Panels with Nonstationary Multifactor Error Structures

G. Kapetanios, M. Hashem Pesaran and T. Yamagata

August 2006

CWPE 0651

Panels with Nonstationary Multifactor Error Structures[*]

G. Kapetanios
Queen Mary, University of London

M. Hashem Pesaran
Cambridge University and Trinity College, Cambridge

T. Yamagata
Cambridge University

July 31, 2006

Abstract

The presence of cross-sectionally correlated error terms invalidates much inferential theory of panel data models. Recently work by Pesaran (2006) has suggested a method which makes use of cross-sectional averages to provide valid inference for stationary panel regressions with multifactor error structure. This paper extends this work and examines the important case where the unobserved common factors follow unit root processes and could be cointegrated. It is found that the presence of unit roots does not affect most theoretical results which continue to hold irrespective of the integration and the cointegration properties of the unobserved factors. This finding is further supported for small samples via an extensive Monte Carlo study. In particular, the results of the Monte Carlo study suggest that the cross-sectional average based method is robust to a wide variety of data generation processes and has lower biases than all of the alternative estimation methods considered in the paper.

Keywords: Cross Section Dependence, Large Panels, Unit Roots, Principal Components, Common Correlated Effects.

JEL-Classification: C12, C13, C33.

1 Introduction

Panel data sets have been increasingly used in economics to analyze complex economic phenomena. One of their attractions is the ability to use an extended data set to obtain information about parameters of interest which are assumed to have common values across panel units. Most of the work carried out on panel data has usually assumed some form of cross sectional independence to derive the theoretical properties of various inferential procedures. However, such assumptions are often suspect and as a result recent advances in the literature have focused on estimation of panel data models subject to error cross sectional dependence.

A number of different approaches have been advanced for this purpose. In the case of spatial data sets where a natural immutable distance measure is available the dependence is often captured through "spatial lags" using techniques familiar from the time series literature. In economic applications, spatial techniques are often adapted using alternative measures of "economic distance". This approach is exemplified in work by Lee and Pesaran (1993), Conley and Dupor (2003), Conley and Topa (2002) and Pesaran, Schuermann, and Weiner (2004), as well as the literature on spatial econometrics recently surveyed by Anselin (2001). In the case of panel data models where the cross section dimension (N) is small (typically $N < 10$) and the time series dimension (T) is large the standard approach is to treat the equations from the different cross section units as a system of seemingly unrelated regression equations (SURE) and then estimate the system by the Generalized Least Squares (GLS) techniques.

In the case of panels with large cross section dimension, SURE approach is not practical and has led a number of investigators to consider unobserved factor models, where the cross section error correlations are defined in terms of the factor loadings. Use of factor models is not new in economics and dates back to the pioneering work of Stone (1947) who applied the principal components (PC) analysis of Hotelling to US macroeconomic time series over the period 1922-1938 and was able to demonstrate that three factors (namely total income, its rate of change and a time trend) explained over 97 per cent of the total variations of all the 17 macro variables that he had considered. Until recently, subsequent applications of the PC approach to economic times series has been primarily in finance. See, for example, Chamberlain and Rothschild (1983), Connor and Korajzcyk (1986) and Connor and Korajzcyk (1988). But more recently the unobserved factor models have gained popularity for forecasting with a large number of variables as advocated by Stock and Watson (2002).

The factor model is used very much in the spirit of the original work by Stone, in order to summarize the empirical content of a large number of macroeconomics variables by a small set of factors which, when estimated using principal components, is then used for further modelling and/or forecasting. A related literature on dynamic factor models has also been put forward by Forni and Reichlin (1998) and Forni, Hallin, Lippi, and Reichlin (2000).

Recent uses of factor models in forecasting focuses on consistent estimation of unobserved factors and their loadings. Related theoretical advances by Bai and Ng (2002) and Bai (2003) are also concerned with estimation and selection of unobserved factors and do not consider the estimation and inference problems in standard panel data models where the objects of interest are slope coefficients of the conditioning variables (regressors). In such panels the unobserved factors are viewed as nuisance variables, introduced primarily to model the cross section dependencies of the error terms in a parsimonious manner relative to the SURE formulation.

Despite these differences knowledge of factor models could still be useful for the analysis of panel data models if it is believed that the errors might be cross sectionally correlated. Disregarding the possible factor structure of the errors in panel data models can lead to inconsistent parameter estimates and incorrect inference. Coakley, Fuertes, and Smith (2002) suggest a possible solution to the problem using the method of Stock and Watson (2002). But, as Pesaran (2006) shows, the PC approach proposed by Coakley, Fuertes, and Smith (2002) can still yield inconsistent estimates. Pesaran (2006) suggests a new approach by noting that linear combinations of the unobserved factors can be well approximated by cross section averages of the dependent variable and the observed regressors. This leads to a new set of estimators, referred to as the Common Correlated Effects estimators, that can be computed by running standard panel regressions augmented with the cross section averages of the dependent and independent variables. The CCE procedure is applicable to panels with a single or multiple unobserved factors so long as the number of unobserved factors is fixed.

In this paper we extend the analysis of Pesaran (2006) to the case where one or more of the unobserved common factors could be integrated of order 1, or $I(1)$. This also allows for the possibility of cointegration amongst the $I(1)$ factors and does not require an a priori knowledge of the number of unobserved factors, or their integration or cointegration properties. It is only required that the number of unobserved factors remain fixed as the sample size is increased. The theoretical findings of the paper are further supported for small samples via an extensive Monte Carlo study. In particular, the results of the Monte Carlo study

clearly show that the CCE estimator is robust to a wide variety of data generation processes and has lower biases than all of the alternative estimation methods considered in the paper.

The structure of the paper is as follows: Section 2 provides an overview of the method suggested by Pesaran (2006) in the case of stationary factor processes. Section 3 provides the theoretical framework of the analysis of nonstationarity. In this section the theoretical properties of the various estimators are presented. Section 4 presents an extensive Monte Carlo study, and Section 5 concludes.

2 Panel Data Models with Observed and Unobserved Common Effects

In this section we review the methodology introduced in Pesaran (2006). Let y_{it} be the observation on the ith cross section unit at time t for $i = 1, 2, ..., N$; $t = 1, 2, ..., T$, and suppose that it is generated according to the following linear heterogeneous panel data model

$$y_{it} = \boldsymbol{\alpha}'_i \mathbf{d}_t + \boldsymbol{\beta}'_i \mathbf{x}_{it} + \boldsymbol{\gamma}'_i \mathbf{f}_t + \varepsilon_{it}, \tag{1}$$

where $\mathbf{d}_t$ is a $n \times 1$ vector of observed common effects, which is partitioned as $\mathbf{d}_t = (\mathbf{d}'_{1t}, \mathbf{d}'_{2t}, \mathbf{d}'_{3t})'$ where $\mathbf{d}_{1t}$ is a $n_1 \times 1$ vector of deterministic components such as intercepts or seasonal dummies, $\mathbf{d}_{2t}$ is a $n_2 \times 1$ vector of unit root stochastic observed common effects and $\mathbf{d}_{3t}$ is a $n_3 \times 1$ vector of stationary stochastic observed common effects, with $n = n_1 + n_2 + n_3$, $\mathbf{x}_{it}$ is a $k \times 1$ vector of observed individual-specific regressors on the ith cross section unit at time t, $\mathbf{f}_t$ is the $m \times 1$ vector of unobserved common effects, and ε_{it} are the individual-specific (idiosyncratic) errors assumed to be independently distributed of $(\mathbf{d}_t, \mathbf{x}_{it})$. The unobserved factors, $\mathbf{f}_t$, could be correlated with $(\mathbf{d}_t, \mathbf{x}_{it})$, and to allow for such a possibility the following specification for the individual specific regressors will be considered

$$\mathbf{x}_{it} = \mathbf{A}'_i \mathbf{d}_t + \boldsymbol{\Gamma}'_i \mathbf{f}_t + \mathbf{v}_{it}, \tag{2}$$

where $\mathbf{A}_i$ and $\boldsymbol{\Gamma}_i$ are $n \times k$ and $m \times k$, factor loading matrices with fixed components, $\mathbf{v}_{it}$ are the specific components of $\mathbf{x}_{it}$ distributed independently of the common effects and across i, but assumed to follow general covariance stationary processes. In this paper we allow for one or more of the common factors, $\mathbf{d}_t$ and $\mathbf{f}_t$, to be $I(1)$.

Combining (1) and (2) we now have

$$\underset{(k+1)\times 1}{\mathbf{z}_{it}} = \begin{pmatrix} y_{it} \\ \mathbf{x}_{it} \end{pmatrix} = \underset{(k+1)\times n}{\mathbf{B}'_i} \underset{n\times 1}{\mathbf{d}_t} + \underset{(k+1)\times m}{\mathbf{C}'_i} \underset{m\times 1}{\mathbf{f}_t} + \underset{(k+1)\times 1}{\mathbf{u}_{it}}, \tag{3}$$

4

where

$$\mathbf{u}_{it} = \begin{pmatrix} \varepsilon_{it} + \boldsymbol{\beta}_i' \mathbf{v}_{it} \\ \mathbf{v}_{it} \end{pmatrix}, \tag{4}$$

$$\mathbf{B}_i = \begin{pmatrix} \boldsymbol{\alpha}_i & \mathbf{A}_i \end{pmatrix} \begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ \boldsymbol{\beta}_i & \mathbf{I}_k \end{pmatrix}, \mathbf{C}_i = \begin{pmatrix} \boldsymbol{\gamma}_i & \boldsymbol{\Gamma}_i \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ \boldsymbol{\beta}_i & \mathbf{I}_k \end{pmatrix}, \tag{5}$$

$\mathbf{I}_k$ is an identity matrix of order k, and the rank of $\mathbf{C}_i$ is determined by the rank of the $m \times (k+1)$ matrix of the unobserved factor loadings

$$\tilde{\boldsymbol{\Gamma}}_i = \begin{pmatrix} \boldsymbol{\gamma}_i & \boldsymbol{\Gamma}_i \end{pmatrix}. \tag{6}$$

As discussed in Pesaran (2006), the above set up is sufficiently general and renders a variety of panel data models as special cases. In the panel literature with T small and N large, the primary parameters of interest are the means of the individual specific slope coefficients, $\boldsymbol{\beta}_i$, $i = 1, 2, ..., N$. The common factor loadings, $\boldsymbol{\alpha}_i$ and $\boldsymbol{\gamma}_i$, are generally treated as nuisance parameters. In cases where both N and T are large, it is also possible to consider consistent estimation of the factor loadings. The presence of unobserved factors in (1) implies that estimation of $\boldsymbol{\beta}_i$ and its cross sectional mean cannot be undertaken using standard methods. Pesaran (2006) has suggested using cross section averages of y_{it} and $\mathbf{x}_{it}$ as proxies for the unobserved factors in (1). To see why such an approach could work, consider simple cross section averages of the equations in (3)[1]

$$\bar{\mathbf{z}}_t = \bar{\mathbf{B}}' \mathbf{d}_t + \bar{\mathbf{C}}' \mathbf{f}_t + \bar{\mathbf{u}}_t, \tag{7}$$

where

$$\bar{\mathbf{z}}_t = \frac{1}{N} \sum_{i=1}^{N} \mathbf{z}_{it}, \bar{\mathbf{u}}_t = \frac{1}{N} \sum_{i=1}^{N} \mathbf{u}_{it},$$

and

$$\bar{\mathbf{B}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{B}_i, \bar{\mathbf{C}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{C}_i. \tag{8}$$

Suppose that

$$Rank(\bar{\mathbf{C}}) = m \le k + 1, \text{ for all } N. \tag{9}$$

Then, we have

$$\mathbf{f}_t = \left(\bar{\mathbf{C}} \bar{\mathbf{C}}' \right)^{-1} \bar{\mathbf{C}} \left(\bar{\mathbf{z}}_t - \bar{\mathbf{B}}' \mathbf{d}_t - \bar{\mathbf{u}}_t \right). \tag{10}$$

But since

$$\bar{\mathbf{u}}_t \overset{q.m.}{\to} \mathbf{0}, \text{ as } N \to \infty, \text{ for each } t, \tag{11}$$

and

$$\bar{\mathbf{C}} \overset{p}{\to} \mathbf{C} = \tilde{\boldsymbol{\Gamma}} \begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ \boldsymbol{\beta} & \mathbf{I}_k \end{pmatrix}, \text{ as } N \to \infty, \tag{12}$$

---

[1]Pesaran (2006) considers cross section weighted averages that are more general. But to simplify the exposition we confine our discussion to simple averages throughout.

where

$$\tilde{\boldsymbol{\Gamma}} = (E(\boldsymbol{\gamma}_i), E(\boldsymbol{\Gamma}_i)) = (\boldsymbol{\gamma}, \boldsymbol{\Gamma}). \tag{13}$$

It follows, assuming that $\text{Rank}(\tilde{\boldsymbol{\Gamma}}) = m$, that

$$\mathbf{f}_t - (\mathbf{C}\mathbf{C}')^{-1}\mathbf{C}\left(\bar{\mathbf{z}}_t - \bar{\mathbf{B}}'\mathbf{d}_t\right) \xrightarrow{p} \mathbf{0}, \text{ as } N \to \infty.$$

This suggests using $\bar{\mathbf{h}}_t = (\mathbf{d}_t', \bar{\mathbf{z}}_t')'$ as observable proxies for $\mathbf{f}_t$, and is the basic insight that lies behind the Common Correlated Effects estimators developed in Pesaran (2006). It is further shown that the CCE estimation procedure in fact holds even if $\tilde{\boldsymbol{\Gamma}}$ turns out to be rank deficient.

We now discuss the two estimators for the means of the individual specific slope coefficients proposed by Pesaran (2006). One is the Mean Group (MG) estimator proposed in Pesaran and Smith (1995) and the other is a generalization of the fixed effects estimator that allows for the possibility of cross section dependence. The former is referred to as the "Common Correlated Effects Mean Group" (CCEMG) estimator, and the latter as the "Common Correlated Effects Pooled" (CCEP) estimator.

The CCEMG estimator is a simple average of the individual CCE estimators, $\hat{\mathbf{b}}_i$ of $\boldsymbol{\beta}_i$,

$$\hat{\mathbf{b}}_{MG} = N^{-1}\sum_{i=1}^{N}\hat{\mathbf{b}}_i. \tag{14}$$

where

$$\hat{\mathbf{b}}_i = (\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i)^{-1}\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{y}_i, \tag{15}$$

$\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, ..., \mathbf{x}_{iT})'$, $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{iT})'$, $\bar{\mathbf{M}}$ is defined by

$$\bar{\mathbf{M}} = \mathbf{I}_T - \bar{\mathbf{H}}\left(\bar{\mathbf{H}}'\bar{\mathbf{H}}\right)^{-1}\bar{\mathbf{H}}', \tag{16}$$

$\bar{\mathbf{H}} = (\mathbf{D}, \bar{\mathbf{Z}})$, $\mathbf{D}$ and $\bar{\mathbf{Z}}$ being, respectively, the $T \times n$ and $T \times (k+1)$ matrices of observations on $\mathbf{d}_t$ and $\bar{\mathbf{z}}_t$.

Efficiency gains from pooling of observations over the cross section units can be achieved when the individual slope coefficients, $\boldsymbol{\beta}_i$, are the same. Such a pooled estimator of $\boldsymbol{\beta}$, denoted by CCEP, is given by

$$\hat{\mathbf{b}}_P = \left(\sum_{i=1}^{N}\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i\right)^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{y}_i. \tag{17}$$

3 Theoretical Properties of CCE Estimators in Non-stationary Panel Data Models

The following assumptions will be used in the derivation of the asymptotic properties of the CCE estimators.

Assumption 1 (non-stationary common effects): The $(n_2 + m) \times 1$ vector of stochastic common effects, $\mathfrak{g}_t = (\mathbf{d}'_{2t}, \mathbf{f}'_t)'$, is a multivariate unit root process given by

$$\mathfrak{g}_t = \mathfrak{g}_{t-1} + \boldsymbol{\zeta}_{\mathfrak{g}t}$$

where $\boldsymbol{\zeta}_{\mathfrak{g}t}$ is a $(n_2 + m) \times 1$ vector of L_2 stationary near epoque dependent (NED) processes of size $1/2$, distributed independently of the individual-specific errors, $\varepsilon_{it'}$ and $\mathbf{v}_{it'}$ for all i, t and t'.

Assumption 2 (stationary common effects): The $n_3 \times 1$ vector of common effects, $\mathbf{d}_{3t}$, is a covariance stationary process given by

$$\mathbf{d}_{3t} = \sum_{\ell=0}^{\infty} \mathbf{J}_\ell \boldsymbol{\psi}_{t-\ell}, \tag{18}$$

where the matrices $\mathbf{J}_\ell$ satisfy the condition

$$\sum_{\ell=0}^{\infty} \ell^s ||\mathbf{J}_\ell|| < \infty, \tag{19}$$

for some $s \geq 1/2$, and the $\boldsymbol{\psi}_t$ are distributed independently of the individual-specific errors, $\varepsilon_{it'}$ and $\mathbf{v}_{it'}$ for all i, t and t'.

Assumption 3 (individual-specific errors): The individual specific errors ε_{it} and $\mathbf{v}_{jt}$ are distributed independently for all i, j and t. For each i, ε_{it} is serially uncorrelated with mean zero, a finite variance $\sigma_i^2 < K$, and a finite fourth-order cumulant. $\mathbf{v}_{it}$ follows a linear stationary process with absolute summable autocovariances given by

$$\mathbf{v}_{it} = \sum_{\ell=0}^{\infty} \mathbf{S}_{i\ell} \boldsymbol{\nu}_{i,t-\ell}, \tag{20}$$

where ν_{it} are $k \times 1$ vectors of identically, independently distributed (*IID*) random variables with mean zero, the variance matrix, $\mathbf{I}_k$, and finite fourth-order cumulants. In particular, the $k \times k$ coefficient matrices $\mathbf{S}_{i\ell}$ satisfy the condition

$$\sum_{\ell=0}^{\infty} \ell^s ||\mathbf{S}_{i\ell}|| < \infty, \tag{21}$$

for all i and some $s \geq 1/2$.

Assumption 4 (factor loadings): The unobserved factor loadings, $\boldsymbol{\gamma}_i$ and $\boldsymbol{\Gamma}_i$, are independently and identically distributed across i, and of the individual specific errors, ε_{jt} and $\mathbf{v}_{jt}$, the common factors, $\mathbf{g}_t = (\mathbf{d}'_{2t}, \mathbf{f}'_t)$, for all i, j and t with fixed means $\boldsymbol{\gamma}$ and $\boldsymbol{\Gamma}$, respectively, and finite variances. In particular,

$$\boldsymbol{\gamma}_i = \boldsymbol{\gamma} + \boldsymbol{\eta}_i, \ \ \boldsymbol{\eta}_i \backsim IID\left(\mathbf{0}, \boldsymbol{\Omega}_{\boldsymbol{\eta}}\right), \text{ for } i = 1, 2, ..., N, \tag{22}$$

where $\boldsymbol{\Omega}_{\boldsymbol{\eta}}$ is a $m \times m$ symmetric non-negative definite matrix, and $\|\boldsymbol{\gamma}\| < K$, $\|\boldsymbol{\Gamma}\| < K$, and $\|\boldsymbol{\Omega}_{\boldsymbol{\eta}}\| < K$.

Assumption 5 (random slope coefficients): The slope coefficients, $\boldsymbol{\beta}_i$, follow the random coefficient model

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{v}_i, \ \ \boldsymbol{v}_i \backsim IID\left(\mathbf{0}, \boldsymbol{\Omega}_{\boldsymbol{v}}\right), \text{ for } i = 1, 2, ..., N, \tag{23}$$

where $\|\boldsymbol{\beta}\| < K$, $\|\boldsymbol{\Omega}_{\boldsymbol{v}}\| < K$, $\boldsymbol{\Omega}_{\boldsymbol{v}}$ is a $k \times k$ symmetric non-negative definite matrix, and the random deviations, $\boldsymbol{v}_i$, are distributed independently of $\boldsymbol{\gamma}_j, \boldsymbol{\Gamma}_j, \varepsilon_{jt}, \mathbf{v}_{jt}$, and $\mathbf{g}_t$ for all i, j and t.

Assumption 6: (identification of $\boldsymbol{\beta}_i$ and $\boldsymbol{\beta}$): Consider the cross section averages of the individual specific variables, $\mathbf{z}_{it}$, defined by $\bar{\mathbf{z}}_t = \frac{1}{N}\sum_{j=1}^N \mathbf{z}_{jt}$, and let

$$\bar{\mathbf{M}} = \mathbf{I}_T - \bar{\mathbf{H}}\left(\bar{\mathbf{H}}'\bar{\mathbf{H}}\right)^- \bar{\mathbf{H}}', \tag{24}$$

and

$$\mathbf{M}_g = \mathbf{I}_T - \mathbf{G}\left(\mathbf{G}'\mathbf{G}\right)^- \mathbf{G}', \tag{25}$$

where $\mathbf{G} = (\mathbf{D}, \mathbf{F})$, $\mathbf{D} = (\mathbf{d}_1, \mathbf{d}_2, ..., \mathbf{d}_T)'$, $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_T)'$ are $T \times n$ and $T \times m$ data matrices on observed and unobserved common factors, respectively, $\bar{\mathbf{Z}} = (\bar{\mathbf{z}}_1, \bar{\mathbf{z}}_2, ..., \bar{\mathbf{z}}_T)'$ is the $T \times (k+1)$ matrix of observations on the cross section averages, and $\left(\bar{\mathbf{H}}'\bar{\mathbf{H}}\right)^-$ and $(\mathbf{G}'\mathbf{G})^-$ denote the generalized inverses of $\bar{\mathbf{H}}'\bar{\mathbf{H}}$ and $\mathbf{G}'\mathbf{G}$, respectively. Also denote the $T \times k$ observation matrix on individual specific regressors by $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, ..., \mathbf{x}_{iT})'$.

6a: (identification of $\boldsymbol{\beta}_i$): The $k \times k$ matrices $\hat{\boldsymbol{\Psi}}_{iT} = T^{-1}\left(\mathbf{X}'_i\bar{\mathbf{M}}\mathbf{X}_i\right)$ and $\boldsymbol{\Psi}_{ig} = T^{-1}\left(\mathbf{X}'_i\mathbf{M}_g\mathbf{X}_i\right)$ are non-singular and $\hat{\boldsymbol{\Psi}}_{iT}^{-1}$ and $\boldsymbol{\Psi}_{ig}^{-1}$ have finite second order moments, for all i.

6b: (identification of $\boldsymbol{\beta}$): The $k \times k$ pooled observation matrix $\hat{\boldsymbol{\Psi}}_{NT}$ defined by

$$\hat{\boldsymbol{\Psi}}_{NT} = \frac{1}{N}\sum_{i=1}^N \left(\frac{\mathbf{X}'_i\bar{\mathbf{M}}\mathbf{X}_i}{T}\right) \tag{26}$$

is non-singular.

Remark 1 *Note that Assumption 3 is slightly stronger than Assumption 2 of Pesaran (2006) which requires that*

$$Var\left(\mathbf{v}_{it}\right) = \sum_{\ell=0}^{\infty} \mathbf{S}_{i\ell}\mathbf{S}'_{i\ell} = \boldsymbol{\Sigma}_i \leq \mathbf{K} < \infty, \tag{27}$$

for all i and some constant matrix $\mathbf{K}$, where $\boldsymbol{\Sigma}_i$ is a positive definite matrix, rather than (21). Further note that (21) implies that $\mathbf{v}_{it}$ are L_2 stationary near epoque dependent processes of size 1/2.

Remark 2 *Assumption 1 implicitly allows for cointegration among the m unobserved factors. To see this it is useful to impose some further structure on $\boldsymbol{\zeta}_{2\mathfrak{g}t}$ where $\boldsymbol{\zeta}_{\mathfrak{g}t} = (\boldsymbol{\zeta}'_{1\mathfrak{g}t}, \boldsymbol{\zeta}'_{2\mathfrak{g}t})'$ is a partition of $\boldsymbol{\zeta}_{\mathfrak{g}t}$ comformable to the partition of $\mathfrak{g}_t$ in terms of $\mathbf{d}_{2t}$ and $\mathbf{f}_t$. If we let $\boldsymbol{\zeta}_{2\mathfrak{g}t} = \sum_{\ell=0}^{\infty} \mathbf{W}_\ell \varpi_{t-\ell}$ where $\varpi_{t-\ell}$ are i.i.d. with finite variance and $\sum_{\ell=0}^{\infty} \ell ||\mathbf{W}_\ell|| < \infty$ then if $\sum_{\ell=0}^{\infty} \mathbf{W}_\ell$ is of reduced rank equal to r, then $\boldsymbol{\zeta}_{2\mathfrak{g}t}$ exhibits cointegration. Further $\mathbf{f}_t$ may be represented by a set of r stationary components and $m - r$ random walk components. This representation will be of use in the proofs of our results. Note that the above MA representation of $\boldsymbol{\zeta}_{2\mathfrak{g}t}$ implies that $\boldsymbol{\zeta}_{2\mathfrak{g}t}$ is a vector of L_2 stationary near epoque dependent (NED) processes of size 1 and hence satisfy assumption 1[2].*

For each i and $t = 1, 2, ..., T$, writing the model in matrix notation we have

$$\mathbf{y}_i = \boldsymbol{D}\boldsymbol{\alpha}_i + \mathbf{X}_i\boldsymbol{\beta}_i + \boldsymbol{F}\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i, \tag{28}$$

where $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iT})'$, and as set out in Assumption 5, $\mathbf{D} = (\mathbf{d}_1, \mathbf{d}_2, ..., \mathbf{d}_T)'$ and $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_T)'$. Using (28) in (15) we have

$$\hat{\mathbf{b}}_i - \boldsymbol{\beta}_i = \left(\frac{\mathbf{X}'_i\bar{\mathbf{M}}\mathbf{X}_i}{T}\right)^{-1}\left(\frac{\mathbf{X}'_i\bar{\mathbf{M}}\mathbf{F}}{T}\right)\boldsymbol{\gamma}_i + \left(\frac{\mathbf{X}'_i\bar{\mathbf{M}}\mathbf{X}_i}{T}\right)^{-1}\left(\frac{\mathbf{X}'_i\bar{\mathbf{M}}\varepsilon_i}{T}\right), \tag{29}$$

which shows the direct dependence of $\hat{\mathbf{b}}_i$ on the unobserved factors through $T^{-1}\mathbf{X}'_i\bar{\mathbf{M}}\mathbf{F}$. To examine the properties of this component, writing (2) and (7) in matrix notations, we first note that

$$\mathbf{X}_i = \mathbf{G}\boldsymbol{\Pi}_i + \mathbf{V}_i, \tag{30}$$

and

$$\bar{\mathbf{H}} = \mathbf{G}\bar{\mathbf{P}} + \bar{\mathbf{U}}^*, \tag{31}$$

where $\boldsymbol{\Pi}_i = (\mathbf{A}'_i, \boldsymbol{\Gamma}'_i)'$, $\mathbf{V}_i = (\mathbf{v}_{i1}, \mathbf{v}_{i2}, ..., \mathbf{v}_{iT})'$,

$$\underset{(n+m)\times(n+k+1)}{\bar{\mathbf{P}}} = \begin{pmatrix} \mathbf{I}_n & \bar{\mathbf{B}} \\ \mathbf{0} & \bar{\mathbf{C}} \end{pmatrix}, \ \bar{\mathbf{U}}^* = (\mathbf{0}, \bar{\mathbf{U}}), \tag{32}$$

Using Lemma 1 which forms the basis of all the theoretical results of this paper and assuming that the rank condition (9) is satisfied, it follows that

$$\frac{\mathbf{X}'_i\bar{\mathbf{M}}\mathbf{F}}{T} = \frac{\mathbf{X}'_i\bar{\mathbf{M}}_g\mathbf{F}}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right) + O_p\left(\frac{1}{N}\right), \tag{33}$$

---

[2]Our analysis allows for cointegration in $\mathbf{d}_{2t}$ as well. However, for the sake of simplicity we do not explicitly analyse this case which is a straightforward extension of the results presented in this paper.

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i}{T} = \frac{\mathbf{X}_i'\bar{\mathbf{M}}_g\mathbf{X}_i}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right) + O_p\left(\frac{1}{N}\right),\tag{34}$$

and

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\boldsymbol{\varepsilon}_i}{T} = \frac{\mathbf{X}_i'\bar{\mathbf{M}}_g\boldsymbol{\varepsilon}_i}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right) + O_p\left(\frac{1}{N}\right).\tag{35}$$

Note that $\mathbf{F} \subset \mathbf{G}$ and hence that $T^{-1}\mathbf{X}_i'\bar{\mathbf{M}}_g\mathbf{F} = \mathbf{0}$. If the rank condition does not hold then again by Lemma 1 it follows that

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{F}}{T} = \frac{\mathbf{X}_i'\bar{\mathbf{M}}_q\mathbf{F}}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right) + O_p\left(\frac{1}{N}\right),\tag{36}$$

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i}{T} = \frac{\mathbf{X}_i'\bar{\mathbf{M}}_q\mathbf{X}_i}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right) + O_p\left(\frac{1}{N}\right),\tag{37}$$

and

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\boldsymbol{\varepsilon}_i}{T} = \frac{\mathbf{X}_i'\bar{\mathbf{M}}_q\boldsymbol{\varepsilon}_i}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right) + O_p\left(\frac{1}{N}\right).\tag{38}$$

where

$$\bar{\mathbf{M}}_q = \mathbf{I}_T - \bar{\mathbf{Q}}\left(\bar{\mathbf{Q}}'\bar{\mathbf{Q}}\right)^{-}\bar{\mathbf{Q}}', \text{ with } \bar{\mathbf{Q}} = \mathbf{G}\bar{\mathbf{P}}.\tag{39}$$

Using the above results, noting that $T^{-1}\mathbf{X}_i'\bar{\mathbf{M}}_q\mathbf{X}_i = O_p(1)$, and assuming that the rank condition (9) is satisfied we have

$$\hat{\mathbf{b}}_i - \boldsymbol{\beta}_i = \left(\frac{\mathbf{X}_i'\mathbf{M}_g\mathbf{X}_i}{T}\right)^{-1}\left(\frac{\mathbf{X}_i'\mathbf{M}_g\boldsymbol{\varepsilon}_i}{T}\right) + O_p\left(\frac{1}{\sqrt{NT}}\right) + O_p\left(\frac{1}{N}\right).\tag{40}$$

Since $\boldsymbol{\varepsilon}_i$ is independently distributed of $\mathbf{X}_i$ and $\mathbf{G} = (\mathbf{D}, \mathbf{F})$, then for a fixed T, $\lim_{N\to\infty} E\left(\hat{\mathbf{b}}_i - \boldsymbol{\beta}_i\right) = \mathbf{0}$. The finite-$T$ distribution of $\hat{\mathbf{b}}_i - \boldsymbol{\beta}_i$ will be free of nuisance parameters as $N \to \infty$, but will depend on the probability density of $\boldsymbol{\varepsilon}_i$. For N and T sufficiently large, the distribution of $\sqrt{T}\left(\hat{\mathbf{b}}_i - \boldsymbol{\beta}_i\right)$ will be asymptotically normal if the rank condition (9) is satisfied and if N and T are of the same order of magnitudes, namely, if $T/N \to \kappa$ as N and $T \to \infty$, where κ is a positive finite constant. To see why this additional condition is needed, using (40) note that

$$\sqrt{T}\left(\hat{\mathbf{b}}_i - \boldsymbol{\beta}_i\right) = \left(\frac{\mathbf{X}_i'\mathbf{M}_g\mathbf{X}_i}{T}\right)^{-1}\frac{\mathbf{X}_i'\mathbf{M}_g\boldsymbol{\varepsilon}_i}{\sqrt{T}} + O_p\left(\frac{\sqrt{T}}{N}\right) + O_p\left(\frac{1}{\sqrt{N}}\right),\tag{41}$$

and the asymptotic distribution of $\sqrt{T}\left(\hat{\mathbf{b}}_i - \boldsymbol{\beta}_i\right)$ will be free of nuisance parameters only if $\sqrt{T}/N \to 0$, as $(N, T) \xrightarrow{j} \infty$. For this condition to be satisfied it is sufficient that $T/N \to \kappa$, as $(N, T) \xrightarrow{j} \infty$, where κ is a finite non-negative constant as N and $T \to \infty$.

In the case where there are no cointegrating relations amongst the elements of $\mathbf{f}_t$, and $\mathbf{d}_{3t} = \mathbf{0}$, the results simplify since by Lemma 2

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{F}}{T} = \frac{\mathbf{X}_i'\bar{\mathbf{M}}_g\mathbf{F}}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right),\tag{42}$$

10

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i}{T} = \frac{\mathbf{X}_i'\bar{\mathbf{M}}_g\mathbf{X}_i}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right), \tag{43}$$

and

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\varepsilon_i}{T} = \frac{\mathbf{X}_i'\bar{\mathbf{M}}_g\varepsilon_i}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right). \tag{44}$$

Hence, the condition $\sqrt{T}/N \to 0$, as $(N,T) \overset{j}{\to} \infty$ will not be needed for the validity of the asymptotic distribution of $\sqrt{T}\left(\hat{\mathbf{b}}_i - \boldsymbol{\beta}_i\right)$ for large N and T. The following theorem provides a formal statement of these results and the associated asymptotic distributions in the case where the rank condition (9) is satisfied.

Theorem 1 *Consider the panel data model (1) and (2) and suppose that $\|\boldsymbol{\beta}_i\| < K$, $\|\boldsymbol{\Pi}_i\| < K$, Assumptions 1-5 hold. Let $\sqrt{T}/N \to 0$, as $(N,T) \overset{j}{\to} \infty$, and the rank condition (9) be satisfied. (a) - (N-asymptotic) The common correlated effects estimator, $\hat{\mathbf{b}}_i$, defined by (15) is unbiased for a fixed $T > n + 2k + 1$ and $N \to \infty$, in the sense that $\lim_{N\to\infty}E\left(\hat{\mathbf{b}}_i\right) = \boldsymbol{\beta}_i$. Under the additional assumption that $\varepsilon_{it} \sim IIDN(0, \sigma_i^2)$,*

$$\hat{\mathbf{b}}_i - \boldsymbol{\beta}_i \overset{d}{\to} N(\mathbf{0}, \boldsymbol{\Sigma}_{T,b_i}), \tag{45}$$

as $N \to \infty$, where

$$\boldsymbol{\Sigma}_{T,b_i} = T^{-1}\sigma_i^2\boldsymbol{\Psi}_{ig}^{-1}, \quad \boldsymbol{\Psi}_{ig} = T^{-1}\left(\mathbf{X}_i'\mathbf{M}_g\mathbf{X}_i\right), \tag{46}$$

$$\mathbf{M}_g = \mathbf{I}_T - \mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}', \tag{47}$$

and $\mathbf{G} = (\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_T) = (\mathbf{F}, \mathbf{D})$. (b) - (Joint asymptotic) As $(N,T) \overset{j}{\to} \infty$ (in no particular order), $\hat{\mathbf{b}}_i$ is a consistent estimator of $\boldsymbol{\beta}_i$. Then, as $(N,T) \overset{j}{\to} \infty$

$$\sqrt{T}\left(\hat{\mathbf{b}}_i - \boldsymbol{\beta}_i\right) \overset{d}{\to} N(\mathbf{0}, \boldsymbol{\Sigma}_{b_i}), \tag{48}$$

where

$$\boldsymbol{\Sigma}_{b_i} = \sigma_i^2\boldsymbol{\Sigma}_i^{-1}. \tag{49}$$

An asymptotically unbiased estimator of $\boldsymbol{\Sigma}_{T,b_i}$, as $N \to \infty$ for a fixed $T > n + 2k + 1$, is given by:

$$\hat{\boldsymbol{\Sigma}}_{T,b_i} = \hat{\sigma}_i^2\left(\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i\right)^{-1}, \tag{50}$$

where

$$\hat{\sigma}_i^2 = \frac{\left(\mathbf{y}_i - \mathbf{X}_i\hat{\mathbf{b}}_i\right)'\bar{\mathbf{M}}\left(\mathbf{y}_i - \mathbf{X}_i\hat{\mathbf{b}}_i\right)}{T - (n + m + k)}. \tag{51}$$

In the case where $(N,T) \overset{j}{\to} \infty$, a consistent estimator of $\boldsymbol{\Sigma}_{b_i}$ is given by

$$\hat{\boldsymbol{\Sigma}}_{b_i} = \mathring{\sigma}_i^2\left(\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i}{T}\right)^{-1}, \tag{52}$$

where

$$\mathring{\sigma}_i^2 = \frac{\left(\mathbf{y}_i - \mathbf{X}_i\hat{\mathbf{b}}_i\right)' \bar{\mathbf{M}} \left(\mathbf{y}_i - \mathbf{X}_i\hat{\mathbf{b}}_i\right)}{T - (n + 2k + 1)}. \tag{53}$$

Remark 3 *It is worth noting that despite the fact that under our Assumptions* $\mathbf{f}_t$, y_{it} *and* $\mathbf{x}_{it}$ *are* $I(1)$ *and cointegrated, in the results of Theorem 1 the rate of convergence of* $\hat{\mathbf{b}}_i$ *to* $\boldsymbol{\beta}_i$ *as* $(N,T) \xrightarrow{j} \infty$ *is* $\sqrt{T}$ *and not* T. *It is helpful to develop some intuition behind this result. Since for* N *sufficiently large* $\mathbf{f}_t$ *can be well approximated by the cross section averages we might as well consider the case where* $\mathbf{f}_t$ *is observed. Abstracting from* $\mathbf{d}_t$, *and without loss of generality substitute (2) in (1) to obtain*

$$y_{it} = \boldsymbol{\beta}_i' \left(\boldsymbol{\Gamma}_i'\mathbf{f}_t + \mathbf{v}_{it}\right) + \boldsymbol{\gamma}_i'\mathbf{f}_t + \varepsilon_{it} = \boldsymbol{\vartheta}_i'\mathbf{f}_t + \varkappa_{it}$$

where $\boldsymbol{\vartheta}_i = \boldsymbol{\Gamma}_i\boldsymbol{\beta}_i + \boldsymbol{\gamma}_i$ *and* $\varkappa_{it} = \varepsilon_{it} + \boldsymbol{\beta}_i'\mathbf{v}_{it}$. *It is clear that if* $\mathbf{f}_t$ *is* $I(1)$, *as postulated under our assumptions, then for all values of* $\boldsymbol{\beta}_i$, $\varkappa_{it}$ *is* $I(0)$ *and* y_{it} *and* $\mathbf{f}_t$ *will be* $I(1)$ *and cointegrated. Hence, it is easily seen that* $\boldsymbol{\vartheta}_i$ *can be estimated superconsistently. However, the OLS estimator of* $\boldsymbol{\beta}_i$ *need not be superconsistent. To see this note that* $\boldsymbol{\beta}_i$ *can equivalently be estimated by regressing the residuals of* y_{it} *on* $\mathbf{f}_t$ *on the residuals of* $\mathbf{x}_{it}$ *on* $\mathbf{f}_t$. *Both these sets of residuals are stationary processes and the resulting estimator of* $\boldsymbol{\beta}_i$ *will be* $\sqrt{T}$-consistent.

Remark 4 *When the rank condition, (9), is not satisfied consistent estimation of the individual slope coefficients is not possible.*

3.1   Pooled Estimators

We now examine the asymptotic properties of the pooled estimators. Focusing first on the MG estimator, and using (29) we have

$$\sqrt{N}\left(\hat{\mathbf{b}}_{MG} - \boldsymbol{\beta}\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{v}_i + \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\Psi}}_{iT}^{-1}\left(\frac{\sqrt{N}\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{F}}{T}\right)\boldsymbol{\gamma}_i +$$
$$\frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\Psi}}_{iT}^{-1}\left(\frac{\sqrt{N}\mathbf{X}_i'\bar{\mathbf{M}}\varepsilon_i}{T}\right), \tag{54}$$

In the case where the rank condition (9) is satisfied, again by Lemmas 1-3 we have

$$\frac{\sqrt{N}\left(\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{F}\right)}{T} = O_p\left(\frac{1}{\sqrt{T}}\right),$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\Psi}}_{iT}^{-1}\left(\frac{\sqrt{N}\mathbf{X}_i'\bar{\mathbf{M}}\varepsilon_i}{T}\right) = O_p\left(\frac{1}{\sqrt{T}}\right),$$

Thus

$$\sqrt{N}\left(\hat{\mathbf{b}}_{MG} - \boldsymbol{\beta}\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{v}_i + O_p\left(\frac{1}{\sqrt{T}}\right).$$

Hence

$$\sqrt{N}\left(\hat{\mathbf{b}}_{MG} - \boldsymbol{\beta}\right) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_{MG}), \text{ as } (N,T) \xrightarrow{j} \infty. \tag{55}$$

The variance estimator for $\boldsymbol{\Sigma}_{MG}$ suggested by Pesaran (2006) is given by

$$\hat{\boldsymbol{\Sigma}}_{MG} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\mathbf{b}}_i - \hat{\mathbf{b}}_{MG}\right)\left(\hat{\mathbf{b}}_i - \hat{\mathbf{b}}_{MG}\right)', \tag{56}$$

can be used. It can also be shown along identical lines to Pesaran (2006) that the mean group estimator is valid even if the rank condition is not satisfied. The following theorem summarises the results for the mean group estimator.

Theorem 2 *Consider the panel data model (1) and (2) and suppose that Assumptions 1-5 hold. Then the Common Correlated Effects Mean Group estimator, $\mathbf{b}_{MG}$ defined by (14), is asymptotically (for a fixed T and as $N \to \infty$) unbiased for $\boldsymbol{\beta}$, and as $(N,T) \xrightarrow{j} \infty$*

$$\sqrt{N}\left(\hat{\mathbf{b}}_{MG} - \boldsymbol{\beta}\right) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_{MG}),$$

where $\boldsymbol{\Sigma}_{MG}$ is consistently estimated by (56).

This theorem does not require that the rank condition, (9), holds for any number, m, of unobserved factors so long as m is fixed, and does not impose any restrictions on the relative rates of expansion of N and T. But in the case where the rank condition is satisfied Assumption 3 can be relaxed and the factor loadings, $\boldsymbol{\gamma}_i$, need not follow the random coefficient model. It would be sufficient that they are bounded.

Moving to the pooled estimator we have the following. $\hat{\mathbf{b}}_P$ defined by (17), can be written as

$$\sqrt{N}\left(\hat{\mathbf{b}}_P - \boldsymbol{\beta}\right) = \left(\frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i}{T}\right)^{-1}\left[\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{X}_i'\bar{\mathbf{M}}(\mathbf{X}_i\boldsymbol{v}_i + \boldsymbol{\varepsilon}_i)}{T} + \mathbf{q}_{NT}\right]. \tag{57}$$

where

$$\mathbf{q}_{NT} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{(\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{F})\boldsymbol{\gamma}_i}{T}. \tag{58}$$

Assuming random coefficients we note that $\boldsymbol{\gamma}_i = \bar{\boldsymbol{\gamma}} + \boldsymbol{\eta}_i - \bar{\boldsymbol{\eta}}$, where $\bar{\boldsymbol{\eta}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\eta}_i$. Hence

$$\mathbf{q}_{NT} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{F}}{T}\right)(\bar{\boldsymbol{\gamma}} - \bar{\boldsymbol{\eta}}) + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{F}}{T}\right)\boldsymbol{\eta}_i.$$

Recall that when the rank condition is not satisfied $T^{-1}\left(\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{F}\right) = O_p(1)$. But since $\bar{\mathbf{X}}'\bar{\mathbf{M}} = \mathbf{0}$, we have

$$\sum_{i=1}^{N} \mathbf{X}_i'\bar{\mathbf{M}}\mathbf{F}\left(\bar{\boldsymbol{\gamma}} - \bar{\boldsymbol{\eta}}\right) = N\bar{\mathbf{X}}'\bar{\mathbf{M}}\mathbf{F}\left(\bar{\boldsymbol{\gamma}} - \bar{\boldsymbol{\eta}}\right) = \mathbf{0},$$

and it follows that

$$\mathbf{q}_{NT} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{F}}{T}\right)\boldsymbol{\eta}_i \tag{59}$$

$$= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\bar{\mathbf{M}}_q\mathbf{F}}{T}\right)\boldsymbol{\eta}_i + O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{\sqrt{T}}\right),$$

Substituting this result in (57), and making use of (34) and (35) we have

$$\sqrt{N}\left(\hat{\mathbf{b}}_P - \boldsymbol{\beta}\right) = \left(\frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{X}_i'\bar{\mathbf{M}}_q\mathbf{X}_i}{T}\right)^{-1}\left[\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{X}_i'\bar{\mathbf{M}}_q(\mathbf{X}_i\boldsymbol{v}_i + \boldsymbol{\varepsilon}_i + \mathbf{F}\boldsymbol{\eta}_i)}{T}\right] + \tag{60}$$
$$O_p\left(\frac{1}{\sqrt{T}}\right).$$

Hence, as $(N,T) \xrightarrow{j} \infty$

$$\sqrt{N}\left(\hat{\boldsymbol{b}} - \boldsymbol{\beta}\right) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_P^*),$$

where

$$\boldsymbol{\Sigma}_P^* = \boldsymbol{\Psi}^{*-1}\mathbf{R}^*\boldsymbol{\Psi}^{*-1}, \tag{61}$$

$$\boldsymbol{\Psi}^* = \lim_{N\to\infty}\left(N^{-1}\sum_{i=1}^{N}\boldsymbol{\Sigma}_{iq}\right), \quad \mathbf{R}^* = \lim_{N\to\infty}\left[N^{-1}\sum_{i=1}^{N}\left(\boldsymbol{\Sigma}_{iq}\boldsymbol{\Omega}_v\boldsymbol{\Sigma}_{iq} + \mathbf{Q}_{if}\boldsymbol{\Omega}_\eta\mathbf{Q}_{if}'\right)\right], \tag{62}$$

and $\boldsymbol{\Sigma}_{iq}$ and $\mathbf{Q}_{if}$ are defined by

$$\boldsymbol{\Sigma}_{iq} = p\lim_{T\to\infty}\left(T^{-1}\mathbf{X}_i'\bar{\mathbf{M}}_q\mathbf{X}_i\right) \text{ and } \mathbf{Q}_{if} = p\lim_{T\to\infty}\left(T^{-1}\mathbf{X}_i'\bar{\mathbf{M}}_q\mathbf{F}\right). \tag{63}$$

The variance estimator for $\boldsymbol{\Sigma}_P^*$ suggested by Pesaran (2006) and given by

$$\hat{\boldsymbol{\Sigma}}_P^* = \hat{\boldsymbol{\Psi}}^{*-1}\hat{\mathbf{R}}^*\hat{\boldsymbol{\Psi}}^{*-1}, \tag{64}$$

where

$$\hat{\boldsymbol{\Psi}}^* = N^{-1}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i}{T}\right), \tag{65}$$

$$\hat{\mathbf{R}}^* = \frac{1}{(N-1)}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i}{T}\right)\left(\hat{\mathbf{b}}_i - \hat{\mathbf{b}}_{MG}\right)\left(\hat{\mathbf{b}}_i - \hat{\mathbf{b}}_{MG}\right)'\left(\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i}{T}\right). \tag{66}$$

Again we summarise these results in the following theorem.

Theorem 3 *Consider the panel data model (1) and (2) and suppose that Assumptions 1-5 hold. Then the Common Correlated Effects Pooled estimator, $\hat{\mathbf{b}}_P$, defined by (17) is asymptotically unbiased for $\boldsymbol{\beta}$, and as $(N, T) \xrightarrow{j} \infty$ we have*

$$\sqrt{N} \left(\hat{\mathbf{b}}_P - \boldsymbol{\beta} \right) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_P^*),$$

where $\boldsymbol{\Sigma}_P^*$ is given by (61)-(63).

Overall we see that despite a number of differences in the above analysis, especially in terms of the results given in (33)-(35), compared to the results in Pesaran (2006), the conclusions are remarkably similar when the factors are assumed to follow unit root processes rather that weakly stationary ones.

4 Monte Carlo Design and Evidence

In this section we provide Monte Carlo evidence on the small sample properties of the CCEMG and the CCEP estimators. We also consider the two alternative principal component augmentation approaches discussed in Kapetanios and Pesaran (2006). The first PC approach applies the Bai and Ng (2002) procedure to $\mathbf{z}_{it} = (y_{it}, \mathbf{x}_{it}')'$ to obtain consistent estimates of the unobserved factors, and then uses the estimated factors to augment the regression (1), and thus produces consistent estimates of $\boldsymbol{\beta}$. We consider both pooled and mean group versions of this estimator which we refer to as $PC1POOL$ and $PC1MG$. The second PC approach consists of extracting the principal component estimates of the unobserved factors from y_{it} and $\mathbf{x}_{it}$ separately, regressing y_{it} and $\mathbf{x}_{it}$ on their respective factor estimates separately, and then applying the standard pooled and mean group estimators, with no cross-sectional dependence adjustments, to the residuals of these regressions. We refer to the estimators based on this approach as $PC2POOL$ and $PC2MG$, respectively.

The experimental design of the Monte Carlo study is closely related to the one used in Pesaran (2006). Consider the following data generating process (DGP):

$$y_{it} = \alpha_{i1}d_{1t} + \beta_{i1}x_{1it} + \beta_{i2}x_{2it} + \gamma_{i1}f_{1t} + \gamma_{i2}f_{2t} + \varepsilon_{it}, \tag{67}$$

and

$$x_{ijt} = a_{ij1}d_{1t} + a_{ij2}d_{2t} + \gamma_{ij1}f_{1t} + \gamma_{ij3}f_{3t} + \mathrm{v}_{ijt}, \quad j = 1, 2, \tag{68}$$

for $i = 1, 2, ..., N$, and $t = 1, 2, ..., T$. This DGP is a restricted version of the general linear model considered in Pesaran (2006), and sets $n = k = 2$, and $m = 3$, with $\boldsymbol{\alpha}_i' = (\alpha_{i1}, 0)$, $\boldsymbol{\beta}_i' = (\beta_{i1}, \beta_{i2})$, and $\boldsymbol{\gamma}_i' = (\gamma_{i1}, \gamma_{i2}, 0)$, and

$$\mathbf{A}_i' = \begin{pmatrix} a_{i11} & a_{i12} \\ a_{i21} & a_{i22} \end{pmatrix}, \quad \boldsymbol{\Gamma}_i' = \begin{pmatrix} \gamma_{i11} & 0 & \gamma_{i13} \\ \gamma_{i21} & 0 & \gamma_{i23} \end{pmatrix}.$$

15

The observed common factors and the individual specific errors of $\mathbf{x}_{it}$ are generated as independent stationary AR(1) processes with zero means and unit variances:

$$d_{1t} = 1, \ d_{2t} = \rho_d d_{2,t-1} + v_{dt}, \ t = -49, ...1, ..., T,$$

$$v_{dt} \sim IIDN(0, 1 - \rho_d^2), \ \rho_d = 0.5, \ d_{2,-50} = 0,$$

$$\mathrm{v}_{ijt} = \rho_{vij}\mathrm{v}_{ijt-1} + \upsilon_{ijt}, \ t = -49, ...1, ..., T,$$

$$\upsilon_{ijt} \sim IIDN\left(0, 1 - \rho_{vij}^2\right), \ \mathrm{v}_{ji,-50} = 0,$$

and

$$\rho_{vij} \sim IIDU\left[0.05, 0.95\right], \ \text{for } j = 1, 2,$$

but the unobserved common factors are generated as non-stationary processes:

$$f_{jt} = f_{jt-1} + v_{fj,t}, \ \text{for } j = 1, 2, 3, \ t = -49, .., 0, .., T,$$

$$v_{fj,t} \sim IIDN(0, 1), \ f_{j,-50} = 0, \ \text{for } j = 1, 2, 3.$$

The first 50 observations are discarded.

To illustrate the robustness of the CCE and PC estimators to the dynamics of the individual specific errors of y_{it}, these are generated as the (cross sectional) mixture of stationary heterogeneous AR(1) and MA(1) errors. Namely,

$$\varepsilon_{it} = \rho_{i\varepsilon}\varepsilon_{i,t-1} + \sigma_i\sqrt{1 - \rho_{i\varepsilon}^2}\omega_{it}, \ i = 1, 2, ..., N_1, \ t = -49, .., 0, .., T,$$

and

$$\varepsilon_{it} = \frac{\sigma_i}{\sqrt{1 + \theta_{i\varepsilon}^2}}\left(\omega_{it} + \theta_{i\varepsilon}\omega_{i,t-1}\right), \ i = N_1 + 1, ..., N, \ t = -49, .., 0, .., T,$$

where N_1 is the nearest integer of $N/2,$

$$\omega_{it} \sim IIDN\left(0, 1\right), \ \sigma_i^2 \sim IIDU\left[0.5, 1.5\right], \ \rho_{i\varepsilon} \sim IIDU\left[0.05, 0.95\right], \ \theta_{i\varepsilon} \sim IIDU\left[0, 1\right].$$

ρ_{vij}, $\rho_{i\varepsilon}$, $\theta_{i\varepsilon}$ and σ_i are not changed across replications. The first 49 observations are discarded. The factor loadings of the observed common effects, α_{i1}, and $vec(\mathbf{A}_i) = (a_{i11}, a_{i21}, a_{i12}, a_{i22})'$ are generated as $IIDN(1, 1)$, and $IIDN(0.5\tau_4, 0.5 \ \mathbf{I}_4)$, where $\tau_4 = (1, 1, 1, 1)'$, and are not changed across replications. They are treated as fixed effects. The parameters of the unobserved common effects in the $\mathbf{x}_{it}$ equation are generated independently across replications as

$$\mathbf{\Gamma}_i' = \begin{pmatrix} \gamma_{i11} & 0 & \gamma_{i13} \\ \gamma_{i21} & 0 & \gamma_{i23} \end{pmatrix} \sim IID \begin{pmatrix} N\left(0.5, 0.50\right) & 0 & N\left(0, 0.50\right) \\ N\left(0, 0.50\right) & 0 & N\left(0.5, 0.50\right) \end{pmatrix}.$$

For the parameters of the unobserved common effects in the y_{it} equation, $\boldsymbol{\gamma}_i$, we considered two different sets that we denote by $\mathcal{A}$ and $\mathcal{B}$. Under set $\mathcal{A}$, $\boldsymbol{\gamma}_i$ are drawn such that the rank condition is satisfied, namely

$$\gamma_{i1} \sim IIDN\left(1, 0.2\right), \ \gamma_{i2\mathcal{A}} \sim IIDN\left(1, 0.2\right), \ \gamma_{i3} = 0,$$

and

$$E\left(\tilde{\mathbf{\Gamma}}_{i\mathcal{A}}\right) = \left(E\left(\boldsymbol{\gamma}_{i\mathcal{A}}\right), E\left(\mathbf{\Gamma}_i\right)\right) = \begin{pmatrix} 1 & 0.5 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0.5 \end{pmatrix}.$$

Under set $\mathcal{B}$

$$\gamma_{i1} \sim IIDN\left(1, 0.2\right), \ \gamma_{i2\mathcal{B}} \sim IIDN\left(0, 1\right), \gamma_{i3} = 0,$$

so that

$$E\left(\tilde{\mathbf{\Gamma}}_{i\mathcal{B}}\right) = \left(E\left(\boldsymbol{\gamma}_{i\mathcal{B}}\right), E\left(\mathbf{\Gamma}_i\right)\right) = \begin{pmatrix} 1 & 0.5 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.5 \end{pmatrix},$$

and the rank condition is *not* satisfied. For each set we conducted two different experiments:

- **Experiment 1** examines the case of heterogeneous slopes with $\beta_{ij} = 1 + \eta_{ij}$, $j = 1, 2$, and $\eta_{ij} \sim IIDN(0, 0.04)$, across replications.

- **Experiment 2** considers the case of homogeneous slopes with $\boldsymbol{\beta}_i = \boldsymbol{\beta} = (1, 1)'$.

The two versions of experiment 1 will be denoted by $1\mathcal{A}$ and $1\mathcal{B}$, and those of experiment 2 by $2\mathcal{A}$ and $2\mathcal{B}$. For this Monte Carlo study we also computed the CCEMG and the CCEP estimators as well as the associated "infeasible" estimators (MG and Pooled) that include f_{1t} and f_{2t} in the regressions of y_{it} on $(d_{1t}, \mathbf{x}_{it})$, and the "naive" estimators that excludes these factors. The naive estimators illustrate the extent of bias and size distortions that can occur if the error cross section dependence is ignored.

In relation to the infeasible pooled estimator, it is important to note that this estimator although unbiased under all the four sets of experiments, it need not be efficient since in these experiments the slope coefficients, $\boldsymbol{\beta}_i$, and/or error variances, σ_i^2, differ across i. As a result the CCE or PC augmented estimators may in fact dominate the infeasible estimator in terms of RMSE, particularly in the case of experiments $1\mathcal{A}$ and $1\mathcal{B}$ where the slopes as well as the error variances are allowed to vary across i.

Another important consideration worth bearing in mind when comparing the CCE and the PC type estimators is the fact that the computation of the PC augmented estimators assumes that $m = 3$, the number of unobserved factors, is known. In practice, m might be difficult to estimate accurately particularly when N or T happen to be smaller than 50.

By contrast the CCE type estimators are valid for any fixed m and do not require a prior estimate for m.

Each experiment was replicated 2000 times for the (N, T) pairs with $N, T = 20, 30, 50, 100, 200$. In what follows we shall focus on β_1 (the cross section mean of β_{i1}). Results for β_2 are very similar and will not be reported. Finally, for completeness we state below the exact formulae for the variance estimators used for the different estimators. The non-parametric variance estimators of the mean group estimators, $\tilde{\mathbf{b}}_{MG} = N^{-1} \sum_{i=1}^{N} \tilde{\mathbf{b}}_i$, are computed as

$$\widehat{Var}(\tilde{\mathbf{b}}_{MG}) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \left(\tilde{\mathbf{b}}_i - \tilde{\mathbf{b}}_{MG} \right) \left(\tilde{\mathbf{b}}_i - \tilde{\mathbf{b}}_{MG} \right)', \tag{69}$$

where

$$\tilde{\mathbf{b}}_i = \left(\mathbf{X}_i' \tilde{\mathbf{M}}_x \mathbf{X}_i \right)^{-1} \mathbf{X}_i' \tilde{\mathbf{M}}_x \tilde{\mathbf{M}}_y \mathbf{y}_i,$$

$$\tilde{\mathbf{M}}_x = \mathbf{I}_T - \tilde{\mathbf{H}}_x \left(\tilde{\mathbf{H}}_x' \tilde{\mathbf{H}}_x \right)^{-} \tilde{\mathbf{H}}_x', \quad \tilde{\mathbf{M}}_y = \mathbf{I}_T - \tilde{\mathbf{H}}_y \left(\tilde{\mathbf{H}}_y' \tilde{\mathbf{H}}_y \right)^{-} \tilde{\mathbf{H}}_y'.$$

For the CCEMG estimator, $\tilde{\mathbf{H}}_x = \tilde{\mathbf{H}}_y = \bar{\mathbf{H}} = (\mathbf{D}, \bar{\mathbf{Z}})$, so that $\tilde{\mathbf{b}}_i = \hat{\mathbf{b}}_i$, which is defined by (15); for the PC1MG estimator, $\tilde{\mathbf{H}}_x = \tilde{\mathbf{H}}_y = \hat{\mathbf{F}}_z$, where $\hat{\mathbf{F}}_z$ is a $T \times (n + m)$ matrix of extracted factors from $\mathbf{Z}_i$ for all i, together with observed common factors; for the PC2MG estimator $\tilde{\mathbf{H}}_x = \hat{\mathbf{F}}_x$ and $\tilde{\mathbf{H}}_y = \hat{\mathbf{F}}_y$, where $\hat{\mathbf{F}}_x$ and $\hat{\mathbf{F}}_y$ are $T \times (n_x + m_x)$ and $T \times (n_y + m_y)$ matrices of extracted factors from $\mathbf{X}_i$ and $\mathbf{y}_i$ respectively for all i, together with the observed common factors with n_x and n_y being the number of observed common factors in $\mathbf{X}_i$ and $\mathbf{y}_i$ respectively, and m_x and m_y defined similarly; for the infeasible mean group estimator, $\tilde{\mathbf{H}}_x = \tilde{\mathbf{H}}_y = \mathbf{F}_y$, which is a $T \times m_y$ matrix of unobserved factors in $\mathbf{y}_i$; for the naive mean group estimator, $\tilde{\mathbf{H}}_x = \tilde{\mathbf{H}}_y = \mathbf{D}$. Next, the non-parametric variance of the pooled estimator, $\tilde{\mathbf{b}}_P$, is computed as

$$\widehat{Var}(\tilde{\mathbf{b}}_P) = N^{-1} \tilde{\boldsymbol{\Psi}}^{-1} \tilde{\mathbf{R}} \tilde{\boldsymbol{\Psi}}^{-1}, \tag{70}$$

where

$$\tilde{\mathbf{b}}_P = \left(\sum_{i=1}^{N} \mathbf{X}_i' \tilde{\mathbf{M}}_x \mathbf{X}_i \right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i' \tilde{\mathbf{M}}_x \tilde{\mathbf{M}}_y \mathbf{y}_i,$$

$$\tilde{\boldsymbol{\Psi}} = N^{-1} \sum_{i=1}^{N} \left(\frac{\mathbf{X}_i' \tilde{\mathbf{M}}_x \mathbf{X}_i}{T} \right),$$

$$\tilde{\mathbf{R}} = \frac{1}{(N-1)} \sum_{i=1}^{N} \left(\frac{\mathbf{X}_i' \tilde{\mathbf{M}}_x \mathbf{X}_i}{T} \right) \left(\tilde{\mathbf{b}}_i - \tilde{\mathbf{b}}_{MG} \right) \left(\tilde{\mathbf{b}}_i - \tilde{\mathbf{b}}_{MG} \right)' \left(\frac{\mathbf{X}_i' \tilde{\mathbf{M}}_x \mathbf{X}_i}{T} \right).$$

4.1 Results

Results of experiments $1\mathcal{A}$, $2\mathcal{A}$, $1\mathcal{B}$, $2\mathcal{B}$ are summarized in Tables 1 to 4, respectively. We also provide results for the naive estimator (that excludes the unobserved factors or their

estimates) and the infeasible estimator (that includes the unobserved factors as additional regressors) for comparison purposes. But for the sake of brevity we include the simulation results for these estimators only for experiment 1$\mathcal{A}$ reported in Table 1.

As can be seen from Table 1 the naive estimator is substantially biased, performs very poorly and is subject to large size distortions; an outcome that continues to apply in the case of other experiments (not reported here). In contrast, the feasible CCE estimators perform well, have bias that are close to the bias of the infeasible estimators, show little size distortions even for relatively small values of N and T, and their RMSE falls steadily with increases in N and/or T. These results are quite similar to the results presented in Pesaran (2006), and illustrate the robustness of the CCE estimators to the presence of unit roots in the unobserved common factors. This is important since it obviates the need for pre-testing involving unobserved factors.

The CCE estimators perform well, in both heterogeneous and homogeneous cases, and irrespective of whether the rank condition is satisfied, although the CCE estimators with rank deficiency have sightly higher RMSEs than those with full rank. The RMSEs of the CCE estimators of Tables 1 and 3 (heterogeneous case) are higher than those reported in Tables 2 and 4 for the homogeneous case. The sizes of the t-test based on the CCE estimators are very close to the nominal 5% level. In the case of full rank, the power of the tests for the CCE estimators are much higher than in the rank deficient case. Finally, not surprisingly the power of the tests for the CCE estimators in the homogeneous case is higher than that in the heterogeneous case.

It is also important to note that the small sample properties of the CCE estimator does not seem to be much affected by the residual serial correlation of the idiosyncratic errors, ε_{it}. The robustness of the CCE estimator to the short run dynamics is particularly helpful in practice where typically little is known about such dynamics. In fact a comparison of the results for the CCEP estimator with the infeasible counterpart given in Table 1 shows that the former can even be more efficient (in the RMSE sense). For example the RMSE of the CCEP for $N = T = 50$ is 3.97 whilst the RMSE of the infeasible pooled estimator is 4.31. This might seem counter intuitive at first, but as indicated above the infeasible estimator does not take account of the residual serial correlation of the idiosyncratic errors, but the CCE estimator does allow for such possibilities indirectly through the use of the cross section averages that partly embody the serial correlation properties of $\mathbf{f}_t$ and ε_{it}'s.

Consider now the PC augmented estimators and recall that they are computed assuming the true number of common factors is known. The results summarized in Tables 1-4 bear some resemblance to those presented in Kapetanios and Pesaran (2006). The bias and RMSEs of the PC1POOL and PC1MG estimators improve as both N and T increase, but

the t-tests based on these estimators substantially over-reject the null hypothesis. The PC2POOL and PC2MG estimators perform even worse. The biases of the PC estimators are always larger in absolute value than the respective biases of the CCE estimators. The size distortion of the PC augmented estimators is particularly pronounced in the case of the experiments $1\mathcal{A}$ and $2\mathcal{A}$ (in Tables 1 and 2) where the full rank conditions are met. It is also interesting that in the case of some of the experiments the bias distortions of the tests based on the PC augmented estimators do not improve even for relatively large N and T. An interesting distinction arises when comparing results for experiments $1\mathcal{A}$ and $1\mathcal{B}$. For $1\mathcal{A}$ (heterogeneous slopes and full rank) results are very poor for small values of N and T but improve considerably as N rises and less perceptibly as T rises. For experiment $1\mathcal{B}$ (heterogeneous slopes and rank deficient) results are much better for small values of N and T. Finally, it is worth noting that in both cases the performance of the PC estimators actually get worse when N is small and kept small but T rises. This may be related to the fact that the accuracy of the factor estimates depends on the minimum of N and T.

5   Conclusions

Recently, there has been increased focus in the panel data literature on problems arising in estimation and inference when the standard assumption that the errors of the panel regression are cross-sectionally uncorrelated, is violated. When the errors of a panel regression are cross-sectionally correlated then standard estimation methods do not necessarily produce consistent estimates of the parameters of interest. An influential strand of the relevant literature provides a convenient parametrisation of the problem in terms of a factor model for the error terms.

Pesaran (2006) adopts an error multifactor structure and suggests new estimators that take into account cross-sectional dependence, making use of cross-sectional averages of the dependent and explanatory variables. However, he focusses on the case of weakly stationary factors that could be restrictive in some applications. This paper provides a formal extension of the results of Pesaran (2006) to the case where the unobserved factors are allowed to follow unit root processes. It is shown that the main results of Pesaran continue to hold in this more general case. This is certainly of interest given the fact that usually there is a large difference between results obtained for unit root and stationary processes. When we consider the small sample properties of the new estimators, we observe that again the results accord with the conclusions reached in the stationary case, lending further support to the use of the CCE estimators irrespective of the order of integration of the data observed.

References

ANSELIN, L. (2001): "Spatial Econometrics," in *A Companion to Theoretical Econometrics*, ed. by B. Baltagi. Blackwell, Oxford.

BAI, J. (2003): "Inferential Theory for Factor Models of Large Dinensions," *Econometrica*, 71, 135–173.

BAI, J., AND S. NG (2002): "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70, 191–221.

CHAMBERLAIN, G., AND M. ROTHSCHILD (1983): "Arbitrage, Factor Structure and Mean-Variance Analysis in Large Asset Markets," *Econometrica*, 51, 1305–1324.

COAKLEY, J., A. FUERTES, AND R. P. SMITH (2002): "A Principal Components Approach to Cross-Section Dependence in Panels," *Mimeo, Birkbeck College, University of London.*

CONLEY, T. G., AND B. DUPOR (2003): "A Spatial Analysis of Sectoral Complementarity," *Journal of Political Economy*, 111, 311–352.

CONLEY, T. G., AND G. TOPA (2002): "Socio-economic Distance and Spatial Patterns in Unemployment," *Journal of Applied Econometrics*, 17, 303–327.

CONNOR, G., AND R. KORAJZCYK (1986): "Performance Measurement with the Arbitrage Pricing Theory: A New Framework for Analysis," *Journal of Financial Economics*, 15, 373–394.

———— (1988): "Risk and Return in an Equilibrium APT: Application to a new test methodology," *Journal of Financial Economics*, 21, 255–289.

DAVIDSON, J., AND R. DE JONG (2000): "The Functional Central Limit Theorem and Weak Convergence to Stochastic Integrals I: Weakly Dependent Processes," *Econometric Theory*, 16, 621–642.

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.

FORNI, M., AND L. REICHLIN (1998): "Let's Get Real: A Factor Analytical Approach to Disaggregated Business Cycle Dynamics," *Review of Economic Studies*, 65, 453–473.

KAPETANIOS, G., AND M. H. PESARAN (2006): "Alternative Approaches to Estimation and Inference in Large Multifactor Panels: Small Sample Results with an Application to

Modelling of Asset Returns," in *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis*, ed. by G. D. A. Phillips, and E. Tzavalis. Cambridge University Press, Cambridge.

LEE, K. C., AND M. H. PESARAN (1993): "The Role of Sectoral Interactions in Wage Determination in the UK Economy," *The Economic Journal*, 103, 21–55.

PESARAN, M. H. (2006): "Estimation and Inference in Large Heterogeneous panels with a Multifactor Error Structure," *Econometrica*, 74, 967–1012.

PESARAN, M. H., T. SCHUERMANN, AND S. M. WEINER (2004): "Modeling Regional Interdependencies using a Global Error-Correcting Macroeconomic Model," *Journal of Business Economics and Statistics, (with Discussions and a Rejoinder)*, 22, 129–181.

PESARAN, M. H., AND R. P. SMITH (1995): "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics*, 68, 79–113.

STOCK, J. H., AND M. W. WATSON (2002): "Macroeconomic Forecasting Using Diffusion Indices," *Journal of Business and Economic Statistics*, 20, 147–162.

STONE, R. (1947): "On the Interdependence of Blocks of Transactions," *Supplement of the Journal of the Royal Statistical Society*, 9, 1–45.

6   Appendix

Lemma 1 *Under Assumptions 1-5 and if $\sum_{\ell=0}^{\infty} \mathbf{W}_\ell$ is of reduced rank where $\mathbf{W}_\ell$, $\ell = 0, ...,$ is defined in Remark 2, then as $(N,T) \overset{j}{\to} \infty$,*

$$\frac{\mathbf{X}_i' \bar{\mathbf{M}} \mathbf{X}_i}{T} = \frac{\mathbf{X}_i' \bar{\mathbf{M}}_q \mathbf{X}_i}{T} + O_p \left(\frac{1}{\sqrt{NT}} \right) + O_p \left(\frac{1}{N} \right), \tag{71}$$

$$\frac{\mathbf{X}_i' \bar{\mathbf{M}} \mathbf{F}}{T} = \frac{\mathbf{X}_i' \bar{\mathbf{M}}_q \mathbf{F}}{T} + O_p \left(\frac{1}{\sqrt{NT}} \right) + O_p \left(\frac{1}{N} \right), \tag{72}$$

$$\frac{\mathbf{X}_i' \bar{\mathbf{M}} \boldsymbol{\varepsilon}_i}{T} = \frac{\mathbf{X}_i' \bar{\mathbf{M}}_q \boldsymbol{\varepsilon}_i}{T} + O_p \left(\frac{1}{\sqrt{NT}} \right) + O_p \left(\frac{1}{N} \right). \tag{73}$$

Proof. We start by proving (71). The proof for (72) and (73) follow similarly. We need to determine the order of probability of $\frac{\mathbf{X}_i' \bar{\mathbf{M}} \mathbf{X}_i}{T} - \frac{\mathbf{X}_i' \bar{\mathbf{M}}_q \mathbf{X}_i}{T}$. We split the problem in two parts. By remark 2, $\mathbf{f}_t$ may be represented, via a common trends representation, by a set of r stationary components and $m - r$ random walk components. Without loss of generality we disregard the deterministic component $\mathbf{d}_{1t}$ in the analysis. Let $\mathbf{K}_1$ denote the matrix of observations on the nonstationary components of the common trends representation of $\mathbf{f}_t$ and the nonstationary components of $\mathbf{d}_t$. Let $\mathbf{K}_2$ denote the matrix of observations on the stationary components of the common trends representation of $\mathbf{f}_t$ and the stationary components of $\mathbf{d}_t$. Note that the transformations needed to get $(\mathbf{K}_1, \mathbf{K}_2)$ from $\mathbf{G}$ simply involve nonsingular rotations of $\mathbf{f}_t$. Let

$$\bar{\mathbf{H}}_j = \mathbf{K}_j \bar{\mathbf{P}}_j + \bar{\mathbf{U}}^*,$$

$$\bar{\mathbf{M}}_{jq} = \mathbf{I}_T - \bar{\mathbf{Q}}_j \left(\bar{\mathbf{Q}}_j' \bar{\mathbf{Q}}_j \right)^{-} \bar{\mathbf{Q}}_j',$$

$$\bar{\mathbf{M}}_j = \mathbf{I}_T - \bar{\mathbf{H}}_j \left(\bar{\mathbf{H}}_j' \bar{\mathbf{H}}_j \right)^{-} \bar{\mathbf{H}}_j,$$

with $\bar{\mathbf{Q}}_j = \mathbf{K}_j \bar{\mathbf{P}}_j$ for $j = 1, 2$ where $\bar{\mathbf{P}}_j$ is a nonsingular transformation of a partition of $\bar{\mathbf{P}}$ comformable to $(\mathbf{K}_1, \mathbf{K}_2)$. Then, the order of probability of $\frac{\mathbf{X}_i' \bar{\mathbf{M}} \mathbf{X}_i}{T} - \frac{\mathbf{X}_i' \bar{\mathbf{M}}_q \mathbf{X}_i}{T}$ is equal to the maximum of the orders of probability of $\frac{\mathbf{X}_i' \bar{\mathbf{M}}_1 \mathbf{X}_i}{T} - \frac{\mathbf{X}_i' \bar{\mathbf{M}}_{1q} \mathbf{X}_i}{T}$ and $\frac{\tilde{\mathbf{X}}_i' \bar{\mathbf{M}}_2 \tilde{\mathbf{X}}_i}{T} - \frac{\tilde{\mathbf{X}}_i' \bar{\mathbf{M}}_{2q} \tilde{\mathbf{X}}_i}{T}$ where $\tilde{\mathbf{X}}_i$ are the residuals of $\mathbf{X}_i$ when regressed on $\mathbf{K}_1$. To see this note that

$$\frac{\mathbf{X}_i' \bar{\mathbf{M}} \mathbf{X}_i}{T} = \frac{\tilde{\mathbf{X}}_i' \tilde{\mathbf{M}}_2 \tilde{\mathbf{X}}_i}{T},$$

and

$$\frac{\mathbf{X}_i' \bar{\mathbf{M}}_q \mathbf{X}_i}{T} = \frac{\tilde{\mathbf{X}}_{qi}' \tilde{\mathbf{M}}_{2q} \tilde{\mathbf{X}}_{qi}}{T},$$

where $\tilde{\mathbf{X}}_i = \bar{\mathbf{M}}_1 \mathbf{X}_i$, $\tilde{\mathbf{X}}_{qi} = \bar{\mathbf{M}}_{1q} \mathbf{X}_i$,

$$\tilde{\mathbf{M}}_2 = \mathbf{I}_T - \bar{\mathbf{M}}_1 \bar{\mathbf{H}}_2 \left(\bar{\mathbf{H}}_2' \bar{\mathbf{M}}_1 \bar{\mathbf{H}}_2 \right)^{-} \bar{\mathbf{H}}_2' \bar{\mathbf{M}}_1,$$

23

and

$$\tilde{\mathbf{M}}_{2q} = \mathbf{I}_T - \bar{\mathbf{M}}_{1q}\bar{\mathbf{Q}}_2\left(\bar{\mathbf{Q}}_2'\bar{\mathbf{M}}_{1q}\bar{\mathbf{Q}}_2\right)^-\bar{\mathbf{Q}}_2'\bar{\mathbf{M}}_{1q}.$$

Then

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i}{T} - \frac{\mathbf{X}_i'\bar{\mathbf{M}}_q\mathbf{X}_i}{T} \le C\left(\left\|\frac{\mathbf{X}_i'\left(\bar{\mathbf{M}}_1 - \bar{\mathbf{M}}_{1q}\right)\tilde{\mathbf{M}}_2\tilde{\mathbf{X}}_i}{T}\right\| + \left\|\frac{\tilde{\mathbf{X}}_i'\left(\tilde{\mathbf{M}}_2 - \tilde{\mathbf{M}}_{2q}\right)\tilde{\mathbf{X}}_i}{T}\right\|\right),$$

for some positive constant, C. The desired result then follows easily from Lemma 2, Lemmas A.2 and A.3 of Pesaran (2006). $\blacksquare$

Lemma 2 *Under Assumptions 1-5 and if $\sum_{\ell=0}^{\infty}\mathbf{W}_\ell$ is of full rank where $\mathbf{W}_\ell$, $\ell = 0, ...,$ is defined in Remark 2, then as $(N,T) \xrightarrow{j} \infty$,*

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i}{T} = \frac{\mathbf{X}_i'\bar{\mathbf{M}}_q\mathbf{X}_i}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right), \tag{74}$$

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{F}}{T} = \frac{\mathbf{X}_i'\bar{\mathbf{M}}_q\mathbf{F}}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right), \tag{75}$$

$$\frac{\mathbf{X}_i'\bar{\mathbf{M}}\boldsymbol{\varepsilon}_i}{T} = \frac{\mathbf{X}_i'\bar{\mathbf{M}}_q\boldsymbol{\varepsilon}_i}{T} + O_p\left(\frac{1}{\sqrt{NT}}\right). \tag{76}$$

Proof. For simplicity but without loss of generality we assume that $\mathbf{d}_{3t}$ is empty. We start by proving (74). Throughout the proof C_i, $i = 1, ...,$ denote different positive constants. We need to determine the order of probability of $\frac{\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i}{T} - \frac{\mathbf{X}_i'\bar{\mathbf{M}}_q\mathbf{X}_i}{T}$. But this is equal to

$$\frac{\mathbf{X}_i'\bar{\mathbf{H}}\left(\bar{\mathbf{H}}'\bar{\mathbf{H}}\right)^-\bar{\mathbf{H}}'\mathbf{X}_i}{T} - \frac{\mathbf{X}_i'\bar{\mathbf{Q}}\left(\bar{\mathbf{Q}}'\bar{\mathbf{Q}}\right)^-\bar{\mathbf{Q}}'\mathbf{X}_i}{T} =$$

$$\left(\frac{\mathbf{X}_i'\bar{\mathbf{H}}}{T^{3/2}}\right)\left(\frac{\bar{\mathbf{H}}'\bar{\mathbf{H}}}{T^2}\right)^-\left(\frac{\bar{\mathbf{H}}'\mathbf{X}_i}{T^{3/2}}\right) - \left(\frac{\mathbf{X}_i'\bar{\mathbf{Q}}}{T^{3/2}}\right)\left(\frac{\bar{\mathbf{Q}}'\bar{\mathbf{Q}}}{T^2}\right)^-\left(\frac{\bar{\mathbf{Q}}'\mathbf{X}_i}{T^{3/2}}\right) \le$$

$$C_1\left\|\frac{\mathbf{X}_i'\bar{\mathbf{H}}}{T^{3/2}} - \frac{\mathbf{X}_i'\bar{\mathbf{Q}}}{T^{3/2}}\right\| + C_2\left\|\frac{\bar{\mathbf{H}}'\bar{\mathbf{H}}}{T^2} - \frac{\bar{\mathbf{Q}}'\bar{\mathbf{Q}}}{T^2}\right\|. \tag{77}$$

We examine the first term of (77) first. We have

$$\frac{\mathbf{X}_i'\bar{\mathbf{H}}}{T^{3/2}} - \frac{\mathbf{X}_i'\bar{\mathbf{Q}}}{T^{3/2}} = \frac{\mathbf{X}_i'\bar{\mathbf{U}}}{T^{3/2}},$$

where $\bar{\mathbf{U}} = \frac{1}{N}\sum_{i=1}^N\mathbf{U}_i$. But

$$\frac{\mathbf{X}_i'\bar{\mathbf{U}}}{T^{3/2}} \le C_3\left\|\frac{\mathbf{F}'\bar{\mathbf{U}}}{T^{3/2}} + \frac{\mathbf{D}'\bar{\mathbf{U}}}{T^{3/2}} + \frac{\mathbf{V}_i'\bar{\mathbf{U}}}{T^{3/2}}\right\|, \tag{78}$$

and

$$\frac{\mathbf{V}_i'\bar{\mathbf{U}}}{T^{3/2}} = O_p\left(\frac{1}{T\sqrt{N}}\right) + O_p\left(\frac{1}{N\sqrt{T}}\right), \tag{79}$$

24

by Lemma A.2 of Pesaran (2006). A similar treatment can be applied to the first and second terms of the RHS of (78). We therefore examine the first term only. Consider the ℓth row of $T^{-3/2}\left(\mathbf{F}'\bar{\mathbf{U}}\right)$ and note that it can be written as $T^{-3/2}\left(\sum_{t=1}^{T} f_{\ell t}\bar{\mathbf{u}}_t'\right)$. Consider now the limit of $T^{-3/2}\sum_{t=1}^{T} f_{\ell t}\bar{\mathbf{u}}_t$. First note that Assumptions 1-3 and Remark 1, via Theorem 4.1 of Davidson and De Jong (2000), guarantee that $\mathbf{u}_{it}$ can play the role of the integrator in a functional central limit theorem, i.e. the limit of $T^{-1}\sum_{t=1}^{T} f_{\ell t}\mathbf{u}_{it}'$ is a stochastic integral and therefore $O_p(1)$. Further note that for all finite N as well as when $N \to \infty$, the limit of $T^{-1}\sum_{t=1}^{T} f_{\ell t}\sqrt{N}\bar{\mathbf{u}}_t$ is a stochastic integral since assumption 3 and Remark 1 state that $\mathbf{u}_{it}'$ is a $L_2 - NED$ process of size $1/2$. This assumption, in turn, implies that $\bar{\mathbf{u}}_t$ is a $L_2 - NED$ process of size $1/2$ for all finite N and as $N \to \infty$, and the result holds again via Theorem 4.1 of Davidson and De Jong (2000). To see this we need to show that for some positive integer $s > 0$

$$E\left\{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{u}_{it} - E\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{u}_{it}\bigg| t-s\right)\right\}^2 = O(s^{-1}). \tag{80}$$

But

$$E\left\{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{u}_{it} - E\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{u}_{it}\bigg| t-s\right)\right\}^2 \leq \frac{1}{N}\sum_{i=1}^{N} E\left[\mathbf{u}_{it} - E\left(\mathbf{u}_{it}|t-s\right)\right]^2, \tag{81}$$

and since $E\left[\mathbf{u}_{it} - E\left(\mathbf{u}_{it}|t-s\right)\right]^2 = O(s^{-1})$ for all i, the desired result follows. From the above we also have

$$\frac{\mathbf{F}'\bar{\mathbf{U}}}{T^{3/2}} = O_p\left(\frac{1}{\sqrt{NT}}\right). \tag{82}$$

Similarly,

$$\frac{\mathbf{D}'\bar{\mathbf{U}}}{T^{3/2}} = O_p\left(\frac{1}{\sqrt{NT}}\right), \tag{83}$$

and hence,

$$\frac{\mathbf{X}_i'\bar{\mathbf{U}}}{T^{3/2}} = O_p\left(\frac{1}{\sqrt{NT}}\right). \tag{84}$$

We next examine the second term of (77). We have

$$\frac{\bar{\mathbf{H}}'\bar{\mathbf{H}}}{T^2} - \frac{\bar{\mathbf{Q}}'\bar{\mathbf{Q}}}{T^2} \leq C_4\left\|\frac{\bar{\mathbf{Q}}'\bar{\mathbf{U}}}{T^2} + \frac{\bar{\mathbf{U}}'\bar{\mathbf{U}}}{T^2}\right\|. \tag{85}$$

By Lemma A.2 of Pesaran (2006) it follows that

$$\frac{\bar{\mathbf{U}}'\bar{\mathbf{U}}}{T^2} = O_p\left(\frac{1}{NT}\right). \tag{86}$$

Further,

$$\frac{\bar{\mathbf{Q}}'\bar{\mathbf{U}}}{T^2} = O_p\left(\frac{\mathbf{F}'\bar{\mathbf{U}}}{T^2}\right) + O_p\left(\frac{\mathbf{D}'\bar{\mathbf{U}}}{T^2}\right) = O_p\left(\frac{1}{T\sqrt{N}}\right), \tag{87}$$

by (82). Thus,
$$\frac{\bar{\mathbf{H}}'\bar{\mathbf{H}}}{T^2} - \frac{\bar{\mathbf{Q}}'\bar{\mathbf{Q}}}{T^2} = O_p\left(\frac{1}{T\sqrt{N}}\right). \tag{88}$$

We now consider (75). We have

$$\frac{\mathbf{X}_i'\bar{\mathbf{H}}\left(\bar{\mathbf{H}}'\bar{\mathbf{H}}\right)^{-}\bar{\mathbf{H}}'\mathbf{F}}{T} - \frac{\mathbf{X}_i'\bar{\mathbf{Q}}\left(\bar{\mathbf{Q}}'\bar{\mathbf{Q}}\right)^{-}\bar{\mathbf{Q}}'\mathbf{F}}{T} =$$

$$\left(\frac{\mathbf{X}_i'\bar{\mathbf{H}}}{T^{3/2}}\right)\left(\frac{\bar{\mathbf{H}}'\bar{\mathbf{H}}}{T^2}\right)^{-}\left(\frac{\bar{\mathbf{H}}'\mathbf{F}}{T^{3/2}}\right) - \left(\frac{\mathbf{X}_i'\bar{\mathbf{Q}}}{T^{3/2}}\right)\left(\frac{\bar{\mathbf{Q}}'\bar{\mathbf{Q}}}{T^2}\right)^{-}\left(\frac{\bar{\mathbf{Q}}'\mathbf{F}}{T^{3/2}}\right) \leq$$

$$C_5\left\|\frac{\mathbf{X}_i'\bar{\mathbf{H}}}{T^{3/2}} - \frac{\mathbf{X}_i'\bar{\mathbf{Q}}}{T^{3/2}}\right\| + C_6\left\|\frac{\bar{\mathbf{H}}'\bar{\mathbf{H}}}{T^2} - \frac{\bar{\mathbf{Q}}'\bar{\mathbf{Q}}}{T^2}\right\| + C_7\left\|\frac{\bar{\mathbf{H}}'\mathbf{F}}{T^{3/2}} - \frac{\bar{\mathbf{Q}}'\mathbf{F}}{T^{3/2}}\right\|. \tag{89}$$

We only need to examine the third term of the RHS of (89). But

$$\frac{\bar{\mathbf{H}}'\mathbf{F}}{T^{3/2}} - \frac{\bar{\mathbf{Q}}'\mathbf{F}}{T^{3/2}} = \frac{\bar{\mathbf{U}}'\mathbf{F}}{T^{3/2}}.$$

But then by (82) and (84), (75) follows.

Finally, we consider (76). The treatment follows closely that in (77) and (89). In particular

$$\frac{\mathbf{X}_i'\bar{\mathbf{H}}\left(\bar{\mathbf{H}}'\bar{\mathbf{H}}\right)^{-}\bar{\mathbf{H}}'\boldsymbol{\varepsilon}_i}{T} - \frac{\mathbf{X}_i'\bar{\mathbf{Q}}\left(\bar{\mathbf{Q}}'\bar{\mathbf{Q}}\right)^{-}\bar{\mathbf{Q}}'\boldsymbol{\varepsilon}_i}{T} =$$

$$\left(\frac{\mathbf{X}_i'\bar{\mathbf{H}}}{T^{3/2}}\right)\left(\frac{\bar{\mathbf{H}}'\bar{\mathbf{H}}}{T^2}\right)^{-}\left(\frac{\bar{\mathbf{H}}'\boldsymbol{\varepsilon}_i}{T^{3/2}}\right) - \left(\frac{\mathbf{X}_i'\bar{\mathbf{Q}}}{T^{3/2}}\right)\left(\frac{\bar{\mathbf{Q}}'\bar{\mathbf{Q}}}{T^2}\right)^{-}\left(\frac{\bar{\mathbf{Q}}'\boldsymbol{\varepsilon}_i}{T^{3/2}}\right) \leq$$

$$C_8\left\|\frac{\mathbf{X}_i'\bar{\mathbf{H}}}{T^{3/2}} - \frac{\mathbf{X}_i'\bar{\mathbf{Q}}}{T^{3/2}}\right\| + C_9\left\|\frac{\bar{\mathbf{H}}'\bar{\mathbf{H}}}{T^2} - \frac{\bar{\mathbf{Q}}'\bar{\mathbf{Q}}}{T^2}\right\| + C_{10}\left\|\frac{\bar{\mathbf{H}}'\boldsymbol{\varepsilon}_i}{T^{3/2}} - \frac{\bar{\mathbf{Q}}'\boldsymbol{\varepsilon}_i}{T^{3/2}}\right\|. \tag{90}$$

But,

$$\frac{\bar{\mathbf{H}}'\boldsymbol{\varepsilon}_i}{T^{3/2}} - \frac{\bar{\mathbf{Q}}'\boldsymbol{\varepsilon}_i}{T^{3/2}} = \frac{\bar{\mathbf{U}}'\boldsymbol{\varepsilon}_i}{T^{3/2}}.$$

By Lemma A.2 of Pesaran (2006), it follows that

$$\frac{\bar{\mathbf{U}}'\boldsymbol{\varepsilon}_i}{T^{3/2}} = O_p\left(\frac{1}{T\sqrt{N}}\right) + O_p\left(\frac{1}{N\sqrt{T}}\right).$$

Hence, (76) follows. ∎

Table 1: Small Sample Properties of Common Correlated Effects Type Estimators in the Case of Experiment 1$\mathcal{A}$ (Heterogeneous Slopes + Full Rank)

(N,T)	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0: \beta_1 = 1.00$)					Power (5% level, $H_1: \beta_1 = 0.95$)				
	20	30	50	100	200	20	30	50	100	200	20	30	50	100	200	20	30	50	100	200
CCE Type Estimators																				
CCEMG																				
20	0.05	-0.10	-0.03	0.06	-0.07	9.67	7.89	6.74	5.87	5.54	7.20	6.90	7.15	7.90	7.55	11.65	13.00	16.10	17.50	20.10
30	0.09	-0.01	-0.01	-0.13	0.10	7.69	6.09	5.11	4.54	4.22	6.95	5.30	5.90	6.25	6.35	11.40	14.25	18.05	22.05	26.85
50	-0.19	0.22	-0.11	0.14	-0.04	5.88	4.61	4.01	3.44	3.13	5.70	5.05	6.65	6.20	5.95	15.10	20.40	25.60	34.10	36.65
100	0.00	0.04	0.04	0.03	0.04	4.25	3.46	2.89	2.33	2.27	5.75	5.85	5.25	4.90	6.20	23.35	34.30	44.40	56.00	63.25
200	-0.05	-0.02	-0.03	0.05	0.00	3.07	2.49	2.01	1.72	1.51	4.40	5.15	4.90	5.60	5.10	35.55	52.65	68.70	83.65	90.50
CCEP																				
20	0.18	0.00	-0.05	-0.01	-0.13	8.75	7.67	6.85	6.32	6.21	7.70	8.10	7.30	8.05	7.15	12.75	13.50	16.05	16.80	18.30
30	-0.17	-0.12	0.09	-0.15	0.13	7.10	5.99	5.32	4.78	4.46	7.55	6.25	6.75	6.65	6.45	12.40	15.00	19.30	20.65	26.90
50	0.00	0.18	-0.07	0.12	-0.01	5.33	4.51	3.97	3.47	3.22	6.80	6.20	5.90	6.35	6.45	17.45	22.15	26.40	32.90	36.25
100	0.00	0.09	0.03	0.00	0.02	3.78	3.25	2.85	2.34	2.28	5.70	5.65	5.60	5.15	6.25	28.15	37.40	44.80	55.20	61.75
200	-0.07	-0.04	-0.05	0.05	0.00	2.71	2.29	1.95	1.70	1.53	5.10	4.35	5.05	4.70	4.75	44.75	56.80	70.30	83.55	89.75
Principal Component Estimators, Augmented																				
PC1MG																				
20	-12.27	-11.15	-10.30	-8.87	-8.90	17.09	14.81	13.24	11.51	11.55	22.55	25.35	30.05	33.40	37.40	12.15	12.95	13.30	12.70	13.75
30	-9.25	-7.86	-6.46	-5.72	-5.25	13.55	10.84	8.98	7.80	7.15	20.60	20.90	21.65	24.75	24.70	10.75	8.25	7.35	7.40	6.75
50	-6.84	-5.05	-3.89	-3.01	-3.12	10.10	7.79	5.86	4.67	4.47	19.95	17.65	16.25	14.95	17.90	8.70	8.20	7.65	11.40	9.75
100	-4.78	-3.21	-2.03	-1.57	-1.45	7.44	5.34	3.68	2.87	2.72	20.10	16.80	11.45	9.75	11.10	9.55	12.15	20.25	28.85	36.75
200	-4.31	-2.54	-1.39	-0.81	-0.78	6.39	4.19	2.60	1.93	1.71	25.20	17.95	10.95	8.15	7.65	13.85	21.95	42.85	67.65	77.15
PC1POOL																				
20	-11.97	-11.04	-10.35	-9.09	-9.23	15.88	14.38	13.07	11.59	12.07	25.50	28.35	32.05	34.45	38.95	12.05	14.10	14.90	14.55	14.90
30	-8.86	-7.66	-6.34	-5.73	-5.37	12.48	10.45	8.89	7.80	7.34	21.45	23.75	22.05	24.70	25.50	11.00	8.80	7.55	7.95	6.35
50	-6.20	-4.86	-3.81	-3.07	-3.19	9.06	7.52	5.72	4.73	4.54	21.40	18.75	16.00	16.05	18.90	8.55	9.55	8.10	10.90	9.65
100	-4.36	-3.00	-2.01	-1.60	-1.49	6.61	5.01	3.61	2.88	2.74	21.05	16.85	11.25	9.35	10.80	11.25	14.55	20.85	27.90	36.30
200	-3.62	-2.32	-1.36	-0.81	-0.79	5.39	3.81	2.51	1.91	1.73	25.15	17.60	10.50	7.80	7.80	16.35	26.75	45.45	68.00	76.15
Principal Component Estimators, Orthogonalised																				
PC2MG																				
20	-31.26	-27.06	-24.01	-22.67	-23.11	32.83	28.34	25.00	23.44	23.83	86.50	88.45	91.25	95.20	97.40	74.10	73.95	75.80	82.05	88.20
30	-25.50	-21.21	-18.27	-16.69	-16.33	26.82	22.25	19.13	17.35	16.92	86.85	87.10	89.10	93.35	95.95	70.15	67.80	66.10	69.25	74.70
50	-20.65	-16.23	-13.32	-11.41	-10.89	21.68	17.06	13.98	11.95	11.37	90.15	88.35	88.80	89.05	91.70	70.80	60.25	52.20	45.80	46.10
100	-16.17	-12.44	-9.69	-7.61	-6.60	16.87	12.97	10.18	7.99	7.02	93.65	93.30	89.75	87.50	83.30	72.35	56.20	37.60	19.30	13.60
200	-14.61	-10.78	-8.12	-5.79	-4.59	15.11	11.19	8.45	6.08	4.85	98.95	97.85	95.45	90.75	83.75	79.65	60.20	33.30	10.00	6.75
PC2POOL																				
20	-31.97	-27.47	-24.27	-23.18	-24.19	33.39	28.69	25.23	23.99	24.99	91.00	90.70	93.20	95.55	98.50	80.65	78.60	78.80	83.35	90.45
30	-26.32	-21.51	-18.24	-16.83	-16.75	27.53	22.48	19.13	17.51	17.37	91.35	90.40	89.70	93.35	96.15	78.50	71.80	66.65	70.65	76.90
50	-21.22	-16.35	-13.17	-11.35	-10.99	22.10	17.15	13.82	11.91	11.48	95.05	90.90	88.95	88.20	91.70	79.65	63.80	52.95	46.20	48.25
100	-16.77	-12.52	-9.62	-7.55	-6.60	17.43	13.06	10.11	7.95	7.03	97.95	95.05	90.50	86.45	82.30	80.90	60.80	38.10	18.30	14.25
200	-15.16	-10.91	-8.00	-5.66	-4.53	15.67	11.33	8.34	5.96	4.79	99.75	98.45	95.95	89.35	82.50	88.65	65.85	33.35	8.40	6.30

(Table 1 Continued)

	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0 : \beta_1 = 1.00$)					Power (5% level, $H_1 : \beta_1 = 0.95$)				
CCE Type Estimators																				
Infeasible Estimators (including f_{1t} and f_{2t})																				
(N,T)	**20**	**30**	**50**	**100**	**200**	**20**	**30**	**50**	**100**	**200**	**20**	**30**	**50**	**100**	**200**	**20**	**30**	**50**	**100**	**200**
	Infeasible MG																			
20	0.01	-0.19	-0.08	0.15	-0.08	7.21	6.33	5.62	4.98	4.76	6.40	6.20	6.80	5.95	6.50	12.75	15.35	16.85	19.70	20.40
30	0.02	-0.14	0.01	-0.02	0.12	5.91	4.95	4.43	3.97	3.87	6.50	5.80	6.05	5.30	5.90	16.15	18.05	23.35	25.20	28.80
50	-0.10	0.07	-0.06	0.14	-0.04	4.48	3.75	3.39	3.09	2.94	6.45	5.25	5.90	5.25	5.20	21.70	27.35	31.45	38.45	40.25
100	0.01	0.07	0.02	0.00	0.04	3.16	2.78	2.49	2.15	2.14	5.50	5.15	5.45	4.70	5.45	36.85	46.15	55.10	62.50	66.65
200	-0.07	0.04	-0.07	0.06	0.01	2.22	1.93	1.69	1.57	1.44	4.85	5.00	5.00	5.60	4.70	59.15	72.85	82.25	90.40	92.75
	Infeasible Pooled																			
20	0.15	-0.13	-0.15	-0.26	-0.21	7.30	6.96	6.92	7.11	7.40	6.40	6.80	6.60	7.00	5.10	13.70	13.75	14.55	14.10	12.65
30	-0.20	-0.15	0.22	-0.07	0.27	6.23	5.78	5.79	5.89	6.61	7.05	5.90	7.00	5.25	5.70	15.70	15.35	18.95	16.70	16.60
50	0.12	0.07	-0.08	0.21	0.02	4.61	4.40	4.31	4.71	5.02	5.70	5.80	5.50	6.25	5.00	22.20	22.55	23.65	25.50	21.00
100	-0.05	0.07	0.09	0.06	0.00	3.30	3.26	3.12	3.30	3.52	5.25	5.60	5.20	5.20	5.30	33.45	38.20	38.85	36.75	32.30
200	-0.08	0.06	-0.12	0.07	-0.02	2.35	2.22	2.20	2.45	2.49	4.95	4.70	4.50	5.85	4.70	56.15	62.10	59.50	59.05	52.20
Naïve Estimators (excluding f_{1t} and f_{2t})																				
	Naïve MG																			
20	22.18	23.13	26.82	29.96	32.62	31.76	32.97	37.37	41.49	47.04	32.05	32.95	34.85	35.45	31.50	41.00	42.65	43.50	41.95	38.05
30	22.23	25.06	28.36	31.33	34.01	30.51	33.31	37.87	41.46	45.32	40.45	44.10	46.65	43.85	39.45	51.00	53.95	57.45	52.20	47.15
50	22.21	23.91	25.65	29.61	33.64	29.75	31.12	32.75	37.73	42.66	55.80	59.30	58.00	59.25	54.75	68.30	70.85	70.30	69.20	65.05
100	21.97	23.92	26.76	30.04	32.88	28.40	30.02	32.97	36.39	40.06	71.20	75.25	77.90	78.60	75.25	81.05	84.35	85.95	85.85	83.20
200	22.15	24.09	27.49	30.09	33.23	27.87	29.44	32.80	35.71	39.34	81.85	86.00	87.85	88.05	87.95	88.75	91.95	92.30	92.90	92.05
	Naïve Pooled																			
20	25.25	26.60	31.27	33.59	34.84	35.30	37.01	42.66	45.42	47.67	42.15	43.65	47.75	45.20	44.50	52.50	52.65	55.95	53.40	51.95
30	25.76	29.39	32.45	35.37	35.46	35.48	39.13	42.70	45.97	46.81	51.55	56.70	57.65	59.55	56.20	61.05	66.60	66.55	67.75	64.55
50	26.54	28.75	30.39	34.01	35.88	35.61	37.39	39.05	44.04	45.93	64.75	67.15	69.25	70.35	69.35	73.55	76.25	78.25	78.65	77.45
100	25.81	28.47	31.30	33.15	34.91	34.39	36.76	39.90	41.79	44.27	75.85	78.90	81.35	79.30	80.15	85.10	86.55	88.05	86.65	86.40
200	25.95	28.32	31.89	33.65	34.11	34.20	36.21	39.63	42.39	42.68	83.45	86.25	87.70	87.40	87.20	89.95	91.90	93.55	92.20	92.20

Notes: The DGP is $y_{it} = \alpha_{i1}d_{1t} + \beta_{i1}x_{1it} + \beta_{i2}x_{2it} + \gamma_{i1}f_{1t} + \gamma_{i2}f_{2t} + \varepsilon_{it}$ with $\varepsilon_{it} = \rho_{i\varepsilon}\varepsilon_{i,t-1} + \sigma_i(1 - \rho_{i\varepsilon}^2)^{1/2}\omega_{it}$, $i = 1, 2, ..., [N/2]$, and $\varepsilon_{it} = \sigma_i(1 + \theta_{i\varepsilon}^2)^{-1/2}(\omega_{it} + \theta_{i\varepsilon}\omega_{i,t-1})$, $i = [N/2] + 1, ..., N$, $\omega_{it} \sim IIDN(0,1)$, $\sigma_i^2 \sim IIDU[0.5, 1.5]$, $\rho_{i\varepsilon} \sim IIDU[0.05, 0.95]$, $\theta_{i\varepsilon} \sim IIDU[0,1]$. Regressors are generated by $x_{ijt} = a_{ij1}d_{1t} + a_{ij2}d_{2t} + \gamma_{ij1}f_{1t} + \gamma_{ij3}f_{3t} + v_{ijt}$, $j = 1, 2$, for $i = 1, 2, ..., N$. $d_{1t} = 1$, $d_{2t} = 0.5d_{2,t-1} + v_{dt}$, $v_{dt} \sim IIDN(0, 1 - 0.5^2)$, $d_{2,-50} = 0$; $f_{jt} = f_{jt-1} + v_{fj,t}$, $v_{fj,t} \sim IIDN(0,1)$, $f_{j,-50} = 0$, for $j = 1, 2, 3$; $v_{ijt} = \rho_{vij}v_{ijt-1} + v_{ijt}$, $v_{ijt} \sim IIDN(0, 1 - \rho_{vij}^2)$, $v_{ij,-50} = 0$ and $\rho_{vij} \sim IIDU[0.05, 0.95]$ for $j = 1, 2$, for $t = -49, ..., T$ with the first 50 observations discarded; $\alpha_{i1} \sim IIDN(1,1)$; $a_{ij\ell} \sim IIDN(0.5, 0.5)$ for $j = 1, 2$, $\ell = 1, 2$; γ_{i11} and $\gamma_{i23} \sim IIDN(0.5, 0.50)$, γ_{i13} and $\gamma_{i21} \sim IIDN(0, 0.50)$; γ_{i1} and $\gamma_{i2} \sim IIDN(1, 0.2)$; $\beta_{ij} = 1 + \eta_{ij}$ with $\eta_{ij} \sim IIDN(0, 0.04)$ for $j = 1, 2$. ρ_{vij}, $\rho_{i\varepsilon}$, $\theta_{i\varepsilon}$, σ_i^2, α_{i1}, $a_{ij\ell}$ for $j = 1, 2$, $\ell = 1, 2$ are fixed across replications. CCEMG and CCEP are defined by (14) and (17), and their variance estimators are defined by (56) and (64), respectively. The variance estimators of all other mean group and pooled estimators are defined by (69) and (70), respectively. The PC type estimators are computed assuming the number of unobserved factors, $m = 3$, is known. All experiments are based on 2000 replications.

Table 2: Small Sample Properties of Common Correlated Effects Type Estimators
in the Case of Experiment 2$\mathcal{A}$ (Homogeneous Slopes + Full Rank)

(N,T)	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0 : \beta_1 = 1.00$)					Power (5% level, $H_1 : \beta_1 = 0.95$)				
	20	30	50	100	200	20	30	50	100	200	20	30	50	100	200	20	30	50	100	200
CCE Type Estimators																				
CCEMG																				
20	0.05	-0.15	0.02	-0.15	0.09	8.45	6.29	5.10	3.78	3.14	7.15	6.40	6.80	6.75	6.85	11.70	13.80	21.75	31.25	47.90
30	-0.14	0.12	0.04	0.03	0.00	6.44	5.11	3.80	2.67	2.07	6.05	6.75	7.25	6.40	6.45	12.70	20.45	30.70	50.90	71.60
50	0.08	-0.06	0.02	0.05	0.03	5.08	3.79	2.80	1.94	1.39	6.10	5.90	4.85	5.40	5.35	18.00	26.90	44.45	75.65	95.00
100	-0.04	-0.08	0.06	-0.04	-0.01	3.59	2.76	2.02	1.35	0.98	4.55	5.50	6.05	5.10	6.10	28.30	43.00	72.35	95.20	99.90
200	0.06	-0.02	0.03	0.01	0.00	2.83	2.05	1.52	1.00	0.68	5.60	4.45	6.35	5.20	5.70	44.20	67.95	91.90	99.90	100.00
CCEP																				
20	0.18	0.00	0.03	-0.14	0.08	6.95	5.56	4.94	3.98	3.74	6.60	6.75	7.30	6.75	6.80	14.25	16.25	25.25	33.70	46.25
30	-0.14	0.14	0.07	0.01	0.01	5.20	4.50	3.55	2.67	2.26	5.10	5.90	7.25	6.25	6.40	15.25	24.55	34.90	52.95	70.70
50	0.05	0.07	-0.02	0.04	0.03	4.08	3.29	2.56	1.84	1.39	5.40	5.40	5.45	6.20	5.30	24.60	34.35	51.70	78.65	95.00
100	-0.02	-0.04	0.06	-0.04	-0.01	2.87	2.37	1.78	1.24	0.93	5.60	6.20	6.40	5.25	5.95	41.65	58.35	81.85	97.80	100.00
200	0.07	-0.03	0.01	0.02	0.00	2.17	1.63	1.32	0.92	0.65	5.60	3.95	5.70	5.60	5.35	65.25	84.40	96.95	100.00	100.00
Principal Component Estimators, Augmented																				
PC1MG																				
20	-12.34	-11.39	-10.06	-9.44	-8.84	16.76	14.48	12.18	11.24	10.92	24.55	32.55	39.50	58.50	71.95	13.05	15.30	13.80	19.00	21.20
30	-9.35	-7.83	-6.39	-5.66	-5.34	12.96	10.55	8.18	6.93	6.15	22.10	26.10	32.55	46.60	68.25	9.55	10.60	8.40	10.25	10.10
50	-7.05	-5.28	-3.81	-3.08	-3.17	10.12	7.38	5.11	3.86	3.73	23.55	24.05	25.30	35.65	56.85	10.40	9.10	9.45	20.80	30.00
100	-5.00	-3.45	-2.04	-1.64	-1.57	7.19	5.16	3.14	2.20	1.90	22.60	22.00	16.50	22.45	35.70	9.90	14.45	31.05	64.90	91.50
200	-4.23	-2.65	-1.27	-0.87	-0.79	6.27	4.11	2.13	1.37	1.06	28.05	22.90	14.90	16.85	21.95	16.15	30.55	67.35	97.55	100.00
PC1POOL																				
20	-11.78	-11.12	-9.89	-9.43	-8.93	15.09	13.70	11.86	11.20	10.77	28.20	37.15	46.35	64.10	76.75	14.80	17.20	17.65	21.80	24.55
30	-8.55	-7.35	-6.10	-5.58	-5.35	11.37	9.59	7.66	6.79	6.18	25.60	29.35	35.60	49.60	71.00	10.65	9.95	8.95	10.10	10.30
50	-6.39	-4.86	-3.74	-3.05	-3.19	8.82	6.71	4.87	3.77	3.81	26.60	26.60	28.30	36.55	59.50	10.95	10.25	10.05	22.35	31.90
100	-4.42	-3.23	-1.98	-1.61	-1.56	6.14	4.68	2.89	2.10	1.87	26.80	25.60	19.70	24.80	38.70	12.65	19.40	40.40	73.05	93.65
200	-3.57	-2.37	-1.21	-0.84	-0.78	5.19	3.57	1.93	1.30	1.03	32.05	25.30	16.60	17.35	24.10	24.55	42.80	79.25	98.80	100.00
Principal Component Estimators, Orthogonalised																				
PC2MG																				
20	-31.24	-27.21	-23.95	-22.96	-22.95	32.64	28.30	24.70	23.46	23.37	89.70	92.75	96.35	99.60	100.00	78.30	82.20	85.40	95.90	98.25
30	-25.74	-21.23	-18.28	-16.52	-16.52	26.93	22.12	18.86	16.89	16.81	90.55	93.60	97.60	99.60	100.00	78.20	76.50	82.60	90.75	97.65
50	-20.76	-16.51	-13.40	-11.39	-10.92	21.63	17.17	13.81	11.69	11.12	94.65	95.85	98.35	99.85	100.00	78.65	73.60	70.95	73.80	86.75
100	-16.31	-12.50	-9.58	-7.70	-6.67	16.92	12.94	9.90	7.88	6.81	96.60	97.65	98.25	99.80	99.95	79.25	68.50	52.60	40.60	32.25
200	-14.51	-10.80	-8.05	-5.85	-4.57	14.98	11.13	8.28	5.98	4.66	99.50	99.65	99.30	99.85	99.95	83.90	69.50	47.75	14.70	13.00
PC2POOL																				
20	-31.95	-27.52	-24.18	-23.50	-24.05	33.04	28.47	24.87	24.02	24.56	95.80	96.50	98.25	99.80	100.00	87.95	87.00	89.80	96.95	99.30
30	-26.27	-21.47	-18.38	-16.67	-16.96	27.25	22.29	18.91	17.06	17.28	96.35	96.25	98.95	99.75	100.00	86.75	83.60	86.75	92.40	98.35
50	-21.31	-16.46	-13.29	-11.34	-11.05	22.05	17.04	13.68	11.63	11.26	98.80	98.00	99.00	99.90	100.00	89.35	81.05	76.05	74.65	88.35
100	-16.95	-12.65	-9.52	-7.62	-6.67	17.50	13.05	9.81	7.80	6.81	99.40	99.40	99.50	99.95	100.00	90.45	78.30	58.35	41.55	32.65
200	-15.07	-10.92	-7.93	-5.72	-4.52	15.52	11.24	8.15	5.85	4.60	99.90	99.95	99.90	99.90	99.95	94.80	80.20	51.05	14.00	15.30

Notes: The DGP is the same as that of Table 1, except $\beta_{ij} = 1$ for all i and j, $i = 1, 2, ..., N$, $j = 1, 2$. See notes to Table 1.

Table 3: Small Sample Properties of Common Correlated Effects Type Estimators in the Case of Experiment 1$\mathcal{B}$ (Heterogeneous Slopes + Rank Deficient)

	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0 : \beta_1 = 1.00$)					Power (5% level, $H_1 : \beta_1 = 0.95$)				
CCE Type Estimators																				
(N,T)	20	30	50	100	200	20	30	50	100	200	20	30	50	100	200	20	30	50	100	200
CCEMG																				
20	0.33	-0.19	0.20	0.14	0.23	15.02	13.90	12.61	13.35	13.78	6.80	6.90	6.75	6.60	7.20	9.40	8.95	10.15	10.15	10.15
30	0.30	0.14	0.09	-0.17	0.35	12.91	12.03	10.70	10.07	10.59	5.50	6.80	5.25	6.15	4.80	8.40	10.05	9.45	10.35	11.65
50	-0.15	0.63	-0.20	-0.17	0.02	9.82	8.46	7.87	7.42	7.34	5.80	5.10	6.10	5.75	5.90	9.75	12.90	13.40	14.00	15.20
100	0.25	0.13	0.27	0.00	0.06	7.01	6.55	5.85	5.25	5.01	5.75	5.95	5.45	5.45	6.10	14.50	17.75	21.65	22.65	27.30
200	0.05	-0.11	-0.17	-0.07	-0.05	5.35	4.65	4.15	3.61	3.31	4.80	5.05	4.75	5.15	4.55	19.45	23.70	29.75	37.25	43.45
CCEP																				
20	0.48	0.06	-0.04	0.16	0.10	13.13	12.81	12.21	13.57	15.30	6.75	7.40	7.00	6.65	6.75	9.90	10.20	10.40	10.35	10.25
30	-0.23	-0.06	0.18	-0.25	0.43	11.48	10.70	10.39	9.95	11.04	6.10	6.90	5.70	6.00	5.50	9.05	9.95	10.55	10.25	10.60
50	0.00	0.48	-0.18	-0.17	-0.02	8.42	7.57	7.23	7.22	7.22	5.25	5.90	6.25	5.30	5.50	11.40	14.05	14.15	14.35	15.20
100	0.11	0.18	0.24	-0.06	0.05	5.87	5.72	5.27	4.87	4.98	5.10	6.00	5.40	4.95	6.00	17.25	19.60	23.50	23.55	27.00
200	0.04	-0.10	-0.16	-0.04	-0.03	4.35	3.99	3.75	3.30	3.15	5.40	4.70	5.25	4.10	3.95	25.75	28.50	34.50	41.10	46.05
Principal Component Estimators, Augmented																				
PC1MG																				
20	-8.33	-8.02	-7.93	-7.27	-7.44	14.50	12.28	10.98	9.59	9.25	13.35	16.10	21.45	24.50	28.80	7.20	7.35	7.20	7.90	8.00
30	-5.36	-4.97	-4.70	-4.58	-4.27	10.67	8.66	7.30	6.53	5.96	10.00	11.45	13.80	17.60	18.75	5.85	5.40	5.35	5.40	5.60
50	-3.15	-2.86	-2.99	-2.59	-2.63	7.44	5.94	5.15	4.33	4.07	7.95	8.25	12.35	11.65	14.30	6.40	6.50	8.70	12.45	11.50
100	-1.35	-1.27	-1.33	-1.26	-1.17	5.04	4.01	3.24	2.68	2.55	6.80	8.00	7.85	7.40	8.40	13.50	18.90	24.85	31.90	42.30
200	-0.85	-0.77	-0.72	-0.58	-0.62	3.56	2.83	2.18	1.83	1.64	4.95	6.35	6.05	6.75	6.30	23.40	36.25	53.50	72.15	80.20
PC1POOL																				
20	-7.93	-7.85	-7.87	-7.23	-7.32	12.94	11.59	10.53	9.43	9.14	13.45	16.70	22.35	25.30	28.35	7.50	7.30	7.60	7.65	7.35
30	-5.33	-4.94	-4.61	-4.54	-4.21	9.81	8.23	7.13	6.50	5.92	11.10	12.55	13.25	17.70	18.40	6.55	5.90	5.60	5.70	5.35
50	-2.98	-2.78	-2.92	-2.63	-2.65	6.62	5.67	5.01	4.37	4.12	8.00	9.85	11.35	12.15	14.20	7.40	7.90	8.40	11.70	11.50
100	-1.41	-1.23	-1.36	-1.31	-1.21	4.37	3.66	3.19	2.71	2.57	6.80	7.25	7.50	7.70	9.15	15.85	21.20	26.30	31.30	40.80
200	-0.82	-0.77	-0.75	-0.60	-0.64	3.02	2.57	2.12	1.82	1.67	6.05	6.35	6.65	5.90	6.75	29.20	42.25	56.25	71.40	78.95
Principal Component Estimators, Orthogonalised																				
PC2MG																				
20	-30.74	-26.62	-23.99	-22.56	-23.18	32.50	27.94	25.06	23.36	23.92	82.50	87.55	90.45	95.10	97.55	70.50	72.50	74.75	81.15	87.80
30	-24.89	-20.85	-18.19	-16.68	-16.43	26.28	21.96	19.03	17.35	17.03	84.15	86.05	90.35	93.60	96.05	67.80	65.45	66.75	69.85	75.65
50	-19.61	-15.65	-13.13	-11.38	-10.85	20.69	16.50	13.80	11.93	11.34	87.30	86.20	87.95	89.65	91.70	65.10	57.00	51.20	45.00	46.35
100	-15.19	-11.94	-9.58	-7.57	-6.60	15.96	12.53	10.09	7.96	7.03	91.80	90.95	88.90	87.35	82.65	65.05	50.60	37.45	18.45	14.20
200	-13.64	-10.37	-7.98	-5.75	-4.58	14.16	10.80	8.33	6.05	4.85	98.50	96.95	95.00	90.60	83.65	72.00	55.40	31.15	9.30	6.80
PC2POOL																				
20	-31.30	-27.08	-24.30	-23.16	-24.26	32.81	28.29	25.32	23.99	25.10	88.95	89.95	93.15	96.00	98.75	79.05	77.70	79.10	83.00	90.25
30	-25.55	-21.11	-18.14	-16.84	-16.86	26.80	22.15	18.99	17.53	17.49	90.25	88.25	90.40	94.00	96.30	75.35	68.95	66.20	70.00	77.15
50	-19.99	-15.74	-13.00	-11.35	-10.96	20.92	16.56	13.65	11.91	11.46	93.30	89.05	88.30	88.65	91.70	74.15	60.30	51.40	45.55	47.50
100	-15.72	-11.98	-9.51	-7.52	-6.60	16.43	12.56	10.01	7.91	7.03	96.00	92.55	89.40	86.50	81.95	74.30	54.45	36.55	17.60	14.00
200	-14.09	-10.47	-7.86	-5.63	-4.52	14.61	10.90	8.21	5.93	4.79	99.30	97.95	95.20	88.75	81.90	82.55	60.20	32.10	8.25	6.15

Notes: The DGP is the same as that of Table 1, except $\gamma_{i2} \sim IIDN(0,1)$, so that the rank condition is not satisfied. See notes to Table 1.

Table 4: Small Sample Properties of Common Correlated Effects Type Estimators in the Case of Experiment 2ℬ (Homogeneous Slopes + Rank Deficient)

	Bias (×100)					Root Mean Square Errors (×100)					Size (5% level, $H_0 : \beta_1 = 1.00$)					Power (5% level, $H_1 : \beta_1 = 0.95$)				
CCE Type Estimators																				
(N,T)	20	30	50	100	200	20	30	50	100	200	20	30	50	100	200	20	30	50	100	200
CCEMG																				
20	-0.28	-0.26	0.41	-0.31	0.73	14.45	12.85	12.02	12.07	13.47	7.35	5.45	6.40	6.70	6.00	9.35	9.15	10.95	11.55	10.90
30	-0.11	0.07	0.09	0.45	-0.05	11.99	10.78	9.82	9.52	10.33	5.20	5.90	5.95	6.50	6.55	7.85	10.50	12.40	14.35	14.90
50	0.00	0.23	-0.07	-0.02	0.00	9.01	7.97	7.62	6.79	6.72	5.05	4.80	5.00	5.45	4.95	9.40	12.20	15.75	17.60	21.15
100	0.14	-0.08	-0.12	-0.03	0.06	6.66	5.92	5.16	4.78	4.56	4.65	5.40	5.60	4.60	6.35	15.10	18.15	23.95	28.50	34.85
200	0.14	0.11	0.01	-0.17	-0.07	5.13	4.45	3.88	3.27	3.34	5.45	5.10	5.45	4.65	5.15	22.35	28.80	36.60	44.75	56.70
CCEP																				
20	-0.12	-0.19	0.35	-0.26	0.66	12.66	11.53	11.56	12.12	15.07	7.45	7.00	7.55	6.35	6.50	9.85	10.00	12.60	12.65	11.50
30	-0.09	0.05	0.06	0.39	0.03	10.00	9.57	9.26	9.36	11.05	5.55	5.75	6.80	6.70	6.75	9.90	11.70	13.30	15.20	14.50
50	-0.14	0.39	-0.08	0.01	0.03	7.29	6.92	6.84	6.58	6.79	4.95	5.25	5.45	5.60	4.85	11.25	15.60	16.65	19.95	20.40
100	0.20	-0.13	-0.11	-0.05	0.04	5.44	4.97	4.55	4.45	4.39	4.80	5.35	5.40	4.95	6.05	20.60	22.65	28.35	31.40	36.80
200	0.19	0.11	-0.08	-0.13	-0.07	3.97	3.71	3.35	2.96	3.09	5.25	5.15	5.05	5.00	5.60	31.95	38.45	44.30	50.70	60.40
Principal Component Estimators, Augmented																				
PC1MG																				
20	-8.22	-8.36	-7.93	-7.89	-7.48	13.75	11.78	10.07	9.07	8.40	13.30	18.95	27.70	46.00	58.60	7.25	7.80	8.20	9.55	9.40
30	-5.40	-5.01	-4.70	-4.59	-4.52	10.11	8.07	6.49	5.54	5.03	10.35	13.45	19.70	33.45	52.05	5.95	5.45	6.45	5.80	4.95
50	-3.24	-3.09	-2.81	-2.66	-2.66	6.96	5.35	4.22	3.39	3.04	9.40	11.05	15.90	25.40	44.85	6.85	7.80	11.60	22.45	35.50
100	-1.53	-1.45	-1.28	-1.33	-1.28	4.54	3.40	2.57	1.96	1.64	6.10	7.75	10.25	15.80	26.30	13.10	20.95	41.60	71.75	95.70
200	-0.80	-0.76	-0.67	-0.64	-0.63	3.40	2.39	1.74	1.22	0.94	6.00	5.45	7.75	10.75	15.70	27.15	46.80	79.60	98.65	100.00
PC1POOL																				
20	-7.56	-7.83	-7.57	-7.61	-7.12	11.45	10.60	9.32	8.59	7.81	14.05	21.70	31.20	48.25	57.85	6.40	8.35	7.55	8.90	8.35
30	-5.00	-4.76	-4.50	-4.43	-4.35	8.47	7.20	6.04	5.25	4.82	11.70	15.75	21.45	35.10	51.40	6.00	5.70	6.50	5.65	5.40
50	-2.89	-2.77	-2.75	-2.60	-2.61	5.66	4.63	3.92	3.26	2.97	8.90	11.55	17.50	27.45	44.70	7.85	9.60	12.95	24.65	38.05
100	-1.38	-1.39	-1.26	-1.31	-1.27	3.67	2.95	2.29	1.87	1.60	7.80	8.70	10.50	17.90	28.50	19.90	28.70	51.45	79.70	97.35
200	-0.72	-0.74	-0.67	-0.63	-0.63	2.54	1.93	1.55	1.14	0.91	6.25	5.90	9.00	11.80	17.10	43.25	64.40	89.05	99.45	100.00
Principal Component Estimators, Orthogonalised																				
PC2MG																				
20	-30.58	-26.66	-23.75	-22.85	-23.07	32.14	27.74	24.51	23.36	23.51	87.40	91.15	95.95	99.60	99.95	75.80	79.65	85.50	94.95	98.85
30	-25.10	-20.76	-18.15	-16.58	-16.60	26.40	21.70	18.75	16.97	16.90	89.25	92.45	97.20	99.75	100.00	74.35	74.15	80.85	90.90	98.10
50	-19.66	-15.83	-13.16	-11.36	-10.91	20.58	16.50	13.60	11.67	11.12	92.45	94.35	97.90	99.75	100.00	72.90	69.30	69.35	72.45	86.90
100	-15.37	-11.97	-9.44	-7.65	-6.66	16.01	12.43	9.75	7.84	6.80	95.30	96.70	98.00	99.70	99.95	74.25	63.40	50.10	40.00	32.15
200	-13.52	-10.36	-7.93	-5.82	-4.57	14.01	10.70	8.17	5.96	4.66	98.80	99.15	99.40	99.90	99.95	76.85	63.50	45.30	14.80	13.35
PC2POOL																				
20	-31.34	-27.06	-24.05	-23.47	-24.10	32.52	28.01	24.76	24.00	24.63	95.25	95.90	97.85	99.95	100.00	86.20	86.30	89.00	97.50	99.40
30	-25.58	-20.90	-18.28	-16.76	-17.05	26.61	21.72	18.83	17.15	17.38	95.30	95.80	98.85	99.85	100.00	84.80	81.60	85.35	93.05	98.50
50	-20.12	-15.76	-13.06	-11.33	-11.05	20.90	16.33	13.46	11.63	11.26	97.65	97.50	98.90	99.75	100.00	85.10	77.55	73.60	75.00	88.75
100	-15.88	-12.09	-9.36	-7.58	-6.66	16.46	12.51	9.65	7.76	6.80	99.10	99.00	99.45	99.85	99.95	86.15	72.45	55.15	39.85	32.95
200	-13.95	-10.45	-7.81	-5.69	-4.52	14.41	10.78	8.03	5.82	4.60	99.70	99.85	99.75	99.85	100.00	90.10	74.75	49.05	13.20	15.30

Notes: The DGP is the same as that of Table 1, except $\gamma_{i2} \sim IIDN(0,1)$, so that the rank condition is not satisfied, and $\beta_{ij} = 1$ for all i and j, $i = 1, 2, ..., N$, $j = 1, 2$. See notes to Table 1.