econometrics

MDPI

Article

A Conditional Approach to Panel Data Models with Common Shocks

Giovanni Forchini [1,*] and Bin Peng [2]

Received: 15 September 2015; Accepted: 6 January 2016; Published: 12 January 2016
Academic Editor: Kerry Patterson

[1] School of Economics, Ground Floor AD Building, University of Surrey, Guildford, Surrey GU2 7XH, UK
[2] Economics Discipline Group, University of Technology Sydney, Sydney 2007, Australia; Bin.Peng@uts.edu.au
* Correspondence: Giovanni.Forchini@gmail.com; Tel.: +44-(0)1483-68-2772

Abstract: This paper studies the effects of common shocks on the OLS estimators of the slopes' parameters in linear panel data models. The shocks are assumed to affect both the errors and some of the explanatory variables. In contrast to existing approaches, which rely on using results on martingale difference sequences, our method relies on conditional strong laws of large numbers and conditional central limit theorems for conditionally-heterogeneous random variables.

Keywords: factor structure; common shocks; conditional independence; conditional central limit theorem

JEL: C23

1. Introduction

The effects of common shocks, which may be macroeconomic, technological, institutional, political, environmental, health related, sociological, *etc.* (e.g., [1]), have been recently investigated by various authors, including, among others, [1–6]. There are several examples in economics where common shocks may affect the analysis:

- Accounting for technological and sociological shocks is extremely important when explaining the healthcare attainments of different countries in terms of, say, their per capita health expenditures and educational attainments (e.g., [7]).
- Financial and political shocks are likely to be relevant when explaining the differences in individual countries' exchange rate ratios (*i.e.*, the ratio between purchasing power parity relative to the U.S., say, and the nominal exchange rate relative to the U.S.) in terms of their per capita GDP measured in purchasing power parity—the Balassa-Samuelson hypothesis (e.g., [8]).
- Finance, political, environmental and industry-specific shocks impact the models of executive compensation in which the latter is explained by returns on assets, stock returns, the level of responsibility and gender.
- In the cross-country cross-industry analysis of returns to R & D, both global shocks (e.g., the recent financial crisis) and local shocks (e.g., spillovers between a limited group of industries or countries) may be fundamental in explaining output, as well as the explanatory variables (*cf.* [9]).

Other detailed examples (e.g., consumption model and asset pricing model) are provided in Appendix A of [10].

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 factors in the errors of the model (e.g., [3,4]) for which consistent estimation of the parameters of interest could be done by maximum likelihood procedures (e.g., [11]). Coakley, Fuertes and Smith [12] 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, or a combination of the two (see among others [1,5]), and would thus induce endogeneity requiring more sophisticated estimation procedures. A recent survey can be found in [13].

Andrews [1] has studied the conditions for the consistency of the OLS estimator in a cross-section regression with common shocks, which are captured by the sigma algebra generated by the factors. Andrews [1] assumes that the observed variables are independent and identically distributed given such sigma algebra. He shows that the OLS estimator is consistent if and only if the errors and the regressors are uncorrelated given the sigma algebra generated by unobserved factors and shows that the (possibly re-centred) OLS estimator has an asymptotic mixed normal distribution. However, tests on the coefficients can be constructed using classical distributions under the null if the OLS estimator is consistent. Andrews [1] also extends his results to the OLS and the fixed effect estimators in panel data with fixed T.

Recently, Kuersteiner and Prucha [14] have extended the work of Andrews [1] by deriving a stable central limit theorem for sample moments under weaker assumptions and have established limiting distributions of GMMand maximum likelihood estimators for general models in which unobservable factors may induce cross-sectional heterogeneity, but do not affect the regressors. The approach of Kuersteiner and Prucha is based on a generalization of Corollary 3.1 of [15] on martingale difference sequences, which allows them to deal with sequentially exogenous regressors. Kao, Trapani and Urga [16] also employ Corollary 3.1 of [15] to investigate panel data models with factor structures.

Although central limit theorems for martingale difference sequences are very powerful tools in time series, their application in cross-sections and panel data with a fixed time dimension is not fully intuitive, and the assumptions employed may be cumbersome. The fundamental reason for this is that in such models, there is no natural order of the observations, and the assumptions must be formulated to guarantee the validity of the derived results for all possible permutations of the sequences of observations.

This paper proposes an alternative approach to that of [1,14] to study estimators of linear panel data models in which the errors and the regressors are affected by common shocks represented by common factors. In contrast to the work of [1,14], our work employs a conditional strong law of large numbers (e.g., [17–20]) and a conditional central limit theorem (e.g., [19–24]) from which stable convergence follows.

Conditional strong laws of large numbers and conditional central limit theorems are very similar to their standard counterparts and, therefore, are familiar and intuitive to econometricians. Similarly, the assumptions on which they are based are simple, and one does not have to worry about establishing the validity of the results under all possible permutations of the sequences of observations. We will show in a companion paper that the approach can be used to analyse panel data models with endogeneity due to both simultaneity and factor structures in the errors and the explanatory variables (e.g., [25]).

The approach that we suggest is based on two steps:

1. formulation of the assumptions concerning the unobservable and heterogeneous variables conditional on the sigma algebra capturing the common shocks;
2. application of conditional strong laws of large numbers and conditional central limit theorems to establish the limits of the estimator of interest conditional on the common shocks, from which the unconditional distribution can be obtained.

This approach is used in Section 2 to study the OLS estimator for the slope coefficients in a panel data model with homogeneous slopes and in Section 3 to investigate a model with heterogeneous slopes. Section 4 briefly discusses a fixed effects model, and Section 5 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:

$$\begin{aligned}
\underset{T\times 1}{\boldsymbol{y}_i} &= \underset{T\times 1}{\boldsymbol{\tau}} + \underset{T\times k}{\boldsymbol{Z}_i}\,\underset{k\times 1}{\boldsymbol{\alpha}_0} + \underset{T\times p}{\boldsymbol{X}_i}\,\underset{p\times 1}{\boldsymbol{\beta}_0} + \underset{T\times 1}{\boldsymbol{u}_i}\,, \\
\boldsymbol{u}_i &= \underset{T\times m}{\boldsymbol{F}_T}\,\underset{m\times 1}{\gamma_i} + \underset{T\times 1}{\boldsymbol{\varepsilon}_i}\,, \\
\boldsymbol{X}_i &= \boldsymbol{F}_T\,\underset{m\times p}{\boldsymbol{\Gamma}_i} + \underset{T\times p}{\boldsymbol{V}_i}\,.
\end{aligned} \tag{1}$$

The observed regressors are split into two groups: those that are not affected by common shocks (e.g., unit characteristics, such as gender, race, age, *etc.*), $\boldsymbol{Z}_i$, and those that may be affected by common shocks, $\boldsymbol{X}_i$. The parameters associated with the regressors, $\boldsymbol{\alpha}_0$ and $\boldsymbol{\beta}_0$, and the constant vector $\boldsymbol{\tau}$ are the same for $i = 1, \ldots, N$. The common shocks are captured by the matrix of unobserved common factors, $\boldsymbol{F}_T$, (*cf.* [1]); γ_i and $\boldsymbol{\Gamma}_i$ are factor loadings; $\boldsymbol{\varepsilon}_i$ is a purely idiosyncratic random vector with zero mean and arbitrary covariance matrix, which may depend on i; and $\boldsymbol{V}_i$ represents the values of the regressors that would be observed in the absence of common shocks. Factors, factor loadings, $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{V}_i$ are not observed. The factor structure generates cross-sectional heterogeneity in the error term of (1). This also creates correlation between errors $\boldsymbol{u}_i$ and regressors $\boldsymbol{X}_i$. Notice also that the impact of the unobserved shocks is different for each unit depending on the realizations of the factor loadings Γ_i and γ_i.

It may help to think of an example where a model like (1) is applicable. Suppose we are interested in estimating the healthcare attainments of different countries in terms of, say, their per capita health expenditures and educational attainments. In this case, y_i contains a measure of the educational attainment of country i over T years; Z_i is a vector containing a measure of educational attainments over the T periods for country i, and X_i is per capita health expenditure over the same period. There is also a dummy for each time period whose coefficients are in τ. The common shocks are represented by new procedures, drugs, surgical techniques, *etc.* The shocks are not observed by the econometrician. They directly affect the healthcare attainments. However, they also affect health expenditures over time for the different countries. Therefore, the observed health expenditure also includes the common shocks.

All variables are defined on a probability space $(\Omega, \mathcal{A}, P)$. The sigma algebra generated by the random vector $\boldsymbol{F}_T$ is denoted by $\mathcal{F}$. 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 that differ only on sets of probability zero are regarded as equivalent. We will regard conditioning on $\mathcal{F}$ as conditioning on the factors $\boldsymbol{F}_T$. In the rest of the paper, $\|\cdot\|$ denotes the Euclidean norm for a vector and the Frobenius norm for a matrix.

We now introduce assumptions on both the observed and the unobserved variables. These are adapted from [1], but allow for heterogeneity conditional on $\mathcal{F}$. 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. The following assumptions state that the unobservables are independent sequences of independent random quantities given the factors.

Assumption:
Let δ be a positive constant and Δ be $\mathcal{F}$-measurable and such that $\Delta \leq \infty$ a.s.

- **C1** $\{\boldsymbol{\varepsilon}_i, i \geq 1\}$ is a sequence of conditionally-independent random vectors given $\mathcal{F}$, $E[\boldsymbol{\varepsilon}_i|\mathcal{F}] = \boldsymbol{0}$ a.s. and $E[\|\boldsymbol{\varepsilon}_i\|^{1+\delta}|\mathcal{F}] < \Delta$ a.s.
- **C2** $\{(\boldsymbol{Z}_i, \boldsymbol{V}_i), i \geq 1\}$ is a sequence of conditionally-independent random matrices given $\mathcal{F}$ with $E[\|(\boldsymbol{Z}_i, \boldsymbol{V}_i)\|^{2+\delta}|\mathcal{F}] < \Delta$ a.s.
- **C3** $\{\gamma_i, i \geq 1\}$ is a sequence of conditionally-independent random vectors given $\mathcal{F}$, $E[\gamma_i|\mathcal{F}] = \gamma$ a.s., where γ is $\mathcal{F}$-measurable, and $E[\|\gamma_i\|^{1+\delta}|\mathcal{F}] < \Delta$ a.s.

C4 $\{\boldsymbol{\Gamma}_i, i \geq 1\}$ is a sequence of conditionally-independent random matrices given $\mathcal{F}$, $E[\boldsymbol{\Gamma}_i|\mathcal{F}] = \boldsymbol{\Gamma}$ a.s., where $\boldsymbol{\Gamma}$ is $\mathcal{F}$-measurable, and $E[\|\boldsymbol{\Gamma}_i\|^{2+\delta}|\mathcal{F}] < \Delta$ a.s.

C5 $\{\boldsymbol{\varepsilon}_i, i \geq 1\}$, $\{(\boldsymbol{Z}_i, \boldsymbol{V}_i), i \geq 1\}$, $\{\gamma_i, i \geq 1\}$ and $\{\boldsymbol{\Gamma}_i, i \geq 1\}$ are conditionally independent of each other given $\mathcal{F}$.

C6 $E\left[\frac{1}{N}\sum\limits_{i=1}^{N}(\boldsymbol{Z}_i - \bar{\boldsymbol{Z}}, \boldsymbol{V}_i - \bar{\boldsymbol{V}})'(\boldsymbol{Z}_i - \bar{\boldsymbol{Z}}, \boldsymbol{V}_i - \bar{\boldsymbol{V}})|\mathcal{F}\right]$ is uniformly positive definite a.s., where $\bar{\boldsymbol{Z}} = \frac{1}{N}\sum\limits_{i=1}^{N}\boldsymbol{Z}_i$ and $\bar{\boldsymbol{V}} = \frac{1}{N}\sum\limits_{i=1}^{N}\boldsymbol{V}_i$.

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 $\boldsymbol{\varepsilon}_i$, are assumed to be purely idiosyncratic, and the $(\boldsymbol{Z}_i, \boldsymbol{V}_i)$ are assumed to form a sequence of independent random vectors given the factors. Since we interpret $\boldsymbol{V}_i$ as a vector of regressors, which would be observed if the common shocks would not affect the regressors, we assume that these form an independent sequence of events that are heterogeneous and may be correlated with $\boldsymbol{Z}_i$. Notice that γ and $\boldsymbol{\Gamma}$ may be constant or may be functions of the factors.

The factor loadings in both the regressors and the errors are assumed to be independent conditional on $\mathcal{F}$, but not necessarily identically distributed. We will consider a violation of this assumption later. Notice that Assumption C5 only requires independence conditional on $\mathcal{F}$, but does not require the factor loadings in the regressors and errors to be independent unconditionally. Section 7 of [1,19,26] gives a thorough discussion of the relationship between conditional and unconditional independence.

Andrews [1] considers a similar model for $T = 1$ in which $\boldsymbol{\varepsilon}_i$, $(\boldsymbol{Z}_i, \boldsymbol{V}_i)$, γ_i, and $\boldsymbol{\Gamma}_i$ are conditionally independent of each other given $\mathcal{F}$. This implies that these random vectors and matrices are exchangeable. It is easy to see that exchangeable random variables are identically distributed (but not necessarily independent) unconditionally. On the other hand, Assumption C5 implies that $\boldsymbol{\varepsilon}_i$, $(\boldsymbol{Z}_i, \boldsymbol{V}_i)$, γ_i, and $\boldsymbol{\Gamma}_i$ may be dependent and non-identically distributed unconditionally.

Assumption C6 is needed for the application of the conditional weak law of large numbers. It ensures that the OLS estimator of the slope parameters has an asymptotic non-singular normal distribution conditional on the sigma algebra $\mathcal{F}$.

The OLS estimator of the slope parameters is

$$\hat{\boldsymbol{\theta}} = \left(\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{\mathcal{X}}}\right)^{-1}\left(\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{y}_i - \bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{y}}\right) \tag{2}$$

where $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\alpha}}', \hat{\boldsymbol{\beta}}')'$, $\boldsymbol{\mathcal{X}}_i = (\boldsymbol{Z}_i, \boldsymbol{X}_i)$, $\bar{\boldsymbol{\mathcal{X}}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i$ and $\bar{\boldsymbol{y}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{y}_i$.

Theorem 1. *Under Assumptions C1–C6, as $N \to \infty$, $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\alpha}}', \hat{\boldsymbol{\beta}}')'$ is unbiased, and $\hat{\boldsymbol{\theta}} \to \boldsymbol{\theta}_0$ a.s., where $\boldsymbol{\theta}_0 = (\boldsymbol{\alpha}_0', \boldsymbol{\beta}_0')'$*

Theorem 1 shows that, for fixed T, the estimator $\hat{\boldsymbol{\theta}}$ is unbiased and consistent as N tends to infinity. For the case where the factors affect all of the regressors, unbiasedness is also noticed for the estimator in (2) by [27], where the factor loadings in the regressors and the errors are assumed to be mutually independent (which is stronger than Assumption C5).

To obtain the asymptotic distribution of $\hat{\boldsymbol{\theta}}$, we need slightly stronger versions of Assumptions C1 and C3 requiring the existence of higher order moments.

Assumption:
Let δ be a positive constant and Δ be $\mathcal{F}$-measurable and such that $\Delta \leq \infty$ a.s.

N1 $\{\boldsymbol{\varepsilon}_i, i \geq 1\}$ is a sequence of conditionally-independent random vectors given $\mathcal{F}$. $E[\boldsymbol{\varepsilon}_i|\mathcal{F}] = \mathbf{0}$ a.s. and $E[\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'|\mathcal{F}] = \boldsymbol{\Sigma}_{\varepsilon_i}$ a.s. with $\boldsymbol{\Sigma}_{\varepsilon_i}$ being $\mathcal{F}$-measurable uniformly in i. Moreover, $E[\|\boldsymbol{\varepsilon}_i\|^{2+\delta}|\mathcal{F}] < \Delta$ a.s.
N3 $\{\boldsymbol{\gamma}_i, i \geq 1\}$ is a sequence of conditionally-independent random vectors given $\mathcal{F}$. $E[\boldsymbol{\gamma}_i|\mathcal{F}] = \boldsymbol{\gamma}$ a.s. and $\mathrm{Cov}[\boldsymbol{\gamma}_i|\mathcal{F}] = \boldsymbol{\Sigma}_{\gamma_i}$ a.s. with $\boldsymbol{\Sigma}_{\gamma_i}$ being $\mathcal{F}$-measurable uniformly in i. Moreover, $E[\|\boldsymbol{\gamma}_i\|^{2+\delta}|\mathcal{F}] < \Delta$ a.s.

Theorem 2. *Under Assumptions N1, C2, N3 and C4–C6, conditional on $\mathcal{F}$, as $N \to \infty$*

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \to_D \boldsymbol{B}^{-1}(\boldsymbol{F}_T)\boldsymbol{C}^{1/2}(\boldsymbol{F}_T)N(\mathbf{0}, \boldsymbol{I}_{p+k}),$$

where $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\alpha}}', \hat{\boldsymbol{\beta}}')'$, $\boldsymbol{\theta}_0 = (\boldsymbol{\alpha}_0', \boldsymbol{\beta}_0')'$,

$$\boldsymbol{B}(\boldsymbol{F}_T) = \lim_{N\to\infty}\left(\frac{1}{N}\sum_{i=1}^{N} E[\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i|\mathcal{F}] - E[\bar{\boldsymbol{\mathcal{X}}}'|\mathcal{F}]E[\bar{\boldsymbol{\mathcal{X}}}|\mathcal{F}]\right), \tag{3}$$

$$\boldsymbol{C}(\boldsymbol{F}_T) = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}})'(\boldsymbol{\Sigma}_{\varepsilon_i} + \boldsymbol{F}_T\boldsymbol{\Sigma}_{\gamma_i}\boldsymbol{F}_T')(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}})|\mathcal{F}\right]. \tag{4}$$

Theorem 2 shows that, for fixed T, $\hat{\boldsymbol{\theta}}$ has a normal asymptotic distribution conditional on $\mathcal{F}$. Since the factors are unobservable, one needs to remove the conditioning on them by averaging them out. Thus, the unconditional asymptotic distribution of the OLS estimator of the slope parameters is covariance matrix mixed normal with mixing density given by the density function of the unobserved factors. Notice also that Theorem 2 implies that $\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ converges $\mathcal{F}$-stably.

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

$$\hat{\boldsymbol{B}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{\mathcal{X}}} \to \boldsymbol{B}(\boldsymbol{F}_T) \quad \text{a.s.} \tag{5}$$

For $\boldsymbol{C}(\boldsymbol{F}_T)$, we need more restrictive versions of Assumptions C2 and C4 requiring the existence of higher order moments.

Assumption:
Let δ be a positive constant and Δ be $\mathcal{F}$-measurable and such that $\Delta \leq \infty$ a.s.

CM2 $\{(\boldsymbol{Z}_i, \boldsymbol{V}_i), i \geq 1\}$ is a sequence of conditionally-independent random matrices given $\mathcal{F}$ with $E[\|(\boldsymbol{Z}_i, \boldsymbol{V}_i)\|^{4+\delta}|\mathcal{F}] < \Delta$ a.s.
CM4 $\{\boldsymbol{\Gamma}_i, i \geq 1\}$ is a sequence of conditionally-independent random matrices given $\mathcal{F}$ with $E[\boldsymbol{\Gamma}_i|\mathcal{F}] = \boldsymbol{\Gamma}$, $E[\|\boldsymbol{\Gamma}_i\|^{4+\delta}|\mathcal{F}] < \Delta$ a.s.

Lemma 1. *Given Assumption N1, CM2, N3, CM4 and C5–C6, conditional on $\mathcal{F}$, as $N \to \infty$*

$$\begin{aligned}\hat{\boldsymbol{C}} =& \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}})'\left[\boldsymbol{y}_i - \bar{\boldsymbol{y}} - (\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}})\hat{\boldsymbol{\theta}}\right]\left[\boldsymbol{y}_i - \bar{\boldsymbol{y}} - (\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}})\hat{\boldsymbol{\theta}}\right]'(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}) \\ &\to \boldsymbol{C}(\boldsymbol{F}_T), \quad a.s.\end{aligned} \tag{6}$$

An asymptotic version of the F-test conditional on $\mathcal{F}$ for the null hypothesis that $H_0 : \boldsymbol{R\theta}_0 = \boldsymbol{r}$ against the alternative hypothesis $H_1 : \boldsymbol{R\theta}_0 \neq \boldsymbol{r}$ can be easily constructed condition on $\mathcal{F}$

$$N(\boldsymbol{R}\hat{\boldsymbol{\theta}} - \boldsymbol{r})'(\boldsymbol{R}\hat{\boldsymbol{B}}^{-1}\hat{\boldsymbol{C}}\hat{\boldsymbol{B}}^{-1}\boldsymbol{R}')^{-1}(\boldsymbol{R}\hat{\boldsymbol{\theta}} - \boldsymbol{r}) \to_D \chi^2(q), \tag{7}$$

where $\boldsymbol{R}$ is a known and fixed $q \times (p+k)$ matrix of rank $q < (p+k)$ and $\boldsymbol{r}$ is a known and fixed $q \times 1$ vector.

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

Assumption:

D5 $\{\varepsilon_i, i \geq 1\}$, $\{(\boldsymbol{Z}_i, \boldsymbol{V}_i), i \geq 1\}$, and $\{(\gamma_i, \boldsymbol{\Gamma}_i), i \geq 1\}$ are conditionally-independent of each other given $\mathcal{F}$.

Assumption D5 differs from C5 because it allows the factor loading in the regressors $\boldsymbol{\Gamma}_i$ and those in the errors γ_i to be correlated conditional on $\mathcal{F}$ for each i. This means that the endogeneity induced by the factor structure persists even conditioning on the factors. Therefore, the OLS estimator of the slope parameters will be biased, as shown by the following theorem.

Theorem 3. *Under Assumptions C1–C4, D5 and C6, conditional on $\mathcal{F}$, as $N \to \infty$*

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + \boldsymbol{B}^{-1}(\boldsymbol{F}_T)\left(\boldsymbol{0}',\ \boldsymbol{\phi}(\boldsymbol{F}_T)'\right)' \quad a.s.,$$

where $\boldsymbol{B}(\boldsymbol{F}_T)$ is defined in (3) and $\boldsymbol{\phi}(\boldsymbol{F}_T) = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} E[\boldsymbol{\Gamma}_i'\boldsymbol{F}_T'\boldsymbol{F}_T\gamma_i|\mathcal{F}] - \boldsymbol{\Gamma}'\boldsymbol{F}_T'\boldsymbol{F}_T\gamma$.

Notice that by replacing Assumption C5 with Assumption D5, the estimator of $\hat{\boldsymbol{\theta}}$ has an asymptotic bias conditional on $\mathcal{F}$, which depends in a complicated way on the distribution of the factors and of the factor loadings. This implies that unconditionally, the estimator of $\boldsymbol{\theta}_0 = (\boldsymbol{\alpha}_0',\ \boldsymbol{\beta}_0')'$ has a non-degenerate non-standard asymptotic distribution. The intuition behind this result is as follows: since the factor loadings are correlated among themselves, even conditioning on the factors, endogeneity is present even when we condition on $\mathcal{F}$.

3. Heterogeneous Slopes

In this section, we consider a more general case, where the coefficients of $\boldsymbol{Z}_i$ and $\boldsymbol{X}_i$ are allowed to be different for each unit. Precisely, the model is

$$\begin{aligned}
\underset{T\times 1}{\boldsymbol{y}_i} &= \underset{T\times 1}{\boldsymbol{\tau}} + \underset{T\times k}{\boldsymbol{Z}_i}\underset{k\times 1}{\boldsymbol{\alpha}_i} + \underset{T\times p}{\boldsymbol{X}_i}\underset{p\times 1}{\boldsymbol{\beta}_i} + \underset{T\times 1}{\boldsymbol{u}_i}, \\
\boldsymbol{u}_i &= \underset{T\times m}{\boldsymbol{F}_T}\underset{m\times 1}{\gamma_i} + \underset{T\times 1}{\boldsymbol{\varepsilon}_i}, \\
\boldsymbol{X}_i &= \boldsymbol{F}_T \underset{m\times p}{\boldsymbol{\Gamma}_i} + \underset{T\times p}{\boldsymbol{V}_i}, \\
(\boldsymbol{\alpha}_i',\ \boldsymbol{\beta}_i') &= (\boldsymbol{\alpha}_0',\ \boldsymbol{\beta}_0') + (\boldsymbol{\eta}_{1i}',\ \boldsymbol{\eta}_{2i}').
\end{aligned} \tag{8}$$

where $(\boldsymbol{\eta}_{1i}',\ \boldsymbol{\eta}_{2i}')$'s are random variables. We are interested in inference about the mean of the unit-specific coefficients $(\boldsymbol{\alpha}_i',\ \boldsymbol{\beta}_i')$. These parameters are estimated using the OLS estimator defined in (8) in the previous sections. Some further assumptions are needed.

Assumptions:
Let δ be a positive constant and Δ be $\mathcal{F}$-measurable and such that $\Delta \leq \infty$ a.s.

H5 $\{(\boldsymbol{\eta}_{1i}',\ \boldsymbol{\eta}_{2i}'), i \geq 1\}$, $\{\varepsilon_i, i \geq 1\}$, $\{(\boldsymbol{Z}_i, \boldsymbol{V}_i), i \geq 1\}$, $\{\gamma_i, i \geq 1\}$ and $\{\boldsymbol{\Gamma}_i, i \geq 1\}$ are conditionally-independent of each other given $\mathcal{F}$.

H7 $\{(\eta'_{1i}, \eta'_{2i}), i \geq 1\}$ is a sequence of conditionally-independent random matrices given $\mathcal{F}$ with $E[(\eta'_{1i}, \eta'_{2i})'|\mathcal{F}] = \mathbf{0}$, $E[\|(\eta'_{1i}, \eta'_{2i})\|^{1+\delta}|\mathcal{F}] < \Delta$ a.s.

NH7 $\{(\eta'_{1i}, \eta'_{2i}), i \geq 1\}$ is a sequence of conditionally-independent random matrices given $\mathcal{F}$ with $E[(\eta'_{1i}, \eta'_{2i})'|\mathcal{F}] = \mathbf{0}$, $E[(\eta'_{1i}, \eta'_{2i})'(\eta'_{1i}, \eta'_{2i})|\mathcal{F}] = \boldsymbol{\Sigma}_\eta$ and $E[\|(\eta'_{1i}, \eta'_{2i})\|^{2+\delta}|\mathcal{F}] < \Delta$ a.s.

Assumption H5 extends Assumption C5 by requiring that $\{(\eta'_{1i}, \eta'_{2i}), i \geq 1\}$ is independent of all other variables conditional on $\mathcal{F}$.

The next result gives the distributional properties for the OLS estimator for the slope parameters in (8).

Theorem 4. *Let $\boldsymbol{B}(\boldsymbol{F}_T)$ and $\boldsymbol{C}(\boldsymbol{F}_T)$ be as in (3) and (4). As $N \to \infty$*

1. *Under Assumptions C1–C4, H5, C6 and H7, $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\alpha}}', \hat{\boldsymbol{\beta}}')'$ is unbiased and consistent.*
2. *Under Assumptions N1, CM2, N3, CM4, H5, C6 and NH7, conditional on $\mathcal{F}$*

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \to_D \boldsymbol{B}^{-1}(\boldsymbol{F}_T)\left(\boldsymbol{C}(\boldsymbol{F}_T) + \boldsymbol{C}^*(\boldsymbol{F}_T)\right)^{1/2} N(\mathbf{0}, \boldsymbol{I}_{p+k}),$$

where $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\alpha}}', \hat{\boldsymbol{\beta}}')'$, $\boldsymbol{\theta}_0 = (\boldsymbol{\alpha}'_0, \boldsymbol{\beta}'_0)'$, and

$$\boldsymbol{C}^*(\boldsymbol{F}_T) = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} E\left[(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}})'\boldsymbol{\mathcal{X}}_i\boldsymbol{\Sigma}_\eta\boldsymbol{\mathcal{X}}'_i(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}})|\mathcal{F}\right].$$

Theorem 4 shows that for a fixed T, the OLS 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 $\boldsymbol{C}^*(\boldsymbol{F}_T)$. Thus, the effect of random coefficients on the asymptotic properties of the OLS estimator is just an increase in the conditional variance. It also follows from Theorem 4 that $\hat{\theta}$ converges $\mathcal{F}$-stably to a covariance matrix mixed normal random vector. Notice that Theorem 4 reduces to Theorems 1 and 2 if (η'_{1i}, η'_{2i})'s are identically zero.

In order to construct tests on $\boldsymbol{\theta}_0 = (\boldsymbol{\alpha}'_0, \boldsymbol{\beta}'_0)'$, we need to find a statistic that converges to $\boldsymbol{B}^{-1}(\boldsymbol{F}_T)\left(\boldsymbol{C}(\boldsymbol{F}_T) + \boldsymbol{C}^*(\boldsymbol{F}_T)\right)\boldsymbol{B}^{-1}(\boldsymbol{F}_T)$ conditional on $\mathcal{F}$ as N tends to infinity. This is given in the following lemma.

Lemma 2. *Under Assumptions N1, CM2, N3, CM4, H5, C6 and HN7, as $N \to \infty$,*

$$\hat{\boldsymbol{C}} \to_P \boldsymbol{C}(\boldsymbol{F}_T) + \boldsymbol{C}^*(\boldsymbol{F}_T) \quad a.s.$$

conditional on $\mathcal{F}$, where $\hat{\boldsymbol{C}}$ is given in (6).

Tests of hypotheses can then be constructed as outlined in the previous section.

4. A Fixed Effects Model

We now briefly discuss the fixed effects model

$$\underset{(T\times1)}{\boldsymbol{y}_i} = \underset{(T\times1)}{\boldsymbol{\tau}} + \underset{(T\times1)}{\mathbf{1}_T}\underset{(1\times1)}{\boldsymbol{\theta}_i} + \underset{(T\times k)}{\boldsymbol{Z}_i}\underset{(k\times1)}{\boldsymbol{\alpha}_0} + \underset{(T\times p)}{\boldsymbol{X}_i}\underset{(p\times1)}{\boldsymbol{\beta}_0} + \boldsymbol{u}_i, \tag{9}$$

where $\mathbf{1}_T$ denotes a $(T \times 1)$ vector of ones and $\boldsymbol{\theta}_i$ denotes the unit specific fixed effect. Let $\boldsymbol{C}_{T\times(T-1)}$ be a matrix, such that $\boldsymbol{C}'\mathbf{1}_T = \mathbf{0}$, $\boldsymbol{C}'\boldsymbol{C} = \boldsymbol{I}_{T-1}$ and $\boldsymbol{C}\boldsymbol{C}' = \boldsymbol{M}_{\mathbf{1}_T}$, where $\boldsymbol{M}_{\mathbf{1}_T}$ is the usual projection matrix in the space orthogonal to $\mathbf{1}_T$. Pre-multiplying (9) by $\boldsymbol{C}$, we obtain

$$\boldsymbol{C}'\boldsymbol{y}_i = \boldsymbol{C}'\boldsymbol{\tau} + \boldsymbol{C}'\left(\boldsymbol{Z}_i, \boldsymbol{X}_i\right)\begin{pmatrix}\boldsymbol{\alpha}_0\\ \boldsymbol{\beta}_0\end{pmatrix} + \boldsymbol{C}'\boldsymbol{u}_i, \tag{10}$$

which has the same form as the model described in (1). Notice that the assumptions involving $\boldsymbol{C}'\boldsymbol{u}_i$ and $\boldsymbol{C}'(\boldsymbol{Z}_i, \boldsymbol{X}_i)$ follow from sub-additivity of the norm, and the fact that $\boldsymbol{C}$ is a finite matrix of full rank can be established from the assumptions above.

The results of the previous sections, including those on heterogeneous slopes, can be applied to this case with obvious changes of notation. Thus,

- The fixed effect estimator is consistent if the factor loadings $\boldsymbol{\gamma}_i$ and $\boldsymbol{\Gamma}_i$ are independent. In this case, once standardised, the fixed effect estimator is asymptotically normal given $\mathcal{F}$, and thus, it has an asymptotic covariance matrix mixed normal distribution. In this case, standard t- and F-tests on the slope coefficients have standard asymptotic distributions under the null hypothesis.
- If the factor loadings in the error and the regressors are not independent, then the fixed effect estimator has a non-degenerated asymptotic distribution.

5. Conclusions

This paper has considered a panel data model with both homogeneous and heterogeneous slopes, with multi-factor error structures in the errors and the regressors. The method employed has relied on an approach based on a conditional strong law of large numbers and a conditional central limit theorem, which are similar to the results with which econometricians are familiar.

The model assumptions have been formulated conditional on the sigma algebra generated by the factors, and it has been shown that the OLS estimator of the slope parameters is consistent in both the homogeneous and heterogeneous case if the factor loadings in the regressors and the errors are independent conditional on the factors. It this case, the OLS estimator has an asymptotic mixed normal distribution, but t- and F-tests have standard distributions under the null hypothesis. The fixed effects model was also discussed.

Acknowledgments: This research was partially supported by an Australian Research Council Grant DP0985432. We would like to thank three referees for very helpful comments and suggestions.

Author Contributions: The authors contributed equally to the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix: Proofs

Theorem A1 *(conditional Markov strong law of large numbers): Let* $\{z_i : i \geq 1\}$ *be a sequence of* $\mathcal{F}$*-independent random variables with conditional means* $E[z_i|\mathcal{F}]$ *for* $i = 1, 2, \ldots$ *If for some scalar* $0 < \delta \leq 1$, $\sum_{i=1}^{\infty} \frac{1}{i^{1+\delta}} E\left[|z_i - E[z_i|\mathcal{F}]|^{1+\delta} |\mathcal{F}\right] < \infty$ *a.s., then conditional on* $\mathcal{F}$, $\frac{1}{n}\sum_{i=1}^{n}(z_i - E[z_i|\mathcal{F}]) \to 0$ *a.s.*

Theorem A2 *(conditional Liapounov central limit theorem): Let* $\{z_i : 1 \leq i \leq n\}$ *be a sequence of* $\mathcal{F}$*-independent random variables with conditional means* $E[z_i|\mathcal{F}]$*, conditional variances* $\sigma_i^2 = E\left[(z_i - E[z_i|\mathcal{F}])^2 |\mathcal{F}\right]$ *and* $E\left[|z_i|^{2+\delta} |\mathcal{F}\right] < \Delta$ *a.s. for* $i = 1, 2, \ldots$ *and* Δ *arbitrary* $\mathcal{F}$*-measurable, where* $\Delta < \infty$ *a.s. and* $\delta > 0$*. If there is* η*, which is* $\mathcal{F}$*-measurable and such that* $\bar{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 > \eta > 0$ *a.s., then conditional on* $\mathcal{F}$, $\frac{1}{\bar{\sigma}_n\sqrt{n}}\sum_{i=1}^{n}(z_i - E[z_i|\mathcal{F}]) \to^D N(0,1)$ *a.s. Moreover,* $\frac{1}{\bar{\sigma}_n\sqrt{n}}\sum_{i=1}^{n}(z_i - E[z_i|\mathcal{F}]) \to^D N(0,1)$ *(*$\mathcal{F}$*-stably).*

The detailed proofs of Theorems A1 and A2 are provided in [25].

Proof of Theorem 1:

From the definition of the OLS estimator and (1), write

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + \left(\sum_{i=1}^{N} \boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{\mathcal{X}}}\right)^{-1}\left(\sum_{i=1}^{N} \boldsymbol{\mathcal{X}}_i'\varepsilon_i - \bar{\boldsymbol{\mathcal{X}}}'\bar{\varepsilon}\right) \tag{11}$$

By Assumptions C1–C5, conditional on $\mathcal{F}$, unbiasedness is straightforward. Thus, the details are omitted.

We then show that conditional on $\mathcal{F}$, $\bar{\boldsymbol{\mathcal{X}}} - E[\bar{\boldsymbol{\mathcal{X}}}|\mathcal{F}] \to 0$ a.s. Since $\boldsymbol{\mathcal{X}}_i = (\mathbf{Z}_i, \boldsymbol{X}_i)$, write

$$\bar{\boldsymbol{\mathcal{X}}} - E[\bar{\boldsymbol{\mathcal{X}}}|\mathcal{F}] = \left(\bar{\mathbf{Z}} - E[\bar{\mathbf{Z}}|\mathcal{F}],\ \bar{\boldsymbol{V}} - E[\bar{\boldsymbol{V}}|\mathcal{F}] + \boldsymbol{F}_T(\bar{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma})\right),$$

where $\bar{\mathbf{Z}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i$, $\bar{\boldsymbol{V}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{V}_i$ and $\bar{\boldsymbol{\Gamma}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i$.

Assumption C2 implies that the components of $\boldsymbol{\mathcal{X}}_i = (\mathbf{Z}_i, \boldsymbol{X}_i)$ form sequences of independent random variables with finite means and satisfy the conditions for Theorem A1; thus, $\bar{\mathbf{Z}} - E[\bar{\mathbf{Z}}|\mathcal{F}] \to 0$ a.s. and $\bar{\boldsymbol{V}} - E[\bar{\boldsymbol{V}}|\mathcal{F}] \to 0$ a.s. conditional on $\mathcal{F}$. Similarly, we can conclude that $\bar{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma} \to 0$ a.s. conditional on $\mathcal{F}$. Thus, conditional on $\mathcal{F}$, $\bar{\boldsymbol{\mathcal{X}}} - E[\bar{\boldsymbol{\mathcal{X}}}|\mathcal{F}] \to 0$ a.s.

We now focus on $\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i$. Each term in the sum is a $(p+k)\times(p+k)$ matrix. Therefore, let $\boldsymbol{\zeta}_1$ and $\boldsymbol{\zeta}_2$ be arbitrary $(p+k)\times 1$ vectors. Then, $\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\zeta}_1'\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i\boldsymbol{\zeta}_2$ is a sum of independent random variables satisfying the following inequality a.s.:

$$\begin{aligned}
&E\left[\left|\boldsymbol{\zeta}_1'\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i\boldsymbol{\zeta}_2\right|^{1+\delta}|\mathcal{F}\right] \le \|\boldsymbol{\zeta}_1\|^{1+\delta}\|\boldsymbol{\zeta}_2\|^{1+\delta} E\left[\|\boldsymbol{\mathcal{X}}_i\|^{2+2\delta}|\mathcal{F}\right]\\
&\le \|\boldsymbol{\zeta}_1\|^{1+\delta}\|\boldsymbol{\zeta}_2\|^{1+\delta} E\left[(\|(\mathbf{Z}_i, \boldsymbol{V}_i)\| + \|\boldsymbol{F}_T\|\,\|\boldsymbol{\Gamma}_i\|)^{2+2\delta}|\mathcal{F}\right]\\
&\le 2^{1+\delta}\|\boldsymbol{\zeta}_1\|^{1+\delta}\|\boldsymbol{\zeta}_2\|^{1+\delta}\left(E\left[\|(\mathbf{Z}_i, \boldsymbol{V}_i)\|^{2+2\delta}|\mathcal{F}\right] + \|\boldsymbol{F}_T\|^{2+2\delta} E\left[\|\boldsymbol{\Gamma}_i\|^{2+2\delta}|\mathcal{F}\right]\right),
\end{aligned}$$

where the last terms is uniformly bounded a.s. because of Assumptions C2 and C4. Thus, conditional on $\mathcal{F}$, $\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i - \frac{1}{N}\sum_{i=1}^{N}E[\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i|\mathcal{F}] \to 0$ a.s. Further, notice that by Assumption C6, $E\left[\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{\mathcal{X}}}|\mathcal{F}\right]$, is a.s. positive definite uniformly.

Similar to the above, we can show $\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\varepsilon}_i - \bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{\varepsilon}} \to \mathbf{0}$ a.s. Thus, the result is proven.

Proof of Theorem 2:

To prove conditional normality, we write

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) = \left(\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{\mathcal{X}}}\right)^{-1}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\varepsilon}_i - \sqrt{N}\bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{\varepsilon}}\right). \tag{12}$$

We know already that conditional on $\mathcal{F}$

$$\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{\mathcal{X}}} - E\left[\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{\mathcal{X}}}\Big|\mathcal{F}\right] \to \mathbf{0} \quad \text{a.s.}$$

We now focus on $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\left(\boldsymbol{\varepsilon}_i - \bar{\boldsymbol{\varepsilon}}\right)$ and write

$$\begin{aligned}
\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\varepsilon}_i - \sqrt{N}\bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{\varepsilon}} \quad &= \quad \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i - E\left[\bar{\boldsymbol{\mathcal{X}}}|\mathcal{F}\right]\right)'\left(\boldsymbol{\varepsilon}_i - \boldsymbol{F}_T\boldsymbol{\gamma}\right)\\
&\quad + \left(\bar{\boldsymbol{\mathcal{X}}} - E\left[\bar{\boldsymbol{\mathcal{X}}}|\mathcal{F}\right]\right)'\sqrt{N}\left(\bar{\boldsymbol{\varepsilon}} - \boldsymbol{F}_T\boldsymbol{\gamma}\right).
\end{aligned}$$

We will now show that the last term can be neglected. In fact, we already know that $\bar{\boldsymbol{\mathcal{X}}} - E\left[\boldsymbol{\mathcal{X}}|\mathcal{F}\right] \to 0$ a.s. Thus, we need to prove that conditional on $\mathcal{F}$

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Sigma}_{\varepsilon_i} + \boldsymbol{F}_T\left(\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Sigma}_{\gamma_i}\right)\boldsymbol{F}_T'\right)^{-1/2}\left(\bar{\boldsymbol{\varepsilon}} - \boldsymbol{F}_T\boldsymbol{\gamma}\right) \to_D N\left(\mathbf{0}, \boldsymbol{I}_T\right). \tag{13}$$

Let $\boldsymbol{\kappa}_i = \boldsymbol{\varepsilon}_i - \boldsymbol{F}_T\boldsymbol{\gamma}$ 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 $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\kappa}_i$. Let $\boldsymbol{\zeta}$ be an arbitrary $T \times 1$ vector and focus on $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\zeta}'\boldsymbol{\kappa}_i$. We will now verify the conditions for the validity of Theorem A2. Firstly, note that

$$E\left[\boldsymbol{\zeta}'\boldsymbol{\kappa}_i|\mathcal{F}\right] = 0 \quad \text{and} \quad E\left[\left(\boldsymbol{\zeta}'\boldsymbol{\kappa}_i\right)^2|\mathcal{F}\right] = \boldsymbol{\zeta}'\left(\boldsymbol{\Sigma}_{\varepsilon_i} + \boldsymbol{F}_T\boldsymbol{\Sigma}_{\gamma_i}\boldsymbol{F}_T'\right)\boldsymbol{\zeta}.$$

Notice also that

$$\begin{aligned} E\left[\left|\boldsymbol{\zeta}'\boldsymbol{\kappa}_i\right|^{2+\delta}|\mathcal{F}\right] &\leq \|\boldsymbol{\zeta}\|^{2+\delta} E\left[\left(\|\boldsymbol{\varepsilon}_i\| + \|\boldsymbol{F}_T(\boldsymbol{\gamma}_i - \boldsymbol{\gamma})\|\right)^{2+\delta}|\mathcal{F}\right] \\ &\leq \|\boldsymbol{\zeta}\|_2^{2+\delta} 2^{1+\delta}\left(E\left[\|\boldsymbol{\varepsilon}_i\|^{2+\delta}|\mathcal{F}\right] + \|\boldsymbol{F}_T\|_2^{2+\delta} E\left[\|\boldsymbol{\gamma}_i - \boldsymbol{\gamma}\|_2^{2+\delta}|\mathcal{F}\right]\right). \end{aligned}$$

Based on the above, (13) has been proven. Thus,

$$\left(\bar{\boldsymbol{\mathcal{X}}} - E\left[\bar{\boldsymbol{\mathcal{X}}}|\mathcal{F}\right]\right)'\sqrt{N}\left(\bar{\boldsymbol{\varepsilon}} - \boldsymbol{F}_T\boldsymbol{\gamma}\right) \quad \text{a.s.}$$

Similar to the proof of (13), we can show that:

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i - E\left[\bar{\boldsymbol{\mathcal{X}}}|\mathcal{F}\right]\right)'\left(\boldsymbol{\varepsilon}_i - \boldsymbol{F}_T\boldsymbol{\gamma}\right) \to_D \boldsymbol{B}^{-1}(\boldsymbol{F}_T)\boldsymbol{C}^{1/2}(\boldsymbol{F}_T)N(\boldsymbol{0}, \boldsymbol{I}_{p+k}).$$

Thus, Theorem 2 is proven.

Lemma A1 *Suppose that $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \to \boldsymbol{0}$ a.s. Given Assumptions N1, CM2, N3, CM4 and C5, the following results hold conditional on $\mathcal{F}$ as $N \to \infty$:*

1. $\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\varepsilon}_i - \bar{\boldsymbol{\varepsilon}}\right)\left(\boldsymbol{\varepsilon}_i - \bar{\boldsymbol{\varepsilon}}\right)'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right) \to \boldsymbol{C}\left(\boldsymbol{F}_T\right)$ *a.s.*
2. $\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\varepsilon}_i - \bar{\boldsymbol{\varepsilon}}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right) \to \boldsymbol{0}$ *a.s.*
3. $\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right) \to \boldsymbol{0}$ *a.s.*

If also Assumption H5 and HN7 hold, then

1. $\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)\boldsymbol{\eta}_i'\boldsymbol{\mathcal{X}}_i'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right) \to \boldsymbol{0}$ *a.s.*
2. $\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\varepsilon}_i - \bar{\boldsymbol{\varepsilon}}\right)\boldsymbol{\eta}_i'\boldsymbol{\mathcal{X}}_i'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right) \to \boldsymbol{0}$ *a.s.*
3. $\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right) \to \boldsymbol{C}^*\left(\boldsymbol{F}_T\right)$ *a.s.*

Proof of Lemma A1:

The proofs are similar to those given for Theorem 1, thus omitted.

Proof of Lemma 1.

Write

$$\hat{\boldsymbol{C}} = \frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\varepsilon}_i - \bar{\boldsymbol{\varepsilon}} + \left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)\right)\left(\boldsymbol{\varepsilon}_i - \bar{\boldsymbol{\varepsilon}} + \left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)\right)'\left(\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}\right).$$

Then, the proof follows from Results 1–3 of Lemma A1.

Proof of Theorem 3:

Write

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + \left(\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{\mathcal{X}}_i - \bar{\boldsymbol{\mathcal{X}}}'\bar{\boldsymbol{\mathcal{X}}}\right)^{-1}\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\left(\boldsymbol{F}_T\boldsymbol{\gamma}_i - \boldsymbol{F}_T\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_i - \bar{\boldsymbol{\varepsilon}}\right).$$

Under Assumption D5, the factor loadings are not independent. This affects only the term

$$\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mathcal{X}}_i'\boldsymbol{F}_T\boldsymbol{\gamma}_i = \frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix}\boldsymbol{Z}_i' \\ \boldsymbol{V}_i'\end{pmatrix}\boldsymbol{F}_T\boldsymbol{\gamma}_i + \begin{pmatrix}\boldsymbol{0} \\ \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i'\boldsymbol{F}_T'\boldsymbol{F}_T\boldsymbol{\gamma}_i\end{pmatrix}.$$

Let $\boldsymbol{\zeta}$ be an arbitrary $(p+k)\times 1$ vector, and consider $\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\zeta}'\begin{pmatrix}\boldsymbol{Z}_i' \\ \boldsymbol{V}_i'\end{pmatrix}\boldsymbol{F}_T\boldsymbol{\gamma}_i$. Then,

$$E\left[\boldsymbol{\zeta}'\begin{pmatrix}\boldsymbol{Z}_i' \\ \boldsymbol{V}_i'\end{pmatrix}\boldsymbol{F}_T\boldsymbol{\gamma}_i|\mathcal{F}\right] = \boldsymbol{\zeta}'E\left[\begin{pmatrix}\boldsymbol{Z}_i' \\ \boldsymbol{V}_i'\end{pmatrix}|\mathcal{F}\right]\boldsymbol{F}_T\boldsymbol{\gamma}$$

and

$$E\left[\left(\zeta'\left(\begin{array}{c}\boldsymbol{Z}_i'\\ \boldsymbol{V}_i'\end{array}\right)\boldsymbol{F}_T\gamma_i\right)^{1+\delta}|\mathcal{F}\right]\leq\|\zeta\|^{1+\delta}E\left[\|(\boldsymbol{Z}_i,\boldsymbol{V}_i)\|^{1+\delta}|\mathcal{F}\right]\cdot E\left[\|\boldsymbol{F}_T\gamma_i\|^{1+\delta}|\mathcal{F}\right],$$

which is uniformly bounded by Assumptions C2 and C3. Thus, we can conclude that

$$\frac{1}{N}\sum_{i=1}^{N}\left(\begin{array}{c}\boldsymbol{Z}_i'\\ \boldsymbol{V}_i'\end{array}\right)\boldsymbol{F}_T\gamma_i-\left(\begin{array}{c}E\left[\bar{\boldsymbol{Z}}|\mathcal{F}\right]'\boldsymbol{F}_T\gamma\\ E\left[\bar{\boldsymbol{V}}|\mathcal{F}\right]'\boldsymbol{F}_T\gamma\end{array}\right)\to\mathbf{0}\quad\text{a.s.}$$

conditional on $\mathcal{F}$. Similarly, it is easy to show that conditional on $\mathcal{F}$

$$\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i'\boldsymbol{F}_T'\boldsymbol{F}_T\gamma_i-\frac{1}{N}\sum_{i=1}^{N}E\left[\boldsymbol{\Gamma}_i'\boldsymbol{F}_T'\boldsymbol{F}_T\gamma_i|\mathcal{F}\right]\to\mathbf{0}\quad\text{a.s.}$$

Thus, conditional on $\mathcal{F}$, $\hat{\boldsymbol{\theta}}\to\boldsymbol{\theta}_0+\boldsymbol{B}^{-1}\left(\boldsymbol{F}_T\right)\left(\begin{array}{c}\mathbf{0}\\ \boldsymbol{\phi}\left(\boldsymbol{F}_T\right)\end{array}\right)$ a.s.

Proof of Theorem 4:

The proof is the same as those given for Theorems 1 and 2, and it is omitted.

Proof of Lemma 2:

Write

$$\begin{aligned}\hat{C} &= \frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)'\left(\varepsilon_i-\bar{\varepsilon}+\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\right)\right)\left(\varepsilon_i-\bar{\varepsilon}+\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\right)\right)'\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)\\ &\quad+\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)'\left(\varepsilon_i-\bar{\varepsilon}+\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\right)\right)\boldsymbol{\eta}_i'\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)\\ &\quad+\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)\boldsymbol{\eta}_i\left(\varepsilon_i-\bar{\varepsilon}+\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\right)\right)'\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)\\ &\quad+\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)'\left(\boldsymbol{\mathcal{X}}_i-\bar{\boldsymbol{\mathcal{X}}}\right)\end{aligned}$$

The first line is proven in Lemma 1. The other lines follow from Lemma A1.

References

1. Andrews, D.W.K. Cross-section regression with common shocks. *Econometrica* **2005**, *73*, 1551–1585.
2. Bai, J. Panel data models with interactive fixed effects. *Econometrica* **2009**, *77*, 1229–1279.
3. Case, A.C. Spatial patterns in household demand. *Econometrica* **1991**, *59*, 953–965.
4. Conley, T. G. GMM estimation with cross sectional dependence. *J. Econom.* **1999**, *92*, 1–45.
5. Pesaran, M.H. Estimation and inference in large heterogeneous panels with a multi-factor error structure. *Econometrica* **2006**, *74*, 967–1012.
6. Woutersen, T. *Robustness against Incidental Parameters*; Working paper; The University of Western Ontario Department of Economics: London, Canada, 2004.
7. Evans, D.; Tandon, A.; Murray, C.; Lauer, J. *The Comparative Efficiency of National Health Systems in Producing Health: An Analysis of 191 Countries*; GPE Discussion Paper No. 29; World Health Organization: Geneva, Switzerland, 2000.
8. De Boeck, M.; Slok, T. Interpreting real exchange rate movements in transition countries. *J. Int. Econ.* **2006**, *68*, 368–383.
9. Eberhardt, M.; Helmers, C.; Strauss, H. Do spillovers matter when estimating private returns to R & D? *Rev. Econ. Stat.* **2013**, *95*, 436–448.
10. Ahn, S.C.; Lee, Y. H.; Schmidt, P. GMM estimation of linear panel data models with time-varying individual effects. *J. Econom.* **2001**, *101*, 219–255.

11. Robertson, D.; Symons, J. Maximum likelihood factor analysis with rank-deficient sample covariance matrices. *J. Multivar. Anal.* **2007**, *98*, 813–828.
12. Coakley, J.; Fuertes, A.M.; Smith, R. *A Principal Components Approach to Cross-Section Dependence in Panels*; Discussion Paper 01/2002; Birkbeck College: London, UK, 2002.
13. Sarafidis, V.; Wansbeek, T. Cross-sectional dependence in panel data analysis. *Econom. Rev.* **2012**, *31*, 483–531.
14. Kuersteiner, G.M.; Prucha, I.R. Limit theory for panel data models with cross sectional dependence and sequential exogeneity. *J. Econom.* **2013**, *174*, 107–126.
15. Hall, P.; Heyde, C.C. *Martingale Limit Theory and Its Application*; Academic Press: New York, NY, USA; London, UK, 1980.
16. Kao, C.W.; Trapani, L.; Urga, G. Asymptotics for panel models with common shocks. *Econom. Rev.* **2012**, *31*, 390–439.
17. Cabrera, M.O.; Rosalsky, A.; Volodin, A. Some theorems on conditional mean convergence and conditional almost sure convergence for randomly weighted sums of dependent random variables. *Test* **2012**, *21*, 369–385.
18. Majerek, D.; Nowak, W.; Zięba, W. Conditional strong law of large numbers. *Int. J. Pure Appl. Math.* **2005**, *20*, 143–156.
19. Prakasa Rao, B.L.S. Conditional independence, conditional mixing and conditional association. *Ann. Inst. Stat. Math.* **2009**, *61*, 441–460.
20. White, H. *Asymptotic Theory for Econometricians*; Academic Press: Orlando, FL, USA; London, UK, 1984.
21. Dedecker, J.; Merlevede, F. Necessary and sufficient conditions for the conditional central limit theorem. *Ann. Probab.* **2002**, *30*, 1044–1081.
22. Grzenda, W.; Zięba, W. Conditional central limit theorem. *Int. Math. Forum* **2008**, *31*, 1521–1528.
23. Rényi, A. On stable sequences of events. *Indian J. Stat. Ser. A* **1963**, *25*, 293–302.
24. Yuan, D.M.; Wei, L.R.; Lei, L. Conditional central limit theorems for a sequence of conditional independent random variables. *J. Korean Math. Soc.* **2014**, *51*, 1–15.
25. Forchini, G.; Jiang, B.; Peng, B. *Common shocks in panels with endogenous regressors*; Working paper 08/15; Department of Econometrics and Business Statistics, Monash University: Clayton, Victoria, Australia, 2015.
26. Phillips, P.C.B. Conditional and unconditional statistical independence. *J. Econom.* **1988**, *38*, 341–348.
27. Coakley, J.; Fuertes, A.M.; Smith, R. Unobserved heterogeneity in panel time series models. *Comput. Stat. Data Anal.* **2006**, *50*, 2361–2380.