 (0)	3 (0)
40	150	3 (0)	3 (0)	3 (0)	2.92 (0.14)	3 (0.01)	3 (0)	3 (0)
40	200	3 (0)	3 (0)	3 (0)	2.91 (0.16)	3.01 (0.03)	3 (0)	3 (0)
40	300	3 (0)	3 (0)	3 (0)	2.94 (0.11)	3 (0)	3 (0)	3 (0)
50	40	3.58 (0.64)	3.04 (0.04)	3.03 (0.03)	2.82 (0.37)	3.02 (0.04)	2.99 (0.01)	2.99 (0)
50	50	3.17 (0.17)	3 (0)	3.01 (0.01)	2.89 (0.22)	3.02 (0.03)	2.99 (0)	3 (0)
50	100	3 (0)	3 (0)	3 (0)	2.87 (0.24)	3 (0)	3 (0)	3 (0)
50	150	3 (0)	3 (0)	3 (0)	2.9 (0.19)	3 (0)	3 (0)	3 (0)
50	200	3 (0)	3 (0)	3 (0)	2.9 (0.18)	3.01 (0.02)	3 (0)	3 (0)
50	300	3 (0)	3 (0)	3.03 (0.09)	2.92 (0.15)	3.01 (0.02)	3 (0)	3 (0)
100	40	3.04 (0.04)	3 (0)	3 (0)	2.94 (0.12)	3.01 (0.02)	3 (0)	3 (0)
100	50	3 (0)	3 (0)	3 (0)	2.91 (0.16)	3.02 (0.03)	3 (0)	3 (0)
100	100	3 (0)	3 (0)	3 (0)	2.89 (0.21)	3.01 (0.02)	3 (0)	3 (0)
100	150	3 (0)	3 (0)	3 (0)	2.88 (0.22)	3.01 (0.02)	3 (0)	3 (0)
100	200	3 (0)	3 (0)	3.1 (0.1)	2.91 (0.16)	3.01 (0.01)	3 (0)	3 (0)
100	300	3 (0)	3 (0)	3 (0)	2.88 (0.24)	3 (0.01)	3 (0)	3 (0)
150	40	3 (0)	3 (0)	3 (0)	2.91 (0.16)	3.03 (0.06)	3 (0)	3 (0)
150	50	3 (0)	3 (0)	3 (0)	2.9 (0.19)	3.01 (0.02)	3 (0)	3 (0)
150	100	3 (0)	3 (0)	3.1 (0.1)	2.93 (0.12)	3.01 (0.02)	3 (0)	3 (0)
150	150	3 (0)	3 (0)	3 (0)	2.92 (0.15)	3.01 (0.02)	3 (0)	3 (0)
150	200	3 (0)	3 (0)	3.1 (0.1)	2.93 (0.12)	3.01 (0.04)	3 (0)	3 (0)
150	300	3 (0)	3 (0)	3 (0)	2.91 (0.17)	3.02 (0.04)	3 (0)	3 (0)
200	40	3 (0)	3 (0)	3 (0)	2.89 (0.2)	3 (0.02)	3 (0)	3 (0)
200	50	3 (0)	3 (0)	3 (0)	2.89 (0.2)	3.01 (0.03)	3 (0)	3 (0)
200	100	3 (0)	3 (0)	3 (0)	2.9 (0.18)	3.02 (0.04)	3 (0)	3 (0)
200	150	3 (0)	3 (0)	3 (0)	2.89 (0.21)	3 (0)	3 (0)	3 (0)
200	200	3 (0)	3 (0)	3 (0)	2.92 (0.15)	3 (0.01)	3 (0)	3 (0)
200	300	3 (0)	3 (0)	3 (0)	2.94 (0.12)	3.01 (0.02)	3 (0)	3 (0)
300	40	3 (0)	3 (0)	3 (0)	2.89 (0.21)	3 (0)	3 (0)	3 (0)
300	50	3 (0)	3 (0)	3 (0)	2.88 (0.23)	3 (0.02)	3 (0)	3 (0)
300	100	3 (0)	3 (0)	3 (0)	2.91 (0.17)	3 (0)	3 (0)	3 (0)
300	150	3 (0)	3 (0)	3 (0)	2.93 (0.13)	3 (0)	3 (0)	3 (0)
300	200	3 (0)	3 (0)	3 (0)	2.9 (0.18)	3 (0)	3 (0)	3 (0)
300	300	3 (0)	3 (0)	3 (0)	2.9 (0.18)	3.01 (0.02)	3 (0)	3 (0)

Table 6: DGP4: Estimtion of number of factors

N	T	BN02a	BN02b	ABC	Ona09	Ona10	AH_ER	AH_GR
40	40	6	6	4.17	1.25	3.59	4.73	2.59
		(9)	(9)	(5.95)	(6.53)	(7.62)	(6.33)	(2.38)
40	50	6	6	4.69	1.04	3.82	5.05	2.59
		(9)	(9)	(6.73)	(6.96)	(7.9)	(6.8)	(2.18)
40	100	6	6	4.7	0.86	3.8	5.32	2.45
		(9)	(9)	(6.88)	(6.64)	(8.18)	(7.38)	(2.18)
40	150	6	6	5.03	0.74	4.13	5.55	2.45
		(9)	(9)	(7.41)	(6.69)	(8.55)	(7.79)	(2.11)
40	200	6	6	5.18	0.81	4.12	5.65	2.3
		(9)	(9)	(7.72)	(6.55)	(8.61)	(8.01)	(2.27)
40	300	6	6	5.41	0.78	4.32	5.78	2.36
		(9)	(9)	(7.81)	(6.36)	(8.77)	(8.28)	(2.17)
50	40	6	6	4.76	1	1.06	4.25	2.79
		(9)	(9)	(7.12)	(6.78)	(7.65)	(4.96)	(2.22)
50	50	6	6	5	0.86	0.83	4.09	2.71
		(9)	(9)	(7.54)	(6.88)	(7.85)	(4.55)	(2.11)
50	100	6	6	5.39	0.61	0.51	4.43	2.6
		(9)	(9)	(8.11)	(7.35)	(8.2)	(5.3)	(2.06)
50	150	6	6	5.89	0.67	0.44	4.59	2.56
		(9)	(9)	(8.85)	(7.08)	(8.38)	(5.61)	(2.01)
50	200	6	6	5.77	0.7	0.39	4.7	2.41
		(9)	(9)	(8.63)	(6.97)	(8.29)	(5.96)	(2.26)
50	300	6	6	5.73	0.73	0.24	4.99	2.42
		(9)	(9)	(8.63)	(6.72)	(8.65)	(6.25)	(1.97)
100	40	6	6	0.45	0.9	0.16	3.46	2.59
		(9)	(9)	(8.45)	(7.3)	(8.29)	(3.69)	(2.19)
100	50	6	6	0	0.81	0.08	3.38	2.6
		(9)	(9)	(9)	(7.43)	(8.58)	(3.4)	(2.15)
100	100	6	6	5.95	0.67	0.04	3.58	2.62
		(9)	(9)	(8.95)	(7.61)	(8.82)	(3.85)	(2.16)
100	150	6	6	5.95	0.66	0.01	3.43	2.28
		(9)	(9)	(8.95)	(7.71)	(8.91)	(4)	(2.26)
100	200	6	6	6	0.72	0	3.55	2.36
		(9)	(9)	(9)	(7.85)	(8.98)	(3.99)	(2.3)
100	300	6	6	6	0.73	0	3.52	2.21
		(9)	(9)	(9)	(7.67)	(9)	(4)	(2.26)
150	40	6	6	0.85	0.9	0.29	2.78	2.48
		(9)	(9)	(6.77)	(7.28)	(7.87)	(3.08)	(2.2)
150	50	6	6	0.99	0.88	0.2	2.73	2.48
		(9)	(9)	(7.87)	(7.42)	(8.21)	(2.81)	(2.16)
150	100	6	6	0.35	0.63	0.18	2.5	2.32
		(9)	(9)	(7.35)	(7.62)	(8.29)	(2.45)	(1.95)
150	150	6	6	5.66	0.71	0.16	2.49	2.29
		(9)	(9)	(8.62)	(7.51)	(8.31)	(2.24)	(1.72)
150	200	6	6	5.56	0.75	0.14	2.54	2.41
		(9)	(9)	(8.52)	(7.36)	(8.44)	(1.94)	(1.68)
150	300	6	6	5.66	0.68	0.11	2.42	2.39
		(9)	(9)	(8.62)	(7.29)	(8.56)	(1.35)	(1.18)
200	40	6	6	3.58	0.96	0.45	2.39	2.4
		(9)	(9)	(6.3)	(7.13)	(7.34)	(2.41)	(1.93)
200	50	6	6	3.51	0.77	0.44	2.36	2.32
		(9)	(9)	(6.11)	(7.49)	(7.35)	(2.17)	(1.86)
200	100	6	6	3.83	0.81	0.75	2.41	2.45
		(9)	(9)	(5.91)	(7.35)	(6.52)	(1.51)	(1.24)
200	150	6	6	3.92	0.86	0.93	2.51	2.58
		(9)	(9)	(5.14)	(7.48)	(6.08)	(0.92)	(0.8)
200	200	6	6	3.62	0.93	1.13	2.57	2.67
		(9)	(9)	(3.9)	(7.23)	(5.49)	(0.75)	(0.57)
200	300	6	6	3.16	1.22	2.41	2.87	2.89
		(9)	(9)	(1.84)	(5.92)	(1.69)	(0.19)	(0.15)
300	40	6	6	4.28	0.82	0.14	2.9	2.56
		(9)	(9)	(7.16)	(7.23)	(8.48)	(2.94)	(2.19)
300	50	6	6	4.07	0.79	0.04	2.85	2.39
		(9)	(9)	(6.99)	(7.6)	(8.78)	(3.07)	(2.14)
300	100	6	6	5.3	0.7	0.03	2.95	2.46
		(9)	(9)	(8.3)	(7.53)	(8.85)	(3.07)	(2.27)
300	150	6	6	5.6	0.66	0.02	2.82	2.36
		(9)	(9)	(8.6)	(7.78)	(8.87)	(2.98)	(2.18)
300	200	6	6	5.95	0.79	0.02	2.89	2.45
		(9)	(9)	(8.95)	(7.44)	(8.91)	(2.94)	(2.23)
300	300	6	6	6	0.52	0	3.22	2.47
		(9)	(9)	(9)	(7.99)	(9)	(3.25)	(2.18)

Table 7: DGP5 a: Estimtion of number of factors

N	T	BN02a	BN02b	ABC	Ona09	Ona10	AH_ER	AH_GR
40	**40**	6	5.98	3.53	2.02	4.78	4.41	3.76
		(9)	(8.91)	(3.95)	(6.6)	(7.09)	(4.59)	(2.38)
40	**50**	6	5.99	3.51	2.18	5.12	4.74	3.86
		(9)	(8.99)	(3.79)	(7.13)	(7.62)	(5.28)	(2.53)
40	**100**	6	6	4.78	2.09	5.55	5.2	3.94
		(9)	(9)	(4.3)	(7.63)	(8.36)	(6.43)	(2.36)
40	**150**	6	6	4.86	1.82	5.48	5.34	4.07
		(9)	(9)	(3.92)	(7.74)	(8.6)	(6.93)	(2.52)
40	**200**	6	6	4.95	1.59	5.78	5.51	4.03
		(9)	(9)	(4.35)	(7.83)	(8.75)	(7.35)	(2.48)
40	**300**	6	6	4.96	1.46	5.74	5.56	4.17
		(9)	(9)	(4.16)	(7.66)	(8.77)	(7.44)	(2.67)
50	**40**	6	5.99	4.28	1.38	3.13	3.74	3.29
		(9)	(8.99)	(3.52)	(6.91)	(5.98)	(3.34)	(1.61)
50	**50**	6	6	4.33	1.21	2.87	3.89	3.32
		(9)	(9)	(4.31)	(6.86)	(6.27)	(3.55)	(1.66)
50	**100**	6	6	4.66	0.83	2.51	4.01	3.33
		(9)	(9)	(4.32)	(7.12)	(6.74)	(3.44)	(1.33)
50	**150**	6	6	4.61	0.69	2.23	3.98	3.33
		(9)	(9)	(4.15)	(7.45)	(6.81)	(3.32)	(1.08)
50	**200**	6	6	4.79	0.5	2.14	4.01	3.3
		(9)	(9)	(4.41)	(7.46)	(6.88)	(3.38)	(1.04)
50	**300**	6	6	4.57	0.43	1.87	3.94	3.16
		(9)	(9)	(3.85)	(7.69)	(6.12)	(3.21)	(0.81)
100	**40**	6	6	3.34	1.45	2.78	2.82	2.92
		(9)	(9)	(0.78)	(5.68)	(0.89)	(0.36)	(0.26)
100	**50**	6	6	3.3	1.57	2.9	2.91	2.98
		(9)	(9)	(0.7)	(5.13)	(0.44)	(0.13)	(0.13)
100	**100**	6	6	3.1	2.07	2.99	2.99	2.99
		(9)	(9)	(0.1)	(3.17)	(0.03)	(0)	(0)
100	**150**	6	6	3	2.32	3	2.99	2.99
		(9)	(9)	(0)	(2.3)	(0)	(0)	(0)
100	**200**	6	6	3.39	2.61	3	2.99	3
		(9)	(9)	(0.91)	(1.36)	(0)	(0)	(0)
100	**300**	6	6	4.5	2.79	3	3	3
		(9)	(9)	(4.22)	(0.64)	(0)	(0)	(0)
150	**40**	6	6	3.15	1.86	3.01	2.94	2.97
		(9)	(9)	(0.23)	(3.89)	(0.11)	(0.07)	(0.02)
150	**50**	6	6	3.29	2.22	3.05	2.97	2.99
		(9)	(9)	(0.43)	(2.8)	(0.11)	(0.02)	(0.01)
150	**100**	6	6	3	2.78	3	3	3
		(9)	(9)	(0.19)	(0.6)	(0.01)	(0)	(0)
150	**150**	6	6	3	2.9	3	3	3
		(9)	(9)	(4.37)	(0.23)	(0)	(0)	(0)
150	**200**	6	6	3.3	2.9	3	3	3
		(9)	(9)	(7.81)	(0.2)	(0)	(0)	(0)
150	**300**	6	6	3.1	2.89	3	3	3
		(9)	(9)	(8.2)	(0.2)	(0)	(0)	(0)
200	**40**	5.98	5.95	3.1	2.36	3.03	2.98	2.99
		(8.94)	(8.75)	(0.12)	(2.11)	(0.04)	(0.02)	(0)
200	**50**	6	5.98	3.09	2.69	3.02	2.99	2.99
		(9)	(8.94)	(0.11)	(1.18)	(0.03)	(0)	(0)
200	**100**	6	6	3.29	2.91	3	2.99	3
		(9)	(9)	(0.77)	(0.2)	(0.02)	(0)	(0)
200	**150**	6	6	3	2.92	3	3	3
		(9)	(9)	(8.05)	(0.16)	(0.01)	(0)	(0)
200	**200**	6	6	3	2.92	3	3	3
		(9)	(9)	(8.6)	(0.16)	(0)	(0)	(0)
200	**300**	6	6	3	2.92	3	3	3
		(9)	(9)	(8.52)	(0.16)	(0)	(0)	(0)
300	**40**	6	6	3	1.92	3	2.95	2.97
		(9)	(9)	(0.16)	(4)	(0.07)	(0.07)	(0.02)
300	**50**	6	6	3.12	2.22	3.03	2.98	2.99
		(9)	(9)	(0.18)	(2.91)	(0.05)	(0.01)	(0)
300	**100**	6	6	3	2.87	3	3	3
		(9)	(9)	(0.35)	(0.44)	(0)	(0)	(0)
300	**150**	6	6	3.1	2.92	3	3	3
		(9)	(9)	(7.9)	(0.18)	(0)	(0)	(0)
300	**200**	6	6	3.79	2.94	3	3	3
		(9)	(9)	(8.15)	(0.1)	(0)	(0)	(0)
300	**300**	6	6	5.82	2.96	3	3	3
		(9)	(9)	(8.26)	(0.08)	(0)	(0)	(0)

Table 8: DGP5 b: Estimtion of number of factors

N	T	BN02a	BN02b	ABC	Ona09	Ona10	AH_ER	AH_GR
40	40	6	6	3.95	1.27	3.46	4.74	2.67
		(9)	(9)	(9)	(6.57)	(7.65)	(5.93)	(2.31)
40	50	6	6	4.29	1.13	3.28	4.86	2.64
		(9)	(9)	(6.43)	(6.55)	(7.84)	(6.36)	(2.36)
40	100	6	6	4.73	0.81	3.82	5.31	2.64
		(9)	(9)	(6.87)	(6.48)	(8.38)	(7.26)	(2.14)
40	150	6	6	4.65	0.7	3.94	5.55	2.45
		(9)	(9)	(6.79)	(6.79)	(8.51)	(7.86)	(2.08)
40	200	6	6	5.1	0.7	4.32	5.69	2.38
		(9)	(9)	(7.64)	(6.78)	(8.71)	(8.05)	(2.2)
40	300	6	6	5.63	0.74	3.98	5.68	2.4
		(9)	(9)	(8.21)	(6.61)	(8.73)	(8.04)	(2.1)
50	40	6	6	4.76	0.99	1.08	4.22	2.77
		(9)	(9)	(7.14)	(6.85)	(7.68)	(4.81)	(2.22)
50	50	6	6	4.78	0.91	0.93	4.05	2.63
		(9)	(9)	(7.22)	(6.77)	(7.77)	(4.67)	(2.2)
50	100	6	6	5.56	0.69	0.46	4.57	2.59
		(9)	(9)	(8.44)	(7.13)	(8.27)	(5.51)	(2.13)
50	150	6	6	5.45	0.68	0.43	4.5	2.53
		(9)	(9)	(8.15)	(7.14)	(8.31)	(5.49)	(2.14)
50	200	6	6	5.86	0.66	0.3	4.72	2.52
		(9)	(9)	(8.74)	(7)	(8.53)	(5.81)	(2.01)
50	300	6	6	5.96	0.8	0.41	4.85	2.37
		(9)	(9)	(8.92)	(6.64)	(8.44)	(6.14)	(2.12)
100	40	6	6	0.6	0.82	0.17	3.24	2.48
		(9)	(9)	(8.6)	(7.57)	(8.27)	(3.59)	(2.27)
100	50	6	6	0.18	0.73	0.14	3.25	2.5
		(9)	(9)	(8.64)	(7.46)	(8.33)	(3.45)	(2.27)
100	100	6	6	5.85	0.72	0.04	3.49	2.42
		(9)	(9)	(8.85)	(7.7)	(8.78)	(3.81)	(2.22)
100	150	6	6	6	0.71	0.02	3.36	2.24
		(9)	(9)	(9)	(7.52)	(8.9)	(3.83)	(2.3)
100	200	6	6	5.95	0.59	0.01	3.45	2.41
		(9)	(9)	(8.95)	(7.72)	(8.95)	(3.72)	(2.29)
100	300	6	6	6	0.79	0	3.62	2.29
		(9)	(9)	(9)	(7.53)	(8.98)	(4.28)	(2.39)
150	40	6	6	0.41	0.73	0.32	2.83	2.47
		(9)	(9)	(7.17)	(7.49)	(7.79)	(2.88)	(2.08)
150	50	6	6	0.61	0.84	0.22	2.68	2.35
		(9)	(9)	(7.57)	(7.49)	(8.06)	(2.89)	(2.2)
150	100	6	6	0.35	0.73	0.15	2.52	2.27
		(9)	(9)	(8.35)	(7.49)	(8.31)	(2.55)	(1.91)
150	150	6	6	5.5	0.78	0.17	2.36	2.17
		(9)	(9)	(8.5)	(7.43)	(8.26)	(2.29)	(1.91)
150	200	6	6	5.66	0.8	0.16	2.45	2.36
		(9)	(9)	(8.62)	(7.54)	(8.35)	(1.92)	(1.54)
150	300	6	6	5.65	0.77	0.12	2.43	2.34
		(9)	(9)	(8.65)	(7.5)	(8.47)	(1.77)	(1.43)
200	40	6	6	3.01	0.9	0.42	2.37	2.28
		(9)	(9)	(5.57)	(7.24)	(7.43)	(2.31)	(1.93)
200	50	6	6	3.41	0.93	0.48	2.41	2.43
		(9)	(9)	(5.97)	(7.28)	(7.28)	(2.15)	(1.83)
200	100	6	6	3.14	0.92	0.63	2.34	2.41
		(9)	(9)	(4.74)	(7.4)	(6.82)	(1.55)	(1.34)
200	150	6	6	3.65	0.87	0.93	2.45	2.54
		(9)	(9)	(5.01)	(7.3)	(6.05)	(1.01)	(0.85)
200	200	6	6	3.8	0.84	1.33	2.62	2.67
		(9)	(9)	(4.7)	(7.12)	(4.88)	(0.7)	(0.58)
200	300	6	6	3.31	1.03	1.89	2.75	2.79
		(9)	(9)	(2.89)	(6.78)	(3.21)	(0.38)	(0.3)
300	40	6	6	0.67	0.85	0.13	2.78	2.44
		(9)	(9)	(7.61)	(7.43)	(8.42)	(2.83)	(2.13)
300	50	6	6	0.31	0.84	0.15	2.59	2.3
		(9)	(9)	(7.27)	(7.38)	(8.3)	(2.77)	(2.12)
300	100	6	6	0.23	0.72	0.13	2.58	2.31
		(9)	(9)	(8.17)	(7.74)	(8.53)	(2.78)	(2.27)
300	150	6	6	5.6	0.59	0.08	2.39	2.2
		(9)	(9)	(8.6)	(7.51)	(8.63)	(2.59)	(2.15)
300	200	6	6	5.45	0.66	0.09	2.4	2.27
		(9)	(9)	(8.45)	(7.61)	(8.59)	(2.33)	(2)
300	300	6	6	5.65	0.67	0.07	2.28	2.11
		(9)	(9)	(8.65)	(7.56)	(8.67)	(2.28)	(1.97)

Table 9: DGP10: Estimtion of number of factors

		Static Factors							Dynamic Factors								
N	**T**	BNra	BNrb	ABCr	Ona09r	Ona10r	AHERr	AHGRr	BN07a	BN07b	Ona09	HLq	BPq	JORq	SWa	SWb
50	**50**	9.01	7.05	6.26	1.81	6	2.1	3.45	1	1	1.53	2.66	2.08	3.56	6	6
		(13.28)	(7.41)	(0.66)	(21.17)	(0.22)	(19.42)	(12.6)	(1)	(1)	(0.46)	(2.58)	(0.08)	(7.26)	(16)	(16)
50	**100**	6.3	5.12	6.26	1.47	6	1.99	4.22	1	1	1.52	2.7	2.18	2.83	6	6
		(9.16)	(9.79)	(0.5)	(22.58)	(0)	(20)	(8.87)	(1)	(1)	(0.48)	(2.62)	(0.18)	(5.81)	(16)	(16)
50	**150**	4.77	4.1	6.36	1.17	6.01	2.35	4.57	1	1	1.54	2.04	2.29	2.3	6	6
		(9.95)	(10.63)	(0.68)	(24.02)	(0.01)	(18.25)	(7.12)	(1)	(1)	(0.45)	(0.08)	(0.29)	(4.41)	(16)	(16)
50	**200**	4.04	3.63	6.22	1.21	6.01	2.2	4.8	1	1	1.56	2.04	2.38	2.37	6	6
		(10.51)	(10.86)	(0.3)	(24.12)	(0.01)	(19)	(6)	(1)	(1)	(0.44)	(0.08)	(0.38)	(4.66)	(16)	(16)
50	**300**	3.03	2.84	6.24	1.12	6	2.17	5.45	1	1	1.51	2.02	2.49	2	6	6
		(12.09)	(12.6)	(0.32)	(24.28)	(0)	(19.12)	(2.75)	(1)	(1)	(0.48)	(0.06)	(0.49)	(3.7)	(16)	(16)
100	**50**	7.21	6.16	6.24	1.54	5.99	2.35	4.05	1	1	1.59	3.02	2	3.82	6	6
		(13.02)	(11.81)	(0.76)	(22.28)	(0.01)	(18.21)	(9.64)	(1)	(1)	(0.4)	(3.9)	(0)	(8.61)	(16)	(16)
100	**100**	4.97	3.98	6.16	1.15	6.01	2.85	5.07	1	1	1.56	2.02	2.01	3.03	6	6
		(13.78)	(12.97)	(0.32)	(24.29)	(0.01)	(15.75)	(4.62)	(1)	(1)	(0.43)	(0.02)	(0.01)	(6.71)	(16)	(16)
100	**150**	4.02	3.4	6.26	1.13	6	3.32	5.77	1	1	1.54	2	2.02	2.6	6	6
		(13.41)	(12.97)	(0.54)	(24.33)	(0)	(13.37)	(1.12)	(1)	(1)	(0.45)	(0.08)	(0.02)	(5.52)	(16)	(16)
100	**200**	3.31	2.9	6.1	1.08	6	3.72	5.77	1	1	1.53	2	2.03	2.14	6	6
		(13.38)	(13.52)	(0.1)	(24.53)	(0)	(11.37)	(1.12)	(1)	(1)	(0.46)	(0.04)	(0.03)	(4.26)	(16)	(16)
100	**300**	2.75	2.54	6.14	1.02	6	4.12	6	1	1	1.59	2.04	2.08	1.98	6	6
		(13.59)	(13.87)	(0.22)	(24.87)	(0)	(9.37)	(0)	(1)	(1)	(0.4)	(0.04)	(0.08)	(3.82)	(16)	(16)
150	**50**	6.8	6.23	6.12	1.61	6.02	2.72	4.22	1	1	1.6	3.84	2	3.77	6	6
		(15.88)	(14.71)	(0.16)	(22.03)	(0.08)	(16.38)	(8.88)	(1)	(1)	(0.39)	(7.36)	(0)	(8.62)	(16)	(16)
150	**100**	4.94	4.31	6.1	1.22	6.01	3.2	5.4	1	1	1.58	2.1	2	3.12	6	6
		(15.69)	(14.59)	(0.22)	(23.83)	(0.02)	(14)	(3)	(1)	(1)	(0.41)	(0.34)	(0)	(7)	(16)	(16)
150	**150**	3.95	3.3	6.06	1.03	6	3.62	5.72	1	1	1.54	2	2	2.84	6	6
		(15.13)	(14.36)	(0.06)	(24.83)	(0)	(11.87)	(1.37)	(1)	(1)	(0.45)	(0)	(0)	(6.26)	(16)	(16)
150	**200**	3.61	3.04	6.08	1	6	4.4	5.92	1	1	1.56	1.98	2	2.44	6	6
		(14.07)	(13.76)	(0.08)	(25)	(0)	(8)	(0.37)	(1)	(1)	(0.44)	(0.02)	(0)	(5.17)	(16)	(16)
150	**300**	2.83	2.54	6.06	1	6	5.1	6	1	1	1.55	2.02	2	2.08	6	6
		(14.36)	(14.38)	(0.06)	(24.95)	(0)	(4.5)	(0)	(1)	(1)	(0.44)	(0.02)	(0)	(4.17)	(16)	(16)
200	**50**	6.17	5.79	6.32	1.39	5.99	2.65	4.45	1	1	1.51	2.96	2	3.84	6	6
		(15.86)	(14.86)	(1.12)	(23.13)	(0.01)	(16.75)	(7.75)	(1)	(1)	(0.48)	(3.84)	(0)	(8.84)	(16)	(16)
200	**100**	5.02	4.52	6.04	1.03	6	3.52	5.32	1	1	1.55	2.08	2	3.09	6	6
		(16.85)	(15.8)	(0.04)	(24.73)	(0)	(12.37)	(3.37)	(1)	(1)	(0.44)	(0.36)	(0)	(6.88)	(16)	(16)
200	**150**	4.08	3.61	6.04	1.03	6.02	4.22	5.82	1	1	1.57	2.08	2	2.44	6	6
		(16.61)	(15.18)	(0.08)	(24.78)	(0.08)	(8.87)	(0.87)	(1)	(1)	(0.43)	(0.08)	(0)	(5.14)	(16)	(16)
200	**200**	3.54	2.96	6.08	1.05	6	4.47	5.97	1	1	1.57	2	2	2.22	6	6
		(15.18)	(14.4)	(0.08)	(24.7)	(0)	(7.62)	(0.12)	(1)	(1)	(0.42)	(0)	(0)	(4.55)	(16)	(16)
200	**300**	2.82	2.5	6.02	1.03	6	5.27	6	1	1	1.58	2	2	2.04	6	6
		(14.81)	(14.64)	(0.02)	(24.78)	(0)	(3.62)	(0)	(1)	(1)	(0.42)	(0)	(0)	(4.02)	(16)	(16)
300	**50**	6.4	6.16	6.32	1.34	5.98	2.77	4.42	1	1	1.59	3.12	2	3.78	6	6
		(17.86)	(17.02)	(1)	(23.28)	(0.02)	(16.13)	(7.87)	(1)	(1)	(0.4)	(4.48)	(0)	(8.63)	(16)	(16)
300	**100**	4.79	4.47	6.16	1.1	6	3.57	5.52	1	1	1.56	2.02	2	3.07	6	6
		(17)	(16.28)	(0.6)	(24.45)	(0)	(12.12)	(2.37)	(1)	(1)	(0.43)	(0.1)	(0)	(6.9)	(16)	(16)
300	**150**	3.95	3.59	6	1.08	6	4.72	5.82	1	1	1.54	2.04	2	2.54	6	6
		(16.27)	(16.09)	(0)	(24.53)	(0)	(6.37)	(0.87)	(1)	(1)	(0.45)	(0.04)	(0)	(5.46)	(16)	(16)
300	**200**	4.19	3.8	6	1	6	5	5.97	1	1	1.57	1.98	2	2.13	6	6
		(17.47)	(16.42)	(0)	(24.95)	(0)	(5)	(0.12)	(1)	(1)	(0.42)	(0.02)	(0)	(4.32)	(16)	(16)
300	**300**	3.15	2.65	6	1.03	6	5.65	6	1	1	1.57	2	2	2	6	6
		(15.49)	(14.77)	(0)	(24.78)	(0)	(1.75)	(0)	(1)	(1)	(0.43)	(0)	(0)	(3.96)	(16)	(16)

Table 10: DGP9: Estimtion of number of factors

		Static Factors							Dynamic Factors							
N	**T**	**BNra**	**BNrb**	**ABCr**	**Ona09r**	**Ona10r**	**AHERr**	**AHGRr**	**BN07a**	**BN07b**	**Ona09**	**HLq**	**BPq**	**JORq**	**SWa**	**SWb**
50	**50**	9.94	8.57	6.12	4.39	6.01	5.75	5.92	1	1	3.92	1.96	2.02	5.58	6	6
		(18.48)	(11.07)	(0.2)	(9.48)	(0,16)	(0.89)	(0.17)	(1)	(1)	(8.86)	(0.06)	(0.02)	(13.51)	(16)	(16)
50	**100**	8.27	7.66	6.22	4.71	6.01	5.97	6	1	1	4.51	1.98	2.01	5.77	6	6
		(10.14)	(6.9)	(0.22)	(5.83)	(0,01)	(0.09)	(0)	(1)	(1)	(10.6)	(0.02)	(0.01)	(14.68)	(16)	(16)
50	**150**	7.23	6.93	6.08	4.82	6.01	5.99	6	1	1	4.52	2	2.02	5.83	6	6
		(4.78)	(3.32)	(0.08)	(5)	(0,01)	(0)	(0)	(1)	(1)	(10.62)	(0)	(0.02)	(14.93)	(16)	(16)
50	**200**	7	6.75	6.1	4.9	6.01	5.99	6	1	1	4.5	2	2.03	5.87	6	6
		(3.45)	(2.35)	(0.1)	(5.03)	(0,01)	(0)	(0)	(1)	(1)	(10.48)	(0)	(0.03)	(15.17)	(16)	(16)
50	**300**	6.45	6.34	6.2	4.96	6.01	6	6	1	1	4.39	2	2.05	5.86	6	6
		(1.22)	(0.81)	(0.24)	(4.5)	(0,01)	(0)	(0)	(1)	(1)	(10.11)	(0)	(0.05)	(15.1)	(16)	(16)
100	**50**	8.8	8.28	6.2	4.51	6.01	5.92	5.96	1	1	4.08	2.08	2	5.75	6	6
		(13.82)	(10.87)	(0.48)	(7.33)	(0,03)	(0.28)	(0.15)	(1)	(1)	(9.47)	(0.32)	(0)	(14.49)	(16)	(16)
100	**100**	7.54	7.03	6.02	4.79	6.01	6	6	1	1	4.49	1.98	2	5.92	6	6
		(7.51)	(4.63)	(0.02)	(4.5)	(0,01)	(0)	(0)	(1)	(1)	(10.49)	(0.02)	(0)	(15.49)	(16)	(16)
100	**150**	7.23	6.85	6.02	4.88	6	6	6	1	1	4.6	2	2	5.94	6	6
		(5.81)	(3.53)	(0.02)	(4.36)	(0)	(0)	(0)	(1)	(1)	(10.96)	(0)	(0)	(15.63)	(16)	(16)
100	**200**	6.71	6.48	6.04	5.07	6.01	6	6	1	1	4.66	2	2	5.97	6	6
		(2.98)	(1.72)	(0.04)	(4.95)	(0,01)	(0)	(0)	(1)	(1)	(11.18)	(0)	(0)	(15.85)	(16)	(16)
100	**300**	6.29	6.18	6	4.99	6	6	6	1	1	4.51	2	2	5.98	6	6
		(0.85)	(0.44)	(0)	(4.48)	(0)	(0)	(0)	(1)	(1)	(10.67)	(0)	(0)	(15.86)	(16)	(16)
150	**50**	8.67	8.38	6.24	4.49	6.01	5.94	5.99	1	1	4.32	2	2	5.84	6	6
		(14.01)	(12.32)	(0.84)	(6.64)	(0,06)	(0.17)	(0)	(1)	(1)	(10.08)	(0)	(0)	(14.99)	(16)	(16)
150	**100**	7.68	7.31	6	4.65	6.01	6	6	1	1	4.49	1.98	2	5.96	6	6
		(8.54)	(6.42)	(0)	(4.52)	(0,01)	(0)	(0)	(1)	(1)	(10.54)	(0.02)	(0)	(15.77)	(16)	(16)
150	**150**	7.15	6.74	6	4.72	6.01	6	6	1	1	4.56	2	2	5.98	6	6
		(5.77)	(3.3)	(0)	(4.08)	(0,01)	(0)	(0)	(1)	(1)	(10.6)	(0)	(0)	(15.91)	(16)	(16)
150	**200**	6.87	6.53	6.02	4.87	6	6	6	1	1	4.55	2	2	5.98	6	6
		(4.08)	(2.04)	(0.02)	(4.89)	(0)	(0)	(0)	(1)	(1)	(10.75)	(0)	(0)	(15.86)	(16)	(16)
150	**300**	6.34	6.18	6.02	4.96	6.01	6	6	1	1	4.61	2	2	5.99	6	6
		(1.25)	(0.55)	(0.02)	(4.44)	(0,01)	(0)	(0)	(1)	(1)	(11.07)	(0)	(0)	(15.98)	(16)	(16)
200	**50**	8.55	8.32	6.2	4.57	6.01	5.93	5.99	1	1	4.18	2	2	5.84	6	6
		(13.58)	(12.2)	(0.76)	(6.63)	(0,05)	(0.23)	(0)	(1)	(1)	(9.69)	(0)	(0)	(15.01)	(16)	(16)
200	**100**	7.84	7.56	6	4.74	6.01	5.99	6	1	1	4.61	2	2	5.96	6	6
		(9.71)	(7.91)	(0)	(4.64)	(0,01)	(0)	(0)	(1)	(1)	(10.91)	(0)	(0)	(15.74)	(16)	(16)
200	**150**	7.02	6.72	6.02	4.75	6.01	6	6	1	1	4.6	2	2	5.98	6	6
		(5.15)	(3.51)	(0.02)	(5.75)	(0,01)	(0)	(0)	(1)	(1)	(10.97)	(0)	(0)	(15.9)	(16)	(16)
200	**200**	6.88	6.5	6	4.84	6.01	6	6	1	1	4.49	2	2	5.99	6	6
		(4.32)	(2.1)	(0)	(5.28)	(0,01)	(0)	(0)	(1)	(1)	(10.53)	(0)	(0)	(15.95)	(16)	(16)
200	**300**	6.44	6.21	6	5.03	6	6	6	1	1	4.61	2	2	5.99	6	6
		(1.79)	(0.62)	(0)	(4.65)	(0)	(0)	(0)	(1)	(1)	(10.98)	(0)	(0)	(15.98)	(16)	(16)
300	**50**	8.68	8.53	6.1	4.48	6	5.94	5.97	1	1	4.16	2	2	5.89	6	6
		(14.51)	(13.61)	(0.5)	(7.24)	(0,01)	(0.2)	(0.09)	(1)	(1)	(9.31)	(0)	(0)	(15.31)	(16)	(16)
300	**100**	7.75	7.54	6	4.67	6.01	6	6	1	1	4.49	2	2	5.97	6	6
		(9.43)	(8.01)	(0)	(5.32)	(0,03)	(0)	(0)	(1)	(1)	(10.52)	(0)	(0)	(15.79)	(16)	(16)
300	**150**	7.15	6.98	6.02	4.71	6.01	6	6	1	1	4.51	2	2	5.99	6	6
		(6.15)	(5.25)	(0.02)	(5.64)	(0,01)	(0)	(0)	(1)	(1)	(10.61)	(0)	(0)	(15.94)	(16)	(16)
300	**200**	6.84	6.62	6	4.79	6.01	6	6	1	1	4.57	2	2	5.99	6	6
		(4.37)	(3.24)	(0)	(5.67)	(0,01)	(0)	(0)	(1)	(1)	(10.78)	(0)	(0)	(15.95)	(16)	(16)
300	**300**	6.64	6.33	6	4.82	6.01	6	6	1	1	4.5	2	2	6	6	6
		(3.3)	(1.31)	(0)	(4.17)	(0,01)	(0)	(0)	(1)	(1)	(10.73)	(0)	(0)	(16)	(16)	(16)

Figure 1: DGP6a: Estimtion of number of factors

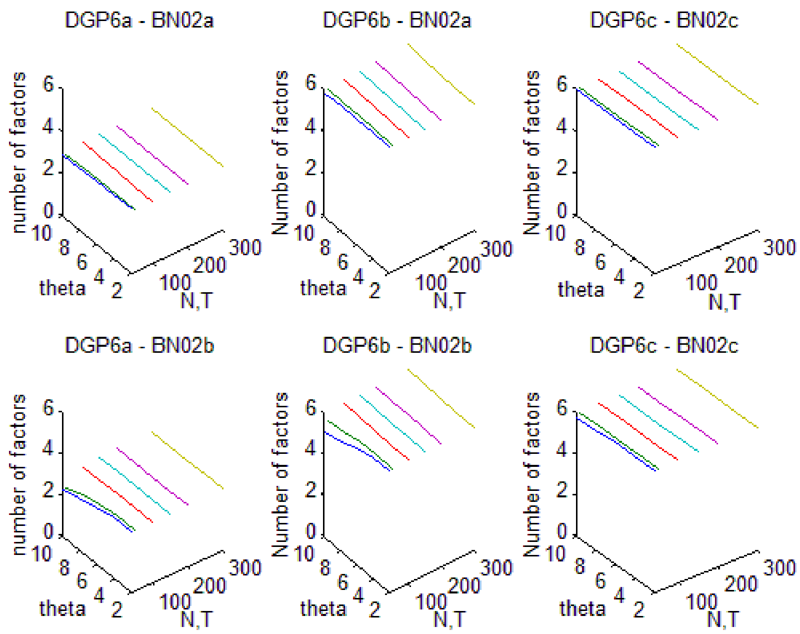

Figure 2: DGP6b: Estimtion of number of factors

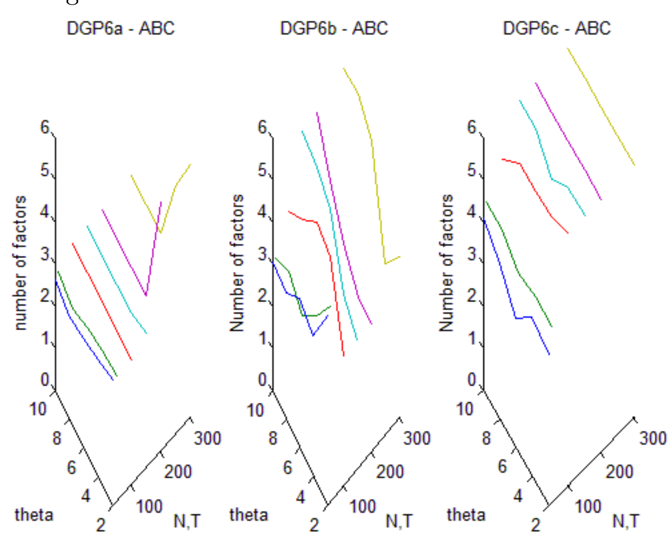

Figure 3: DGP6c: Estimtion of number of factors

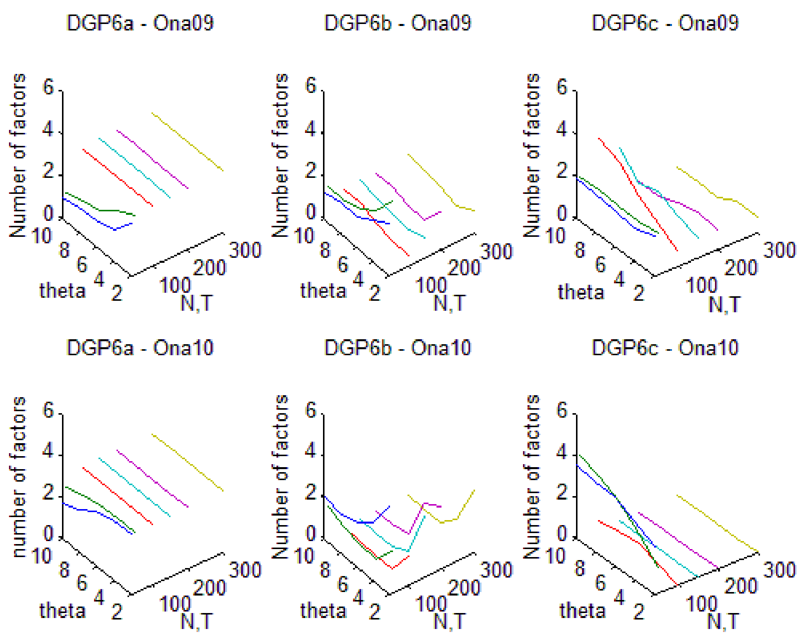

Figure 4: DGP6d: Estimtion of number of factors

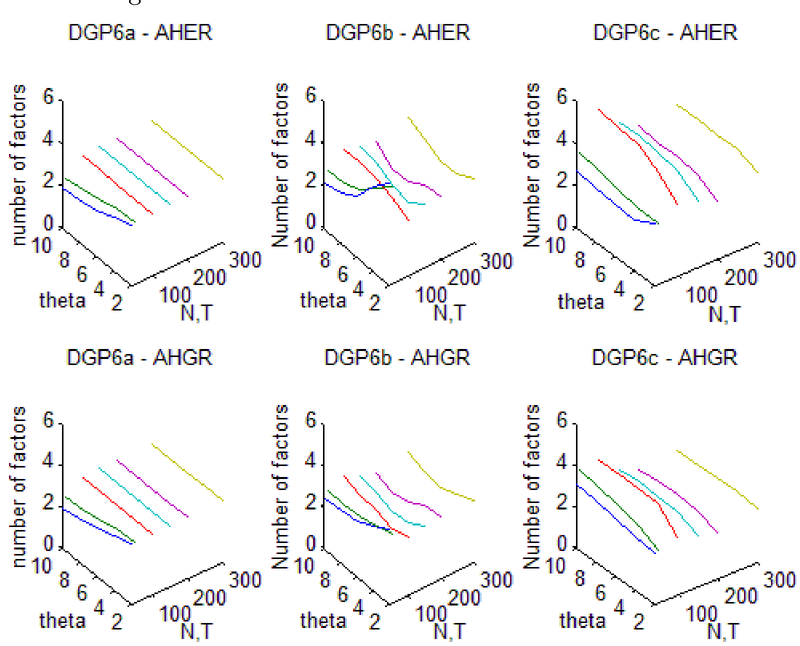

Figure 5: DGP7a: Estimtion of number of factors

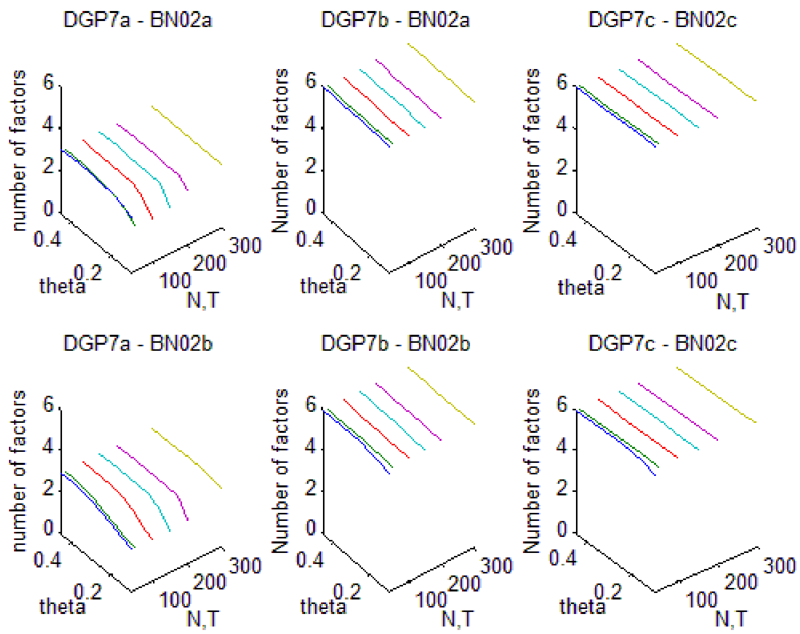

Figure 6: DGP7b: Estimtion of number of factors

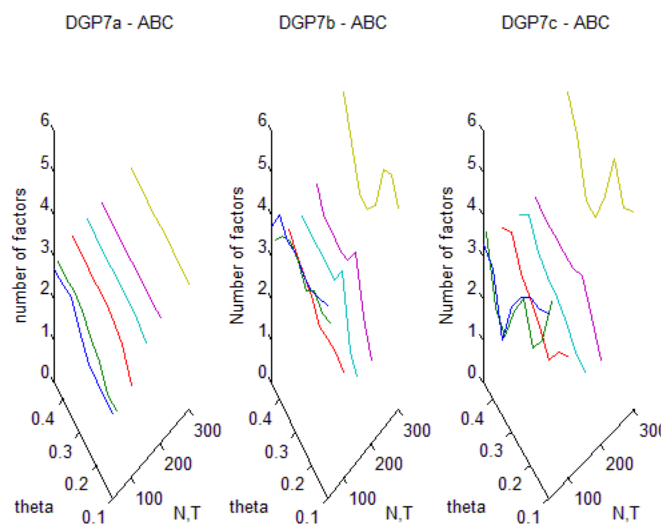

Figure 7: DGP7c: Estimtion of number of factors

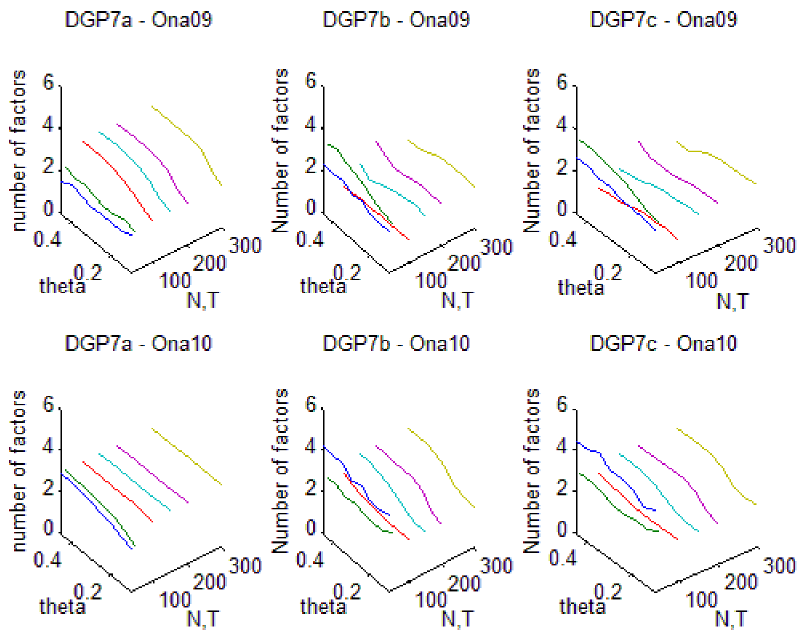

Figure 8: DGP7d: Estimtion of number of factors

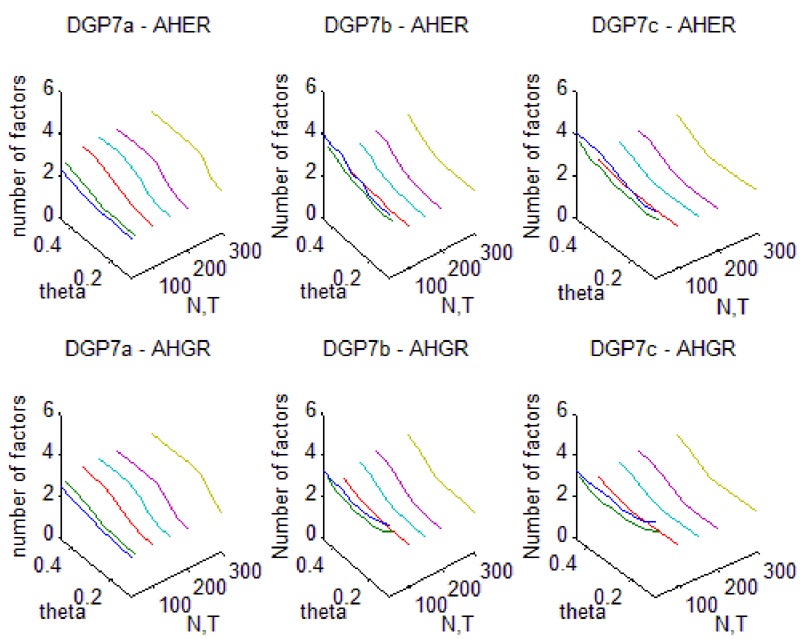

Figure 9: DGP8a: Estimtion of number of factors

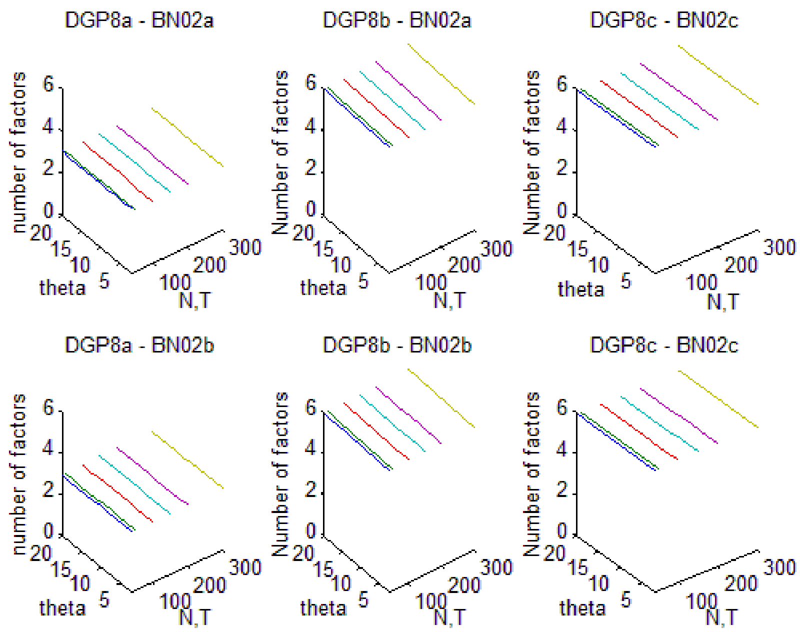

Figure 10: DGP8b: Estimtion of number of factors

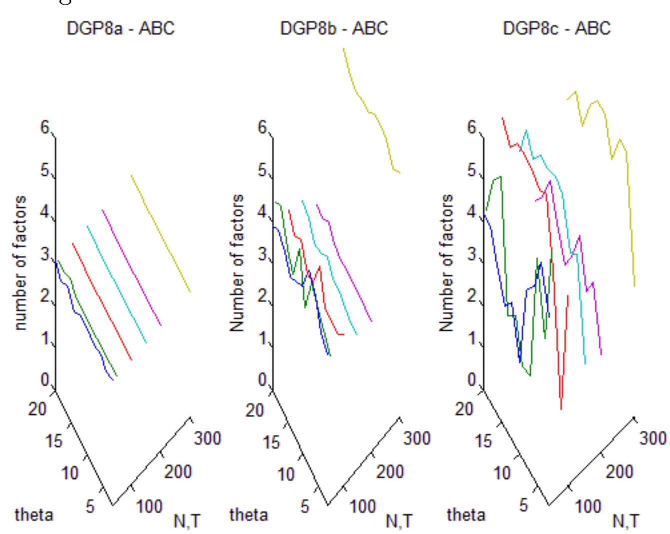

Figure 11: DGP8c: Estimtion of number of factors

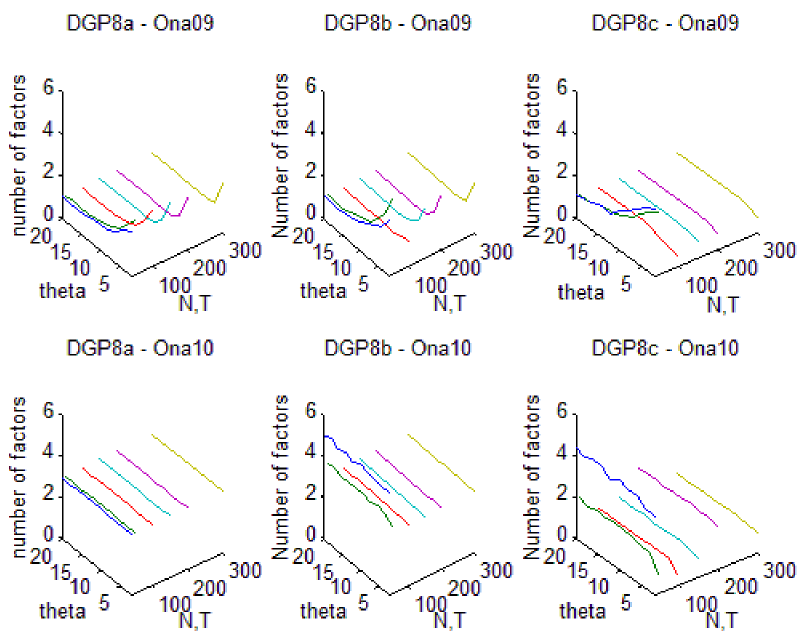

Figure 12: DGP8d: Estimtion of number of factors

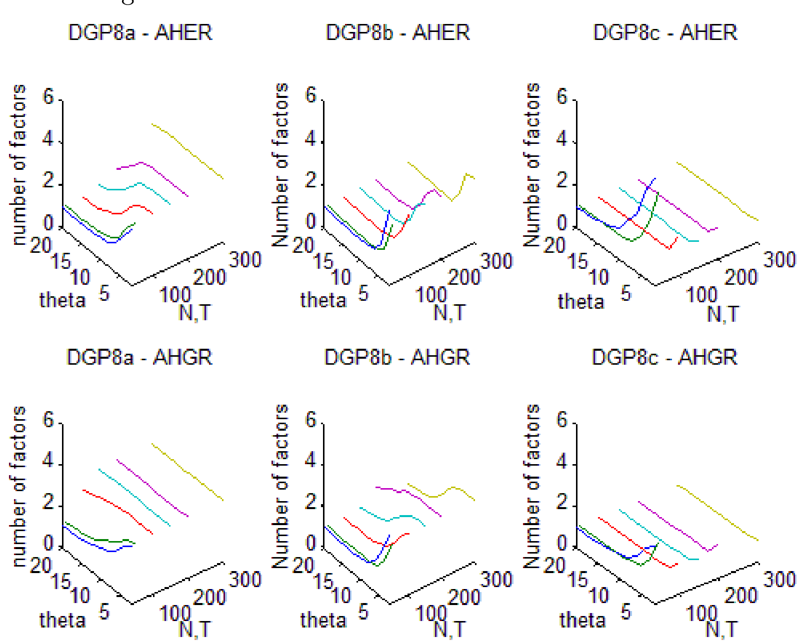

Figure 13: DGP11a: Estimtion of number of static factors

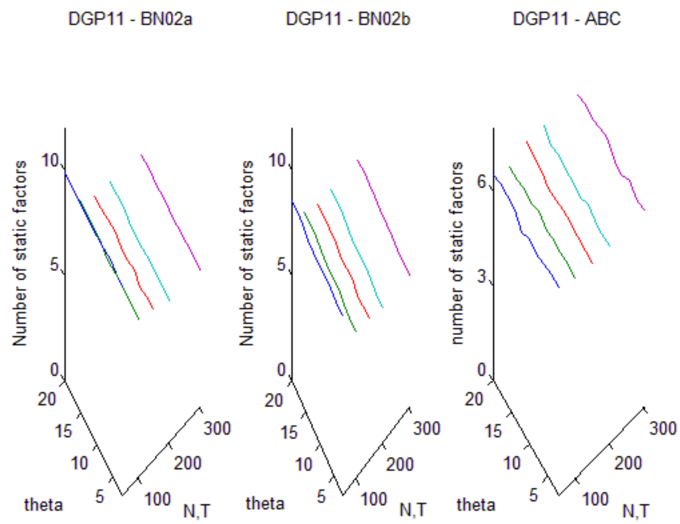

Figure 14: DGP11b: Estimtion of number of static factors

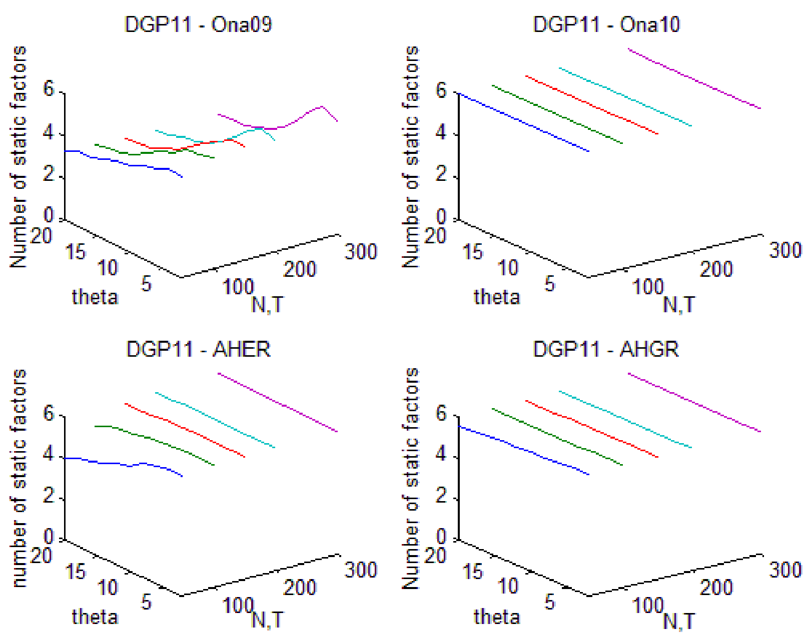

Figure 15: DGP11c: Estimtion of number of dynamic factors

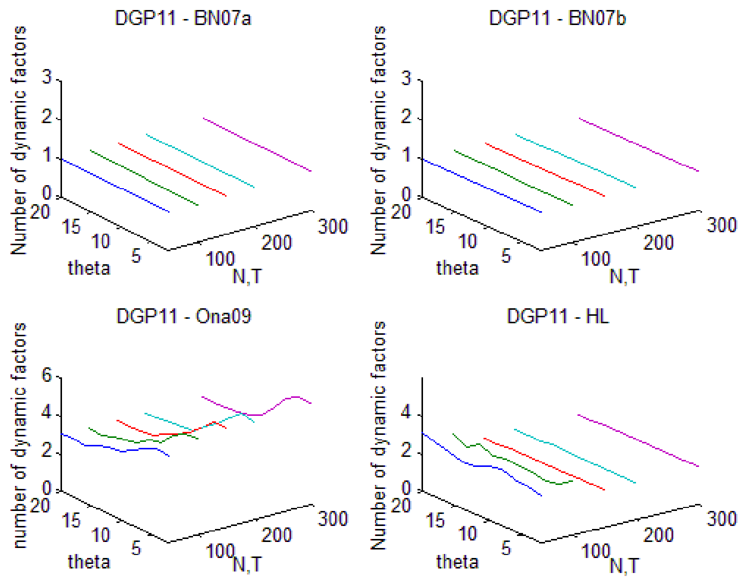

Figure 16: DGP11d: Estimtion of number of dynamic factors

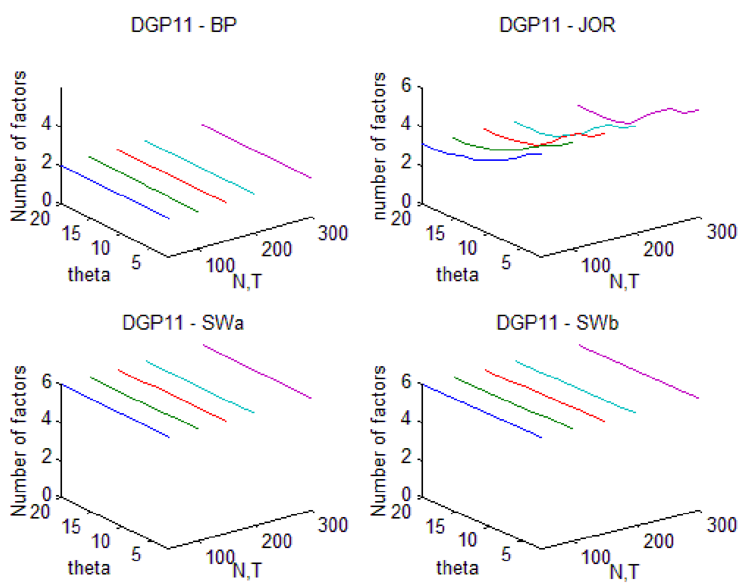

Figure 17: DGP12a: Estimtion of number of static factors

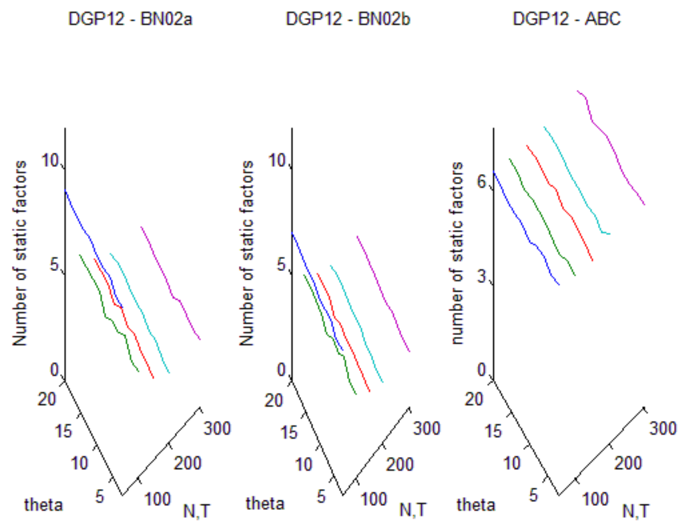

Figure 18: DGP12b: Estimtion of number of static factors

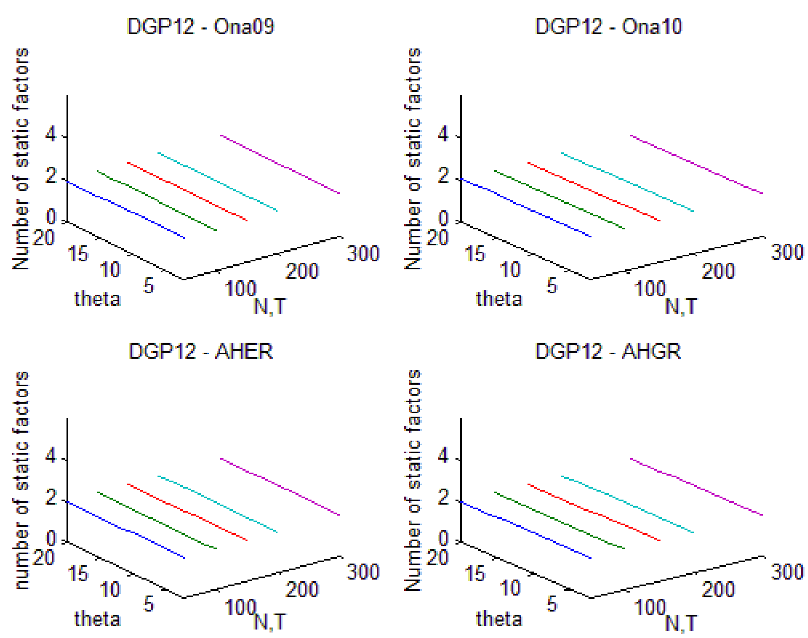

Figure 19: DGP12c: Estimtion of number of dynamic factors

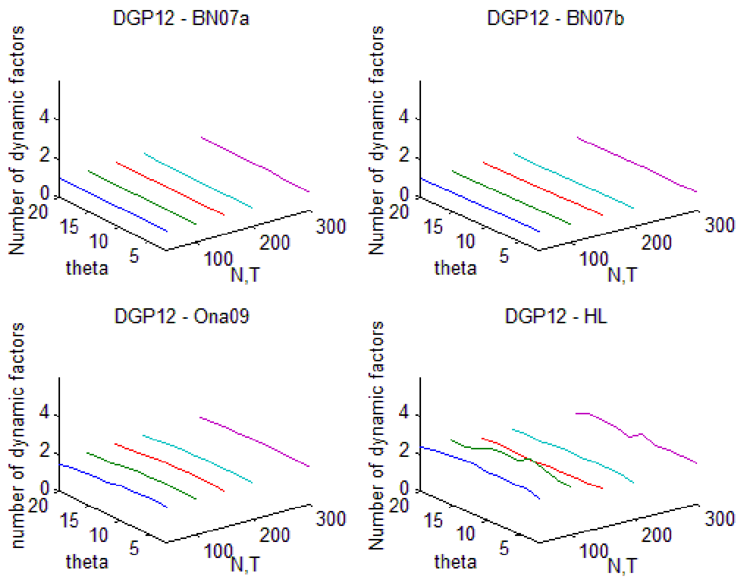

Figure 20: DGP12d: Estimtion of number of dynamic factors