Department of Economics

Dynamic Factor Extraction of Cross-Sectional Dependence in Panel Unit Root Tests

George Kapetanios

Working Paper No. 509 February 2004 ISSN 1473-0278

Queen Mary
University of London

Dynamic Factor Extraction of Cross-Sectional Dependence In Panel Unit Root Tests

George Kapetanios*
Queen Mary, University of London

February 10, 2004

Abstract

Recently, considerable emphasis has been placed on the problems arising out of cross-sectional dependence in panel unit root tests. This paper adopts the factor based cross-sectional dependence paradigm of Bai and Ng (2004) but suggests alternative factor extraction methods. Some theoretical results for these methods are provided. Further, a detailed Monte Carlo study of these methods for multiple and persistent factors is undertaken. It is found that results are radically different to the serially uncorrelated single factor case. Tests perform much worse and in some cases it is preferable not to correct at all for cross-sectional dependence.
JEL Codes: C32, C33
Keywords: Panel Unit Root Tests, Factor Models, Subspace Algorithms

1 Introduction

Over the past decade the problem of testing for unit roots in heterogeneous panels has attracted a great of deal attention. Papers that deal with the

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

problem include Choi (2001), Chang (2000), Im, Pesaran, and Shin (2003), Levin, Lin, and Lu (2002) and Maddal and Wu (1999). Baltagi and Kao (2000) provide a review. This literature, however, assumed that the individual time series in the panel were cross- sectionally independently distributed. While it was recognized that this was a restrictive assumption, particularly in the context of cross-country (region) regressions, it was thought that cross-sectionally de-meaning the series before application of the panel unit root test could partly deal with the problem. However, it was clear that cross-section de-meaning could not work in general where pair-wise cross-section covariances of the error terms differed across the individual series. Recognizing this deficiency new panel unit root tests have been proposed in the literature by Bai and Ng (2004), Chang (2002), Harvey and Bates (2002), Moon and Perron (2004), Phillips and Sul (2002) and Pesaran (2003). Chang (2002) proposes a non-linear instrumental variable approach to deal with the cross section dependence of a general form and establishes that individual Dickey-Fuller (DF) or the Augmented DF (ADF) statistics are asymptotically independent when an integrable function of the lagged dependent variables are used as instruments. The test proposed by Harvey and Bates (2002) is also valid for general specifications of error cross correlations, but is limited as it requires the parameters to be the same across all the series.

Bai and Ng (2004), Moon and Perron (2004), Phillips and Sul (2002) and Pesaran (2003) avoid the restrictive nature of cross section de-meaning procedures by allowing the common factors to have differential effects on different cross section units. In the context of a residual one-factor model Phillips and Sul (2002) show that in the presence of cross section dependence the standard panel unit root tests are no longer asymptotically similar, and propose an orthogonalization procedure which eliminates the common factors before preceding to the application of standard panel unit root tests. Pesaran (2003) considers a factor model but does not use a factor estimation method but rather augments standard Dickey Fuller regressions with the cross sectional average of the series, thus accounting for the factor effects.

This paper adopts a similar approach to the work of Bai and Ng (2004) but considers a number of extensions. In particular, we consider the possibility of alternative factor extraction and estimation methods to the approximately dynamic approach of Stock and Watson (1999). The alternative methods we consider are the dynamic principal component method developed in a series of papers by Forni and Reichlin (1996, 1998); Forni, Hallin, Lippi, and Reichlin (2000, 2004) and the parametric state space dynamic approach of Kapetanios and Marcellino (2003). In that vein we provide some theoretical results for general factor estimation methods. The second, and perhaps most important, contribution is a Monte Carlo study that compares the tests based on alternative factor extraction methods. A significant element of that analysis is the consideration of multiple dynamic factors. Previous simulation work in the literature has overwhelmigly concetrated on single serially uncorrelated factors to introduce cross sectional dependence in panel datasets . However, it is clear that this is both restrictive and unlikely to hold in practice. Factors that underlie cross dependence in dynamic panels are likely to be serially correlated and in many cases more than one factors will be needed to absorb cross sectional dependence. Once we consider this extended setup, it is clear that the performance of the various factor based methods varies considerably in a number of dimensions. These issues are discussed.

The paper is structured as follows: Section 2 provides some theoretical results for general factor estimation methods. Section 3 introduces the alternative factor estimation methods we consider. Section 4 describes the design and results of our extensive Monte Carlo study. Finally, Section 5 concludes.

2 Theoretical Results

Consider a sample of N cross sections observed over T time periods. Let the stochastic process $y_{i,t}$ be generated by

$$y_{i,t} = \phi_i y_{i,t-1} + \eta_{i,t}, \quad i = 1, \dots, N, \quad t = 1, \dots, T \tag{1}$$

where initial values $y_{i,0}$ are given. We are interested in testing the null hypothesis of $\phi_i = 1$ for all i. Rewritting (1) as

$$\Delta y_{i,t} = \beta_i y_{i,t-1} + \eta_{i,t} \tag{2}$$

where $\beta_i = \phi_i - 1$, the null hypothesis becomes

$$H_0 : \beta_i = 0, \quad \forall i \tag{3}$$

This is an illustrative framework for our analysis, which can be extended in a variety of ways, such as, e.g., including deterministic terms, without affecting its essence. We consider the following structure for $\eta_{i,t}$

$$\eta_{i,t} = f_t' \gamma_i + \epsilon_{i,t}$$

where $f_t = (\ f_{1,t},\ f_{2,t}, ...,\ f_{k,t})'$ is a $k \times 1$ vector of factor variables are time t. Define $f = (\ f_1,\ f_2, ...,\ f_T)'$. We assume the existence of an estimator for f_t when applied to $\Delta y_{i,t}$ under the assumption that γ_i are unknown. Any of the available estimators in the literature may be used, such as those proposed by Stock and Watson (2002), Kapetanios and Marcellino (2003) or Forni, Hallin, Lippi, and Reichlin (2004). This implicitly imposes conditions on the nature of the idiosyncratic shock $\epsilon_{i,t}$. These conditions vary depending on the factor estimation method used. Note that our setup differs from that in Bai and Ng (2004). In their setup the common and the idiosyncratic part of the series are allowed to have different orders of integration. We do not consider this possibility. Although we can amend our analysis to allow for this, we do not do so because we have a different focus. In particular, in common with a large part of the literature, we view the cross correlation mainly as a nuisance feature to be eliminated prior to carrying out panel unit root tests. It is clear that interest may focus on the characteristics of the common and idiosyncratic parts. Further, if the common part is I(1) but the idiosyncratic part is not then the outcome of the unit root analysis will be altered due to the removal of cross sectional correlation. We abstract from these issues for two reasons. Firstly, the study of the theoretical properties of the alternative factor extraction methods we consider are not as fully developped as that of principal components for nonstationary processes. Secondly, our Monte Carlo analysis will show that all methods do not perform very well even for the simple setup we consider, when stationary dynamics enter the factors. It seems likely that these results extend to more complicated setups.

We make the following assumptions.

Assumption 1 $\left\| \hat{f}_t - H f_t \right\| = O_p(T^{-\alpha}), \forall t,\ \alpha > 1/2$

Assumption 2 $\ni T_0$ *such that* $\forall T > T_0,\ 1/T f' f$ *and* $1/T \hat{f}' \hat{f}$ *are positive definite matrices.*

Assumption 3 $E \left\| f_t \right\|^r < \infty$, *for some* $r > 2$

Assumption 4 $\{\epsilon_{i,t}\}_{t=1}^T$ *are weakly dependent sequences, with zero means and heterogeneous variances* σ_i^2*, such that assumption 1 holds. These sequences are independent across* i*. Also* $E|\epsilon_{i,t}|^r < \infty$*, for some* $r > 2$

Assumption 5 *γ_i are uniformly bounded over N*

For a weaker theoretical result we also make the following alternative assumption to assumption 1.

Assumption 6 $\left\| \hat{f}_t - H f_t \right\| = o_p(1), \forall t$

The most remarkable of the above assumptions are assumptions 1 and 5. They are high level assumptions and require a considerably large set of more primitive assumptions to hold. These depend on the factor extraction method used. We adopt these assumptions for simplicity and refer the reader to the original papers that developed the estimation methods for a full list and discussion of the assumptions needed.

The following procedure which is very similar in spirit to the one proposed by Bai and Ng (2004) may be used for removing the cross sectional dependence in 1.

- Difference the data $y_{i,t}$ to obtain $\Delta y_{i,t}$
- Estimate f_t by any method that satisfies the neccesary assumptions. Define $\Delta y^*_{i,t} = \Delta y_{i,t} - \hat{f}'_t \hat{\gamma}_i$
- Cumulate $\Delta y^*_{i,t}$ to obtain $y^*_{i,t}$

Any panel unit root test for panel datasets with no cross sectional dependence may then be applied on $y^*_{i,t}$. We have the following theoretical results.

Theorem 1 *Define $\tilde{y}_{i,t} = \sum_{s=1}^{t} \epsilon_{i,s}$. Denote the individual Dickey-Fuller test statistic from $y^*_{i,t}$ by t^*_i and from $\tilde{y}_{i,t}$ by $\tilde{t}_i$. Under Assumptions 1-5, $t^*_i - \tilde{t}_i = o_p(1)$*

Theorem 2 *Under assumptions 2-6 , $\tilde{t}_i$ and $\tilde{t}_j$ are asymptotically uncorrelated for $j \neq i$. Further, under the null hypothesis, $\tilde{t}_i \Rightarrow \frac{\int W(r)dW(r)}{\left(\int W(r)^2 dr\right)^{1/2}}$ where $W(r)$ denotes standard Brownian motion.*

These theorems are proven in the appendix. Theorems 1 and 2 extedn the analysis of Bai and Ng (2004) on principal components to any factor estimation method that satisfies the neccesary assumptions. Theorem 1 provides a further modest extension to the results of Bai and Ng (2004) by showing that DF test statistics based on $y^*_{i,t}$ not only have the same distribution as those based on data with no cross sectional correlation but also converge to them in probability. Note that as long as $\sqrt{N} > T$ the pricipal component method satisfies the necessary assumptions for Thoerem 1 to hold. It is clear that Theorems 1 and 2 may be used to justify the construction of panel unit root tests based on $y^*_{i,t}$. Finally, we note that there may be modest interest of indepenent nature in the proof of Theorem 2 which uses a strong approximation result for Brownian motions previously utilised by Park (2002) and Kapetanios (2003) in the context of the bootstrap.

3 Factor Extraction Methods

We do not discuss the method of principal components as it is well known in the literature We briefly discuss the other two methods.

3.1 Dynamic Principal Components

Frequency domain analysis of the dynamic factor model was recently proposed by Forni and Reichlin (1996, 1998), Forni, Hallin, Lippi, and Reichlin (2000, 2004) (FHLR henceforth). The model they adopt is

$$x_{it} = b_i^{'}(L)u_t + \xi_{it}, \quad i \in N, t \in Z$$

where x_{it} is a stationary univariate random variable, u_t is a $q \times 1$ vector of common factors, $\chi_{it} = x_{it} - \xi_{it}$ is the common component of x_{it}, and ξ_{it} is its idiosyncratic component. More precisely, u_t is an orthonormal white noise process, so that $var(u_{jt}) = 1$, $cov(u_t, u_{t-k}) = 0$, and $cov(u_{jt}, u_{st-k}) = 0$ for any $j \neq s$, t and k. $\xi_n = \{\xi_{1t}, ..., \xi_{nt}\}^{'}$ is a wide sense stationary process for any n, and $cov(\xi_{jt}, u_{st-k}) = 0$ for any j, s, t and k. $b_i(L)$ is a $q \times 1$ vector of square summable, bilateral filters, for any i. Hence, $x_{nt} = \{x_{1t}, ..., x_{nt}\}^{'}$ is also a stationary vector process. FHLR also require χ_{nt}, ξ_{nt}, and therefore x_{nt}, to have rational spectral density matrices, Σ_n^{χ}, Σ_n^{ξ}, and Σ_n^{x}, respectively. To achieve (asymptotic) identification, they assume that the first (largest) idiosyncratic dynamic eigenvalue, λ_{n1}^{ξ}, is uniformly bounded, and that the first (largest) q common dynamic eigenvalues, $\lambda^{\chi}_{n1}, ..., \lambda^{\chi}_{nq}$, diverge, where dynamic eigenvalues are the eigenvalues of the spectral density matrix, see e.g. Brillinger (1981). In words, the former condition limits the effects of ξ_{it} on other cross-sectional units. The latter, instead, requires u_t to affect infinitely many units. The static version of this model was analyzed, among others, by Chamberlain and Rothschild (1983), and Connor and Korajczyk (1986, 1993). When the idiosyncratic components are uncorrelated across units the model is usually referred to as an exact static model, otherwise it is approximate. Sargent and Sims (1977) studied a dynamic factor model for a limited number of units. Further developments were due to , Stock and Watson (1991) and Quah and Sargent (1993), but all these methods are not suited when n, the number of variables, is larger than 50-60, while the procedure by FHLR can handle hundreds of variables. We will refer to the procedure by FHLR as dynamic principle component analysis (DPCA). Note that the standard code provided by FHLR provides estimates of the common component rather than the serially uncorrelated factor u_t.

As our methods are crucially dependent on the use of factors, we note the following. It is clear that the common component is spanned by $u_{t+s}, \ldots, u_t, \ldots u_{t-p}$, where s is the maximum lead and p is the maximum lag of $b_i(L)$ over i. This is reflected in the fact that the $T \times N$ matrix $\hat{\chi}$, denoting the estimated common component is of rank $s+p+1$. Hence, we define the factor estimate from this method to be the $s+p+1$ left singular vectors associated with the largest $s+p+1$ singular values from the singular value decomposition of $\hat{\chi}$. We note that by Forni, Hallin, Lippi, and Reichlin (2004) the rate of convergence of the common component (and hence the factor estimate) is at best $T^{1/4}$. Hence, only Theorem 2 is of relevance for this method.

3.2 Subspace Factor Estimation

A method of estimating factors from large datasets using a state pace representation of the data has been suggested by Kapetanios and Marcellino (2003). We give details of this method below.

3.2.1 The Basic State Space Model

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

$$\begin{aligned} x_{Nt} &= Cf_t + D^* u_t, \quad t = 1, \ldots, T \\ f_t &= Af_{t-1} + B^* v_{t-1}, \end{aligned} \tag{4}$$

where x_{Nt} is an N-dimensional vector of stationary zero-mean variables observed at time t, f_t is a k-dimensional vector of unobserved states (factors) at time t, and u_t and v_t are multivariate, mutually uncorrelated, standard orthogonal white noise sequences of dimension, respectively, N and k. D^* is assumed to be nonsingular. The aim of the analysis is to obtain estimates of the states f_t, for $t = 1, \ldots, T$. Notice that the factors are driven by lagged errors. This is an important hypothesis for the methodology. This hypothesis is not considered restrictive in the state space model literature, see e.g. Hannan and Diestler (1988).

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

$$\begin{aligned} x_{Nt} &= C_1 f_t + C_2 f_{t-1} + D^* u_t, \quad t = 1, \ldots, T \\ f_t &= Af_{t-1} + B^* v_{t-1}, \end{aligned} \tag{5}$$

This model can be written as

$$\begin{aligned} x_{Nt} &= (C_1, C_2)\tilde{f}_t + D^* u_t, \quad t = 1, \ldots, T \\ \tilde{f}_t &= \begin{pmatrix} f_t \\ f_{t-1} \end{pmatrix} = \begin{pmatrix} A & 0 \\ I & 0 \end{pmatrix} \begin{pmatrix} f_{t-1} \\ f_{t-2} \end{pmatrix} + \begin{pmatrix} B^* & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_{t-1} \\ 0 \end{pmatrix}, \end{aligned} \tag{6}$$

which is a special case of the specification in (4), even though by not taking into account the particular structure of the A matrix and the reduced rank of the error process we are losing in terms of efficiency.

A large literature exists on the identification issues related with the state space representation given in (4). An extensive discussion may be found in

Hannan and Diestler (1988). In particular, they show in Chapter 1 that (4) is equivalent to the prediction error representation of the state space model given by

$$\begin{aligned} x_{Nt} &= Cf_t + Du_t, \quad t = 1, \ldots, T \\ f_t &= Af_{t-1} + Bu_{t-1}. \end{aligned} \tag{7}$$

This form will be used for the derivation of our estimation algorithm. Note that as at this stage the number of series, N, is large but fixed we need to impose no conditions on the structure of C. Conditions on this matrix will be discussed later when we consider the case of N tending to infinity and possibly correlated idiosyncratic errors.

3.2.2 Subspace Estimators

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

To address this issue, Kapetanios and Marcellino (2003) exploit subspace algorithms, which avoid expensive iterative techniques by relying on matrix algebraic methods, and can be used to provide estimates for the factors as well as the parameters of the state space representation.

There are many subspace algorithms, and vary in many respects, but a unifying characteristic is their view of the state as the interface between the past and the future in the sense that the best linear prediction of the future of the observed series is a linear function of the state. A review of existing subspace algorithms is given by Bauer (1998) in an econometric context. Another review with an engineering perspective may be found in Overshee and Moor (1996). An economic application of the algorithm may be found in Kapetanios (2004).

The starting point of most subspace algorithms is the following representation of the system which follows from the state space representation in (7) and the assumed nonsingularity of D.

$$X_t^f = \mathcal{O}\mathcal{K}X_t^p + \mathcal{E}E_t^f \tag{8}$$

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

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

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

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

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

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

It is important to note that the choice of the weighting matrices is not crucial for the asymptotic properties of the estimation method. This is because the choice does not affect neither the consistency nor the rate of convergence of the factor estimates. They are only required to be nonsingular. So an alternative possibility is to simply use identity matrices instead of the covariance matrices. It is this possibility we follow in the Monte Carlo study.

A second point to note is that consistent estimation of the factor space requires that p increases at a rate greater than $ln(T)^\alpha$, for some $\alpha > 1$ that depends on the maximum eigenvalue of A, but at a rate lower than $T^{1/3}$. A simplified condition for the lower bound would be to set it to $T^{1/r}$ for any $r > 3$. For consistency s is also required to be set so as to satisfy $sN \geq k$. As N is usually going to be very large for the applications we have in mind, this restriction is not binding.

3.2.3 Dealing with Large Datasets

Up to now we have outlined a method for estimating factors which requires the number of observations to be larger than the number of elements in X_t^p. Given the work on principal components and FHLR this is rather restrictive. Kapetanios and Marcellino (2003) therefore suggest a modification of the methodology to allow the number of series be larger than the number of observations.

The problem arises because the least squares estimate of $\mathcal{F}$ is not uniquely defined due to rank deficiency of $X^{p'}X^p$. We do not neccesarily want an estimate of $\mathcal{F}$ but an estimate of the states $X^p\mathcal{K}'$. That could be obtained if we had an estimate of $X^p\mathcal{F}'$ and used an SVD of that. But it is well known (see e.g. Magnus and Neudecker (1988)) that although $\hat{\mathcal{F}}$ may not be estimable, $X^p\mathcal{F}'$ always is using least squares methods. In particular, the least squares estimate of $\widehat{X^p\mathcal{F}'}$ is given by

$$\widehat{X^p\mathcal{F}'} = X^p(X^{p'}X^p)^+X^{p'}X^f \tag{9}$$

where $X^f = (X_1^f, \ldots, X_T^f)$ and A^+ denotes the unique Moore-Penrose inverse of matrix A. However, when the row dimension of X^p is smaller than its column dimension, $X^p(X^{p'}X^p)^+X^{p'} = I$ implying that $\widehat{X^p\mathcal{F}'} = X^f$. A decomposition of X^f is then easily seen to be similar, but not identical, to the eigenvalue decomposition of the covariance matrix of X^f which is the SW principle component method. We will refer to this method as SSS0. This method is static, abstracting from the fact that s may be larger than 1, thereby leading to a decomposition involving leads of x_{Nt}.

Alternative solutions exist to this problem. In particular, note that we are after a subspace decomposition of the estimate of the fitted value $X^p\mathcal{F}'$. Essentially, we are after a reduced rank approximation of $X^p\mathcal{F}'$ and several possibilities exist. The main requirement is that as the assumed rank (number of factors) tends to the full rank of the estimate of the fitted value, the approximation should tend to the estimated fitted value

$\widehat{X^p\mathcal{F}'} = X^p(X^{p'}X^p)^+X^{p'}X^f = X^f$. The alternative decomposition we suggest is a SVD on $X^{f'}X^p(X^{p'}X^p)^+ = \hat{U}\hat{S}\hat{V}'$. Then the estimated factors are given by $\hat{\mathcal{K}}X_t^p$ where $\hat{\mathcal{K}}$ is obtained as before but using the SVD of $X^{f'}X^p(X^{p'}X^p)^+$. We choose to set both weighting matrices to the identity matrix in this case. We also refer to this decomposition as SSS, because it is simply a generalisation of the original method and if $Np < T$ it reduces to that method. As k tends to $\min(Ns, Np)$ the set of factor estimates tends to the OLS estimated fitted value X^f. This method needs to be judged in terms of its small sample properties in approximating (linear combinations of) the true factors, and the simulations in the next section indicate it performs satisfactorily in the current context.

We conclude this section by noting the asymptotic properties of the factor estimates as discussed in Kapetanios and Marcellino (2003). The factor estimates are consistent and asymptotically normal with rates of convergence at most $T^{1/2-\delta}$, $\delta > 0$ and at least $T^{1/3}$. Hence, once again only Theorem 2 is applicable for this method.

4 Monte Carlo Study

4.1 Monte Carlo Setup

One of the main motivations of this paper is to examine whether the factor based methods for removing cross-sectional dependence do what they are designed to do under general setups. In particular, the focus of our work is on the size of the Im, Pesaran, and Shin (2003) test when prior removal of cross-sectional dependence has taken place using any of three methods of factor estimation. The theoretical results we have presented suggest that since all three methods consistently estimate the factors, they should be able to correct for cross-sectional dependence.

We consider the following setup.

$$y_{i,t} = \phi_i y_{i,t-1} + \eta_{i,t}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T \tag{10}$$

$$\eta_{i,t} = f_t'\gamma_i + \epsilon_{i,t}$$

$$f_{j,t} = \rho_j f_{j,t-1} + \varepsilon_{j,t}$$

where $E(\varepsilon_{j,t}^2) = \sigma_\varepsilon^2$. Let m denote the number of factors. Then we consider: $N \in \{5, 10, 20, 30, 50\}$, $T \in \{30, 50, 100, 200, 400\}$, $m \in \{1, 2\}$, $\rho_j \in$

$\{0.95, 0.7, 0.1\}$ and $\sigma_\varepsilon^2 \in \{1, 3\}$. Perhaps the most crucial dimension is the persistence in the factor process. Note that all these experiments are size experiments as long as $\phi_i = 1$. A note is in order for the SSS method. The theory is developped for predetermined factors, i.e. factors that are determined at time $t-1$. This is reflected by (4) where the error term of the factor equation is dated at time $t-1$. This assumption is not considered restrictive in the state space model literature. Yet, the specification we use for the simulations allows for factors that are determined at time t. This brings us in line with the nonparametric context of SW and FHLR. The simulations show that, even for this setup which is not captured by SSS, the method works well.

Although we have not discussed the treatment of constants and trends in the theoretical part of the paper, it is imperative that the Monte Carlo analysis is of empirical relevance. Hence, for the Monte Carlo study we accomodate the presence of constants and trends using the standard analysis of the Dickey-Fuller test as follows. We do not consider the version of the Dickey Fuller test with neither constant nor trend. For the case of constant only we note that in the Dickey-Fuller setup the value of the constant is equal to zero under the null hypothesis. Hence, we construct $y_{i,t}^*$ as described in Section 2. Then, we apply to $y_{i,t}^*$ the constant-only version of the Dickey-Fuller test. For the case of constant and trend we demean the differenced data and store the vector of means. Then, we extract the factor as described in Section 2. We then have the choice of either cumulating with or without adding the vector of means of the differenced data to construct $y_{i,t}^*$. Both methods are asymptotically equivalent. We choose not to add the mean to be closer to the treatment of Bai and Ng (2003). We then apply the constant and trend version of the Dickey-Fuller test to $y_{i,t}^*$. Once the Dickey-Fuller statistics have been constructed we apply the Im, Pesaran, and Shin (2003) test.

The errors $\epsilon_{i,t}$ and $\varepsilon_{i,t}$ are standard normal random variables. $Var(\epsilon_{i,t}) = 1$. $\gamma_i \sim N(0, 1)$. Initial conditions are set to zero throughout. 1000 replications for each experiment have been carried out. The nominal significance level is set to 95%.

4.2 Monte Carlo Results

All results are presented in Tables 1 to 8. Tests 0,1,2 and 3 are respectively the unadjusted test and the tests based on the SSS method, principal components and DPCA respectively. The Tables are revealing. Whereas for all cases which are close to those previously considered in the literature (i.e. $\rho_j = 0.1$) the methods using factor estimates to remove cross correlation work better than the unadjusted test, this is not necessarily the case for cases with more persistent factors.

Both the constant-only and the constant and trend versions of the DF test produce similar results. Test 0 performs badly as expected as it overrejects substantially and the problem gets worse as either N or T rises. Interestingly, the overrejection is more pronounced for $\rho_j = 0.95, 0.1$ than for the middle value $\rho_i = 0.7$. Things also get worse as σ_ε^2 rises. Tests 1, 2 and 3 seem to have relatively similar behaviour. Test 1 seems to be better behaved than 2 and 3 for low N.

The remarkable results concern the relative performance of the tests with Test 0 and also along the following two dimensions: persistence and number of factors. For low persistence the factor extraction helps and provides well behaved tests. All factor extraction methods achieve similar results. As the persistence is increased we see that Tests 1,2 and 3 start performing much worse. While their performance improves for larger T, it is clear that in terms of rejection probabiities under the null hypothesis one is better off not correcting for cross sectional dependence when the factors are highly serially correlated. Test 0 is much better bahaved than the tests based on factor extraction. Nevertheless, its performance worsens with T indicating that asymptotically this test is inappropriate. However, even for $T = 400$ Test 0 dominates the other tests. Similar but less dramatic results are obtained for lower persistence.

Interesting results are obtained when multiple factors are considered. There, the performance of the tests seems to be much more uniform across different persistence structures. Again, the clear superiority of factor based methods that correct for cross sectional dependence over the standard test seems to be in doubt. In summary, factor based methods work well only for low levels of factor persistence and one factor. Departure from this possibly unrealistic setup may lead to less accurate inference compared to the standard panel unit root tests. Of course, asymptotically, correcting for cross sectional dependence is essential. It seems though that in small samples the case for correcting for cross sectional dependence, at least using factor based methods, is less convincing.

5 Conclusion

Following the large literature on panel unit root tests, the need to correct for the presence of cross sectional dependence is clear both from theoretical and Monte Carlo results. A number of methods have been suggested in the literature. A significant part of the discussion focuses on removing factors from the data and carrying panel unit root tests on the 'residuals' from this factor extraction.

This paper aims to fill two gaps. The first is the consideration of factor extraction techniques alternative to principal components for removing factors. A theoretical analysis of general factor estimation methods is provided. Details on two alternative estimation methods are givrn. These are the Kapetanios and Marcellino (2003) and Forni, Hallin, Lippi, and Reichlin (2004) factor estimation methods. Secondly, little work has been done on whether the dynamic nature and number of factors are of relevance to the performance of the factor extraction methods. We provide a detailed Monte Carlo study of these issues. Unlike the case of serially uncorrelated factor, persistent factors cannot be extracted easily and the resulting panel unit root seems to suffer because of that. In a number of cases not correcting for cross sectional dependence seems to be preferable to the existing factor based correction methods. The number of factors is of importance as well. Multiple factors are less easily extracted. In summary, more work is needed if these methods are to dominate in all cases the performance of simple panel unit root tests.

6 Appendix

6.1 Proof of Theorem 1

Assumption 1 implies that

$$1/T \sum_t \left\| \hat{f}_t - H f_t \right\|^2 = O_p(T^{-2\alpha}), \tag{11}$$

Then,

$$\Delta y^*_{i.t} = \Delta y_{i.t} - \hat{f}'_t \hat{\gamma}_i = \epsilon_{i,t} + f'_t \gamma_i - \hat{f}'_t \hat{\gamma}_i = \epsilon_{i,t} + f'_t H' H^{-1'} \gamma_i - \hat{f}'_t \hat{\gamma}_i = \tag{12}$$

$$\epsilon_{i,t} + (f'_t H' - \hat{f}'_t) H^{-1'} \gamma_i - \hat{f}'_t (\hat{\gamma}_i - H^{-1'} \gamma_i)$$

where $\hat{\gamma}_i = (\hat{f}'\hat{f})^{-1} \hat{f}' \Delta y_i$ and $y^*_{i.t} = \sum_{i=1}^{t} \Delta y^*_{i.t}$
We wish to show that

$$t^*_i - \tilde{t}_i = o_p(1) \tag{13}$$

First, we show that

$$\left\| \hat{\gamma}_i - H^{-1'} \gamma_i \right\| = O_p(T^{-\alpha}) \tag{14}$$

We first show that

$$\left\| (1/T \hat{f}' \hat{f})^{-1} - (1/T H' f' f H)^{-1} \right\| = O_p(T^{-\alpha}) \tag{15}$$

To do that we have

$$\left\| (1/T \hat{f}' \hat{f})^{-1} - (1/T H' f' f H)^{-1} \right\| \leq \left\| (1/T \hat{f}' \hat{f})^{-1} \right\| \tag{16}$$

$$\left\| \Big((1/T H' f' f H) - (1/.T \hat{f}' \hat{f}) \Big) \right\| \left\| (1/T H' f' f H)^{-1} \right\|$$

By assumption 2 on the positive definiteness of $(1/.T\hat{f}'\hat{f})^{-1}$ and $(1/TH'f'fH)^{-1} \forall T > T_0$, we only need to show that the middle term of the RHS of (16) is $O_p(T^{-\alpha})$.
Thus we need to show that

$$\left\| 1/T \sum_{t=1}^{T} \hat{f}_t \hat{f}'_t - 1/T \sum_{t=1}^{T} H f_t f'_t H' \right\| = O_p(T^{-\alpha}) \tag{17}$$

But

$$\left\| 1/T \sum_{t=1}^{T} \hat{f}_t \hat{f}'_t - 1/T \sum_{t=1}^{T} H f_t f'_t H' \right\| \leq \left\| 1/T \sum_{t=1}^{T} \hat{f}_t \hat{f}'_t - 1/T \sum_{t=1}^{T} H f_t \hat{f}'_t \right\| +$$

$$\left\| 1/T \sum_{t=1}^{T} H f_t \hat{f}_t' - 1/T \sum_{t=1}^{T} H f_t f_t' H' \right\|$$

It then suffices to show that the following is $O_p(T^{-\alpha})$

$$1/T \sum_{t=1}^{T} H f_t \hat{f}_t' - 1/T \sum_{t=1}^{T} H f_t f_t' H' = 1/T \sum_{t=1}^{T} H f_t (\hat{f}_t' - f_t' H') \leq \tag{18}$$

$$\left(1/T \sum_{t=1}^{T} ||H f_t||^2 \right)^{1/2} \left(1/T \sum_{t=1}^{T} ||\hat{f}_t' - f_t' H'||^2 \right)^{1/2}$$

But by Assumption 3 $\left(1/T \sum_{t=1}^{T} ||H f_t||^2 \right)^{1/2}$ is $O_p(1)$ and hence the result holds.

We can similarly show that

$$\left\| 1/T \hat{f}' \Delta y_i - 1/T f' H' \Delta y_i \right\| = O_p(T^{-\alpha}) \tag{19}$$

We have to show that

$$\left\| 1/T \sum_{t=1}^{T} \hat{f}_t \Delta y_t' - 1/T \sum_{t=1}^{T} H f_t \Delta y_t' \right\| = O_p(T^{-\alpha}) \tag{20}$$

But

$$1/T \sum_{t=1}^{T} \hat{f}_t \Delta y_t' - 1/T \sum_{t=1}^{T} H f_t \Delta y_t' = 1/T \sum_{t=1}^{T} (\hat{f}_t - H f_t) \Delta y_t' \leq \tag{21}$$

$$\left(1/T \sum_{t=1}^{T} \left\| \hat{f}_t - H f_t \right\|^2 \right)^{1/2} \left(1/T \sum_{t=1}^{T} \|\Delta y_t'\|^2 \right)^{1/2}$$

which follows readily from Assumption 1

The above imply that $1/t \sum_{j=1}^{t} |\Delta \tilde{y}_{i,j} - \Delta y_{i,j}^*|^2 = O_p(t^{-2\alpha})$ and then we have that

$$1/T^2 \sum_t |\tilde{y}_{i,t} - y^*_{i,t}|^2 \leq 1/T^2 \sum_t \left| \sum_{i=1}^{t} (\Delta \tilde{y}_{i,j} - \Delta y^*_{i,j}) \right|^2 \leq \tag{22}$$

$$1/T^2 \sum_{t=1}^{T} t^2 (1/t) \sum_{i=1}^{t} (\Delta \tilde{y}_{i,j} - \Delta y^*_{i,j})^2 = o_p(1)$$

Next we wish to show that

$$T(\hat{\alpha}_i - \tilde{\alpha}_i) = o_p(1) \tag{23}$$

where $\hat{\alpha}_i = (y_i^{*\prime} y_i^*)^{-1} y_i^{*\prime} \Delta y_i^*$ and $\tilde{\alpha}_i = (\tilde{y}_i' \tilde{y}_i)^{-1} \tilde{y}_i' \Delta \tilde{y}_i$ which would imply $t_i^* - t_i = o_p(1)$.

If we show that

$$\frac{1}{T^2} \sum_{i=1}^{T} y^{*2}_{i,t} - \frac{1}{T^2} \sum_{i=1}^{T} \tilde{y}^2_{i,t} = o_p(1) \tag{24}$$

and

$$\frac{1}{T} \sum_{i=1}^{T} y^*_{i,t} \Delta y^*_{i,t} - \frac{1}{T} \sum_{i=1}^{T} \tilde{y}_{i,t} \Delta \tilde{y}_{i,t} = o_p(1) \tag{25}$$

we get the result.

We have that

$$\frac{1}{T^2} \sum_{i=1}^{T} y_t^{*2} - \frac{1}{T^2} \sum_{i=1}^{T} \tilde{y}_t^2 = \left(\frac{1}{T^2} \sum_{i=1}^{T} y_t^{*2} - \frac{1}{T^2} \sum_{i=1}^{T} y_t^* \tilde{y}_t \right) + \left(\frac{1}{T^2} \sum_{i=1}^{T} y_t^* \tilde{y}_t - \frac{1}{T^2} \sum_{i=1}^{T} \tilde{y}_t^2 \right) \tag{26}$$

Concentrating on the first term of the RHS of (26) we have

$$\frac{1}{T^2} \sum_{i=1}^{T} y_t^{*2} - \frac{1}{T^2} \sum_{i=1}^{T} y_t^* \tilde{y}_t = \frac{1}{T^2} \sum_{i=1}^{T} y_t^* (y_t^* - \tilde{y}_t) \leq \tag{27}$$

$$\left(1/T^2 \sum_{t=1}^{T} |y_t^*|^2 \right)^{1/2} \left(1/T^2 \sum_{t=1}^{T} |y_t^* - \tilde{y}_t|^2 \right)^{1/2}$$

But $\left(1/T^2 \sum_{t=1}^{T} |y_t^*|^2\right)^{1/2} = O_p(1)$ by assumption 4 and $\left(1/T^2 \sum_{t=1}^{T} |y_t^* - \tilde{y}_t|^2\right)^{1/2} = o_p(1)$ by (22). Hence, the whole term is $o_p(1)$. Similarly for the second term of the RHS of (26). By similar arguments we get that (25) holds. More specifically we get that

$$\frac{1}{T}\sum_{i=1}^{T} y_{i,t}^* \Delta y_{i,t}^* - \frac{1}{T}\sum_{i=1}^{T} \tilde{y}_{i,t} \Delta \tilde{y}_{i,t} = \left(\frac{1}{T}\sum_{i=1}^{T} y_{i,t}^* \Delta y_{i,t}^* - \frac{1}{T}\sum_{i=1}^{T} y_{i,t}^* \Delta \tilde{y}_{i,t}\right) + \quad (28)$$
$$\left(\frac{1}{T}\sum_{i=1}^{T} y_{i,t}^* \Delta \tilde{y}_{i,t} - \frac{1}{T}\sum_{i=1}^{T} \tilde{y}_{i,t} \Delta \tilde{y}_{i,t}\right)$$

Then

$$\frac{1}{T}\sum_{i=1}^{T} y_{i,t}^* \Delta y_{i,t}^* - \frac{1}{T}\sum_{i=1}^{T} y_{i,t}^* \Delta \tilde{y}_{i,t} = \frac{1}{T}\sum_{i=1}^{T} y_t^* (\Delta y_{i,t}^* - \Delta \tilde{y}_{i,t}) \leq \quad (29)$$
$$\left(1/T \sum_{t=1}^{T} |y_t^*|^2\right)^{1/2} \left(1/T \sum_{t=1}^{T} |\Delta y_{i,t}^* - \Delta \tilde{y}_{i,t}|^2\right)^{1/2}$$

But $\left(1/T \sum_{t=1}^{T} |y_t^*|^2\right)^{1/2} = O_p(T^{1/2})$, $\left(1/T \sum_{t=1}^{T} |\Delta y_{i,t}^* - \Delta \tilde{y}_{i,t}|^2\right)^{1/2} = O_p(T^{-\alpha})$ and so the required result holds.

6.2 Proof of Theorem 2

We now wish to prove the weaker result that for any consistent factor estimate i.e.

$$\left\| \hat{f}_t - H f_t \right\| = o_p(1), \forall t \quad (30)$$

then

$$W_T^*(r) \equiv \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \Delta y_{i.t}^* \Rightarrow W(r) \quad (31)$$

We will also show that

$$\frac{1}{T}\sum_{t=1}^{T}\Delta y^*_{i.t}\Delta y^*_{j.t} = o_p(1) \tag{32}$$

These two facts would imply that unit root tests on the constructed dataset would have the same distribution as if no cross correlation arose and would prove Theorem 2. Following the proof of Theorem 1 we can easily get that

$$\left\|\hat{\gamma}_i - H^{-1'}\gamma_i\right\| = o_p(1) \tag{33}$$

We start by proving (32) $\forall i \neq j$

$$\frac{1}{T}\sum_{t=1}^{T}\Delta y^*_{i.t}\Delta y^*_{j.t} = \frac{1}{T}\sum_{t=1}^{T}\left[\epsilon_{i,t} + (f_t'H' - \hat{f}_t')H^{-1'}\gamma_i - \hat{f}_t'(\hat{\gamma}_i - H^{-1'}\gamma_i)\right] \tag{34}$$

$$\left[\epsilon_{j,t} + (f_t'H' - \hat{f}_t')H^{-1'}\gamma_j - \hat{f}_t'(\hat{\gamma}_j - H^{-1'}\gamma_j)\right] =$$

$$\frac{1}{T}\sum_{t=1}^{T}(f_t'H'-\hat{f}_t')H^{-1'}\gamma_i(f_t'H'-\hat{f}_t')H^{-1'}\gamma_j - \frac{2}{T}\sum_{t=1}^{T}(f_t'H'-\hat{f}_t')H^{-1'}\gamma_i\hat{f}_t'(\hat{\gamma}_j-H^{-1'}\gamma_j)$$

$$\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t'(\hat{\gamma}_i-H^{-1'}\gamma_i)\hat{f}_t'(\hat{\gamma}_j-H^{-1'}\gamma_j)+\frac{1}{T}\sum_{t=1}^{T}\epsilon_{i,t}(f_t'H'-\hat{f}_t')H^{-1'}\gamma_j-\frac{1}{T}\sum_{t=1}^{T}\epsilon_{i,t}\hat{f}_t'(\hat{\gamma}_j-H^{-1'}\gamma_j)+$$

$$\frac{1}{T}\sum_{t=1}^{T}\epsilon_{j,t}(f_t'H' - \hat{f}_t')H^{-1'}\gamma_i - \frac{1}{T}\sum_{t=1}^{T}\epsilon_{j,t}\hat{f}_t'(\hat{\gamma}_i - H^{-1'}\gamma_i) + \frac{1}{T}\sum_{t=1}^{T}\epsilon_{i,t}\epsilon_{j,t}$$

All terms of the above apart from the last can easily shown to be $o_p(1)$ by Liapunov's inequality and (30), (33). The last is $o_p(1)$ by assumption. Hence the resut is proven. Moving on to (31), we will show that $E|\Delta y^*_{i.t}|^r < \infty$, for some $r > 2$. This, by Sakhanenko (1980) and Park (2002) implies that

$$P\left(\sup_r |W^*_T(r) - W(r)|\right) = o(1) \tag{35}$$

which implies (31). By assumption 4 $E|\epsilon_{i,t}|^r < \infty$, for some $r > 2$. Then we need to show that

$$\frac{1}{T}\sum_{t=1}^{T}|\Delta y^*_{i.t}|^r < \infty$$

We have

$$\frac{1}{T}\sum_{t=1}^{T}|\Delta y^*_{i.t}|^r = \frac{1}{T}\sum_{t=1}^{T}\left|\epsilon_{i,t} + (f'_tH' - \hat{f}'_t)H^{-1'}\hat{\gamma}_i - f'_t(\hat{\gamma}_i - H^{-1'}\gamma_i)\right|^r \leq$$

$$\frac{1}{T}\sum_{t=1}^{T}|\epsilon_{i,t}|^r + \frac{1}{T}\sum_{t=1}^{T}\left|(f'_tH' - \hat{f}'_t)H^{-1'}\hat{\gamma}_i\right|^r + \frac{1}{T}\sum_{t=1}^{T}\left|f'_t(\hat{\gamma}_i - H^{-1'}\gamma_i)\right|^r$$

By the Law of Large Numbers we have

$$\frac{1}{T}\sum_{t=1}^{T}|\epsilon_{i,t}|^r \xrightarrow{p.} E|\epsilon_{i,t}|^r$$

Also

$$\frac{1}{T}\sum_{t=1}^{T}\left|f'_t(\hat{\gamma}_i - H^{-1'}\gamma_i)\right|^r \leq |\hat{\gamma}_i - H^{-1'}\gamma_i|^r\frac{1}{T}\sum_{t=1}^{T}|f'_t|^r = o_p(1)$$

by (33), the Law of Large Numbers and Assumption 3. Finally,

$$\frac{1}{T}\sum_{t=1}^{T}\left|(f'_tH' - \hat{f}'_t)H^{-1'}\hat{\gamma}_i\right|^r \leq |\hat{\gamma}_i|^r\,\frac{1}{T}\sum_{t=1}^{T}\left|(f'_tH' - \hat{f}'_t)\right|^r = o_p(1)$$

by (30), boundeness of $H^{-1'}\gamma_i$ and (33).

References

Bai, J., and S. Ng (2004): "A PANIC Attack on Unit Roots and Cointegration," *Econometrica*, Forthcoming.

Baltagi, B. H., and C. Kao (2000): "Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey," *Advances in Econometrics*, 15, 7–51.

Bauer, D. (1998): "Some Asymptotic Theory for the Estimation of Linear Systems Using Maximum Likelihood Methods or Subspace Algorithms," Institute f. Econometrics, Operation Research and System Theory, TU Wien, Austria, PhD Thesis.

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

Chamberlain, G., and M. Rothschild (1983): "Arbitrage Factor Structure and and Mean Variance Analysis of Large Asset Markets," *Econometrica*, 51, 1281–1304.

Chang, Y. (2000): "Testing for Stationarity in Heterogeneous Panel Data," *Econometrics Journal*, 3, 148–161.

——— (2002): "Nonlinear IV Unit Root Tests in Panels with Cross Sectional Dependence," *Journal of Econometrics*, 110, 261–292.

Choi, I. (2001): "Unit Root Tests For Panel Data," *Journal of Internation Money and Banking*, 20, 249–272.

Connor, G., and R. A. Korajczyk (1986): "Performance Measurement with the Arbitrage Pricing Theory," *Journal of Financial Economics*, 15, 373–394.

——— (1993): "A Test for the Number Factors in an Approximate Factor Model," *Journal of Finance*, 48, 1263–1291.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000): "The Generalised Factor Model: Identification and Estimation," *Review of Economics and Statistics*, 82, 540–554.

——— (2004): "The Generalised Factor Model: Consistency and Rates," *Journal of Econometrics*, Forthcoming.

Forni, M., and L. Reichlin (1996): "Dynamic Common Factors in Large Cross Sections," *Empirical Economics*, 21, 27–42.

——— (1998): "Let's Get Real: A Dynamic Factor Analytical Approach to the Disaggregated Business Cycle," *Review of Economic Studies*, 65, 453–474.

Hannan, E., and M. Diestler (1988): *The Statistical Theory of Linear Systems*. Wiley.

Harvey, A. C., and D. Bates (2002): "Multivariate Unit Root Tests and Testing for Convergence," *University of Cambridge DAE Working Paper No. 0301.*

Im, K. S., M. H. Pesaran, and Y. Shin (2003): "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics*, 115(1), 53–74.

Kapetanios, G. (2003): "A Bootstrap Invariance Principle for Highly Nonstationary Long Memory Processes," *Queen Mary, University of London Mimeo.*

——— (2004): "A note on modelling core inflation for the UK using a new dynamic factor estimation method and a large disaggregated price index dataset," *Economics Letters*, Forthcoming.

Kapetanios, G., and M. Marcellino (2003): "A Comparison of Estimation Methods for Dynamic Factor Models of Large Dimensions," *Queen Mary, University of London Working Paper No. 489.*

Levin, A., C. F. Lin, and C. S. Lu (2002): "Bootstrap Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," *Journal of Econometrics*, 108, 1–24.

Maddal, G. S., and S. Wu (1999): "A Comparative Study of Unit Root Tests in Panel Data and a New Simple Test," *Oxford Bulettin of Economics and Statistics*, 61, 631–652.

Magnus, J., and H. Neudecker (1988): *Matrix Differential Calculus with Applications to Statistics and Econometrics*. Wiley.

MOON, H. R., AND B. PERRON (2004): "Testing for a Unit Root in Panels with Dynamic Factors," *Journal of Econometrics*, Forthcoming.

OVERSHEE, P. V., AND B. D. MOOR (1996): *Subspace Identification for Linear Systems*. Kluwer.

PARK, J. Y. (2002): "An Invariance Principle for Sieve Bootstrap in Time Series," *Econometric Theory*, 18, 469–490.

PESARAN, M. H. (2003): "A Simple Panel Unit Root Test in the Presence of Cross section Dependence," *Cambridge University DAE Working Paper, number 0346*.

PHILLIPS, P. C. B., AND D. SUL (2002): "Dynamic Panel Estimation and Homogeneity Testing Under Cross Sectional Dependence," *Cowles Foundation Discussion Paper No. 1362*.

QUAH, D., AND T. J. SARGENT (1993): "A Dynamic Index Model for Large Cross-Sections," in *Business Cycles, Indicators and Forecasting*, ed. by J. H. Stock, and M. W. Watson. University of Chicago Press for the NBER.

SAKHANENKO, A. (1980): "On Unimprovable Estimates of the Rate of Convergence in Invariance Principle," in *Nonparametric Statistical Inference*, no. 32 in Colloquia Mathematica Societatis Janos Bolyai, pp. 779–783.

SARGENT, T. J., AND C. A. SIMS (1977): "Business Cycle Modelling Without Pretending to Have Too Much Economic Theory," in *New Methods in Business Cycle Research*, ed. by C. A. Sims. Federal Researve Bank of Minneapolis.

STOCK, J. H., AND M. W. WATSON (1991): "A Probability Model of the Coincident Economic Indicators," in *Leading Economic Indicators: New Aproaches and Forecasting Records*, ed. by K. Lahiri, and G. H. Moore. Cambridge University Press.

——— (2002): "Macroeconomic Forecasting Using difucions Indices," *Journal of Business and Economic Statistics*, 20, 147–162.

Table 1: Test 0, DF 2																	
m	σ_ε^2		$\rho_j = 0.95$					$\rho_j = 0.7$					$\rho_j = 0.1$				
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	6.6	10.8	16.9	19.2	21.3	5.5	6.7	6.8	6.0	5.8	6.1	4.8	6.7	6.1	5.8
	1	10	8.8	11.2	19.0	24.8	22.3	8.4	9.6	8.6	7.9	8.3	9.5	9.9	8.7	11.4	6.1
		20	9.3	14.7	20.7	22.9	26.4	8.4	10.7	10.3	10.7	9.3	12.2	12.0	12.3	12.5	12.3
		30	10.4	15.0	22.7	26.3	25.5	12.5	10.1	15.4	13.9	12.7	14.0	14.2	15.3	15.3	13.4
		50	9.5	16.6	25.2	24.6	27.4	12.1	13.9	13.8	16.0	13.2	20.5	21.4	19.5	21.6	22.2
1		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	7.2	12.7	19.0	19.8	22.6	8.2	7.8	8.9	9.0	8.0	8.0	7.2	8.8	6.3	7.8
	3	10	9.7	12.9	19.3	24.6	25.2	10.2	10.9	13.1	11.1	10.5	12.5	11.9	14.5	12.8	13.2
		20	11.0	14.7	23.0	26.3	26.1	12.3	13.2	14.1	12.3	13.2	18.7	19.3	17.2	18.9	18.1
		30	12.9	17.9	23.3	27.8	26.1	13.2	13.9	15.2	14.7	13.1	24.0	22.9	23.1	23.9	22.7
		50	13.4	20.8	23.4	28.0	26.6	15.3	14.4	15.5	16.9	14.0	25.5	26.9	28.3	25.9	26.9
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	2.8	5.5	13.5	17.1	20.2	5.7	4.2	5.6	6.2	5.2	7.2	4.1	5.6	6.5	5.6
	1	10	3.2	6.0	12.9	19.6	19.0	4.7	8.0	7.5	7.2	6.3	7.3	8.6	9.5	9.7	6.9
		20	3.5	7.6	15.6	20.8	24.4	6.2	7.2	6.5	7.7	7.4	13.3	15.8	13.8	12.8	12.2
		30	4.0	9.3	16.1	21.3	19.3	7.5	8.2	8.4	9.2	8.9	17.2	16.5	17.7	17.0	13.9
		50	5.5	7.8	16.1	21.7	26.7	8.5	8.9	10.2	11.9	10.1	17.9	21.0	21.6	19.4	20.2
2		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	4.6	9.6	13.0	16.5	19.7	4.6	4.6	6.9	8.3	7.4	9.4	6.6	7.8	7.7	8.3
	3	10	4.0	9.1	14.8	20.8	21.1	6.7	7.6	9.1	7.4	6.5	14.2	10.3	10.3	13.1	12.0
		20	5.5	9.4	13.9	20.5	24.3	8.4	10.6	10.4	9.8	7.5	18.1	17.9	19.1	19.0	16.7
		30	7.4	9.2	16.5	22.7	22.7	10.3	9.6	8.9	10.8	9.6	19.8	21.0	22.0	21.9	21.8
		50	4.3	10.8	16.5	22.6	23.7	11.6	10.3	11.0	12.1	13.6	27.6	25.2	24.5	24.7	25.6

Table 2: Test 0, DF 3																	
m	σ_ε^2		$\rho_j = 0.95$					$\rho_j = 0.7$					$\rho_j = 0.1$				
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	8.0	6.1	6.2	7.7	10.3	1.2	1.8	2.8	1.0	1.9	5.6	4.5	4.8	6.0	5.7
	1	10	7.2	7.8	8.3	10.0	11.0	2.8	2.1	2.3	2.0	2.9	8.9	7.7	7.5	8.5	8.2
		20	9.6	9.0	7.4	11.4	13.0	3.4	3.0	3.2	3.5	3.5	10.7	11.9	9.6	10.6	11.9
		30	10.5	12.3	9.9	13.6	12.1	4.4	3.0	3.6	3.4	3.5	15.7	12.7	13.8	12.8	12.4
		50	11.5	11.8	10.1	13.2	15.5	3.1	4.2	5.5	4.3	4.3	19.1	18.1	17.3	18.4	19.6
1		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	7.7	8.8	9.4	11.3	10.6	1.6	3.4	2.3	3.9	2.5	8.1	6.3	8.5	6.6	8.2
	3	10	11.3	11.1	10.6	11.4	13.0	4.0	3.4	3.3	3.2	3.7	11.4	10.5	12.2	11.2	11.4
		20	10.9	11.2	12.2	12.7	17.5	5.2	4.4	3.6	4.3	5.1	18.1	17.1	17.0	18.5	16.6
		30	12.8	11.3	12.3	12.8	14.6	5.4	4.3	4.3	6.2	5.0	21.7	18.8	18.6	19.4	20.7
		50	12.3	14.1	12.1	13.8	17.2	6.8	6.4	4.7	5.3	6.3	22.7	21.3	23.4	23.1	23.5
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	5.0	3.8	4.8	6.0	8.0	0.6	1.1	0.6	1.1	1.2	6.8	4.1	4.0	5.7	5.9
	1	10	5.0	5.1	6.4	7.1	9.7	0.9	0.9	0.9	1.3	0.9	7.3	6.6	7.8	7.0	8.4
		20	5.9	4.6	5.6	7.8	8.9	1.3	1.3	1.3	1.6	1.4	12.1	12.1	10.3	9.4	10.6
		30	5.2	5.3	5.7	7.5	9.3	1.6	1.5	1.1	1.3	1.6	14.3	12.6	14.3	13.3	14.5
		50	6.3	5.7	7.0	7.2	9.6	1.1	2.2	1.5	1.7	2.2	16.7	15.7	14.1	14.3	18.5
2		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	5.8	5.4	6.1	7.5	8.8	1.3	1.1	1.6	0.9	1.2	7.0	6.9	5.8	7.6	5.0
	3	10	5.0	4.8	5.3	6.5	8.7	1.7	1.4	1.6	1.1	1.4	10.0	9.4	7.3	8.2	7.8
		20	8.7	6.3	6.0	8.0	8.8	0.9	2.0	1.4	1.7	1.0	18.6	13.3	14.6	12.8	14.3
		30	7.7	6.5	6.7	8.3	9.5	2.3	2.6	1.9	2.3	2.4	17.4	17.1	15.3	14.9	16.8
		50	8.1	5.6	8.5	8.4	9.3	3.0	1.6	1.8	2.8	2.8	22.7	22.6	20.5	19.5	20.0

Table 3: Test 1, DF 2

m	σ_ε^2		$\rho_j = 0.95$					$\rho_j = 0.7$					$\rho_j = 0.1$				
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	11.5	9.2	5.3	5.1	4.0	5.2	5.5	4.4	5.1	5.3	4.2	3.3	5.5	6.4	5.6
	1	10	26.1	24.7	12.5	9.0	7.8	11.4	6.8	6.3	7.1	8.3	7.8	5.1	4.5	9.9	7.0
		20	45.5	41.6	30.1	14.3	10.1	15.4	10.6	6.2	8.7	8.2	6.8	5.6	6.1	8.4	11.1
		30	54.3	51.7	38.2	21.8	13.3	19.9	13.0	10.0	8.8	11.7	7.1	7.2	6.7	6.9	11.8
		50	63.9	56.4	48.1	28.8	16.2	25.5	18.5	11.7	7.0	8.9	8.2	6.4	6.3	5.6	7.8
1		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	17.5	11.0	8.6	5.5	5.1	7.4	5.7	6.3	7.8	8.1	4.5	5.9	7.5	5.8	7.6
	3	10	34.3	29.2	16.8	9.8	9.8	10.8	9.3	8.8	11.6	12.9	6.3	4.7	8.6	11.3	13.3
		20	48.4	42.2	30.0	21.1	17.0	14.9	11.3	9.4	13.6	18.3	6.2	5.3	5.2	14.6	17.6
		30	55.0	48.5	36.0	27.9	21.7	20.9	15.6	11.1	15.5	21.7	6.0	5.1	5.9	14.4	22.6
		50	60.5	59.3	44.1	26.4	25.7	25.9	18.2	12.0	8.7	18.7	7.3	6.4	5.3	5.3	19.1
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	11.4	12.8	16.2	16.5	17.0	5.0	4.2	3.7	4.4	4.1	5.3	3.6	5.0	6.0	5.6
	1	10	21.0	19.1	21.0	21.5	21.5	11.0	10.9	8.5	8.3	7.3	9.2	7.1	7.4	8.2	8.8
		20	20.7	23.0	23.8	24.0	23.5	14.1	13.9	10.3	9.1	9.7	12.5	11.4	11.5	11.2	11.2
		30	26.5	26.1	25.0	23.3	21.7	14.6	14.9	12.0	10.9	9.1	13.4	16.1	14.7	14.5	13.0
		50	27.8	24.0	25.8	26.5	28.0	17.7	14.7	15.2	13.1	13.3	18.5	21.7	19.7	17.2	20.1
2		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	12.3	15.7	17.1	18.5	19.5	5.3	5.0	4.9	5.3	5.3	5.6	4.6	7.0	8.2	7.6
	3	10	18.6	19.0	20.9	23.5	19.8	12.0	11.3	8.7	9.2	8.1	13.0	9.6	9.6	13.3	12.8
		20	22.7	22.4	25.4	24.2	27.0	17.3	14.8	14.0	12.7	10.1	18.2	17.7	18.5	19.7	20.2
		30	25.6	24.7	26.6	24.4	26.3	16.3	15.0	15.7	14.1	12.0	22.8	22.7	23.7	18.6	22.5
		50	23.3	25.3	25.4	27.5	27.1	19.9	18.4	16.8	15.1	15.5	29.9	29.3	26.2	27.8	23.3

Table 4: Test 1, DF 3

m	σ_ε^2		$\rho_j = 0.95$					$\rho_j = 0.7$					$\rho_j = 0.1$				
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	8.5	4.5	2.9	2.9	3.1	5.1	2.9	2.7	3.0	4.8	4.8	3.2	3.5	5.1	5.3
	1	10	17.8	16.4	10.7	6.0	4.5	9.3	6.5	5.6	5.6	6.9	6.7	5.2	5.0	6.3	6.9
		20	30.9	30.9	28.5	13.8	10.6	17.4	12.7	8.0	6.8	10.5	6.9	5.0	6.4	7.7	13.1
		30	40.2	42.4	37.5	21.5	12.9	22.3	16.4	10.4	7.4	12.7	7.3	5.9	5.0	6.2	11.3
		50	50.3	47.4	46.3	37.5	23.6	32.7	20.7	13.7	8.8	11.8	8.2	5.3	6.2	5.9	10.9
1		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	10.5	7.6	5.7	3.7	3.6	6.5	4.6	4.9	6.2	6.0	4.5	4.0	5.7	5.3	9.5
	3	10	20.9	20.4	11.2	7.5	5.9	11.9	11.4	6.9	8.1	13.1	6.1	6.0	8.6	9.3	12.7
		20	35.6	32.1	28.6	18.1	13.8	18.2	11.3	8.7	11.8	17.2	6.2	6.3	5.8	16.1	16.9
		30	42.5	42.7	35.7	26.4	20.1	24.5	18.6	11.7	16.2	21.5	7.5	6.4	4.9	15.0	21.0
		50	54.7	51.8	46.6	34.5	31.4	31.9	21.4	12.3	7.5	21.9	9.2	6.7	6.4	5.5	23.9
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	4.8	4.4	5.4	6.7	8.6	2.1	1.6	1.0	1.2	1.7	3.2	2.4	3.0	5.1	5.6
	1	10	7.0	8.7	9.8	10.7	10.3	5.0	3.8	2.0	3.0	2.6	10.2	6.5	4.8	7.6	9.6
		20	13.0	9.4	11.8	11.5	13.9	6.6	5.4	3.5	2.8	2.5	11.7	9.9	11.4	8.6	11.7
		30	11.4	13.3	12.2	11.9	13.5	8.6	5.6	4.6	4.3	3.2	15.4	14.4	11.7	12.1	14.5
		50	13.4	10.9	13.0	12.2	13.7	10.6	7.3	5.5	5.6	3.9	22.3	19.5	18.4	17.0	18.7
2		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	6.3	6.6	7.1	8.4	8.8	3.3	1.7	0.7	1.3	1.7	5.9	4.9	4.6	6.9	6.3
	3	10	10.5	11.9	10.0	10.4	13.9	5.5	4.5	3.5	2.7	2.1	12.0	9.3	9.9	9.9	10.4
		20	12.7	12.1	14.1	12.8	14.0	8.9	6.1	6.5	2.9	3.3	17.0	16.0	16.3	13.8	17.0
		30	15.0	13.6	12.7	13.0	14.6	9.5	8.4	6.6	4.5	4.5	22.5	22.2	18.7	15.6	19.7
		50	15.7	14.1	12.6	15.7	15.3	10.6	10.2	5.8	6.9	5.2	27.9	26.0	21.2	23.0	22.5

Table 5: Test 2, DF 2

m	σ_ε^2		$\rho_j = 0.95$					$\rho_j = 0.7$					$\rho_j = 0.1$				
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	23.0	18.6	15.5	11.4	8.1	8.7	8.3	8.3	7.3	8.1	6.6	7.4	8.5	5.9	7.7
	1	10	33.3	32.9	20.8	16.0	10.5	11.0	7.2	7.5	6.8	6.4	5.7	6.1	6.6	6.9	5.9
		20	50.4	42.9	30.0	17.6	10.1	13.4	9.6	7.1	7.4	5.9	6.7	5.1	6.5	6.0	5.2
		30	58.4	52.8	39.8	22.1	12.1	19.4	13.0	10.2	7.1	7.2	6.8	6.6	6.0	6.1	6.2
		50	67.3	61.2	50.5	29.7	17.0	26.6	19.0	10.4	7.5	5.3	6.4	5.5	5.1	5.8	4.8
1		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	25.0	19.5	16.5	11.6	9.5	10.4	8.3	8.5	7.5	8.6	7.5	6.9	7.8	6.5	7.1
	3	10	37.1	33.3	23.2	13.6	10.1	12.5	11.2	6.5	6.2	6.1	6.4	5.3	5.3	6.3	6.1
		20	53.0	45.0	29.3	18.9	12.4	14.9	10.4	8.5	8.3	6.2	5.4	4.2	5.3	6.2	5.1
		30	58.1	50.3	38.4	21.5	12.1	20.3	15.0	11.4	6.4	4.1	6.7	5.3	6.9	4.7	4.8
		50	65.2	62.0	46.1	26.8	16.1	25.6	17.8	11.2	7.7	6.8	7.3	5.7	5.5	4.7	5.6
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	13.3	10.6	17.6	17.7	18.7	5.7	6.6	5.0	5.5	7.0	7.3	6.6	6.3	6.6	5.8
	1	10	17.8	16.1	18.4	18.7	21.6	8.9	7.6	8.6	7.6	8.5	8.1	8.3	7.3	6.8	7.2
		20	22.8	17.8	20.0	22.6	23.8	12.8	11.4	9.2	9.1	9.3	11.1	10.5	10.9	9.3	11.5
		30	30.9	22.5	21.2	21.8	22.0	12.6	11.4	10.4	10.9	11.0	13.4	15.9	14.0	14.8	10.2
		50	31.9	25.6	18.7	23.4	25.3	18.1	13.3	12.2	13.7	13.7	19.1	20.8	18.8	16.3	19.9
2		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	11.1	14.1	17.8	20.3	20.9	5.6	5.2	7.7	7.9	7.7	8.0	9.4	7.3	8.5	8.8
	3	10	19.1	15.4	19.6	24.6	21.9	9.1	9.6	9.7	11.2	11.6	13.7	12.8	9.7	13.3	10.5
		20	22.9	21.6	21.2	22.8	26.2	11.4	11.3	13.1	12.2	12.5	16.3	17.7	19.3	16.3	15.1
		30	26.3	23.1	21.3	24.2	25.4	13.5	12.7	14.0	14.6	12.8	21.9	20.6	21.8	19.3	20.0
		50	27.9	24.1	23.8	22.7	28.0	18.3	16.7	15.3	15.8	15.5	30.7	26.2	24.5	26.1	25.3

Table 6: Test 2, DF 3

m	σ_ε^2		$\rho_j = 0.95$					$\rho_j = 0.7$					$\rho_j = 0.1$				
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	14.2	14.1	15.0	11.7	12.7	10.5	9.9	9.6	7.5	8.9	8.8	6.6	7.3	6.9	7.7
	1	10	20.5	21.4	18.6	16.3	12.5	13.1	8.9	9.0	7.7	7.4	7.9	7.7	7.3	7.4	6.4
		20	31.3	34.3	31.0	19.7	17.1	16.2	13.7	8.5	7.4	7.9	6.2	4.9	5.1	6.2	6.6
		30	41.0	42.2	35.7	25.0	18.8	22.4	17.4	9.9	7.2	7.8	6.4	5.3	6.8	6.5	5.8
		50	50.3	49.1	48.3	37.2	25.6	29.7	23.4	13.4	8.7	8.0	8.3	5.0	5.3	5.7	7.4
1		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	15.9	13.9	15.1	13.7	12.4	11.1	10.1	8.1	8.3	8.5	9.1	7.2	7.8	7.3	6.8
	3	10	19.4	23.3	19.9	15.6	10.8	14.7	10.4	7.8	8.1	8.2	6.9	5.9	6.1	4.9	6.8
		20	33.5	33.4	29.1	20.6	16.6	18.2	10.5	8.7	7.0	7.1	5.4	5.8	5.0	5.8	5.8
		30	43.3	42.6	37.6	27.6	18.0	22.6	16.6	11.2	7.5	7.7	6.3	4.8	5.5	4.4	5.7
		50	53.2	51.9	47.7	35.4	24.2	31.1	20.1	12.5	8.1	8.9	9.0	6.2	5.7	5.7	7.0
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	6.4	6.0	5.1	7.4	9.5	2.6	3.4	1.8	1.8	2.0	8.5	7.2	7.2	7.2	7.5
	1	10	7.8	6.9	7.7	9.1	9.9	3.5	2.9	1.5	2.5	3.0	7.7	7.8	8.1	6.1	8.1
		20	13.0	9.0	9.2	9.9	12.3	6.6	2.9	3.1	3.7	3.4	12.0	9.3	11.2	9.4	9.1
		30	13.3	11.1	10.5	8.4	11.7	7.3	4.3	2.6	3.7	3.5	15.8	13.8	11.9	12.0	14.0
		50	15.9	11.8	11.1	9.0	11.1	9.4	4.9	4.8	3.9	3.6	19.3	18.1	16.9	15.2	18.6
2		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	6.0	6.8	7.1	9.3	9.6	2.8	1.7	1.3	1.5	3.8	9.3	8.4	9.0	8.8	7.1
	3	10	8.3	9.4	8.5	9.0	12.8	3.6	3.0	2.4	4.1	3.6	11.4	11.8	9.5	10.3	10.0
		20	11.3	12.0	10.2	10.3	12.0	6.2	4.6	4.6	4.6	3.9	17.6	15.6	15.1	15.3	14.8
		30	13.7	12.7	11.6	10.1	14.5	8.8	6.8	4.1	5.2	5.2	21.3	19.9	20.9	18.0	18.9
		50	15.8	12.0	10.9	12.6	13.3	9.1	6.1	5.1	5.2	5.4	26.2	25.4	19.2	23.7	22.5

Table 7: Test 3, DF 2

m	σ_ε^2		$\rho_j = 0.95$					$\rho_j = 0.7$					$\rho_j = 0.1$				
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	21.2	17.4	15.4	11.4	8.2	7.9	8.2	8.5	7.2	8.1	6.1	7.6	7.8	5.8	7.7
	1	10	32.6	32.1	20.7	16.0	10.6	9.6	7.5	6.9	6.8	6.3	5.4	5.6	6.9	6.9	6.0
		20	49.2	41.7	29.5	17.5	10.3	11.2	9.0	6.9	7.4	6.0	6.5	4.9	6.4	6.0	5.2
		30	55.7	51.8	39.6	22.0	12.1	16.5	12.4	10.0	6.8	7.1	6.1	6.1	6.0	5.9	5.9
		50	65.7	60.5	50.0	29.4	16.9	23.3	17.8	10.2	7.5	5.2	6.0	5.6	5.0	5.8	4.9
1		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	24.5	19.3	16.4	11.5	9.5	9.7	8.4	8.5	7.6	8.6	7.4	7.1	8.1	6.6	7.1
	3	10	36.1	32.8	22.7	13.7	10.1	12.3	11.4	6.4	6.3	6.1	5.7	4.9	5.2	6.1	6.1
		20	52.9	44.6	29.5	18.9	12.4	14.1	10.2	8.5	8.2	6.2	5.2	4.1	5.4	6.2	5.1
		30	56.6	50.1	38.1	21.6	12.2	19.3	14.5	11.3	6.5	4.1	6.0	5.0	6.8	4.8	5.0
		50	64.5	61.9	46.0	26.7	16.2	24.3	17.4	11.3	7.7	6.8	7.0	5.7	5.6	4.7	5.7
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	12.8	10.6	17.6	17.6	18.9	5.3	6.1	5.0	5.9	7.2	7.6	6.6	6.7	6.7	6.2
	1	10	17.1	16.7	18.6	18.7	21.2	8.6	7.5	8.7	7.9	8.4	7.2	8.2	8.1	7.9	7.8
		20	22.5	17.4	20.0	22.8	23.9	11.1	11.5	9.3	9.3	9.5	11.1	11.2	13.3	11.0	11.6
		30	27.7	22.5	21.3	21.9	21.8	11.1	11.2	10.5	11.0	11.4	14.8	17.5	15.2	17.0	12.4
		50	29.8	22.9	18.6	23.3	25.1	15.2	14.1	12.7	13.6	14.0	18.8	22.6	20.1	18.4	21.4
2		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	11.3	13.8	17.8	20.3	20.8	5.7	5.3	7.5	7.8	7.5	8.4	10.0	7.3	8.5	9.7
	3	10	18.9	15.7	18.7	24.8	21.7	9.6	9.7	9.7	11.2	11.6	14.1	13.0	11.7	14.2	10.8
		20	23.2	21.6	21.3	22.8	26.3	11.3	11.3	13.2	12.8	12.8	15.9	19.8	23.0	19.0	17.1
		30	25.2	24.0	20.9	24.2	25.6	12.4	12.6	14.7	14.5	13.8	22.4	22.6	25.0	21.6	22.5
		50	27.3	22.4	23.6	22.6	28.4	16.6	17.6	15.5	15.8	16.9	30.5	29.4	28.4	28.9	28.7

Table 8: Test 3, DF 3

m	σ_ε^2		$\rho_j = 0.95$					$\rho_j = 0.7$					$\rho_j = 0.1$				
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	14.3	14.3	14.7	11.9	12.8	10.4	9.8	9.0	7.3	9.5	9.2	6.7	7.1	7.4	8.1
	1	10	19.4	20.2	18.1	16.1	12.3	13.2	8.6	9.1	7.5	7.5	7.6	7.9	7.1	7.1	6.3
		20	29.6	32.8	30.8	19.7	17.0	15.8	12.7	8.3	7.4	7.9	6.2	4.6	5.3	6.1	6.6
		30	39.4	41.4	35.7	24.9	18.7	20.7	17.1	9.8	7.1	7.8	6.4	5.4	6.8	6.5	6.0
		50	48.8	48.4	48.4	37.1	25.6	28.7	22.3	13.2	8.7	8.1	8.1	4.9	5.3	5.7	7.4
1		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	16.3	13.8	15.0	13.5	12.5	11.1	9.9	8.0	8.3	8.5	9.4	6.9	7.9	7.3	6.3
	3	10	19.0	23.1	19.8	15.5	10.7	14.2	10.3	8.1	8.3	8.2	6.6	6.1	6.1	4.9	6.9
		20	33.3	33.2	29.0	20.7	16.6	18.0	10.2	8.6	7.0	7.1	5.1	5.6	4.9	5.7	6.0
		30	42.8	42.4	37.6	27.4	18.0	22.1	16.4	11.1	7.4	7.7	6.8	4.8	5.5	4.3	5.7
		50	53.0	51.8	47.7	35.4	24.2	30.7	19.8	12.5	8.1	8.7	8.9	6.1	5.7	5.7	7.1
		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	6.9	6.0	5.3	7.3	9.5	2.5	2.9	1.9	1.8	2.1	8.3	7.3	7.5	7.5	8.3
	1	10	8.3	6.5	7.7	9.0	9.9	3.7	2.6	1.9	2.4	3.0	7.8	9.0	8.3	7.2	8.8
		20	12.7	8.6	8.5	9.9	12.0	6.9	2.8	3.2	3.8	3.5	12.8	11.5	13.0	11.9	10.4
		30	12.0	10.8	10.1	8.5	11.2	7.2	3.8	2.7	3.8	3.8	16.4	15.9	14.3	14.2	15.4
		50	14.9	11.7	10.5	8.6	11.0	9.2	4.9	4.9	3.9	4.0	22.3	21.0	19.5	17.9	21.0
2		N/T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
		5	5.9	6.8	7.3	9.1	9.6	3.1	2.0	1.4	1.5	4.0	9.1	9.1	9.4	9.3	7.2
	3	10	8.6	9.1	8.2	9.1	12.5	4.2	3.3	2.1	4.5	3.7	13.7	13.7	11.9	12.3	11.8
		20	11.6	11.7	9.8	10.1	11.6	6.5	4.5	4.9	4.8	4.1	19.6	18.2	17.8	20.2	17.8
		30	13.1	12.2	11.3	9.8	13.9	8.6	6.4	3.9	5.0	5.5	24.6	22.4	26.1	22.0	22.6
		50	14.8	11.1	10.2	12.0	13.3	9.9	6.4	5.5	6.0	5.9	30.1	28.9	25.3	28.4	26.5

Queen Mary
University of London

This working paper has been produced by
the Department of Economics at
Queen Mary, University of London

Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 5096 or Fax: +44 (0)20 8983 3580
Email: j.conner@qmul.ac.uk
Website: www.econ.qmul.ac.uk/papers/wp.htm