



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/139172/>

Version: Accepted Version

Article:

Kapetanios, G., Serlenga, Laura and Shin, Yongcheol (2021) Estimation and Inference for Multi-dimensional Heterogeneous Panel Datasets with Hierarchical Multi-factor Error Structure. *Journal of Econometrics*. pp. 504-531. ISSN: 0304-4076

<https://doi.org/10.1016/j.jeconom.2020.04.011>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Estimation and Inference for Multi-dimensional Heterogeneous Panel Datasets with Hierarchical Multi-factor Error Structure*

George Kapetanios
King's College London

Laura Serlenga
University of Bari

Yongcheol Shin
University of York

Abstract

Given the growing availability of large datasets and following recent research trends on multi-dimensional modelling, we develop three dimensional (3D) panel data models with hierarchical error components that allow for strong cross-sectional dependence through unobserved heterogeneous global and local factors. We propose consistent estimation procedures by extending the common correlated effects (CCE) estimation approach proposed by Pesaran (2006). The standard CCE approach needs to be modified in order to account for the hierarchical factor structure in 3D panels. Further, we provide the associated asymptotic theory, including new nonparametric variance estimators. The validity of the proposed approach is confirmed by Monte Carlo simulation studies. We also demonstrate the empirical usefulness of the proposed approach through an application to a 3D panel gravity model of bilateral export flows.

JEL Classification: C13, C33, F14, F45.

Key Words: Multi-dimensional Panel Data Models, Cross-sectional Error Dependence, Unobserved Heterogeneous Global and Local Factors, Multilateral Resistance, The Gravity Model of Bilateral Export Flows.

*We are mostly grateful to the guest-editor, Vasilis Sarafidis, and two anonymous referees for their helpful comments. We are also grateful to In Choi, Terrence Cheung, Matthew Greenwood-Nimmo, Taehwan Kim, Younghoon Lee, Hosung Lim, Hashem Pesaran, Younggi Shin, and the conference delegates at the AMES Conference, Hong Kong, 2-6 June 2017, the 23rd International Panel Data Conference, Thessaloniki, 7-8 July 2017 and the ESEM Conference, Lisbon, 20-25 August 2017, and seminar participants at Universities of Adelaide, Melbourne, Queensland, Sogang, Technology Sydney, and Yonsei, and the Bank of Korea for their helpful comments. The usual disclaimer applies.

1 Introduction

In recent years there has been a rapid emergence of big datasets, that has fueled a burgeoning literature on multi-dimensional panel data econometric analysis. These arise in a variety of ways. For example, they can arise by refining individual unit information (e.g., sectors within countries), by considering origin-destination flow data (e.g., trade and migration flows), by matching and combining different cross-sectional datasets (e.g., employer-employee databases) and by multi-dimensional interactive data (e.g., social networking information). Such a multi-dimensional dataset allows and, possibly, requires a rich treatment of error components, which can accommodate heterogeneous fixed effects, spatial or cross-sectional interdependence, endogeneity (or reverse causality) and other related features. All these issues eventually call for new econometric methods, though the proper formulation of multi-dimensional models is a non-trivial and challenging issue.¹

A few studies have attempted to extend the two-way error components specification to the multi-dimensional setting. Baltagi, Egger, and Pfaffermayr (2003) propose such a specification with fixed exporter-time, importer-time, and country-pair effects, referred to as the country-time fixed effects (CTFE) specification. This approach has been applied to a number of bilateral flows, such as trade, foreign direct investment, or migration flows, between countries or regions (see, e.g., Feenstra (2004) and Baltagi, Egger, and Pfaffermayr (2015)). Balazsi, Matyas, and Wansbeek (2018) generalise the three-dimensional (3D) within estimator for the 3D panel data models with CTFEs while Balazsi, Baltagi, Matyas, and Pus (2016) propose a sequence of GLS estimators for 3D models under the assumption that the unobserved effects are random and mutually uncorrelated. They discuss several extensions to deal with dynamic and unbalanced panels, which are often characterised with a high percentage of zero observations and which have different numbers of observations across different dimensions such as limited observations in one dimension (e.g., regions/sectors or time periods) and very large numbers of observations in other dimensions (e.g., individuals).

However, most existing studies have so far neglected an important issue of controlling cross-sectional error dependence in 3D models, despite pervasive evidence of the presence of strong cross-sectional dependence (CSD) in 2D panels (e.g., Pesaran (2015)). In this regard, Kapetanios, Mastromarco, Serlenga, and Shin (2017) (KMSS) propose to combine unobserved heterogeneous global factors with the CTFE specification and show that standard estimator are biased in the presence of nonzero correlation between factors and regressors. KMSS propose consistent estimation procedures by extending the common correlated effects (CCE) estimation in 2D panels advanced by Pesaran (2006). Though KMSS is the first paper to

¹Recently, Matyas (2017) collects and presents a number of the econometric foundations and applications of multi-dimensional panels, that aim to fill the gap in the literature.

accommodate strong CSD within multi-dimensional panels, they do not consider hierarchical factor structures that are, as we explore, of relevance for multi-dimensional panel datasets.

The above discussion suggests that it is important to develop more realistic specifications for accommodating CSD within multi-dimensional datasets and that such specifications should involve several layers of components to account for effects along different dimensions. Similar issues have been examined extensively in the international business cycle literature, especially in relation to the separate identification of regional and global factors. Kose, Otrok, and Whiteman (2003) propose a Bayesian method based on state space representations. Alternatively, Breitung and Eickmeier (2016) suggest a two-step approach based on canonical correlations for multi-level factor models, which are subject to blocks of zero restrictions on the factor loadings matrix.

Following this research trend, we propose a 3D error components specification with hierarchical multi-factor structures, which permits us to accommodate a more flexible degree of CSD along different panel dimensions. Specifically, our approach can be regarded as a generalisation of the CTFE model, which has been applied to the structural gravity trade models. Our approach also shares similar motivations with the spatial econometric methodologies that explicitly introduce spatial structures for origin and destination effects, e.g., LeSage and Llano (2016). We also allow all factor hierarchies to be correlated with the regressors in order to deal with potentially important endogeneity issues.

After providing this hierarchical error component structure, we develop consistent estimation procedures by extending the CCE approach of Pesaran (2006). In order to account for the hierarchical factor structure of the 3D model, we need to modify the CCE approach by augmenting the model with cross-section averages of dependent variable and regressors, respectively, over single and double dimensions of cross-section units. These local and global averages are shown to provide valid proxies for unobserved factor hierarchies. We have also provided the associated asymptotic theory. We conduct Monte Carlo experiments and find evidence that the small sample properties of the 3D-CCE estimators are in line with theoretical predictions. On the contrary, the conventional CCE estimator tends to display severe biases and unreliable inference.

We apply the 3DCCE estimation techniques to a trade dataset over the period 1970-2013 (44 years), with two alternative control groups: the 210 country-pairs among 15 EU countries consisting of 11 Euro plus 4 control non-Euro countries, and the 342 country-pairs among 19 countries including 4 additional non-EU countries. Focussing on the impacts of CEE and EMU dummies (equal to one when both countries belong to the European Community or when both adopt the same currency), we find in the case with 15 EU countries that both impacts on export flows are rather modest at 5 and 3%, which are well below those reported (7 to 10%) in the 2D panels by recent studies, e.g., Mastromarco, Serlenga, and Shin (2016),

and Gunnella, Mastromarco, Serlenga, and Shin (2015). Next, for the enlarged dataset with the larger control group of 8 countries, we observe that the impacts of EMU and CEE on exports are reduced from 3% to 1.5% and from 5% to 3%, respectively. These smaller effects might indicate a trade diversion effect between the Euro and non-Euro area. Overall, we may conclude that the trade increase within the Euro area may partly reflect a continuation of a long-run historical trend rather than specific EU economic integration policies.

The paper proceeds in 6 Sections. Section 2 introduces the 3D panel data models with three-way error components, that follow a hierarchical multi-factor structure. Section 3 develops consistent estimation procedures, referred to as the 3DCCE estimator. We derive the asymptotic distributions of both mean group and pooled 3DCCE estimators. Section 4 provides Monte Carlo simulation results. We present the empirical results for the gravity model of EU export flows in Section 5. Section 6 concludes. Mathematical proofs are relegated to an Appendix while the construction of the dataset is summarised in the Data Appendix. Additional Monte Carlo simulation results are provided in the Supplement.

2 The Model

We consider the three-dimensional (3D) heterogeneous panel data model given by

$$y_{ijt} = \beta'_{ij} \mathbf{x}_{ijt} + \delta'_{ij} \mathbf{d}_t + e_{ijt}, \quad i = 1, \dots, N, \quad j = 1, \dots, N, \quad t = 1, \dots, T, \quad (1)$$

where y_{ijt} is the dependent variable observed across three indices, i being the origin cross-section unit, j the destination unit at period t (say, the export from country i to country j at time t), \mathbf{x}_{ijt} is the $m_x \times 1$ vector of covariates observed across three indices and \mathbf{d}_t is the $m_d \times 1$ vector of observed common effects including deterministic components such as constant and trend. β_{ij} and δ_{ij} are the $m_x \times 1$ and $m_d \times 1$ vectors of parameters.

We extend the multi-factor structure analysed by Pesaran (2006) into 3D panels. In particular, we allow e_{ijt} to follow the hierarchical multi-factor structure:

$$e_{ijt} = \gamma'_{ij} \mathbf{f}_t + \gamma'_{oj} \mathbf{f}_{iot} + \gamma'_{io} \mathbf{f}_{ojt} + \varepsilon_{ijt}, \quad (2)$$

where \mathbf{f}_t , \mathbf{f}_{iot} and \mathbf{f}_{ojt} are respectively $m_f \times 1$, $m_{\bullet o} \times 1$ and $m_{o\bullet} \times 1$ vectors of unobserved common effects, γ_{ij} , γ_{oj} and γ_{io} are $m_f \times 1$, $m_{\bullet o} \times 1$ and $m_{o\bullet} \times 1$ vectors of heterogeneous loadings, and ε_{ijt} are idiosyncratic errors.

Unobserved factors, \mathbf{f}_t , \mathbf{f}_{ojt} and \mathbf{f}_{iot} , are likely to be correlated with \mathbf{x}_{ijt} . Thus, we consider the following data generating process for \mathbf{x}_{ijt} :

$$\mathbf{x}_{ijt} = \mathcal{D}_{ij} \mathbf{d}_t + \Gamma_{ij} \mathbf{f}_t + \Gamma_{oj} \mathbf{f}_{iot} + \Gamma_{io} \mathbf{f}_{ojt} + \mathbf{v}_{ijt}, \quad (3)$$

where \mathcal{D}_{ij} is the $(m_x \times m_d)$ parameter matrix on observed common effects, $\mathbf{\Gamma}_{ij}$, $\mathbf{\Gamma}_{oj}$ and $\mathbf{\Gamma}_{io}$ are $(m_x \times m_f)$, $(m_x \times m_{\bullet o})$, $(m_x \times m_{o\bullet})$ factor loading matrices, and \mathbf{v}_{ijt} are the idiosyncratic errors.

Combining (1)-(3), we have:

$$\mathbf{z}_{ijt} = \begin{pmatrix} y_{ijt} \\ \mathbf{x}_{ijt} \end{pmatrix} = \mathbf{\Xi}_{ij} \mathbf{d}_t + \mathbf{\Phi}_{ij} \mathbf{f}_t + \mathbf{\Phi}_{i,oj} \mathbf{f}_{iot} + \mathbf{\Phi}_{io,j} \mathbf{f}_{ojt} + \mathbf{u}_{ijt} \quad (4)$$

where

$$\mathbf{\Xi}_{ij} = \begin{pmatrix} \boldsymbol{\delta}'_{ij} + \boldsymbol{\beta}'_{ij} \mathcal{D}_{ij} \\ \mathcal{D}_{ij} \end{pmatrix}, \mathbf{\Phi}_{ij} = \begin{pmatrix} \boldsymbol{\gamma}'_{ij} + \boldsymbol{\beta}'_{ij} \mathbf{\Gamma}_{ij} \\ \mathbf{\Gamma}_{ij} \end{pmatrix}, \mathbf{\Phi}_{i,oj} = \begin{pmatrix} \boldsymbol{\gamma}'_{oj} + \boldsymbol{\beta}'_{ij} \mathbf{\Gamma}_{oj} \\ \mathbf{\Gamma}_{oj} \end{pmatrix}, \mathbf{\Phi}_{io,j} = \begin{pmatrix} \boldsymbol{\gamma}'_{io} + \boldsymbol{\beta}'_{ij} \mathbf{\Gamma}_{io} \\ \mathbf{\Gamma}_{io} \end{pmatrix} \quad (5)$$

$$\mathbf{u}_{ijt} = \begin{pmatrix} \varepsilon_{ijt} + \boldsymbol{\beta}'_{ij} \mathbf{v}_{ijt} \\ \mathbf{v}_{ijt} \end{pmatrix}.$$

Notice that the ranks of $\mathbf{\Phi}_{ij}$, $\mathbf{\Phi}_{i,oj}$ and $\mathbf{\Phi}_{io,j}$ are determined by the ranks of the following matrices:

$$\tilde{\mathbf{\Gamma}}_{ij} = \begin{pmatrix} \boldsymbol{\gamma}'_{ij} \\ \mathbf{\Gamma}_{ij} \end{pmatrix}_{(m_x+1) \times m_f}, \quad \tilde{\mathbf{\Gamma}}_{io} = \begin{pmatrix} \boldsymbol{\gamma}'_{io} \\ \mathbf{\Gamma}_{io} \end{pmatrix}_{(m_x+1) \times m_{\bullet o}}, \quad \tilde{\mathbf{\Gamma}}_{oj} = \begin{pmatrix} \boldsymbol{\gamma}'_{oj} \\ \mathbf{\Gamma}_{oj} \end{pmatrix}_{(m_x+1) \times m_{o\bullet}}.$$

Throughout we assume that $\|\mathbf{\Xi}_{ij}\|$, $\|\mathbf{\Phi}_{ij}\|$, $\|\mathbf{\Phi}_{i,oj}\|$ and $\|\mathbf{\Phi}_{io,j}\|$ or their expectations are bounded (see Assumption 3 below).

For each (i, j) , we rewrite (1) and (4) in matrix notation:

$$\mathbf{y}_{ij} = \mathbf{X}_{ij} \boldsymbol{\beta}_{ij} + \mathbf{D} \boldsymbol{\delta}_{ij} + \mathbf{F} \boldsymbol{\gamma}_{ij} + \mathbf{F}_{io} \boldsymbol{\gamma}_{oj} + \mathbf{F}_{oj} \boldsymbol{\gamma}_{io} + \boldsymbol{\varepsilon}_{ij}, \quad (6)$$

$$\mathbf{z}_{ij} = \mathbf{D} \mathbf{\Xi}_{ij} + \mathbf{F} \mathbf{\Phi}_{ij} + \mathbf{F}_{io} \mathbf{\Phi}_{i,oj} + \mathbf{F}_{oj} \mathbf{\Phi}_{io,j} + \mathbf{u}_{ij}, \quad (7)$$

where

$$\mathbf{y}_{ij} = \begin{bmatrix} y_{ij1} \\ \vdots \\ y_{ijT} \end{bmatrix}_{T \times 1}, \quad \mathbf{X}_{ij} = \begin{bmatrix} \mathbf{x}'_{ij1} \\ \vdots \\ \mathbf{x}'_{ijT} \end{bmatrix}_{T \times m_x}, \quad \mathbf{D} = \begin{bmatrix} \mathbf{d}'_1 \\ \vdots \\ \mathbf{d}'_T \end{bmatrix}_{T \times m_d}, \quad \mathbf{z}_{ij} = \begin{bmatrix} \mathbf{z}'_{ij1} \\ \vdots \\ \mathbf{z}'_{ijT} \end{bmatrix}_{T \times (m_x+1)}, \quad (8)$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{f}'_1 \\ \vdots \\ \mathbf{f}'_T \end{bmatrix}_{T \times m_f}, \quad \mathbf{F}_{io} = \begin{bmatrix} \mathbf{f}'_{io1} \\ \vdots \\ \mathbf{f}'_{ioT} \end{bmatrix}_{T \times m_{\bullet o}}, \quad \mathbf{F}_{oj} = \begin{bmatrix} \mathbf{f}'_{oj1} \\ \vdots \\ \mathbf{f}'_{ojT} \end{bmatrix}_{T \times m_{o\bullet}}, \quad \boldsymbol{\varepsilon}_{ij} = \begin{bmatrix} \varepsilon_{ij1} \\ \vdots \\ \varepsilon_{ijT} \end{bmatrix}_{T \times 1}, \quad \mathbf{u}_{ij} = \begin{bmatrix} \mathbf{u}'_{ij1} \\ \vdots \\ \mathbf{u}'_{ijT} \end{bmatrix}_{T \times (m_x+1)}$$

We develop the estimation and inference theory for $E(\boldsymbol{\beta}_{ij}) = \boldsymbol{\beta}$, but we also discuss the circumstance under which the individual coefficients, $\boldsymbol{\beta}_{ij}$ can be consistently estimated. We make the following assumptions:

Assumption 1. Common Effects: The $(m_d + m_f + Nm_{\bullet o} + Nm_{o\bullet}) \times 1$ vector of common factors $\mathbf{g}_t = (\mathbf{d}'_t, \mathbf{f}'_t, \mathbf{f}'_{1ot}, \dots, \mathbf{f}'_{Not}, \mathbf{f}'_{o1t}, \dots, \mathbf{f}'_{oNt})'$, is covariance stationary with absolute

summable autocovariances, distributed independently of $\varepsilon_{ijt'}$ and $\mathbf{v}_{ijt'}$ for all N, i, j, t and t' . $\mathbf{f}_t, \mathbf{f}_{1ot}, \dots, \mathbf{f}_{Not}, \mathbf{f}_{o1t}, \dots, \mathbf{f}_{oNt}$ are zero mean process and are mutually uncorrelated.

Assumption 2. Individual-specific Errors: ε_{ijt} and $\mathbf{v}_{ijt'}$ are distributed independently for all i, j, t and t' , and they are distributed independently of \mathbf{x}_{ijt} and \mathbf{d}_t .

Assumption 3. Factor Loadings: The factor loadings are independently and identically distributed across (i, j) , and of the individual-specific errors ε_{ijt} and \mathbf{v}_{ijt} , the common factors \mathbf{g}_t for all i, j and t , with finite means and finite variances. In particular, we have:

$$\gamma_{ij} = \gamma_{\bullet\bullet} + \boldsymbol{\eta}_{ij}, \quad \gamma_{i\circ} = \gamma_{\bullet\circ} + \boldsymbol{\eta}_{i\circ}, \quad \gamma_{\circ j} = \gamma_{\bullet\bullet} + \boldsymbol{\eta}_{\circ j}, \quad (9)$$

$$\Gamma_{ij} = \Gamma_{\bullet\bullet} + \boldsymbol{\xi}_{ij}, \quad \Gamma_{i\circ} = \Gamma_{\bullet\circ} + \boldsymbol{\xi}_{i\circ}, \quad \Gamma_{\circ j} = \Gamma_{\bullet\bullet} + \boldsymbol{\xi}_{\circ j}, \quad (10)$$

where $\boldsymbol{\eta}_{ij} \sim iid(0, \boldsymbol{\Omega}_{\boldsymbol{\eta}_{ij}})$, $\boldsymbol{\xi}_{ij} \sim iid(0, \boldsymbol{\Omega}_{\boldsymbol{\xi}_{ij}})$, $\boldsymbol{\eta}_{i\circ} \sim iid(0, \boldsymbol{\Omega}_{\boldsymbol{\eta}_{i\circ}})$, $\boldsymbol{\xi}_{i\circ} \sim iid(0, \boldsymbol{\Omega}_{\boldsymbol{\xi}_{i\circ}})$, $\boldsymbol{\eta}_{\circ j} \sim iid(0, \boldsymbol{\Omega}_{\boldsymbol{\eta}_{\circ j}})$ and $\boldsymbol{\xi}_{\circ j} \sim iid(0, \boldsymbol{\Omega}_{\boldsymbol{\xi}_{\circ j}})$. Further, $\|\gamma_{\bullet\bullet}\| < K$, $\|\gamma_{\bullet\circ}\| < K$, $\|\gamma_{\circ\bullet}\| < K$, $\|\Gamma_{\bullet\bullet}\| < K$, $\|\Gamma_{\bullet\circ}\| < K$, and $\|\Gamma_{\circ\bullet}\| < K$ for some positive constant $K < \infty$.

Assumption 4. Random Slope Coefficients: β_{ij} follow the random coefficient specification:

$$\beta_{ij} = \boldsymbol{\beta} + \boldsymbol{\nu}_{i\circ} + \boldsymbol{\nu}_{\circ j} + \boldsymbol{\nu}_{ij} \text{ with } \boldsymbol{\nu}_{i\circ} \sim iid(\mathbf{0}, \boldsymbol{\Omega}_{\boldsymbol{\nu}_{i\circ}}), \boldsymbol{\nu}_{\circ j} \sim iid(\mathbf{0}, \boldsymbol{\Omega}_{\boldsymbol{\nu}_{\circ j}}), \boldsymbol{\nu}_{ij} \sim iid(\mathbf{0}, \boldsymbol{\Omega}_{\boldsymbol{\nu}_{ij}}) \quad (11)$$

where $\|\boldsymbol{\beta}\| < K$ and $\boldsymbol{\nu}_{ij}, \boldsymbol{\nu}_{i\circ}, \boldsymbol{\nu}_{\circ j}$ are distributed independently of one another, and of $\gamma_{ij}, \Gamma_{ij}, \varepsilon_{ijt}, \mathbf{v}_{ijt}$ and \mathbf{g}_t for all i, j and t .

Assumption 5. Identification of β_{ij} and $\boldsymbol{\beta}$: Let

$$\bar{\mathbf{z}}_t = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \mathbf{z}_{ijt}, \quad \bar{\mathbf{z}}_{i\circ t} = \frac{1}{N} \sum_{j=1}^N \mathbf{z}_{ijt}, \quad \bar{\mathbf{z}}_{\circ j t} = \frac{1}{N} \sum_{i=1}^N \mathbf{z}_{ijt}, \quad (12)$$

and $\bar{\mathbf{Z}}_{ij} = (\bar{\mathbf{Z}}, \bar{\mathbf{Z}}_{i\circ}, \bar{\mathbf{Z}}_{\circ j})$ and $\bar{\mathbf{H}}_{ij} = (\mathbf{D}, \bar{\mathbf{Z}}_{ij})$, where

$$\bar{\mathbf{Z}}_{T \times (m_x + 1)} = \begin{bmatrix} \bar{\mathbf{z}}'_1 \\ \vdots \\ \bar{\mathbf{z}}'_T \end{bmatrix}, \quad \bar{\mathbf{Z}}_{i\circ T \times (m_x + 1)} = \begin{bmatrix} \bar{\mathbf{z}}'_{i\circ 1} \\ \vdots \\ \bar{\mathbf{z}}'_{i\circ T} \end{bmatrix}, \quad \bar{\mathbf{Z}}_{\circ j T \times (m_x + 1)} = \begin{bmatrix} \bar{\mathbf{z}}'_{\circ j 1} \\ \vdots \\ \bar{\mathbf{z}}'_{\circ j T} \end{bmatrix},$$

Further,²

$$\bar{\mathbf{M}}_{ij} = \mathbf{I}_T - \bar{\mathbf{H}}_{ij} \left(\bar{\mathbf{H}}'_{ij} \bar{\mathbf{H}}_{ij} \right)^{-1} \bar{\mathbf{H}}'_{ij}. \quad (13)$$

(i) Identification of β_{ij} : The $m_x \times m_x$ matrices, $\bar{\boldsymbol{\Psi}}_{ij,T} = T^{-1} (\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij})$ are nonsingular, and $\bar{\boldsymbol{\Psi}}^{-1}_{ij,T}$ have finite second-order moments for all (i, j) .

(ii) Identification of $\boldsymbol{\beta}$: The $m_x \times m_x$ matrix, $N^{-2} \sum_{i=1}^N \sum_{j=1}^N \bar{\boldsymbol{\Psi}}_{ij,T}$ is nonsingular.

²In the case where $(\bar{\mathbf{H}}'_{ij} \bar{\mathbf{H}}_{ij})$ does not have full rank, the following analysis can still be carried out using a generalize inverse as in Pesaran (2006).

Remark 1 *The factors are assumed to have zero mean for simplicity. Any means can be subsumed in δ_{ij} . Further, they are assumed mutually uncorrelated to ensure that cross-sectional averages of local factors converge to zero. This is another simplifying assumption, since some weak cross-sectional dependence across local factors could be allowed, in exact analogy to weak cross sectional dependence across idiosyncratic shocks.*

Remark 2 *The weights are not necessarily unique, but they do not affect the asymptotic results advanced in this paper (see also Pesaran (2006)). We focus, for simplicity, on equal weights, $1/N$. Alternatively, economic distance-based or time-varying measures such as trade weights or input-output shares could be considered (e.g. Pesaran, Schuermann, and Weiner (2004); Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012)). The number of observed factors, m_d and the number of individual-specific regressors, m_x are assumed fixed. The number of unobserved factors, $m = m_f + m_{\bullet\circ} + m_{\circ\bullet}$, is assumed fixed, but need not to be known.*

Using (4), we represent the hierarchical cross-section averages in (12) as follows:

$$\bar{z}_t = \bar{\Xi}_{\circ\circ} d_t + \bar{\Phi}_{\circ\circ} f_t + \bar{\Phi}_{\bullet\circ} f_{\bullet\circ t} + \bar{\Phi}_{\circ\bullet} f_{\circ\bullet t} + \bar{u}_t, \quad (14)$$

$$\bar{z}_{i\circ t} = \bar{\Xi}_{i\circ} d_t + \bar{\Phi}_{i\circ} f_t + \bar{\Phi}_{\bullet\circ} f_{i\circ t} + \bar{\Phi}_{i\circ\bullet} f_{\circ\bullet t} + \bar{u}_{i\circ t}, \quad (15)$$

$$\bar{z}_{\circ j t} = \bar{\Xi}_{\circ j} d_t + \bar{\Phi}_{\circ j} f_t + \bar{\Phi}_{\circ j\bullet} f_{\bullet\circ t} + \bar{\Phi}_{\bullet\circ} f_{\circ j t} + \bar{u}_{\circ j t}, \quad (16)$$

where

$$\begin{aligned} \bar{\Xi}_{\circ\circ} &= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \Xi_{ij}, & \bar{\Xi}_{i\circ} &= \frac{1}{N} \sum_{j=1}^N \Xi_{ij}, & \bar{\Xi}_{\circ j} &= \frac{1}{N} \sum_{i=1}^N \Xi_{ij}, \\ \bar{\Phi}_{\circ\circ} &= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \Phi_{ij}, & \bar{\Phi}_{i\circ} &= \frac{1}{N} \sum_{j=1}^N \Phi_{ij}, & \bar{\Phi}_{\circ j} &= \frac{1}{N} \sum_{i=1}^N \Phi_{ij}, \end{aligned} \quad (17)$$

$$\bar{\Phi}_{\bullet\circ} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \Phi_{i,\circ j}, \quad \bar{\Phi}_{\circ\bullet} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \Phi_{i\circ,j}, \quad (18)$$

$$\bar{u}_t = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N u_{ijt}, \quad \bar{u}_{i\circ t} = \frac{1}{N} \sum_{j=1}^N u_{ijt}, \quad \bar{u}_{\circ j t} = \frac{1}{N} \sum_{i=1}^N u_{ijt}$$

$$f_{\bullet\circ t} = \frac{1}{N} \sum_{j=1}^N f_{\circ j t}, \quad f_{\circ\bullet t} = \frac{1}{N} \sum_{i=1}^N f_{i\circ t}$$

Combining (14)-(16), we have:

$$\bar{z}_{ijt} = \bar{\Xi}_{ij} d_t + \bar{\Phi}_{ij} f_{ijt} + \bar{u}_{ijt}, \quad (19)$$

where

$$\begin{aligned} \bar{\mathbf{z}}_{ijt} &= \begin{bmatrix} \bar{\mathbf{z}}_t \\ \bar{\mathbf{z}}_{i\circ t} \\ \bar{\mathbf{z}}_{\circ j t} \end{bmatrix}, \quad \mathbf{f}_{ijt} = \begin{bmatrix} \mathbf{f}_t \\ \mathbf{f}_{i\circ t} \\ \mathbf{f}_{\circ j t} \end{bmatrix}, \quad \bar{\mathbf{u}}_{ijt} = \begin{bmatrix} \bar{\mathbf{u}}_t + \bar{\Phi}_{\circ\bullet} \mathbf{f}_{\bullet\circ t} + \bar{\Phi}_{\bullet\circ} \mathbf{f}_{\circ\bullet t} \\ \bar{\mathbf{u}}_{i\circ t} + \bar{\Phi}_{i\circ} \mathbf{f}_{\circ\bullet t} \\ \bar{\mathbf{u}}_{\circ j t} + \bar{\Phi}_{\circ j} \mathbf{f}_{\bullet\circ t} \end{bmatrix} \\ \bar{\Xi}_{ij} &= \begin{bmatrix} \bar{\Xi}_{\circ\circ} \\ \bar{\Xi}_{i\circ} \\ \bar{\Xi}_{\circ j} \end{bmatrix}, \quad \bar{\Phi}_{ij} = \begin{bmatrix} \bar{\Phi}_{\circ\circ} & \mathbf{0} & \mathbf{0} \\ \bar{\Phi}_{i\circ} & \bar{\Phi}_{\bullet\bullet} & \mathbf{0} \\ \bar{\Phi}_{\circ j} & \mathbf{0} & \bar{\Phi}_{\bullet\bullet} \end{bmatrix}. \end{aligned} \quad (20)$$

Using (5) and (11), the elements of $\bar{\Phi}_{ij}$ can be represented as follows:

$$\bar{\Phi}_{\circ\circ} = \tilde{\mathbf{B}} \tilde{\Gamma}_{\circ\circ} + \begin{pmatrix} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (\boldsymbol{\nu}_{i\circ} + \boldsymbol{\nu}_{\circ j} + \boldsymbol{\nu}_{ij})' \Gamma_{ij} \\ \mathbf{0} \end{pmatrix} \quad (21)$$

$$\bar{\Phi}_{\bullet\bullet} = \tilde{\mathbf{B}} \tilde{\Gamma}_{\bullet\bullet} + \begin{pmatrix} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (\boldsymbol{\nu}_{i\circ} + \boldsymbol{\nu}_{\circ j} + \boldsymbol{\nu}_{ij})' \Gamma_{\circ j} \\ \mathbf{0} \end{pmatrix} \quad (22)$$

$$\bar{\Phi}_{\bullet\bullet} = \tilde{\mathbf{B}} \tilde{\Gamma}_{\bullet\bullet} + \begin{pmatrix} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (\boldsymbol{\nu}_{i\circ} + \boldsymbol{\nu}_{\circ j} + \boldsymbol{\nu}_{ij})' \Gamma_{i\circ} \\ \mathbf{0} \end{pmatrix} \quad (23)$$

$$\bar{\Phi}_{i\circ} = \tilde{\mathbf{B}} \tilde{\Gamma}_{i\circ} + \begin{pmatrix} \frac{1}{N} \sum_{j=1}^N (\boldsymbol{\nu}_{i\circ} + \boldsymbol{\nu}_{\circ j} + \boldsymbol{\nu}_{ij})' \Gamma_{ij} \\ \mathbf{0} \end{pmatrix} \quad (24)$$

$$\bar{\Phi}_{\circ j} = \tilde{\mathbf{B}} \tilde{\Gamma}_{\circ j} + \begin{pmatrix} \frac{1}{N} \sum_{i=1}^N (\boldsymbol{\nu}_{i\circ} + \boldsymbol{\nu}_{\circ j} + \boldsymbol{\nu}_{ij})' \Gamma_{ij} \\ \mathbf{0} \end{pmatrix} \quad (25)$$

where

$$\tilde{\mathbf{B}} = \begin{pmatrix} 1 & \boldsymbol{\beta}' \\ 0 & \mathbf{I}_k \end{pmatrix}, \quad \tilde{\Gamma}_{\circ\circ} = \begin{pmatrix} \bar{\gamma}'_{\circ\circ} \\ \bar{\Gamma}_{\circ\circ} \end{pmatrix}, \quad \tilde{\Gamma}_{\bullet\bullet} = \begin{pmatrix} \bar{\gamma}'_{\bullet\bullet} \\ \bar{\Gamma}_{\bullet\bullet} \end{pmatrix}, \quad \tilde{\Gamma}_{\bullet\bullet} = \begin{pmatrix} \bar{\gamma}'_{\bullet\bullet} \\ \bar{\Gamma}_{\bullet\bullet} \end{pmatrix}, \quad \tilde{\Gamma}_{i\circ} = \begin{pmatrix} \bar{\gamma}'_{i\circ} \\ \bar{\Gamma}_{i\circ} \end{pmatrix}, \quad \tilde{\Gamma}_{\circ j} = \begin{pmatrix} \bar{\gamma}'_{\circ j} \\ \bar{\Gamma}_{\circ j} \end{pmatrix}$$

and $\bar{\Gamma}_{\circ\circ}$, $\bar{\gamma}_{\circ\circ}$, $\bar{\Gamma}_{\bullet\bullet}$, $\bar{\gamma}_{\bullet\bullet}$, $\bar{\Gamma}_{\bullet\bullet}$, $\bar{\gamma}_{\bullet\bullet}$, $\bar{\Gamma}_{i\circ}$, $\bar{\gamma}_{i\circ}$, $\bar{\Gamma}_{\circ j}$ and $\bar{\gamma}_{\circ j}$ are defined similarly to $\bar{\Phi}_{\circ\circ}$, $\bar{\Phi}_{i\circ}$, $\bar{\Phi}_{\circ j}$, $\bar{\Phi}_{\bullet\bullet}$ and $\bar{\Phi}_{\bullet\bullet}$ in (17) and (18).

Suppose that the rank condition holds:

$$\text{Rank}(\bar{\Phi}_{ij}) = m \text{ for all } (ij). \quad (26)$$

Then, we obtain from (19):

$$\mathbf{f}_{ijt} = \left(\bar{\Phi}'_{ij} \bar{\Phi}_{ij} \right)^{-1} \bar{\Phi}'_{ij} (\bar{\mathbf{z}}_{ijt} - \bar{\Xi}_{ij} \mathbf{d}_t - \bar{\mathbf{u}}_{ijt}) \quad (27)$$

It is easily seen, by applying Lemma 1 of Pesaran (2006), to $\bar{\mathbf{u}}_t$, $\bar{\mathbf{u}}_{i\circ t}$, $\bar{\mathbf{u}}_{\circ j t}$, $\mathbf{f}_{\bullet\circ t}$ and $\mathbf{f}_{\circ\bullet t}$, that for each t , as $N \rightarrow \infty$,

$$\bar{\mathbf{u}}_{ijt} = O_p \left(\frac{1}{\sqrt{N}} \right).$$

Therefore, we establish that

$$\mathbf{f}_{ijt} - \left(\bar{\Phi}'_{ij} \bar{\Phi}_{ij} \right)^{-1} \bar{\Phi}'_{ij} (\bar{\mathbf{z}}_{ijt} - \bar{\Xi}_{ij} \mathbf{d}_t) = O_p \left(\frac{1}{\sqrt{N}} \right).$$

This suggests that we can use $\bar{\mathbf{h}}_{ijt} = (\mathbf{d}'_t, \bar{\mathbf{z}}'_{ijt})'$ as observable proxies for \mathbf{f}_{ijt} . Then, we can consistently estimate the individual slope coefficients, β_{ij} and their means β by augmenting the regression, (1) with \mathbf{d}_t and the cross-section averages $\bar{\mathbf{z}}_{ijt}$. This estimator is referred to as the 3 dimensional common correlated effect (3DCCCE) estimator.

If the rank condition (26) does not hold, we need to show that $\frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}(\mathbf{F}\gamma_{ij} + \mathbf{F}_{io}\gamma_{oj} + \mathbf{F}_{oj}\gamma_{io})$ converge to zero. In particular, we need to show that

$$p \lim_{N,T \rightarrow \infty} \frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}\mathbf{F}\bar{\Phi}'_{\infty} = 0, \quad p \lim_{N,T \rightarrow \infty} \frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}\mathbf{F}_{io}\bar{\Phi}'_{\bullet} = 0, \quad \text{and} \quad p \lim_{N,T \rightarrow \infty} \frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}\mathbf{F}_{oj}\bar{\Phi}'_{\bullet} = 0. \quad (28)$$

Define $\bar{\mathbf{Z}}_{ij}^{(1)} = (\mathbf{D}, \mathbf{F}\bar{\Phi}'_{\infty}, \bar{\mathbf{Z}}_{io}, \bar{\mathbf{Z}}_{oj})$ and $\bar{\mathbf{M}}_{ij}^{(1)} = \mathbf{I}_T - \bar{\mathbf{Z}}_{ij}^{(1)}(\bar{\mathbf{Z}}_{ij}^{(1)'}\bar{\mathbf{Z}}_{ij}^{(1)})^{-1}\bar{\mathbf{Z}}_{ij}^{(1)'}$. The first term in (28) can be expressed as (see Lemmas 1-3 in Pesaran (2006)):

$$\frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}\mathbf{F}\bar{\Phi}'_{\infty} = \frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}^{(1)}\mathbf{F}\bar{\Phi}'_{\infty} + O_p\left(\frac{1}{N^2}\right) + O_p\left(\frac{1}{N\sqrt{T}}\right) = O_p\left(\frac{1}{N^2}\right) + O_p\left(\frac{1}{N\sqrt{T}}\right),$$

since $\frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}^{(1)}\mathbf{F}\bar{\Phi}'_{\infty} = 0$. Using (21), it is straightforward to establish that

$$\begin{aligned} \frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}\mathbf{F}\left(\bar{\Gamma}'_{\infty}\beta + \bar{\gamma}_{\infty} + \frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\Gamma'_{ij}(\nu_{io} + \nu_{oj} + \nu_{ij})\right) &= O_p\left(\frac{1}{N^2}\right) + O_p\left(\frac{1}{N\sqrt{T}}\right), \\ \frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}\mathbf{F}\bar{\Gamma}'_{\infty} &= O_p\left(\frac{1}{N^2}\right) + O_p\left(\frac{1}{N\sqrt{T}}\right). \end{aligned}$$

This implies that

$$\frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}\mathbf{F}\bar{\gamma}_{\infty} = O_p\left(\frac{1}{N^2}\right) + O_p\left(\frac{1}{N\sqrt{T}}\right) - \left(\frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}\mathbf{F}\right)\left(\frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\Gamma'_{ij}(\nu_{io} + \nu_{oj} + \nu_{ij})\right). \quad (29)$$

Notice that $\frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}\mathbf{F} = O_p(1)$ and

$$\begin{aligned} \frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\Gamma'_{ij}(\nu_{io} + \nu_{oj} + \nu_{ij}) &= \frac{1}{N}\sum_{i=1}^N\left(\frac{1}{N}\sum_{j=1}^N\Gamma'_{ij}\right)\nu_{io} + \frac{1}{N}\sum_{j=1}^N\left(\frac{1}{N}\sum_{i=1}^N\Gamma'_{ij}\right)\nu_{oj} \\ &+ \frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\Gamma'_{ij}\nu_{ij} = O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{N}\right) \end{aligned} \quad (30)$$

Therefore, we have:

$$\frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}\mathbf{F}\bar{\gamma}_{\infty} = O_p\left(\frac{1}{N^2}\right) + O_p\left(\frac{1}{N\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{N}\right) = O_p\left(\frac{1}{\sqrt{N}}\right) + o_p\left(\frac{1}{\sqrt{N}}\right).$$

Similarly, we can show that the second and third terms in (28) can be expressed as

$$\frac{1}{T}\mathbf{X}'_{ij}\bar{\mathbf{M}}_{ij}\mathbf{F}_{io}\bar{\gamma}_{\bullet} = O_p\left(\frac{1}{\sqrt{N}}\right) + o_p\left(\frac{1}{\sqrt{N}}\right)$$

$$\frac{1}{T} \mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{F}_{\bullet o j} \bar{\boldsymbol{\gamma}}_{\bullet o} = O_p \left(\frac{1}{\sqrt{N}} \right) + o_p \left(\frac{1}{\sqrt{N}} \right).$$

Letting $\mathbf{Q}_{ij} = \left(\mathbf{D}, \mathbf{F} \bar{\boldsymbol{\Phi}}'_{\bullet o}, \mathbf{F}_{i \bullet} \bar{\boldsymbol{\Phi}}'_{\bullet \bullet}, \mathbf{F}_{\bullet o j} \bar{\boldsymbol{\Phi}}'_{\bullet o} \right)$ and $\mathbf{M}_{Q,ij} = \mathbf{I}_T - \mathbf{Q}_{ij} (\mathbf{Q}'_{ij} \mathbf{Q}_{ij})^{-1} \mathbf{Q}'_{ij}$, we have:

$$\frac{1}{T} \mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{F}_{ij} = \frac{1}{T} \mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{ij} + O_p \left(\frac{1}{N} \right) + O_p \left(\frac{1}{\sqrt{NT}} \right), \quad \forall i, j \quad (31)$$

$$\frac{1}{T} \mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \boldsymbol{\varepsilon}_{ij} = \frac{1}{T} \mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \boldsymbol{\varepsilon}_{ij} + O_p \left(\frac{1}{N} \right) + O_p \left(\frac{1}{\sqrt{NT}} \right), \quad \forall i, j \quad (32)$$

and

$$\frac{1}{T} \mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij} = \frac{1}{T} \mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij} + O_p \left(\frac{1}{N} \right) + O_p \left(\frac{1}{\sqrt{NT}} \right), \quad \forall i, j. \quad (33)$$

If the rank condition holds, then the above results hold by replacing \mathbf{Q}_{ij} and $\mathbf{M}_{Q,ij}$ with \mathbf{F}_{ij} and $\mathbf{M}_{F,ij}$ where $\mathbf{M}_{F,ij} = \mathbf{I}_T - \mathbf{F}_{ij} (\mathbf{F}'_{ij} \mathbf{F}_{ij})^{-1} \mathbf{F}'_{ij}$.

Remark 3 *It is non-trivial to develop an appropriate error components specification for accommodating strong CSD within the multi-dimensional dataset since such setup would involve several layers of specification. Similar issues have been examined intensively in the international business cycle literature, especially in terms of the separate identification of regional and global factors. Kose, Otrok, and Whiteman (2003) propose the Bayesian method based on the state space representation, and Moench, Ng, and Potter (2013) extend this approach. Beck, Hubrich, and Marcellino (2016) propose a sequential principal component method to extract factors from overlapping data blocks to identify aggregate, sectoral, country-specific and regional components of price changes in the euro area. Breitung and Eickmeier (2016) suggest a sequential least squares algorithm and a two-step approach based on canonical correlations for multi-level factor models, which are subject to blocks of zero restrictions on the factor loadings matrix. Choi, Kim, Kim, and Kwark (2016) propose a sequential principal component estimation procedure to identify global and country factors. Notice, however, that the main difference between these approaches and ours lies in that we develop the appropriate 3D panel data regression between dependent variable and regressors with hierarchical cross section dependence rather than extracting global and regional factors.*

Remark 4 *The 3D error components model, (2) with the hierarchical multi-factor structure can be viewed as the generalisation of the CTFE specification in (68) of Section 5. Given that most existing studies have neglected the presence of cross-sectional error dependence in the 3D models, KMSS have proposed the error components specification in (69), that combines unobserved global factors with CTFEs. However, the current paper is substantially different from the KMSS approach. KMSS introduce global factors only, and combine the Pesaran (2006) PCCE estimator with the 3D within estimation proposed by Balazsi, Matyas, and Wansbeek*

(2018). Through the hierarchical factor structure, we now provide a more parsimonious and structural specification, and develop a novel 3D CCE estimator. Furthermore, our approach can share similar motivations with the spatial econometric methodology that explicitly introduces spatially-structured origin and destination effects, in such a way that regions treated as origins (destinations) exhibit similar effects to neighbors of origins (destinations), e.g. LeSage and Llano (2016). We allow both factors to be correlated with the regressors in order to deal with endogeneity issues related to the (omitted) time-varying determinants such as globalisation or business cycle factors, which are difficult to measure in practice.³

Remark 5 *An extension to the general case*

$$y_{ijt} = \beta'_{ij} \mathbf{x}_{ijt} + \beta'_j \mathbf{x}_{it} + \beta'_i \mathbf{x}_{jt} + \delta'_{ij} \mathbf{d}_t + u_{ijt}$$

would be straightforward, though we stick to the simpler setting for tractability. See Section 5.2 on the empirical specification, (70).

3 3D Common Correlated Effects Estimator

In this Section we present the main asymptotic results.

3.1 Individual Specific Coefficients

The 3DCCE estimator of the individual slope coefficients, β_{ij} is given by

$$\hat{\mathbf{b}}_{ij} = (\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij})^{-1} \mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{y}_{ij} \quad (34)$$

where $\bar{\mathbf{M}}_{ij}$ is defined in (13). Using (6) in (34), we show the dependence of $\hat{\mathbf{b}}_{ij}$ on the unobserved factors as follows:

$$\hat{\mathbf{b}}_{ij} - \beta_{ij} = (\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij})^{-1} \mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} (\mathbf{F} \gamma_{ij} + \mathbf{F}_{io} \gamma_{oj} + \mathbf{F}_{oj} \gamma_{io}) + (\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij})^{-1} \mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \boldsymbol{\varepsilon}_{ij} \quad (35)$$

$$\begin{aligned} &= \left(\frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T} \right)^{-1} \frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{F}_{ij}}{T} \gamma_{ij}^* + \left(\frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T} \right)^{-1} \frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \boldsymbol{\varepsilon}_{ij}}{T} \\ &= \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T} \right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{ij}}{T} \gamma_{ij}^* + \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T} \right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \boldsymbol{\varepsilon}_{ij}}{T} \\ &+ O_p \left(\frac{1}{N} \right) + O_p \left(\frac{1}{\sqrt{NT}} \right) \end{aligned}$$

³Suppose that y_{ijt} is the export flow from country i to country j at time t (see the application in Section 5). In this case the global factors \mathbf{f}_t may proxy the globalisation trends or worldwide multilateral resistance, and thus affect all of the bilateral pairs (i, j) . On the other hand \mathbf{f}_{ojt} and \mathbf{f}_{iot} represent local factors associated with the flows of the i -th (exporter) or j -th (importer) unit.

where $\mathbf{M}_{Q,ij} = \mathbf{I}_T - \mathbf{Q}_{ij} (\mathbf{Q}'_{ij} \mathbf{Q}_{ij})^{-1} \mathbf{Q}'_{ij}$ with $\mathbf{Q}_{ij} = (\mathbf{D}, \mathbf{F} \bar{\boldsymbol{\Phi}}'_{\bullet\bullet}, \mathbf{F}_{i\circ} \bar{\boldsymbol{\Phi}}'_{\bullet\bullet}, \mathbf{F}_{\circ j} \bar{\boldsymbol{\Phi}}'_{\bullet\bullet})$, $\mathbf{F}_{ij} = (\mathbf{F}, \mathbf{F}_{i\circ}, \mathbf{F}_{\circ j})$ and $\boldsymbol{\gamma}_{ij}^* = (\boldsymbol{\gamma}'_{ij}, \boldsymbol{\gamma}'_{i\circ}, \boldsymbol{\gamma}'_{\circ j})'$.

We analyse the individual components of $\left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T}\right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{ij}}{T} \boldsymbol{\gamma}_{ij}^*$. Using (30), we first show:

$$\begin{aligned} & \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T}\right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{ij}}{T} \boldsymbol{\gamma}_{ij}^* = \\ & \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T}\right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F} (\bar{\boldsymbol{\gamma}}_{\bullet\bullet} + (\boldsymbol{\eta}_{ij} - \bar{\boldsymbol{\eta}}_{\bullet\bullet}))}{T} = \boldsymbol{\chi}_{ij,\bullet\bullet} + \boldsymbol{\chi}_{ij} \end{aligned} \quad (36)$$

where we use $\boldsymbol{\gamma}_{ij} = \bar{\boldsymbol{\gamma}}_{\bullet\bullet} + (\boldsymbol{\eta}_{ij} - \bar{\boldsymbol{\eta}}_{\bullet\bullet})$ from (9). Using (29), we have:

$$\begin{aligned} \boldsymbol{\chi}_{ij,\bullet\bullet} &= \left[\left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T}\right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}}{T} \right] (-\bar{\boldsymbol{\Gamma}}'_{\nu,\bullet\bullet} - \bar{\boldsymbol{\eta}}_{\bullet\bullet}), \\ \boldsymbol{\chi}_{ij} &= \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T}\right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F} \boldsymbol{\eta}_{ij}}{T} \end{aligned}$$

where

$$\begin{aligned} \bar{\boldsymbol{\Gamma}}_{\nu,\bullet\bullet} &= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (\boldsymbol{\nu}_{i\circ} + \boldsymbol{\nu}_{\circ j} + \boldsymbol{\nu}_{ij})' \boldsymbol{\Gamma}_{ij} \\ &= \frac{1}{N} \sum_{i=1}^N \boldsymbol{\nu}'_{i\circ} \left(\frac{1}{N} \sum_{j=1}^N \boldsymbol{\Gamma}_{ij}\right) + \frac{1}{N} \sum_{j=1}^N \boldsymbol{\nu}'_{\circ j} \left(\frac{1}{N} \sum_{i=1}^N \boldsymbol{\Gamma}_{ij}\right) + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\nu}'_{ij} \boldsymbol{\Gamma}_{ij} \\ &= \frac{1}{N} \sum_{i=1}^N \boldsymbol{\nu}'_{i\circ} \bar{\boldsymbol{\Gamma}}_{i\circ} + \frac{1}{N} \sum_{j=1}^N \boldsymbol{\nu}'_{\circ j} \bar{\boldsymbol{\Gamma}}_{\circ j} + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\nu}'_{ij} \boldsymbol{\Gamma}_{ij}. \end{aligned}$$

Similarly, we show that

$$\begin{aligned} & \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T}\right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ} \boldsymbol{\gamma}_{\circ j}}{T} = \\ & \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T}\right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ} (\bar{\boldsymbol{\gamma}}_{\bullet\bullet} + (\boldsymbol{\eta}_{\circ j} - \bar{\boldsymbol{\eta}}_{\bullet\bullet}))}{T} = \boldsymbol{\chi}_{ij,\bullet\bullet} + \boldsymbol{\chi}_{ij,\circ j} \end{aligned} \quad (37)$$

where

$$\begin{aligned} \boldsymbol{\chi}_{ij,\bullet\bullet} &= \left[\left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T}\right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right] (-\bar{\boldsymbol{\Gamma}}'_{i,\nu,\bullet\bullet} - \bar{\boldsymbol{\eta}}_{\bullet\bullet}), \\ \boldsymbol{\chi}_{ij,\circ j} &= \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T}\right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ} \boldsymbol{\eta}_{\circ j}}{T} \end{aligned}$$

$$\bar{\boldsymbol{\Gamma}}_{i,\nu,\bullet\bullet} = \frac{1}{N} \sum_{j=1}^N (\boldsymbol{\nu}_{i\circ} + \boldsymbol{\nu}_{\circ j} + \boldsymbol{\nu}_{ij})' \boldsymbol{\Gamma}_{\circ j} = \boldsymbol{\nu}'_{i\circ} \bar{\boldsymbol{\Gamma}}_{\bullet\bullet} + \frac{1}{N} \sum_{j=1}^N \boldsymbol{\nu}'_{\circ j} \boldsymbol{\Gamma}_{\circ j} + \frac{1}{N} \sum_{j=1}^N \boldsymbol{\nu}'_{ij} \boldsymbol{\Gamma}_{\circ j},$$

Finally,

$$\begin{aligned} & \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T} \right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{oj} \gamma_{io}}{T} = \\ & \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T} \right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{oj} (\bar{\gamma}_{\bullet\bullet} + (\boldsymbol{\eta}_{io} - \bar{\boldsymbol{\eta}}_{\bullet\bullet}))}{T} = \boldsymbol{\chi}_{ij,\bullet\bullet} + \boldsymbol{\chi}_{ij,io} \end{aligned} \quad (38)$$

where

$$\begin{aligned} \boldsymbol{\chi}_{ij,\bullet\bullet} &= \left[\left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T} \right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{oj}}{T} \right] \left(-\bar{\boldsymbol{\Gamma}}'_{j,\nu,\bullet\bullet} - \bar{\boldsymbol{\eta}}_{\bullet\bullet} \right), \\ \boldsymbol{\chi}_{ij,io} &= \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T} \right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{oj} \boldsymbol{\eta}_{io}}{T} \\ \bar{\boldsymbol{\Gamma}}_{j,\nu,\bullet\bullet} &= \frac{1}{N} \sum_{i=1}^N (\boldsymbol{\nu}_{io} + \boldsymbol{\nu}_{oj} + \boldsymbol{\nu}_{ij})' \boldsymbol{\Gamma}_{io} = \frac{1}{N} \sum_{i=1}^N \boldsymbol{\nu}'_{io} \boldsymbol{\Gamma}_{io} + \boldsymbol{\nu}'_{oj} \bar{\boldsymbol{\Gamma}}_{\bullet\bullet} + \frac{1}{N} \sum_{i=1}^N \boldsymbol{\nu}'_{ij} \boldsymbol{\Gamma}_{io}. \end{aligned}$$

Using the above results in (35), we establish that

$$\hat{\boldsymbol{\beta}}_{ij} - \boldsymbol{\beta}_{ij} = \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T} \right)^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \boldsymbol{\varepsilon}_{ij}}{T} + o_p(1).$$

For N and T sufficiently large, the distribution of $\sqrt{T} (\hat{\boldsymbol{\beta}}_{ij} - \boldsymbol{\beta}_{ij})$ will be asymptotically normal if $\sqrt{T}/N \rightarrow 0$ as $N, T \rightarrow \infty$.

In the case where the rank condition (26) is satisfied, we summarise the above analyses and provide the asymptotic distribution in Theorem 1.

Theorem 1 *Consider the 3D heterogeneous panel data model, (1)-(3). Suppose that Assumptions 1-4 and 5(a) hold. Then, the 3DCCE estimator of the individual slope coefficients given by (34) is consistent. Further, as $N, T \rightarrow \infty$ and $\sqrt{T}/N \rightarrow 0$,*

$$\sqrt{T} (\hat{\boldsymbol{\beta}}_{ij} - \boldsymbol{\beta}_{ij}) \rightarrow^d N(0, \mathbf{V}_{ij}), \quad (39)$$

where $\mathbf{V}_{ij} = \boldsymbol{\Sigma}_{v,ij}^{-1} \boldsymbol{\Sigma}_{ij\varepsilon} \boldsymbol{\Sigma}_{v,ij}^{-1}$, $\boldsymbol{\Sigma}_{v,ij} = \text{Var}(\mathbf{v}_{ijt})$, $\boldsymbol{\Sigma}_{ij\varepsilon} = p \lim_{T \rightarrow \infty} \left[\frac{\mathbf{X}'_{ij} \mathbf{M}_{F,ij} \boldsymbol{\Omega}_{ij\varepsilon} \mathbf{M}_{F,ij} \mathbf{X}_{ij}}{T} \right]$ with $\mathbf{M}_{F,ij} = \mathbf{I}_T - \mathbf{F}_{ij} (\mathbf{F}'_{ij} \mathbf{F}_{ij})^{-1} \mathbf{F}'_{ij}$, and $\boldsymbol{\Omega}_{ij\varepsilon} = E(\boldsymbol{\varepsilon}'_{ij} \boldsymbol{\varepsilon}_{ij})$.

3.2 3D Common Correlated Effects Mean Group Estimator

In this section we assume that the parameters of interest are the cross-sectional means of $\boldsymbol{\beta}_{ij}$ defined by (11), and consider the mean group (MG) estimator proposed by Pesaran and Smith (1995). We call it the 3D-CCEMG estimator, which is given by an average of the individual 3DCCE estimators $\hat{\boldsymbol{\beta}}_{ij}$ in (34):

$$\hat{\boldsymbol{\beta}}_{MG} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \hat{\boldsymbol{\beta}}_{ij}. \quad (40)$$

Under Assumption 4 and using (35), we have:

$$\sqrt{N} \left(\hat{\mathbf{b}}_{MG} - \boldsymbol{\beta} \right) = \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\nu}_{ij} + \frac{1}{\sqrt{N}} \sum_{i=1}^N \boldsymbol{\nu}_{i\circ} + \frac{1}{\sqrt{N}} \sum_{j=1}^N \boldsymbol{\nu}_{\circ j} \quad (41)$$

$$+ \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \left(\frac{\sqrt{N} \mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{ij}}{T} \right) \boldsymbol{\gamma}_{ij}^* + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \left(\frac{\sqrt{N} \mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \boldsymbol{\varepsilon}_{ij}}{T} \right) \quad (42)$$

$$= \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\nu}_{ij} + \frac{1}{\sqrt{N}} \sum_{i=1}^N \boldsymbol{\nu}_{i\circ} + \frac{1}{\sqrt{N}} \sum_{j=1}^N \boldsymbol{\nu}_{\circ j} + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \left(\frac{\sqrt{N} \mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \boldsymbol{\varepsilon}_{ij}}{T} \right) \quad (43)$$

$$+ \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij} + \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij,\circ\circ} \quad (44)$$

$$+ \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij,i\circ} + \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij,\bullet\circ} \quad (45)$$

$$+ \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij,\circ j} + \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij,\circ\bullet} \quad (46)$$

$$+ O_p \left(\frac{1}{\sqrt{N}} \right) + O_p \left(\frac{1}{\sqrt{T}} \right),$$

where $\boldsymbol{\Psi}_{Q,ijT} = (T^{-1} \mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij})$ satisfies Assumption 5(a). By analysing each term above, we obtain the following Theorem:

Theorem 2 Consider the 3D heterogeneous panel data model, (1)-(3). Suppose that Assumptions 1-4 and 5(a) hold. Then, the 3D-CCEMG estimator, $\hat{\mathbf{b}}_{MG}$ defined by (40), is consistent. Further, as $N, T \rightarrow \infty$,

$$\sqrt{N} (\bar{\mathbf{b}}_{MG} - \boldsymbol{\beta}) \rightarrow^d N(0, \mathbf{V}_{MG}) \text{ with } \mathbf{V}_{MG} = \boldsymbol{\Omega}_{\nu,\bullet\bullet} + \boldsymbol{\Omega}_{\eta,\bullet\bullet} + \boldsymbol{\Omega}_{\nu,\circ\bullet} + \boldsymbol{\Omega}_{\eta,\circ\bullet} \quad (47)$$

where

$$\boldsymbol{\Omega}_{\nu,\bullet\bullet} = \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E(\mathbf{A}_{1,i,NT} \boldsymbol{\Omega}_{\nu,\bullet\bullet} \mathbf{A}'_{1,i,NT}), \quad \boldsymbol{\Omega}_{\eta,\bullet\bullet} = \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E(\mathbf{A}_{2,i,NT} \boldsymbol{\Omega}_{\eta,\bullet\bullet} \mathbf{A}'_{2,i,NT})$$

$$\boldsymbol{\Omega}_{\nu,\circ\bullet} = \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N E(\mathbf{A}_{1,j,NT} \boldsymbol{\Omega}_{\nu,\circ\bullet} \mathbf{A}'_{1,j,NT}), \quad \boldsymbol{\Omega}_{\eta,\circ\bullet} = \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N E(\mathbf{A}_{2,j,NT} \boldsymbol{\Omega}_{\eta,\circ\bullet} \mathbf{A}'_{2,j,NT})$$

with $\mathbf{A}_{1,i,NT}$, $\mathbf{A}_{2,i,NT}$, $\mathbf{A}_{1,j,NT}$, $\mathbf{A}_{2,j,NT}$ defined in (87)-(90) in the Appendix. Finally, \mathbf{V}_{MG} can be consistently estimated by

$$\hat{\mathbf{V}}_{MG} = \frac{1}{N-1} \sum_{i=1}^N (\hat{\mathbf{b}}_i - \hat{\mathbf{b}}_{MG}) (\hat{\mathbf{b}}_i - \hat{\mathbf{b}}_{MG})' + \frac{1}{N-1} \sum_{j=1}^N (\hat{\mathbf{b}}_j - \hat{\mathbf{b}}_{MG}) (\hat{\mathbf{b}}_j - \hat{\mathbf{b}}_{MG})', \quad (48)$$

where $\hat{\mathbf{b}}_i = \frac{1}{N} \sum_{j=1}^N \hat{\mathbf{b}}_{ij}$ and $\hat{\mathbf{b}}_j = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{b}}_{ij}$.

Theorem 2 establishes an existence of a normal limit. Another important finding is that the dominant terms of $\sqrt{N} (\bar{\mathbf{b}}_{MG} - \boldsymbol{\beta})$ are those that involve $\boldsymbol{\nu}_{i\circ}$, $\boldsymbol{\nu}_{\circ j}$, $\boldsymbol{\eta}_{i\circ}$ and $\boldsymbol{\eta}_{\circ j}$ only, because the terms associated with $\boldsymbol{\nu}_{ij}$ and $\boldsymbol{\eta}_{ij}$ are asymptotically negligible. This explains \sqrt{N} rate of convergence. We also establish a novel result that the nonparametric variance estimator, $\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (\hat{\mathbf{b}}_{ij} - \hat{\mathbf{b}}_{MG}) (\hat{\mathbf{b}}_{ij} - \hat{\mathbf{b}}_{MG})'$, proposed by Pesaran (2006), does not provide a consistent estimator of \mathbf{V}_{MG} , because it gives an equal weight to the terms containing $\boldsymbol{\nu}_{i\circ}$, $\boldsymbol{\nu}_{\circ j}$, $\boldsymbol{\eta}_{i\circ}$ and $\boldsymbol{\eta}_{\circ j}$, and those containing $\boldsymbol{\nu}_{ij}$ and $\boldsymbol{\eta}_{ij}$. As is clear from the proof in the Appendix, the terms associated with $\boldsymbol{\nu}_{i\circ}$, $\boldsymbol{\nu}_{\circ j}$, $\boldsymbol{\eta}_{i\circ}$ and $\boldsymbol{\eta}_{\circ j}$ dominate, providing the sole constituents of the variance estimator. Hence, the consistent nonparametric estimator, $\hat{\mathbf{V}}_{MG}$ in (48) ensures that $\boldsymbol{\nu}_{ij}$ and $\boldsymbol{\eta}_{ij}$ are averaged out by the use of $\hat{\mathbf{b}}_i$ and $\hat{\mathbf{b}}_j$.

Remark 6 *Theorem 2 does not require the rank condition to hold as long as the number of unobserved factors m is fixed. Further, we do not require any restriction on the relative rates of N and T . When the rank condition (26) is satisfied, Assumption 3 can be relaxed such that it would be sufficient that the factor loadings are bounded.*

3.3 3D Common Correlated Effects Pooled Estimator

Consider the special case where the individual coefficients $\boldsymbol{\beta}_{ij}$ are assumed (possibly incorrectly) homogeneous. In this case efficiency gain from pooling can be achieved. We still allow coefficients on observed and unobserved common effects to differ across (ij) . We derive the pooled estimator of $\boldsymbol{\beta}$, referred to as the 3D-CCEP estimator by

$$\hat{\mathbf{b}}_P = \left(\sum_{i=1}^N \sum_{j=1}^N \mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij} \right)^{-1} \sum_{i=1}^N \sum_{j=1}^N \mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{y}_{ij}, \quad (49)$$

Along similar lines of the analysis of the MG estimator, we obtain the following result:

Theorem 3 *Consider the 3D heterogeneous panel data model, (1)-(3). Suppose that Assumptions 1-4 and 5(b) hold. Then, the 3D-CCEP estimator, $\hat{\mathbf{b}}_P$ in (49), is consistent. Further,*

$$\sqrt{N} (\hat{\mathbf{b}}_P - \boldsymbol{\beta}) \rightarrow^d N(0, \boldsymbol{\Psi}^{-1} \mathbf{R} \boldsymbol{\Psi}^{-1}) \quad (50)$$

where

$$\boldsymbol{\Psi} = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\Psi}_{ij} \text{ with } \boldsymbol{\Psi}_{ij} = E \left[\left(\frac{\mathbf{X}_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T} \right)^{-1} \right] \quad (51)$$

$$\mathbf{R} = \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\nu}, \bullet \bullet} + \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\eta}, \bullet \bullet} + \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\nu}, \bullet \circ} + \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\eta}, \bullet \circ}. \quad (52)$$

$$\begin{aligned}\tilde{\Omega}_{\nu,\bullet\bullet} &= \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E \left(\tilde{\mathbf{A}}_{1,i,NT} \Omega'_{\nu\bullet\bullet} \tilde{\mathbf{A}}'_{1,i,NT} \right), & \tilde{\Omega}_{\eta,\bullet\bullet} &= \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E \left(\tilde{\mathbf{A}}_{2,i,NT} \Omega'_{\eta\bullet\bullet} \tilde{\mathbf{A}}'_{2,i,NT} \right) \\ \tilde{\Omega}_{\nu,\circ\bullet} &= \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N E \left(\tilde{\mathbf{A}}_{1,j,NT} \Omega_{\nu\circ\bullet} \tilde{\mathbf{A}}'_{1,j,NT} \right), & \tilde{\Omega}_{\eta,\circ\bullet} &= \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N E \left(\tilde{\mathbf{A}}_{2,j,NT} \Omega_{\eta\circ\bullet} \tilde{\mathbf{A}}'_{2,j,NT} \right)\end{aligned}$$

where $\tilde{\mathbf{A}}_{1,i,NT}$, $\tilde{\mathbf{A}}_{2,i,NT}$, $\tilde{\mathbf{A}}_{1,j,NT}$ and $\tilde{\mathbf{A}}_{2,j,NT}$ are defined in (91)-(92) in the Appendix. The variance $\Psi^{-1} \mathbf{R} \Psi^{-1}$ can be consistently estimated by $\hat{\Psi}^{-1} \hat{\mathbf{R}} \hat{\Psi}^{-1}$ where

$$\hat{\Psi} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T}, \quad (53)$$

$$\begin{aligned}\hat{\mathbf{R}} &= \frac{1}{N} \sum_{i=1}^N \left[\frac{1}{N} \sum_{j=1}^N \left(\frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T} \right) (\hat{\mathbf{b}}_{ij} - \hat{\mathbf{b}}_{MG}) \right] \left[\frac{1}{N} \sum_{j=1}^N (\hat{\mathbf{b}}_{ij} - \hat{\mathbf{b}}_{MG})' \left(\frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T} \right) \right] \\ &+ \frac{1}{N} \sum_{j=1}^N \left[\frac{1}{N} \sum_{i=1}^N \left(\frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T} \right) (\hat{\mathbf{b}}_{ij} - \hat{\mathbf{b}}_{MG}) \right] \left[\frac{1}{N} \sum_{i=1}^N (\hat{\mathbf{b}}_{ij} - \hat{\mathbf{b}}_{MG})' \left(\frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T} \right) \right]\end{aligned} \quad (54)$$

Remark 7 The asymptotic variance matrix of $\hat{\mathbf{b}}_P$ depends on unobserved factors and loadings. It is nevertheless possible to estimate it consistently along lines similar to the case of 3D-CCEMG.

3.4 Special Cases

It is important to investigate whether better convergence rates can be achieved if the hierarchical structure is simplified. We focus on two special cases. The first case sets

$$\text{Condition S1 : } \boldsymbol{\eta}_{i\circ} = \boldsymbol{\eta}_{\circ j} = \boldsymbol{\nu}_{i\circ} = \boldsymbol{\nu}_{\circ j} = \mathbf{0}$$

while the second sets

$$\text{Condition S2 : } \mathbf{F}_{i\circ} = \mathbf{F}_{\circ j} = \mathbf{0}.$$

There may be other permutations but their effects follow easily once we consider S1 and S2. The setup under the more restrictive Condition S2 is similar to that of Pesaran (2006) because there is no hierarchical factor structure. Here, we can treat the dataset as a $T \times N^2$ panel by amalgamating the two cross-section dimensions into one and applying the 2DCCE estimation procedure. Then, it is clear that \sqrt{N} rate will be replaced by N , and all the results of Pesaran (2006) and other papers analysing the CCE estimator hold.

Next, we consider the case where S1 holds but not S2. Then, it is instructive to revisit the decomposition in (41)-(46), which becomes:

$$\begin{aligned} \sqrt{N} \left(\hat{\mathbf{b}}_{MG} - \boldsymbol{\beta} \right) &= \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\nu}_{ij} + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \left(\frac{\sqrt{N} \mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \boldsymbol{\varepsilon}_{ij}}{T} \right) + \\ &\frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij} + \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij,\circ\circ} + \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij,\bullet\bullet} + \\ &\frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij,\circ\bullet} + O_p \left(\frac{1}{\sqrt{N}} \right) + O_p \left(\frac{1}{\sqrt{T}} \right), \end{aligned} \quad (55)$$

From the proof of Theorem 2, we find that the magnitude of all terms on the RHS of (55) is, at most, $\frac{1}{\sqrt{N}}$, as long as $N/T \rightarrow 0$, since $\frac{1}{\sqrt{T}} = o\left(\frac{1}{\sqrt{N}}\right)$, giving a faster rate of convergence. But, normality does not always follow since the $O_p\left(\frac{1}{\sqrt{N}}\right)$ term in RHS of (55) is nonnegligible. Notice that the asymptotic variance estimators proposed in Pesaran (2006) become relevant only if normality holds; in particular,

$$\hat{\mathbf{V}}_{MG} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left(\hat{\mathbf{b}}_{ij} - \hat{\mathbf{b}}_{MG} \right) \left(\hat{\mathbf{b}}_{ij} - \hat{\mathbf{b}}_{MG} \right)', \quad (56)$$

for the mean group estimator, and

$$\hat{\mathbf{R}} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left(\frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T} \right) \left(\hat{\mathbf{b}}_{ij} - \hat{\mathbf{b}}_{MG} \right) \left(\hat{\mathbf{b}}_{ij} - \hat{\mathbf{b}}_{MG} \right)' \left(\frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T} \right), \quad (57)$$

for the pooled estimator.

If Condition S1 is considered too restrictive, as it implies homogeneity for the coefficients of the local factors in e_{ijt} , we may entertain the more general heterogeneous setup given by

$$\gamma_{i\circ} = \gamma_{ij\circ} = \gamma_{\bullet\circ} + \boldsymbol{\eta}_{ij\circ}, \quad \gamma_{\circ j} = \gamma_{\circ ij} = \gamma_{\circ\bullet} + \boldsymbol{\eta}_{\circ ij}.$$

Because of the double cross-sectional averaging, the errors $\boldsymbol{\eta}_{ij\circ}$ and $\boldsymbol{\eta}_{\circ ij}$ are negligible since the terms associated with $\boldsymbol{\chi}_{ij,\bullet\bullet}$ and $\boldsymbol{\chi}_{ij,\circ\bullet}$, in (41)-(46), decay sufficiently fast (as fast as the term associated with $\boldsymbol{\chi}_{ij,\circ\circ}$) to give the same fast convergence rate as under S1.

4 Monte Carlo Study

In this section we examine the small sample properties of the 3D-CCEMG and 3D-CCEP estimators derived by (40) and (49). Following the hierarchical multifactor structure, (1)-(3), we generate y_{ijt} and x_{ijt} as follows:

$$y_{ijt} = \beta_{ij} x_{ijt} + \gamma_{1,ij} f_{1,t} + \gamma_{2,ij} f_{2,t} + \gamma_{1,\circ j} f_{1,\circ t} + \gamma_{2,\circ j} f_{2,\circ t} + \gamma_{1,i\circ} f_{1,\circ t} + \gamma_{2,i\circ} f_{2,\circ t} + \varepsilon_{ijt}, \quad (58)$$

$$x_{ijt} = \Gamma_{1,ij}f_{1,t} + \Gamma_{2,ij}f_{2,t} + \Gamma_{1,oj}f_{1,iot} + \Gamma_{2,oj}f_{2,iot} + \Gamma_{1,io}f_{1,ojt} + \Gamma_{2,io}f_{2,ojt} + v_{ijt}, \quad (59)$$

for $i = 1, \dots, N$, $j = 1, \dots, N$ and $t = 1, \dots, T$. Without loss of generality we set $m_d = 0$, $m_x = 1$, and $m_f = m_{\bullet\bullet} = m_{\bullet\circ} = 2$.

The hierarchical unobserved factors, \mathbf{f}_t , \mathbf{f}_{ojt} , \mathbf{f}_{iot} are generated independently as stationary autoregressive processes with zero mean and unit variance:

$$f_{h,t} = \rho_{f_h}f_{h,t-1} + v_{f_{ht}} \text{ with } v_{f_{ht}} \sim iidN(0, 1 - \rho_{f_h}^2) \text{ for } h = 1, 2$$

$$f_{h,iot} = \rho_{f_{h,io}}f_{h,io,t-1} + v_{f_{h,io}t} \text{ with } v_{f_{h,io}t} \sim iidN(0, 1 - \rho_{f_{h,io}}^2) \text{ for } h = 1, 2 \text{ and } i = 1, \dots, N$$

$$f_{h,ojt} = \rho_{f_{h,oj}}f_{h,oj,t-1} + v_{f_{h,oj}t} \text{ with } v_{f_{h,oj}t} \sim iidN(0, 1 - \rho_{f_{h,oj}}^2) \text{ for } h = 1, 2 \text{ and } j = 1, \dots, N$$

The idiosyncratic errors, ε_{ijt} and v_{ijt} , are generated independently as stationary autoregressive processes with zero mean and unit variance:

$$\varepsilon_{ijt} = \rho_\varepsilon\varepsilon_{ij,t-1} + e_{\varepsilon,ijt} \text{ with } e_{\varepsilon,ijt} \sim iidN(0, 1 - \rho_\varepsilon^2)$$

$$v_{ijt} = \rho_vv_{ij,t-1} + e_{v,ijt} \text{ with } e_{v,ijt} \sim iidN(0, 1 - \rho_v^2)$$

We set $\rho_{f_h} = \rho_{f_{h,io}} = \rho_{f_{h,oj}} = \rho_\varepsilon = \rho_v = 0$ and 0.5.⁴

Next, we consider two experiments: Experiment A with the rank condition (26) being satisfied and Experiment B with the rank condition violated. For x_{ijt} in (59), we draw the factor loadings randomly and independently by

$$\Gamma_{1,ij} \sim iidN(0.5, 0.5) \text{ and } \Gamma_{2,ij} \sim iidN(0, 0.5) \text{ for } i, j = 1, \dots, N$$

$$\Gamma_{1,oj} \sim iidN(0.5, 0.5) \text{ and } \Gamma_{2,oj} \sim iidN(0, 0.5) \text{ for } j = 1, \dots, N$$

$$\Gamma_{1,io} \sim iidN(0.5, 0.5) \text{ and } \Gamma_{2,io} \sim iidN(0, 0.5) \text{ for } i = 1, \dots, N$$

For y_{ijt} in (58), we consider the two experiments. For experiment A, we generate:

$$\gamma_{1,ij} \sim iidN(1, 0.2) \text{ and } \gamma_{2,ij} \sim iidN(1, 0.2) \text{ for } i, j = 1, \dots, N \quad (60)$$

$$\gamma_{1,oj} \sim iidN(1, 0.2) \text{ and } \gamma_{2,oj} \sim iidN(1, 0.2) \text{ for } j = 1, \dots, N$$

$$\gamma_{1,io} \sim iidN(1, 0.2) \text{ and } \gamma_{2,io} \sim iidN(1, 0.2) \text{ for } i = 1, \dots, N.$$

In this case the rank condition (26) is satisfied as follows:

$$E\left(\tilde{\mathbf{\Gamma}}_{ij}^A\right) = E\left(\begin{array}{c} \gamma_{ij}^A \\ \mathbf{\Gamma}_{ij} \end{array}\right) = \left(\begin{array}{cc} 1 & 1 \\ 0.5 & 0 \end{array}\right), E\left(\tilde{\mathbf{\Gamma}}_{oj}^A\right) = \left(\begin{array}{cc} 1 & 1 \\ 0.5 & 0 \end{array}\right), E\left(\tilde{\mathbf{\Gamma}}_{io}^A\right) = \left(\begin{array}{cc} 1 & 1 \\ 0.5 & 0 \end{array}\right)$$

⁴We have also considered the different values of the autoregressive parameters, $(\rho_{f_h}, \rho_{f_{h,io}}, \rho_{f_{h,oj}}, \rho_\varepsilon, \rho_v)$, and found that all those simulation results are qualitatively similar. Additional simulation results will be available upon request.

For experiment B , we generate:

$$\begin{aligned}\gamma_{1,ij} &\sim iidN(1, 0.2) \text{ and } \gamma_{2,ij} \sim iidN(0, 1) \text{ for } i, j = 1, \dots, N \\ \gamma_{1,oj} &\sim iidN(1, 0.2) \text{ and } \gamma_{2,oj} \sim iidN(0, 1) \text{ for } j = 1, \dots, N \\ \gamma_{1,io} &\sim iidN(1, 0.2) \text{ and } \gamma_{2,io} \sim iidN(0, 1) \text{ for } i = 1, \dots, N\end{aligned}\tag{61}$$

Then, the rank condition (26) is violated:

$$E\left(\tilde{\mathbf{\Gamma}}_{ij}^B\right) = E\left(\begin{array}{cc} \gamma_{ij}^{B'} & \\ \mathbf{\Gamma}_{ij} & \end{array}\right) = \begin{pmatrix} 1 & 0 \\ 0.5 & 0 \end{pmatrix}, E\left(\tilde{\mathbf{\Gamma}}_{oj}^B\right) = \begin{pmatrix} 1 & 0 \\ 0.5 & 0 \end{pmatrix}, E\left(\tilde{\mathbf{\Gamma}}_{io}^B\right) = \begin{pmatrix} 1 & 0 \\ 0.5 & 0 \end{pmatrix}$$

For each experiment, we consider two cases: Case 1 with the heterogeneous slopes:

$$\beta_{ij} = \beta + \nu_{io} + \nu_{oj} + \nu_{ij}, \nu_{io} \sim iidN(0, 1), \nu_{oj} \sim iidN(0, 1), \nu_{ij} \sim iidN(0, 1)\tag{62}$$

where ν_{io} , ν_{oj} and ν_{ij} are generated independently of one another, and Case 2 with the homogeneous slopes:

$$\beta_{ij} = \beta = 1.\tag{63}$$

To investigate the small sample properties of the alternative estimators, we consider the three versions of 3D estimators: the $3DCCCE_G$ with $u_{ijt} = \alpha_{ij} + \gamma'_{ij}\mathbf{f}_t + \varepsilon_{ijt}$ where we approximate heterogeneous global factors only by $\bar{\mathbf{z}}_t = (\bar{y}_t, \bar{x}_t)'$; the $3DCCCE_L$ estimator with $u_{ijt} = \alpha_{ij} + \gamma'_{oj}\mathbf{f}_{iot} + \gamma'_{io}\mathbf{f}_{ojt} + \varepsilon_{ijt}$ where we approximate heterogeneous local factors only by $\bar{\mathbf{z}}_{iot}$ and $\bar{\mathbf{z}}_{ojt}$; and the $3DCCCE_{GL}$ estimator with $u_{ijt} = \alpha_{ij} + \gamma'_{oj}\mathbf{f}_{iot} + \gamma'_{io}\mathbf{f}_{ojt} + \gamma'_{ij}\mathbf{f}_t + \varepsilon_{ijt}$ where we approximate heterogeneous global and local factors by $\bar{\mathbf{z}}_t$, $\bar{\mathbf{z}}_{iot}$ and $\bar{\mathbf{z}}_{ojt}$. In each case we consider both mean group and pooled estimators.

For each of these estimators we report the bias, the root mean squared error (RMSE) and the coverage rate at the 95% confidence with 1,000 replications for the (N, T) pairs with $N = \{10, 25, 50, 100\}$ and $T = \{50, 100\}$.

Table 1 shows simulation results for Experiment A (the full rank) with heterogeneous coefficients (Case 1). We find that the biases of our proposed $3DCCCE_{GL}$ estimator are mostly negligible even for the relatively small sample sizes. Further, the performance of both pooled and mean group estimators is shown to be almost identical. On the contrary, the $3DCCCE_G$ estimator suffers from severe biases. Finally, the biases of the $3DCCCE_L$ estimator are still non-negligible even for large N and T , but much smaller than those of $3DCCCE_G$, indicating that the local factors approximations seem to be more effective than the global counterpart. This indicates that the CCE estimator advanced by Pesaran (2006) is not effective in removing correlations between local factors and regressors within 3D panels.⁵ These results provide support for theoretical predictions that the joint approximations of heterogeneous global and

⁵The $3DCCCE_G$ estimator is identical to the $2DCCCE$ estimator only if $\mathbf{F}_{io} = \mathbf{F}_{oj} = 0$ (Condition S2).

local factors can only provide consistent estimation of $E(\beta_{ij})$ in the presence of the hierarchical multifactors.

We also find that the results for RMSE display similar patterns. RMSEs of $3DCCE_L$ and $3DCCE_{GL}$ estimators are significantly lower than those of $3DCCE_G$ estimator. The difference between $3DCCE_L$ and $3DCCE_{GL}$ is mostly negligible, but RMSEs of $3DCCE_{GL}$ tend to decline with the sample size. Turning to the coverage rates, we find that $3DCCE_L$ and the $3DCCE_{GL}$ estimators perform better than $3DCCE_G$ estimator, whose coverage rate tends to under-estimate as N rises. Only coverage rates of $3DCCE_{GL}$ estimator tend to the nominal 95% as N or T rises.

Table 2 presents simulation results for Experiment A (the full rank) with homogeneous coefficients (Case 2). We observe qualitatively similar results for biases of the three estimators in Table 1, confirming that the $3DCCE_{GL}$ estimator is most reliable. RMSEs of $3DCCE_G$ estimator are significantly higher than those of $3DCCE_L$ and $3DCCE_{GL}$ estimators. RMSEs of $3DCCE_{GL}$ are now significantly lower than those of $3DCCE_L$, and fall sharply with the sample size. We do not compute coverage rates for the homogeneous coefficient case because our theory only covers the asymptotic distribution of heterogeneous coefficients similarly to Pesaran (2006) with the only exception of the special case in his Theorem 4. We leave this to future research.

Table 3 presents simulation results for Experiment B (the rank deficiency) with heterogeneous coefficients (Case 1). We obtain qualitatively similar results to Table 1, suggesting that the performance of the estimators is not significantly affected by the rank deficiency while confirming that the $3DCCE_{GL}$ estimator is most reliable. This finding provides a support for the validity of our proposed estimator.

Table 4 presents simulation results for Experiment B (the rank deficiency) with homogeneous coefficients (Case 2). We find qualitatively similar results to Table 2. The $3DCCE_{GL}$ estimator is still most reliable.

We have also conducted the additional simulations by constructing the data under Conditions S1 and S2 described in Section 3.4. These simulation results, available in the Supplement, confirm our predictions that the faster convergence rate can be achieved. Under S1 the bias and RMSE of $3DCCE_G$ are substantially higher whilst they decline significantly for $3DCCE_L$. Importantly, they become almost negligible for $3DCCE_{GL}$, and their coverage rate tends to the nominal 95% as N or T rises. Under both S1 and S2, there is no longer the hierarchical factor structure. The bias and RMSE of $3DCCE_G$ are reduced substantially, confirming that both $3DCCE_G$ and $3DCCE_{GL}$ are most reliable estimators. Further, their coverage rates reach the nominal 95% as N or T rises. We also find that the relative performance of the estimators are not affected significantly by the rank deficiency in both experiments.

We note in passing that the $2DCCE$ estimator fails to properly deal with the hierarchical

cross section dependence structure with heterogenous global and local factors. The biases of the $3DCCE_G$ estimator become negligible when the DGP include global factors only, not the local factors, which is confirmed by the additional simulation results for Cases 5s to 8s reported in Supplement, Section 9. The $3DCCE_L$ shows relatively better results, highlighting a potentially more important role played by the local factors in capturing the cross section dependence within a hierarchical structure. In order to achieve the most reliable and consistent estimation results, however, we should apply the global and local approximations jointly.

Table 1: Simulation results for Case 1 - Full Rank (Experiment A)

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.232	0.251	0.064	0.079	-0.006	0.008
	25	0.263	0.272	0.062	0.070	-0.004	0.004
	50	0.285	0.279	0.077	0.072	0.008	0.002
	100	0.282	0.285	0.074	0.076	0.002	0.003
0.5	10	0.302	0.253	0.115	0.067	0.040	-0.008
	25	0.265	0.289	0.049	0.074	-0.022	0.004
	50	0.301	0.290	0.079	0.068	0.005	-0.006
	100	0.300	0.300	0.077	0.078	-0.002	-0.001
Mean Group Estimator							
0	10	0.219	0.239	0.046	0.063	0.001	0.017
	25	0.249	0.258	0.053	0.061	0.003	0.012
	50	0.268	0.264	0.067	0.062	0.011	0.007
	100	0.265	0.268	0.062	0.065	0.002	0.005
0.5	10	0.292	0.247	0.103	0.056	0.053	0.006
	25	0.254	0.276	0.041	0.063	-0.012	0.010
	50	0.287	0.276	0.068	0.056	0.009	-0.003
	100	0.286	0.287	0.064	0.065	-0.001	0.000
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.544	0.530	0.488	0.474	0.483	0.468
	25	0.389	0.396	0.293	0.296	0.286	0.287
	50	0.348	0.345	0.213	0.216	0.198	0.203
	100	0.316	0.316	0.160	0.156	0.141	0.137
0.5	10	0.557	0.536	0.479	0.475	0.466	0.469
	25	0.390	0.412	0.290	0.300	0.288	0.291
	50	0.369	0.353	0.225	0.211	0.211	0.200
	100	0.332	0.332	0.162	0.163	0.142	0.143
Mean Group Estimator							
0	10	0.528	0.523	0.480	0.469	0.478	0.464
	25	0.379	0.385	0.289	0.292	0.284	0.286
	50	0.333	0.332	0.208	0.212	0.197	0.203
	100	0.301	0.301	0.154	0.151	0.141	0.137
0.5	10	0.544	0.525	0.466	0.463	0.458	0.460
	25	0.382	0.400	0.287	0.297	0.285	0.290
	50	0.356	0.340	0.220	0.207	0.209	0.199
	100	0.319	0.320	0.155	0.157	0.141	0.143
Panel C: Coverage rate at 95 confidence level							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.965	0.953	0.903	0.903	0.895	0.886
	25	0.951	0.947	0.924	0.919	0.924	0.929
	50	0.911	0.898	0.922	0.919	0.940	0.932
	100	0.792	0.775	0.917	0.924	0.947	0.959
0.5	10	0.965	0.973	0.898	0.912	0.892	0.890
	25	0.960	0.952	0.925	0.931	0.935	0.927
	50	0.899	0.907	0.914	0.927	0.935	0.945
	100	0.790	0.745	0.915	0.912	0.952	0.947
Mean Group Estimator							
0	10	0.951	0.939	0.891	0.900	0.888	0.891
	25	0.931	0.928	0.923	0.914	0.925	0.931
	50	0.876	0.863	0.933	0.925	0.946	0.931
	100	0.724	0.713	0.929	0.929	0.947	0.961
0.5	10	0.954	0.954	0.904	0.908	0.903	0.901
	25	0.936	0.926	0.932	0.934	0.935	0.936
	50	0.849	0.850	0.917	0.937	0.933	0.949
	100	0.691	0.677	0.926	0.920	0.949	0.946

Notes: $3DCCE_G$ is the estimator with the global factors approximation only, $3DCCE_L$ is the estimator with the local factors approximation only, and $3DCCE_{GL}$ is the estimator with both global and local factors approximation. We consider both mean group and pooled estimators defined in (40) and (49). The variance of $3DCCE_G$ is estimated by (56) for the mean group and (57) for the pooled estimator. The variances of $3DCCE_L$ and $3DCCE_{GL}$ are given by (48) for the mean group and (53)-(54) for the pooled estimator.

Table 2: Simulation Results for Case 2 - Full Rank (Experiment A)

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.240	0.246	0.067	0.073	-0.001	0.006
	25	0.269	0.270	0.068	0.068	0.001	0.002
	50	0.278	0.278	0.071	0.070	0.001	0.001
	100	0.281	0.282	0.074	0.074	0.000	0.000
0.5	10	0.267	0.261	0.076	0.072	0.002	-0.003
	25	0.292	0.292	0.075	0.075	0.003	0.004
	50	0.300	0.299	0.077	0.076	0.002	0.002
	100	0.304	0.303	0.081	0.080	0.001	0.000
Mean Group Estimator							
0	10	0.230	0.235	0.057	0.062	0.005	0.011
	25	0.255	0.256	0.059	0.060	0.006	0.006
	50	0.262	0.262	0.060	0.060	0.002	0.003
	100	0.265	0.265	0.061	0.062	0.000	0.001
0.5	10	0.256	0.252	0.067	0.064	0.010	0.006
	25	0.279	0.278	0.066	0.065	0.008	0.008
	50	0.286	0.286	0.066	0.066	0.004	0.004
	100	0.289	0.289	0.068	0.067	0.002	0.001
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.268	0.273	0.129	0.136	0.110	0.112
	25	0.274	0.274	0.083	0.082	0.045	0.044
	50	0.279	0.279	0.074	0.074	0.022	0.021
	100	0.282	0.282	0.075	0.075	0.010	0.010
0.5	10	0.293	0.289	0.136	0.137	0.113	0.114
	25	0.296	0.296	0.090	0.088	0.046	0.045
	50	0.301	0.300	0.081	0.080	0.022	0.022
	100	0.304	0.303	0.082	0.081	0.011	0.010
Mean Group Estimator							
0	10	0.255	0.259	0.118	0.121	0.103	0.104
	25	0.259	0.260	0.073	0.073	0.042	0.042
	50	0.263	0.263	0.063	0.064	0.020	0.021
	100	0.265	0.265	0.062	0.063	0.010	0.010
0.5	10	0.279	0.276	0.123	0.122	0.103	0.103
	25	0.283	0.282	0.078	0.077	0.042	0.041
	50	0.286	0.287	0.070	0.069	0.021	0.021
	100	0.289	0.289	0.069	0.068	0.011	0.010

Notes: See notes to Table 1.

Table 3: Simulation results for Case 1 - Rank Deficiency- (Experiment B)

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.220	0.242	0.077	0.099	-0.005	0.018
	25	0.223	0.239	0.050	0.066	-0.017	-0.001
	50	0.248	0.246	0.061	0.061	0.002	0.001
	100	0.260	0.249	0.067	0.057	0.012	0.001
0.5	10	0.242	0.249	0.076	0.084	-0.004	0.003
	25	0.273	0.274	0.076	0.079	0.009	0.012
	50	0.258	0.266	0.049	0.058	-0.012	-0.003
	100	0.272	0.271	0.056	0.055	0.000	-0.002
Mean Group Estimator							
0	10	0.211	0.234	0.040	0.062	0.008	0.030
	25	0.214	0.228	0.028	0.043	-0.006	0.009
	50	0.237	0.235	0.047	0.046	0.009	0.008
	100	0.248	0.238	0.057	0.046	0.010	0.005
0.5	10	0.230	0.241	0.041	0.054	0.006	0.019
	25	0.263	0.263	0.059	0.060	0.023	0.024
	50	0.247	0.254	0.038	0.045	-0.003	0.005
	100	0.260	0.259	0.047	0.046	0.003	0.002
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.509	0.520	0.467	0.469	0.461	0.460
	25	0.364	0.362	0.294	0.279	0.290	0.271
	50	0.313	0.322	0.202	0.216	0.192	0.208
	100	0.297	0.286	0.158	0.152	0.144	0.141
0.5	10	0.517	0.520	0.461	0.463	0.454	0.457
	25	0.397	0.396	0.298	0.297	0.289	0.285
	50	0.326	0.335	0.205	0.212	0.199	0.205
	100	0.306	0.306	0.151	0.152	0.140	0.143
Mean Group Estimator							
0	10	0.503	0.510	0.459	0.457	0.457	0.454
	25	0.358	0.354	0.288	0.274	0.286	0.270
	50	0.304	0.313	0.197	0.211	0.191	0.207
	100	0.287	0.276	0.154	0.148	0.144	0.141
0.5	10	0.501	0.509	0.447	0.452	0.444	0.450
	25	0.388	0.388	0.291	0.291	0.286	0.285
	50	0.317	0.326	0.202	0.208	0.198	0.203
	100	0.295	0.295	0.146	0.149	0.139	0.142
Panel C: Coverage rate at 95 confidence level							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.967	0.957	0.896	0.893	0.891	0.894
	25	0.967	0.971	0.927	0.944	0.925	0.951
	50	0.936	0.909	0.938	0.920	0.949	0.927
	100	0.801	0.833	0.922	0.930	0.941	0.952
0.5	10	0.967	0.971	0.901	0.912	0.897	0.902
	25	0.954	0.955	0.912	0.931	0.926	0.938
	50	0.943	0.912	0.937	0.926	0.942	0.934
	100	0.815	0.795	0.932	0.935	0.958	0.942
Mean Group Estimator							
0	10	0.956	0.949	0.905	0.907	0.898	0.909
	25	0.948	0.949	0.934	0.951	0.935	0.953
	50	0.904	0.879	0.946	0.928	0.953	0.932
	100	0.742	0.768	0.932	0.933	0.949	0.952
0.5	10	0.962	0.965	0.922	0.923	0.919	0.922
	25	0.926	0.928	0.931	0.941	0.931	0.944
	50	0.898	0.869	0.940	0.937	0.940	0.935
	100	0.727	0.724	0.937	0.936	0.957	0.943

Notes: See notes to Table 1.

Table 4: Simulation results for Case 2 - Rank Deficiency (Experiment B)

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.231	0.223	0.076	0.072	0.003	-0.001
	25	0.243	0.243	0.063	0.063	0.002	0.002
	50	0.247	0.248	0.056	0.058	0.000	0.001
	100	0.250	0.251	0.054	0.054	0.000	0.001
0.5	10	0.250	0.250	0.078	0.078	0.004	0.004
	25	0.270	0.266	0.067	0.064	0.004	0.001
	50	0.272	0.271	0.059	0.059	0.001	0.001
	100	0.276	0.275	0.056	0.056	0.001	0.000
Mean Group Estimator							
0	10	0.222	0.215	0.053	0.047	0.012	0.006
	25	0.232	0.234	0.048	0.049	0.008	0.009
	50	0.236	0.237	0.046	0.047	0.005	0.005
	100	0.238	0.239	0.045	0.046	0.002	0.003
0.5	10	0.242	0.246	0.057	0.060	0.013	0.016
	25	0.258	0.256	0.054	0.052	0.011	0.010
	50	0.261	0.260	0.051	0.050	0.007	0.006
	100	0.264	0.264	0.049	0.049	0.004	0.004
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.258	0.251	0.136	0.132	0.111	0.110
	25	0.248	0.247	0.078	0.077	0.045	0.044
	50	0.249	0.249	0.061	0.062	0.023	0.023
	100	0.250	0.251	0.055	0.056	0.012	0.011
0.5	10	0.278	0.276	0.138	0.137	0.115	0.112
	25	0.275	0.271	0.082	0.079	0.047	0.046
	50	0.273	0.273	0.064	0.064	0.024	0.024
	100	0.276	0.276	0.058	0.057	0.012	0.012
Mean Group Estimator							
0	10	0.245	0.239	0.116	0.112	0.103	0.101
	25	0.236	0.238	0.063	0.064	0.042	0.042
	50	0.237	0.238	0.051	0.051	0.021	0.021
	100	0.238	0.239	0.047	0.047	0.011	0.010
0.5	10	0.264	0.266	0.118	0.118	0.105	0.103
	25	0.261	0.259	0.068	0.066	0.042	0.042
	50	0.262	0.261	0.055	0.054	0.022	0.021
	100	0.264	0.264	0.050	0.050	0.011	0.011

Notes: See notes to Table 1.

5 The Gravity Model of Bilateral Export Flows

Anderson and van Wincoop (2003) show that the gravity equation tells that bilateral trade, after controlling for size, depends on the bilateral trade barriers as well as the multilateral resistance (MTR) indices measuring bilateral barrier relative to average trade barriers that both regions face with all their trading partners. The system of the structural gravity equations is derived as follows:

$$X_{ij} = \frac{Y_i Y_j}{Y} \left(\frac{t_{ij}}{\Pi_i P_j} \right)^{1-\sigma} \quad (64)$$

$$\Pi_i^{1-\sigma} = \sum_j \left(\frac{t_{ij}}{P_j} \right)^{1-\sigma} \frac{Y_j}{Y} \quad \text{and} \quad P_j^{1-\sigma} = \sum_i \left(\frac{t_{ij}}{\Pi_i} \right)^{1-\sigma} \frac{Y_i}{Y} \quad (65)$$

where X_{ij} are exports from i to j , Y_i , Y_j and Y are GDP measures for i (exporter), j (importer) and the world, t_{ij} (> 1) is one plus the tariff equivalent of overall trade costs of imports of j from i , σ (> 1) is the elasticity of substitution with CES preference, Π_i is ease of access of exporter i , and P_j is the ease of access of importer j . P_j and Π_i are called inward and outward multilateral resistance.

Omitting MTR induces potentially severe bias (e.g. Baldwin and Taglioni (2006)). We consider the log-linearised specification of (64):⁶

$$\ln X_{ij} = \beta_0 + \beta_1 \ln Y_i + \beta_2 \ln Y_j + \beta_3 \ln t_{ij} + \beta_4 \ln P_i + \beta_5 \ln P_j + \varepsilon_{ij} \quad (66)$$

where P_i and P_j are unobservable CES price indices, and t_{ij} contain both barriers and incentives to trade between i and j . Subsequent research has focused on estimating (66) with directional country-specific fixed effects to control for unobservable MTRs, which are therefore replaced by a vector of N country-specific dummies, μ_i and μ_j .

When we extend the model (66) into the 3D panels:

$$\ln X_{ijt} = \beta_0 + \beta_1 \ln Y_{it} + \beta_2 \ln Y_{jt} + \beta_3 \ln t_{ijt} + \beta_4 \ln P_{it} + \beta_5 \ln P_{jt} + \varepsilon_{ijt}, \quad (67)$$

we should allow MTRs to vary over time. Baltagi, Egger, and Pfaffermayr (2003) propose the following 3D error components:

$$u_{ijt} = \alpha_{ij} + \theta_{it} + \theta_{jt}^* + \varepsilon_{ijt}, \quad (68)$$

which contains bilateral pair-fixed effects α_{ij} as well as origin (exporter) and destination (importer) country-time fixed effects (CTFE) θ_{it} and θ_{jt}^* , respectively. This approach has been popularly applied to measure the impacts of MTRs of the exporters and the importers in the structural gravity studies, e.g. Baltagi, Egger, and Pfaffermayr (2015).

⁶Under the assumption of bilateral trade cost symmetry ($t_{ij} = t_{ji}$) and balanced trade, Anderson and van Wincoop (2003) impose the normalisation, $\Pi_i = P_i$.

However, the main drawback of the CTFE approach lies in the underlying assumption that bilateral trade flows are independent of what happens to the rest of the trading world.⁷ Thus, KMSS propose the more general error components specification:⁸

$$u_{ijt} = \alpha_{ij} + \theta_{it} + \theta_{jt}^* + \pi_{ij}\theta_t + \varepsilon_{ijt}, \quad (69)$$

that attempts to model residual CSD via unobserved heterogeneous global factor θ_t in addition to CTFEs. KMSS develop the two-step consistent estimation procedure, called the 3D-PCCE estimator, by approximating global factors with the double cross-section averages of dependent variable and regressors and then applying the 3D-within transformation.

Following this research trend, in this paper, we have developed the hierarchical multi-factor error components specification, (2), which is more structural and parsimonious than (69). Suppose that y_{ijt} in (1) is the trade (export) volumes from source (export) country i to destination (import) country j at time t . In the hierarchical error components specification (2), \mathbf{f}_t are the global factors such as world-wide technology advancement in transportation that affects every trade volume between source country i and destination country j . For example, if there is technology advancement that substantially reduces transportation cost, then it is likely to increase all trade volumes between any pair of countries. The local factors \mathbf{f}_{iot} and \mathbf{f}_{ojt} represent source and destination country factors, respectively. \mathbf{f}_{iot} are the factors of the i th exporting country (say, Germany) that affect the trade volumes from Germany to all destination countries while \mathbf{f}_{ojt} are the factor of the j th importing country (say, Greece) that affect all trade volumes to Greece from all other countries. If German's export price increases, its export volume to all other countries may decrease whereas if the Greece economy depresses, it may decrease its imports from all other countries. In this regard, local factors can be thought of as capturing unobservable multilateral resistance, an index of inward and outward bilateral trade costs. We then apply the 3DCCCE estimators to the gravity model of bilateral exports.

5.1 The Data

We collect the dataset over the period 1970-2013 (44 years), and consider two control groups: (i) the 210 country-pairs of the EU15 member countries with 11 Euro countries (Austria, Belgium-Luxemburg, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Portugal, Spain) and 4 control non-Euro countries (Denmark, Norway, Sweden, the UK), and (ii)

⁷To control for such multilateral cross-sectional correlations across trade flows, Behrens, Ertur, and Kock (2012) develop the cross-section spatial econometric specification. A number of studies have also established an importance of taking into account multilateral resistance, trade costs and bilateral heterogeneity in 2D panels, e.g. Serlenga and Shin (2007), Mastromarco, Serlenga, and Shin (2016) and Gunnella, Mastromarco, Serlenga, and Shin (2015).

⁸KMSS also propose the simpler error components specification, $u_{ijt} = \alpha_{ij} + \pi_{ij}\theta_t + \varepsilon_{ijt}$ in which case the standard CCE estimator in 2D panels can be applied, as discussed in Section 3.4.

the 342 country-pairs among 19 countries with the EU15 countries and 4 additional control non-EU OECD countries (Australia, Canada, Japan and the US). We collect the bilateral export flows from IMF. The data starts from 1970 as information on Germany are unavailable in the 60s. See the Data Appendix for details about the variables.⁹

Our sample period consists of several important economic integrations, such as the European Monetary System in 1979 and the Single Market in 1993, all of which can be regarded as promoting intra-EU trades.

5.2 Empirical Specifications

We consider the following 3D panel gravity specification:

$$EXP_{ijt} = \beta_0 + \beta_1 CEE_{ijt} + \beta_2 EMU_{ijt} + \beta_3 SIM_{ijt} + \beta_4 RLF_{ijt} + \beta_5 RER_{ijt} + \beta_6 GDP_{it} + \beta_7 GDP_{jt} + \gamma_1 DIS_{ij} + \gamma_2 BOR_{ij} + \gamma_3 LAN_{ij} + u_{ijt} \quad (70)$$

where the dependent variable, EXP_{ijt} is the log of export flow from country i to country j at time t , CEE and EMU are dummies for European community membership and for European monetary union, SIM is the logarithm of an index that captures the relative size of two countries and bounded between zero (absolute divergence) and 0.5 (equal size), RLF is the logarithm of the absolute value of the difference between per capita GDPs of trading countries, RER represents the logarithm of real exchange rates, GDP_{it} and GDP_{jt} are logged GDPs of exporter and importer, and finally the logarithm of geographical distance (DIS) and the dummies for common language (LAN) and for common border (BOR) represent time-invariant bilateral barriers.

We apply four estimators to (70); namely the two-way within estimator with $u_{ijt} = \alpha_{ij} + \theta_t + \varepsilon_{ijt}$ and the three versions of 3D-CCEP estimators¹⁰ respectively with $u_{ijt} = \alpha_{ij} + \gamma'_{ij}\mathbf{f}_t + \varepsilon_{ijt}$, $u_{ijt} = \alpha_{ij} + \gamma'_{oj}\mathbf{f}_{iot} + \gamma'_{io}\mathbf{f}_{ojt} + \varepsilon_{ijt}$ and $u_{ijt} = \alpha_{ij} + \gamma'_{oj}\mathbf{f}_{iot} + \gamma'_{io}\mathbf{f}_{ojt} + \gamma'_{ij}\mathbf{f}_t + \varepsilon_{ijt}$. We report the CD test results applied to the residuals and estimates of the CSD exponent (denoted α) respectively for the four estimation methods.¹¹

⁹There are no missing data so we consider the balanced panel. In previous studies, Serlenga and Shin (2007) collect the trade flows data from OECD and estimate a panel gravity model for 182 country-pairs among 14 EU countries (excluding Norway) over the period 1960-2001 (42 years) while Gunnella, Mastromarco, Serlenga, and Shin (2015) and KMSS cover the longer period 1960-2008 (49 years). The IMF data provide more complete information with less missing values than OECD trade data.

¹⁰Here we do not report the results for the 3D-CCEMG estimator, as they are mostly unreliable. This is mainly due to the fact that the substantial number of individual estimation results tend to become unreliable, say 90 cases out of 210 for the EU15 sample and 130 times of 342 for the EU19 sample.

¹¹We evaluate the CD test statistic and estimate the CSD exponent sequentially. For each specification, we estimate residuals by $\mathbf{e}_{ij} = \bar{\mathbf{M}}_{ij}\mathbf{y}_{ij} - \bar{\mathbf{M}}_{ij}\mathbf{x}_{ij}\hat{\beta}$, where $\bar{\mathbf{M}}_{ij}$ is defined in (13). Following KMSS, we compute a modified counterpart of a CD test by Pesaran (2015), and apply the 3D extension of the estimation technique by Bailey, Kapetanios, and Pesaran (2016). We also evaluate the confidence band for $\hat{\alpha}$ by applying the test statistic in (B47) in Bailey, Kapetanios, and Pesaran (2016)'s Supplementary Appendix VI.

We focus on the impacts of the two dummy variables, CEE and EMU (equal to one when both countries belong to the European Community or when both adopt the same currency). Both are expected to exert a positive impact on bilateral export flows. The empirical evidence on the common currency effects is rather mixed (e.g. Rose (2001), de Nardis and Vicarelli (2003) and Frankel (2008)), though recent studies by Mastromarco, Serlenga, and Shin (2016), and Gunnella, Mastromarco, Serlenga, and Shin (2015) that control for strong CSD via unobserved heterogeneous factors in 2D panels, find modest but significant effects (7 to 10%) of the euro on intra-EU trade flows.

5.3 Estimation Results for the 15 EU Countries

Table 5 reports the panel gravity estimation results for the 210 country-pairs among the 15 EU countries, with 11 Euro and 4 non-Euro countries, over the period 1970-2013 (44 years).¹² The FE estimator suffers from strong CSD while the 3D-CCEP estimators display much lower CSD. The CD diagnostic test fails to reject the null of weak CSD for both $3DCCEP_L$ and $3DCCEP_{GL}$. This is also supported by the smaller CSD exponent estimates, $\hat{\alpha}$ for $3DCCEP_L$ (0.624) and $3DCCEP_{GL}$ (0.609), which are close to a moderate range of weak CSD.¹³

On the basis of these results we focus on the results for $3DCCEP_{GL}$ with the lowest CSD. All the coefficients are significant and their signs are consistent with *a priori* expectations. The effect of the foreign GDP on the export is substantially higher than that of the home GDP. The effects of SIM and RER are positive while a depreciation of the home currency (increase in RER) leads to a significant increase in exports. In particular, the real export boosting effect of similarity in size suggests that the intra-industry trade is the main part of the trade in the EU countries, in line with studies on the developed countries, e.g. Markusen and Maskus (2001), and OECD (2002, 2010). More importantly, we find that the impacts of EMU and CEE are significant, but substantially smaller than FE estimates. Both Euro and CEE impacts drop sharply from 0.099 and 0.074 to 0.03 and 0.05, respectively. This evidence is in line with the thesis that the potential trade-creating effects of the Euro should be viewed in the long-

¹²Theoretically, we should employ the entire set of cross-section averages to approximate heterogeneous global and local factors. In practice, however, this might raise an issue of multicollinearity. In order to avoid the curse of dimensionality and obtain the economically sensible estimation results, we conduct the search for an optimal subset of cross sectional averages. Furthermore, in the aftermath of the global financial crisis, the export flows display a negative average growth from 2009 to 2012, as shown below:

	70/80	80/90	90/00	00/10	10/13
Average Growth Export (EU15 + 4 OECD countries)	7.06	6.25	4.35	2.16	-0.34
Average Growth Export (EU15)	8.86	7.37	3.92	2.82	-2.05

Such a decline is more evident for the EU15 countries. Hence, we add t^2 as an observed factor, which helps to capture the confounding effect of the crisis.

¹³Bailey, Kapetanios, and Pesaran (2016) show that $\alpha \in [1/2, 3/4)$ represent a moderate degree of CSD.

run historical and multilateral perspectives rather than simply focusing on the formation of a monetary union as an isolated event, see also Gunnella, Mastromarco, Serlenga, and Shin (2015). On the contrary other estimation results tend to be rather unreliable.¹⁴

Table 5: Estimation Results for 15 EU Countries

	FE	$3DCCEP_G$	$3DCCEP_L$	$3DCCEP_{GL}$
GDP_i	1.517 (0.044)	0.230 (0.036)	0.023 (0.037)	0.342 (0.124)
GDP_j	0.953 (0.044)	1.478 (0.037)	0.779 (0.057)	1.498 (0.031)
SIM	-0.045 (0.060)	0.639 (0.069)	-0.012 (0.056)	0.197 (0.075)
RLF	0.030 (0.006)	-0.002 (0.005)	0.002 (0.002)	0.006 (0.004)
RER	0.012 (0.008)	-0.046 (0.007)	0.016 (0.004)	0.103 (0.010)
EMU	0.099 (0.016)	0.030 (0.003)	0.012 (0.003)	0.030 (0.003)
CEE	0.074 (0.014)	0.066 (0.007)	0.007 (0.007)	0.050 (0.013)
CD stat	206.6	4.67	2.33	2.72
α	0.91 (0.90-0.93)	0.78 (0.72-0.84)	0.62 (0.59-0.66)	0.61 (0.57-0.65)

Notes: The dependent variable is log of export. FE is the two-way fixed effect estimator. $3DCCEP_G$ is the CCEP estimator with only the global factors approximated by $\mathbf{f}_t = \{\overline{EXP}_{..t}, \overline{GDP}_{..t}, \overline{SIM}_{..t}, \overline{RLF}_{..t}, \overline{CEE}_{..t}, t, t^2\}$. $3DCCEP_L$ is the CCEP estimator with only the local factors approximated by $\mathbf{f}_{iot} = \{EXP_{i.t}, GDP_{i.t}\}$ and $\mathbf{f}_{ojt} = \{SIM_{.jt}, RLF_{.jt}\}$. $3DCCEP_{GL}$ is the CCEP estimator with both global and local factors approximated by $\mathbf{f}_t = \{\overline{EXP}_{..t}, \overline{GDP}_{..t}, \overline{SIM}_{..t}, \overline{RLF}_{..t}, \overline{cee}_{..t}, t, t^2\}$ and $\mathbf{f}_{iot} = \{\overline{SIM}_{i.t}, \overline{RLF}_{i.t}, \overline{RER}_{i.t}\}$. Standard errors inside parenthesis. CD test refers to testing the null hypothesis of residual cross-section independence or weak dependence. α is the estimate of CSD exponent with 90% confidence bands inside parenthesis.

The $3DCCEP$ estimator wipes out the time invariant regressors. Following the 2-step approach in Serlenga and Shin (2007), we now estimate γ 's by the between estimator:

$$d_{ijt} = \alpha_{ij} + \gamma_1 DIS_{ij} + \gamma_2 BOR_{ij} + \gamma_3 LAN_{ij} + u_{ijt} \quad (71)$$

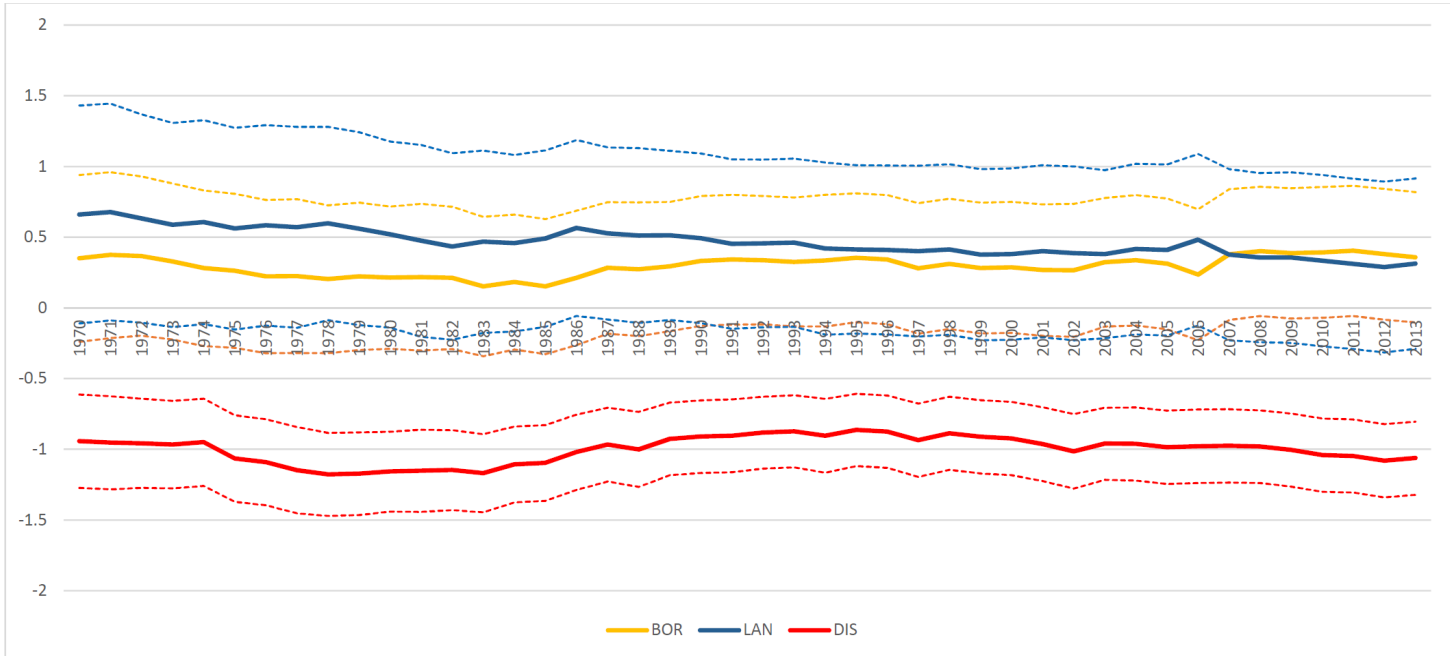
where $d_{ijt} = y_{ijt} - \hat{\beta}'\mathbf{x}_{ijt}$ with $\hat{\beta}$ being the $3DCCEP_{GL}$ estimator. We then propose to test the validity of the following hypothesis: if the Euro had a positive effect on the EU trade by reducing bilateral trade barriers and eliminating exchange-related transactions costs, this might have caused a decrease in trade impacts of bilateral barriers (e.g. Cafiso (2011)). Consequently, we examine whether the trend line of their coefficients is more downward-sloping after 1999. A declining trend will support the hypothesis that the Euro helps to

¹⁴For the FE, the impacts of home GDP on exports is surprisingly larger than the foreign impact while Euro and CEE impacts seem to be rather high. The RER coefficient is significantly negative for $3DCCEP_G$. Finally, the CEE impact is insignificant while the Euro impact is almost negligible for $3DCCEP_L$.

promote more EU integration. To this end we estimate (71) by the cross-section regression at each time period and produce the time-varying coefficients of γ in Figure 1.¹⁵

The border effect has been declining until the mid 1980s, and then becomes stable except the slight dip during the global financial crisis in 2007, albeit statistically insignificant over the whole period. The language effect has been decreasing steadily until the end of 1980's. This downward trend may reflect the progressive lessening of restrictions on labor mobility within the EU, that encouraged migration and reduced the relative importance of cultural difference. In the 90s and especially after the introduction of the Euro in 1999, both language and border effects became flat, suggesting that the EU integration may reach near-completion stage. This is in line with the currency union formation hypothesis by Frankel (2005) that countries, which decide to join a currency union, are self-selected on the basis of distinctive features shared by members (e.g. enhancing standards of harmonisation and reducing regulatory barriers). On the other hand, the distance effect started to rise slightly after 1999.

Figure 1: Time-varying Effects of Bilateral Trade Barriers for 15 EU countries



Notes: The time-varying effects of bilateral trade barriers (*DIS*, *BOR* and *LAN*) are estimated in two steps: we estimate the panel gravity model (70) by the $3DCCEP_{GL}$ estimator, and then estimate (71) by the between estimator at each time period. We also provide 90% confidence intervals.

¹⁵In the presence of endogenous bilateral barriers such as language dummies, the between estimator would be biased. It would be worthwhile to develop an extended Hausman and Taylor (1981) IV estimation, which has been developed in the 2D panels with cross-sectionally correlated errors by Serlenga and Shin (2007). But, it is beyond scope of the current paper, because such extension would involve the complex layers of correlations between regressors and hierarchical unobserved effects (e.g. Balazsi, Bun, Chan, and Harris (2017)).

5.4 Estimation Results for the 15 EU plus 4 OECD Countries

Table 6 reports the panel gravity estimation results for an enlarged sample of the 342 country-pairs among the 15 EU countries plus four more large non-EU countries (Australia, Canada, Japan and the US) over the period 1970-2013.

We focus on the estimation results for the $3DCCEP_{GL}$ with the lowest CSD. The estimation results are qualitatively similar to those reported in Table 5. All the coefficients are significant, and display expected signs. The effect of the foreign GDP is substantially higher than the home GDP effect, but slightly smaller. The effects of SIM is slightly higher (from 0.2 to 0.22) while RLF becomes negative but negligible (reflecting that the impact of RLF on total flows, the sum of inter- and intra-industry flows, might not be always unambiguous). The impact of RER shows a stronger terms of trade effect (from 0.10 to 0.18). When adding more heterogeneous control countries, we observe that (i) the impacts of GDPs become slightly smaller; (ii) the impact of similarity increases whereas RLF becomes insignificant, suggesting that the additional four non-EU countries tend to display a higher propensity to intra-industry trade (e.g. de Boyrie and Kreinin (2012)); (iii) the impact of the real exchange rate is higher, possibly reflecting the higher volatility in the enlarged sample.

The impacts of EMU and CEE on bilateral export flows are still significant, though their magnitudes become smaller than those with the 15 EU countries, namely from 3% to 1.5% and from 5% to 3%, respectively. These smaller effects for the enlarged sample with 4 additional non-EU control countries might reflect the trade diversion between the Euro and non-Euro area. The effects of the EMU will differ with respect to the selected control group and depend on the composition of treatment and control groups (e.g. Baier and Bergstrand (2009)).¹⁶ On the contrary, other estimators still provide rather misleading results. In particular, the FE estimation provides an opposite result that both impacts increase rather substantially, i.e. from 0.099 and 0.074 to 0.258 and 0.161. This finding suggests that the declining effect of the euro is further exacerbated due to the trade diversion effect in the enlarged OECD sample.

Figure 2 displays the time varying estimates of γ in (71), using the 3DCCEP estimator with both local and global factors approximation. Border and language effects display similar patterns to the case with the 15 EU countries. Here, we do not observe any evidence in favour of the Euro effect on trade integration, once again consistent with the currency union formation hypothesis by Frankel (2005). On the contrary, the effect of distance has been rather increasing over the whole period. This is broadly consistent with the meta-study by

¹⁶Santos Silva and Tenreyro (2010) find a (insignificant) positive effect of the EMU for the 12 EU countries, but a (significant) negative effect when including more (heterogeneous) countries. This indicates that sharing a common currency may have a small and potentially negative effect on bilateral trade when the control group is larger.

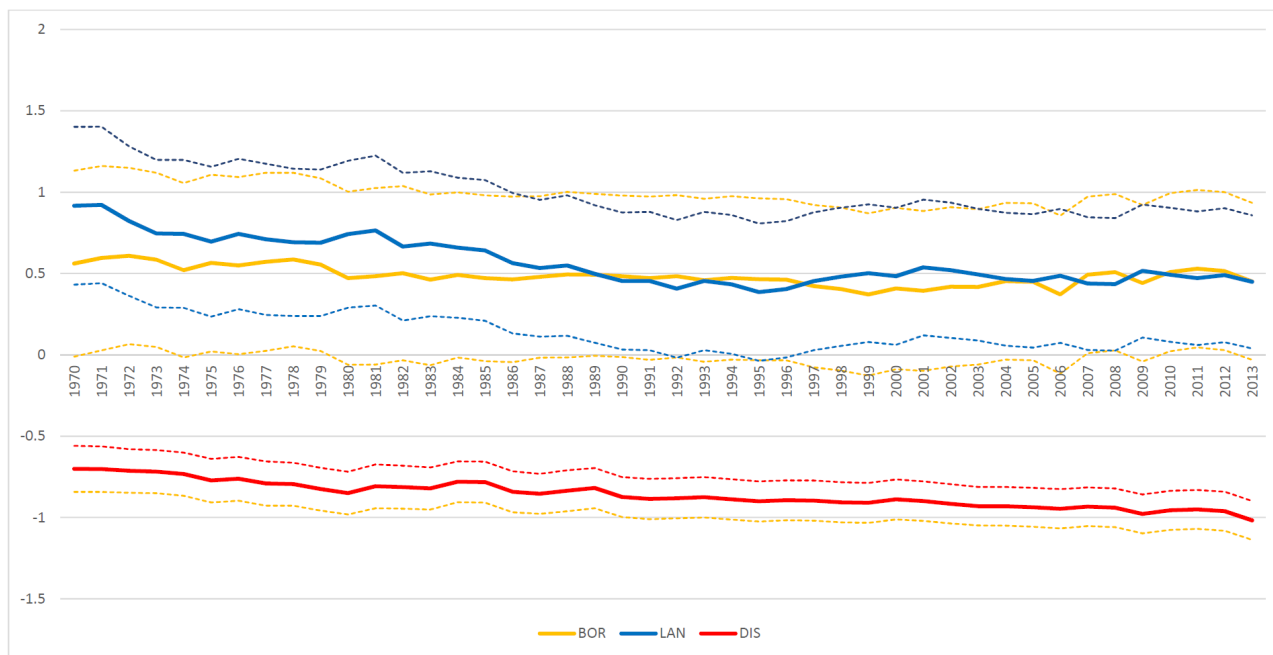
Table 6: Estimation Results for 15 EU plus 4 OECD countries

	FE	$3DCCEP_G$	$3DCCEP_L$	$3DCCEP_{GL}$
GDP_i	1.066 (0.019)	0.531 (0.016)	0.069 (0.010)	0.169 (0.055)
GDP_j	0.904 (0.020)	1.419 (0.02)	1.262 (0.015)	1.417 (0.017)
SIM	0.332 (0.029)	0.109 (0.021)	0.100 (0.013)	0.220 (0.023)
RLF	0.027 (0.004)	-0.008 (0.002)	0.010 (0.001)	-0.004 (0.002)
RER	0.058 (0.008)	0.086 (0.004)	0.074 (0.002)	0.179 (0.005)
EMU	0.258 (0.009)	0.012 (0.003)	0.012 (0.001)	0.014 (0.002)
CEE	0.161 (0.028)	0.021 (0.017)	0.007 (0.001)	0.030 (0.012)
CD stat	243.33	3.272	2.331	3.201
α	0.90 (0.88-0.92)	0.74 (0.69-0.76)	0.65 (0.61-0.69)	0.62 (0.57-0.66)

Notes: The dependent variable is the log of export. $3DCCEP_G$ is the CCEP estimator with only the global factors approximated by $\mathbf{f}_t = \{\overline{EXP}_{..t}, \overline{GDP}_{..t}, \overline{SIM}_{..t}, \overline{RLF}_{..t}, \overline{CEE}_{..t}, t\}$. $3DCCEP_L$ is the CCEP estimator with only the local factors approximated by $\mathbf{f}_t = \mathbf{f}_{iot} = \{\overline{EXP}_{i.t}, \overline{GDP}_{i.t}\}$ and $\mathbf{f}_{ojt} = \{\overline{SIM}_{.jt}, \overline{RLF}_{.jt}\}$. $3DCCEP_{GL}$ is the CCEP estimator with both global and local factors approximated by $\mathbf{f}_t = \{\overline{EXP}_{..t}, \overline{GDP}_{..t}, \overline{SIM}_{..t}, \overline{RLF}_{..t}, \overline{CEE}_{..t}, t\}$ and $\mathbf{f}_{iot} = \{\overline{SIM}_{i.t}, \overline{RLF}_{i.t}, \overline{RER}_{i.t}\}$. Standard errors inside parenthesis. CD test refers to testing the null hypothesis of residual cross-section independence or weak dependence. α is the estimate of CSD exponent with 90% confidence bands inside parenthesis.

Disdier and Head (2008), who document that the trade elasticity with respect to distance has not steadily declined over time, but rather increased recently.

Figure 2: Time-varying effects of Bilateral Trade Barriers for 15 EU plus 4 OECD countries



Notes: see the notes to Figure 1.

6 Conclusions

Given the growing availability of large panel datasets which contain information along multiple dimensions, the recent literature on panel data has focused mainly on extending two-way error component models to a multi-dimensional setting. We extend this literature, develop novel error components specifications, and propose estimation techniques to accommodate cross-sectional error dependence within 3D panel data models.

We develop a consistent estimation procedure, called the 3DCCE estimator, extending the work of Pesaran (2006) and derive its asymptotic properties. We establish that the standard CCE estimator fails to provide consistent estimation for the 3D panel data models with hierarchical multi-factor error structure. The empirical usefulness and superiority of the 3DCCE estimator are demonstrated via Monte Carlo studies and an empirical application to the 3D panel gravity model of EU export flows.

At this stage, it seems appropriate to mention a number of potential extensions and generalisations. First, it is of interest to analyse 4D or higher dimensional models. However, the proper development of hierarchical structure of such an extension and its accompanying estimators would be more challenging. Second, our proposed approach can be easily extended to dynamic panels. Finally and more importantly, following the recent studies by Bai and Li

(2015) and Shi and Lee (2017), it is of interest to develop more challenging and realistic models by combining both spatial-based and factor-based elements within 3D or higher dimensional models.

7 Appendix

Here we provide a complete proof for Theorems 2 and 3.

7.1 Proof of Theorem 2

We analyse each of the terms in (41)-(46). By independence of $\boldsymbol{\varepsilon}_{ij}$ across (i, j) , we note that

$$\begin{aligned} \text{Var} \left[\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \left(\frac{\sqrt{N} \mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \boldsymbol{\varepsilon}_{ij}}{T} \right) \right] &= \text{Var} \left[\frac{1}{\sqrt{TN}^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \boldsymbol{\varepsilon}_{ij}}{\sqrt{T}} \right) \right] \\ &= \frac{1}{TN} \text{Var} \left[\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \boldsymbol{\varepsilon}_{ij}}{\sqrt{T}} \right) \right] = O\left(\frac{1}{TN}\right) \end{aligned}$$

We now examine $\boldsymbol{\chi}_{ij}$, $\boldsymbol{\chi}_{ij,io}$ and $\boldsymbol{\chi}_{ij,oj}$ defined in (36)-(38). By independence of $\boldsymbol{\eta}_{ij}$ across (i, j) , we have:

$$\text{Var} \left[\frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij} \right] = \frac{1}{N} \text{Var} \left[\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F} \boldsymbol{\eta}_{ij}}{T} \right] = O\left(\frac{1}{N}\right) \quad (72)$$

Next,

$$\text{Var} \left[\frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij,io} \right] = \text{Var} \left[\frac{1}{N^{1/2}} \sum_{i=1}^N \left[\frac{1}{N} \sum_{j=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{oj}}{T} \right] \boldsymbol{\eta}_{io} \right] = O(1) \quad (73)$$

and further,

$$\frac{1}{N^{1/2}} \sum_{i=1}^N \left[\frac{1}{N} \sum_{j=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{oj}}{T} \right] \boldsymbol{\eta}_{io} = \frac{1}{N^{1/2}} \sum_{i=1}^N \mathbf{A}_{i\circ,NT} \boldsymbol{\eta}_{io} \xrightarrow{d} N(0, \boldsymbol{\Omega}_{\eta,\bullet\circ}) \quad (74)$$

where

$$\mathbf{A}_{i\circ,NT} = \frac{1}{N} \sum_{j=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{oj}}{T} \quad (75)$$

$$\boldsymbol{\Omega}_{\eta,\bullet\circ} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E(\mathbf{A}_{i\circ,NT} \boldsymbol{\Omega}_{\eta,\bullet\circ} \mathbf{A}'_{i\circ,NT})$$

Similarly, we have:

$$\text{Var} \left[\frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij,oj} \right] = \text{Var} \left\{ \frac{1}{N^{1/2}} \sum_{j=1}^N \left[\frac{1}{N} \sum_{i=1}^N \boldsymbol{\Psi}_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{io}}{T} \right] \boldsymbol{\eta}_{oj} \right\} = O(1) \quad (76)$$

and

$$\frac{1}{N^{1/2}} \sum_{j=1}^N \left[\frac{1}{N} \sum_{i=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i_o}}{T} \right] \boldsymbol{\eta}_{oj} = \frac{1}{N^{1/2}} \sum_{j=1}^N \mathbf{A}_{oj,NT} \boldsymbol{\eta}_{oj} \xrightarrow{d} N(0, \boldsymbol{\Omega}_{\eta, \bullet\bullet}) \quad (77)$$

where

$$\mathbf{A}_{oj,NT} = \frac{1}{N} \sum_{i=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i_o}}{T} \quad (78)$$

$$\boldsymbol{\Omega}_{\eta, \bullet\bullet} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N E \left(\mathbf{A}_{oj,NT} \boldsymbol{\Omega}_{\eta_{o\bullet}} \mathbf{A}'_{oj,NT} \right).$$

We next examine $\boldsymbol{\chi}_{ij, \bullet\bullet}$, $\boldsymbol{\chi}_{ij, \bullet\bullet}$ and $\boldsymbol{\chi}_{ij, \bullet\bullet}$, defined in (36)-(38). For $\boldsymbol{\chi}_{ij, \bullet\bullet}$, we notice that

$$\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left[\Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}}{T} \right] \xrightarrow{p} \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N E \left[\Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}}{T} \right]$$

and

$$\sqrt{N} \left(\bar{\boldsymbol{\Gamma}}'_{\nu, \bullet\bullet} + \bar{\boldsymbol{\eta}}_{\bullet\bullet} \right) = \frac{1}{\sqrt{N}} \sum_{i=1}^N \bar{\boldsymbol{\Gamma}}'_i \boldsymbol{\nu}_{i_o} + \frac{1}{\sqrt{N}} \sum_{j=1}^N \bar{\boldsymbol{\Gamma}}'_j \boldsymbol{\nu}_{oj} + \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\Gamma}'_{ij} \boldsymbol{\nu}_{ij} + \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\eta}_{ij}.$$

But,

$$\frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\Gamma}'_{ij} \boldsymbol{\nu}_{ij} = O_p \left(\frac{1}{N^{1/2}} \right), \quad \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\eta}_{ij} = O_p \left(\frac{1}{N^{1/2}} \right) \quad (79)$$

and

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \bar{\boldsymbol{\Gamma}}'_i \boldsymbol{\nu}_{i_o} \xrightarrow{d} N(0, \boldsymbol{\Omega}_{\Gamma\nu, \bullet\bullet}), \quad \frac{1}{\sqrt{N}} \sum_{j=1}^N \bar{\boldsymbol{\Gamma}}'_j \boldsymbol{\nu}_{oj} \xrightarrow{d} N(0, \boldsymbol{\Omega}_{\Gamma\nu, \bullet\bullet}) \quad (80)$$

where

$$\boldsymbol{\Omega}_{\Gamma\nu, \bullet\bullet} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E \left(\bar{\boldsymbol{\Gamma}}'_i \boldsymbol{\Omega}_{\nu, \bullet\bullet} \bar{\boldsymbol{\Gamma}}_i \right), \quad \boldsymbol{\Omega}_{\Gamma\nu, \bullet\bullet} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N E \left(\bar{\boldsymbol{\Gamma}}'_j \boldsymbol{\Omega}_{\nu, \bullet\bullet} \bar{\boldsymbol{\Gamma}}_j \right).$$

Hence, we have:

$$\frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \boldsymbol{\chi}_{ij, \bullet\bullet} \xrightarrow{d} N \left[0, \mathbf{A}_{F1} \left(\boldsymbol{\Omega}_{\Gamma\nu, \bullet\bullet} + \boldsymbol{\Omega}_{\Gamma\nu, \bullet\bullet} \right) \mathbf{A}'_{F1} \right], \quad (81)$$

where

$$\mathbf{A}_{F1} = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N E \left[\Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}}{T} \right].$$

Next, noting that

$$\left(\bar{\boldsymbol{\Gamma}}'_{i, \nu, \bullet\bullet} + \bar{\boldsymbol{\eta}}_{\bullet\bullet} \right) = \bar{\boldsymbol{\Gamma}}'_{\bullet\bullet} \boldsymbol{\nu}_{i_o} + \frac{1}{N} \sum_{j=1}^N \boldsymbol{\Gamma}'_{oj} \boldsymbol{\nu}_{oj} + \frac{1}{N} \sum_{j=1}^N \boldsymbol{\Gamma}'_{oj} \boldsymbol{\nu}_{ij} + \frac{1}{N} \sum_{j=1}^N \boldsymbol{\eta}_{oj},$$

we have for $\chi_{ij,\bullet\bullet}$:

$$\begin{aligned}
\frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \chi_{ij,\bullet\bullet} &= \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \left[\Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right] \left(\bar{\Gamma}'_{i,\nu,\bullet\bullet} - \bar{\eta}_{\bullet\bullet} \right) \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^N \left[\frac{1}{N} \sum_{j=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right] \bar{\Gamma}'_{\bullet\bullet} \boldsymbol{\nu}_{i\circ} \\
&\quad + \frac{1}{N^{3/2}} \left(\sum_{i=1}^N \sum_{j=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right) \left(\frac{1}{N} \sum_{j=1}^N \Gamma'_{\circ j} \boldsymbol{\nu}_{ij} \right) \\
&\quad + \left\{ \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right\} \times \left[\frac{1}{\sqrt{N}} \sum_{j=1}^N \Gamma'_{\circ j} \boldsymbol{\nu}_{\circ j} - \frac{1}{\sqrt{N}} \sum_{j=1}^N \boldsymbol{\eta}_{\circ j} \right]
\end{aligned}$$

But,

$$\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left[\Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right] \rightarrow^p \mathbf{A}_{F2}, \quad (82)$$

where

$$\mathbf{A}_{F2} = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N E \left[\Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right]$$

Further,

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \left[\frac{1}{N} \sum_{j=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right] \bar{\Gamma}'_{\bullet\bullet} \boldsymbol{\nu}_{i\circ} \rightarrow^d N(0, \boldsymbol{\Sigma}_{\Gamma,\bullet\bullet}), \quad (83)$$

where

$$\begin{aligned}
\boldsymbol{\Sigma}_{\Gamma,\bullet\bullet} &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E \left\{ \left[\frac{1}{N} \sum_{j=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right] \bar{\Gamma}'_{\bullet\bullet} \boldsymbol{\Omega}_{\nu,\bullet\bullet} \bar{\Gamma}_{\bullet\bullet} \left[\frac{1}{N} \sum_{j=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right]' \right\}, \\
&\quad \frac{1}{N^{1/2}} \sum_{j=1}^N \Gamma'_{\circ j} \boldsymbol{\nu}_{\circ j} \rightarrow^d N(0, \boldsymbol{\Gamma}'_{\bullet\bullet} \boldsymbol{\Omega}_{\nu,\bullet\bullet} \boldsymbol{\Gamma}_{\bullet\bullet}), \quad (84)
\end{aligned}$$

$$\frac{1}{N^{3/2}} \left(\sum_{i=1}^N \sum_{j=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right) \left(\frac{1}{N} \sum_{j=1}^N \Gamma'_{\circ j} \boldsymbol{\nu}_{ij} \right) = O_p \left(\frac{1}{N^{1/2}} \right), \quad (85)$$

$$\frac{1}{\sqrt{N}} \sum_{j=1}^N \boldsymbol{\eta}_{\circ j} \rightarrow^d N(0, \boldsymbol{\Omega}_{\eta,\bullet\bullet}). \quad (86)$$

Hence,

$$\frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \chi_{ij,\bullet\bullet} \rightarrow^d N \left[0, \mathbf{A}_{F2} (\boldsymbol{\Gamma}'_{\bullet\bullet} \boldsymbol{\Omega}_{\nu,\bullet\bullet} \boldsymbol{\Gamma}_{\bullet\bullet} + \boldsymbol{\Omega}_{\eta,\bullet\bullet}) \mathbf{A}'_{F2} + \boldsymbol{\Sigma}_{\Gamma,\bullet\bullet} \right],$$

Similarly, for $\chi_{ij,\bullet\bullet}$, we derive the result as follows:

$$\frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \chi_{ij,\bullet\bullet} \rightarrow^d N [0, \mathbf{A}_{F3} (\mathbf{\Gamma}'_{\bullet\bullet} \mathbf{\Omega}_{\nu_{\bullet\bullet}} \mathbf{\Gamma}_{\bullet\bullet} + \mathbf{\Omega}_{\eta_{\bullet\bullet}}) \mathbf{A}'_{F3} + \mathbf{\Sigma}_{\Gamma,\bullet\bullet}],$$

where

$$\mathbf{A}_{F3} = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N E \left[\Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{j\circ}}{T} \right]$$

$$\mathbf{\Sigma}_{\Gamma,\bullet\bullet} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N E \left\{ \left[\frac{1}{N} \sum_{i=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{\circ j}}{T} \right] \bar{\mathbf{\Gamma}}'_{\bullet\bullet} \mathbf{\Omega}_{\nu_{\bullet\bullet}} \bar{\mathbf{\Gamma}}_{\bullet\bullet} \left[\frac{1}{N} \sum_{i=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{\circ j}}{T} \right]' \right\},$$

Keeping the asymptotically dominant terms only, we obtain the main result as follows

$$\sqrt{N} (\hat{\mathbf{b}}_{MG} - \boldsymbol{\beta}) = \frac{1}{\sqrt{N}} \sum_{i=1}^N [\mathbf{A}_{1,i,NT} \boldsymbol{\nu}_{i\circ} + \mathbf{A}_{2,i,NT} \boldsymbol{\eta}_{i\circ}] + \frac{1}{\sqrt{N}} \sum_{j=1}^N [\mathbf{A}_{1,j,NT} \boldsymbol{\nu}_{\circ j} + \mathbf{A}_{2,j,NT} \boldsymbol{\eta}_{\circ j}] + o_p(1),$$

where

$$\begin{aligned} \mathbf{A}_{1,i,NT} &= 1 + \left[\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left(\Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}}{T} \right) \right] \bar{\mathbf{\Gamma}}'_i \\ &+ \left[\frac{1}{N} \sum_{j=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right] \bar{\mathbf{\Gamma}}'_{\bullet\bullet} + \left[\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left(\Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{j\circ}}{T} \right) \right] \mathbf{\Gamma}'_{i\circ}, \end{aligned} \quad (87)$$

$$\begin{aligned} \mathbf{A}_{1,j,NT} &= 1 + \left[\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left(\Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}}{T} \right) \right] \bar{\mathbf{\Gamma}}'_j \\ &+ \left[\frac{1}{N} \sum_{i=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{\circ j}}{T} \right] \bar{\mathbf{\Gamma}}'_{\bullet\bullet} + \left[\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left(\Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T} \right) \right] \mathbf{\Gamma}'_{\circ j}, \end{aligned} \quad (88)$$

$$\mathbf{A}_{2,i,NT} = \mathbf{A}_{i\circ,NT} + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{i\circ}}{T}, \quad (89)$$

$$\mathbf{A}_{2,j,NT} = \mathbf{A}_{\circ j,NT} + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \Psi_{Q,ijT}^{-1} \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{\circ j}}{T}. \quad (90)$$

7.2 Proof of Theorem 3

To develop the asymptotic properties of $\hat{\mathbf{b}}_P$, we proceed from (49) as follows:

$$\sqrt{N} (\hat{\mathbf{b}}_P - \boldsymbol{\beta}) =$$

$$\left(\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T} \right)^{-1} \left[\begin{aligned} & \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} (\mathbf{X}_{ij} (\boldsymbol{\nu}_{io} + \boldsymbol{\nu}_{oj} + \boldsymbol{\nu}_{ij}) + \boldsymbol{\varepsilon}_{ij})}{T} \\ & + \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} (\mathbf{F} \boldsymbol{\gamma}_{ij} + \mathbf{F}_{io} \boldsymbol{\gamma}_{oj} + \mathbf{F}_{oj} \boldsymbol{\gamma}_{io})}{T} \end{aligned} \right]$$

Similarly to the analysis for $\hat{\mathbf{b}}_{MG}$, we have:

$$\begin{aligned} & \left(\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} \mathbf{X}_{ij}}{T} \right)^{-1} \left[\begin{aligned} & \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} (\mathbf{X}_{ij} (\boldsymbol{\nu}_{io} + \boldsymbol{\nu}_{oj} + \boldsymbol{\nu}_{ij}) + \boldsymbol{\varepsilon}_{ij})}{T} + \\ & \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \bar{\mathbf{M}}_{ij} (\mathbf{F} \boldsymbol{\gamma}_{ij} + \mathbf{F}_{io} \boldsymbol{\gamma}_{oj} + \mathbf{F}_{oj} \boldsymbol{\gamma}_{io})}{T} \end{aligned} \right] \\ & = \left(\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{X}_{ij}}{T} \right)^{-1} \left[\begin{aligned} & \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} (\mathbf{X}_{ij} (\boldsymbol{\nu}_{io} + \boldsymbol{\nu}_{oj}) + \boldsymbol{\varepsilon}_{ij})}{T} + \\ & \frac{1}{N^{3/2}} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} (\mathbf{F} \boldsymbol{\gamma}_{ij} + \mathbf{F}_{io} \boldsymbol{\gamma}_{oj} + \mathbf{F}_{oj} \boldsymbol{\gamma}_{io})}{T} \end{aligned} \right] + o_p(1). \end{aligned}$$

Then, using the analysis of the terms in (41)-(46) we get the asymptotic normality result.

Finally, defining

$$\begin{aligned} \tilde{\mathbf{A}}_{1,i,NT} &= 1 + \left[\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}}{T} \right) \right] \bar{\boldsymbol{\Gamma}}'_{io} \\ &+ \left[\frac{1}{N} \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{io}}{T} \right] \bar{\boldsymbol{\Gamma}}'_{\bullet} + \left[\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{j\circ}}{T} \right) \right] \boldsymbol{\Gamma}'_{io}, \end{aligned} \quad (91)$$

$$\begin{aligned} \tilde{\mathbf{A}}_{1,j,NT} &= 1 + \left[\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}}{T} \right) \right] \bar{\boldsymbol{\Gamma}}'_{oj} \\ &+ \left[\frac{1}{N} \sum_{i=1}^N \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{oj}}{T} \right] \bar{\boldsymbol{\Gamma}}'_{\bullet} + \left[\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left(\frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{io}}{T} \right) \right] \boldsymbol{\Gamma}'_{oj}, \end{aligned}$$

$$\tilde{\mathbf{A}}_{2,i,NT} = \tilde{\mathbf{A}}_{io,NT} - \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{io}}{T},$$

$$\tilde{\mathbf{A}}_{2,j,NT} = \tilde{\mathbf{A}}_{oj,NT} - \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{oj}}{T}.$$

$$\tilde{\mathbf{A}}_{io,NT} = \frac{1}{N} \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{oj}}{T}$$

$$\tilde{\mathbf{A}}_{oj,NT} = \frac{1}{N} \sum_{j=1}^N \frac{\mathbf{X}'_{ij} \mathbf{M}_{Q,ij} \mathbf{F}_{io}}{T}, \quad (92)$$

we can derive the asymptotic variance estimator following the proof of and the discussions surrounding Theorem 2.

8 Data Appendix

We describe how the variables are constructed. All variables are converted in constant dollar prices with 2000 as the base year. Bilateral exports are defined as logarithms of real export, $X_{hft}^R = (X_{hft}^N / XPI_{US}) \times 100$, where X_{hft}^N are bilateral exports measured in millions of current US dollars; XPI_{US} are the US export price indices.

The explanatory variables can be divided in two categories: time-varying and time-invariant variables. First, we consider the following time-varying variables:

The (log of) home and foreign GDPs are defined as $\ln(GDP_{ht}^R)$ and $\ln(GDP_{ft}^R)$, respectively, where GDP_{ht}^R and GDP_{ft}^R are gross domestic products at constant dollar of home and foreign countries. GDP's are originally expressed in million Euro for the twelve countries that joined the European Monetary Union (Austria, Belgium-Luxemburg, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Portugal, Spain) and in millions of current national currency for Australia, Canada, Denmark, Japan, Norway, Sweden, the UK and the US (GDP^N). For the former twelve countries the European GDP deflator has been used whilst for the latter eight countries the original nominal values of GDP have been deflated by the GDP deflator ($GDPD$, 2000 = 100) of the respective countries. We also convert GDPs in US dollar at the exchange rate of 2000 (mean over period) in order to exclude the effect of a dollar depreciation or appreciation as $GDP_{ht}^R = (GDP_{ht}^N / GDPD_{ht}) \times (US\$/NC_h)_{2000} \times 100$, where NC_h stands for national currency of the home country.

SIM: a measure of countries' similarity in size constructed as

$$SIM_{hft} = \ln \left[1 - \left(\frac{GDP_{ht}^R}{GDP_{ht}^R + GDP_{ft}^R} \right)^2 - \left(\frac{GDP_{ft}^R}{GDP_{ft}^R + GDP_{ht}^R} \right)^2 \right];$$

RLF: a measure of countries' difference in relative factor endowment calculated as

$$RLF_{hft} = \ln |PGDP_{ft}^R - PGDP_{ht}^R|,$$

where $PGDP$ is per capita GDP.

RER: the real exchange rate. Real exchange rates in constant dollars at 2000 are defined as $RER_{hft} = NER_{hft} \times XPI_{US}$, where NER_{hft} is nominal exchange rate between currencies h and f in year t in terms of the U.S. dollars.

CEE: a dummy for European Community which is equal to one when both countries belong to the European Community.

EMU: a dummy for the European Monetary Union which is equal to one when both trading partners adopt the same currency.

Next, the time-invariant variables are:

LAN: a dummy for common language which is equal to one when both countries speak the same official language.

BOR: a dummy for common border which is equal to one when the trading partners share a border.

DIS: the (log of) distance, where the distance (DIS_{hf}) between countries is measured as the (log) of great circle distance between the most populated cities in kilometers.

The data sources are as follows: we collect bilateral nominal export data (X^N) from the IMF-Direction of Trade Statistics (DOTS); export indices and *NER* from the IMF-IFS database; GDP deflators, GDP, per capita GDP (already converted in constant dollars) from the World Bank-World Development Indicators; and finally, the time-invariant variables from CEPII.

9 Supplement

We present further Monte Carlo studies to discuss the special cases presented in Section 3.4. In particular, we are interested in understanding if a better convergence rates can be achieved in case the hierarchical structure of our model is modified or simplified. We mainly discuss of the performance of the proposed estimators under the validity of two conditions: Condition S1 ($\boldsymbol{\eta}_{i_0} = \boldsymbol{\eta}_{o_j} = \boldsymbol{\nu}_{i_0} = \boldsymbol{\nu}_{o_j} = \mathbf{0}$) and Condition S2 ($\mathbf{F}_{i_0} = \mathbf{F}_{o_j} = 0$). Hence, we focus on the following eight cases of interest:

Case 1s Condition S1 - $\boldsymbol{\eta}_{i_0} = \boldsymbol{\eta}_{o_j} = \boldsymbol{\nu}_{i_0} = \boldsymbol{\nu}_{o_j} = \mathbf{0}$ - and full rank. In this case (62) becomes

$$\beta_{ij} = \beta + \nu_{ij}, \quad \nu_{ij} \sim iidN(0, 1)$$

and (60) is replaced by

$$\begin{aligned} \gamma_{1,ij} &\sim iidN(1, 0.2) \text{ and } \gamma_{2,ij} \sim iidN(1, 0.2) \text{ for } i, j = 1, \dots, N \\ \gamma_{1,o_j} &= 1; \text{ and } \gamma_{2,o_j} = 1 \text{ for } j = 1, \dots, N \\ \gamma_{1,i_0} &= 1 \text{ and } \gamma_{2,i_0} = 1 \text{ for } i = 1, \dots, N \end{aligned}$$

Case 2s Condition S1 ($\boldsymbol{\eta}_{i_0} = \boldsymbol{\eta}_{o_j} = \boldsymbol{\nu}_{i_0} = \boldsymbol{\nu}_{o_j} = \mathbf{0}$) and rank deficiency. Here, (62) becomes

$$\beta_{ij} = \beta + \nu_{ij}, \quad \nu_{ij} \sim iidN(0, 1)$$

and (61) is replaced by

$$\begin{aligned} \gamma_{1,ij} &\sim iidN(1, 0.2) \text{ and } \gamma_{2,ij} \sim iidN(0, 0.2) \text{ for } i, j = 1, \dots, N \\ \gamma_{1,o_j} &= 1 \text{ and } \gamma_{2,o_j} = 0 \text{ for } j = 1, \dots, N \\ \gamma_{1,i_0} &= 1 \text{ and } \gamma_{2,i_0} = 0 \text{ for } i = 1, \dots, N \end{aligned}$$

Case 3s Full homogeneity, where $\boldsymbol{\eta}_{ij} = \boldsymbol{\eta}_{i_0} = \boldsymbol{\eta}_{o_j} = \boldsymbol{\nu}_{ij} = \boldsymbol{\nu}_{i_0} = \boldsymbol{\nu}_{o_j} = 0$ and full rank. In this case (62) becomes

$$\beta_{ij} = 1$$

and (60) is replaced by

$$\begin{aligned} \gamma_{1,ij} &= 1 \text{ and } \gamma_{2,ij} = 1 \text{ for } i, j = 1, \dots, N \\ \gamma_{1,o_j} &= 1 \text{ and } \gamma_{2,o_j} = 1 \text{ for } j = 1, \dots, N \\ \gamma_{1,i_0} &= 1 \text{ and } \gamma_{2,i_0} = 1 \text{ for } i = 1, \dots, N \end{aligned}$$

Case 4s Full homogeneity, where $\boldsymbol{\eta}_{ij} = \boldsymbol{\eta}_{i_0} = \boldsymbol{\eta}_{o_j} = \boldsymbol{\nu}_{ij} = \boldsymbol{\nu}_{i_0} = \boldsymbol{\nu}_{o_j} = 0$ and rank deficiency. In this case (62) becomes

$$\beta_{ij} = 1$$

and (61) is replaced by

$$\begin{aligned}\gamma_{1,ij} &= 1 \text{ and } \gamma_{2,ij} = 0 \text{ for } i, j = 1, \dots, N \\ \gamma_{1,oj} &= 1 \text{ and } \gamma_{2,oj} = 0 \text{ for } j = 1, \dots, N \\ \gamma_{1,i0} &= 1 \text{ and } \gamma_{2,i0} = 0 \text{ for } i = 1, \dots, N\end{aligned}$$

Case 5s Condition S1, where $\boldsymbol{\eta}_{i0} = \boldsymbol{\eta}_{oj} = \boldsymbol{\nu}_{i0} = \boldsymbol{\nu}_{oj} = \mathbf{0}$, Condition S2, where $\mathbf{F}_{i0} = \mathbf{F}_{oj} = 0$, and full rank. Here, the DGP (58)-(59) is replaced by

$$\begin{aligned}y_{ijt} &= \beta_{ij}x_{ijt} + \gamma_{1,ij}f_{1,t} + \gamma_{2,ij}f_{2,t} + \varepsilon_{ijt}, \\ x_{ijt} &= \Gamma_{1,ij}f_{1,t} + \Gamma_{2,ij}f_{2,t} + v_{ijt};\end{aligned}$$

(62) becomes

$$\beta_{ij} = \beta + \nu_{ij}, \quad \nu_{ij} \sim iidN(0, 1)$$

and condition (60) is replaced by

$$\begin{aligned}\gamma_{1,ij} &\sim iidN(1, 0.2) \text{ and } \gamma_{2,ij} \sim iidN(1, 0.2) \text{ for } i, j = 1, \dots, N \\ \gamma_{1,oj} &= 1 \text{ and } \gamma_{2,oj} = 1 \text{ for } j = 1, \dots, N \\ \gamma_{1,i0} &= 1 \text{ and } \gamma_{2,i0} = 1 \text{ for } i = 1, \dots, N\end{aligned}$$

Case 6s Condition S1, where $\boldsymbol{\eta}_{i0} = \boldsymbol{\eta}_{oj} = \boldsymbol{\nu}_{i0} = \boldsymbol{\nu}_{oj} = \mathbf{0}$, and Condition S2, where $\mathbf{F}_{i0} = \mathbf{F}_{oj} = 0$, and rank deficiency. Here, the DGP (58)-(59) becomes

$$\begin{aligned}y_{ijt} &= \beta_{ij}x_{ijt} + \gamma_{1,ij}f_{1,t} + \gamma_{2,ij}f_{2,t} + \varepsilon_{ijt}, \\ x_{ijt} &= \Gamma_{1,ij}f_{1,t} + \Gamma_{2,ij}f_{2,t} + v_{ijt},\end{aligned}$$

(62) becomes

$$\beta_{ij} = \beta + \nu_{ij}, \quad \nu_{ij} \sim iidN(0, 1)$$

and (61) is replaced by

$$\begin{aligned}\gamma_{1,ij} &\sim iidN(1, 0.2) \text{ and } \gamma_{2,ij} \sim iidN(0, 0.2) \text{ for } i, j = 1, \dots, N \\ \gamma_{1,oj} &= 1 \text{ and } \gamma_{2,oj} = 0 \text{ for } j = 1, \dots, N \\ \gamma_{1,i0} &= 1 \text{ and } \gamma_{2,i0} = 0 \text{ for } i = 1, \dots, N\end{aligned}$$

Case 7s Full homogeneity, where $\boldsymbol{\eta}_{ij} = \boldsymbol{\eta}_{i0} = \boldsymbol{\eta}_{oj} = \boldsymbol{\nu}_{ij} = \boldsymbol{\nu}_{i0} = \boldsymbol{\nu}_{oj} = 0$, Condition S2, where $\mathbf{F}_{i0} = \mathbf{F}_{oj} = 0$, and full rank. Here, the DGP (58)-(59) becomes

$$y_{ijt} = \beta_{ij}x_{ijt} + \gamma_{1,ij}f_{1,t} + \gamma_{2,ij}f_{2,t} + \varepsilon_{ijt},$$

$$x_{ijt} = \Gamma_{1,ij}f_{1,t} + \Gamma_{2,ij}f_{2,t} + v_{ijt},$$

with

$$\beta_{ij} = 1$$

and (60) is replaced by

$$\gamma_{1,ij} = 1 \text{ and } \gamma_{2,ij} = 1 \text{ for } i, j = 1, \dots, N$$

$$\gamma_{1,oj} = 1 \text{ and } \gamma_{2,oj} = 1 \text{ for } j = 1, \dots, N$$

$$\gamma_{1,i\circ} = 1 \text{ and } \gamma_{2,i\circ} = 1 \text{ for } i = 1, \dots, N$$

Case 8s Full homogeneity, where $\boldsymbol{\eta}_{ij} = \boldsymbol{\eta}_{i\circ} = \boldsymbol{\eta}_{oj} = \boldsymbol{\nu}_{ij} = \boldsymbol{\nu}_{i\circ} = \boldsymbol{\nu}_{oj} = 0$, Condition S2, where $\mathbf{F}_{i\circ} = \mathbf{F}_{oj} = 0$, and rank deficiency. Here, (58)-(59) become

$$y_{ijt} = \beta_{ij}x_{ijt} + \gamma_{1,ij}f_{1,t} + \gamma_{2,ij}f_{2,t} + \varepsilon_{ijt},$$

$$x_{ijt} = \Gamma_{1,ij}f_{1,t} + \Gamma_{2,ij}f_{2,t} + v_{ijt},$$

with

$$\beta_{ij} = 1$$

and (61) is replaced by

$$\gamma_{1,ij} = 1 \text{ and } \gamma_{2,ij} = 0 \text{ for } i, j = 1, \dots, N$$

$$\gamma_{1,oj} = 1 \text{ and } \gamma_{2,oj} = 0 \text{ for } j = 1, \dots, N$$

$$\gamma_{1,i\circ} = 1 \text{ and } \gamma_{2,i\circ} = 0 \text{ for } i = 1, \dots, N$$

9.1 Simulation Results

Table 7 reports results for Case 1s. The bias turns out to be similar for the $3DCCE_G$, reduced for $3DCCE_L$ and rather negligible for $3DCCE_{GL}$. In all cases, both bias and the RMSE decline with the N dimension and are rather stable over T . The coverage rates are mostly above the nominal level and, in the case of $3DCCE_{GL}$, approach the 95% as N grows. Similar patterns are shown in Table 8 when Condition S1 holds and the rank condition is not satisfied, Case 2s.

Table 9 and Table 10 report results assuming $\boldsymbol{\eta}_{ij} = \boldsymbol{\eta}_{i\circ} = \boldsymbol{\eta}_{oj} = \boldsymbol{\nu}_{ij} = \boldsymbol{\nu}_{i\circ} = \boldsymbol{\nu}_{oj} = 0$ (full homogeneity) with full rank and rank deficiency, Cases 3s and 4s, respectively. The patterns of bias and RMSE are similar to those shown in Table 7 and 8.

Tables 11 and 12 report results imposing $\boldsymbol{\eta}_{i\circ} = \boldsymbol{\eta}_{oj} = \boldsymbol{\nu}_{i\circ} = \boldsymbol{\nu}_{oj} = 0$ (Condition S1) and $\mathbf{F}_{i\circ} = \mathbf{F}_{oj} = 0$ (Condition S2) with full rank and rank deficiency, Cases 5s and 6s, respectively. As discussed, if S2 holds, the setup reverts to that of Pesaran (2006) and no hierarchical

structure exists. Indeed, in those cases the patterns of bias and RMSE improve significantly for $3DCCE_G$. The coverage rates of $3DCCE_G$, $3DCCE_L$ and $3DCCE_{GL}$ approach the 95% as N grows.

Tables 13 and 14 report results setting $\boldsymbol{\eta}_{ij} = \boldsymbol{\eta}_{i\circ} = \boldsymbol{\eta}_{\circ j} = \boldsymbol{\nu}_{ij} = \boldsymbol{\nu}_{i\circ} = \boldsymbol{\nu}_{\circ j} = 0$ (full homogeneity) and $\mathbf{F}_{i\circ} = \mathbf{F}_{\circ j} = 0$ (Condition S2) with full rank and rank deficiency, Cases 7s and 8s, respectively. Under those assumptions also the performances of the $3DCCE_L$ improve significantly, especially when the rank condition is not satisfied.

Table 7: Simulation results for Case 1s

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.246	0.250	0.071	0.075	0.001	0.006
	25	0.269	0.268	0.067	0.067	0.001	-0.001
	50	0.277	0.278	0.070	0.070	-0.001	0.000
	100	0.282	0.282	0.074	0.075	0.000	0.001
0.5	10	0.270	0.272	0.080	0.082	0.004	0.008
	25	0.292	0.290	0.074	0.073	0.001	0.001
	50	0.297	0.299	0.075	0.077	0.000	0.001
	100	0.303	0.303	0.080	0.081	0.001	0.001
Mean Group Estimator							
0	10	0.235	0.239	0.060	0.064	0.007	0.012
	25	0.253	0.253	0.058	0.058	0.005	0.004
	50	0.261	0.262	0.059	0.060	0.001	0.002
	100	0.265	0.266	0.062	0.063	0.001	0.001
0.5	10	0.263	0.263	0.072	0.073	0.014	0.015
	25	0.279	0.278	0.065	0.065	0.006	0.006
	50	0.284	0.285	0.065	0.066	0.002	0.003
	100	0.289	0.289	0.067	0.067	0.001	0.001
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.275	0.275	0.137	0.131	0.115	0.107
	25	0.273	0.272	0.081	0.080	0.045	0.043
	50	0.278	0.279	0.074	0.074	0.021	0.021
	100	0.282	0.283	0.075	0.076	0.010	0.011
0.5	10	0.297	0.297	0.140	0.139	0.114	0.111
	25	0.296	0.294	0.087	0.087	0.044	0.045
	50	0.298	0.300	0.079	0.080	0.021	0.021
	100	0.304	0.304	0.082	0.082	0.011	0.010
Mean Group Estimator							
0	10	0.261	0.261	0.125	0.118	0.108	0.099
	25	0.257	0.257	0.072	0.071	0.042	0.040
	50	0.262	0.263	0.063	0.063	0.020	0.020
	100	0.265	0.266	0.063	0.063	0.010	0.010
0.5	10	0.285	0.285	0.127	0.126	0.105	0.103
	25	0.282	0.282	0.076	0.078	0.040	0.042
	50	0.285	0.286	0.068	0.069	0.020	0.020
	100	0.289	0.289	0.068	0.068	0.010	0.010
Panel C: Coverage rate at 95 confidence level							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.451	0.425	0.848	0.848	0.859	0.891
	25	0.000	0.000	0.661	0.657	0.917	0.927
	50	0.000	0.000	0.142	0.118	0.931	0.933
	100	0.000	0.000	0.000	0.000	0.948	0.937
0.5	10	0.423	0.380	0.836	0.810	0.879	0.880
	25	0.000	0.000	0.604	0.603	0.941	0.925
	50	0.000	0.000	0.110	0.090	0.946	0.943
	100	0.000	0.000	0.000	0.000	0.943	0.955
Mean Group Estimator							
0	10	0.405	0.390	0.875	0.870	0.896	0.898
	25	0.000	0.000	0.668	0.659	0.930	0.927
	50	0.000	0.000	0.141	0.163	0.946	0.936
	100	0.000	0.000	0.000	0.000	0.943	0.943
0.5	10	0.346	0.343	0.855	0.838	0.886	0.885
	25	0.000	0.000	0.616	0.630	0.926	0.924
	50	0.000	0.000	0.131	0.118	0.929	0.944
	100	0.000	0.000	0.000	0.000	0.943	0.950

Notes: See notes to Table 1.

Table 8: Simulation results for Case 2s

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.230	0.228	0.077	0.074	0.004	0.001
	25	0.242	0.243	0.062	0.063	0.000	0.001
	50	0.247	0.247	0.057	0.056	0.000	-0.001
	100	0.252	0.251	0.054	0.054	0.000	0.000
0.5	10	0.250	0.256	0.075	0.082	-0.001	0.009
	25	0.269	0.267	0.067	0.064	0.004	0.001
	50	0.272	0.272	0.060	0.059	0.001	0.001
	100	0.276	0.276	0.057	0.056	0.001	0.000
Mean Group Estimator							
0	10	0.223	0.218	0.054	0.049	0.013	0.008
	25	0.232	0.232	0.047	0.047	0.007	0.007
	50	0.236	0.236	0.046	0.046	0.004	0.004
	100	0.239	0.239	0.046	0.045	0.002	0.002
0.5	10	0.243	0.249	0.056	0.063	0.011	0.018
	25	0.259	0.257	0.055	0.052	0.012	0.009
	50	0.261	0.261	0.050	0.051	0.006	0.006
	100	0.264	0.264	0.050	0.049	0.004	0.003
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.258	0.254	0.137	0.131	0.111	0.107
	25	0.246	0.247	0.077	0.077	0.045	0.044
	50	0.249	0.248	0.062	0.061	0.024	0.022
	100	0.252	0.251	0.055	0.055	0.011	0.011
0.5	10	0.277	0.281	0.135	0.140	0.111	0.113
	25	0.274	0.271	0.081	0.078	0.045	0.045
	50	0.274	0.273	0.064	0.064	0.022	0.023
	100	0.277	0.276	0.058	0.058	0.012	0.012
Mean Group Estimator							
0	10	0.261	0.261	0.125	0.118	0.108	0.099
	25	0.257	0.257	0.072	0.071	0.042	0.040
	50	0.262	0.263	0.063	0.063	0.020	0.020
	100	0.239	0.239	0.047	0.047	0.010	0.010
0.5	10	0.285	0.285	0.127	0.126	0.105	0.103
	25	0.282	0.282	0.076	0.078	0.040	0.042
	50	0.285	0.286	0.068	0.069	0.020	0.020
	100	0.265	0.264	0.051	0.050	0.011	0.010
Panel C: Coverage rate at 95 confidence level							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.479	0.500	0.817	0.838	0.882	0.899
	25	0.001	0.000	0.678	0.669	0.929	0.928
	50	0.000	0.000	0.273	0.281	0.934	0.939
	100	0.000	0.000	0.002	0.003	0.943	0.945
0.5	10	0.420	0.390	0.826	0.802	0.883	0.878
	25	0.001	0.001	0.627	0.677	0.945	0.928
	50	0.000	0.000	0.249	0.247	0.947	0.950
	100	0.000	0.000	0.005	0.002	0.937	0.947
Mean Group Estimator							
0	10	0.439	0.468	0.890	0.887	0.916	0.906
	25	0.000	0.001	0.755	0.762	0.936	0.915
	50	0.000	0.000	0.370	0.346	0.934	0.931
	100	0.000	0.000	0.003	0.007	0.940	0.942
0.5	10	0.341	0.332	0.855	0.869	0.885	0.901
	25	0.000	0.000	0.726	0.708	0.928	0.934
	50	0.000	0.000	0.273	0.293	0.941	0.939
	100	0.000	0.000	0.002	0.002	0.936	0.929

Notes: See notes to Table 1.

Table 9: Simulation results for Case 3s

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.246	0.246	0.067	0.068	0.000	0.000
	25	0.270	0.269	0.068	0.068	0.000	0.000
	50	0.277	0.278	0.071	0.071	0.000	0.000
	100	0.282	0.282	0.075	0.075	0.000	0.000
0.5	10	0.267	0.266	0.074	0.074	0.000	0.000
	25	0.289	0.290	0.074	0.074	0.000	0.000
	50	0.299	0.299	0.077	0.078	0.000	0.000
	100	0.303	0.303	0.081	0.081	0.000	0.000
Mean Group Estimator							
0	10	0.235	0.236	0.063	0.064	0.000	0.000
	25	0.255	0.256	0.060	0.060	0.000	0.000
	50	0.262	0.262	0.061	0.061	0.000	0.000
	100	0.266	0.265	0.063	0.063	0.000	0.000
0.5	10	0.257	0.256	0.070	0.069	0.000	0.000
	25	0.277	0.277	0.066	0.065	0.000	0.000
	50	0.286	0.285	0.066	0.067	0.000	0.000
	100	0.289	0.289	0.068	0.068	0.000	0.000
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.249	0.249	0.069	0.069	0.011	0.008
	25	0.271	0.270	0.069	0.068	0.005	0.003
	50	0.277	0.278	0.071	0.072	0.003	0.002
	100	0.282	0.282	0.075	0.075	0.002	0.001
0.5	10	0.270	0.269	0.076	0.076	0.013	0.009
	25	0.290	0.290	0.075	0.074	0.006	0.004
	50	0.300	0.299	0.078	0.078	0.003	0.002
	100	0.303	0.303	0.082	0.082	0.002	0.001
Mean Group Estimator							
0	10	0.239	0.239	0.065	0.065	0.012	0.008
	25	0.256	0.256	0.061	0.061	0.005	0.004
	50	0.262	0.262	0.061	0.061	0.003	0.002
	100	0.266	0.265	0.063	0.063	0.002	0.001
0.5	10	0.261	0.260	0.073	0.071	0.014	0.010
	25	0.278	0.278	0.067	0.066	0.006	0.004
	50	0.286	0.286	0.067	0.067	0.003	0.002
	100	0.289	0.289	0.068	0.068	0.002	0.001

Notes: See notes to Table 1.

Table 10: Simulation results for Case 4s

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.228	0.228	0.062	0.062	0.000	0.000
	25	0.243	0.242	0.055	0.055	0.000	0.000
	50	0.248	0.248	0.053	0.053	0.000	0.000
	100	0.251	0.251	0.052	0.052	0.000	0.000
0.5	10	0.249	0.250	0.066	0.066	0.000	0.000
	25	0.267	0.267	0.059	0.059	0.000	0.000
	50	0.272	0.273	0.056	0.056	0.000	0.000
	100	0.276	0.276	0.055	0.055	0.000	0.000
Mean Group Estimator							
0	10	0.220	0.220	0.059	0.059	0.001	0.000
	25	0.232	0.232	0.051	0.051	0.000	0.000
	50	0.236	0.236	0.048	0.048	0.000	0.000
	100	0.239	0.239	0.047	0.047	0.000	0.000
0.5	10	0.242	0.243	0.064	0.064	0.000	0.000
	25	0.256	0.257	0.055	0.055	0.000	0.000
	50	0.261	0.261	0.052	0.052	0.000	0.000
	100	0.264	0.264	0.050	0.050	0.000	0.000
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.229	0.229	0.063	0.063	0.011	0.007
	25	0.243	0.243	0.056	0.056	0.004	0.003
	50	0.248	0.248	0.053	0.053	0.002	0.001
	100	0.251	0.251	0.052	0.052	0.001	0.001
0.5	10	0.251	0.252	0.068	0.067	0.012	0.009
	25	0.267	0.267	0.059	0.059	0.005	0.003
	50	0.272	0.273	0.056	0.056	0.002	0.002
	100	0.276	0.276	0.055	0.055	0.001	0.001
Mean Group Estimator							
0	10	0.221	0.221	0.060	0.060	0.011	0.007
	25	0.233	0.233	0.051	0.052	0.005	0.003
	50	0.237	0.237	0.048	0.048	0.002	0.002
	100	0.239	0.239	0.047	0.047	0.001	0.001
0.5	10	0.244	0.245	0.066	0.065	0.013	0.009
	25	0.257	0.257	0.055	0.055	0.005	0.004
	50	0.261	0.261	0.052	0.052	0.003	0.002
	100	0.264	0.264	0.050	0.050	0.001	0.001

Notes: See notes to Table 1.

Table 11: Simulation results for Case 5s

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	-0.002	-0.004	0.013	0.008	0.000	-0.004
	25	-0.001	-0.001	0.002	0.002	-0.001	0.000
	50	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	100	0.000	0.000	0.000	0.000	0.000	0.000
0.5	10	-0.007	0.002	0.008	0.016	-0.007	0.002
	25	0.000	-0.002	0.003	0.001	0.001	-0.002
	50	-0.001	0.001	0.000	0.002	-0.001	0.001
	100	0.000	0.000	0.000	0.000	0.000	0.000
Mean Group Estimator							
0	10	-0.002	-0.004	-0.001	-0.004	-0.001	-0.004
	25	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	50	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	100	0.000	0.000	0.000	0.000	0.000	0.000
0.5	10	-0.006	0.002	-0.005	0.003	-0.006	0.002
	25	0.000	-0.001	0.001	-0.001	0.000	-0.002
	50	-0.001	0.001	-0.001	0.001	-0.001	0.001
	100	0.000	0.000	0.000	0.000	0.000	0.000
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.103	0.102	0.108	0.106	0.105	0.103
	25	0.041	0.042	0.041	0.042	0.041	0.042
	50	0.020	0.021	0.020	0.021	0.020	0.021
	100	0.010	0.010	0.011	0.010	0.010	0.010
0.5	10	0.105	0.104	0.107	0.109	0.106	0.105
	25	0.042	0.041	0.043	0.042	0.043	0.041
	50	0.021	0.021	0.021	0.021	0.021	0.021
	100	0.011	0.010	0.011	0.010	0.011	0.010
Mean Group Estimator							
0	10	0.101	0.101	0.102	0.102	0.102	0.101
	25	0.040	0.041	0.040	0.042	0.040	0.041
	50	0.019	0.021	0.019	0.021	0.019	0.021
	100	0.010	0.010	0.010	0.010	0.010	0.010
0.5	10	0.102	0.102	0.102	0.103	0.102	0.103
	25	0.041	0.040	0.041	0.040	0.041	0.041
	50	0.020	0.020	0.020	0.020	0.020	0.020
	100	0.010	0.010	0.010	0.010	0.010	0.010
Panel C: Coverage rate at 95 confidence level							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.948	0.944	0.918	0.924	0.869	0.886
	25	0.944	0.943	0.942	0.936	0.926	0.911
	50	0.963	0.941	0.960	0.939	0.953	0.929
	100	0.943	0.951	0.941	0.946	0.933	0.946
0.5	10	0.943	0.951	0.918	0.903	0.873	0.865
	25	0.948	0.952	0.948	0.941	0.919	0.921
	50	0.960	0.945	0.961	0.937	0.947	0.930
	100	0.944	0.949	0.941	0.947	0.935	0.946
Mean Group Estimator							
0	10	0.951	0.951	0.921	0.918	0.913	0.914
	25	0.943	0.950	0.938	0.939	0.940	0.937
	50	0.960	0.939	0.962	0.931	0.960	0.931
	100	0.946	0.954	0.944	0.953	0.948	0.950
0.5	10	0.951	0.951	0.924	0.918	0.910	0.915
	25	0.950	0.950	0.944	0.942	0.945	0.940
	50	0.963	0.955	0.957	0.947	0.956	0.948
	100	0.949	0.956	0.946	0.954	0.946	0.952

Notes: See notes to Table 1.

Table 12: Simulation results for Case 6s

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.002	-0.004	0.044	0.038	0.002	-0.004
	25	-0.001	-0.001	0.021	0.022	-0.001	0.000
	50	0.001	0.001	0.013	0.013	0.001	0.001
	100	0.000	0.000	0.006	0.006	0.000	0.000
0.5	10	0.004	0.003	0.044	0.046	0.003	0.003
	25	0.001	0.001	0.023	0.023	0.001	0.000
	50	0.000	0.000	0.012	0.013	0.000	0.001
	100	0.000	0.000	0.007	0.007	0.000	0.001
Mean Group Estimator							
0	10	0.002	-0.004	0.002	-0.004	0.001	-0.004
	25	0.000	0.000	-0.001	0.000	0.000	0.000
	50	0.001	0.001	0.001	0.001	0.001	0.001
	100	0.000	0.000	0.000	0.000	0.000	0.000
0.5	10	0.004	0.000	0.005	0.002	0.005	0.001
	25	0.001	0.001	0.001	0.001	0.001	0.001
	50	0.000	0.001	0.000	0.000	0.000	0.001
	100	0.000	0.000	0.000	0.000	0.000	0.001
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.111	0.110	0.115	0.114	0.106	0.105
	25	0.045	0.046	0.048	0.050	0.042	0.044
	50	0.023	0.022	0.025	0.025	0.021	0.021
	100	0.011	0.011	0.012	0.013	0.011	0.011
0.5	10	0.113	0.111	0.119	0.117	0.109	0.107
	25	0.045	0.045	0.050	0.049	0.043	0.042
	50	0.023	0.022	0.026	0.025	0.022	0.021
	100	0.011	0.011	0.013	0.013	0.011	0.010
Mean Group Estimator							
0	10	0.099	0.102	0.099	0.101	0.099	0.101
	25	0.040	0.042	0.040	0.042	0.040	0.042
	50	0.021	0.020	0.021	0.020	0.021	0.020
	100	0.010	0.011	0.010	0.010	0.010	0.010
0.5	10	0.103	0.101	0.104	0.101	0.104	0.101
	25	0.041	0.041	0.040	0.041	0.040	0.041
	50	0.021	0.021	0.021	0.021	0.021	0.021
	100	0.010	0.010	0.010	0.010	0.010	0.010
Panel C: Coverage rate at 95 confidence level							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.944	0.950	0.889	0.882	0.884	0.879
	25	0.949	0.930	0.914	0.888	0.915	0.903
	50	0.955	0.947	0.897	0.888	0.951	0.937
	100	0.957	0.953	0.904	0.913	0.951	0.945
0.5	10	0.947	0.943	0.886	0.892	0.876	0.862
	25	0.956	0.955	0.904	0.899	0.936	0.930
	50	0.952	0.950	0.906	0.906	0.928	0.921
	100	0.950	0.949	0.903	0.901	0.936	0.942
Mean Group Estimator							
0	10	0.951	0.944	0.929	0.922	0.927	0.918
	25	0.950	0.940	0.946	0.932	0.946	0.932
	50	0.953	0.958	0.947	0.953	0.949	0.952
	100	0.953	0.946	0.954	0.945	0.952	0.946
0.5	10	0.950	0.949	0.924	0.917	0.921	0.914
	25	0.960	0.961	0.949	0.947	0.951	0.948
	50	0.945	0.946	0.943	0.937	0.946	0.938
	100	0.946	0.958	0.945	0.952	0.943	0.950

Notes: See notes to Table 1.

Table 13: Simulation results for Case 7s

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.001	-0.001	0.005	0.003	0.001	-0.001
	25	0.000	0.000	0.001	0.001	0.000	0.000
	50	0.000	0.000	0.000	0.000	0.000	0.000
	100	0.000	0.000	0.000	0.000	0.000	0.000
0.5	10	0.000	0.000	0.004	0.004	0.000	0.000
	25	0.001	0.000	0.001	0.001	0.001	0.000
	50	0.000	0.000	0.000	0.000	0.000	0.000
	100	0.000	0.000	0.000	0.000	0.000	0.000
Mean Group Estimator							
0	10	0.001	-0.001	0.006	0.004	0.001	-0.001
	25	0.000	0.000	0.001	0.001	0.000	0.000
	50	0.000	0.000	0.000	0.000	0.000	0.000
	100	0.000	0.000	0.000	0.000	0.000	0.000
0.5	10	0.000	0.000	0.005	0.005	-0.001	0.000
	25	0.001	0.000	0.002	0.001	0.001	0.000
	50	0.000	0.000	0.000	0.001	0.000	0.000
	100	0.000	0.000	0.000	0.000	0.000	0.000
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.014	0.010	0.015	0.011	0.016	0.011
	25	0.006	0.004	0.006	0.004	0.007	0.004
	50	0.003	0.002	0.003	0.002	0.003	0.002
	100	0.002	0.001	0.002	0.001	0.002	0.001
0.5	10	0.018	0.013	0.019	0.013	0.021	0.014
	25	0.008	0.005	0.008	0.005	0.008	0.005
	50	0.004	0.003	0.004	0.003	0.004	0.003
	100	0.002	0.001	0.002	0.001	0.002	0.001
Mean Group Estimator							
0	10	0.014	0.010	0.016	0.011	0.017	0.011
	25	0.006	0.004	0.006	0.004	0.007	0.004
	50	0.003	0.002	0.003	0.002	0.003	0.002
	100	0.002	0.001	0.002	0.001	0.002	0.001
0.5	10	0.018	0.013	0.020	0.014	0.021	0.014
	25	0.008	0.005	0.008	0.005	0.008	0.005
	50	0.004	0.003	0.004	0.003	0.004	0.003
	100	0.002	0.001	0.002	0.001	0.002	0.001

Notes: See notes to Table 1.

Table 14: Simulation results for Case 8s

		$3DCCE_G$		$3DCCE_L$		$3DCCE_{GL}$	
Panel A: Bias							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.000	0.000	0.013	0.013	0.000	0.000
	25	0.000	0.000	0.006	0.006	0.000	0.000
	50	0.000	0.000	0.003	0.003	0.000	0.000
	100	0.000	0.000	0.002	0.002	0.000	0.000
0.5	10	0.000	0.000	0.014	0.013	0.001	0.000
	25	-0.001	0.000	0.005	0.006	-0.001	0.000
	50	0.000	0.000	0.003	0.003	0.000	0.000
	100	0.000	0.000	0.002	0.002	0.000	0.000
Mean Group Estimator							
0	10	0.000	0.000	0.014	0.014	0.000	0.000
	25	0.000	0.000	0.007	0.006	0.000	0.000
	50	0.000	0.000	0.003	0.003	0.000	0.000
	100	0.000	0.000	0.002	0.002	0.000	0.000
0.5	10	0.001	0.000	0.014	0.014	0.001	0.000
	25	-0.001	0.000	0.005	0.006	-0.001	0.000
	50	0.000	0.000	0.004	0.003	0.000	0.000
	100	0.000	0.000	0.002	0.002	0.000	0.000
Panel B: RMSE							
Pooled Estimator							
ρ	N/T	50	100	50	100	50	100
0	10	0.013	0.009	0.020	0.016	0.016	0.010
	25	0.005	0.004	0.008	0.007	0.006	0.004
	50	0.003	0.002	0.004	0.004	0.003	0.002
	100	0.001	0.001	0.002	0.002	0.001	0.001
0.5	10	0.016	0.011	0.023	0.018	0.019	0.013
	25	0.007	0.005	0.009	0.008	0.007	0.005
	50	0.003	0.002	0.005	0.004	0.004	0.003
	100	0.002	0.001	0.002	0.002	0.002	0.001
Mean Group Estimator							
0	10	0.014	0.009	0.021	0.017	0.016	0.011
	25	0.005	0.004	0.009	0.007	0.006	0.004
	50	0.003	0.002	0.004	0.004	0.003	0.002
	100	0.001	0.001	0.002	0.002	0.002	0.001
0.5	10	0.016	0.012	0.023	0.019	0.019	0.014
	25	0.007	0.005	0.009	0.008	0.008	0.005
	50	0.003	0.002	0.005	0.004	0.004	0.003
	100	0.002	0.001	0.003	0.002	0.002	0.001

Notes: See notes to Table 1.

References

- ACEMOGLU, D., V. M. CARVALHO, A. OZDAGLAR, AND A. TAHBAZ-SALEHI (2012): “The Network Origins of Aggregate Fluctuations,” *Econometrica*, 80(5), 1977–2016.
- ANDERSON, J. E., AND E. VAN WINCOOP (2003): “Gravity with Gravitas: A Solution to the Border Puzzle,” *American Economic Review*, 93(1), 170–192.
- BAI, J., AND K. LI (2015): “Dynamic Spatial Panel Data Models with Common Shocks,” Working papers, Columbia University.
- BAIER, S., AND J. BERGSTRAND (2009): “Estimating the Effects of Free Trade Agreements on International Trade Flows using Matching Econometrics,” *Journal of International Economics*, 77(1), 63–76.
- BAILEY, N., G. KAPETANIOS, AND M. H. PESARAN (2016): “Exponent of Cross-Sectional Dependence: Estimation and Inference,” *Journal of Applied Econometrics*, 31(6), 929–960.
- BALAZSI, L., B. H. BALTAGI, L. MATYAS, AND D. PUS (2016): “Modelling Multi-dimensional Panel Data: A Random Effects Approach,” *Central European University*.
- BALAZSI, L., M. BUN, F. CHAN, AND M. N. HARRIS (2017): “Models with Endogenous Regressors,” in *The Econometrics of Multi-dimensional Panels*, ed. by L. Matyas. Springer: Berlin.
- BALAZSI, L., L. MATYAS, AND T. WANSBEEK (2018): “The Estimation of Multi-dimensional Fixed Effects Panel Data Models,” *Econometric Reviews*, 37(3), 212–227.
- BALDWIN, R., AND D. TAGLIONI (2006): “Gravity for Dummies and Dummies for Gravity Equations,” NBER Working Papers 12516, National Bureau of Economic Research, Inc.
- BALTAGI, B. H., P. EGGER, AND M. PFAFFERMAYR (2003): “A Generalized Design for Bilateral Trade Flow Models,” *Economics Letters*, 80(3), 391–397.
- (2015): “Panel Data Gravity Models of International Trade,” in *The Oxford Handbook of Panel Data*, ed. by B. H. Baltagi, pp. 608–641. Oxford University Press: New York.
- BECK, G. W., K. HUBRICH, AND M. MARCELLINO (2016): “On the Importance of Sectoral and Regional Shocks for Price Setting,” *Journal of Applied Econometrics*, 31(7), 1234–1253.
- BEHRENS, K., C. ERTUR, AND W. KOCK (2012): “Dual Gravity: Using Spatial Econometrics to Control For Multilateral Resistance,” *Journal of Applied Econometrics*, 27, 773–794.

- BREITUNG, J., AND S. EICKMEIER (2016): “Analyzing International Business and Financial Cycles using Multi-Level Factor Models: A Comparison of Alternative Approaches,” in *Dynamic Factor Models*, chap. 5, pp. 177–214. Emerald Group Publishing Limited.
- CAFISO, G. (2011): “Rose effect versus border effect: the Euro’s impact on trade,” *Applied Economics*, 43(13), 1691–1702.
- CHOI, I., D. KIM, Y. KIM, AND N. KWARK (2016): “A multilevel factor model: identification, asymptotic theory and applications,” *Journal of Applied Econometrics*, 33(3), 355–377.
- DE BOYRIE, M., AND M. KREININ (2012): “Intra-Industry Trade Revisited: A Note,” *Open Economies Review*, 23(4), 741–745.
- DE NARDIS, S., AND C. VICARELLI (2003): “Currency Unions and Trade: The Special Case of EMU,” *Review of World Economics*, 139(4), 625–649.
- DISDIER, A., AND K. HEAD (2008): “The Puzzling Persistence of the Distance Effect on Bilateral Trade,” *The Review of Economics and Statistics*, 90(1), 37–48.
- FEENSTRA, R. C. (2004): *Advanced International Trade*. Princeton University Press, Princeton, NJ.
- FRANKEL, J. (2005): “Real Convergence and Euro Adoption in Central and Eastern Europe: Trade and Business Cycle Correlations as Endogenous Criteria for Joining EMU,” in *Euro Adoption in Central and Eastern Europe: Opportunities and Challenges*, ed. by S. Schadler. International Monetary Fund.
- (2008): “The Estimated Effects of the Euro on Trade: Why are They below Historical Effects of Monetary Unions among Smaller Countries?,” NBER Working Papers 14542, National Bureau of Economic Research, Inc.
- GUNNELLA, V., C. MASTROMARCO, L. SERLENGA, AND Y. SHIN (2015): “The Euro Effects on Intra-EU Trade Flows and Balances: Evidence from the Cross Sectionally Correlated Panel Gravity Models,” Discussion paper, University of York.
- HAUSMAN, J. A., AND W. E. TAYLOR (1981): “Panel Data and Unobservable Individual Effects,” *Econometrica*, 49(6), 1377–1398.
- KAPETANIOS, G., C. MASTROMARCO, L. SERLENGA, AND Y. SHIN (2017): “Modelling in the Presence of Cross-sectional Error Dependence,” in *The Econometrics of Multi-dimensional Panels*, ed. by L. Matyas, pp. 291–322. Springer: Berlin.

- KOSE, A., C. OTROK, AND C. WHITEMAN (2003): “International Business Cycles: World, Region, and Country-Specific Factors,” *American Economic Review*, 93(4), 1216–1239.
- LESAGE, J. P., AND C. LLANO (2016): “A Spatial Interaction Model with Spatially Structured Origin and Destination Effects,” in *Spatial Econometric Interaction Modelling*, ed. by G. Arbia, and R. Patuelli, pp. 171–198. Springer: Berlin.
- MARKUSEN, J., AND K. MASKUS (2001): “A Unified Approach to Intra-Industry Trade and Direct Foreign Investment,” NBER Working Papers 8335, National Bureau of Economic Research, Inc.
- MASTROMARCO, C., L. SERLENGA, AND Y. SHIN (2016): “Multilateral Resistance and Euro Effects on Trade Flows,” in *Spatial Econometric Interaction Modelling*, ed. by G. Arbia, and R. Patuelli, pp. 253–278. Springer: Berlin.
- MATYAS, L. (2017): *The Econometrics of Multi-dimensional Panels: Theory and Applications*, Advanced Studies in Theoretical and Applied Econometrics. Springer: Berlin.
- MOENCH, E., S. NG, AND S. POTTER (2013): “Dynamic Hierarchical Factor Models,” *The Review of Economics and Statistics*, 95(5), 1811–1817.
- PESARAN, H. M., T. SCHUERMAN, AND S. M. WEINER (2004): “Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model,” *Journal of Business and Economic Statistics*, 22(2), 129–162.
- PESARAN, M. H. (2006): “Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure,” *Econometrica*, 74(4), 967–1012.
- (2015): “Testing Weak Cross-Sectional Dependence in Large Panels,” *Econometric Reviews*, 34(6), 1089–1117.
- PESARAN, M. H., AND R. SMITH (1995): “Estimating Long-run Relationships from Dynamic Heterogeneous Panels,” *Journal of Econometrics*, 68(1), 79–113.
- ROSE, A. K. (2001): “Currency Unions and Trade: the Effect is Large,” *Economic Policy*, 16(33), 449–461.
- SANTOS SILVA, J. M. C., AND S. TENREYRO (2010): “Currency Unions in Prospect and Retrospect,” *Annual Review of Economics*, 2(1), 51–74.
- SERLENGA, L., AND Y. SHIN (2007): “Gravity Models of Intra-EU Trade: Application of the CCEP-HT Estimation in Heterogeneous Panels with Unobserved Common Time-Specific Factors,” *Journal of Applied Econometrics*, 22(2), 361–381.

SHI, W., AND L. LEE (2017): “Spatial Dynamic Panel Data Models with Interactive Fixed Effects,” *Journal of Econometrics*, 197(2), 323–347.