Discussion Papers in Economics

CONSISTENT ESTIMATION OF PANEL DATA MODELS WITH A MULTI- FACTOR ERROR STRUCTURE

By

Bin Peng
(Monash University)
&
Giovanni Forchini
(University of Surrey)

DP 01/12

Department of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web www.econ.surrey.ac.uk
ISSN: 1749-5075

Consistent Estimation of Panel Data Models with a Multi-factor Error Structure[1]

18 December 2011

Bin Peng[2]
Monash University

Giovanni Forchini[3]
University of Surrey

Abstract:

This paper considers the panel data model with a multifactor structure in both the errors and the regressors which was studied by Pesaran ("Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure, *Econometrica*, 2006). Estimators are proposed that are consistent for fixed T as N tends to infinity. By allowing T to be fixed some of the assumptions imposed by Pesaran are relaxed and, at the same time, some of the complexities of the large N and T asymptotics are bypassed. A small Monte Carlo simulation shows that these new estimators are very accurate for very small values of T .

---

[1] This research was partially supported by Australian Research Council grant DP0985432.

[2] Address for correspondence: Bin Peng, Department of Econometrics and Business Statistics, Monash University, Clayton, Victoria 3800, Australia. E-mail: Bing.Peng@monash.edu

[3] Address for correspondence: Giovanni Forchini, School of Economics, Ground Floor AD Building, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom. E-mail: G.Forchini@surrey.ac.uk.

1. Introduction

Cross sectional dependence in panel data model leads to inefficiency of estimators based on ordinary least squares. Special forms of cross sectional dependence for which asymptotically efficient estimation is possible through the use generalized least squares procedures have been investigated by various authors, including, among others, Case (1991), Conley (1999), Pesaran (2006) and Bai (2009) (see also the references therein).

A parsimonious way of dealing with cross sectional dependence in panel data uses factors in the errors. Consistent estimation of these models can be done by maximum likelihood procedures when the regressors are not correlated with the factors (e.g. among others Robertson and Symons (2000)). Coakley, Fuertes and Smith (2001) suggest an estimation procedure based on principal components applied to the residuals and claim that their approach yields a consistent estimator even if the regressors are correlated with the factors in the error term. Recently, Pesaran (2006) notices that the estimation procedure suggested by Coakley, Fuertes and Smith (2001) is, in general, inconsistent and proposes a new approach that yields consistent estimators when both the N and T dimensions tend to infinity. These results have been extended by Bai (2009) to set-ups that allow for a more complex dependence of the regressors on the unknown factors and factor loadings, and by Su and Jin (2010) and Huang (2010) to semiparametric models. Tests for cross sectional dependence have been studied by Pesaran (2004), Hoyos and Sarafidis (2006) and Sarafidis, Yamagata and Robertson (2009). Dynamic panel data models with factor structures have been considered by Phillips and Sul (2003), Phillips and Sul (2007), Sarafidis and Robertson (2009) and Sarafidis (2009).

The existing literature for a static panel data model with error factor structure such as the one studied by Pesaran (2006) and Bai (2009) relies on large T and N asymptotics because, under the assumption of stationarity, it is possible to find a consistent estimator of the covariance matrix of the N errors at each $i = 1,2,...,N$. Thus, it is possible to use a feasible generalized least squares approach to obtain consistent estimators of the parameters of interest (e.g. Robertson and Symons (2000)). The approaches of both Pesaran (2006) and Bai (2009) present several technical challenges because it is difficult to establish joint limits in both N and T (c.f. Phillips and Moon (1999)). In fact, reliance on large T results forces the various authors to impose very strong assumptions about the stationarity (e.g. Pesaran (2006) and Bai (2009)) or the precise nature of the nonstationarity (e.g. Bai, Kao and Ng (2009) and Bai and Ng (2010)) of various unobserved quantities in the time dimension. Only a few large N-fixed T results are available but are very weak: Pesaran (2006) shows that the common correlated effect estimators are asymptotically unbiased for fixed T.

This paper investigates a variant of the model proposed by Pesaran (2006) when T is fixed and N is large. The approach has two advantages over existing ones. First, regarding T as fixed allows us to considerably weaken the assumptions on the behaviour of the various stochastic terms which are commonly imposed in the time dimension. For example we do not need to impose that the T individual specific errors follow linear stochastic processes with absolute summable autocovariances as in Assumption 2 of Pesaran (2006). Analogously, we do not need to assume covariance stationarity with absolute summable autocovariances for the factors (e.g. Assumption 1 of Pesaran (2006)). Our results will be valid irrespective of the time-series assumption in the T dimension imposed. Second, the large N asymptotic we use is very simple and standard and does not require mastering the technical complexities of the large N and T asymptotics studied by Phillips and Moon (1999).

Three estimators are proposed in this paper. Two of them are generalized method of moments estimators. All three of them are standard fixed effects estimators in which the N and T dimensions are interchanged. As such they are very easy to compute. They have two advantages over the estimators suggested by Pesaran (2006). They are unbiased for fixed N and T despite the presence of correlation between the errors and the regressors, and they can be calculated for any T. In contrast, Pesaran's estimators can be calculated only if T is large in comparison to the dimension of the interest parameters. This permits the use of panel data models with error factor structure in a microeconometric context where the number of regressors usually exceeds T by a large amount.

The structure of the paper is as follows. Section 2 derives unbiased and consistent estimators for a panel data model with homogeneous slopes. Section 3 extends these results to a model with heterogeneous slopes. Section 4 discusses a simple simulation exercise and Section 5 concludes. All proofs are in the appendix.

2. Homogeneous slopes

A simple panel data model with cross-sectional dependence and correlation between the errors and the regressors that allows us to illustrate our methodology is as follows:

(1) $$\underset{(T\times 1)}{y_i} = \underset{(T\times p)}{x_i}\underset{(p\times 1)}{\beta_0} + e_i$$

(2) $$\underset{(T\times 1)}{e_i} = \underset{(T\times m)}{F_T}\underset{(m\times 1)}{\gamma_i} + \varepsilon_i$$

(3) $$x_i = F_T \underset{(m\times p)}{\Gamma_i} + v_i$$

where $F_T = (f_1, f_2, ..., f_T)'$ denotes a $(T \times m)$ matrix of unobserved common factors effects, γ_i is a $(m \times 1)$ vector of factor loadings and ε_i is a purely idiosyncratic random vector with zero mean and constant co-variance matrix. The error vector v_i is distributed independently of the common factors across *i*. Finally, Γ_i is a $(m \times p)$ factor loading matrix.

Equation (1) is a standard panel data model in which the parameter β_0 is assumed to be the same for every $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$ (hence the homogeneous slopes). Notice, however, that both e_i and x_i depend on the factors F_T, so that, in contrast to a standard model, the errors and the regressors in (1) are correlated and standard estimators are inconsistent. Equations (1), (2) and (3) are a special case of a model suggested by Pesaran (2006), for which β_0 changes across i, that we will consider in detail in the next section.

We will focus on the case where N is large and T is small since no consistent estimator is available for this case in the existing literature. We now make the following assumptions.

Assumption 1

i. *The ε_i's are identically, independently distributed (i.i.d.) random vectors with zero mean vector and covariance matrix Σ_ε.*

ii. *The $vec(v_i')$'s are i.i.d. random vectors with mean vector zero and covariance matrix Σ_v.*

iii. *The ε_i and v_i are independent for every $i = 1,2,...,N$.*

iv. *ε_i and v_i are independent of the factors F_T.*

v. $E\|\varepsilon_i\|^4 \le \sigma < \infty$ *and* $E\|v_i\|^4 \le \sigma < \infty$.

Assumption 2

i. *F_T is a random matrix.*

ii. *The γ_i's are i.i.d. with mean vector $\gamma \neq 0$ and covariance matrix Σ_γ for $i = 1,2,...,N$. Moreover they are independent of F_T, v_i and ε_i.*

Assumption 3

i. *The $vec(\Gamma_i')$'s are i.i.d. with mean $vec(\Gamma')$ and covariance matrix Σ_Γ and have bounded fourth order moments $E\|\Gamma_i\|^4 \le M < \infty$ for $i = 1,2,...,N$.*

ii. *The unobserved factor loadings γ_i and Γ_i are uncorrelated. Moreover, the Γ_i's are independent of F_T, v_i, ε_i and γ_i.*

iii. *Γ and γ are not orthogonal in the sense that $\Gamma' F_T' F_T \gamma \neq 0$ where $E(\Gamma_i) = \Gamma \neq 0$.*

Assumptions 1, 2 and 3 are standard in the literature (c.f. Pesaran (2006) and Sarafidis and Robertson (2009)). They ensure that all the errors cross sectional dependence is captured by the factor structure only. They also imply that the correlation between e_i and x_i is only due to the common factors. Dropping the assumption of independence between ε_i and v_i would have serious consequences and, for example, it would not be possible to find unbiased estimators for β_0.

Assumption 1 does not impose any restrictions on the time series property of the panel (c.f. Pesaran (2006) who assumes that the series of t follows a linear stationary process with absolute summable autocovariances). Therefore, for each i the components of ε_i and v_i could follow stationary or nonstationary processes.

The OLS estimator is not consistent under these assumptions. To see this notice that the OLS estimator is

$$\hat{\beta}_{OLS} = \left(\frac{1}{N}\sum_{i=1}^{N} x_i ' x_i\right)^{-1} \left(\frac{1}{N}\sum_{i=1}^{N} x_i ' y_i\right)$$

$$= \beta_0 + \left(\frac{1}{N}\sum_{i=1}^{N} x_i ' x_i\right)^{-1} \left(\frac{1}{N}\sum_{i=1}^{N} \Gamma_i ' F_T ' F_T \gamma_i + \frac{1}{N}\sum_{i=1}^{N} v_i F_F \gamma_i + \frac{1}{N}\sum_{i=1}^{N} x_i ' \varepsilon_i\right).$$

It is easy to check (see Lemma 1 in the Appendix) that, as $N \to \infty$, $\frac{1}{N}\sum_{i=1}^{N} x_i ' \varepsilon_i \to^P 0$, the term $\frac{1}{N}\sum_{i=1}^{N} x_i ' x_i$ is bounded and converges in probability to a fixed symmetric positive definite matrix , $\frac{1}{N}\sum_{i=1}^{N} v_i ' F_T \gamma_i \to^P 0$ and $\frac{1}{N}\sum_{i=1}^{N} \Gamma_i ' F_T ' F_T \gamma_i \to^P \Gamma ' F_T ' F_T \gamma \neq 0$ under Assumption 1, 2 and 3. Therefore, the OLS estimator is not consistent.

Assumption 3 could be replaced by a more general assumption in which the Γ_i 's are either fixed or independently distributed random variables with mean $E(\Gamma_i) = \overline{\Gamma}_i$ and $\overline{\Gamma}_i \neq 0$ for some $i = 1,2,...,N$. However, in these situations, the OLS estimator of β_0 may or may not be consistent. For example, if the quantities $\sum_{i=1}^{N} \Gamma_i '$ (for fixed Γ_i 's) or $\sum_{i=1}^{N} \overline{\Gamma}_i '$ (for random Γ_i 's) are $o(N)$, the error cross sectional dependence does not affect the consistency of the OLS estimator[4].

To develop consistent estimators of β_0 when $N \to \infty$ and T is fixed, we notice that, given Assumptions 1, 2 and 3, the following conditional moments hold

---

[4] This can be seen in the case where the Γ_i 's are random variable as follows. Under Assumptions 1 and 2, if $\left\|\overline{\Gamma}_i\right\| < \infty$ and if $\lim_{N\to\infty} \left(1/N^2\right)\sum_{i=1}^{n} \left\|Var(\Gamma_i)\right\| < \infty$, then Chebychev's week law of large number implies that $\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} \Gamma_i ' f_t f_t ' \gamma_i - \left(\frac{1}{N}\sum_{i=1}^{N} \overline{\Gamma}_i '\right) \sum_{t=1}^{T} \left(\frac{1}{T} f_t f_t '\right) \gamma \to^P 0$. Notice that $\frac{1}{N}\sum_{i=1}^{N} \overline{\Gamma}_i '$ converges to zero if $\sum_{i=1}^{N} \overline{\Gamma}_i ' = o(N)$ and in this case the OLS estimator of β_0 is consistent. The argument can be modified of the case in which the Γ_i 's are fixed.

(4) $$E\left[x_i{}'e_i \mid F_T\right] = E\left[x_i{}'e_j \mid F_T\right] = \Gamma' F_T{}' F_T \gamma$$

$i, j = 1,2,...,N$. Unfortunately, replacing these population moments with naive empirical moments does not lead to a consistent estimator of β_0. The intuition goes as follows. The empirical moments are written as

(5) $$\frac{1}{N}\sum_{i=1}^{N} x_i{}'\left(e_i - e_j\right) = 0$$

where j is fixed as a reference individual. When N is allowed to tend to infinity with fixed T, the left hand side of (5) does not converge to the zero vector and the empirical moment conditions will not generate consistent estimators. In fact, if we focus on equation (5) we can write:

$$\frac{1}{N}\sum_{i=1}^{N} x_i{}'\left(e_i - F_T\gamma_i\right) - \frac{1}{N}\sum_{i=1}^{N} x_i{}'\left(e_j - F_T\gamma\right) + \frac{1}{N}\sum_{i=1}^{N} x_i{}'\left(F_T\gamma_i - F_T\gamma\right)$$

$$= \frac{1}{N}\sum_{i=1}^{N} x_i{}'\varepsilon_i - \left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)'\left(e_j - F_T\gamma\right) + \frac{1}{N}\sum_{i=1}^{N} x_i{}'F_T\left(\gamma_i - \gamma\right)$$

$$= o_p(1) - \left(\Gamma' F_T{}' + o_p(1)\right)\left(e_j - F_T\gamma\right) + o_p(1)$$

$$= -\Gamma' F_T{}'\left(e_j - F_T\gamma\right) + o_p(1),$$

where the third line above follows from Lemma 1 in the Appendix. The presence of e_j implies that the empirical moments condition (5) does not vanish asymptotically (instead it converges to a non-degenerate random variable) and, as a consequence, it does not yield a consistent estimator of β_0.

Since the inconsistencies arise from the presence in (5) of a term (i.e. e_j) which is not affected when N is allowed to grow to infinity, it is natural to replace it with a term which has the same expectation in finite samples but is such that the empirical moment conditions are $o_p(1)$ as $N \to \infty$. This paper explores two alternative ways of achieving this.

The first method involves averaging the moments over both i and j,

$$\frac{1}{N}\sum_{j=1}^{N}\frac{1}{N}\sum_{i=1}^{N} x_i{}'\left(e_i - e_j\right) = 0.$$

Simple algebra shows that these moment conditions can be written as

(6) $$\frac{1}{N}\sum_{i=1}^{N} x_i{}'\left(e_i - \overline{e}\right) = 0,$$

where $\bar{e} = \frac{1}{N}\sum_{i=1}^{N} e_i$. Notice that the left-hand side of (6) satisfies

$$\frac{1}{N}\sum_{i=1}^{N} x_i'\left(e_i - F_T\gamma_i\right) - \frac{1}{N}\sum_{i=1}^{N} x_i'\left(\bar{e} - F_T\gamma\right) + \frac{1}{N}\sum_{i=1}^{N} x_i'\left(F_T\gamma_i - F_T\gamma\right)$$

$$= \frac{1}{N}\sum_{i=1}^{N} x_i'\varepsilon_i - \left(\frac{1}{N}\sum_{i=1}^{N} x_i'\right)\left(\bar{e} - F_T\gamma\right) + \frac{1}{N}\sum_{i=1}^{N} x_i' F_T\left(\gamma_i - \gamma\right)$$

$$= o_p(1) - \left(\Gamma' F_T' + o_p(1)\right) o_p(1) + o_p(1)$$

$$= o_p(1).$$

The third line in the above equation follows from Lemma 1 in the Appendix. Let $\bar{y} = (1/N)\sum_{i=1}^{N} y_i$ and $\bar{x} = (1/N)\sum_{i=1}^{N} x_i$. The mean-standardised moments (MSM) estimator is

$$\hat{\beta}_{MSM} = \left[\sum_{i=1}^{N} x_i'\left(x_i - \bar{x}\right)\right]^{-1} \sum_{i=1}^{N} x_i'\left(y_i - \bar{y}\right). \tag{7}$$

The MSM estimator is the ordinary least squares estimator for the transformed model

$$y_i - \bar{y} = \left(x_i - \bar{x}\right)\beta_0 + e_i - \bar{e},$$

Thus, it is a fixed effects estimator for the model

$$y_i = \alpha + x_i\beta_0 + e_i$$

where α is a $T \times 1$ vector of fixed effects (notice that the i and t dimensions are interchanged from the "standard" fixed effects estimator).

An alternative estimator of β_0 can be obtained by choosing $j = i-1$,

$$\frac{1}{N}\sum_{i=2}^{N} x_i'\left(e_i - e_{i-1}\right) = 0, \tag{8}$$

and as N tends to infinity the left-hand side is

$$\frac{1}{N}\sum_{i=2}^{N} x_i'\left(e_i - F_T\gamma_i\right) - \frac{1}{N}\sum_{i=2}^{N} x_i'\left(e_{i-1} - F_T\gamma_{i-1}\right) + \frac{1}{N}\sum_{i=2}^{N} x_i'\left(F_T\gamma_i - F_T\gamma_{i-1}\right)$$

$$= \frac{1}{N}\sum_{i=2}^{N} x_i'\varepsilon_i - \frac{1}{N}\sum_{i=2}^{N} x_i'\varepsilon_{i-1} + \frac{1}{N}\sum_{i=2}^{N} x_i' F_T\left(\gamma_i - \gamma\right) - \frac{1}{N}\sum_{i=2}^{N} x_i' F_T\left(\gamma_{i-1} - \gamma\right)$$

$$= o_p(1).$$

The third line above follows from the second one using the results of Lemma 1 in the Appendix. The differenced moments (DM) estimator is thus

(9) $$\hat{\beta}_{DM} = \left[\sum_{i=2}^{N} x_i '(x_i - x_{i-1}) \right]^{-1} \sum_{i=2}^{N} x_i '(y_i - y_{i-1}).$$

The DM estimator is close to a first difference (FD) estimator of the form

(10) $$\hat{\beta}_{FD} = \left[\sum_{i=2}^{N} (x_i - x_{i-1})'(x_i - x_{i-1}) \right]^{-1} \sum_{i=2}^{N} (x_i - x_{i-1})'(y_i - y_{i-1}).$$

Notice again that this is not the "standard" fixed effects estimator because i and t have been interchanged. The distributions of these three estimators are given in the following theorem.

Theorem 1. *Given assumptions 1, 2 and 3 the following results hold for the MSM, DM and FD estimators.*

(1) $\hat{\beta}_{MSM}$, $\hat{\beta}_{DM}$ *and* $\hat{\beta}_{FD}$ *are unbiased.*

(2) $\hat{\beta}_{MSM}$, $\hat{\beta}_{DM}$ *and* $\hat{\beta}_{FD}$ *are consistent.*

(3) Their asymptotic distributions are respectively:

a. $\sqrt{N}\left(\hat{\beta}_{MSM} - \beta_0\right) | F_T \to^D N\left(0, \Sigma(F_T)\right)$;

b. $\sqrt{N}\left(\hat{\beta}_{DM} - \beta_0\right) | F_T \to^D N\left(0, 2 \cdot \Sigma(F_T)\right)$;

c. $\sqrt{N}\left(\hat{\beta}_{FD} - \beta_0\right) | F_T \to^D N\left(0, (3/2) \cdot \Sigma(F_T)\right)$.

where

$$\Sigma(F_T) = Q(F_T)^{-1}\left(A(F_T) + B(F_T)\right) Q(F_T)^{-1}$$

$$Q(F_T) = \sum_{t=1}^{T} \left(f_t '\otimes I_p\right) \Sigma_\Gamma \left(f_t \otimes I_p\right) + \Sigma_{v0}$$

$$\Sigma_{v0} = E\left[v_i 'v_i\right] = \sum_{t=1}^{T} E\left[v_{it} v_{it} '\right]$$

$$A(F_T) = E\left[(\Gamma_i - \Gamma)' F_T '\left(F_T \Sigma_\gamma F_T '+ \Sigma_\varepsilon\right) F_T (\Gamma_i - \Gamma) | F_T\right]$$

$$B(F_T) = E\left[v_i '\left(F_T \Sigma_\gamma F_T '+ \Sigma_\varepsilon\right) v_i | F_T\right].$$

Theorem 1 shows that the MSM, DM and FD estimators are unbiased and consistent for fixed T , and that their asymptotic distributions conditional on the factors are normal.

Both the DM and FD estimators, although unbiased and consistent, are not asymptotically efficient, since their asymptotic covariance matrices are multiples (with coefficient of proportionality larger than 1) of the asymptotic covariance matrix for the MSM estimator. Essentially, the DM and FD estimators are equivalent to the MSM estimator calculated respectively on half and two thirds of the sample only.

The asymptotic covariance matrix conditional on F_T can be estimated as follows. Let

$$\widehat{\Omega(F_T)} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})'\left(y_i - \bar{y} - (x_i - \bar{x})\hat{\beta}_{MSM}\right)\left(y_i - \bar{y} - (x_i - \bar{x})\hat{\beta}_{MSM}\right)'(x_i - \bar{x})$$

$$= \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})'(e_i - \bar{e})(e_i - \bar{e})'(x_i - \bar{x})$$

(11)
$$-\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})'(e_i - \bar{e})\left(\hat{\beta}_{MSM} - \beta_0\right)'(x_i - \bar{x})'(x_i - \bar{x})$$

$$-\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})'(x_i - \bar{x})\left(\hat{\beta}_{MSM} - \beta_0\right)(e_i - \bar{e})'(x_i - \bar{x})$$

$$+\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})'(x_i - \bar{x})\left(\hat{\beta}_{MSM} - \beta_0\right)\left(\hat{\beta}_{MSM} - \beta_0\right)'(x_i - \bar{x})'(x_i - \bar{x}).$$

From Lemma A.2 we have that

$$\widehat{\Omega(F_T)} = \frac{1}{N}\sum_{i=1}^{N}(x_i - F_T\Gamma)'(e_i - F_T\gamma)(e_i - F_T\gamma)'(x_i - F_T\Gamma) + o_p(1)$$

$$\to^P E\left[(x_i - F_T\Gamma)'(e_i - F_T\gamma)(e_i - F_T\gamma)'(x_i - F_T\Gamma) \mid F_T\right].$$

Therefore, a consistent estimator of the conditional covariance matrix is

(12)
$$\widehat{\Sigma(F_T)} = \left[\frac{1}{N}\sum_{i=1}^{N}x_i'(x_i - \bar{x})\right]^{-1}\widehat{\Omega(F_T)}\left[\frac{1}{N}\sum_{i=1}^{N}x_i'(x_i - \bar{x})\right]^{-1} \to^P \Sigma(F_T).$$

Notice that the use of the asymptotic normality in Theorem 1 and the estimator of the asymptotic conditional covariance matrix in (12) implies a frequentist framework in which the available sample is one realization of an infinite sequence of identical experiments in which F_T does not change. If F_T is random, as we have assumed in Assumption 2, then different values of F_T will be realized in this infinite sequence of experiments. Thus the relevant distribution for the estimators we consider is not the one conditional on F_T but the unconditional one, in which F_T has been averaged out. Employing ≡

to denote equality in distribution (e.g. Phillips (1989)), the asymptotic distributions of $\hat{\beta}_{MSM}$, $\hat{\beta}_{DM}$ and $\hat{\beta}_{FD}$ are given in the following corollary.

Corollary 1. *Given Assumptions 1,2, and 3 the asymptotic distributions of* $\hat{\beta}_{MSM}$, $\hat{\beta}_{DM}$ *and* $\hat{\beta}_{FD}$ *are*

$$\sqrt{N}\left(\hat{\beta}_{MSM}-\beta_0\right)\to^D X,$$

$$\sqrt{N}\left(\hat{\beta}_{DM}-\beta_0\right)\to^D \sqrt{2}\times X,$$

and

$$\sqrt{N}\left(\hat{\beta}_{FD}-\beta_0\right)\to^D \sqrt{3/2}\times X$$

where

$$X\equiv\int N\left(0,\Sigma\left(F_T\right)\right)pdf\left(F_T\right)dF_T.$$

Removing the conditioning on F_T, the asymptotic distributions of the MSM, DM and FD estimators are covariance-matrix-mixed normal with mixing density given by the density function of the factors (c.f. Corollary 1). Notice that the mixing density is the distribution of F_T, which is unknown.

Before considering a random coefficient model we deal with the problem of hypothesis testing in this set-up. Even if the relevant distributions for the MSM, DM and FD estimators are the unconditional ones which are nonstandard, tests of hypotheses can be constructed as usual. An asymptotic version of the F test conditional on F_T for the null hypothesis that $H_0: R\beta_0 = r$ against the alternative hypothesis $H_0: R\beta_0 \neq r$, where R is a known and fixed $q\times p$ matrix of rank $q<p$ and r is a known and fixed $q\times 1$ vector can be easily constructed since

$$N\left(R\hat{\beta}_{MSM}-r\right)'\left(R\widehat{\Sigma\left(F_T\right)}R'\right)^{-1}\left(R\hat{\beta}_{MSM}-r\right)|\,F_T\to^d \chi^2\left(q\right), \tag{13}$$

under the null hypothesis. Notice that the chi-square random variable on the right-hand-side does not depend on F_T, so that the left-hand-side of (13) will converge to a $\chi^2\left(q\right)$ unconditionally.

Asymptotic versions of the t-test can be constructed similarly. If we denote by $\widehat{\sigma_i^2\left(F_T\right)}$ the element in position ii of the matrix $\widehat{\Sigma\left(F_T\right)}$, we can test the null hypothesis that the i-th component of β_0, β_{0i} equals a fixed value r by noting that under the null hypothesis $\dfrac{\hat{\beta}_{MSM,i}-r}{\sqrt{\widehat{\sigma_i^2\left(F_T\right)}}}|\,F_T\to N\left(0,1\right)$.

Once again the limiting distribution under the null hypothesis does not depend on F_T so that $\frac{\hat{\beta}_{MSM,i}-r}{\sqrt{\widehat{\sigma_i^2(F_T)}}} \to N(0,1)$. We summarise this in the following corollary

Corollary 2. *Let R be a known and fixed $q \times p$ matrix of rank $q < p$ and r be a known and fixed $q \times 1$ vector. Given Assumptions 1, 2, and 3, if the null hypothesis $H_0 : R\beta_0 = r$ holds then*

$$N\left(R\hat{\beta}_{MSM}-r\right)'\left(R\widehat{\Sigma(F_T)}R'\right)^{-1}\left(R\hat{\beta}_{MSM}-r\right) \to^d \chi^2(q). \quad (14)$$

Moreover, if $q = 1$

$$\frac{R\hat{\beta}_{MSM}-r}{\sqrt{R\widehat{\Sigma(F_T)}R'}} \to N(0,1), \quad (15)$$

where $\widehat{\Sigma(F_T)}$ is defined in (12).

Notice that the distributions of the two test statistics above under the null hypothesis do not depend on F_T, however, they do depend on F_T under the alternative hypothesis.

3. Heterogeneous slopes

In this section, we consider a more general case, where the coefficient β_0 is allowed to be different for each individual. Precisely, the model is

$$\underset{(T\times 1)}{y_i} = \underset{(T\times p)}{x_i}\underset{(p\times 1)}{\beta_i} + e_i \quad (16)$$

$$\underset{(T\times 1)}{e_i} = \underset{(T\times m)}{F_T}\underset{(m\times 1)}{\gamma_i} + \varepsilon_i \quad (17)$$

$$x_i = F_T \underset{(m\times p)}{\Gamma_i} + v_i \quad (18)$$

$$\beta_i = \beta_0 + \eta_i . \quad (19)$$

We are interested in inference about the mean of the individual-specific slope coefficients. This random coefficients model was considered by Pesaran (2006) and is expected to satisfy Assumptions 1 to 3 as well as Assumption 4 below.

Assumption 4.

The η_i's are i.i.d. with mean zero and covariance matrix Σ_η. Moreover they are independent of f_t, v_i, Γ_i, ε_i and γ_i. Moreover $\left\|\Sigma_\eta\right\| < M < \infty$.

Once again this assumption is standard in the literature (e.g. Pesaran (2006)). Notice also that Pesaran (2006) specifies equation (16) as

$$(20) \qquad y_i = D_T' \underset{(n\times 1)}{\alpha_i} + x_i\beta_i + e_i .$$

and (18) as

$$(21) \qquad x_i = D_T' \underset{(n\times p)}{A_i} + F_T\Gamma_i + v_i ,$$

where $D_T = \left[d_1, d_2, \ldots \underset{(n\times 1)}{d_T} \right]$. Since T is fixed and we are conditioning on the factors we can always rewrite this as in equations (16) to (19) with F_T, γ_i and Γ_i replaced, respectively, by $\left[D_T', F_T\right]$, $\begin{bmatrix}\alpha_i \\ \gamma_i\end{bmatrix}$ and $\begin{bmatrix}A_i \\ \Gamma_i\end{bmatrix}$. When removing the conditioning on F_T we will just have to take into account that this now has two components and we will still condition on part of it (i.e. D_T).

Pesaran (2006) suggested a procedure whereby β_i is estimated for each fixed T by

$$(22) \qquad \hat{b}_i = \left(x_i'\bar{M}_w x_i\right)^{-1} x_i'\bar{M}_w y_i$$

where $\bar{M}_w = I_T - \bar{H}\left(\bar{H}'\bar{H}\right)^{-} \bar{H}'$, $\bar{H} = \left[\bar{y}, \bar{x}\right]$, ($\left(\bar{H}'\bar{H}\right)^{-}$ denotes the generalized inverse of $\bar{H}'\bar{H}$), then β_0 is estimated by the "common correlated effects mean group" (CCEMG) estimator

$$(23) \qquad \hat{b}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{b}_i ,$$

or the "common correlated effects pooled" (CCEP) estimator

$$(24) \qquad \hat{b}_{CCEP} = \left(\sum_{i=1}^{N} x_i'\bar{M}_w x_i\right)^{-1} \sum_{i=1}^{N} x_i'\bar{M}_w y_i .$$

For simplicity we have taken the aggregating and the pooling weights to be the same in (22) and (24).

These two estimators have some obvious disadvantages. The first one requires T to be larger than p to guarantee that $\hat{b}_i$ can be calculated for every i. Both estimators involve $\bar{y}$ and $\bar{x}$ through the projection matrix $\bar{M}_w$, and this makes it difficult to determine their small sample properties. Finally, and most importantly, their properties are known only if both T and N are large.

The MSM, DM and FD estimators defined in (7), (9) and (10) are very easy to calculate and allow one to overcome the reliance on large T results. In fact, our results also apply to the case where $T=1$. By considering T fixed, we do not need to make any assumption about the time series properties of the data.

Theorem 2 gives some of the distributional properties for the MSM, DM and FD estimators.

Theorem 2. *Given assumptions 1, 2, 3 and 4 the following results hold for the MSM, DM and FD estimators.*

(1) The MSM, DM and FD estimators are unbiased;

(2) They are consistent, and

(3) Their asymptotic distributions are respectively:

a. $$\sqrt{N}\left(\hat{\beta}_{MSM}-\beta_0\right)|F_T \to^D N\left(0,\Sigma(F_T)+Q(F_T)^{-1}C(F_T)Q(F_T)^{-1}\right)$$

b. $$\sqrt{N}\left(\hat{\beta}_{DM}-\beta_0\right)|F_T \to^D N\left(0,2\cdot\Sigma(F_T)+Q(F_T)^{-1}\left(C(F_T)+D(F_T)\right)Q(F_T)^{-1}\right)$$

c. $$\sqrt{N}\left(\hat{\beta}_{FD}-\beta_0\right)|F_T \to^D N\left(0,\frac{3}{2}\cdot\Sigma(F_T)+Q(F_T)^{-1}\left(C(F_T)+\frac{1}{2}D(F_T)\right)Q(F_T)^{-1}\right).$$

where

$$\Sigma(F_T)=Q(F_T)^{-1}\left(A(F_T)+B(F_T)\right)Q(F_T)^{-1}$$

$$Q(F_T)=\sum_{t=1}^{T}\left(f_t'\otimes I_p\right)\Sigma_\Gamma\left(f_t\otimes I_p\right)+\Sigma_{v0}$$

$$A(F_T)=E\left[(\Gamma_i-\Gamma)'F_T'\left(F_T\Sigma_\gamma F_T'+\Sigma_\varepsilon\right)F_T(\Gamma_i-\Gamma)|F_T\right]$$

$$B(F_T)=E\left[v_i'\left(F_T\Sigma_\gamma F_T'+\Sigma_\varepsilon\right)v_i|F_T\right]$$

$$C(F_T)=E\left[\left(F_T(\Gamma_i-\Gamma)+v_i\right)'\left(F_T\Gamma_i+v_i\right)\Sigma_\eta\left(F_T\Gamma_i+v_i\right)'\left(F_T(\Gamma_i-\Gamma)+v_i\right)|F_T\right]$$

$$D(F_T)=E\left[\left(F_T(\Gamma_i-\Gamma)+v_i\right)'\left(F_T\Gamma_{i-1}+v_{i-1}\right)\Sigma_\eta\left(F_T\Gamma_{i-1}+v_{i-1}\right)'\left(F_T(\Gamma_i-\Gamma)+v_i\right)|F_T\right],$$

and Σ_{v0} *is defined in Theorem 1.*

Theorem 2 shows that for a fixed T, all three of the MSM, DM and FD estimators are unbiased, consistent and asymptotically normal given F_T. The MSM estimator is asymptotically more efficient than the DM and FD estimators in terms of conditional variance when N is large. Notice that for this case, the asymptotic covariance matrix depends up to the fourth moments of Γ_i and v_i. As argued in the previous section, it is the marginal distribution which is relevant, so removing the conditioning on F_T, we obtain the following corollary.

Corollary 3. *Given assumptions 1,2, 3 and 4 the asymptotic distributions of the MSM, DM and FD estimators are*

$$\sqrt{N}\left(\hat{\beta}_{MSM}-\beta_0\right)\rightarrow^D \int N\left(0,\Sigma\left(F_T\right)+Q\left(F_T\right)^{-1}C\left(F_T\right)Q\left(F_T\right)^{-1}\right)pdf\left(F_T\right)dF_T,$$

$$\sqrt{N}\left(\hat{\beta}_{DM}-\beta_0\right)\rightarrow^D \int N\left(0,2\cdot\Sigma\left(F_T\right)+Q\left(F_T\right)^{-1}\left(C\left(F_T\right)+D\left(F_T\right)\right)Q\left(F_T\right)^{-1}\right)pdf\left(F_T\right)dF_T,$$

and

$$\sqrt{N}\left(\hat{\beta}_{FD}-\beta_0\right)\rightarrow^D \int N\left(0,\frac{3}{2}\cdot\Sigma\left(F_T\right)+Q\left(F_T\right)^{-1}\left(C\left(F_T\right)+\frac{1}{2}D\left(F_T\right)\right)Q\left(F_T\right)^{-1}\right)pdf\left(F_T\right)dF_T.$$

Thus, the three estimators are asymptotically covariance-matrix-mixed normal. Notice that Theorem 2 and Corollary 3 reduce to Theorem 1 and Corollary 1 if the η_i are identically zero.

In order to construct tests on β_0 we need to find a consistent estimator of the asymptotic covariance matrix $\Omega\left(F_T\right)=A\left(F_T\right)+B\left(F_T\right)+C\left(F_T\right)$. Such a consistent estimator of $\widehat{\Omega\left(F_T\right)}$ is given in the following theorem.

Theorem 3. *Given Assumptions 1-4 a consistent estimator of* $\Omega\left(F_T\right)$ *is*

$$\text{(25)}\quad \widehat{\Omega\left(F_T\right)}=\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\overline{x}\right)'\left(y_i-\overline{y}-\left(x_i-\overline{x}\right)\hat{\beta}_{MSM}\right)\left(y_i-\overline{y}-\left(x_i-\overline{x}\right)\hat{\beta}_{MSM}\right)'\left(x_i-\overline{x}\right),$$

so that a consistent estimator of asymptotic covariance matrix of the MSM estimator is

$$\text{(26)}\quad \widehat{\Xi\left(F_T\right)}=\left[\frac{1}{N}\sum_{i=1}^{N}x_i'\left(x_i-\overline{x}\right)\right]^{-1}\widehat{\Omega\left(F_T\right)}\left[\frac{1}{N}\sum_{i=1}^{N}x_i'\left(x_i-\overline{x}\right)\right]^{-1}.$$

Asymptotic tests of hypothesis can be constructed as outlined in the previous section.

Corollary 4. *If the conditions stated in Corollary 2 and Assumption 4 apply, then (14) and (15) hold with* $\widehat{\Sigma\left(F_T\right)}$ *replaced by* $\widehat{\Xi\left(F_T\right)}$ *in (26).*

4. Monte Carlo Study

This section provides some Monte Carlo evidence on the properties of the MSM, DM and FD and comparisons with the CCEMG and CCEP estimators. To do this we consider a simplified version of data generating process (DGP) used by Pesaran (2006) in his Section 7. Precisely, we assume the DGP is

(27) $$y_{it} = \alpha_{i1} d_{1t} + \beta_{i1} x_{i1t} + \beta_{i2} x_{i2t} + e_{it} \,,$$

(28) $$e_{it} = \gamma_{i1} f_{1t} + \gamma_{i2} f_{2t} + \gamma_{i3} f_{3t} + \varepsilon_{it}$$

and

(29) $$x_{ijt} = a_{ij1} d_{1t} + a_{ij2} d_{2t} + \Gamma_{ij1} f_{1t} + \Gamma_{ij2} f_{2t} + \Gamma_{ij3} f_{3t} + v_{ijt} \,,$$

$j = 1,2$, $i = 1,2,...,N$ and $t = 1,...,T$. Notice that we consider the case where β has only two components to allow a comparison with the results of Pesaran (2006) because the CCEMG cannot be calculated when T is less or equal to the dimension of β .

The common factors and the errors in (29) are generated as follows

$$f_{jt} = \rho_{fj} f_{jt-1} + i.i.d.N(0,1) \text{ for } j = 1,2, \; t = -49,...,0,...,T$$

$$f_{j,-50} = 0$$

$$v_{ijt} = \rho_{vij} v_{ijt-1} + i.i.d.N(0,1) \text{ for } v, j = 1,2, \; t = -49,...,0,...,T$$

$$v_{ij,-50} = 0, \quad \rho_{vij} \sim i.i.d.U[0.05, 0.95].$$

The errors in (28) are generated as AR(1) process

$$\varepsilon_{it} = \rho_{i\varepsilon} \varepsilon_{i,t-1} + \sigma_i \left(1 - \rho_{i\varepsilon}^2\right)^{1/2} \zeta_{it}$$

for $i = 1,...,[N/2]$, $t = 0,...,T$, and as MA(1)

$$\varepsilon_{it} = \sigma_i \left(1 + \theta_{i\varepsilon}{}^2\right)^{-1/2} \left(\zeta_{it} + \theta_{i\varepsilon} \zeta_{it-1}\right)$$

for $i = [N/2] + 1,...,N$ where $\zeta_{it} \sim i.i.d.N(0,1)$, $\sigma_i^2 \sim i.i.d.U(0.5,1.5)$, $\rho_{i\varepsilon} \sim i.i.d.U[0.05,0.95]$ and $\theta_{i\varepsilon} \sim i.i.d.U[0,1]$

The factor loadings in (29) are generated independently as $\Gamma_{ijp} \sim i.i.d.N\left(\mu_{jp}, \sigma_\Gamma^2\right)$ and $\mu_{jp} \sim i.i.d.U(-.5,1.5)$ for $p = 1,2,3$, and those in (28) as $\gamma_{i1} \sim i.i.d.N\left(1, \sigma_\gamma^2\right)$, $\gamma_{i2} \sim i.i.d.N\left(1, \sigma_\gamma^2\right)$,

$\gamma_{i3} = 0$, $\sigma_{\Gamma}^2 = 0.7$ and $\sigma_{\gamma}^2 = 0.5$. Finally, for the slopes coefficients, we consider the case where $\beta_{01} = \beta_{02} = 1$ and $\beta_{ij} = \beta_{0j} + \eta_i$ where $\eta_i \sim i.i.d.N(0, 0.4)$ and we will report information for the case where $\sigma_{\eta}^2 = 0$ (the model of Section 2) and $\sigma_{\eta}^2 > 0$ (the random coefficient model of Section 3).

We will report results for some combinations of the parameters (further results are available from the Authors' webpages). Tables 1 to 4 impose $\alpha_{i1} = 0$, $a_{ij1} = 0$ and $a_{ij2} = 0$. However, we present results for a model with fixed effects in Tables 5-8. In this case we generate the fixed effects as

$$\begin{aligned} &d_{1t} = 1 \\ &d_{2t} = .5 d_{2t-1} + i.i.d.N\left(0, 1 - (.5)^2\right) \quad t = -49, ..., 0, ..., T \\ &d_{2,-50} = 0 \end{aligned}$$

and $\alpha_{it} \sim i.i.d.N(1,1)$ and the a_{ijk} for $k = 1, 2$ are independent $N(0.5, 0.5)$ as in Pesaran (2006).

We report the bias and mean square error (MSE) for the MSM, DM and FD estimators and for the CCEMG and CCEP estimators of Pesaran (2006). We also report results for a two-sided "t test" for $H_0 : \beta_{01} = 1$. The Monte Carlo experiments are based on 10,000 replications and are summarized in Tables 1 to 8.

4.1 Bias and MSE

The bias for the MSM, DM and FD estimators is very small for every sample size as one would expect from the fact that they are unbiased. In contrast, the bias of the CCEP and CCEMG is small only for large T. For very small T (say less than 5) these two estimators either cannot be computed or have bias comparable or larger than that of the OLS estimator. The CCEMG estimator has a considerably larger bias than all other estimators for small T.

The MSE of the MSM, DM and FD decreases as N becomes large but does not seem to be strongly affected by T. The MSM estimator performs better than the DM and FD estimators for all sample sizes as expected from our theoretical results in Theorems 1 and 2. The MSE of both the CCEP and CCEMG decreases as either N or T grows and it can be extremely large for small T.

The presence of fixed effects (Tables 5 to 8) does not affect the bias or the MSE of the MSM, DM and FD estimator but considerably affects the CCEP and the CCEMG estimators. As remarked by Pesaran (2006), the CCEP estimator performs better than the CCEMG estimator for small T. Simulations (that are not reported here but are available on the Authors' website) show that the bias and the MSE of the MSM, DM, and FD estimators are not affected by either the dimension of β_0 or the number of factors. Pesaran's CCEP and CCEMG estimators, on the other hand, are severely affected.

4.2 Size and power

Tables 1 to 4 report a t-type test for $H_0 : \beta_{01} = 1$ against $H_0 : \beta_{01} \neq 1$ constructed using the MSM, CCEP and CCEMG estimators. The test based on the MSM estimator tends to be slightly oversized but the empirical size is very close to the theoretical 5% size for $N > 100$. Also the size for the test based on the MSM does not seem to be affected by T. The power of the MSM based test seems to increase with T.

The tests based on CCEP and CCEMG estimators are sometimes considerably oversized and sometimes undersized when T is small. The size for the CCEP based test can be as large as 0.35 for small T. For the CCEMG based test the size is closer to the nominal size. For $T > 10$ the empirical size is very close to the nominal one for the tests based on both estimators. Their power is very much affected by T and for small T both tests have power smaller than the size.

Simulations available from the Authors' website show that size and power of the test based on the MSM estimator are not affected by either number of factors or the dimension of the parameter β_0, but both size and power of the tests based on the CCEP and the CCEMG estimators are influenced by these in a significant way.

5. Conclusions

This paper has analysed a panel data model with both homogeneous and heterogeneous slopes and with multifactor error structure in which the factors also affect the regressors under conditions that are weaker than those of Pesaran (2006). Consistent estimators have been proposed for the case where the time dimension is fixed and N tends to infinity. These estimators exist for every $T \geq 1$ and are very simple to compute. As $N \to \infty$ they have a nonstandard asymptotic distribution which is covariance-matrix mixed normal with mixing density given by the unknown distribution of the error factors. However, tests on the interest parameters can be constructed using standard t- and F-procedures. A Monte Carlo simulation has shown that these estimators clearly outperform the CCEP and CCEMG estimators of Pesaran (2006) when the time dimension is small in comparison to the dimension of the slope parameter.

Appendix: technical results

We first present a lemma giving all the basic results used in the appendix. The proofs are simple and can be found in the authors' webpage.

Lemma A.1. *Given Assumptions 1, 2, 3 and 4 the following results hold as* $N \to \infty$*:*

a) $\frac{1}{N}\sum_{i=1}^{N} v_i'\varepsilon_i \to^P 0$

b) $\frac{1}{N}\sum_{i=1}^{N} \Gamma_i' F_T'\varepsilon_i \to^P 0$;

c) $\frac{1}{N}\sum_{i=1}^{N} x_i' F_T\gamma_i \to^P \Gamma' F_T' F_T\gamma$;

d) $\frac{1}{N}\sum_{i=1}^{N} x_t \to^P F_T\Gamma$;

e) $\frac{1}{N}\sum_{i=1}^{N} x_i' x_i \to^P \sum_{t=1}^{T}\left(f_t'\otimes I_p\right)\Sigma_\Gamma\left(f_t\otimes I_p\right)+\Gamma' F_T' F_T\Gamma+\Sigma_{v0}$;

f) $\frac{1}{N}\sum_{i=1}^{N} x_i'\left(x_i-\bar{x}\right) \to^P \sum_{t=1}^{T}\left(f_t'\otimes I_p\right)\Sigma_\Gamma\left(f_t\otimes I_p\right)+\Sigma_{v0}$;

g) $\frac{1}{N}\sum_{i=2}^{N} x_i'\left(x_i-x_{i-1}\right) \to^P \sum_{t=1}^{T}\left(f_t'\otimes I_p\right)\Sigma_\Gamma\left(f_t\otimes I_p\right)+\Sigma_{v0}$;

h) $\frac{1}{N}\sum_{i=1}^{N} e_i \to^P F_T\gamma$;

i) $\frac{1}{N}\sum_{i=2}^{N} x_i' x_i\eta_i \to^P 0$;

and Σ_{v0} *is defined in Theorem 1.*

Lemma A.2. *Given Assumptions 1, 2, 3 and 4 the following results hold*

a) $\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\bar{x}\right)'\left(e_i-\bar{e}\right)\left(e_i-\bar{e}\right)'\left(x_i-\bar{x}\right)=\frac{1}{N}\sum_{i=1}^{N}\left(x_i-F_T\Gamma\right)'\left(e_i-F_T\gamma\right)\left(e_i-F_T\gamma\right)'\left(x_i-F_T\Gamma\right)+o_p(1)$;

b) $\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\bar{x}\right)'\left(e_i-\bar{e}\right)\left(\hat{\beta}_{MSM}-\beta_0\right)'\left(x_i-\bar{x}\right)'\left(x_i-\bar{x}\right)=o_p(1)$;

c) $\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\bar{x}\right)'\left(x_i-\bar{x}\right)\left(\hat{\beta}_{MSM}-\beta_0\right)\left(\hat{\beta}_{MSM}-\beta_0\right)'\left(x_i-\bar{x}\right)'\left(x_i-\bar{x}\right)=o_p(1)$

d) $\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\bar{x}\right)'\left(x_i-\bar{x}\right)\left(\beta_0-\hat{\beta}_{MSM}\right)\eta_i'\left(x_i-\bar{x}\right)'\left(x_i-\bar{x}\right)=o_p(1)$;

e) $\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\bar{x}\right)'\left(e_i-\bar{e}\right)\eta_i'\left(x_i-\bar{x}\right)'\left(x_i-\bar{x}\right)=o_p(1)$

f) $\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\bar{x}\right)'\left(\left(x_i-\bar{x}\right)\eta_i\right)\left(\left(x_i-\bar{x}\right)\eta_i\right)'\left(x_i-\bar{x}\right)=E\left[\left(x_i-F_T\Gamma\right)'\left(x_i-F_T\Gamma\right)\Sigma_\eta\left(x_i-F_T\Gamma\right)'\left(x_i-F_T\Gamma\right)\right]+o_p(1)$

Proof

Let

$$z_i = x_i - F_T \Gamma ,$$

$$w_i = e_i - F_T \gamma$$

$$\overline{z} = \overline{x} - F_T \Gamma$$

$$\overline{w} = \overline{e} - F_T \gamma$$

and let z_i^p and $\overline{z}^p$ be the p-th columns of z_i and $\overline{z}$ respectively. Then the (p,q) component of $\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})'(e_i - \overline{e})(e_i - \overline{e})'(x_i - \overline{x})$ can be written as

$$\frac{1}{N}\sum_{i=1}^{N}\left(z_i^p - \overline{z}^p\right)'(w_i - \overline{w})(w_i - \overline{w})'\left(z_i^q - \overline{z}^q\right) = \frac{1}{N}\sum_{i=1}^{N} z_i^{p\prime} w_i w_i' z_i^q$$

+sum of terms involving either $\overline{z}^l$ or $\overline{w}$.

The terms involving either $\overline{z}^l$ or $\overline{w}$ are $o_p(1)$ so that a) is proved.

Next, we prove b). Using the same notation as above we can write the (p,q) component of $\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})'(e_i - \overline{e})\left(\hat{\beta}_{MSM} - \beta_0\right)'(x_i - \overline{x})'(x_i - \overline{x})$ as

$$\frac{1}{N}\sum_{i=1}^{N}\left(z_i^p - \overline{z}^p\right)'(w_i - \overline{w})\left(\hat{\beta}_{MSM} - \beta_0\right)'(z_i - \overline{z})'\left(z_i^q - \overline{z}^q\right)$$

$$= \left(\frac{1}{N}\sum_{i=1}^{N}\left(z_i^p - \overline{z}^p\right)'(w_i - \overline{w})\left(z_i^q - \overline{z}^q\right)'(z_i - \overline{z})\right)\left(\hat{\beta}_{MSM} - \beta_0\right) = o_p(1)$$

provided $\frac{1}{N}\sum_{i=1}^{N}\left(z_i^p - \overline{z}^p\right)'(w_i - \overline{w})\left(z_i^q - \overline{z}^q\right)'\left(z_i^r - \overline{z}^r\right)$ is $o_p(1)$. Notice that the leading term with the largest number of terms in the summation is

$$\frac{1}{N}\sum_{i=1}^{N}\left(z_i^p\right)' w_i \left(z_i^q\right)' z_i \ .$$

All other terms contain either $\overline{w}$ or $\overline{z}^p$ and are $o_p(1)$. This is a sum of i.i.d. random variables with mean zero if $\left(z_i^p\right)\left(z_i^q\right)' z_i$ have finite mean. This follows from Assumption 1.v and 3.i.

Similarly for c), the (p,q) component of

$$\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})'(x_i - \overline{x})\left(\hat{\beta}_{MSM} - \beta_0\right)\left(\hat{\beta}_{MSM} - \beta_0\right)'(x_i - \overline{x})'(x_i - \overline{x})$$

is

$$\frac{1}{N}\sum_{i=1}^{N}\left(z_i^{\,p}-\overline{z}^{\,p}\right)'\left(z_i-\overline{z}\right)\left(\hat{\beta}_{MSM}-\beta_0\right)\left(\hat{\beta}_{MSM}-\beta_0\right)'\left(z_i-\overline{z}\right)'\left(z_i^{\,q}-\overline{z}^{\,q}\right)$$

$$=\left(\hat{\beta}_{MSM}-\beta_0\right)'\left(\frac{1}{N}\sum_{i=1}^{N}\left(z_i-\overline{z}\right)'\left(z_i^{\,p}-\overline{z}^{\,p}\right)\left(z_i^{\,q}-\overline{z}^{\,q}\right)'\left(z_i-\overline{z}\right)\right)\left(\hat{\beta}_{MSM}-\beta_0\right)=o_p\left(1\right)$$

provided $\frac{1}{N}\sum_{i=1}^{N}\left(z_i-\overline{z}\right)'\left(z_i^{\,p}-\overline{z}^{\,p}\right)\left(z_i^{\,q}-\overline{z}^{\,q}\right)'\left(z_i-\overline{z}\right)$ converges in probability to a finite quantity. This follows as above using Assumption 1.v and 3.i.

For d) we can proceed similarly to c) and find that

$$\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)\left(\beta_0-\hat{\beta}_{MSM}\right)\eta_i'\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)=o_p\left(1\right).$$

This is guaranteed by Assumptions 1.v, 3.i-ii and 4.

e) $\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\overline{x}\right)'\left(e_i-\overline{e}\right)\eta_i'\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)=o_p\left(1\right)$ as in the previous part of the Lemma.

f) Write

$$\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\overline{x}\right)'\left(\left(x_i-\overline{x}\right)\eta_i\right)\left(\left(x_i-\overline{x}\right)\eta_i\right)'\left(x_i-\overline{x}\right)$$

$$=\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)\left(\eta_i\eta_i'-\Sigma_\eta\right)\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)+\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)\Sigma_\eta\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right).$$

The term $\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)\left(\eta_i\eta_i'-\Sigma_\eta\right)\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)$ is uncorrelated across i and has the same mean. By using Kinchin's Weak Law of Large Numbers,

$$\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)\left(\eta_i\eta_i'-\Sigma_\eta\right)\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)\to^{P}0.$$

Similarly to Lemma A.2 (a),

$$\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)\Sigma_\eta\left(x_i-\overline{x}\right)'\left(x_i-\overline{x}\right)\to^{P}E\left[\left(x_i-F_T\Gamma\right)'\left(x_i-F_T\Gamma\right)\Sigma_\eta\left(x_i-F_T\Gamma\right)'\left(x_i-F_T\Gamma\right)\right].$$

Proof of Theorem 1

From the definition of the MSM estimator and equation (1) we can write

$$\begin{aligned}\hat{\beta}_{MSM}&=\beta_0+\left[\sum_{j=1}^{N}x_j'\left(x_j-\overline{x}\right)\right]^{-1}\sum_{i=1}^{N}x_i'\left(e_i-\overline{e}\right)\\&=\beta_0+\left[\sum_{j=1}^{N}x_j'\left(x_j-\overline{x}\right)\right]^{-1}\sum_{i=1}^{N}x_i'\left(F_T\left(\gamma_i-\gamma\right)-F_T\left(\overline{\gamma}-\gamma\right)+\varepsilon_i-\overline{\varepsilon}\right),\end{aligned}\quad(30)$$

From assumption 2.ii and 3.ii the γ_i's are i.i.d. with mean γ and are independent of v_i and Γ_i and from assumptions 1.iii and 3.ii ε_i is also independent of v_i and Γ_i. Thus, we can write

$$E\left[\hat{\beta}_{MSM} \mid F_T\right] = \beta_0 + \sum_{i=1}^{N} E\left[\left[\sum_{j=1}^{N} x_j'\left(x_j - \overline{x}\right)\right]^{-1} x_i' \mid F_T\right] E\left[F_T\left(\gamma_i - \gamma\right) - F_T\left(\overline{\gamma} - \gamma\right) + \varepsilon_i - \overline{\varepsilon} \mid F_T\right]$$
$$= \beta_0.$$

Therefore, the MSM estimator is unbiased given F_T, and it is thus unbiased unconditionally. Consistency follows from noticing that

$$\hat{\beta}_{MSM} = \beta_0 + \left[\frac{1}{N}\sum_{i=1}^{N} x_i'\left(x_i - \overline{x}\right)\right]^{-1}\left(\frac{1}{N}\sum_{i=1}^{N} x_i' F_T \gamma_i - \left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)' F_T \overline{\gamma} + \frac{1}{N}\sum_{i=1}^{N} x_i' \varepsilon_i - \left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)' \overline{\varepsilon}\right) \to^{P} \beta_0$$

using c), e), b), f) and a) in Lemma 1.

To prove asymptotic normality conditional on F_t we write

$$\sqrt{N}\left(\hat{\beta}_{MSM} - \beta_0\right) = \left[\frac{1}{N}\sum_{i=1}^{N} x_i'\left(x_i - \overline{x}\right)\right]^{-1} \times \frac{1}{\sqrt{N}}\sum_{i=1}^{N} x_i'\left(e_i - \overline{e}\right).$$

The term on the right-hand side is

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N} x_i'\left(e_i - \overline{e}\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(x_i - F_T\Gamma\right)'\left(e_i - \overline{e}\right) + \Gamma' F_T' \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(e_i - \overline{e}\right)$$

$$= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(x_i - F_T\Gamma\right)'\left(e_i - \overline{e}\right)$$

$$= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(x_i - F_T\Gamma\right)'\left(e_i - F_T\gamma\right) - \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(x_i - F_T\Gamma\right)\right)'\left(\overline{e} - F_T\gamma\right)$$

$$= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(x_i - F_T\Gamma\right)'\left(e_i - F_T\gamma\right) + o_p(1),$$

where the last line follows from the fact that $\overline{\gamma} - \gamma = o_p(1)$, $\overline{\varepsilon} = o_p(1)$ and

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(x_i - F_T\Gamma\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(F_T\left(\Gamma_i - \Gamma\right) + v_i\right) = O_p(1).$$

The terms in the remaining sum are i.i.d. with mean zero and covariance matrix

$$E\left[\left(x_i - F_T\Gamma\right)'\left(e_i - F_T\gamma\right)\left(e_i - F_T\gamma\right)'\left(x_i - F_T\Gamma\right) \mid F_T\right] = E\left[\left(x_i - F_T\Gamma\right)' E\left[\left(e_i - F_T\gamma\right)\left(e_i - F_T\gamma\right)' \mid F_T\right]\left(x_i - F_T\Gamma\right) \mid F_T\right]$$

The expectation in the middle is

$$E\left[(e_i - F_T\gamma)(e_i - F_T\gamma)' \mid F_T\right] = E\left[\left(F_T(\gamma_i - \gamma) + \varepsilon_i\right)\left(F_T(\gamma_i - \gamma) + \varepsilon_i\right)' \mid F_T\right]$$

$$= E\left[F_T(\gamma_i - \gamma)(\gamma_i - \gamma)' F_T' \mid F_T\right] + E\left[\varepsilon_i \varepsilon_i' \mid F_T\right]$$

$$= F_T \Sigma_\gamma F_T' + \Sigma_\varepsilon .$$

So that

$$E\left[(x_i - F_T\Gamma)'(e_i - F_T\gamma)(e_i - F_T\gamma)'(x_i - F_T\Gamma) \mid F_T\right]$$

$$= E\left[\left(F_T(\Gamma_i - \Gamma) + v_i\right)'\left(F_T\Sigma_\gamma F_T' + \Sigma_\varepsilon\right)\left(F_T(\Gamma_i - \Gamma) + v_i\right) \mid F_T\right]$$

$$= A(F_T) + B(F_T)$$

where $A(F_T)$ and $B(F_T)$ are defined in Theorem 1. The result follows from Assumptions 1.ii and 2.ii and 3.i-ii.

Now we focus on the DM estimator. Proceeding as in the previous case we have

$$\hat{\beta}_{DM} = \beta_0 + \left(\sum_{i=2}^{N} x_i'(x_i - x_{i-1})\right)^{-1} \sum_{i=2}^{N} x_i'(e_i - e_{i-1}).$$

The expected value is

$$E\left[\hat{\beta}_{DM} \mid F_T\right] = \beta_0 + E\left[\left(\sum_{i=2}^{N} x_i'(x_i - x_{i-1})\right)^{-1} \sum_{i=2}^{N} x_i'(e_i - e_{i-1}) \mid F_T\right]$$

$$= \beta_0 + \sum_{i=2}^{N} E\left[\left(\sum_{j=2}^{N} x_j'(x_j - x_{j-1})\right)^{-1} x_i' \mid F_T\right] E\left[e_i - e_{i-1} \mid F_T\right]$$

$$= \beta_0$$

because of the independent of v_i and Γ_i from ε_i and γ_i. To show consistency we write

$$\hat{\beta}_{DM} = \beta_0 + \left(\sum_{i=2}^{N} x_i'(x_i - x_{i-1})\right)^{-1} \left[\sum_{i=2}^{N} x_i' F_T \gamma_i + \sum_{i=2}^{N} x_i' \varepsilon_i - \sum_{i=2}^{N} x_i' F_T \gamma_{i-1} - \sum_{i=2}^{N} x_i' \varepsilon_{i-1}\right] \to^{P} \beta_0$$

where the convergence of the various components can be easily obtained from the assumptions using the weak law of large numbers . To show asymptotic normality we write

$$\sqrt{N}\left(\hat{\beta}_{DM} - \beta_0\right) = \left(\frac{1}{N}\sum_{i=2}^{N} x_i'(x_i - x_{i-1})\right)^{-1} \times \frac{1}{\sqrt{N}} \sum_{i=2}^{N} x_i'(e_i - e_{i-1}).$$

The last term on the right-hand side is sum of terms of an m-dependent process with $m=1$ so that

$$\frac{1}{\sqrt{N}}\sum_{i=2}^{N} x_i '\left(e_i - e_{i-1}\right) \to^{D} N\left(0, \Sigma_{DM}\right)$$

where

$$\Sigma_{DM} = Var\Big(x_i'\left(e_i - e_{i-1}\right)\Big) + Cov\Big(x_2'\left(e_2 - e_1\right), x_3'\left(e_3 - e_2\right)\Big) + Cov\Big(x_3'\left(e_3 - e_2\right), x_2'\left(e_2 - e_1\right)\Big)$$

(e.g. DasGupta (2008)).

The components of Σ_{DM} are

$$Var\Big(x_i'\left(e_i - e_{i-1}\right) | F_T\Big) = E\Big[x_i'\left(e_i - e_{i-1}\right)\left(e_i - e_{i-1}\right)' x_i | F_T\Big]$$

$$= E\Big[x_i' E\Big[\left(e_i - e_{i-1}\right)\left(e_i - e_{i-1}\right)' | F_T\Big] x_i | F_T\Big]$$

$$= 2E\Big[x_i'\left(F_T \Sigma_\gamma F_T' + \Sigma_\varepsilon\right) x_i | F_T\Big]$$

$$Cov\Big(x_2'\left(e_2 - e_1\right), x_3'\left(e_3 - e_2\right) | F_T\Big) = E\Big[x_2'\left(e_2 - e_1\right)\left(e_3 - e_2\right)' x_3 | F_T\Big]$$

$$= E\Big[x_2' E\Big[\left(e_2 - e_1\right)\left(e_3 - e_2\right)' | F_T\Big] x_3 | F_T\Big]$$

$$= E\Big[\Big(\Gamma_2' F_T' + v_2'\Big)\Big(-F_T \Sigma_\gamma F_T' - \Sigma_\varepsilon\Big)\left(\Gamma_3 F_T + v_3\right)\Big]$$

$$= -\Gamma' F_T'\Big(F_T \Sigma_\gamma F_T' + \Sigma_\varepsilon\Big)\Gamma F_T$$

$$= Cov\Big(x_3'\left(e_3 - e_2\right), x_2'\left(e_2 - e_1\right) | F_T\Big).$$

The result follows from noting that $\Sigma_{DM} = 2\left(A\left(F_T\right) + B\left(F_T\right)\right)$ and g) of Lemma A.1.

For the FD estimator unbiasedness and consistency follow as in the previous two cases. For asymptotic normality, we write

$$\sqrt{N}\left(\hat{\beta} - \beta_0\right) = \left[\frac{1}{N}\sum_{i=1}^{N}\left(x_i - x_{i-1}\right)'\left(x_i - x_{i-1}\right)\right]^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(x_i - x_{i-1}\right)'\left(e_i - e_{i-1}\right).$$

The first term is $\frac{1}{N}\sum_{i=2}^{N} x_i'(x_i - x_{i-1}) + \frac{1}{N}\sum_{i=2}^{N} x_{i-1}'(x_{i-1} - x_i) \to^P 2Q(F_T)$ and for the second term we use a central limit theorem of m-dependent process, so that $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(x_i - x_{i-1})'(e_i - e_{i-1}) \to N(0, E(F_T))$ as N tends to infinity. The asymptotic covariance matrix is

$$E(F_T) = Var\left[(x_i - x_{i-1})'(e_i - e_{i-1}) \mid F_T\right] + E\left[(x_2 - x_1)'(e_2 - e_1)(e_3 - e_2)'(x_3 - x_2) \mid F_T\right]$$

$$+E\left[(x_3 - x_2)'(e_3 - e_2)(e_2 - e_1)'(x_2 - x_1) \mid F_T\right]$$

$$= 2E\left[(x_i - x_{i-1})'(F_T\Sigma_\gamma F_T' + \Sigma_\varepsilon)(x_i - x_{i-1}) \mid F_T\right] - E\left[(x_2 - x_1)'(F_T\Sigma_\gamma F_T' + \Sigma_\varepsilon)(x_3 - x_2) \mid F_T\right]$$

$$-E\left[(x_3 - x_2)'(F_T\Sigma_\gamma F_T' + \Sigma_\varepsilon)(x_2 - x_1) \mid F_T\right].$$

The three terms can be evaluated to give

$$E\left[(x_i - x_{i-1})'(F_T\Sigma_\gamma F_T' + \Sigma_\varepsilon)(x_i - x_{i-1}) \mid F_T\right] = 2\left(A(F_T) + B(F_T)\right)$$

$$E\left[(x_2 - x_1)'(F_T\Sigma_\gamma F_T' + \Sigma_\varepsilon)(x_3 - x_2) \mid F_T\right] = -\left(A(F_T) + B(F_T)\right),$$

and the result of the theorem follows.

Proof of Theorem 2

We start from the MSM estimator

$$\hat{\beta}_{MSM} = \left[\sum_{i=1}^{N} x_i'(x_i - \overline{x})\right]^{-1} \sum_{i=1}^{N} x_i'\left((x_i - \overline{x})\beta_0 + e_i - \overline{e} + x_i\eta_i - \frac{1}{N}\sum_{j=1}^{N} x_j\eta_j\right)$$

$$= \beta_0 + \left[\sum_{i=1}^{N} x_i'(x_i - \overline{x})\right]^{-1}\left(\sum_{i=1}^{N} x_i'(e_i - \overline{e}) + \sum_{i=1}^{N}(x_i - \overline{x})' x_i\eta_i\right).$$

Unbiasedness follows from the independence of x_i and e_i and x_i and η_i given F_T,

$$E\left[\hat{\beta}_{MSM} \mid F_T\right] = \beta_0 + \sum_{i=1}^{N} E\left[\left[\sum_{j=1}^{N} x_j'\left(x_j - \bar{x}\right)\right]^{-1} x_i' \mid F_T\right] E\left[e_i - \bar{e} \mid F_T\right]$$

$$+\sum_{i=1}^{N} E\left[\left[\sum_{j=1}^{N} x_j'\left(x_j - \bar{x}\right)\right]^{-1} \left(x_i - \bar{x}\right)' x_i \mid F_T\right] E\left[\eta_i \mid F_T\right]$$

$$= \beta_0.$$

To show consistency we need the following results

$$\frac{1}{N}\sum_{i=1}^{N} x_i'\left(e_i - \bar{e}\right) = \frac{1}{N}\sum_{i=1}^{N} x_i' e_i - \bar{x} \cdot \bar{e}$$

$$= \frac{1}{N}\sum_{i=1}^{N} x_i' F_T \gamma_i + \frac{1}{N}\sum_{i=1}^{N} x_i' \varepsilon_i - \bar{x} \cdot \bar{e}$$

$$\to^{P} \Gamma' F_T' F_T \gamma + 0 - \Gamma' F_T' F_T \gamma = 0$$

which follows from Lemma A.1 c), d), h), b) and a), and

$$\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)' x_i \eta_i = \frac{1}{N}\sum_{i=1}^{N} x_i' x_i \eta_i - \bar{x}'\left(\frac{1}{N}\sum_{i=1}^{N} x_i \eta_i\right) \to^{P} 0 .$$

This follows of Lemma A.1 i) and the fact that

$$\frac{1}{N}\sum_{i=1}^{N} x_i \eta_i = \frac{1}{N}\sum_{i=1}^{N} F_T \Gamma_i \eta_i + \frac{1}{N}\sum_{i=1}^{N} v_i \eta_i$$

where i.i.d. terms with zero means are summed. Consistency follows from the two results above and f) in Lemma A.1,

$$\hat{\beta}_{MSM} \to^{P} \beta_0 + \left[\sum_{t=1}^{T}\left(f_t' \otimes I_p\right)\Sigma_{\Gamma}\left(f_t \otimes I_p\right) + \Sigma_{v0}\right]^{-1}\left(0+0\right) = \beta_0 .$$

For asymptotic normality we write

$$\sqrt{N}\left(\hat{\beta}_{MSM} - \beta_0\right) = \left[\frac{1}{N}\sum_{i=1}^{N} x_i'\left(x_i - \bar{x}\right)\right]^{-1} \frac{1}{\sqrt{N}}\left(\sum_{i=1}^{N} x_i'\left(e_i - \bar{e}\right) + \sum_{i=1}^{N}\left(x_i - \bar{x}\right)' x_i \eta_i\right).$$

Notice that

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\begin{pmatrix} x_i'(e_i-\bar{e}) \\ (x_i-\bar{x})'x_i\eta_i \end{pmatrix} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\begin{pmatrix} (x_i-F_T\Gamma+F_T\Gamma)'(e_i-\bar{e}) \\ (x_i-F_T\Gamma-(\bar{x}-F_T\Gamma))'x_i\eta_i \end{pmatrix}$$

$$= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\begin{pmatrix} (x_i-F_T\Gamma)'(e_i-\bar{e}) \\ (x_i-F_T\Gamma)'x_i\eta_i \end{pmatrix} - \begin{pmatrix} 0 \\ (\bar{x}-F_T\Gamma)'\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(x_i\eta_i) \end{pmatrix}$$

$$= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\begin{pmatrix} (x_i-F_T\Gamma)'(e_i-F_T\gamma) \\ (x_i-F_T\Gamma)'x_i\eta_i \end{pmatrix} + o_p(1).$$

This follows from the fact that conditional on F_T the $x_i\eta_i$ forms an sequence of i.i.d. random variables with mean zero ($E[x_i\eta_i \mid F_T]=0$) and covariance matrix

$$E[x_i\eta_i\eta_i'x_i' \mid F_T] = E\left[x_iE[\eta_i\eta_i' \mid F_T]x_i' \mid F_T\right] = E\left[x_i\Sigma_\eta x_i' \mid F_T\right] = E\left[F_T\Gamma_i\Sigma_\eta\Gamma_i'F_T' + v_i\Sigma_\eta v_i' \mid F_T\right]$$

$$= F_TE\left[\Gamma_i\Sigma_\eta\Gamma_i'\right]F_T' + E\left[v_i\Sigma_\eta v_i'\right],$$

we have

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}x_i\eta_i \mid F_T \to^D N\left(0, F_TE\left[\Gamma_i\Sigma_\eta\Gamma_i'\right]F_T' + E\left[v_i\Sigma_\eta v_i'\right]\right) .$$

Since $E\left[(x_i-F_T\Gamma)'(e_i-F_T\gamma)\eta_i'x_i'(x_i-F_T\Gamma) \mid F_T\right]=0$,

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\begin{pmatrix} (x_i-F_T\Gamma)'(e_i-F_T\gamma) \\ (x_i-F_T\Gamma)'x_i\eta_i \end{pmatrix} \mid F_T \to^D N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} A(F_T)+B(F_T) & 0 \\ 0 & C(F_T) \end{pmatrix}\right)$$

so that

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(x_i-F_T\Gamma)'(e_i-F_T\gamma) + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(x_i-F_T\Gamma)'x_i\eta_i \mid F_T \to^D N\left(0, A(F_T)+B(F_T)+C(F_T)\right),$$

and the part of the theorem concerning the MSM estimator is proved.

For the DM model we have

$$\hat{\beta}_{DM} = \beta_0 + \left[\sum_{i=2}^{N}x_i'(x_i-x_{i-1})\right]^{-1}\left(\sum_{i=2}^{N}x_i'(e_i-e_{i-1}) + \sum_{i=2}^{N}x_i'(x_i\eta_i - x_{i-1}\eta_{i-1})\right).$$

Unbiasedness follows as before. Consistency follows from results in Theorem 1 and Lemma A.1 i) and the fact which can be proved similarly to Lemma A.1 i) $\frac{1}{N}\sum_{i=2}^{N}x_i'x_i\eta_{i-1} \to^P 0$. To obtain asymptotic normality we notice that

$$\sqrt{N}\left(\hat{\beta}_{DM}-\beta_0\right)=\left[\frac{1}{N}\sum_{i=2}^{N}x_i'\left(x_i-x_{i-1}\right)\right]^{-1}\left(\frac{1}{\sqrt{N}}\sum_{i=2}^{N}x_i'\left(e_i-e_{i-1}\right)+\frac{1}{\sqrt{N}}\sum_{i=2}^{N}x_i'\left(x_i\eta_i-x_{i-1}\eta_{i-1}\right)\right).$$

Since $E\left[x_i'\left(e_i-e_{i-1}\right)\left(x_i\eta_i-x_{i-1}\eta_{i-1}\right)'x_i\mid F_T\right]=0$, using the central limit theorem for m-dependent processes with $m=1$, we obtain

$$\frac{1}{\sqrt{N}}\sum_{i=2}^{N}\begin{pmatrix}x_i'\left(e_i-e_{i-1}\right)\\ x_i'\left(x_i\eta_i-x_{i-1}\eta_{i-1}\right)\end{pmatrix}\to^{D}N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}2A\left(F_T\right)+2B\left(F_T\right) & 0\\ 0 & V_1\left(F_T\right)\end{pmatrix}\right).$$

where

$$V_1\left(F_T\right)=Var\left(x_i'\left(x_i\eta_i-x_{i-1}\eta_{i-1}\right)\mid F_T\right)$$

$$+Cov\left(x_2'\left(x_2\eta_2-x_1\eta_1\right),x_3'\left(x_3\eta_3-x_2\eta_2\right)\mid F_T\right)+Cov\left(x_3'\left(x_3\eta_3-x_2\eta_2\right),x_2'\left(x_2\eta_2-x_1\eta_1\right)\mid F_T\right)$$

$$=E\left[x_i'\left(x_i\eta_i-x_{i-1}\eta_{i-1}\right)\left(\eta_i'x_i'-\eta_{i-1}'x_{i-1}'\right)x_i\mid F_T\right]$$

$$+E\left[x_2'\left(x_2\eta_2-x_1\eta_1\right)\left(x_3\eta_3-x_2\eta_2\right)'x_3\mid F_T\right]+E\left[x_3'\left(x_3\eta_3-x_2\eta_2\right)\left(x_2\eta_2-x_1\eta_1\right)'x_2\mid F_T\right]$$

$$=E\left[x_i'x_i\eta_i\eta_i'x_i'x_i\mid F_T\right]-E\left[x_i'x_i\eta_i\eta_{i-1}'x_{i-1}'x_i\mid F_T\right]-E\left[x_i'x_{i-1}\eta_{i-1}\eta_i'x_i'x_i\mid F_T\right]$$

$$+E\left[x_i'x_{i-1}\eta_{i-1}\eta_{i-1}'x_{i-1}'x_i\mid F_T\right]-E\left[x_2'x_2\eta_2\eta_2'x_2'x_3\mid F_T\right]-E\left[x_3'x_2\eta_2\eta_2'x_2'x_2\mid F_T\right]$$

$$=E\left[x_i'x_i\Sigma_\eta x_i'x_i\mid F_T\right]+E\left[x_i'x_{i-1}\Sigma_\eta x_{i-1}'x_i\mid F_T\right]-E\left[x_2'x_2\Sigma_\eta x_2'x_3\mid F_T\right]-E\left[x_3'x_2\Sigma_\eta x_2'x_2\mid F_T\right]$$

$$=E\left[\left(x_i-F_T\Gamma\right)'x_i\Sigma_\eta x_i'\left(x_i-F_T\Gamma\right)\mid F_T\right]+E\left[x_i'x_{i-1}\Sigma_\eta x_{i-1}'x_i\mid F_T\right]-E\left[\Gamma'F_T'x_i\Sigma_\eta x_i'F_T\Gamma\mid F_T\right]$$

$$=E\left[\left(x_i-F_T\Gamma\right)'x_i\Sigma_\eta x_i'\left(x_i-F_T\Gamma\right)\mid F_T\right]+\left(E\left[x_i'x_{i-1}\Sigma_\eta x_{i-1}'x_i\mid F_T\right]-E\left[\Gamma'F_T'x_{i-1}\Sigma_\eta x_{i-1}'F_T\Gamma\mid F_T\right]\right)$$

$$=E\left[\left(x_i-F_T\Gamma\right)'x_i\Sigma_\eta x_i'\left(x_i-F_T\Gamma\right)\mid F_T\right]+E\left[\left(x_i-F_T\Gamma\right)'x_{i-1}\Sigma_\eta x_{i-1}'\left(x_i-F_T\Gamma\right)\mid F_T\right]$$

$$=C\left(F_T\right)+D\left(F_T\right).$$

The Theorem now follows from this.

Similarly for the FD estimator we have

$$\hat{\beta}_{FD} = \beta_0 + \left[\sum_{i=2}^{N}(x_i - x_{i-1})'(x_i - x_{i-1})\right]^{-1}\sum_{i=2}^{N}(x_i - x_{i-1})'(x_i\eta_i + e_i - x_{i-1}\eta_{i-1} - e_{i-1}).$$

Unbiasedness and consistency follow as in the previous cases. To obtain asymptotic normality we write

$$\sqrt{N}\left(\hat{\beta}_{FD} - \beta_0\right) = \left[\frac{1}{N}\sum_{i=2}^{N}(x_i - x_{i-1})'(x_i - x_{i-1})\right]^{-1}\left(\frac{1}{\sqrt{N}}\sum_{i=2}^{N}(x_i - x_{i-1})'(e_i - e_{i-1}) + \frac{1}{\sqrt{N}}\sum_{i=2}^{N}(x_i - x_{i-1})'(x_i\eta_i - x_{i-1}\eta_{i-1})\right)$$

Notice that $E\left[(x_i - x_{i-1})'(e_i - e_{i-1})(x_i\eta_i - x_{i-1}\eta_{i-1})'(x_i - x_{i-1})\right] = 0$ so that by the central limit theorem for m-dependent processes we have

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\begin{pmatrix}(x_i - x_{i-1})'(e_i - e_{i-1})\\(x_i - x_{i-1})'(x_i\eta_i - x_{i-1}\eta_{i-1})\end{pmatrix} \to^{D} N\left(0, \begin{pmatrix}E(F_T) & 0\\0 & V_2(F_T)\end{pmatrix}\right),$$

where $V_2(F_T)$ is now defined as

$$V_2(F_T) = E\left[(x_i - x_{i-1})'(x_i\eta_i - x_{i-1}\eta_{i-1})(x_i\eta_i - x_{i-1}\eta_{i-1})'(x_i - x_{i-1})\right]$$

$$+2Cov\left((x_2 - x_1)'(x_2\eta_2 - x_1\eta_1), (x_3 - x_2)'(x_3\eta_3 - x_2\eta_2)\right)$$

$$= 2E\left[(x_i - F_T\Gamma)'x_i\Sigma_\eta x_i'(x_i - F_T\Gamma)\right] + 2E\left[(x_{i-1} - F_T\Gamma)'x_i\Sigma_\eta x_i'(x_{i-1} - F_T\Gamma)\right]$$

$$+2E\left[(x_i - F_T\Gamma)'x_i\Sigma_\eta x_i'(x_i - F_T\Gamma)\right]$$

$$= 4E\left[(x_i - F_T\Gamma)'x_i\Sigma_\eta x_i'(x_i - F_T\Gamma)\right] + 2E\left[(x_{i-1} - F_T\Gamma)'x_i\Sigma_\eta x_i'(x_{i-1} - F_T\Gamma)\right]$$

$$= 4C(F_T)$$

$$+2E\left[\left(F_T(\Gamma_{i-1} - \Gamma) + v_{i-1}\right)'(F_T\Gamma_i + v_i)\Sigma_\eta(F_T\Gamma_i + v_i)'\left(F_T(\Gamma_{i-1} - \Gamma) + v_{i-1}\right)\right]$$

$$= 4C(F_T) + 2D(F_T),$$

where last line follows from the independence of the Γ's and v's for individual i and $i-1$. The result follows.

Proof of Theorem 3.

The result follows from Lemma A.2 by writing

$$\widehat{\Omega(F_T)} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})'\left((x_i - \overline{x})\left(\beta_0 - \hat{\beta}_{MSM}\right) + (e_i - \overline{e})\right)\left((x_i - \overline{x})\left(\beta_0 - \hat{\beta}_{MSM}\right) + (e_i - \overline{e})\right)'(x_i - \overline{x})$$

$$+\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})'\left((x_i - \overline{x})\left(\beta_0 - \hat{\beta}_{MSM}\right) + (e_i - \overline{e})\right)\left((x_i - \overline{x})\eta_i\right)'(x_i - \overline{x})$$

$$+\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})'\left((x_i - \overline{x})\eta_i\right)\left((x_i - \overline{x})\left(\beta_0 - \hat{\beta}_{MSM}\right) + (e_i - \overline{e})\right)'(x_i - \overline{x})$$

$$+\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})'\left((x_i - \overline{x})\eta_i\right)\left((x_i - \overline{x})\eta_i\right)'(x_i - \overline{x}).$$

Table 1

Small Sample Properties of the MSM, FD, FE, CCEP, CCEMG and OLS estimators

Homogeneous slopes mildly correlated factors

		Bias					MSE					Size (5% level, $H_0 : \beta_1 = 1$)					Power (5% level, $H_1 : \beta_1 \neq 1$)				
β_1	N\T	T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20
MSM	N=20	-0.001	-0.001	-0.001	-0.001	0.000	0.016	0.011	0.008	0.006	0.004	0.115	0.098	0.105	0.098	0.091	0.062	0.071	0.084	0.117	0.144
	50	0.001	-0.001	0.000	0.000	-0.001	0.006	0.004	0.003	0.002	0.001	0.072	0.076	0.072	0.065	0.066	0.100	0.112	0.147	0.203	0.258
	100	0.000	-0.001	0.000	0.001	0.000	0.003	0.002	0.002	0.001	0.001	0.062	0.062	0.059	0.058	0.059	0.154	0.191	0.257	0.368	0.454
	200	-0.001	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.000	0.055	0.056	0.056	0.054	0.056	0.225	0.354	0.446	0.599	0.727
DM	N=20	-0.001	-0.003	0.001	-0.001	0.000	0.035	0.025	0.018	0.012	0.008										
	50	0.002	-0.003	0.000	0.000	-0.001	0.012	0.009	0.006	0.004	0.003										
	100	-0.001	0.000	0.000	0.001	0.000	0.006	0.004	0.003	0.002	0.001										
	200	-0.001	0.001	0.001	-0.001	0.000	0.003	0.002	0.002	0.001	0.001										
FD	N=20	-0.002	-0.002	-0.001	-0.001	0.000	0.024	0.017	0.012	0.008	0.006										
	50	0.001	-0.002	0.000	0.000	-0.001	0.009	0.006	0.005	0.003	0.002										
	100	0.000	0.000	0.000	0.001	0.000	0.004	0.003	0.002	0.002	0.001										
	200	-0.001	0.000	0.001	0.000	0.000	0.002	0.002	0.001	0.001	0.001										
CCEP	N=20	NA	4.859	0.002	0.000	0.000	NA	236869.105	0.017	0.006	0.003	NA	0.119	0.008	0.067	0.067	NA	0.020	0.007	0.085	0.155
	50	NA	0.177	0.000	-0.001	0.000	NA	190.123	0.006	0.002	0.001	NA	0.193	0.006	0.059	0.064	NA	0.032	0.009	0.165	0.312
	100	NA	0.308	0.000	0.000	0.000	NA	1129.204	0.003	0.001	0.001	NA	0.266	0.005	0.050	0.056	NA	0.039	0.013	0.310	0.560
	200	NA	0.437	0.000	0.000	0.000	NA	29642.509	0.001	0.001	0.000	NA	0.349	0.004	0.055	0.054	NA	0.047	0.023	0.556	0.836
CCEMG	N=20	NA	-1.279	-0.326	0.001	0.001	NA	15485.661	4044.353	0.010	0.004	NA	0.069	0.032	0.063	0.069	NA	0.013	0.006	0.058	0.124
	50	NA	0.456	-0.533	0.000	0.000	NA	1607.718	7557.054	0.004	0.001	NA	0.070	0.024	0.055	0.058	NA	0.011	0.005	0.102	0.256
	100	NA	0.282	0.558	0.000	0.000	NA	95057.996	887.079	0.002	0.001	NA	0.070	0.022	0.048	0.055	NA	0.015	0.004	0.188	0.466
	200	NA	0.588	-0.827	0.000	0.000	NA	9696.891	3360.061	0.001	0.000	NA	0.083	0.021	0.052	0.052	NA	0.017	0.004	0.335	0.766
OLS	N=20	0.114	0.120	0.125	0.131	0.135	0.061	0.052	0.047	0.042	0.039										
	50	0.115	0.118	0.126	0.133	0.136	0.047	0.043	0.040	0.038	0.036										
	100	0.108	0.120	0.124	0.131	0.134	0.042	0.041	0.038	0.035	0.035										
	200	0.109	0.117	0.126	0.130	0.135	0.041	0.038	0.037	0.035	0.034										

Notes: $\sigma_\eta^2 = 0$, $\rho_{fj} = 0.1$

Table 2

Small Sample Properties of the MSM, FD, FE, CCEP, CCEMG and OLS estimators

Homogeneous slopes moderately correlated in factors

β_1		Bias					MSE					Size (5% level, $\beta_1 = 1$)					Power (5% level, $\beta_1 = 0.95$)				
		T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20
MSM	N=20	-0.001	-0.002	-0.001	-0.001	0.000	0.018	0.013	0.010	0.007	0.005	0.119	0.108	0.107	0.103	0.095	0.062	0.068	0.077	0.107	0.128
	50	0.001	-0.001	0.000	0.000	-0.001	0.006	0.005	0.004	0.003	0.002	0.074	0.074	0.076	0.068	0.069	0.092	0.107	0.131	0.173	0.214
	100	0.000	-0.001	0.000	0.001	0.000	0.003	0.002	0.002	0.001	0.001	0.065	0.063	0.060	0.062	0.055	0.147	0.173	0.227	0.309	0.382
	200	-0.001	0.000	0.000	0.000	0.000	0.002	0.001	0.001	0.001	0.000	0.057	0.060	0.057	0.055	0.055	0.246	0.309	0.387	0.508	0.631
DM	N=20	-0.002	-0.004	0.001	0.000	0.000	0.040	0.029	0.022	0.015	0.011										
	50	0.002	-0.002	-0.001	0.001	-0.001	0.013	0.010	0.008	0.005	0.004										
	100	0.000	0.000	0.000	0.001	0.000	0.006	0.005	0.004	0.003	0.002										
	200	-0.001	0.001	0.001	-0.001	0.000	0.003	0.002	0.002	0.001	0.001										
FD	N=20	-0.001	-0.002	-0.001	-0.001	-0.001	0.026	0.020	0.015	0.010	0.007										
	50	0.001	-0.002	-0.001	0.000	-0.001	0.010	0.008	0.006	0.004	0.003										
	100	0.000	0.000	0.000	0.001	0.000	0.005	0.004	0.003	0.002	0.001										
	200	-0.001	0.000	0.000	-0.001	0.000	0.002	0.002	0.001	0.001	0.001										
CCEP	N=20	NA	0.209	0.001	0.000	0.000	NA	109.246	0.019	0.007	0.004	NA	0.123	0.011	0.066	0.069	NA	0.019	0.006	0.084	0.143
	50	0.130	-0.587	0.000	-0.001	-0.001	34.862	4473.610	0.007	0.002	0.001	0.140	0.192	0.006	0.059	0.062	0.026	0.032	0.011	0.159	0.299
	100	-0.068	-0.336	0.000	0.000	0.000	234.917	792.777	0.003	0.001	0.001	0.199	0.275	0.005	0.050	0.057	0.035	0.042	0.013	0.299	0.536
	200	NA	-0.205	0.000	0.000	0.000	NA	1540.047	0.002	0.001	0.000	NA	0.368	0.005	0.056	0.051	NA	0.048	0.025	0.538	0.824
CCEMG	N=20	NA	4.059	0.165	0.001	0.001	NA	73864.215	1593.559	0.011	0.004	NA	0.072	0.032	0.064	0.068	NA	0.011	0.005	0.059	0.122
	50	NA	0.228	-3.196	-0.001	-0.001	NA	2559.604	63766.362	0.004	0.001	NA	0.070	0.025	0.054	0.057	NA	0.013	0.004	0.105	0.248
	100	NA	-4.114	3.190	0.000	0.000	NA	200537.028	90693.652	0.002	0.001	NA	0.077	0.022	0.053	0.053	NA	0.017	0.004	0.180	0.458
	200	NA	-0.828	0.678	0.000	0.000	NA	11510.035	5708.849	0.001	0.000	NA	0.088	0.021	0.052	0.052	NA	0.014	0.003	0.335	0.757
OLS	N=20	0.119	0.125	0.130	0.139	0.143	0.069	0.061	0.055	0.050	0.046										
	50	0.120	0.123	0.132	0.140	0.144	0.054	0.050	0.047	0.045	0.043										
	100	0.114	0.126	0.130	0.138	0.143	0.049	0.047	0.045	0.042	0.041										
	200	0.115	0.122	0.131	0.138	0.143	0.047	0.044	0.043	0.042	0.040										

Notes: $\sigma_\eta^2 = 0$, $\rho_{fj} = 0.5$

Table 3

Small Sample Properties of the MSM, FD, FE, CCEP, CCEMG and OLS estimators

Heterogeneous slopes mildly correlated in factors

β_1		Bias					MSE					Size (5% level, $\beta_1 = 1$)					Power (5% level, $\beta_1 = 0.95$)				
		T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20
MSM	N=20	0.000	0.004	0.003	-0.003	-0.001	0.088	0.075	0.061	0.050	0.041	0.142	0.134	0.129	0.118	0.105	0.026	0.028	0.030	0.031	0.030
	50	0.001	-0.001	0.001	0.000	-0.001	0.036	0.030	0.026	0.020	0.017	0.085	0.090	0.088	0.081	0.078	0.033	0.036	0.040	0.043	0.043
	100	0.000	-0.001	0.001	0.000	0.000	0.018	0.015	0.013	0.010	0.009	0.072	0.069	0.064	0.062	0.062	0.043	0.045	0.053	0.056	0.060
	200	-0.001	0.002	0.001	0.002	-0.001	0.009	0.008	0.006	0.005	0.004	0.061	0.061	0.059	0.054	0.055	0.060	0.069	0.078	0.092	0.098
DM	N=20	0.005	0.005	0.000	-0.002	-0.001	0.154	0.126	0.097	0.075	0.060										
	50	0.002	-0.001	0.000	-0.001	-0.001	0.058	0.046	0.038	0.029	0.024										
	100	0.000	-0.003	0.001	0.001	-0.001	0.028	0.023	0.018	0.014	0.012										
	200	-0.002	0.002	0.001	0.001	-0.001	0.014	0.012	0.009	0.007	0.006										
FD	N=20	0.002	0.006	0.000	-0.003	-0.001	0.111	0.094	0.075	0.060	0.048										
	50	0.001	-0.001	0.001	0.000	0.000	0.047	0.037	0.032	0.024	0.020										
	100	0.000	-0.002	0.001	0.001	0.000	0.023	0.019	0.015	0.012	0.010										
	200	-0.002	0.002	0.002	0.002	-0.001	0.012	0.010	0.008	0.006	0.005										
CCEP	N=20	NA	0.181	0.002	-0.003	0.000	NA	294.643	0.077	0.049	0.039	NA	0.100	0.029	0.083	0.084	NA	0.013	0.007	0.027	0.031
	50	NA	1.329	-0.001	0.000	0.001	NA	18462.732	0.034	0.022	0.018	NA	0.141	0.018	0.063	0.061	NA	0.021	0.008	0.037	0.044
	100	NA	0.805	0.002	0.001	0.000	NA	4217.675	0.018	0.012	0.010	NA	0.192	0.015	0.054	0.057	NA	0.031	0.009	0.057	0.057
	200	NA	0.341	0.001	0.001	0.000	NA	99.225	0.009	0.007	0.005	NA	0.277	0.015	0.051	0.052	NA	0.038	0.015	0.083	0.092
CCEMG	N=20	NA	-0.282	0.288	-0.001	0.001	NA	2582.768	3453.806	0.031	0.024	NA	0.060	0.035	0.073	0.074	NA	0.011	0.004	0.027	0.033
	50	NA	1.062	0.004	0.000	0.000	NA	9866.830	415.336	0.012	0.009	NA	0.067	0.026	0.059	0.058	NA	0.013	0.004	0.046	0.052
	100	NA	-1.826	-0.837	0.002	0.000	NA	40801.656	2902.374	0.006	0.005	NA	0.072	0.026	0.052	0.055	NA	0.013	0.004	0.075	0.086
	200	NA	1.047	-0.137	0.000	0.001	NA	10802.576	2443.533	0.003	0.002	NA	0.077	0.023	0.054	0.051	NA	0.017	0.004	0.122	0.155
OLS	N=20	0.108	0.126	0.129	0.131	0.136	0.115	0.106	0.091	0.080	0.073										
	50	0.113	0.120	0.124	0.133	0.134	0.071	0.065	0.060	0.054	0.049										
	100	0.109	0.115	0.125	0.132	0.137	0.054	0.050	0.047	0.044	0.042										
	200	0.108	0.116	0.127	0.132	0.136	0.046	0.044	0.042	0.039	0.038										

Notes: $\sigma_\eta^2 = 0.4$, $\rho_{fj} = 0.1$

Table 4

Small Sample Properties of the MSM, FD, FE, CCEP, CCEMG and OLS estimators

Heterogeneous slopes moderately correlated in factors

β_1		Bias					MSE					Size (5% level, $\beta_1 = 1$)					Power (5% level, $\beta_1 = 0.95$)				
		T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20
MSM	N=20	0.000	0.005	0.003	-0.002	0.000	0.095	0.083	0.069	0.056	0.046	0.138	0.140	0.130	0.122	0.110	0.029	0.028	0.029	0.032	0.028
	50	0.001	-0.001	0.001	0.000	-0.001	0.039	0.033	0.029	0.022	0.019	0.087	0.094	0.089	0.080	0.075	0.031	0.035	0.034	0.039	0.042
	100	0.000	-0.002	0.001	0.000	0.000	0.019	0.017	0.014	0.011	0.009	0.072	0.069	0.065	0.060	0.063	0.037	0.040	0.050	0.052	0.056
	200	-0.001	0.002	0.001	0.002	-0.001	0.010	0.008	0.007	0.006	0.005	0.060	0.062	0.058	0.055	0.056	0.060	0.066	0.073	0.083	0.089
DM	N=20	0.006	0.006	0.001	-0.001	-0.001	0.177	0.144	0.115	0.089	0.071										
	50	0.002	-0.001	0.001	-0.002	-0.002	0.064	0.052	0.045	0.033	0.027										
	100	0.001	-0.004	0.001	0.001	-0.001	0.031	0.026	0.021	0.016	0.013										
	200	-0.002	0.003	0.002	0.001	-0.001	0.016	0.013	0.011	0.008	0.007										
FD	N=20	0.002	0.006	0.000	-0.001	0.000	0.121	0.105	0.085	0.069	0.055										
	50	0.001	-0.001	0.001	-0.001	0.000	0.051	0.042	0.036	0.027	0.023										
	100	0.001	-0.002	0.000	0.000	0.000	0.025	0.021	0.017	0.014	0.011										
	200	-0.001	0.002	0.002	0.002	-0.001	0.013	0.011	0.009	0.007	0.006										
CCEP	N=20	0.082	-1.183	0.003	-0.004	-0.001	194.238	12114.255	0.076	0.048	0.039	0.072	0.105	0.025	0.080	0.084	0.009	0.012	0.006	0.027	0.030
	50	NA	3.039	-0.002	-0.001	0.001	NA	78470.707	0.033	0.021	0.017	NA	0.147	0.020	0.062	0.062	NA	0.022	0.007	0.036	0.044
	100	NA	0.238	0.002	0.001	0.001	NA	143.562	0.017	0.011	0.009	NA	0.208	0.014	0.056	0.056	NA	0.033	0.012	0.056	0.059
	200	0.129	0.216	0.001	0.001	0.000	72.405	114.163	0.009	0.006	0.005	0.215	0.285	0.014	0.051	0.052	0.036	0.042	0.018	0.085	0.094
CCEMG	N=20	NA	17.086	-0.544	-0.001	0.001	NA	3671775.446	1641.101	0.031	0.024	NA	0.066	0.034	0.074	0.073	NA	0.011	0.005	0.026	0.034
	50	NA	1.889	0.545	0.000	0.000	NA	50227.456	2279.704	0.012	0.009	NA	0.068	0.025	0.057	0.058	NA	0.013	0.004	0.046	0.053
	100	NA	-0.525	0.974	0.002	0.000	NA	3701.143	15963.319	0.006	0.005	NA	0.075	0.023	0.053	0.055	NA	0.016	0.004	0.074	0.087
	200	NA	-1.171	-4.362	0.000	0.001	NA	81669.960	237091.094	0.003	0.002	NA	0.084	0.022	0.051	0.051	NA	0.016	0.004	0.125	0.157
OLS	N=20	0.114	0.131	0.132	0.138	0.145	0.127	0.119	0.103	0.091	0.082										
	50	0.119	0.123	0.129	0.139	0.143	0.079	0.073	0.068	0.061	0.056										
	100	0.114	0.121	0.129	0.140	0.145	0.061	0.057	0.054	0.051	0.049										
	200	0.113	0.121	0.132	0.139	0.144	0.053	0.051	0.048	0.046	0.044										

Notes: $\sigma_{\eta}^{2} = 0.4$, $\rho_{fj} = 0.5$

Table 5

Small Sample Properties of the MSM, FD, FE, CCEP, CCEMG and OLS estimators

Homogeneous slopes with fixed effects and mildly correlated in factors

β_1		Bias					MSE					Size (5% level, $\beta_1 = 1$)					Power (5% level, $\beta_1 = 0.95$)				
		T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20
MSM	N=20	0.000	0.000	-0.001	0.000	0.000	0.021	0.016	0.011	0.008	0.005	0.110	0.108	0.101	0.104	0.097	0.050	0.059	0.071	0.090	0.108
	50	0.000	0.000	0.000	0.000	0.000	0.008	0.006	0.004	0.003	0.002	0.077	0.073	0.072	0.068	0.069	0.076	0.088	0.113	0.147	0.187
	100	0.000	0.000	0.000	0.000	0.000	0.004	0.003	0.002	0.002	0.001	0.063	0.062	0.059	0.063	0.058	0.119	0.146	0.182	0.256	0.338
	200	0.000	0.000	0.000	0.000	0.000	0.002	0.002	0.001	0.001	0.001	0.053	0.061	0.057	0.060	0.057	0.209	0.250	0.316	0.433	0.564
DM	N=20	0.003	-0.001	0.001	0.001	0.000	0.046	0.034	0.024	0.016	0.012										
	50	0.000	0.001	0.000	0.000	-0.002	0.016	0.012	0.009	0.006	0.004										
	100	0.001	0.000	0.000	0.000	0.000	0.008	0.006	0.004	0.003	0.002										
	200	0.000	0.000	0.000	0.000	0.000	0.004	0.003	0.002	0.001	0.001										
FD	N=20	0.001	0.000	0.000	0.001	0.000	0.031	0.024	0.017	0.012	0.008										
	50	0.000	0.000	0.000	0.000	-0.001	0.012	0.009	0.006	0.004	0.003										
	100	0.000	0.000	0.000	0.000	0.000	0.006	0.004	0.003	0.002	0.002										
	200	0.000	0.000	0.000	0.000	0.000	0.003	0.002	0.002	0.001	0.001										
CCEP	N=20	0.347	0.076	2.748	0.001	0.000	439.403	780.924	42906.914	0.009	0.004	0.091	0.112	0.148	0.058	0.072	0.012	0.013	0.020	0.057	0.130
	50	-0.022	0.965	0.665	0.000	0.000	599.157	5552.816	549.577	0.003	0.001	0.141	0.168	0.223	0.051	0.059	0.022	0.021	0.028	0.110	0.262
	100	-0.985	0.465	-1.293	0.000	0.000	5560.135	620.242	14332.405	0.002	0.001	0.197	0.235	0.314	0.053	0.048	0.026	0.033	0.042	0.207	0.482
	200	NA	-0.318	0.608	0.001	0.000	NA	1602.590	451.980	0.001	0.000	NA	0.308	0.414	0.048	0.051	NA	0.038	0.047	0.393	0.776
CCEMG	N=20	NA	0.855	-0.417	0.002	0.001	NA	9750.956	5776.420	0.020	0.004	NA	0.058	0.077	0.063	0.069	NA	0.011	0.014	0.044	0.114
	50	NA	0.329	-24.743	0.001	0.000	NA	5647.516	6591885.939	0.008	0.002	NA	0.055	0.075	0.053	0.055	NA	0.011	0.013	0.070	0.224
	100	NA	-0.282	0.454	-0.001	0.000	NA	1285.805	10351.837	0.004	0.001	NA	0.061	0.073	0.051	0.053	NA	0.012	0.015	0.113	0.423
	200	NA	-0.871	-0.502	0.001	0.000	NA	17680.130	17104.326	0.002	0.000	NA	0.069	0.079	0.053	0.050	NA	0.012	0.016	0.188	0.704
OLS	N=20	0.140	0.143	0.153	0.160	0.165	0.076	0.065	0.056	0.050	0.046										
	50	0.135	0.148	0.151	0.161	0.163	0.058	0.054	0.048	0.044	0.041										
	100	0.136	0.147	0.152	0.162	0.163	0.053	0.051	0.045	0.043	0.040										
	200	0.139	0.144	0.153	0.159	0.163	0.052	0.048	0.044	0.041	0.039										

Notes: $\sigma_\eta^2 = 0, \rho_{fj} = 0.1$

Table 6

Small Sample Properties of the MSM, FD, FE, CCEP, CCEMG and OLS estimators

Homogeneous slopes with fixed effects and moderately correlated in factors

β_1		Bias					MSE					Size (5% level, $\beta_1 = 1$)					Power (5% level, $\beta_1 = 0.95$)				
		T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20
MSM	N=20	0.001	0.000	-0.001	0.000	0.000	0.023	0.019	0.013	0.009	0.006	0.117	0.115	0.107	0.107	0.097	0.047	0.056	0.065	0.086	0.097
	50	0.000	0.000	0.001	0.001	0.000	0.008	0.007	0.005	0.003	0.002	0.077	0.074	0.075	0.065	0.067	0.076	0.083	0.104	0.138	0.165
	100	0.000	0.000	0.000	0.000	0.000	0.004	0.003	0.002	0.002	0.001	0.063	0.062	0.061	0.065	0.058	0.114	0.133	0.164	0.227	0.302
	200	0.000	0.000	0.000	0.000	0.000	0.002	0.002	0.001	0.001	0.001	0.053	0.060	0.058	0.060	0.056	0.194	0.236	0.282	0.389	0.513
DM	N=20	0.004	-0.001	0.001	0.001	0.001	0.050	0.039	0.028	0.019	0.013										
	50	0.001	0.001	0.000	0.001	-0.001	0.017	0.014	0.010	0.007	0.005										
	100	0.001	0.000	0.000	0.000	0.001	0.009	0.007	0.005	0.004	0.002										
	200	0.000	0.000	0.000	0.000	0.000	0.004	0.003	0.003	0.002	0.001										
FD	N=20	0.002	0.000	0.000	0.000	0.001	0.034	0.027	0.020	0.014	0.009										
	50	0.000	0.001	0.000	0.000	-0.001	0.013	0.010	0.007	0.005	0.004										
	100	0.000	0.000	0.000	0.000	0.000	0.006	0.005	0.004	0.003	0.002										
	200	0.000	0.000	0.000	0.000	0.000	0.003	0.002	0.002	0.001	0.001										
CCEP	N=20	NA	2.064	0.080	0.001	0.000	NA	28485.226	573.913	0.009	0.004	NA	0.116	0.144	0.057	0.072	NA	0.013	0.020	0.059	0.126
	50	NA	0.047	0.427	0.000	0.000	NA	1310.049	547.586	0.004	0.001	NA	0.173	0.222	0.055	0.058	NA	0.026	0.031	0.114	0.256
	100	-0.022	0.232	0.314	0.000	0.000	219.879	222.976	110.733	0.002	0.001	0.203	0.234	0.311	0.055	0.049	0.025	0.031	0.037	0.199	0.475
	200	NA	-0.190	0.155	0.001	0.000	NA	7503.862	540.959	0.001	0.000	NA	0.314	0.411	0.048	0.050	NA	0.036	0.045	0.382	0.763
CCEMG	N=20	NA	0.304	0.211	0.002	0.001	NA	1891.045	627.141	0.020	0.004	NA	0.059	0.080	0.061	0.066	NA	0.010	0.014	0.042	0.115
	50	NA	-1.163	0.417	0.001	0.000	NA	20923.011	754.038	0.008	0.002	NA	0.058	0.073	0.052	0.053	NA	0.010	0.016	0.066	0.217
	100	NA	7.930	1.016	0.000	0.000	NA	537593.523	1639.355	0.004	0.001	NA	0.057	0.079	0.052	0.052	NA	0.010	0.013	0.108	0.415
	200	NA	-7.872	-0.586	0.001	0.000	NA	234267.507	1906.952	0.002	0.000	NA	0.059	0.073	0.049	0.047	NA	0.014	0.013	0.185	0.699
OLS	N=20	0.136	0.142	0.150	0.159	0.166	0.083	0.074	0.064	0.057	0.052										
	50	0.133	0.145	0.148	0.160	0.164	0.064	0.061	0.054	0.050	0.046										
	100	0.134	0.143	0.148	0.161	0.165	0.058	0.056	0.051	0.049	0.045										
	200	0.138	0.141	0.149	0.158	0.164	0.057	0.054	0.050	0.046	0.044										

Notes: $\sigma_{\eta}^{2} = 0$, $\rho_{fj} = 0.5$

Table 7

Small Sample Properties of the MSM, FD, FE, CCEP, CCEMG and OLS estimators

Heterogeneous slopes with fixed effects and mildly correlated in factors

		Bias					MSE					Size (5% level, $\beta_1 = 1$)					Power (5% level, $\beta_1 = 0.95$)				
β_1		T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20
MSM	N=20	-0.001	0.000	-0.001	0.001	-0.001	0.058	0.047	0.037	0.029	0.024	0.131	0.127	0.115	0.120	0.109	0.031	0.034	0.039	0.038	0.045
	50	0.002	-0.001	-0.002	0.001	0.001	0.023	0.019	0.015	0.012	0.010	0.083	0.079	0.080	0.078	0.072	0.043	0.044	0.045	0.057	0.066
	100	-0.002	0.000	0.000	-0.001	-0.001	0.011	0.009	0.008	0.006	0.005	0.071	0.063	0.067	0.063	0.064	0.056	0.066	0.070	0.085	0.091
	200	-0.002	0.001	0.000	-0.001	-0.001	0.006	0.005	0.004	0.003	0.002	0.062	0.058	0.056	0.061	0.057	0.080	0.098	0.113	0.130	0.153
DM	N=20	-0.003	0.000	0.000	0.001	-0.001	0.110	0.084	0.063	0.047	0.037										
	50	0.001	-0.001	-0.004	0.001	0.000	0.039	0.031	0.024	0.019	0.014										
	100	-0.002	0.000	-0.001	-0.001	-0.002	0.019	0.016	0.012	0.009	0.007										
	200	-0.001	0.001	0.000	-0.001	-0.001	0.009	0.008	0.006	0.005	0.004										
FD	N=20	-0.001	0.001	-0.001	0.001	-0.001	0.077	0.061	0.047	0.036	0.029										
	50	0.002	-0.002	-0.003	0.001	0.000	0.030	0.025	0.019	0.015	0.012										
	100	-0.003	0.001	-0.001	-0.001	-0.002	0.015	0.012	0.010	0.007	0.006										
	200	-0.002	0.000	0.000	-0.001	-0.001	0.007	0.006	0.005	0.004	0.003										
CCEP	N=20	NA	0.152	-0.047	0.000	0.000	NA	684.853	4494.555	0.028	0.019	NA	0.105	0.128	0.080	0.083	NA	0.011	0.019	0.033	0.041
	50	0.351	0.808	-0.466	0.000	0.001	603.594	2080.515	3087.051	0.011	0.007	0.135	0.148	0.200	0.061	0.056	0.016	0.023	0.028	0.053	0.069
	100	2.464	-0.045	0.239	0.000	0.000	55035.415	158.327	251.113	0.005	0.004	0.168	0.203	0.280	0.054	0.054	0.023	0.028	0.039	0.081	0.115
	200	-0.039	-0.026	0.259	-0.001	-0.001	246.483	485.279	32.793	0.003	0.002	0.236	0.285	0.376	0.050	0.059	0.036	0.033	0.041	0.135	0.183
CCEMG	N=20	NA	-0.184	1.338	0.001	0.000	NA	3906.317	38784.105	0.031	0.014	NA	0.054	0.072	0.071	0.073	NA	0.008	0.013	0.031	0.046
	50	NA	2.333	0.339	0.000	0.001	NA	20879.500	1070.369	0.012	0.006	NA	0.059	0.078	0.054	0.058	NA	0.011	0.013	0.045	0.076
	100	NA	0.154	-0.291	-0.001	0.000	NA	6762.931	2115.024	0.006	0.003	NA	0.060	0.074	0.057	0.054	NA	0.011	0.010	0.074	0.135
	200	NA	-7.110	-0.552	-0.001	-0.001	NA	408412.981	3063.274	0.003	0.001	NA	0.065	0.075	0.051	0.053	NA	0.014	0.015	0.125	0.241
OLS	N=20	0.137	0.147	0.154	0.161	0.164	0.103	0.090	0.078	0.068	0.062										
	50	0.141	0.146	0.152	0.160	0.164	0.072	0.064	0.056	0.051	0.048										
	100	0.137	0.148	0.154	0.157	0.161	0.061	0.056	0.051	0.044	0.042										
	200	0.136	0.149	0.151	0.158	0.161	0.055	0.053	0.046	0.043	0.040										

Notes: $\sigma_\eta^2 = 0.4$, $\rho_{fj} = 0.1$

Table 8

Small Sample Properties of the MSM, FD, FE, CCEP, CCEMG and OLS estimators

Hetrogeneous slopes with fixed effects and moderately correlated in factors

		Bias					MSE					Size (5% level, $\beta_1 = 1$)					Power (5% level, $\beta_1 = 0.95$)				
β_1		T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20	T=2	3	5	10	20
MSM	N=20	-0.001	0.000	-0.001	0.001	-0.001	0.063	0.052	0.041	0.032	0.027	0.134	0.131	0.118	0.122	0.113	0.032	0.031	0.037	0.038	0.042
	50	0.001	-0.001	-0.002	0.001	0.001	0.024	0.020	0.017	0.013	0.011	0.085	0.081	0.084	0.079	0.075	0.043	0.038	0.044	0.053	0.061
	100	-0.002	0.001	0.000	-0.001	-0.001	0.012	0.010	0.009	0.006	0.005	0.073	0.066	0.066	0.064	0.062	0.053	0.062	0.067	0.078	0.088
	200	-0.001	0.001	0.000	-0.001	-0.001	0.006	0.005	0.004	0.003	0.003	0.060	0.059	0.055	0.062	0.057	0.078	0.092	0.104	0.121	0.143
DM	N=20	-0.004	0.000	0.000	0.001	-0.002	0.120	0.098	0.074	0.054	0.042										
	50	0.000	-0.001	-0.004	0.001	0.001	0.041	0.035	0.028	0.021	0.016										
	100	-0.002	0.000	-0.001	-0.001	-0.002	0.021	0.018	0.014	0.010	0.008										
	200	-0.001	0.001	0.000	-0.001	-0.001	0.010	0.008	0.007	0.005	0.004										
FD	N=20	-0.001	0.001	-0.001	0.001	-0.001	0.084	0.068	0.053	0.041	0.033										
	50	0.001	-0.002	-0.003	0.001	0.000	0.032	0.027	0.022	0.017	0.013										
	100	-0.003	0.001	0.000	-0.001	-0.002	0.016	0.014	0.011	0.008	0.007										
	200	-0.002	0.000	0.000	-0.001	-0.001	0.008	0.007	0.005	0.004	0.003										
CCEP	N=20	-0.310	-18.934	-0.089	0.000	-0.001	1417.608	3600626.163	359.474	0.028	0.019	0.092	0.104	0.134	0.082	0.084	0.011	0.013	0.016	0.031	0.044
	50	1.378	0.040	0.220	0.000	0.001	14911.087	326.390	186.860	0.011	0.007	0.131	0.155	0.201	0.060	0.056	0.018	0.020	0.027	0.054	0.071
	100	0.200	-0.020	0.559	0.000	0.000	203.477	207.730	572.911	0.005	0.004	0.173	0.208	0.277	0.055	0.054	0.023	0.028	0.038	0.082	0.113
	200	0.201	1.041	0.095	-0.001	-0.001	597.865	6211.793	680.937	0.003	0.002	0.241	0.286	0.364	0.052	0.057	0.035	0.036	0.045	0.137	0.188
CCEMG	N=20	NA	-0.349	0.240	0.000	0.000	NA	1189.546	533.846	0.031	0.014	NA	0.057	0.074	0.069	0.073	NA	0.009	0.009	0.032	0.045
	50	NA	1.985	0.948	0.001	0.001	NA	53239.473	3309.992	0.012	0.006	NA	0.058	0.070	0.058	0.058	NA	0.011	0.014	0.044	0.074
	100	NA	0.624	-0.331	0.000	-0.001	NA	2769.728	1801.759	0.006	0.003	NA	0.058	0.076	0.054	0.053	NA	0.012	0.013	0.079	0.133
	200	NA	-0.085	-0.239	-0.001	-0.001	NA	28272.389	18550.552	0.003	0.001	NA	0.064	0.074	0.049	0.054	NA	0.011	0.014	0.122	0.241
OLS	N=20	0.135	0.143	0.151	0.159	0.165	0.112	0.099	0.087	0.075	0.069										
	50	0.139	0.143	0.149	0.158	0.165	0.078	0.071	0.063	0.057	0.053										
	100	0.134	0.146	0.151	0.156	0.161	0.065	0.063	0.057	0.050	0.048										
	200	0.134	0.146	0.148	0.157	0.162	0.060	0.058	0.052	0.049	0.045										

Notes: $\sigma_\eta^2 = 0.4$, $\rho_{fj} = 0.5$

References

Bai, J. (2009). "Panel Data Models with Interactive Fixed Effects." *Econometrica* 77(4): 1229-1279.

Bai, J., C. Kao and S. Ng (2009). "Panel Cointegration with Global Stochastic Trends." *Journal of Econometrics* 149: 82-99.

Bai, J. and S. Ng (2010). "Panel Unit Root Tests with Cross-Sectional Dependence: A Further Investigation." *Econometric Theory* 26: 1088-1114.

Case, A. C. (1991). "Spatial Patterns in Household Demand." *Econometrica* 59: 953-965.

Coakley, J., A.-M. Fuertes and R. Smith (2001). "A Principal Components Approach to Cross-Section Dependence in Panels." *Birkbeck College Discussion Paper, 01/2002*.

Conley, T. G. (1999). "Gmm Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92: 1-45.

DasGupta, A. (2008). *Asymptotic Theory of Statistics and Probability*. New York, Springer.

Hoyos, R. E. D. and V. Sarafidis (2006). "Testing for Cross-Sectional Dependence in Panel Data Models." *Stata Journal* 6(4): 482-496.

Huang, X. (2010). "Nonparametric Estimation in Large Panels with Cross Sectional Dependence." *Econometric Reviews* Forthcoming.

Pesaran, M. H. (2004). "General Diagnostic Tests for Cross Section Dependence in Panels." *CESifo Working Paper No 1229*.

Pesaran, M. H. (2006). "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure." *Econometrica* 74(4): 967-1012.

Phillips, P. C. B. and H. R. Moon (1999). "Linear Regression Theory for Nonstationary Panel Data." *Economerica* 67: 1057-1111.

Phillips, P. C. B. and D. Sul (2003). "Dynamic Panel Estimation and Homogeneity Testing under Cross Sectional Dependence." *Econometrics Journal* 6: 217-259.

Phillips, P. C. B. and D. Sul (2007). "Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence." *Journal of Econometrics* 137: 162-188.

Robertson, D. and J. Symons (2000). "Factor Residuals in Sur Regressions: Estimating Panels Allowing for Cross Sectional Dependence." *CEP discussion paper No. 0473, Centre for Economic Performance, LSE, London*.

Sarafidis, V. (2009). "Gmm Estimation of Short Dynamic Panel Data Models with Error Cross-Sectional Dependence." *MPRA paper No. 25176*.

Sarafidis, V. and D. Robertson (2009). "The Impact of Error Cross-Sectional Dependence in Short Dynamic Panel Estimation". *Econometrics Journal* 12(1): 68-81.

Sarafidis, V., T. Yamagata and D. Robertson (2009). "A Test of Cross Section Dependence for a Linear Dynamic Panel Model with Regressors." *Journal of Econometrics* 143(2): 149-161.

Su, L. and S. Jin (2010). "Sieve Estimation of Panel Data Models with Cross Sectional Dependence." *Journal of Econometrics* Forthcoming.