UNIVERSITY OF TECHNOLOGY SYDNEY

ECONOMICS DISCIPLINE GROUP

UTS BUSINESS SCHOOL

WORKING PAPER NO. 20 March 2014

Consistent Estimation of Panel Data Models with a Multifactor Error Structure when the Cross Section Dimension is Large

Bin Peng
Giovanni Forchini

ISSN: 2200-6788
http://www.business.uts.edu.au/economics/

Consistent Estimation of Panel Data Models with a Multifactor Error Structure when the Cross Section Dimension is Large[1]

30 Jan 2014

Bin Peng[2]

University of Technology, Sydney

Giovanni Forchini[3]

University of Surrey

Abstract

The paper studies a panel data models with a multifactor structure in both the errors and the regressors in a microeconometric setting in which the time dimension is fixed and possibly very small. An estimator is proposed that is consistent for fixed T as N tends to infinity and that does not impose restrictive conditions on the number of factors or the number of regressors or the time series properties of the panel. A small Monte Carlo simulation shows that this estimator is very accurate for very small values of T. Two empirical cases are provided to demonstrate performances of our estimator in practice.

Key words: Panel data model, Cross-Sectional Dependence, Asymptotic Theory

JEL Codes: C10, C13, C23

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

[2] Address for correspondence: Bin Peng, Level 4, 645 Harris St, Ultimo NSW, Australia, 2007. E-mail: Bin.Peng@uts.edu.au.

[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

The effects of common shocks – which may be macroeconomic, technological, institutional, political, environmental, health, sociological, etc. (e.g. Andrews (2005)) - have been recently investigated by various authors including, among others, Case (1991), Conley (1999), Andrews (2005), Pesaran (2006) and Bai (2009) (see also the references therein). These shocks induce cross-sectional dependence in panel data models which is often modelled in a parsimonious way through the use of factors. The earlier contributions allow only for models with factors in the errors (e.g. Case (1991), Conley (1999)) for which consistent estimation of the interest parameters could be done by maximum likelihood procedures (e.g. Robertson and Symons (2000)). Coakley, Fuertes and Smith (2002) suggest an estimation procedure based on principal components applied to the residuals. More recently, it has been noticed by several authors that common shocks would likely affect both the errors and the regressors (see among others Andrews (2005) and Pesaran (2006)) and would thus induce endogeneity requiring more sophisticated estimation procedures.

Most of the current literature has a macroeconometric focus and has mainly developed from the seminal contribution of Pesaran (2006) who proposes estimators which are consistent 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 (2012) and Huang (2013) 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), Coakley, Fuertes and Smith (2006), Sarafidis and Robertson (2009) and Sarafidis (2009).

The two main features of the macroeconometric literature on cross-sectional dependence are: (1) the time dimension of the panel is large and the cross-sectional dimension may be small (e.g. Phillips and Sul (2003)) or more commonly it is also large (e.g. Pesaran (2006), Hoyos and Sarafidis (2006) and Sarafidis, Yamagata and Robertson (2009), Sarafidis and Robertson (2009), and Sarafidis (2009), Bai (2009)); and (2) the number of explanatory variables is usually small. On the other hand, microeconometric models are usually characterized by a very large number of individuals, a large number of explanatory variables, usually much larger than the time dimension which tends to be small. Typically, microeconometric models also involve endogeneity (not caused by the factor structure) and may be semi- or non-parametric. For these models it is also unreasonable to assume that all regressors are affected by shocks such as age, gender, race, etc. Therefore, the solutions offered by the macroeconometric literature are unsuitable and very often not even applicable. Surprisingly, the microeconometric literature dealing with common shocks affecting the errors as well as the regressors is very small.

The main contributions to the microeconometric literature involving a factor structure in the errors as well as correlation between the errors and the regressors are given by Ahn, Lee and Schmidt (2001) and Ahn, Lee and Schmidt (2013), Andrews (2005) and Kuersteiner and Prucha (2013). Ahn, Lee and Schmidt (2001) and Ahn, Lee and Schmidt (2013) study a model in which the factors are fixed (rather than random as in this paper). Regarding the factors as parameters creates identification problems requiring a standardization of the factors which is exploited by Ahn, Lee and Schmidt (2001) and Ahn, Lee and Schmidt (2013) to construct GMM estimators for the slope coefficients. Andrews (2005) and Kuersteiner and Prucha (2013) regard the factors as unobservable random variables and as such their models seem to be suited to capture common shocks. Andrews (2005) gives conditions for consistency of the OLS estimator in a cross-section regression with common shocks and extends his results to some panel data estimators. Kuersteiner and Prucha (2013) extend the work of Andrews (2005) by deriving a central limit theory for sample moments under weaker assumptions and establishes limiting distribution of GMM and maximum likelihood estimators for general models in which the regressors may present cross-sectional correlation.

This paper builds on this literature by considering a linear panel data model in which the errors and the regressors are affected by common shocks represented by common factors. The model is very similar to the one suggested by Pesaran (2006)) but the underlying assumptions are expressed conditionally on the common factors in the same spirit as Andrews (2005) and Kuersteiner and Prucha (2013). We suggest a GMM estimator that it is a standard fixed effects estimator in which the N and T dimensions are interchanged and as such it is very easy to calculate and can be applied in a microeconometric set up in which the number of regressors is large (and possibly much larger than the time series dimension) and no restriction is imposed on the number of factors (c.f. Ahn, Lee and Schmidt (2001), Ahn, Lee and Schmidt (2013)). We will require the regressors to be exogenous even if this condition can be weaken as in Kuersteiner and Prucha (2013) and will analyse a model in which classical endogeneity is present in a companion paper.

The structure of the paper is as follows. Section 2 considers a panel data model with homogeneous slopes, discusses the assumptions, proposes a generalized method of moment estimators and studies its asymptotic properties. Section 3 extends these results to a model with heterogeneous slopes. Section 4 discusses a simple simulation exercise. Two empirical examples are provided in Section 5. Section 6 concludes. All proofs are in the Appendix.

2. Homogeneous slopes

We consider a simple panel data model with cross-sectional dependence and correlation between the errors and the regressors:

(1) $$\underset{(T\times 1)}{y_i} = \underset{(T\times k)}{z_i}\underset{(k\times 1)}{\alpha_0} + \underset{(T\times p)}{x_i}\underset{(p\times 1)}{\beta_0} + e_i$$

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

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

The observed regressors are split into two groups: those which are not affected by common shocks (e.g. individual characteristics such as gender, race, age, etc.), z_i, and those which may be affected by common shocks, x_i. The parameters associated to the regressors, α_0 and β_0, are the same for every $i=1,2,...,N$ and $t=1,2,....,T$ (hence the homogeneous slopes). The common shocks are captured by the unobserved matrix of common factors, $F_T = (f_1, f_2,..., f_T)'$, (c.f. Andrews (2005)); γ_i and Γ_i are a $(m\times 1)$ vector and a $(m\times p)$ matrix of factor loadings; ε_i is a purely idiosyncratic random vector with zero mean and constant covariance matrix; and v_i represents the values of the regressors that would be observed in the absence of common shocks. Factors, factor loadings, v_i and ε_i are not observed. The factor structure generates cross-sectional heterogeneity in the error term of (1). This also creates correlation between errors e_i and regressors x_i so that standard estimators of the parameters in (1) are inconsistent.

We now introduce some assumptions on both the observed and the unobserved variables. We assume that the matrix of factors, which we do not observe, is random and finite with probability one. Since we regard the time dimension as fixed, no other assumptions for the factors are needed.

All variables are defined on a probability space $(\Omega, \mathcal{A}, P)$. The sigma algebra generated by the random vector $vec(F_T)$ is denote by $\mathcal{F} = \{\omega \in \mathcal{A} : vec(F_T)(\omega) \in \mathcal{B}^{Tm}\}$ where $\mathcal{B}^{Tm}$ is the Borel sigma algebra in $\mathbb{R}^{Tm}$. Notice that $\mathcal{F}$ is a sub-algebra of $\mathcal{A}$. Notice also that expectations and probabilities conditional on $\mathcal{F}$ are unique up to a.s. equivalence, so that for example two conditional expectations which differ only on sets with probability zero are regarded as equivalent. We will regard conditioning on $\mathcal{F}$ as conditioning on the factors F_T.

Assumption C.1. The sequence of random vectors $\{\varepsilon_i, i=1,...,n\}$ is conditionally independent given $\mathcal{F}$, $E[\varepsilon_i \mid \mathcal{F}] = 0$ a.s. and $E\left[\|\varepsilon_i\|_2^{1+\delta} \mid \mathcal{F}\right] < \Delta < \infty$ a.s. for some $\delta > 0$, $i=1,....,N$.

Assumption C.2. The sequence of random matrices $\{(z_i, v_i), i=1,...,n\}$ is conditionally independent given $\mathcal{F}$ with $E\left[\|(z_i, v_i)\|_2^{2+\delta} \mid \mathcal{F}\right] < \Delta < \infty$ a.s. for some $\delta > 0$, $i=1,....,N$.

Assumption C.3. The sequence of random vectors $\{\gamma_i, i=1,...,n\}$ is conditionally independent given $\mathcal{F}$, $E[\gamma_i \mid \mathcal{F}] = \gamma$ a.s. and $E\left[\|\gamma_i\|_2^{1+\delta} \mid \mathcal{F}\right] < \Delta < \infty$ a.s. for some $\delta > 0$, $i = 1,....,N$.

Assumption C.4. The sequence of random vectors $\{vec(\Gamma_i), i=1,...,n\}$ is conditionally independent given $\mathcal{F}$, $E\left[vec(\Gamma_i) \mid \mathcal{F}\right] = vec(\Gamma)$ a.s. and $E\left[\|\Gamma_i\|_2^{2+\delta} \mid \mathcal{F}\right] \le M < \infty$ a.s. for some $\delta > 0$, $i = 1,....,N$.

Assumption C.5. The random vectors ε_i, $vec(z_i, v_i)$, γ_i and $vec(\Gamma_i)$ are conditionally independent given $\mathcal{F}$ for all $i = 1,2,...,N$.

Notice that the expectations in the assumptions hold a.s. since they involve conditional expectations which are random variables and may fail on sets of probability zero. The random vectors ε_i's are assumed to be purely idiosyncratic and the (z_i, v_i)'s are assumed to form a sequence of independent observations given the factors. Since we interpret v_i as a vector of regressors which would be observed if the common shocks would not affect the regressors, we need to assume that these form an independent sequence of events which are heterogeneous and may be correlated with z_i.

The factor loadings in both the regressors and the errors are assumed to be independent conditional on $\mathcal{F}$ but not necessarily identically distributed. Notice that Assumption C.5 only requires independence conditional on $\mathcal{F}$ but does not require the factor loadings in the regressors and errors to be independent unconditionally (c.f. Pesaran (2006)). Thus, the covariance matrix between x_i and e_i conditional on $\mathcal{F}$ is zero:

$$\begin{aligned}
\operatorname{cov}(x_i, e_i \mid \mathcal{F}) &= (I_p \otimes F_T)\cdot\left\{E\left[vec(\Gamma_i)\cdot\gamma_i' \mid \mathcal{F}\right] - E\left[vec(\Gamma_i) \mid \mathcal{F}\right]\cdot E\left[\gamma_i' \mid \mathcal{F}\right]\right\}\cdot F_T' \\
&= (I_p \otimes F_T)\cdot\left\{E\left[vec(\Gamma_i) \mid \mathcal{F}\right]\cdot E\left[\gamma_i' \mid \mathcal{F}\right] - E\left[vec(\Gamma_i) \mid \mathcal{F}\right]\cdot E\left[\gamma_i' \mid \mathcal{F}\right]\right\}\cdot F_T' \\
&= 0.
\end{aligned} \tag{4}$$

On the other hand, the unconditional covariance matrix between regressors and errors,

$$\begin{aligned}
\operatorname{cov}(x_i, e_i) &= E\left[(I_p \otimes F_T)vec(\Gamma_i)\cdot\gamma_i' F_T'\right] - E\left[(I_p \otimes F_T)vec(\Gamma_i)\right]E\left[\gamma_i' F_T'\right] \\
&= E\left[(I_p \otimes F_T)E\left[vec(\Gamma_i)\cdot\gamma_i' \mid \mathcal{F}\right]F_T'\right] \\
&\quad - E\left[(I_p \otimes F_T)E\left[vec(\Gamma_i) \mid \mathcal{F}\right]\right]\cdot E\left[E\left[\gamma_i' \mid \mathcal{F}\right]F_T'\right] \\
&= E\left[(I_p \otimes F_T)vec(\Gamma)\cdot\gamma' F_T'\right] - E\left[(I_p \otimes F_T)vec(\Gamma)\right]\cdot E\left[\gamma' F_T'\right],
\end{aligned} \tag{5}$$

is, in general, not zero. Thus Assumptions C.1-C.5 imply endogeneity due to the factor structure. More traditional forms of endogeneity in the presence of factor structures in the errors and the regressors are studied in a companion paper.

Ahn, Lee and Schmidt (2001) consider a similar model to ours but they do not make explicit the dependence of the regressors on the factors or the factor loadings. They regard the factors as fixed parameters so that it is as if they would conduct the analysis conditional on the factors. Our work differs from theirs on three major aspects. Firstly, regarding the factors as parameters implies a frequentist sampling scheme whereby the repeated sampling takes place with the factors unchanged from sample to sample. In our model repeated sampling takes place with the factors F_T also been resampled. Since the factors are not observable, we believe that our sampling scheme is a more realistic way of capturing common shocks. Secondly, as the first line of (4) shows, once we condition on the factors, the correlation between the regressors and the errors must come through the correlation between the regressors and the factor loadings given the factors in their case, and this makes the analysis more complicated. Thirdly, Ahn, Lee and Schmidt (2001) do not allow for cross-sectional dependence among the errors in the sense that

$$\text{(6)} \qquad \operatorname{cov}\left(e_i, e_j \mid \mathcal{F}\right) = F_T \gamma\gamma' F_T' - F_T \gamma\gamma' F_T' = 0,$$

whereas, unconditionally

$$\text{(7)} \qquad \begin{aligned} \operatorname{cov}\left(e_i, e_j\right) &= E\left[\left(F_T\gamma_i + \varepsilon_i\right)\left(F_T\gamma_j + \varepsilon_j\right)'\right] - E\left[F_T\gamma_i + \varepsilon_i\right]E\left[\gamma_j' F_T' + \varepsilon_j'\right] \\ &= E\left[E\left[F_T\gamma_i\gamma_j' F_T' \mid \mathcal{F}\right]\right] - E\left[E\left[F_T\gamma_i \mid \mathcal{F}\right]\right]\cdot E\left[E\left[\gamma_j' F_T' \mid \mathcal{F}\right]\right] \\ &= E\left[F_T\gamma\gamma' F_T'\right] - E\left[F_T\gamma\right]\cdot E\left[\gamma' F_T'\right]. \end{aligned}$$

Regarding the factors as parameters also entails identification problems (c.f. Ahn, Lee and Schmidt (2001)). More details will be discussed below.

Andrews (2005) considers a similar model for $T = 1$ in which ε_i and v_i are identically zero and studies the OLS estimator. He assumes that the factors and the factor loadings are mutually independent and shows that if the factor loadings in the regressors and those in the errors are conditionally independent then the OLS estimator is consistent, and, once renormalized in the usual way, has an asymptotic mixed normal distribution. However, if the two sets of factor loadings are conditionally dependent, the OLS estimator is inconsistent and has an asymptotic mixed normal distribution once it has been re-centred in a suitable way.

The assumptions of Pesaran (2006) are slightly different because he specifies assumptions for the unconditional distributions of the various quantities involved and allows for both T and N to tend to infinity. Thus, Pesaran (2006) imposes stationarity on the distributions of the factors and the components of ε_i and v_i and assumes independence of the components of ε_i and v_i over the cross-sectional dimension. Since the time dimension is fixed in our model, we do not need any assumption on the temporal dependence for ε_i, v_i and factors. We also do not impose that ε_i and v_i are independent unconditionally. Moreover, we impose independence between the factor loadings in the regressors and the errors conditional on the factors but do not make any claim about their unconditional distribution and so may or may not violate Assumption 3 of Pesaran (2006).

To develop consistent estimators of β_0 and γ_0 when $N \to \infty$ and T is fixed, we notice that, given Assumptions C.1-C.5, the following conditional moments hold

$$E\left[(z_i, x_i)' e_i \mid \mathcal{F}\right] = E\left[(z_i, x_i)' e_j \mid \mathcal{F}\right] = \left(E\left[z_i \mid \mathcal{F}\right], E\left[v_i \mid \mathcal{F}\right] + F_T \Gamma\right)' F_T \gamma \tag{8}$$

for $i, j = 1, 2, ..., N$, which can be rewritten as $E\left[(z_i, x_i)'\left(e_i - e_j\right) \mid \mathcal{F}\right] = 0$ a.s. implying the unconditional moments $E\left[(z_i, x_i)'\left(e_i - e_j\right)\right] = 0$. The corresponding empirical moments are

$$\frac{1}{N}\sum_{i=1}^{N}(z_i, x_i)'(e_i - \bar{e}) = 0, \tag{9}$$

where $\bar{e} = \frac{1}{N}\sum_{j=1}^{N} e_j$. Then the resulting estimator of $(\alpha_0', \beta_0')'$ is

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \left(\frac{1}{N}\sum_{i=1}^{N}(z_i, x_i)'\left((z_i, x_i) - (\bar{z}, \bar{x})\right)\right)^{-1} \frac{1}{N}\sum_{i=1}^{N}(z_i, x_i)'(y_i - \bar{y}), \tag{10}$$

where $\bar{y} = (1/N)\sum_{i=1}^{N} y_i$ and $\bar{x} = (1/N)\sum_{i=1}^{N} x_i$. If all the regressors are affected by the factors, the above estimator simplifies to an estimator proposed by Coakley, Fuertes and Smith (2006).

$$\hat{\beta} = \left(\sum_{i=1}^{N} x_i'(x_i - \bar{x})\right)^{-1} \sum_{i=1}^{N} x_i'(y_i - \bar{y}). \tag{11}$$

Notice that the estimator in (10) is the ordinary least squares estimator for the transformed model

$$y_i - \bar{y} = (z_i - \bar{z}, x_i - \bar{x})\begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} + e_i - \bar{e}.$$

Thus, it is a fixed effects estimator for the model

$$y_i = a + (z_i, x_i)\begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} + e_i,$$

where a is a $T \times 1$ vector of fixed effects (notice that the i and t dimensions are interchanged from the “standard” fixed effects estimator). The estimator in (10) can be easily estimated using standard statistical packages.

In order to compare our estimator with the GMM estimator of Ahn, Lee and Schmidt (2001) we now take $m=1$. By standardizing the factors as $F_T=\left(1, f_2, f_3,...., f_T\right)'$ and considering them as parameters, Ahn, Lee and Schmidt (2001) notice that

$$\begin{aligned} E\left[\left(z_i, x_i\right)e_{it} \mid F_T\right] &= E\left[\left(z_i, v_i + F_T\Gamma_i\right)\left(f_t\gamma_i + \varepsilon_{it}\right)\right] \\ &= \left(f_tE\left[z_i \mid F_T\right]\gamma, f_tE\left[v_i \mid F_T\right]\gamma + f_tF_TE\left[\Gamma_i\gamma_i\right]\right) \end{aligned}$$

and

$$\begin{aligned} E\left[\left(z_i, x_i\right) f_t e_{i1} \mid F_T\right] &= E\left[\left(z_i, F_T\Gamma_i + v_i\right) f_t\left(\gamma_i + \varepsilon_{i1}\right)\right] \\ &= \left(f_tE\left[z_i \mid F_T\right]\gamma, f_tE\left[v_i \mid F_T\right]\gamma + f_tF_TE\left[\Gamma_i\gamma_i\right]\right). \end{aligned}$$

Thus $E\left[vec\left(\left(z_i, x_i\right)\left(e_{it} - f_te_{i1}\right)\right) \mid F_T\right]=0$, so that an estimator of α_0 and β_0 can be obtained from the $\left(p+k\right)\left(T-1\right)T$ empirical equivalents to the moments condition

$$E\left[vec\left(\left(z_i, x_i\right)\left(e_{it} - f_te_{i1}\right)\right)\right]=0.$$

For the case where $m>1$, Ahn, Lee and Schmidt (2013) use the same technique to introduce a normalization and exploit this for the estimation of the interest parameter. These are very different from the moment conditions in (8) and closely exploit the standardization of the factors in a standard factor model (for a clear discussion see Geweke and Zhou (1996)) which is not invariant to different ordering of the variables (c.f. Chan, Leon-Gonzalez and Strachan (2013)).

Consistency of our estimator in (10) is given in the following theorem.

Theorem 1. Given Assumptions C.1-C.5, if

$$E\left[\frac{1}{N}\sum_{i=1}^{N}\left(z_i-\overline{z}, v_i-\overline{v}\right)'\left(z_i-\overline{z}, v_i-\overline{v}\right) \mid \mathcal{F}\right] \quad \text{and} \quad E\left[\frac{1}{N}\sum_{i=1}^{N}\left(\Gamma_i-\overline{\Gamma}\right)'F_T'F_T\left(\Gamma_i-\overline{\Gamma}\right) \mid \mathcal{F}\right]$$

are uniformly positive definite a.s., the following results hold:

1) $\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}$ is unbiased;

2) $\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix} \to \begin{pmatrix}\alpha_0\\ \beta_0\end{pmatrix}$ a.s..

Theorem 1 shows that, for fixed T, the estimators $\hat{\alpha}$ and $\hat{\beta}$ are unbiased and consistent as N tends to infinity. For the case where the factors affect all the regressors, unbiasedness was also noted for the estimator in (11) by Coakley, Fuertes and Smith (2006) who assume that the factor loadings in the regressors and the errors are mutually independent (which is stronger than Assumption C.5).

To obtain the asymptotic distribution of $\hat{\alpha}$ and $\hat{\beta}$ we need slightly stronger versions of Assumptions C.1 and C.3 requiring the existence of higher order moments.

Assumption N.1. The sequence of random vectors $\{\varepsilon_i, i=1,...,n\}$ is conditionally independent given $\mathcal{F}$, $E[\varepsilon_i \mid \mathcal{F}]=0$ a.s., $E[\varepsilon_i\varepsilon_i' \mid \mathcal{F}]=\Sigma_{\varepsilon i}$ a.s. and $E\left[\|\varepsilon_i\|^{2+\delta} \mid \mathcal{F}\right]<\Delta<\infty$ for some $\delta>0$ a.s., $i=1,....,N$.

Assumption N.3. The sequence of random vectors $\{\gamma_i, i=1,...,n\}$ is conditionally independent given $\mathcal{F}$, $E[\gamma_i \mid \mathcal{F}]=\gamma$ a.s., $\operatorname{cov}[\gamma_i \mid \mathcal{F}]=\Sigma_{\gamma i}$ a.s. and $E\left[\|\gamma_i\|^{2+\delta} \mid \mathcal{F}\right]<\Delta<\infty$ for some $\delta>0$ a.s., $i=1,....,N$.

Theorem 2. Given Assumptions N.1, C.2, N.3, C.4 and C.5, if

$$E\left[\frac{1}{N}\sum_{i=1}^{N}(z_i-\overline{z},v_i-\overline{v})'(z_i-\overline{z},v_i-\overline{v}) \mid \mathcal{F}\right] \quad \text{and} \quad E\left[\frac{1}{N}\sum_{i=1}^{N}\left(\Gamma_i-\overline{\Gamma}\right)'F_T'F_T\left(\Gamma_i-\overline{\Gamma}\right) \mid \mathcal{F}\right]$$

are uniformly positive definite a.s., the following results hold conditional on $\mathcal{F}$:

$$\sqrt{N}\left(\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}-\begin{pmatrix}\alpha_0\\ \beta_0\end{pmatrix}\right)\to^{D}\left(B(F_T)\right)^{-1}\left(A(F_T)\right)^{1/2}N\left(0,I_{p+k}\right) \text{ (stably),}$$

where

$$A(F_T)=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}E\left[\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\left(\Sigma_{\varepsilon i}+F_T\Sigma_{\gamma i}F_T'\right)\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right) \mid \mathcal{F}\right]$$

$$B(F_T)=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(E[\omega_i'\omega_i \mid \mathcal{F}]-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)'\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right) \mid \mathcal{F}\right)$$

$$\omega_i=(z_i,x_i) \quad \text{and} \quad (\mu_{zN},\tau_{vN})=\frac{1}{N}\sum_{i=1}^{N}E\left[(z_i,v_i) \mid \mathcal{F}\right].$$

Corollary 1. Given Assumptions N.1, C.2, N.3, C.4 and C.5, if

$$E\left[\frac{1}{N}\sum_{i=1}^{N}(z_i-\overline{z},v_i-\overline{v})'(z_i-\overline{z},v_i-\overline{v}) \mid \mathcal{F}\right] \quad \text{and} \quad E\left[\frac{1}{N}\sum_{i=1}^{N}\left(\Gamma_i-\overline{\Gamma}\right)'F_T'F_T\left(\Gamma_i-\overline{\Gamma}\right) \mid \mathcal{F}\right]$$

are uniformly positive definite a.s.,

$$\sqrt{N}\left(\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}-\begin{pmatrix}\alpha_0\\ \beta_0\end{pmatrix}\right)\to^{D} X,$$

where

$$X \equiv \int N\left(0,\left(B\left(F_T\right)\right)^{-1} A\left(F_T\right)\left(B\left(F_T\right)\right)^{-1}\right)\cdot pdf\left(F_T\right) dF_T .$$

Theorem 2 shows that, for fixed T, $\hat{\alpha}$ and $\hat{\beta}$ have a normal asymptotic distribution conditional on the factors. However, removing the conditioning on F_T, the asymptotic distribution of our estimator is covariance-matrix-mixed normal with mixing density given by the density function of the factors (c.f. Corollary 1) which is unknown.

Theorem 2 implies that $\sqrt{N}\left(\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}-\begin{pmatrix}\alpha_0\\ \beta_0\end{pmatrix}\right)$ converges stably over $\mathcal{F}$ in the sense of Renyi (1963): The sequence of random variables $\xi_N = \xi_N(\omega)$, $N = 1,2,...$ is *stable* if for any event $F \in \mathcal{F}$ with $\Pr(F) > 0$, the conditional distribution of ξ_N with respect to F tends to a limiting distribution (see also Kuersteiner and Prucha (2013)). Stability implies the mixed normality result in Corollary 1, which for the case where $T = 1$, is analogous to the results for the ordinary least squares estimators derived by Andrews (2005) (e.g. his Theorem 4). Similar mixed normality results are obtained by Kuersteiner and Prucha (2013).

We now briefly deal with the problem of hypothesis testing in this set-up. Even if the relevant distribution for $\left(\hat{\alpha}', \hat{\beta}'\right)'$ is the unconditional one, which is nonstandard, tests of hypotheses can be constructed as usual. In order to do this we need to be able to "estimate" $A(F_T)$ and $B(F_T)$ conditional on $\mathcal{F}$. From the proof of Theorem 1 we know that

$$\hat{B} = \frac{1}{N}\sum_{i=1}^{N}\left(z_i, x_i\right)'\left(\left(z_i, x_i\right) - \left(\overline{z}, \overline{x}\right)\right) \to B\left(F_T\right) \text{ a.s..}$$

For $A(F_T)$ we need more restrictive versions of Assumptions C.2 and C.4 requiring the existence of higher order moments.

Assumption CM.2. The sequence of random matrices $\left\{\left(z_i, v_i\right), i = 1,...,n\right\}$ is conditionally independent given $\mathcal{F}$ with $E\left[\left\|\left(z_i, v_i\right)\right\|_2^{4+\delta} \mid \mathcal{F}\right] < \Delta < \infty$ a.s. for some $\delta > 0$, $i = 1,....,N$.

Assumption CM.4. The sequence of random vectors $\left\{vec\left(\Gamma_i\right), i = 1,...,n\right\}$ is conditionally independent given $\mathcal{F}$, $E\left[vec\left(\Gamma_i\right) \mid \mathcal{F}\right] = vec\left(\Gamma\right)$ a.s. and $E\left[\left\|\Gamma_i\right\|_2^{4+\delta} \mid \mathcal{F}\right] \le M < \infty$ a.s. for some $\delta > 0$, $i = 1,....,N$.

Lemma 1. Given Assumptions N.1, CM.2, N.3, CM.4 and C.5, then

$$
\begin{aligned}
\hat{A} = \frac{1}{N}\sum_{i=1}^{N}\Bigg\{ & \left((z_i, x_i) - (\bar{z}, \bar{x})\right)' \left(y_i - \bar{y} - \left((z_i, x_i) - (\bar{z}, \bar{x})\right)\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \right) \\
& \cdot \left(y_i - \bar{y} - \left((z_i, x_i) - (\bar{z}, \bar{x})\right)\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \right)' \left((z_i, x_i) - (\bar{z}, \bar{x})\right) \Bigg\} \\
\to A(F_T) \quad & \text{a.s..}
\end{aligned} \tag{12}
$$

An asymptotic version of the F-test conditional on $\mathcal{F}$ for the null hypothesis that $H_0 : R(\alpha_0', \beta_0')' = r$ against the alternative hypothesis $H_0 : R(\alpha_0', \beta_0')' \neq r$, where R is a known and fixed $q \times (p+k)$ matrix of rank $q < p+k$ and r is a known and fixed $q \times 1$ vector can be easily constructed since conditional on $\mathcal{F}$

$$
N\left(R\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} - r \right)' \left(R\hat{B}^{-1}\hat{A}\hat{B}^{-1}R' \right)^{-1} \left(R\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} - r \right) \to^{D} \chi^2(q) \tag{13}
$$

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

Asymptotic version of the t-test can be constructed similarly. If we denote by $\hat{\sigma}_i^2(F_T)$ the element in position (i,i) of the matrix $\hat{B}^{-1}\hat{A}\hat{B}^{-1}$, we can test the null hypothesis that the ith component of $(\alpha_0', \beta_0')'$ equals to a fixed value r by noting that under the null hypothesis conditional on $\mathcal{F}$

$$
\left(\hat{\sigma}_i^2\right)^{-1/2}\left(\hat{\theta}_i - r\right) \to^{D} N(0,1),
$$

where $\hat{\theta}_i$ denotes the ith component of $\hat{\theta} = \left(\hat{\alpha}', \hat{\beta}'\right)'$. Once again the limiting distribution under the null hypothesis does not depend on $\mathcal{F}$ so that

$$
\left(\hat{\sigma}_i^2\right)^{-1/2}\left(\hat{\theta}_i - r\right) \to^{D} N(0,1).
$$

We summarise this in the following corollary.

Corollary 2. Let R be a known and fixed $q \times (p+k)$ matrix of rank $q < p+k$ and r be a known and fixed $q \times 1$ vector. Given Assumptions N.1, CM.2, N.3, CM.4 and C.5, if the null hypothesis $H_0 : R(\delta_0', \beta_0')' = r$ holds then as $N \to \infty$,

(14) $$N\left(R\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}-r\right)'\left(R\hat{B}^{-1}\hat{A}\hat{B}^{-1}R'\right)^{-1}\left(R\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}-r\right)\to^{D}\chi^{2}(q).$$

Moreover, if $q=1$

(15) $$\frac{1}{\sqrt{R\hat{B}^{-1}\hat{A}\hat{B}^{-1}R'}}\left(R\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}-r\right)\to^{D}N(0,1).$$

Notice that the distributions of the two test statistics above under the null hypothesis do not depend on $\mathcal{F}$, however, they do depend on $\mathcal{F}$ under the alternative hypothesis.

We now investigate briefly the effects of dependence between the factor loadings in the regressors and the errors conditional on $\mathcal{F}$ for our estimator.

Assumption D.5. The random vectors ε_i, $vec(z_i,v_i)$ and $\begin{pmatrix}\gamma_i\\ vec(\Gamma_i)\end{pmatrix}$ are conditionally independent given $\mathcal{F}$ for all $i=1,2,...,N$.

Theorem 3. Given Assumptions C1-C.4 and D.5, if

$$E\left[\frac{1}{N}\sum_{i=1}^{N}(z_i-\bar{z},v_i-\bar{v})'(z_i-\bar{z},v_i-\bar{v})\,|\,\mathcal{F}\right]\quad\text{and}\quad E\left[\frac{1}{N}\sum_{i=1}^{N}\left(\Gamma_i-\bar{\Gamma}\right)'F_T'F_T\left(\Gamma_i-\bar{\Gamma}\right)|\,\mathcal{F}\right]$$

are uniformly positive definite a.s., conditional on $\mathcal{F}$

$$\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}\to\begin{pmatrix}\alpha_0\\ \beta_0\end{pmatrix}+B(F_T)^{-1}\begin{pmatrix}0\\ \Delta(F_T)\end{pmatrix}\text{ a.s.,}$$

where $B(F_T)$ is defined in Theorem 2 and

$$\Delta(F_T)=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}E\left[\Gamma_i'F_T'F_T\gamma_i\,|\,\mathcal{F}\right]-\Gamma'F_T'F_T\gamma.$$

Notice that by replacing Assumption C.5 with Assumption D.5, the estimators of both δ_0 and β_0 have an asymptotic bias conditional on $\mathcal{F}$, dependent in a complicated way on the correlation between the factor loadings and the distribution of the factors. This implies that unconditionally, the estimators of both α_0 and β_0 have a non-degenerate non-standard asymptotic distribution: $\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}\to^{D}\begin{pmatrix}\alpha_0\\ \beta_0\end{pmatrix}+B(F_T)^{-1}\begin{pmatrix}0\\ \Delta(F_T)\end{pmatrix}$.

3. Heterogeneous slopes

In this section, we consider a more general case, where the coefficients of z_i and x_i are allowed to be different for each individual. Precisely, the model is

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

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

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

(19) $$\begin{pmatrix}\alpha_i\\ \beta_i\end{pmatrix} = \begin{pmatrix}\alpha_0\\ \beta_0\end{pmatrix} + \begin{pmatrix}\eta_{1i}\\ \eta_{2i}\end{pmatrix},$$

where η_{1i} and η_{2i} are random variables. We are interested in inference about the mean of the individual-specific slope coefficients α_0 and β_0. These parameters are estimated using the estimator (10) derived in the previous section.

Some further assumptions are needed.

Assumption H.5. Conditional on $\mathcal{F}$, $(\eta_{1i}',\eta_{2i}')'$, v_i, ε_i, Γ_i and γ_i are independent for $i=1,\ldots,N$.

Assumption H.6. The sequence of random vectors $\{(\eta_{1i}',\eta_{2i}')', i=1,...,n\}$ is conditionally independent given $\mathcal{F}$ with $E\left[(\eta_{1i}',\eta_{2i}')' \mid \mathcal{F}\right]=0$ and $E\left[\left\|(\eta_{1i}',\eta_{2i}')'\right\|_2^{1+\delta} \mid \mathcal{F}\right] < \Delta < \infty$.

Assumption HN.6. The sequence of random vectors $\{(\eta_{1i}',\eta_{2i}')', i=1,...,n\}$ is conditionally independent given $\mathcal{F}$ with $E\left[(\eta_{1i}',\eta_{2i}')' \mid \mathcal{F}\right]=0$, $E\left[(\eta_{1i}',\eta_{2i}')'(\eta_{1i}',\eta_{2i}') \mid \mathcal{F}\right]=\Sigma_\eta$ and $E\left[\left\|(\eta_{1i}',\eta_{2i}')'\right\|_2^{2+\delta} \mid \mathcal{F}\right] < \Delta < \infty$.

Assumption H.5 extends assumption C.5 by requiring that $(\eta_{1i}',\eta_{2i}')'$ is independent of v_i, ε_i, Γ_i and γ_i conditional on $\mathcal{F}$. This does not require unconditional independence so it is weaker than Assumption 3 of Pesaran (2006).

Theorem 4 gives the distributional properties for our estimator.

Theorem 4. If Assumptions C.1, C.2, C.3, C.4, H.5 and H.6 hold and if

$$E\left[\frac{1}{N}\sum_{i=1}^{N}(z_i-\bar{z},v_i-\bar{v})'(z_i-\bar{z},v_i-\bar{v}) \mid \mathcal{F}\right] \text{ and } E\left[\frac{1}{N}\sum_{i=1}^{N}(\Gamma_i-\bar{\Gamma})'F_T'F_T(\Gamma_i-\bar{\Gamma}) \mid \mathcal{F}\right]$$

are uniformly positive definite a.s., then

(1) $\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$ is unbiased; and

(2) $\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$ is consistent.

If Assumptions N.1, CM.2, N.3, CM.4, H.5 and HN.6 hold, and

$$E\left[\frac{1}{N}\sum_{i=1}^{N}\left(z_i-\bar{z},v_i-\bar{v}\right)'\left(z_i-\bar{z},v_i-\bar{v}\right)\mid \mathcal{F}\right] \quad \text{and} \quad E\left[\frac{1}{N}\sum_{i=1}^{N}\left(\Gamma_i-\bar{\Gamma}\right)'F_T\,'F_T\left(\Gamma_i-\bar{\Gamma}\right)\mid \mathcal{F}\right]$$

are uniformly positive definite a.s., then conditional on $\mathcal{F}$,

$$\sqrt{N}\left(\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}-\begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix}\right)\rightarrow^{D} B\left(F_T\right)^{-1}\left(A\left(F_T\right)+C\left(F_T\right)\right)^{1/2} N\left(0,I_{p+k}\right) \text{ (stably)},$$

where

$$C\left(F_T\right)=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}E\left[\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\omega_i\Sigma_\eta\omega_i\,'\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)\mid \mathcal{F}\right],$$

and ω_i, $\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)$, $A\left(F_T\right)$ and $B\left(F_T\right)$ are defined in Theorem 2.

Theorem 4 shows that for a fixed T our estimator is unbiased, consistent and asymptotically normal conditional on the factors. This is different from the conditional asymptotic distribution given in Theorem 2 because of the presence of the term $C\left(F_T\right)$. Thus, the effect of random coefficients on the asymptotic properties of our estimator is just an increase in the variance.

Since the factors are not observable, it is the marginal distribution which is the relevant distribution, so removing the conditioning on $\mathcal{F}$, we obtain the following corollary.

Corollary 3. Given Assumptions N.1, CM.2, N.3, CM.4, H.5 and H.N6 the following result holds:

$$\sqrt{N}\left(\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}-\begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix}\right)\rightarrow^{D} X$$

$$X\equiv\int N\left(0,B\left(F_T\right)^{-1}\left(A\left(F_T\right)+C\left(F_T\right)\right)B\left(F_T\right)^{-1}\right)\cdot pdf\left(F_T\right)dF_T$$

Thus, $\left(\hat{\alpha}',\hat{\beta}'\right)'$ is asymptotically covariance-matrix-mixed normal. Notice that Theorem 3 and Corollary 3 reduce to Theorems 1 and 2 and Corollary 1 if the η_i's are identically zero.

In order to construct tests on α_0 and β_0 we need to find a consistent estimator of the asymptotic covariance matrix $B(F_T)^{-1}\left(A(F_T)+C(F_T)\right)B(F_T)^{-1}$, which is given in the following theorem.

Lemma 2. If assumptions N.1, CM.2, N.3, CM.4, H.5 and HN.6 hold, then a consistent estimator of $A(F_T)+C(F_T)$ can be obtained by expression (12), so that a consistent estimator of the asymptotic covariance matrix conditional on $\mathcal{F}$ in Theorem 3 is still given by expressions $\hat{B}^{-1}\hat{A}\hat{B}^{-1}$ in the previous section.

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

Corollary 4. Let R be a known and fixed $q\times(p+k)$ matrix of rank $q<p+k$ and r be a known and fixed $q\times 1$ vector. If assumptions N.1, CM.2, N.3, CM.4, H.5 and HN.6 and the null hypothesis $H_0 : R(\alpha_0{}',\beta_0{}')'=r$ hold then (14) and (15) still hold for the model (16)-(19).

4. Monte Carlo Study

This section provides some Monte Carlo evidence on the properties of the generalized methods of moment estimator suggested in the previous sections. To simplify the simulation we assume that all the regressors are affected by the common shocks so that we can compare out estimator with that of Ahn, Lee and Schmidt (2001) and Ahn, Lee and Schmidt (2013) (abbreviated in ALS) as well as with the CCEMG and CCEP estimators of Pesaran (2006). These have been designed for a model similar to ours but for a panel in which both the cross-section and the time dimensions are large. We will also include in the comparison the OLS estimator.

We consider a simplified version of the data generating process (DGP) used by Pesaran (2006) in his Section 7. Precisely, we assume the DGP is

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

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

and

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

where $j=1,2$, $i=1,2,...,N$ 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 than the dimension of β (when $p=T$, the matrix $x_i{}'\bar{M}_w x_i$ may or may not be singular depending on the sample realization of x_i and y_i).

The common factors and the errors in (22) are generated as follows

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

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

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

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

The errors in (21) 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 $i = 1,...,[N/2]$ 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 (22) are generated independently as $\Gamma_{ijp} \sim i.i.d.N\left(\mu_{jp}, \sigma_\Gamma^2\right)$ and $\mu_{jp} \sim i.i.d.U(-0.5,1.5)$ for $p = 1,2,3$ and those in (21) 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 both cases where $\beta_{01} = \beta_{02} = 1$ (homogeneous slopes) and $\beta_{ij} = \beta_{0j} + \eta_i$ where $\eta_i \sim i.i.d.N(0,0.4)$ and $\beta_{01} = \beta_{02} = 1$ (heterogeneous slopes). We will report results for some combinations of the parameters (further results are available from the Authors' webpages).

When present, the fixed effects are generated as

$$d_{1t} = 1$$

$$d_{2t} = 0.5 d_{2t-1} + i.i.d.N\left(0, 1-(0.5)^2\right) \quad t = -49,...,0,...,T$$

$$d_{2,-50} = 0$$

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

Tables 1 to 4 show bias and mean-squared errors for the cases without fixed effects by imposing the restrictions $\xi_{i1} = 0$, $a_{ij1} = 0$ and $a_{ij2} = 0$. The estimation of the model without fixed effects is straightforward for all procedures considered. Fixed effects are introduced in Tables 5 and 6 and are dealt with by the method of Ahn, Lee and Schmidt (2013) and by our procedures by regarding d_{1t} and d_{2t}

as unknown factors. For this reason we do not report any evidence for the case where there are fixed effects and only one factor (since this is treated by Ahn, Lee and Schmidt (2013) and by our method as a multifactor model without fixed effects). The method of Ahn, Lee and Schmidt (2013) was devised for the homogeneous slopes case only. Therefore, for this estimator we will compute the ALS estimator for the heterogeneous slopes case as if the slopes would be homogeneous.

We report the bias and mean square error (MSE) for the our estimator (denoted by GMM), for the CCEMG and CCEP estimators of Pesaran (2006) and for the GMM estimators of Ahn, Lee and Schmidt (2001) and Ahn, Lee and Schmidt (2013) (denoted by ALS). We also report results for two-sided "t test" for $H_0 : \beta_{01} = 1$ and $H_0 : \beta_{01} = 0.95$. The Monte Carlo experiments are based on 10,000 replications.

4.1 Bias and MSE

The bias for our GMM estimator is very small for every sample size as one would expect from the fact that it is 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 ALS estimator cannot be calculated for very small T. Its bias seems to increase as T grows.

The MSE of our GMM estimator and ALS estimator decrease as N becomes large but does not seem to be strongly affected by T. Simulations (that are not reported here but are available on the Author's website) show that our GMM estimator is not affected by either the number of regressors or the number of factors. Pesaran's CCEP and CCEMG estimators, on the other hand, are severely affected. The MSE of both the CCEP and CCEMG decrease as either N or T grows and it can be extremely large for small T.

The presence of fixed effects (Tables 5 and 6) does not affect the bias or the MSE of our estimator but considerably affects the CCEP and the CCEMG estimators. As remarked by Pesaran (2006), the CCEP estimator performs better than the CCEMG.

4.2 Size and power

Tables 1 to 6 report a t-type test for $H_0 : \beta_{01} = 1$ against $H_0 : \beta_{01} \neq 1$ constructed using our, the CCEP and CCEMG estimators. The test based on our 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 our estimator does not seem to be affected by T. The power of the test base on our estimator 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 our 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. Empirical Applications

We will now consider two empirical applications.

5.1. Efficiency of health care

Evans, Tandon, Murray and Lauer (2000) have studied the efficiency of health care delivery by various countries using a dataset available at:

http://people.stern.nyu.edu/wgreene/Text/Edition6/tablelist6.htm,

which contains observations on several variables for 191 countries for the period between 1993 and 1997. The variables considered are (1) a composite measure of health care attainment (DALE), (2) a measure of per capita health expenditure (HEXP) and (3) a measure of educational attainment (HC3). Evans, Tandon, Murray and Lauer (2000) estimated the model

(23)
$$\ln\left(DALE_{it}\right)=\alpha_i+\beta_1\ln\left(HEXP_{it}\right)+\beta_2\ln\left(HC3_{it}\right)+\beta_3\left(\ln\left(HC3_{it}\right)\right)^2+e_{it}.$$

The original dataset is an unbalanced panel. In order to keep the example as simple as possible we transform it into a balanced panel by deleting all countries for which the observations are not available for the whole period 1993-1997. By doing so we are left with 140 countries. We will estimate a random effect and a fixed effects model which would be normally used in this context. It is reasonable to assume that in the period considered the errors may be affected by common shocks that may also affect per capita health expenditure and educational attainment, so that we can specify a model of the form

$$\ln\left(DALE_{it}\right)=\alpha+\beta_1\ln\left(HEXP_{it}\right)+\beta_2\ln\left(HC3_{it}\right)+\beta_3\left(\ln\left(HC3_{it}\right)\right)^2+e_{it},$$

(24)
$$e_{it}=f_t'\gamma_i+\varepsilon_{it},$$

$$\left(\ln\left(HEXP_{it}\right),\ln\left(HC3_{it}\right),\left(\ln\left(HC3_{it}\right)\right)^2\right)=v_i+f_t'\Gamma_i$$

where the factor f_t represents the common shock at time t. This model can be estimated using the our estimator as well as the CCEP and CCEMG estimators and the ALS estimator with one factor. We also estimate the first equation in (24) using the Pooled OLS estimator. The results of the estimation procedures are presented in Table 7.

The estimated coefficients for the fixed and random effects models are very close. The Pooled OLS estimator of the first equation in (24) shows a much larger impact of $\ln\left(HEXP_{it}\right)$ on health care attainment than both the fixed and the random effects estimator. When we estimate the model with factors both in the errors and the regressors we find very different results. The estimates of our method suggest that the quadratic term is insignificant and that the impact on health care attainment of both educational attainment and per capita health expenditure is considerably larger than what would be suggested by both the random and the fixed effects model. The magnitude of the CCEP coefficient estimate is much smaller while that of the CCEMG estimate are much more different from any other estimator. We will therefore ignore the CCEMG estimator. For the CCEP estimator the standard errors are not fully interpretable as measures of variability due to the small time dimension (5). The ALS estimator also produces anomalous results.

All estimators produce an elasticity of health care attainment with respect to health care expenditure. In the range covered by the sample, the elasticity of health care attainment with respect to educational attainment is positive. Figure 1 plots these elasticities over the sample values for $HC3_{it}$. Notice that the fixed effects and the random effects model imply an elasticity of health care attainment with respect to educational attainment which is increasing, while our estimator and pooled OLS estimator imply an elasticity which is decreasing as educational attainment increases. For the ALS estimator such estimated elasticity becomes negative for $HC3_{it} > 4.115$. The policy implications of these different results could be substantial.

5.2 The Balassa-Samuelson effect

The Balassa-Samuelson hypothesis states that the differences in individual countries' exchange rate ratios (i.e. the ratio between purchasing power parity relative to the US, say, and the nominal exchange rate relative to the US) can be explained by the GDP per capita measured in purchasing power parity (e.g. de Boeck and Slok (2006)). In its simplest formulation this can be written as

$$\log(\text{exchange rate ratio}_{it}) = \beta_{1i} + \beta_{2i} \cdot \log(\text{PPP GDP per capita}_{it}) + e_{it} .$$

The data is available from "Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 7.1, Center for International Comparison of Production, Income and Prices at University of Pennsylvania, July 2012" for 188 countries for the period 2001-2010. In the period considered, the global financial has taken place so that common shocks are plausible in this period. The scatter plot of the data reported in Figure 2 shows that the sample is very heterogeneous. Thus, we also restrict the sample to OECD countries. As in the previous empirical application we estimate a fixed effects and a random effect model. We also estimate the slope parameter using the Pooled OLS estimator, the CCEP, CCEMG, the ALS and our estimators. The estimates and standard errors are reported in Table 8. For the dataset comprising all countries, the estimated coefficients vary between-.01322 (CCEP) and 0.4989 (Fixed Effects). For OECD countries the estimated coefficients varies between -0.0147 (ALS) and 0.8490 (Fixed effects). The different estimates reflect the high variability of the sample (see Figure 2). The estimates from the Pooled OLS and our estimator for the sample of OECD countries are very close to the coefficient 0.41 for non-transition countries reported in equation (8) of de Boeck and Slok (2006).

6. 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 both errors and the regressors. A consistent generalized method of moments estimator has been proposed for the case where the time dimension is fixed and N tends to infinity. This estimator exists for every $T \geq 1$ and is very simple to compute. As $N \to \infty$, it has a nonstandard asymptotic distribution which is covariance-matrix mixed normal with mixing density given by the unknown distribution of the shocks. However, tests on the interest parameters can be constructed using standard t- and F-procedures. A Monte Carlo simulation has shown that these estimators outperform the CCEP and CCEMG estimators of Pesaran (2006) and the ALS estimator of Ahn, Lee and Schmidt (2001) and Ahn, Lee and Schmidt (2013) when the time dimension is small. Two empirical applications have also been considered.

References

Ahn, S. C., Y. H. Lee and P. Schmidt (2001). "GMM Estimation of Linear Panel Data Models with Time-Varying Individual Effects." Journal of Econometrics **102**: 219-255.

Ahn, S. C., Y. H. Lee and P. Schmidt (2013). "Panel Data Models with Multiple Time-Varying Individual Effects." Journal of Econometrics **174**(1): 1-14.

Aldous, D. J. and G. K. Eagleson (1978). "On Mixing and Stability of Limit Therems." The Annals of Probability **6**(2): 325-331.

Andrews, D. W. K. (2005). "Cross-Section Regression with Common Shocks." Econometrica **73**(5): 1551-1585.

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

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

Chan, J. C. C., R. Leon-Gonzalez and R. W. Strachan (2013). Invariant Inference and Efficient Computation in the Static Factor Model. CAMA Working Papers.

Coakley, J., A. M. Fuertes and R. Smith (2006). "Unobserved Heterogeneity in Panel Time Series Models." Computational Statistics & Data Analysis **50**(9): 2361-2380.

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

de Boeck, M. and T. Slok (2006). "Interpreting real exchange rate movements in transition countries." Journal of International Economics **68**(2): 368-383.

Evans, D., A. Tandon, C. Murray and J. Lauer (2000). "The comparative efficiency of National health systems in Producing Health: an Analysis of 191 countries." World Health Organization, GPE discussion Paper **29**.
Geweke, J. and G. F. Zhou (1996). "Measuring the pricing error of the arbitrage pricing theory." Review of Financial Studies **9**(2): 557-587.
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. (2013). "Nonparametric Estimation in Large Panels with Cross Sectional Dependence." Econometric Reviews **32**(5): 754-777.
Kuersteiner, G. M. and I. R. Prucha (2013). "Limit theory for panel data models with cross sectional dependence and sequential exogeneity." Journal of Econometrics **174**: 107-126.
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 D. Sul (2003). "Dynamic Panel Estimation and Homogeneity Testing Under Cross Sectional Dependence." Econometrics Journal **6**: 217-259.
Prakasa Rao, B. L. S. (2009). "Conditional independence, conditional mixing and conditional association." Annals of the Institute of Statistical Mathematics **61**: 441-460.
Renyi, A. (1963). "On Stable Sequences of Events." Sankhya: The Indian Journal of Statistics, Series A **25**(3): 293-302.
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"." The 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 (2012). "Sieve Estimation of Panel Data Models with Cross Sectional Dependence." Journal of Econometrics **169**(1): 34-47.

Appendix: technical results

Let $\theta_0 = (\alpha_0', \beta_0')'$, $\hat{\theta} = (\hat{\alpha}', \hat{\beta}')'$, $\omega_i = (z_i, x_i)$ and $\overline{\omega} = (\overline{z}, \overline{x})$ to simplify the notation.

Proof of Theorem 1

From the definition and equation (1) we can write

(25) $$\hat{\theta} = \theta_0 + \left(\frac{1}{N} \sum_{i=1}^{N} \omega_i'(\omega_i - \overline{\omega}) \right)^{-1} \frac{1}{N} \sum_{i=1}^{N} \omega_i'(e_i - \overline{e}).$$

From Assumptions C.1 to C.5, conditional on $\mathcal{F}$, the γ_i's are independent with mean γ and are independent of v_j and Γ_j. Moreover, conditional on $\mathcal{F}$, ε_i is also independent of v_j and Γ_j. Thus, the expected value of the second term on the right hand side of (25) is zero such that $E\left[\hat{\theta} \mid \mathcal{F}\right] = \theta_0$. Therefore, $\hat{\theta}$ is unbiased conditional on $\mathcal{F}$ and it is thus unbiased unconditionally.

To show consistency, we write

$$\frac{1}{N}\sum_{i=1}^{N} \omega_i'(\omega_i - \overline{\omega}) = \frac{1}{N}\sum_{i=1}^{N} \omega_i'\omega_i - \overline{\omega}'\overline{\omega}.$$

We firstly show that conditional on $\mathcal{F}$, $\overline{\omega} - \frac{1}{N}\sum_{i=1}^{N} E[\omega_i \mid \mathcal{F}] \to 0$ a.s.. Since $\omega_i = (z_i, v_i)$, write

$$\begin{aligned}
&\overline{\omega} - \frac{1}{N}\sum_{i=1}^{N} E[\omega_i \mid \mathcal{F}] \\
&= \left(\frac{1}{N}\sum_{i=1}^{N}(z_i - E[z_i \mid \mathcal{F}]), \frac{1}{N}\sum_{i=1}^{N}(v_i - E[v_i \mid \mathcal{F}]) + F_T \frac{1}{N}\sum_{i=1}^{N}(\Gamma_i - E[\Gamma_i \mid \mathcal{F}])\right).
\end{aligned}$$

Assumption C.2 implies that the components of $\omega_i = (z_i, v_i)$ form sequences of independent random variables with finite means and satisfy the conditions for a conditional Markov's strong law of large number (e.g. Prakasa Rao (2009)), thus $\left(\frac{1}{N}\sum_{i=1}^{N}(z_i - E[z_i \mid \mathcal{F}]), \frac{1}{N}\sum_{i=1}^{N}(v_i - E[v_i \mid \mathcal{F}])\right) \to 0$ a.s.. Similarly we can conclude that $\frac{1}{N}\sum_{i=1}^{N}(\Gamma_i - E[\Gamma_i \mid \mathcal{F}]) \to 0$ a.s.. Thus, conditional on $\mathcal{F}$, $\overline{\omega} \to (\mu_{zN}, \tau_{vN} + F_T\Gamma)$ a.s..

We now focus on $\frac{1}{N}\sum_{i=1}^{N} \omega_i'\omega_i$. Each term in the sum is a $(p+k)\times(p+k)$ matrix. So let ζ_1 and ζ_2 be arbitrary $(p+k)\times 1$ vectors. Then $\frac{1}{N}\sum_{i=1}^{N} \zeta_1'\omega_i'\omega_i\zeta_2$ is a sum of independent random variables satisfying the following inequality a.s.:

$$\begin{aligned}
&E\left[\left|\zeta_1'\omega_i'\omega_i\zeta_2\right|^{1+\delta} \mid \mathcal{F}\right] \\
&\le (\zeta_1'\zeta_1)^{\frac{1+\delta}{2}} (\zeta_2'\zeta_2)^{\frac{1+\delta}{2}} E\left[\left\|\omega_i\right\|_2^{2+2\delta} \mid \mathcal{F}\right] \\
&\le (\zeta_1'\zeta_1)^{\frac{1+\delta}{2}} (\zeta_2'\zeta_2)^{\frac{1+\delta}{2}} E\left[\left(\left\|(z_i, v_i)\right\|_2 + \left\|F_T\right\|_2 \left\|\Gamma_i\right\|_2\right)^{2+2\delta} \mid \mathcal{F}\right] \\
&\le 2^{1+\delta} (\zeta_1'\zeta_1)^{\frac{1+\delta}{2}} (\zeta_2'\zeta_2)^{\frac{1+\delta}{2}} \left(E\left[\left\|(z_i, v_i)\right\|_2^{2+2\delta} \mid \mathcal{F}\right] + \left\|F_T\right\|_2^{2+2\delta} E\left[\left\|\Gamma_i\right\|_2^{2+2\delta} \mid \mathcal{F}\right]\right)
\end{aligned}$$

The last terms is uniformly bounded because of Assumption C.2 and C.4. Thus,

$$\frac{1}{N}\sum_{i=1}^{N} \omega_i'\omega_i \to \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} E[\omega_i'\omega_i \mid \mathcal{F}] \text{ a.s.}$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\omega_i'(\omega_i-\bar{\omega})\to\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}E[\omega_i'\omega_i\mid\mathcal{F}]-(\mu_z,\tau_v+F_T\Gamma)'(\mu_z,\tau_v+F_T\Gamma)\text{ a.s..}$$

To show that the right hand side is positive definite a.s., we notice that

$$\begin{aligned}E\left[\frac{1}{N}\sum_{i=1}^{N}\omega_i'(\omega_i-\bar{\omega})\mid\mathcal{F}\right]&=E\left[\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})\mid\mathcal{F}\right]\\&=E\left[\frac{1}{N}\sum_{i=1}^{N}(z_i-\bar{z},v_i-\bar{v})'(z_i-\bar{z},v_i-\bar{v})\mid\mathcal{F}\right]\\&\quad+\begin{pmatrix}0&0\\0&E\left[\frac{1}{N}\sum_{i=1}^{N}(\Gamma_i-\bar{\Gamma})'F_T'F_T(\Gamma_i-\bar{\Gamma})\mid\mathcal{F}\right]\end{pmatrix}.\end{aligned}$$

Since the first matrix is positive definite a.s. and the second is positive semi-definite uniformly in N a.s., their sum is positive definite uniformly in N a.s. and the limit is also positive definite a.s..

We now focus on $\frac{1}{N}\sum_{i=1}^{N}\omega_i'(e_i-\bar{e})=\frac{1}{N}\sum_{i=1}^{N}\omega_i'e_i-\bar{\omega}\bar{e}$. We already know that $\bar{\omega}\to(\mu_{zN},\tau_{vN}+F_T\Gamma)$ a.s. and with a similar argument we can show that $\bar{e}\to F_T\gamma$ a.s.. For the remaining term $\frac{1}{N}\sum_{i=1}^{N}\omega_i'e_i$, we notice that C.5 implies $E[\omega_i'e_i\mid\mathcal{F}]=E[\omega_i'\mid\mathcal{F}]E[e_i\mid\mathcal{F}]=\left(E[(z_i,v_i)'\mid\mathcal{F}]+(0,F_T\Gamma)'\right)F_T\gamma$, so that $\frac{1}{N}\sum_{i=1}^{N}\omega_i'e_i\to(\mu_z,\tau_v+F_T\Gamma)$ a.s. and consistency of the estimator follows.

Theorem 2

To prove conditional normality we write

(26) $$\sqrt{N}\left(\hat{\theta}-\theta_0\right)=\left(\frac{1}{N}\sum_{i=1}^{N}\omega_i'(\omega_i-\bar{\omega})\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\omega_i'(e_i-\bar{e}).$$

We know already that

$$\frac{1}{N}\sum_{i=1}^{N}\omega_i'(\omega_i-\bar{\omega})\to\frac{1}{N}\sum_{i=1}^{N}E[\omega_i'\omega_i\mid\mathcal{F}]-(\mu_z,\tau_v+F_T\Gamma)'(\mu_z,\tau_v+F_T\Gamma)\text{ a.s..}$$

We now focus on $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\omega_i'(e_i-\bar{e})$. Write

$$\begin{aligned}&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\omega_i'(e_i-\bar{e})\\&=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\omega_i'(e_i-F_T\gamma)-\sqrt{N}\bar{\omega}'(\bar{e}-F_T\gamma)\\&=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\omega_i-(\mu_{zN},\mu_{vN}+F_T\Gamma)\right)'(e_i-F_T\gamma)-\left(\bar{\omega}-(\mu_{zN},\mu_{vN}+F_T\Gamma)\right)'\sqrt{N}(\bar{e}-F_T\gamma).\end{aligned}$$

We will now show that the last term can be neglected. In fact, we know already that $\varpi - (\mu_{zN}, \mu_{vN} + F_T\Gamma) \to 0$ a.s.. Precisely we will prove that

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}\Sigma_{\varepsilon i} + F_T\left(\frac{1}{N}\sum_{i=1}^{N}\Sigma_{\gamma i}\right)F_T'\right)^{-1/2}(\overline{e} - F_T\gamma) \to^D N(0, I_T).$$

Let $\kappa_i = e_i - F_T\gamma$ and notice that they form a sequence of independent random variables conditional on $\mathcal{F}$. We can now use the Cramer-Wold device to find the distribution of $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\kappa_i$. Let ζ be an arbitrary $T \times 1$ vector and focus on $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\zeta'\kappa_i$. We will now verify the Liapounov's conditions for the validity of a conditional central limit theorem:

$$E[\zeta'\kappa_i \mid \mathcal{F}] = 0 \quad \text{and} \quad E\left[(\zeta'\kappa_i)^2 \mid \mathcal{F}\right] = \zeta'\left(\Sigma_{\varepsilon i} + F_T\Sigma_{\gamma i}F_T'\right)\zeta .$$

Notice also that

$$\begin{aligned}
E\left[|\zeta'\kappa_i|^{2+\delta} \mid \mathcal{F}\right] &= |\zeta'\zeta|^{\frac{2+\delta}{2}} \cdot E\left[\left\|\varepsilon_i + F_T(\gamma_i - \gamma)\right\|_2^{2+\delta} \mid \mathcal{F}\right] \\
&\le |\zeta'\zeta|^{\frac{2+\delta}{2}} E\left[\left(\|\varepsilon_i\|_2 + \|F_T(\gamma_i - \gamma)\|_2\right)^{2+\delta} \mid \mathcal{F}\right] \\
&\le |\zeta'\zeta|^{\frac{2+\delta}{2}} 2^{1+\delta}\left(E\left[\|\varepsilon_i\|_2^{2+\delta} \mid \mathcal{F}\right] + E\left[\|F_T(\gamma_i - \gamma)\|_2^{2+\delta} \mid \mathcal{F}\right]\right) \\
&\le |\zeta'\zeta|^{\frac{2+\delta}{2}} 2^{1+\delta}\left(E\left[\|\varepsilon_i\|_2^{2+\delta} \mid \mathcal{F}\right] + \|F_T\|_2^{2+\delta} E\left[\|\gamma_i - \gamma\|_2^{2+\delta} \mid \mathcal{F}\right]\right).
\end{aligned}$$

These conditions are satisfied so that

$$\frac{\sqrt{N}\zeta'(\overline{e} - F_T\gamma)}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\zeta'\left(\Sigma_{\varepsilon i} + F_T\Sigma_{\gamma i}F_T'\right)\zeta}} \to^D N(0,1) \tag{27}$$

Moreover, (27) converges stably in the sense of Renyi (1963) for all $F \in \mathcal{F}$. So we can conclude that $\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\Sigma_{\varepsilon i} + F_T\Sigma_{\gamma i}F_T'\right)\right)^{-1/2}(\overline{e} - F_T\gamma) \to^D N(0, I_T)$ and

$$\left(\varpi - (\mu_z, \mu_v + F_T\Gamma)\right)'\sqrt{N}(\overline{e} - F_T\gamma) \to 0 \text{ a.s..}$$

We now focus on the term $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\omega_i - (\mu_{zN}, \tau_{vN} + F_T\Gamma)\right)'(e_i - F_T\gamma)$. Let ζ be an arbitrary $(p+k)\times 1$ vector and $\varpi_i = \zeta'\left(\omega_i - (\mu_{zN}, \tau_{vN} + F_T\Gamma)\right)'$. Thus, we can write

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\zeta'\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\left(e_i-F_T\gamma\right)=\sum_{i=1}^{N}\frac{1}{\sqrt{N}}\varpi_i\kappa_i$$

and notice that $E\left[\varpi_i\kappa_i \mid \mathcal{F}\right]=0$.

We will establish that $E\left[\left(\varpi_i\kappa_i\right)^{2+\delta} \mid \mathcal{F}\right]<\infty$ is bounded uniformly. Notice that

$$E\left[\left(\varpi_i\kappa_i\right)^{2+\delta} \mid \mathcal{F}\right]\le E\left[\left\|\varpi_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]\cdot E\left[\left\|\kappa_i\right\|_2^{2+\delta} \mid \mathcal{F}\right].$$

We have already shown that $E\left[\left\|\kappa_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]$ is bounded so we need to verify $E\left[\left\|\varpi_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]$ is also bounded. Notice

$$\begin{aligned}E\left[\left\|\varpi_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]&=\left(\zeta'\zeta\right)^{\frac{2+\delta}{2}}E\left[\left\|\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right\|_2^{2+\delta} \mid \mathcal{F}\right]\\&\le\left(\zeta'\zeta\right)^{\frac{2+\delta}{2}}2^{1+\delta}\left(E\left[\left\|\omega_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]+\left\|\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right\|_2^{2+\delta}\right).\end{aligned}$$

$E\left[\left\|\omega_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]$ is uniformly bounded because of Assumptions C.2 and C.4. Moreover,

$$\left\|\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right\|_2\le\frac{1}{N}\sum_{i=1}^{N}\left\|E\left[\omega_i \mid \mathcal{F}\right]\right\|_2<\infty .$$

Now we have verified all the moment conditions required to obtain convergence in distribution conditional on $\mathcal{F}$

$$\frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\varpi_i\kappa_i}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}E\left[\left(\varpi_i\kappa_i\right)^2 \mid \mathcal{F}\right]}}\to^{D} N\left(0,1\right)$$

so that conditional on $\mathcal{F}$

(28)
$$\begin{aligned}&\left(\frac{1}{N}\sum_{i=1}^{N}E\left[\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\left(\Sigma_{\varepsilon i}+F_T\Sigma_{\gamma i}F_T'\right)\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right) \mid \mathcal{F}\right]\right)^{-\frac{1}{2}}\\&\times\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\left(e_i-F_T\gamma\right)\to^{D} N\left(0,I_{p+k}\right)\end{aligned}$$

Thus, (28) converges stably in the sense of Renyi (1963) for all $F\in\mathcal{F}$.

Proof of Corollary 1.

It follows from Theorem 1 that the sequence of random variables in (28) is stable. By Theorem 1 of Aldous and Eagleson (1978), (28) and $\left(B\left(F_T\right)\right)^{-1}\left(A\left(F_T\right)\right)^{1/2}$ converge stably in distribution to $\left(N\left(0,I_{p+k}\right),\left(B\left(F_T\right)\right)^{-1}\left(A\left(F_T\right)\right)^{1/2}\right)$. It follows that $\sqrt{N}\left(\hat{\theta}-\theta_0\right)\to X$.

Before proving Lemma 1 we prove some preliminary results.

Lemma A.1 Given Assumptions N.1, CM.2, N.3, CM.4 and C.5, the following results hold conditional on $\mathcal{F}$ as $N\to\infty$:

a) $\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(e_i-\bar{e})(e_i-\bar{e})'(\omega_i-\bar{\omega})\to A(F_T)$ a.s.;

b) $\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(e_i-\bar{e})\left(\hat{\theta}-\theta_0\right)'(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})\to 0$ a.s.;

c) $\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})\left(\hat{\theta}-\theta_0\right)\left(\hat{\theta}-\theta_0\right)'(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})\to 0$ a.s..

If also Assumption H.5 and HN.6 hold then

d) $\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})\left(\hat{\theta}-\theta_0\right)\eta_i'\omega_i'(\omega_i-\bar{\omega})\to 0$ a.s.;

e) $\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(e_i-\bar{e})\eta_i'\omega_i'(\omega_i-\bar{\omega})\to 0$ a.s.;

f) $\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})\eta_i\eta_i'(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})\to C(F_T)$ a.s..

Proof of Lemma A.1

Notice that Assumptions CM.2 and CM.4 imply Assumptions C.2 and C.4.

a) Write

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(e_i-\bar{e})(e_i-\bar{e})'(\omega_i-\bar{\omega})=&\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(e_i-F_T\gamma)(e_i-F_T\gamma)'(\omega_i-\bar{\omega}) && (a.1)\\ &-\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(e_i-F_T\gamma)(\bar{e}-F_T\gamma)'(\omega_i-\bar{\omega}) && (a.2)\\ &-\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\bar{e}-F_T\gamma)(e_i-F_T\gamma)'(\omega_i-\bar{\omega}) && (a.3)\\ &+\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\bar{e}-F_T\gamma)(\bar{e}-F_T\gamma)'(\omega_i-\bar{\omega}). && (a.4)\end{aligned}$$

Focus on (a.1) at first. In the proofs of Theorems 1 and 2, we have shown that conditional on $\mathcal{F}$

$$\frac{1}{N}\sum_{i=1}^{N}\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\left(e_i-F_T\gamma\right)\left(e_i-F_T\gamma\right)'\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)\to A\left(F_T\right)\text{ a.s.}$$

and

$$\bar{\omega}-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\to 0\text{ a.s..}$$

Similarly, we can prove that

$$\frac{1}{N}\sum_{i=1}^{N}\left(e_i-F_T\gamma\right)\left(e_i-F_T\gamma\right)'\to\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\Sigma_{\varepsilon i}+F_T\Sigma_{\gamma i}F_T{}'\right)\text{ a.s..}$$

and

$$\begin{aligned}&\frac{1}{N}\sum_{i=1}^{N}\left(e_i-F_T\gamma\right)\left(e_i-F_T\gamma\right)'\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)\\&\to\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\Sigma_{\varepsilon i}+F_T\Sigma_{\gamma i}F_T{}'\right)E\left[\omega_i\mid\mathcal{F}\right]-\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\Sigma_{\varepsilon i}+F_T\Sigma_{\gamma i}F_T{}'\right)\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\text{ a.s..}\end{aligned}$$

So conditional on $\mathcal{F}$

$$\frac{1}{N}\sum_{i=1}^{N}\left(\omega_i-\bar{\omega}\right)'\left(e_i-F_T\gamma\right)\left(e_i-F_T\gamma\right)'\left(\omega_i-\bar{\omega}\right)\to A\left(F_T\right)\text{ a.s..}$$

Now consider (a.2) and notice that (a.3) is just its transpose. The norm of this term is

$$\begin{aligned}&\left\|\frac{1}{N}\sum_{i=1}^{N}\left(\omega_i-\bar{\omega}\right)'\left(e_i-F_T\gamma\right)\left(\bar{e}-F_T\gamma\right)'\left(\omega_i-\bar{\omega}\right)\right\|_2\\&\le\left\|\bar{e}-F_T\gamma\right\|_2\cdot\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i-\bar{\omega}\right\|_2^2\left\|e_i-F_T\gamma\right\|_2\\&\le\left\|\bar{e}-F_T\gamma\right\|_2\cdot\frac{1}{N}\sum_{i=1}^{N}2\left(\left\|\omega_i\right\|_2^2+\left\|\bar{\omega}\right\|_2^2\right)\left\|e_i-F_T\gamma\right\|_2\\&=2\left\|\bar{e}-F_T\gamma\right\|_2\cdot\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i\right\|_2^2\left\|e_i-F_T\gamma\right\|_2+\left\|\bar{\omega}\right\|_2^2\frac{1}{N}\sum_{i=1}^{N}\left\|e_i-F_T\gamma\right\|_2\right).\end{aligned}$$

Since $E\left[\left\|e_i-F_T\gamma\right\|_2^{1+\delta}\mid\mathcal{F}\right]$ and $E\left[\left\|\omega_i\right\|_2^{2(1+\delta)}\left\|e_i-F_T\gamma\right\|_2^{1+\delta}\mid\mathcal{F}\right]$ are bounded the last term converges a.s. to zero so that $\frac{1}{N}\sum_{i=1}^{N}\left(\omega_i-\bar{\omega}\right)'\left(e_i-F_T\gamma\right)\left(\bar{e}-F_T\gamma\right)'\left(\omega_i-\bar{\omega}\right)\to 0$ a.s..

Similarly, by using the triangle inequality and sub-multiplicativity one obtains for (a.4)

$$\begin{aligned}\left\|\frac{1}{N}\sum_{i=1}^{N}\left(\omega_i-\bar{\omega}\right)'\left(\bar{e}-F_T\gamma\right)\left(\bar{e}-F_T\gamma\right)'\left(\omega_i-\bar{\omega}\right)\right\|_2&\le\left\|\bar{e}-F_T\gamma\right\|_2^2\cdot\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i-\bar{\omega}\right\|_2^2\\&\le\left\|\bar{e}-F_T\gamma\right\|_2^2\cdot 2\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i\right\|_2^2+\left\|\bar{\omega}\right\|_2^2\right).\end{aligned}$$

Based on the above, the first result follows.

b) Write

$$\begin{aligned}
&\left\|\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\overline{\omega})'(e_i-\overline{e})\left(\hat{\theta}-\theta\right)'(\omega_i-\overline{\omega})'(\omega_i-\overline{\omega})\right\|_2\\
&\le\frac{1}{N}\sum_{i=1}^{N}\left\|(\omega_i-\overline{\omega})'(e_i-\overline{e})\left(\hat{\theta}-\theta\right)'(\omega_i-\overline{\omega})'(\omega_i-\overline{\omega})\right\|_2\\
&\le\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i-\overline{\omega}\right\|_2^{\,3}\cdot\left\|e_i-\overline{e}\right\|_2\cdot\left\|\hat{\theta}-\theta\right\|_2\\
&\le\left\|\hat{\theta}-\theta\right\|_2\cdot\frac{1}{N}\sum_{i=1}^{N}\left(\left\|\omega_i\right\|_2+\left\|\overline{\omega}\right\|_2\right)^3\left(\left\|e_i\right\|_2+\left\|\overline{e}\right\|_2\right)\\
&\le\left\|\hat{\theta}-\theta\right\|_2\cdot\frac{1}{N}\sum_{i=1}^{N}2^2\left(\left\|\omega_i\right\|_2^{\,3}+\left\|\overline{\omega}\right\|_2^{\,3}\right)\left(\left\|e_i\right\|_2+\left\|\overline{e}\right\|_2\right).
\end{aligned}$$

Above we have applied the triangle inequality for the first inequality, sub-multiplicativity in the second and the triangle inequality in the third and the c_r inequality in the fourth one. For the last sum to converge we need $E\left[\left(\left\|\omega_i\right\|_2^{\,3}\right)^{1+\delta}\mid\mathcal{F}\right]$ and $E\left[\left\|e_i\right\|_2^{\,1+\delta}\mid\mathcal{F}\right]$ to be bounded. This follows from the assumptions of this lemma. Thus this term goes to zero a.s. because $\hat{\theta}$ is consistent a.s..

c) Write

$$\begin{aligned}
\left\|\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\overline{\omega})'(\omega_i-\overline{\omega})\left(\hat{\theta}-\theta_0\right)\left(\hat{\theta}-\theta_0\right)'(\omega_i-\overline{\omega})'(\omega_i-\overline{\omega})\right\|_2 &\le\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i-\overline{\omega}\right\|_2^{\,4}\left\|\hat{\theta}-\theta\right\|_2^{\,2}\\
&\le\left\|\hat{\theta}-\theta\right\|_2^{\,2}\frac{1}{N}\sum_{i=1}^{N}\left(\left\|\omega_i\right\|_2+\left\|\overline{\omega}\right\|_2\right)^4\\
&\le\left\|\hat{\theta}-\theta\right\|_2^{\,2}2^3\frac{1}{N}\sum_{i=1}^{N}\left(\left\|\omega_i\right\|_2^{\,4}+\left\|\overline{\omega}\right\|_2^{\,4}\right).
\end{aligned}$$

Since $E\left[\left(\left\|\omega_i\right\|_2^{\,4}\right)^{1+\delta}\mid\mathcal{F}\right]$ is uniformly bounded the series $\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i\right\|_2^{\,4}$ converges to a finite quantity a.s.. Thus, the result follows because $\hat{\theta}$ is consistent a.s..

d) Write

$$\begin{aligned}
&\left\|\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\overline{\omega})'(\omega_i-\overline{\omega})\left(\hat{\theta}-\theta_0\right)\eta_i{}'\omega_i{}'(\omega_i-\overline{\omega})\right\|_2\\
&\le\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i-\overline{\omega}\right\|_2^{\,3}\left\|\omega_i\right\|_2\left\|\hat{\theta}-\theta\right\|_2^{\,2}\left\|\eta_i\right\|_2\le\left\|\hat{\theta}-\theta\right\|_2^{\,2}\frac{1}{N}\sum_{i=1}^{N}\left(\left\|\omega_i\right\|_2+\left\|\overline{\omega}\right\|_2\right)^3\left\|\omega_i\right\|_2\left\|\eta_i\right\|_2\\
&\le\left\|\hat{\theta}-\theta\right\|_2^{\,2}2^2\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i\right\|_2^{\,4}\left\|\eta_i\right\|_2+\left\|\overline{\omega}\right\|_2^{\,4}\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i\right\|_2\left\|\eta_i\right\|_2\right)
\end{aligned}$$

Since $E\left[\left\|\omega_i\right\|_2^{4+4\delta} \mid \mathcal{F}\right]$ and $E\left[\left\|\eta_i\right\|_2^{1+\delta} \mid \mathcal{F}\right]$ are uniformly bounded, the series $\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i\right\|_2^4\left\|\eta_i\right\|_2$ and $\frac{1}{N}\sum_{i=1}^{N}\left\|\omega_i\right\|_2\left\|\eta_i\right\|_2$ converge to finite quantities a.s.. Thus, the result follows because $\hat{\theta}$ is consistent a.s..

e) As before, write

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(e_i-\bar{e})\eta_i'\omega_i'(\omega_i-\bar{\omega}) &= \frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(e_i-F_T\gamma)\eta_i'\omega_i'(\omega_i-\bar{\omega}) \\ &\quad -\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\bar{e}-F_T\gamma)\eta_i'\omega_i'(\omega_i-\bar{\omega}).\end{aligned}$$

The term in the second row converges to zero because it satisfies

$$\begin{aligned}\left\|\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\bar{e}-F_T\gamma)\eta_i'\omega_i'(\omega_i-\bar{\omega})\right\|_2 &\le \frac{1}{N}\sum_{i=1}^{N}\left\|(\omega_i-\bar{\omega})'(\bar{e}-F_T\gamma)\eta_i'\omega_i'(\omega_i-\bar{\omega})\right\|_2 \\ &\le \left\|\bar{e}-F_T\gamma\right\|_2\frac{1}{N}\sum_{i=1}^{N}\left\|(\omega_i-\bar{\omega})\right\|_2^2\left\|\omega_i\right\|_2\left\|\eta_i\right\|_2 \\ &\le \left\|\bar{e}-F_T\gamma\right\|_2\frac{1}{N}\sum_{i=1}^{N}\left(\left\|\omega_i\right\|_2+\left\|\bar{\omega}\right\|_2\right)^2\left\|\omega_i\right\|_2\left\|\eta_i\right\|_2 \\ &\le \left\|\bar{e}-F_T\gamma\right\|_2\frac{2}{N}\left(\sum_{i=1}^{N}\left\|\omega_i\right\|_2^3\left\|\eta_i\right\|_2+\left\|\bar{\omega}\right\|_2^2\sum_{i=1}^{N}\left\|\omega_i\right\|_2\left\|\eta_i\right\|_2\right).\end{aligned}$$

It is easy to check that the sums in the last term converge to finite quantities. Since $\bar{e}-F_T\gamma\to 0$ a.s., then $\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\bar{e}-F_T\gamma)\eta_i'\omega_i'(\omega_i-\bar{\omega})\to 0$ a.s..

Similarly, $\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(e_i-F_T\gamma)\eta_i'\omega_i'(\omega_i-\bar{\omega})\to 0$ a.s.. Therefore, the result follows.

f) Now we consider the term $\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})\eta_i\eta_i'(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})$. The result follows by proceeding as a).

Proof of Lemma 1.

Write

$$\begin{aligned}\hat{A}(F_T) &= \frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'\left(e_i-\bar{e}-(\omega_i-\bar{\omega})\left(\hat{\theta}-\theta_0\right)\right)\left(e_i-\bar{e}-(\omega_i-\bar{\omega})\left(\hat{\theta}-\theta_0\right)\right)'(\omega_i-\bar{\omega}) \\ &= \frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(e_i-\bar{e})(e_i-\bar{e})'(\omega_i-\bar{\omega}) \\ &\quad -\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})\left(\hat{\theta}-\theta_0\right)(e_i-\bar{e})'(\omega_i-\bar{\omega}) \\ &\quad -\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(e_i-\bar{e})\left(\hat{\theta}-\theta_0\right)'(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega}) \\ &\quad +\frac{1}{N}\sum_{i=1}^{N}(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})\left(\hat{\theta}-\theta_0\right)\left(\hat{\theta}-\theta_0\right)'(\omega_i-\bar{\omega})'(\omega_i-\bar{\omega})\end{aligned}$$

Then, the lemma follows from Lemma A.1 a), b) and c).

Proof of Theorem 3

Write

$$\hat{\theta}=\theta_0+\left(\frac{1}{N}\sum_{i=1}^{N}\omega_i'(\omega_i-\bar{\omega})\right)^{-1}\frac{1}{N}\sum_{i=1}^{N}\omega_i'\left(F_T\gamma_i-F_T\bar{\gamma}+\varepsilon_i-\bar{\varepsilon}\right).$$

Under assumption D.5 the factor loadings are not independent and this affects only the term

$$\frac{1}{N}\sum_{i=1}^{N}\omega_i'F_T\gamma_i=\frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix} z_i' \\ v_i'+\Gamma_i'F_T' \end{pmatrix}F_T\gamma_i=\frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix} z_i' \\ v_i' \end{pmatrix}F_T\gamma_i+\begin{pmatrix} 0 \\ \frac{1}{N}\sum_{i=1}^{N}\Gamma_i'F_T'F_T\gamma_i \end{pmatrix}.$$

Let ζ be an arbitrary $(p+k)\times 1$ vector and consider $\frac{1}{N}\sum_{i=1}^{N}\zeta'\begin{pmatrix} z_i' \\ v_i' \end{pmatrix}F_T\gamma_i$. Then,

$$E\left[\zeta'\begin{pmatrix} z_i' \\ v_i' \end{pmatrix}F_T\gamma_i \mid \mathcal{F}\right]=\zeta'E\left[\begin{pmatrix} z_i' \\ v_i' \end{pmatrix}\mid \mathcal{F}\right]F_T\gamma$$

and

$$\begin{aligned}E\left[\left(\zeta'\begin{pmatrix} z_i' \\ v_i' \end{pmatrix}F_T\gamma_i\right)^{1+\delta}\mid \mathcal{F}\right] &\leq \left|\zeta'\zeta\right|^{\frac{1+\delta}{2}}E\left[\left\|(z_i,v_i)\right\|_2^{1+\delta}\left\|F_T\gamma_i\right\|_2^{1+\delta}\mid \mathcal{F}\right] \\ &\leq \left|\zeta'\zeta\right|^{\frac{1+\delta}{2}}E\left[\left\|(z_i,v_i)\right\|_2^{1+\delta}\mid \mathcal{F}\right]\cdot E\left[\left\|F_T\gamma_i\right\|_2^{1+\delta}\mid \mathcal{F}\right]\end{aligned}$$

and this term is uniformly bounded by Assumption C.2 and C.3. Thus we can conclude that $\frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix} z_i' \\ v_i' \end{pmatrix}F_T\gamma_i\to\begin{pmatrix} \mu_{zN}'F_T\gamma \\ \tau_{vN}'F_T\gamma \end{pmatrix}$ a.s. conditional on $\mathcal{F}$. Similarly, let ζ be an arbitrary $p\times 1$ vector and consider $\frac{1}{N}\sum_{i=1}^{N}\zeta'\Gamma_i'F_T'F_T\gamma_i$. Since $\left|\zeta'\Gamma_i'F_T'F_T\gamma_i\right| \le \left|\zeta'\Gamma_i'F_T'F_T\Gamma_i\zeta\right|^{\frac{1}{2}}\left|\gamma_i'F_T'F_T\gamma_i\right|^{\frac{1}{2}}$ by the Cauchy-Schwartz inequality and $\left|\zeta'\Gamma_i'F_T'F_T\Gamma_i\zeta\right|^{\frac{1}{2}} \le \left|\zeta'\zeta\right|^{1/2}\left\|F_T\right\|_2\left\|\Gamma_i\right\|_2$ so that

$$E\left[\left|\zeta'\Gamma_i'F_T'F_T\gamma_i\right|^{1+\delta} \mid \mathcal{F}\right] \le \left|\zeta'\zeta\right|^{\frac{1+\delta}{2}}\left\|F_T\right\|_2^{1+\delta}\cdot E\left[\left\|\Gamma_i\right\|_2^{1+\delta} \mid \mathcal{F}\right]\cdot\left\|F_T\right\|_2^{1+\delta}\cdot E\left[\left\|\gamma_i\right\|_2^{1+\delta} \mid \mathcal{F}\right],$$

which is uniformly bounded. Thus, we can conclude that

$$\frac{1}{N}\sum_{i=1}^{N}\Gamma_i'F_T'F_T\gamma_i \to \frac{1}{N}\sum_{i=1}^{N}E\left[\Gamma_i'F_T'F_T\gamma_i \mid \mathcal{F}\right] \text{ a.s..}$$

Similarly we can show that

$$\left(\frac{1}{N}\sum_{i=1}^{N}\omega_i'\right)F_T\overline{\gamma} = \overline{\omega}'F_T\overline{\gamma} \to \begin{pmatrix} \mu_{zN}'F_T\gamma \\ \tau_{vN}'F_T\gamma + \Gamma'F_T'F_T\gamma \end{pmatrix} \text{ a.s.}$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\omega_i'\varepsilon_i - \overline{\omega}'\overline{\varepsilon} \to 0 \text{ a.s..}$$

Thus, conditional on $\mathcal{F}$ $\hat{\theta} \to \theta_0 + B\left(F_T\right)^{-1}\begin{pmatrix} 0 \\ \Delta\left(F_T\right) \end{pmatrix}$ a.s. so that $\hat{\theta}$ is biased conditional on $\mathcal{F}$.

Proof of Theorem 4

Let $\eta_i = \begin{pmatrix} \eta_{1i} \\ \eta_{2i} \end{pmatrix}$. Write

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} + \left[\sum_{i=1}^{N}\omega_i'\left(\omega_i - \overline{\omega}\right)\right]^{-1}\sum_{i=1}^{N}\omega_i'\left(e_i - \overline{e} - \omega_i\eta_i + \frac{1}{N}\sum_{j=1}^{N}\omega_j\eta_j\right)$$

Unbiasedness follows from the independence of x_i and e_i and x_i and η_i conditional on $\mathcal{F}$.

We now consider consistency. In Theorem 1, we have shown that $\frac{1}{N}\sum_{i=1}^{N}\omega_i'\left(e_i - \overline{e}\right) \to 0$ a.s.. For the rest error terms, we just need to focus on $\frac{1}{N}\sum_{i=1}^{N}\omega_i'\omega_i\eta_i$ and $\frac{1}{N}\sum_{i=1}^{N}\omega_i\eta_i$. Firstly let ζ be an arbitrary $\left(p+k\right)\times 1$ vector and write

$$
\begin{aligned}
E\left[\left|\zeta'\omega_i'\omega_i\eta_i\right|^{1+\delta} \mid \mathcal{F}\right] &\le \left(\zeta'\zeta\right)^{\frac{1+\delta}{2}} E\left[\left\|\omega_i'\omega_i\eta_i\right\|_2^{1+\delta} \mid \mathcal{F}\right] \\
&\le \left(\zeta'\zeta\right)^{\frac{1+\delta}{2}} E\left[\left\|\omega_i\right\|_2^{2+2\delta}\left\|\eta_i\right\|_2^{1+\delta} \mid \mathcal{F}\right] \\
&\le \left(\zeta'\zeta\right)^{\frac{1+\delta}{2}} E\left[\left\|\omega_i\right\|_2^{2+2\delta} \mid \mathcal{F}\right]\cdot E\left[\left\|\eta_i\right\|_2^{1+\delta} \mid \mathcal{F}\right].
\end{aligned}
$$

Each term on RHS above is bounded. Thus, $\frac{1}{N}\sum_{i=1}^{N}\omega_i'\omega_i\eta_i \to 0$ a.s.. Similarly, $\frac{1}{N}\sum_{i=1}^{N}\omega_i\eta_i \to 0$ a.s.. Hence, the consistency follows.

For asymptotic normality we write

$$
\sqrt{N}\left(\hat{\theta}-\theta_0\right)=\left[\frac{1}{N}\sum_{i=1}^{N}\omega_i'\left(\omega_i-\overline{\omega}\right)\right]^{-1}\frac{1}{\sqrt{N}}\left(\sum_{i=1}^{N}\omega_i'\left(e_i-\overline{e}\right)+\sum_{i=1}^{N}\left(\omega_i-\overline{\omega}\right)'\omega_i\eta_i\right).
$$

Notice that we have shown that $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\omega_i'\left(e_i-\overline{e}\right)$ converges to a normal distribution in Theorem 1. $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\omega_i'\left(e_i-\overline{e}\right)$ and $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\omega_i-\overline{\omega}\right)'\omega_i\eta_i$ are conditionally uncorrelated, so we only need to focus on $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\omega_i-\overline{\omega}\right)'\omega_i\eta_i$. Write

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\omega_i-\overline{\omega}\right)'\omega_i\eta_i \\
&=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\omega_i\eta_i-\left(\overline{\omega}-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\omega_i\eta_i.
\end{aligned}
$$

We know that $\overline{\omega}-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\to 0$ a.s.. Let ζ be an arbitrary $T\times 1$ vector and notice that $E\left[\zeta'\omega_i\eta_i \mid \mathcal{F}\right]=0$,

$$
E\left[\left|\zeta'\omega_i\eta_i\right|^{2+\delta} \mid \mathcal{F}\right]\le\left|\zeta'\zeta\right|^{\frac{2+\delta}{2}} E\left[\left\|\omega_i\eta_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]\le\left|\zeta'\zeta\right|^{\frac{2+\delta}{2}} E\left[\left\|\omega_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]E\left[\left\|\eta_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]
$$

and this is uniformly bounded. Thus, this satisfies Liapounov's condition conditional on $\mathcal{F}$ and $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\omega_i\eta_i$ converges to a normal distribution stably in the sense of Renyi (1963) for all $F\in\mathcal{F}$, so that

$$
\left(\overline{\omega}-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\omega_i\eta_i\to 0 \text{ a.s..}
$$

For the remaining term, we verify that for any arbitrary $p\times 1$ vector ζ the summation $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\zeta'\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\omega_i\eta_i$ satisfies Liapounov's conditions conditional on $\mathcal{F}$:

$$E\left[\zeta'\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\omega_i\eta_i \mid \mathcal{F}\right]=0$$

and

$$\begin{aligned}
&E\left[\left(\zeta'\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\omega_i\eta_i\right)^{2+\delta} \mid \mathcal{F}\right]\\
&\le\left|\zeta'\zeta\right|^{\frac{2+\delta}{2}}E\left[\left\|\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\omega_i\eta_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]\\
&\le\left|\zeta'\zeta\right|^{\frac{2+\delta}{2}}E\left[\left\|\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\omega_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]\cdot E\left[\left\|\eta_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]\\
&\le\left|\zeta'\zeta\right|^{\frac{2+\delta}{2}}2^{\frac{2+\delta}{2}}\left(E\left[\left\|\omega_i\right\|_2^{4+2\delta} \mid \mathcal{F}\right]+E\left[\left\|\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)'\omega_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]\right)\cdot E\left[\left\|\eta_i\right\|_2^{2+\delta} \mid \mathcal{F}\right].
\end{aligned}$$

In the expression above, $E\left[\left\|\omega_i\right\|_2^{4+2\delta} \mid \mathcal{F}\right]$ is uniformly bounded because of Assumptions CM.2 and CM.4, $E\left[\left\|\eta_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]$ is uniformly bounded because of Assumptions HN.6. Finally,

$$E\left[\left\|\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)'\omega_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]\le\left\|\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right\|_2^{2+\delta}E\left[\left\|\omega_i\right\|_2^{2+\delta} \mid \mathcal{F}\right]<\infty$$

because of assumptions CM.2 and CM.4.

Thus,

$$\frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\zeta'\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\omega_i\eta_i}{\sqrt{\zeta'\left(\frac{1}{N}\sum_{i=1}^{N}E\left[\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right)'\omega_i\Sigma_\eta\omega_i\left(\omega_i-\left(\mu_{zN},\tau_{vN}+F_T\Gamma\right)\right) \mid \mathcal{F}\right]\right)\zeta}}\to N(0,1),$$

so that we have

$$\left(A(F_T)+C(F_T)\right)^{-1/2}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\omega_i'\left(e_i-\bar{e}+\omega_i\eta_i-\frac{1}{N}\sum_{j=1}^{N}\omega_j\eta_j\right)\to^{D}N\left(0,I_{p+k}\right). \tag{29}$$

Thus, (29) converges stably in the sense of Renyi (1963) for all $F\in\mathcal{F}$.

Proof of Corollary 3.

It follows from Theorem 1 that the sequence of random variables in (29) is stable. By Theorem 1 of Aldous and Eagleson (1978), (29) and $\left(B(F_T)\right)^{-1}\left(A(F_T)+C(F_T)\right)^{1/2}$ converge stably in distribution to $\left(N\left(0,I_{p+k}\right),\left(B(F_T)\right)^{-1}\left(A(F_T)+C(F_T)\right)^{1/2}\right)$.

Therefore, it follows that $\sqrt{N}\left(\hat{\theta}-\theta_0\right)\to X$.

Proof of Lemma 2. Write

$$\hat{A}(F_T) = \frac{1}{N}\sum_{i=1}^{N}(\omega_i - \overline{\omega})'\Big((\omega_i - \overline{\omega})\Big(\theta_0 - \hat{\theta}\Big) + (e_i - \overline{e})\Big)\cdot\Big((\omega_i - \overline{\omega})\Big(\theta_0 - \hat{\theta}\Big) + (e_i - \overline{e})\Big)'(\omega_i - \overline{\omega})$$
$$+\frac{1}{N}\sum_{i=1}^{N}(\omega_i - \overline{\omega})'\Big((\omega_i - \overline{\omega})\Big(\theta_0 - \hat{\theta}\Big) + (e_i - \overline{e})\Big)\eta_i\,'(\omega_i - \overline{\omega})'(\omega_i - \overline{\omega})$$
$$+\frac{1}{N}\sum_{i=1}^{N}(\omega_i - \overline{\omega})'(\omega_i - \overline{\omega})\eta_i\Big((\omega_i - \overline{\omega})\Big(\theta_0 - \hat{\theta}\Big) + (e_i - \overline{e})\Big)'(\omega_i - \overline{\omega})$$
$$+\frac{1}{N}\sum_{i=1}^{N}(\omega_i - \overline{\omega})'(\omega_i - \overline{\omega})\eta_i\eta_i\,'(\omega_i - \overline{\omega})'(\omega_i - \overline{\omega}).$$

The first line has been proved in Lemma 1. The others lines are proved in Lemma A.1.

Tables

Table 1

Small Sample Properties of the GMM, CCEP, CCEMG, ALS and OLS estimators

Homogeneous slopes without fixed effects (one factor)

$\hat{\beta}_1$		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
GMM	N=20	0.001	-0.001	0.000	-0.001	0.000	0.019	0.016	0.012	0.009	0.006	0.117	0.120	0.113	0.106	0.098	0.140	0.147	0.152	0.160	0.178
	50	-0.001	-0.001	0.000	0.000	0.001	0.007	0.006	0.004	0.003	0.002	0.079	0.076	0.070	0.070	0.067	0.131	0.144	0.160	0.200	0.231
	100	0.000	0.000	0.001	-0.001	0.000	0.003	0.003	0.002	0.002	0.001	0.065	0.069	0.064	0.065	0.059	0.167	0.196	0.239	0.289	0.365
	200	0.000	0.000	0.000	-0.001	0.000	0.002	0.001	0.001	0.001	0.001	0.059	0.061	0.055	0.057	0.057	0.256	0.304	0.377	0.482	0.578
CCEP	N=20	NA	0.075	-0.002	0.001	0.001	NA	47.679	0.020	0.006	0.003	NA	0.086	0.010	0.069	0.076	NA	0.100	0.013	0.123	0.202
	50	0.101	0.217	0.001	0.000	0.000	51.791	61.402	0.007	0.002	0.001	0.096	0.134	0.007	0.056	0.057	0.114	0.158	0.013	0.193	0.356
	100	0.121	0.026	0.000	0.000	0.000	36.997	89.512	0.004	0.001	0.001	0.140	0.202	0.005	0.051	0.054	0.164	0.233	0.015	0.326	0.604
	200	0.116	-0.067	0.000	0.000	0.000	22.815	164.136	0.002	0.001	0.000	0.190	0.280	0.005	0.055	0.049	0.225	0.315	0.025	0.569	0.874
CCEMG	N=20	NA	0.060	2.267	0.000	0.001	NA	2629.068	86801.903	0.010	0.003	NA	0.050	0.028	0.064	0.064	NA	0.058	0.030	0.103	0.172
	50	NA	-1.609	2.435	0.000	0.000	NA	16350.827	24803.917	0.004	0.001	NA	0.047	0.026	0.057	0.055	NA	0.055	0.027	0.137	0.297
	100	NA	-1.209	-5.726	0.000	0.000	NA	2500.712	321242.695	0.002	0.001	NA	0.052	0.022	0.050	0.049	NA	0.060	0.023	0.209	0.517
	200	NA	-3.118	1.703	0.001	0.000	NA	92326.136	16096.756	0.001	0.000	NA	0.059	0.020	0.056	0.051	NA	0.067	0.021	0.359	0.797
OLS	N=20	0.105	0.110	0.119	0.120	0.127	0.060	0.057	0.054	0.052	0.051										
	50	0.103	0.108	0.115	0.123	0.125	0.043	0.043	0.042	0.043	0.044										
	100	0.103	0.108	0.117	0.124	0.128	0.038	0.039	0.040	0.041	0.042										
	200	0.102	0.110	0.115	0.120	0.125	0.035	0.036	0.037	0.039	0.040										
ALS	N=20	0.059	0.070	0.082	0.100	0.143	0.082	0.073	0.066	0.066	0.078										
	50	0.061	0.070	0.082	0.110	0.155	0.047	0.045	0.046	0.054	0.076										
	100	0.063	0.073	0.084	0.116	0.163	0.036	0.039	0.043	0.053	0.075										
	200	0.061	0.075	0.088	0.120	0.162	0.030	0.036	0.041	0.052	0.074										

Notes: $\sigma_{\eta}^2 = 0.4$, $\rho_{f_1} = 0.5$

Table 2

Small Sample Properties of the GMM, CCEP, CCEMG, ALS and OLS estimators

Heterogeneous slopes without fixed effects (one factor)

$\hat{\beta}_1$		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
GMM	N=20	-0.003	0.000	-0.002	-0.002	0.001	0.094	0.084	0.072	0.059	0.051	0.155	0.154	0.146	0.138	0.129	0.160	0.159	0.151	0.145	0.135
	50	0.000	0.001	0.001	-0.001	-0.002	0.041	0.036	0.030	0.025	0.021	0.101	0.099	0.091	0.085	0.085	0.113	0.112	0.105	0.104	0.103
	100	0.001	-0.001	0.003	-0.001	0.001	0.021	0.018	0.015	0.013	0.011	0.081	0.072	0.069	0.069	0.069	0.096	0.097	0.098	0.097	0.106
	200	0.001	0.000	-0.001	0.000	-0.001	0.011	0.009	0.008	0.007	0.006	0.065	0.064	0.067	0.059	0.060	0.103	0.099	0.109	0.117	0.120
CCEP	N=20	NA	0.048	-0.001	0.001	0.000	NA	136.254	0.075	0.043	0.037	NA	0.081	0.025	0.081	0.088	NA	0.090	0.026	0.092	0.099
	50	NA	0.222	0.000	0.001	-0.001	NA	57.014	0.032	0.020	0.017	NA	0.101	0.020	0.060	0.066	NA	0.116	0.023	0.081	0.093
	100	-0.073	0.185	0.002	-0.002	0.001	361.632	14.270	0.016	0.010	0.009	0.108	0.146	0.015	0.056	0.061	0.128	0.174	0.021	0.088	0.100
	200	NA	-1.210	0.000	0.001	0.000	NA	15957.391	0.008	0.006	0.005	NA	0.214	0.011	0.050	0.052	NA	0.248	0.023	0.121	0.131
CCEMG	N=20	NA	-0.097	1.596	-0.001	0.000	NA	7452.975	20639.163	0.031	0.023	NA	0.046	0.034	0.068	0.075	NA	0.053	0.034	0.080	0.089
	50	NA	6.935	0.502	0.001	0.000	NA	682825.900	14686.696	0.012	0.009	NA	0.053	0.024	0.060	0.057	NA	0.059	0.024	0.083	0.092
	100	NA	0.487	-0.820	-0.001	0.000	NA	3502.847	2003.974	0.006	0.005	NA	0.053	0.024	0.053	0.054	NA	0.062	0.025	0.099	0.123
	200	NA	-2.295	0.476	0.001	0.000	NA	24237.121	3298.454	0.003	0.002	NA	0.051	0.022	0.048	0.053	NA	0.062	0.024	0.152	0.186
OLS	N=20	0.104	0.109	0.116	0.124	0.125	0.124	0.117	0.109	0.097	0.091										
	50	0.105	0.111	0.119	0.123	0.123	0.073	0.069	0.067	0.062	0.060										
	100	0.105	0.109	0.120	0.122	0.128	0.054	0.053	0.052	0.052	0.051										
	200	0.104	0.108	0.115	0.122	0.122	0.044	0.043	0.044	0.044	0.044										
ALS	N=20	0.046	0.071	0.078	0.105	0.142	0.183	0.171	0.159	0.152	0.155										
	50	0.060	0.074	0.089	0.110	0.152	0.091	0.091	0.093	0.096	0.109										
	100	0.060	0.074	0.088	0.115	0.164	0.060	0.064	0.066	0.076	0.095										
	200	0.063	0.076	0.088	0.119	0.160	0.045	0.048	0.053	0.063	0.083										

Notes: $\sigma_{\eta}^2 = 0.4$, $\rho_{f_1} = 0.5$

Table 3

Small Sample Properties of the GMM, CCEP, CCEMG, ALS and OLS estimators

Homogeneous slopes without fixed effects (several factors)

$\hat{\beta}_1$		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
GMM	N=20	0.001	0.000	-0.001	0.000	-0.001	0.018	0.014	0.010	0.007	0.005	0.115	0.113	0.111	0.105	0.088	0.146	0.148	0.153	0.169	0.177
	50	-0.001	0.000	0.004	0.002	0.000	0.006	0.005	0.004	0.003	0.002	0.077	0.074	0.067	0.068	0.066	0.137	0.153	0.180	0.226	0.253
	100	0.000	0.000	0.000	0.000	0.000	0.003	0.002	0.002	0.001	0.001	0.062	0.064	0.076	0.058	0.060	0.184	0.208	0.262	0.326	0.408
	200	0.000	0.000	0.001	0.000	0.000	0.002	0.001	0.001	0.001	0.000	0.057	0.059	0.049	0.053	0.052	0.283	0.337	0.424	0.547	0.659
CCEP	N=20	NA	0.515	0.003	0.001	0.001	NA	642.606	0.019	0.007	0.004	NA	0.123	0.011	0.070	0.072	NA	0.140	0.013	0.122	0.185
	50	NA	0.397	0.002	0.000	0.001	NA	263.666	0.007	0.002	0.001	NA	0.199	0.006	0.058	0.063	NA	0.226	0.013	0.188	0.336
	100	NA	0.103	0.001	-0.001	0.000	NA	79.622	0.003	0.001	0.001	NA	0.274	0.006	0.052	0.058	NA	0.305	0.017	0.314	0.554
	200	NA	0.505	0.000	0.000	0.000	NA	892.562	0.002	0.001	0.000	NA	0.363	0.006	0.056	0.051	NA	0.396	0.028	0.565	0.836
CCEMG	N=20	NA	0.320	-1.904	0.001	0.001	NA	8258.417	14540.520	0.011	0.004	NA	0.068	0.032	0.066	0.066	NA	0.079	0.032	0.103	0.161
	50	NA	-3.395	0.096	0.000	0.001	NA	131127.428	745.211	0.004	0.001	NA	0.070	0.022	0.053	0.057	NA	0.081	0.024	0.138	0.292
	100	NA	-0.298	0.010	-0.001	0.000	NA	10827.129	1506.229	0.002	0.001	NA	0.074	0.025	0.051	0.054	NA	0.085	0.025	0.204	0.492
	200	NA	1.280	1.070	0.000	0.000	NA	16580.928	34000.783	0.001	0.000	NA	0.087	0.023	0.052	0.053	NA	0.100	0.023	0.349	0.770
OLS	N=20	0.121	0.126	0.129	0.143	0.145	0.069	0.063	0.055	0.052	0.047										
	50	0.115	0.125	0.134	0.139	0.143	0.053	0.051	0.048	0.044	0.042										
	100	0.113	0.123	0.133	0.138	0.143	0.049	0.046	0.046	0.042	0.041										
	200	0.117	0.123	0.134	0.139	0.144	0.047	0.045	0.045	0.041	0.041										
ALS	N=20	NA	NA	0.044	0.058	0.074	NA	NA	0.061	0.049	0.049										
	50	NA	NA	0.035	0.050	0.064	NA	NA	0.031	0.032	0.036										
	100	NA	NA	0.028	0.043	0.060	NA	NA	0.021	0.024	0.032										
	200	NA	NA	0.026	0.037	0.052	NA	NA	0.016	0.019	0.027										

Notes: $\sigma_\eta^2 = 0$, $\rho_{f_1} = 0.5$

Table 4

Small Sample Properties of the GMM, CCEP, CCEMG, ALS and OLS estimators

Heterogeneous slopes without fixed effects (several factors)

$\hat{\beta}_1$		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
GMM	N=20	0.004	0.000	0.001	0.001	0.001	0.097	0.081	0.069	0.055	0.046	0.144	0.135	0.135	0.126	0.113	0.148	0.140	0.142	0.136	0.123
	50	0.000	0.001	0.000	0.002	-0.001	0.038	0.033	0.028	0.022	0.019	0.090	0.087	0.084	0.078	0.075	0.100	0.102	0.098	0.099	0.097
	100	-0.001	0.001	0.001	0.000	-0.001	0.020	0.017	0.014	0.011	0.010	0.073	0.067	0.070	0.064	0.064	0.093	0.092	0.100	0.095	0.100
	200	0.002	0.001	0.002	0.001	0.000	0.010	0.008	0.007	0.006	0.005	0.064	0.060	0.058	0.057	0.055	0.102	0.101	0.110	0.117	0.124
CCEP	N=20	NA	0.030	0.000	0.000	0.002	NA	112.576	0.075	0.048	0.039	NA	0.102	0.027	0.087	0.082	NA	0.112	0.029	0.095	0.093
	50	NA	-0.085	0.001	0.002	0.000	NA	230.511	0.032	0.021	0.018	NA	0.146	0.025	0.058	0.063	NA	0.166	0.025	0.079	0.085
	100	NA	0.562	-0.001	0.001	-0.001	NA	612.208	0.017	0.011	0.009	NA	0.200	0.015	0.055	0.059	NA	0.223	0.024	0.088	0.096
	200	NA	-2.365	0.001	0.001	0.000	NA	60073.387	0.008	0.006	0.005	NA	0.288	0.022	0.050	0.052	NA	0.324	0.024	0.114	0.127
CCEMG	N=20	NA	-14.525	0.495	0.001	0.001	NA	2885956.452	4243.719	0.031	0.024	NA	0.062	0.032	0.076	0.075	NA	0.071	0.032	0.085	0.091
	50	NA	0.405	0.620	0.000	0.001	NA	3331.833	6301.987	0.012	0.009	NA	0.067	0.018	0.059	0.061	NA	0.078	0.022	0.087	0.091
	100	NA	0.157	-5.799	0.001	-0.001	NA	1500.517	191855.387	0.006	0.005	NA	0.076	0.025	0.051	0.057	NA	0.088	0.026	0.101	0.115
	200	NA	0.419	-0.870	0.000	0.000	NA	727.155	24766.845	0.003	0.002	NA	0.087	0.011	0.047	0.052	NA	0.101	0.024	0.149	0.182
OLS	N=20	0.122	0.130	0.135	0.145	0.147	0.131	0.118	0.105	0.090	0.082										
	50	0.117	0.126	0.128	0.140	0.144	0.080	0.074	0.066	0.061	0.056										
	100	0.115	0.125	0.132	0.138	0.142	0.063	0.059	0.055	0.050	0.048										
	200	0.118	0.128	0.132	0.137	0.143	0.054	0.053	0.049	0.046	0.044										
ALS	N=20	NA	NA	0.043	0.082	0.095	NA	NA	0.185	0.149	0.140										
	50	NA	NA	0.040	0.071	0.089	NA	NA	0.091	0.083	0.087										
	100	NA	NA	0.037	0.063	0.080	NA	NA	0.053	0.056	0.064										
	200	NA	NA	0.030	0.054	0.070	NA	NA	0.031	0.040	0.047										

Notes: $\sigma_\eta^2 = 0.4$, $\rho_{f_1} = 0.5$

Table 5

Small Sample Properties of the GMM, CCEP, CCEMG, ALS and OLS estimators

Homogeneous slopes with fixed effects (several factors)

$\hat{\beta}_1$		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
GMM	N=20	0.001	0.000	0.002	-0.001	0.000	0.023	0.018	0.014	0.009	0.006	0.119	0.109	0.110	0.100	0.088	0.142	0.136	0.145	0.138	0.156
	50	0.000	0.001	0.000	0.000	0.002	0.008	0.007	0.005	0.003	0.003	0.075	0.076	0.072	0.075	0.072	0.123	0.132	0.141	0.173	0.218
	100	-0.001	-0.001	0.001	0.000	-0.001	0.004	0.003	0.003	0.002	0.001	0.061	0.064	0.064	0.068	0.051	0.142	0.162	0.200	0.254	0.317
	200	-0.001	0.000	0.000	0.000	-0.001	0.002	0.002	0.001	0.001	0.001	0.059	0.052	0.056	0.056	0.056	0.219	0.260	0.325	0.423	0.531
CCEP	N=20	NA	0.446	0.501	0.000	-0.001	NA	959.509	409.452	0.009	0.004	NA	0.113	0.144	0.060	0.074	NA	0.121	0.161	0.096	0.159
	50	-0.351	5.410	0.329	-0.001	0.001	3846.041	207686.735	118.117	0.004	0.001	0.146	0.169	0.230	0.057	0.055	0.161	0.186	0.257	0.144	0.305
	100	0.127	0.154	0.168	0.000	0.000	52.832	221.685	652.458	0.002	0.001	0.204	0.236	0.308	0.050	0.056	0.223	0.257	0.337	0.230	0.504
	200	0.103	0.316	-0.604	0.000	0.000	231.473	1255.066	2841.201	0.001	0.000	0.274	0.318	0.408	0.053	0.048	0.298	0.342	0.444	0.404	0.786
CCEMG	N=20	NA	-7.569	1.918	-0.001	-0.001	NA	294101.859	28849.970	0.021	0.004	NA	0.058	0.078	0.062	0.071	NA	0.065	0.088	0.083	0.144
	50	NA	1.301	0.188	0.000	0.001	NA	15223.355	3235.178	0.008	0.002	NA	0.055	0.076	0.055	0.058	NA	0.064	0.088	0.099	0.260
	100	NA	0.885	0.307	0.000	0.000	NA	17339.190	706.492	0.004	0.001	NA	0.062	0.071	0.054	0.056	NA	0.070	0.084	0.128	0.441
	200	NA	0.215	-1.516	0.001	0.000	NA	687.649	16169.695	0.002	0.000	NA	0.063	0.075	0.047	0.052	NA	0.072	0.089	0.211	0.724
OLS	N=20	0.041	0.063	0.081	0.097	0.112	0.031	0.031	0.031	0.031	0.032										
	50	0.041	0.061	0.080	0.099	0.113	0.017	0.020	0.022	0.026	0.028										
	100	0.038	0.058	0.077	0.100	0.110	0.013	0.016	0.019	0.025	0.026										
	200	0.038	0.057	0.078	0.098	0.110	0.011	0.014	0.018	0.023	0.025										
ALS	N=20	NA	NA	NA	0.051	0.071	NA	NA	NA	0.055	0.056										
	50	NA	NA	NA	0.040	0.065	NA	NA	NA	0.030	0.040										
	100	NA	NA	NA	0.034	0.049	NA	NA	NA	0.021	0.029										
	200	NA	NA	NA	0.023	0.042	NA	NA	NA	0.014	0.024										

Notes: $\sigma_\eta^2 = 0$, $\rho_{f_1} = 0.5$

Table 6

Small Sample Properties of the GMM, CCEP, CCEMG, ALS and OLS estimators

Heterogeneous slopes with fixed effects (several factors)

$\hat{\beta}_1$		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
GMM	N=20	0.003	-0.002	-0.004	0.002	0.003	0.100	0.086	0.069	0.056	0.046	0.143	0.133	0.123	0.118	0.117	0.150	0.138	0.132	0.127	0.130
	50	0.001	0.004	-0.001	0.002	0.003	0.041	0.034	0.028	0.023	0.019	0.089	0.085	0.082	0.080	0.079	0.099	0.097	0.094	0.096	0.102
	100	0.001	0.001	-0.001	-0.001	0.001	0.021	0.018	0.014	0.012	0.009	0.070	0.070	0.067	0.066	0.064	0.088	0.088	0.093	0.096	0.099
	200	-0.001	0.000	-0.001	0.000	0.000	0.010	0.009	0.007	0.006	0.005	0.060	0.060	0.059	0.061	0.059	0.091	0.099	0.102	0.111	0.127
CCEP	N=20	NA	0.038	0.488	0.000	0.002	NA	248.890	1258.497	0.046	0.035	NA	0.098	0.121	0.082	0.087	NA	0.105	0.132	0.088	0.099
	50	NA	0.081	-3.438	0.000	0.001	NA	674.674	86293.365	0.019	0.014	NA	0.142	0.184	0.063	0.065	NA	0.150	0.204	0.084	0.090
	100	NA	0.103	0.241	0.000	0.000	NA	149.836	83.813	0.009	0.007	NA	0.189	0.243	0.059	0.059	NA	0.206	0.268	0.095	0.105
	200	-1.960	0.061	0.320	0.000	0.000	50140.994	806.556	475.424	0.005	0.004	0.227	0.258	0.336	0.055	0.056	0.245	0.278	0.370	0.125	0.153
CCEMG	N=20	NA	-0.307	-0.745	-0.001	0.002	NA	189860.758	4461.696	0.041	0.025	NA	0.054	0.072	0.069	0.074	NA	0.062	0.080	0.077	0.092
	50	NA	-1.464	-0.273	0.000	0.002	NA	26492.400	11595.105	0.016	0.010	NA	0.050	0.069	0.055	0.061	NA	0.057	0.078	0.077	0.095
	100	NA	2.218	0.033	-0.001	0.000	NA	22734.894	2508.048	0.008	0.005	NA	0.054	0.075	0.057	0.055	NA	0.064	0.088	0.092	0.121
	200	NA	0.715	-0.266	0.000	0.000	NA	975.098	31518.476	0.004	0.002	NA	0.065	0.080	0.055	0.054	NA	0.076	0.094	0.126	0.186
OLS	N=20	0.043	0.059	0.078	0.104	0.116	0.102	0.088	0.076	0.072	0.067										
	50	0.038	0.061	0.075	0.103	0.116	0.046	0.043	0.041	0.044	0.043										
	100	0.039	0.058	0.078	0.098	0.112	0.027	0.028	0.030	0.032	0.034										
	200	0.037	0.058	0.077	0.097	0.112	0.018	0.021	0.023	0.027	0.029										
ALS	N=20	NA	NA	NA	0.067	0.092	NA	NA	NA	0.163	0.142										
	50	NA	NA	NA	0.052	0.085	NA	NA	NA	0.081	0.085										
	100	NA	NA	NA	0.040	0.067	NA	NA	NA	0.050	0.058										
	200	NA	NA	NA	0.032	0.057	NA	NA	NA	0.030	0.041										

Notes: $\sigma_\eta^2 = 0.4$, $\rho_{f_1} = 0.5$

Table 7. Estimates and standard errors for the Evans, Tandon, Murray and Lauer (2000) model

Variables	Fixed Effects	Random Effects	Pooled OLS	ALS	GMM	CCEP	CCEMG
const	3.8022	3.7000	3.2128				
	(0.0291)	(0.0304)	(0.0329)				
$\ln\left(HEXP_{it}\right)$	0.0090	0.0129	0.0839	0.4215	0.0840	0.0004	-0.1009
	(0.0031)	(0.0032)	(0.0057)	(0.0150)	(0.0116)	(0.0248)	(0.0600)
$\ln\left(HC3_{it}\right)$	0.0640	0.0790	0.2551	2.9922	0.2556	0.0164	5.4785
	(0.0364)	(0.0357)	(0.0394)	(0.0338)	(0.0831)	(6.4480)	(0.2087)
$\left(\ln\left(HC3_{it}\right)\right)^2$	0.0206	0.03852	-0.0218	-1.0576	-0.0220	-0.0113	12.9946
	(0.0142)	(0.01333)	(0.0138)	(0.0333)	(0.0299)	(3.0906)	(3.3602)

Note: Standard errors in parenthesis

Table 8: Estimates of the relationship between log exchange rate ratio and log PPP per capita GDP

Variables	Fixed Effects	Random Effects	Pooled OLS	ALS	GMM	CCEP
All countries						
log(PPP GDP per capita$_{it}$)	0.4989	0.3626	0.1753	-0.0595	0.1671	-0.1322
	(0.01449)	(0.0117)	(0.0059)	(0.0158)	(0.0208)	(0.0398)
OECD countries						
log(PPP GDP per capita$_{it}$)	0.8490	0.7415	0.4986	-0.0147	0.4614	0.6837
	(0.0331)	(0.0297)	(0.0196)	(0.0212)	(0.0663)	(0.0504)

Note: Standard errors in parenthesis

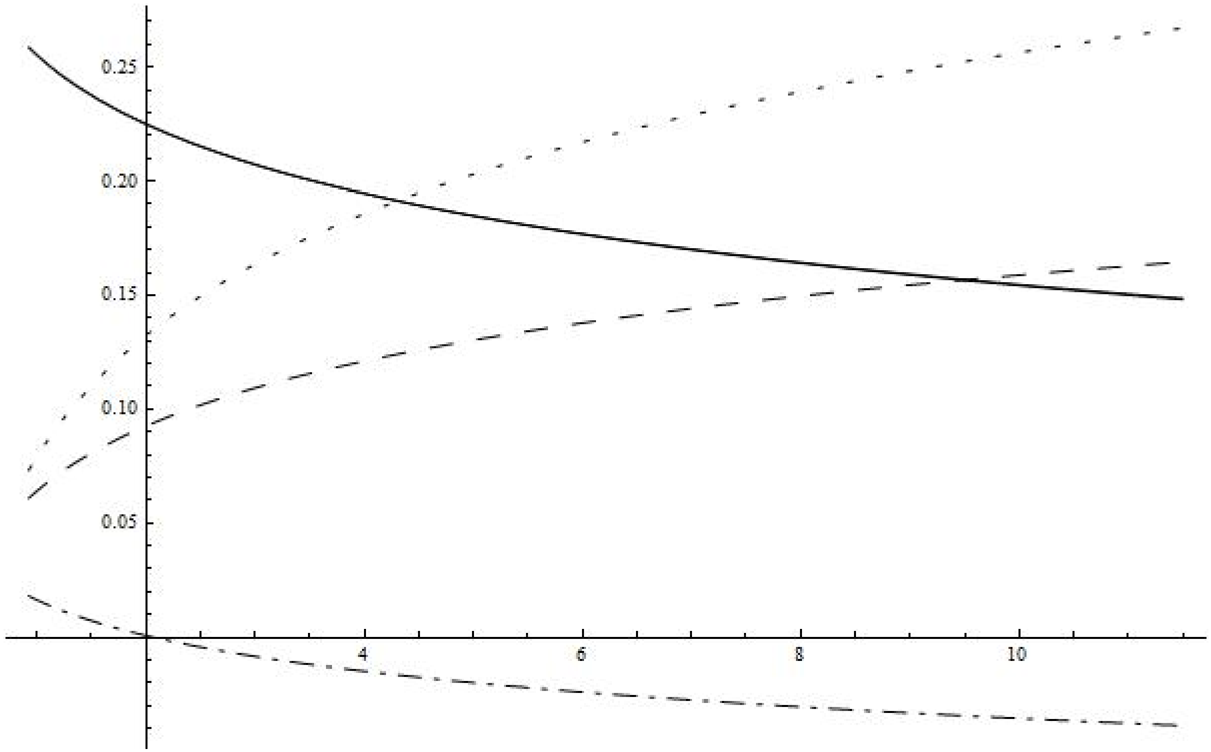

Figure 1: Elasticity of health care attainment with respect to educational attainment (vertical axis) as functions of educational attainments (horizontal axis) calculated for the fixed effects estimator (dashed line), the random effects estimator (dotted line), the GMM estimator (solid line) and the CCEP estimator (dashed-dotted line). For the pooled OLS estimator the elasticity is the same as the one produced by the GMM estimator.

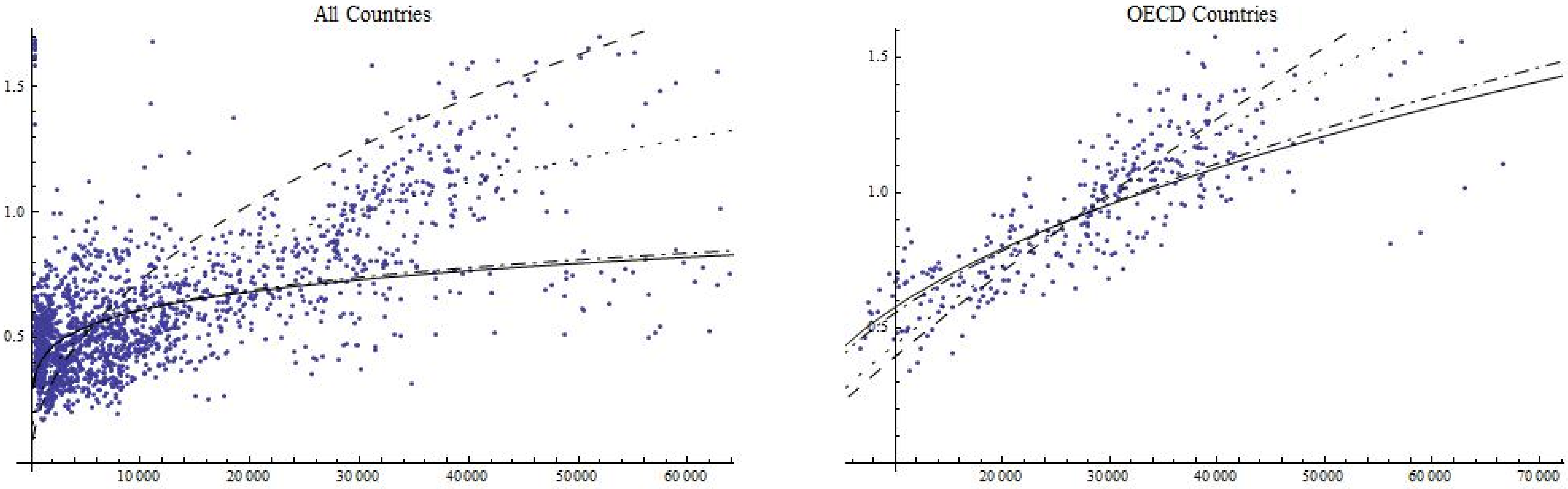

Figure 2: Log exchange rate ratio (vertical axis) vs PPP GDP per capita (horizontal axis): raw data and fitted values from the fixed effects estimator (dashed line), the random effects estimator (dotted line), the GMM estimator (solid line) and the CCEP estimator (dashed-dotted line). For the pooled OLS estimator the elasticity is the same as the one produced by the GMM estimator