DEPARTMENT OF ECONOMICS
WORKING PAPERS

economics.ceu.hu

The Formulation and Estimation of Random Effects Panel Data Models of Trade

by

Laszlo Matyas[1], Cecilia Hornok[2] and Daria Pus[3]

2012/1

[1] Department of Economics, Central European University, matyas@ceu.hu
[2] Department of Economics, Central European University
[3] Department of Economics, Central European University

Abstract

The paper introduces for the most frequently used three-dimensional panel data sets several random effects model specifications. It derives appropriate estimation methods for the balanced and unbalanced cases and deals with some extensions as well. An application is also presented where the bilateral trade of 20 EU countries is analysed for the period 2001-2006. The differences between the fixed and random effects specifications are highlighted through this empirical exercise.

Key words: panel data, multidimensional panel data, random effects, error components model, trade model, gravity model.

JEL classification: C1, C2, C4, F17, F47.

Acknowledgements

Support by Australian Research Council grant DP110103824 is kindly acknowledged, as well as the hospitality of Monash University for Laszlo Matyas and LaTrobe University for Cecilia Hornok.

1. Introduction

The use of multidimensional panel data sets has received momentum the last few years. Especially, three dimensional data bases are becoming readily available and frequently used to analyze different types of economic flows, like capital flows (FDI), or most predominantly, trade relationships (for a recent reviews of the subject see *Anderson* [2010] or *van Bergeijk and Brakman* [2010]). Several model specifications have been proposed in the literature to deal with the heterogeneity of these types of data sets, but all of them treated these heterogeneity factors as fixed effects, i.e., fixed unknown parameters. As it is pretty well understood from the use of "usual" two dimensional panel data sets, the fixed effects formulations are more suited to deal with cases when the panel, at least in one dimension, is short. On the other hand, for large data sets, the random effects specifications seems to be more suited, where the specific effects are considered as random variables, rather than parameters.

In this paper we introduce different types of random effects model specifications which mirror the fixed effects models used so far in the literature to deal with three-dimensional panel data sets (some earlier versions were introduced in *Davis* [2002], and historically the origins can be traced back to *Rao and Kleffe* [1980]), derive proper estimation methods for each of them and analyze their properties under some data problems. Finally, we present a revealing application.

2. Different Heterogeneity Formulations

The most widely used fixed effects model specifications have been proposed by *Baltagi et al.* [2003], *Egger and Pfanffermayr* [2003], *Baldwin and Taglioni* [2006], and *Baier and Bergstrand* [2007]. A simple straightforward direct generalization of the standard fixed effects panel data model (where the usual individuals are in fact the (ij) country pairs in the case of trade data) takes into account bilateral interaction. The model specification is

$$y_{ijt} = \beta' x_{ijt} + \gamma_{ij} + \varepsilon_{ijt} \qquad i = 1, \ldots, N \quad j = 1, \ldots, N, \quad t = 1, \ldots, T$$

where the γ_{ij} are the bilateral specific fixed effects. If the specification is used in a macro trade model, for example, with say 150 countries involved, this explicitly or implicitly, means the estimation of $150 \times 150 = 22,500$ parameters. This looks very much like a textbook over-specification case. Instead we propose, like in a standard panel data context, the use of the much more parsimonious random effects specification

$$y_{ijt} = \beta' x_{ijt} + \mu_{ij} + \varepsilon_{ijt} \qquad i = 1, \ldots, N, \quad j = 1, \ldots, N, \quad t = 1, \ldots, T \qquad (1)$$

where $E(\mu_{ij}) = 0$, the random effects are assumed to be pairwise uncorrelated, and

$$E(\mu_{ij}\mu_{i'j'}) = \begin{cases} \sigma_\mu^2 & i = i' \text{ and } j = j' \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

A natural extension of this model is to include time effects as well

$$y_{ijt} = \beta' x_{ijt} + \mu_{ij} + \lambda_t + \varepsilon_{ijt} \qquad i = 1, \ldots, N \quad j = 1, \ldots, N, \quad t = 1, \ldots, T \tag{3}$$

where $E(\lambda_t) = 0$ and

$$E(\lambda_t \lambda_t') = \begin{cases} \sigma_\lambda^2 & t = t' \\ 0 & \text{otherwise} \end{cases}$$

Another form of heterogeneity is to use individual-time-varying effects. The corresponding specification now is

$$y_{ijt} = \beta' x_{ijt} + u_{jt} + \varepsilon_{ijt} \tag{4}$$

where $E(u_{jt}) = 0$, the random effects are pairwise uncorrelated, and

$$E(u_{ij}u_{j't'}) = \begin{cases} \sigma_u^2 & j = j' \text{ and } t = t' \\ 0 & \text{otherwise} \end{cases}$$

Or alternatively we can also have the following specification

$$y_{ijt} = \beta' x_{ijt} + v_{it} + \varepsilon_{ijt} \tag{5}$$

where $E(v_{it}) = 0$, the random effects are again assumed to be pairwise uncorrelated, and

$$E(v_{it}v_{i't'}) = \begin{cases} \sigma_v^2 & i = i' \text{ and } t = t' \\ 0 & \text{otherwise} \end{cases}$$

The specification containing both the above forms of heterogeneity now is

$$y_{ijt} = \beta' x_{ijt} + v_{it} + u_{jt} + \varepsilon_{ijt} \tag{6}$$

Finally, the model specification which encompasses all above effects is

$$y_{ijt} = \beta' x_{ijt} + \mu_{ij} + v_{it} + u_{jt} + \varepsilon_{ijt} \tag{7}$$

where $E(\mu_{ij}) = 0$, $E(u_{jt}) = 0$, $E(v_{it}) = 0$, all random effects are pairwise uncorrelated, and

$$E(\mu_{ij}\mu_{i'j'}) = \begin{cases} \sigma_\mu^2 & i = i' \text{ and } j = j' \\ 0 & \text{otherwise} \end{cases}$$
$$E(u_{jt}u_{j't'}) = \begin{cases} \sigma_u^2 & j = j' \text{ and } t = t' \\ 0 & \text{otherwise} \end{cases}$$
$$E(v_{it}v_{i't'}) = \begin{cases} \sigma_v^2 & i = i' \text{ and } t = t' \\ 0 & \text{otherwise} \end{cases}$$

These random effects models, just like in the case of the "usual" panel data models, can be estimated (asymptotically) efficiently with (F)GLS (see, for example, *Baltagi, Matyas and Sevestre* [2008]). In order to do so their corresponding covariance matrices need to be derived, and then, also, the respective variance components need to be estimated.

3. Covariance Matrices of the Different Random Effects Specifications

The standard way to estimate these models is with the Feasible GLS (FGLS) estimator. First, we derive the covariance matrix of each of the models introduced in Section 2, then the unknown variance components of these matrices are estimated.

For model (1) let us denote

$$u_{ijt}^\star = \mu_{ij} + \epsilon_{ijt} \tag{8}$$

So for all t observations

$$\begin{aligned} u_{ij}^\star &= \mu_{ij} \otimes l_T + \epsilon_{ij} \\ E\left[u_{ij}^\star {u^\star}'_{ij}\right] &= E\left[(\mu_{ij} \otimes l_T)(\mu_{ij} \otimes l_T')\right] + E\left[\epsilon_{ij}\epsilon_{ij}'\right] \\ &= \sigma_\mu^2 J_T + \sigma_\epsilon^2 I_T \end{aligned}$$

where l_T is the $(T \times 1)$ vector of ones, J_T is the $(T \times T)$ matrix of ones and I_T is the $(T \times T)$ identity matrix. In all the paper matrix J will denote the matrix of ones, with the size in the index, and I the identity matrix, also with the size in the index. Now for individual i

$$\begin{aligned} u_i^\star &= \mu_i \otimes l_T + \epsilon_i \\ E\left[u_i^\star {u^\star}'_i\right] &= E\left[(\mu_i \otimes l_T)(\mu_i' \otimes l_T')\right] + E\left[\epsilon_i\epsilon_i'\right] \\ &= \sigma_\mu^2 I_N \otimes J_T + \sigma_\epsilon^2 I_{NT} \end{aligned}$$

And combining all these results we get for the covariance matrix of model (1)

$$\begin{aligned}
u^{\star} &= \mu \otimes l_T + \epsilon \\
E\left[u^{\star} u^{\star\prime}\right] &= E\left[\left(\mu \otimes l_T\right)\left(\mu' \otimes l_T'\right)\right] + E\left[\epsilon\epsilon'\right] \\
&= \sigma_\mu^2 I_{N^2} \otimes J_T + \sigma_\epsilon^2 I_{N^2T} = \Omega
\end{aligned}$$

Using matrix notation and decomposing the covariance matrix Ω

$$\begin{aligned}
\Omega &= \sigma_\mu^2 I_{N^2} \otimes J_T + \sigma_\epsilon^2 I_{N^2T} \\
&= T\sigma_\mu^2 \left(B_{ij} + \frac{J_{N^2T}}{N^2T}\right) + \sigma_\epsilon^2 \left(W_1 + B_{ij} + \frac{J_{N^2T}}{N^2T}\right) \\
&= \left(T\sigma_\mu^2 + \sigma_\epsilon^2\right) \frac{J_{N^2T}}{N^2T} + \left(T\sigma_\mu^2 + \sigma_\epsilon^2\right) B_{ij} + \sigma_\epsilon^2 W_1
\end{aligned}$$

where

$$W_1 = I_{N^2T} - \left(I_{N^2} \otimes \frac{J_T}{T}\right), \qquad \text{rank}: \ N^2(T-1)$$

and

$$B_{ij} = \left(I_{N^2} \otimes \frac{J_T}{T}\right) - \frac{J_{N^2T}}{N^2T}, \qquad \text{rank}: \ (N^2-1)$$

Like in the usual panel data case, we can use this (through the eigenvalue – eigenvector decomposition) to derive the inverse of Ω (and so saving us to invert a massive matrix in order to use the GLS estimator)

$$\Omega^{-1} = \frac{1}{\left(T\sigma_\mu^2 + \sigma_\epsilon^2\right)} \frac{J_{N^2T}}{N^2T} + \frac{1}{\left(T\sigma_\mu^2 + \sigma_\epsilon^2\right)} B_{ij} + \frac{1}{\sigma_\epsilon^2} W_1$$

This leads us to the GLS estimator

$$\widehat{\beta}_{GLS} = \left[X'(\theta \frac{J_{N^2T}}{N^2T} + \theta B_{ij} + W_1) X\right]^{-1} X' \left(\theta \frac{J_{N^2T}}{N^2T} + \theta B_{ij} + W_1\right) y$$

where

$$\theta = \frac{\sigma_\epsilon^2}{T\sigma_\mu^2 + \sigma_\epsilon^2}$$

After substituting W_1 and B_{ij} in, this becomes

$$\widehat{\beta}_{GLS} = \left[X' \left(I_{N^2T} - (1-\theta) I_{N^2} \otimes \frac{J_T}{T}\right) X\right]^{-1} X' \left(I_{N^2T} - (1-\theta) I_{N^2} \otimes \frac{J_T}{T}\right) y$$

This formula shows that the FGLS estimator is in fact an OLS estimator on the $\tilde{y}_{ijt} = (y_{ijt} - (1-\theta)\sum_t \frac{1}{T} y_{ijt})$ type transformed variables.

Next, deriving likewise the covariance matrix for model (3)

$$\begin{aligned}
u^{\star}_{ij} &= \mu_{ij} \otimes l_T + \lambda + \epsilon_{ij} \\
E\left[u^{\star}_{ij} {u^{\star}}'_{ij}\right] &= E\left[(\mu_{ij} \otimes l_T)(\mu_{ij} \otimes l_T)'\right] + E\left[\lambda\lambda'\right] + E\left[\epsilon_{ij}\epsilon'_{ij}\right] \\
&= \sigma^2_{\mu} J_T + \sigma^2_{\lambda} I_T + \sigma^2_{\epsilon} I_T
\end{aligned}$$

and

$$\begin{aligned}
u^{\star}_{i} &= \mu_{i} \otimes l_T + l_N \otimes \lambda + \epsilon_{i} \\
E\left[u^{\star}_{i} {u^{\star}}'_{i}\right] &= E\left[(\mu_{i} \otimes l_T)(\mu_{i} \otimes l_T)'\right] + E\left[(l_N \otimes \lambda)(l_N \otimes \lambda)'\right] + E\left[\epsilon_{i}\epsilon'_{i}\right] \\
&= \sigma^2_{\mu} I_N \otimes J_T + \sigma^2_{\lambda} J_N \otimes I_T + \sigma^2_{\epsilon} I_{NT}
\end{aligned}$$

so we obtain

$$\begin{aligned}
u^{\star} &= \mu \otimes l_T + l_{N^2} \otimes \lambda + \epsilon \\
E\left[u^{\star} {u^{\star}}'\right] &= E\left[(\mu \otimes l_T)(\mu \otimes l_T)'\right] + E\left[(l_{N^2} \otimes \lambda)(l_{N^2} \otimes \lambda)'\right] + E\left[\epsilon\epsilon'\right] \\
&= \sigma^2_{\mu} I_{N^2} \otimes J_T + \sigma^2_{\lambda} J_{N^2} \otimes I_T + \sigma^2_{\epsilon} I_{N^2T} = \Omega
\end{aligned}$$

Now turning to the inverse of the covariance matrix

$$\begin{aligned}
\Omega &= \sigma^2_{\mu} I_{N^2} \otimes J_T + \sigma^2_{\lambda} J_{N^2} \otimes I_T + \sigma^2_{\epsilon} I_{N^2T} \\
&= T\sigma^2_{\mu}\left(B_{ij} + \frac{J_{N^2T}}{N^2T}\right) + N^2\sigma^2_{\lambda}\left(B_t + \frac{J_{N^2T}}{N^2T}\right) + \sigma^2_{\epsilon}\left(W_2 + B_{ij} + B_t + \frac{J_{N^2T}}{N^2T}\right) \\
&= \left(T\sigma^2_{\mu} + N^2\sigma^2_{\lambda} + \sigma^2_{\epsilon}\right)\frac{J_{N^2T}}{N^2T} + \left(T\sigma^2_{\mu} + \sigma^2_{\epsilon}\right)B_{ij} + \left(N^2\sigma^2_{\lambda} + \sigma^2_{\epsilon}\right)B_t + \sigma^2_{\epsilon} W_2
\end{aligned}$$

where

$$B_t = \left(\frac{J_{N^2}}{N^2} \otimes I_T\right) - \frac{J_{N^2T}}{N^2T}, \qquad \text{rank} : (T-1)$$

and

$$W_2 = I_{N^2T} - \left(I_{N^2} \otimes \frac{J_T}{T}\right) - \left(\frac{J_{N^2}}{N^2} \otimes I_T\right) + \frac{J_{N^2T}}{N^2T}, \qquad \text{rank} : (N^2-1)(T-1)$$

and the inverse of Ω now is

$$\sigma^2_{\epsilon}\Omega^{-1} = \left(\theta_1 \frac{J_{N^2T}}{N^2T} + \theta_2 B_{ij} + \theta_3 B_t + W_2\right)$$

with

$$\theta_1 = \frac{\sigma_\epsilon^2}{T\sigma_\mu^2 + N^2\sigma_\lambda^2 + \sigma_\epsilon^2}$$
$$\theta_2 = \frac{\sigma_\epsilon^2}{T\sigma_\mu^2 + \sigma_\epsilon^2}$$
$$\theta_3 = \frac{\sigma_\epsilon^2}{N^2\sigma_\lambda^2 + \sigma_\epsilon^2}$$

Let us turn now to models (4) and (5) which can be dealt with in a similar way as they are completely symmetric

$$u_{ijt}^\star = u_{jt} + \epsilon_{ijt} \tag{9}$$

$$\begin{aligned} u_{ij}^\star &= u_j + \epsilon_{ij} \\ E\left(u_{ij}^\star {u^\star}'_{ij}\right) &= E\left[u_j u_j'\right] + E\left[\epsilon_{ij}\epsilon_{ij}'\right] = \sigma_u^2 I_T + \sigma_\epsilon^2 I_T \end{aligned}$$

$$\begin{aligned} u_i^\star &= u + \epsilon_i \\ E\left(u_i^\star {u^\star}'_i\right) &= E\left[uu'\right] + E\left[\epsilon_i\epsilon_i'\right] = \sigma_u^2 I_{NT} + \sigma_\epsilon^2 I_{NT} \\ u^\star &= l_N \otimes u + \epsilon \\ E\left(u^\star {u^\star}'\right) &= E\left[\left(l_N \otimes u\right)\left(l_N' \otimes u'\right)\right] + E\left[\epsilon\epsilon'\right] = \sigma_u^2 J_N \otimes I_{NT} + \sigma_\epsilon^2 I_{N^2T} = \Omega \end{aligned}$$

The inverse of the covariance matrix in this case is for model (4)

$$\begin{aligned} \Omega &= \sigma_u^2 J_N \otimes I_{NT} + \sigma_\epsilon^2 I_{N^2T} \\ &= N\sigma_u^2\left(B_{jt} + \frac{J_{N^2T}}{N^2T}\right) + \sigma_\epsilon^2\left(W_3 + B_{jt} + \frac{J_{N^2T}}{N^2T}\right) \\ &= \left(N\sigma_u^2 + \sigma_\epsilon^2\right)\frac{J_{N^2T}}{N^2T} + \left(N\sigma_u^2 + \sigma_\epsilon^2\right)B_{jt} + \sigma_\epsilon^2 W_3 \end{aligned}$$

where

$$W_3 = I_{N^2T} - \left(\frac{J_N}{N} \otimes I_{NT}\right), \qquad \text{rank} : (NT(T-1))$$

$$B_{jt} = \left(\frac{J_N}{N} \otimes I_{NT}\right) - \frac{J_{N^2T}}{N^2T}, \quad \text{rank} : (NT-1)$$

and

$$\sigma_\epsilon^2 \Omega^{-1} = \theta\frac{J_{N^2T}}{N^2T} + \theta B_{jt} + W_3 \qquad \text{with} \quad \theta = \frac{\sigma_\epsilon^2}{N\sigma_u^2 + \sigma_\epsilon^2}$$

For model (5) B_{jt} should be substituted by B_{it} and W_3 by W_4 in the above formulas, where

$$B_{it} = \left(I_N \otimes \frac{J_N}{N} \otimes I_T\right) - \frac{J_{N^2T}}{N^2T}, \quad \text{rank} : (NT-1)$$

and

$$W_4 = I_{N^2T} - \left(I_N \otimes \frac{J_N}{N} \otimes I_T\right)$$

Using the same approach, the covariance matrix for model (6) is

$$u^{\star}_{ijt} = u_{jt} + v_{it} + \epsilon_{ijt}$$

$$u^{\star}_{ij} = u_j + v_i + \epsilon_{ij}$$

$$\begin{aligned} E\left(u^{\star}_{ij} {u^{\star}}'_{ij}\right) &= E\left[u_j u'_j\right] + E\left[v_i v'_i\right] + E\left[\epsilon_{ij}\epsilon'_{ij}\right] \\ &= \sigma^2_u I_T + \sigma^2_v I_T + \sigma^2_\epsilon I_T \end{aligned}$$

$$\begin{aligned} u^{\star}_i &= l_N \otimes v_i + u + \epsilon_i \\ E\left(u^{\star}_i {u^{\star}}'_i\right) &= E\left[(l_N \otimes v_i)\left(l'_N \otimes v'_i\right)\right] + E\left[uu'\right] + E\left[\epsilon_i \epsilon'_i\right] = \\ &= \sigma^2_v J_N \otimes I_T + \sigma^2_u I_{NT} + \sigma^2_\epsilon I_{NT} \end{aligned}$$

and so

$$E\left(u^{\star}{u^{\star}}'\right) = \sigma^2_v (I_N \otimes J_N \otimes I_T) + \sigma^2_u (J_N \otimes I_{NT}) + \sigma^2_\epsilon I_{N^2T} = \Omega$$

Turning now to the inverse of Ω

$$\Omega = N\sigma^2_v \left(B^o_{it} + \frac{J_{N^2}}{N^2} \otimes I_T\right) + N\sigma^2_u \left(B^o_{jt} + \frac{J_{N^2}}{N^2} \otimes I_T\right) + \sigma^2_\varepsilon \left(B^o_{it} + B^o_{jt} + \frac{J_{N^2}}{N^2} \otimes I_T + W_5\right)$$

where

$$B^o_{it} = \left(I_N \otimes \frac{J_N}{N} \otimes I_T\right) - \left(\frac{J_{N^2}}{N^2} \otimes I_T\right), \quad \text{rank} : (N-1)T$$

$$B^o_{jt} = \left(\frac{J_N}{N} \otimes I_{NT}\right) - \left(\frac{J_{N^2}}{N^2} \otimes I_T\right), \quad \text{rank} : (N-1)T$$

and

$$W_5 = I_{N^2T} - \left(\frac{J_N}{N} \otimes I_{NT}\right) - \left(I_N \otimes \frac{J_N}{N} \otimes I_T\right) + \left(\frac{J_{N^2}}{N^2} \otimes I_T\right), \quad \text{rank} : (N-1)^2 T$$

So the inverse is

$$\sigma_\varepsilon^2\Omega^{-1} = \theta_1(\frac{J_{N^2}}{N^2} \otimes I_T) + \theta_2 B^o_{it} + \theta_3 B^o_{jt} + W_5$$

with

$$\begin{aligned}
\theta_1 &= \frac{\sigma_\epsilon^2}{N\sigma_v^2 + N\sigma_u^2 + \sigma_\epsilon^2} \\
\theta_2 &= \frac{\sigma_\epsilon^2}{N\sigma_v^2 + \sigma_\epsilon^2} \\
\theta_3 &= \frac{\sigma_\epsilon^2}{N\sigma_u^2 + \sigma_\epsilon^2}
\end{aligned}$$

This is equivalent to

$$\begin{aligned}
\sigma_\varepsilon^2\Omega^{-1} = I_{N^2T} - (1-\theta_2)\left(I_N \otimes \frac{J_N}{N} \otimes I_T\right) - (1-\theta_3)\left(\frac{J_N}{N} \otimes I_{NT}\right) + \\
+ (1-\theta_2-\theta_3+\theta_1)\left(\frac{J_{N^2}}{N^2} \otimes I_T\right)
\end{aligned}$$

From this last expression we can see that the FGLS estimator is equivalent to the OLS on the transformed variables like

$$\tilde{y}_{ijt} = \left(y_{ijt} - (1-\theta_2)\sum_j \frac{1}{N} y_{ijt} - (1-\theta_3)\sum_i \frac{1}{N} y_{ijt} + (1-\theta_2-\theta_3+\theta_1)\sum_i \sum_j \frac{1}{N^2} y_{ijt}\right)$$

And finally the covariance matrix of the all encompassing model (7) is

$$u^\star_{ijt} = \mu_{ij} + u_{jt} + v_{it} + \epsilon_{ijt} \tag{10}$$

$$\begin{aligned}
u^\star_{ij} &= \mu_{ij} \otimes l_T + u_j + v_i + \epsilon_{ij} \\
E\left(u^\star_{ij} {u^\star}'_{ij}\right) &= E\left[(\mu_{ij} \otimes l_T)(\mu_{ij} \otimes l'_T)\right] + E\left[u_j u'_j\right] + E\left[v_i v'_i\right] + E\left[\epsilon_{ij}\epsilon'_{ij}\right] \\
&= \sigma_\mu^2 J_T + \sigma_u^2 I_T + \sigma_v^2 I_T + \sigma_\epsilon^2 I_T
\end{aligned}$$

$$\begin{aligned}
u^\star_i &= \mu_i \otimes l_T + l_N \otimes v_i + u + \epsilon_i \\
E\left(u^\star_i {u^\star}'_i\right) &= E\left[(\mu_i \otimes l_T)(\mu'_i \otimes l'_T)\right] + E\left[(l_N \otimes v_i)(l'_N \otimes v'_i)\right] + E\left[uu'\right] + E\left[\epsilon_i\epsilon'_i\right] = \\
&= \sigma_\mu^2 I_N \otimes J_T + \sigma_u^2 I_{NT} + \sigma_v^2 J_N \otimes I_T + \sigma_\epsilon^2 I_{NT}
\end{aligned}$$

and so

$$E\left(u^\star {u^\star}'\right) = \sigma_\mu^2(I_{N^2} \otimes J_T) + \sigma_u^2(J_N \otimes I_{NT}) + \sigma_v^2(I_N \otimes J_N \otimes I_T) + \sigma_\epsilon^2 I_{N^2T} = \Omega$$

Similarly as for the previous models, the decomposition of Ω can de carried out like

$$\Omega = T\sigma_\mu^2 \left(B_{ij}^* + B + C + \frac{J_{N^2T}}{N^2T} \right) + N\sigma_u^2 \left(B_{jt}^* + A + C + \frac{J_{N^2T}}{N^2T} \right) +$$
$$+ N\sigma_v^2 \left(B_{it}^* + A + B + \frac{J_{N^2T}}{N^2T} \right) + \sigma_\varepsilon^2 \left(B_{ij}^* + B_{jt}^* + B_{it}^* + A + B + C + \frac{J_{N^2T}}{N^2T} + W_6 \right)$$

where

$$B_{ij}^* = \left(I_{N^2} \otimes \frac{J_T}{T} \right) - B - C - \frac{J_{N^2T}}{N^2T}, \quad \text{rank :}((N-1)^2 + 1)$$

$$B_{jt}^* = \left(\frac{J_N}{N} \otimes I_{NT} \right) - A - C - \frac{J_{N^2T}}{N^2T}, \quad \text{rank :}((N-1)(T-1) + 1)$$

$$B_{it}^* = \left(I_N \otimes \frac{J_N}{N} \otimes I_T \right) - A - B - \frac{J_{N^2T}}{N^2T}, \quad \text{rank :}((N-1)(T-1) + 1)$$

$$W_6 = I_{N^2T} - \left(I_{N^2} \otimes \frac{J_T}{T} \right) - \left(\frac{J_N}{N} \otimes I_{NT} \right) - \left(I_N \otimes \frac{J_N}{N} \otimes I_T \right) + A + B + C + 2\frac{J_{N^2T}}{N^2T}$$
$$\text{rank : ((N} - 1)^2(\text{T} - 1) + 1)$$

$$A = \left(\frac{J_{N^2}}{N^2} \otimes I_T \right), \quad \text{rank : T}$$

$$B = \left(I_N \otimes \frac{J_{NT}}{NT} \right), \quad \text{rank : N}$$

and

$$C = \left(\frac{J_N}{N} \otimes I_N \otimes \frac{J_T}{T} \right), \quad \text{rank : N}$$

So the inverse is

$$\sigma_\varepsilon^2 \Omega^{-1} = \theta_1 B_{ij}^* + \theta_2 B_{jt}^* + \theta_3 B_{it}^* + \theta_4 A + \theta_5 B + \theta_6 C + \theta_7 \frac{J_{N^2T}}{N^2T} + W_6$$

where

$$\theta_1 = \frac{\sigma_\varepsilon^2}{T\sigma_\mu^2 + \sigma_\varepsilon^2}$$

$$\theta_2 = \frac{\sigma_\varepsilon^2}{N\sigma_u^2 + \sigma_\varepsilon^2}$$

$$\theta_3 = \frac{\sigma_\varepsilon^2}{N\sigma_v^2 + \sigma_\varepsilon^2}$$

$$\theta_4 = \frac{\sigma_\varepsilon^2}{N\sigma_u^2 + N\sigma_v^2 + \sigma_\varepsilon^2}$$

$$\theta_5 = \frac{\sigma_\varepsilon^2}{T\sigma_\mu^2 + N\sigma_v^2 + \sigma_\varepsilon^2}$$

$$\theta_6 = \frac{\sigma_\varepsilon^2}{T\sigma_\mu^2 + N\sigma_u^2 + \sigma_\varepsilon^2}$$

and

$$\theta_7 = \frac{\sigma_\varepsilon^2}{T\sigma_\mu^2 + N\sigma_u^2 + N\sigma_v^2 + \sigma_\varepsilon^2}$$

This expression for the inverse is equivalent to

$$\begin{aligned}
\sigma_\varepsilon^2\Omega^{-1} = I_{N^2T} - (1-\theta_1)\left(I_{N^2} \otimes \frac{J_T}{T}\right) - (1-\theta_2)\left(\frac{J_N}{N} \otimes I_{NT}\right) - \\
- (1-\theta_3)\left(I_N \otimes \frac{J_N}{N} \otimes I_T\right) + (1-\theta_2-\theta_3+\theta_4)\left(\frac{J_{N^2}}{N^2} \otimes I_T\right) + \\
+ (1-\theta_1-\theta_3+\theta_5)\left(I_N \otimes \frac{J_{NT}}{NT}\right) + (1-\theta_1-\theta_2+\theta_6)\left(\frac{J_N}{N} \otimes I_N \otimes \frac{J_T}{T}\right) - \\
- (\theta_1+\theta_2+\theta_3-\theta_7-2)\frac{J_{N^2T}}{N^2T}
\end{aligned}$$

This implies that the FGLS estimator is equivalent to the OLS on the transformed variables like

$$\begin{aligned}
\tilde{y}_{ijt} = (y_{ijt} - (1-\theta_1)\sum_t \frac{1}{T}y_{ijt} - (1-\theta_2)\sum_i \frac{1}{N}y_{ijt} - (1-\theta_3)\sum_j \frac{1}{N}y_{ijt} + \\
+ (1-\theta_2-\theta_3+\theta_4)\sum_i\sum_j \frac{1}{N^2}y_{ijt} + (1-\theta_1-\theta_3+\theta_5)\sum_j\sum_t \frac{1}{NT}y_{ijt} + \\
+ (1-\theta_1-\theta_2+\theta_6)\sum_i\sum_t \frac{1}{NT}y_{ijt} - (\theta_1+\theta_2+\theta_3-\theta_7-2)\sum_i\sum_j\sum_t \frac{1}{N^2T}y_{ijt})
\end{aligned}$$

4. Estimation of the Variance Components and the Feasible GLS Estimator

Turning now to the estimation of the variance components of the different models, let us start with model (1)

$$E\left[{u^\star}_{ijt}^2\right] = E\left[(\mu_{ij} + \epsilon_{ijt})^2\right] = E\left[\mu_{ij}^2\right] + E\left[\epsilon_{ijt}^2\right] = \sigma_\mu^2 + \sigma_\epsilon^2 \tag{11}$$

and let us introduce the appropriate Within transformation

$$u^\star_{ijt,within} = u^\star_{ijt} - \bar{u^\star}_{ij} = \epsilon_{ijt} - \bar{\epsilon}_{ij} \tag{12}$$

where $\bar{\epsilon}_{ij} = 1/T \sum_t \epsilon_{ijt}$ and $\bar{u^*}_{ij} = 1/T \sum_t u^*_{ijt}$, so we get

$$E\left[\left(u^\star_{ijt} - \bar{u^\star}_{ij}\right)^2\right] = E\left[(\epsilon_{ijt} - \bar{\epsilon}_{ij})^2\right] = E\left[\epsilon^2_{ijt} - 2\epsilon_{ijt}\frac{1}{T}\sum_{t=1}^{T}\epsilon_{ijt} + \left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_{ijt}\right)^2\right]$$

$$= E\left[\epsilon^2_{ijt}\right] - 2E\left[\epsilon_{ijt}\frac{1}{T}\sum_{t=1}^{T}\epsilon_{ijt}\right] + E\left[\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_{ijt}\right)^2\right]$$

$$= \sigma^2_\epsilon - \frac{2}{T}\sigma^2_\epsilon + \frac{1}{T}\sigma^2_\epsilon = \sigma^2_\epsilon - \frac{1}{T}\sigma^2_\epsilon = \sigma^2_\epsilon\frac{T-1}{T}$$

Let $\hat{u}^*$ be the OLS residual of model (1) and $\hat{u}^*_{within}$ the Within transformation of this residual. Then we can estimate the variance components as

$$\hat{\sigma}^2_\epsilon = \frac{T}{T-1}\hat{u}^{\star'}_{within}\hat{u}^\star_{within}$$

$$\hat{\sigma}^2_\mu = \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\hat{u}^{\star 2}_{ijt} - \hat{\sigma}^2_\epsilon$$

These estimators naturally should be adjusted to the actual degrees of freedom.

Continuing with model (3)

$$E\left[{u^\star}^2_{ijt}\right] = E\left[(\mu_{ij} + \lambda_t + \epsilon_{ijt})^2\right] = E\left[\mu^2_{ij}\right] + E\left[\lambda^2_t\right] + E\left[\epsilon^2_{ijt}\right]$$

$$= \sigma^2_\mu + \sigma^2_\lambda + \sigma^2_\epsilon$$

$$E\left[\left(\frac{1}{T}\sum_{t=1}^{T}{u^\star}_{ijt}\right)^2\right] = E\left[\left(\frac{1}{T}\sum_{t=1}^{T}\mu_{ij} + \lambda_t + \epsilon_{ijt}\right)^2\right]$$

$$= E\left[\mu^2_{ij}\right] + \frac{1}{T^2}E\left[\sum_{t=1}^{T}\lambda^2_t\right] + \frac{1}{T^2}E\left[\sum_{t=1}^{T}\epsilon^2_{ijt}\right]$$

$$= \sigma^2_\mu + \frac{1}{T}\sigma^2_\lambda + \frac{1}{T}\sigma^2_\epsilon$$

and

$$E\left[\left(u^\star_{ijt} - \bar{u^\star}_{ij} - \bar{u^\star}_t + \bar{u^\star}\right)^2\right] = E\left[(\epsilon_{ijt} - \bar{\epsilon}_{ij} - \bar{\epsilon}_t + \bar{\epsilon})^2\right]$$

$$= E\left[\epsilon^2_{ijt}\right] + E\left[\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_{ijt}\right)^2\right] +$$

$$+ E\left[\left(\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\epsilon_{ijt}\right)^2\right] + E\left[\left(\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\epsilon_{ijt}\right)^2\right] -$$

$$
\begin{aligned}
& - 2E\left[\epsilon_{ijt}\frac{1}{T}\sum_{t=1}^{T}\epsilon_{ijt}\right] - 2E\left[\epsilon_{ijt}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\epsilon_{ijt}\right] + \\
& + 2E\left[\epsilon_{ijt}\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\epsilon_{ijt}\right] + 2E\left[\frac{1}{T}\sum_{t=1}^{T}\epsilon_{ijt}\cdot\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\epsilon_{ijt}\right] - \\
& - 2E\left[\frac{1}{T}\sum_{t=1}^{T}\epsilon_{ijt}\cdot\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\epsilon_{ijt}\right] - 2E\left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\epsilon_{ijt}\cdot\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\epsilon_{ijt}\right] \\
& = \sigma_\epsilon^2 + \frac{1}{T}\sigma_\epsilon^2 + \frac{1}{N^2}\sigma_\epsilon^2 + \frac{1}{N^2T}\sigma_\epsilon^2 - \frac{2}{T}\sigma_\epsilon^2 - \frac{2}{N^2}\sigma_\epsilon^2 + \\
& \quad + \frac{2}{N^2T}\sigma_\epsilon^2 + \frac{2}{N^2T}\sigma_\epsilon^2 - \frac{2}{N^2T}\sigma_\epsilon^2 - \frac{2}{N^2T}\sigma_\epsilon^2 = \\
& = \sigma_\epsilon^2\frac{(N-1)(N+1)(T-1)}{N^2T}
\end{aligned}
$$

This leads to the estimation of the variance components

$$
\begin{aligned}
\hat{\sigma}_\epsilon^2 &= \frac{N^2T}{(N-1)(N+1)(T-1)}\hat{u^\star}'_{within}\hat{u^\star}_{within} \\
\hat{\sigma}_\mu^2 &= \frac{1}{N^2T(T-1)}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\left(\sum_{t=1}^{T}\hat{u}_{ijt}^\star\right)^2 - \sum_{t=1}^{T}\left(\hat{u}_{ijt}^\star\right)^2\right)\right) \\
\hat{\sigma}_\lambda^2 &= \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\left(\hat{u}_{ijt}^\star\right)^2 - \hat{\sigma}_\mu^2 - \hat{\sigma}_\epsilon^2
\end{aligned}
$$

Turning now to models (4) and (5)

$$
E\left[{u^\star}_{ijt}^2\right] = E\left[\left(u_{jt}+\epsilon_{ijt}\right)^2\right] = E\left[u_{jt}^2\right] + E\left[\epsilon_{ijt}^2\right] = \sigma_u^2 + \sigma_\epsilon^2 \tag{13}
$$

and the appropriate Within transformation now is

$$
u_{ijt,within}^\star = u_{ijt}^\star - \bar{u^\star}_{jt} = \epsilon_{ijt} - \bar{\epsilon}_{jt} \tag{14}
$$

where $\bar{u}_{jt}^* = 1/N\sum_i u_{ijt}^*$ and $\bar{\epsilon}_{jt} = 1/N\sum_i \epsilon_{ijt}$ and

$$
\begin{aligned}
E\left[\left(u_{ijt}^\star - \bar{u^\star}_{jt}\right)^2\right] &= E\left[\left(\epsilon_{ijt}-\bar{\epsilon}_{jt}\right)^2\right] \\
&= E\left[\epsilon_{ijt}^2 - 2\epsilon_{ijt}\frac{1}{N}\sum_{i=1}^{N}\epsilon_{ijt} + \left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_{ijt}\right)^2\right] \\
&= E\left[\epsilon_{ijt}^2\right] - 2E\left[\epsilon_{ijt}\frac{1}{N}\sum_{i=1}^{N}\epsilon_{ijt}\right] + E\left[\left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_{ijt}\right)^2\right] \\
&= \sigma_\epsilon^2 - \frac{2}{N}\sigma_\epsilon^2 + \frac{1}{N}\sigma_\epsilon^2 = \sigma_\epsilon^2 - \frac{1}{N}\sigma_\epsilon^2 = \sigma_\epsilon^2\frac{N-1}{N}
\end{aligned}
$$

And the estimators for the variance components are

$$\hat{\sigma}_\epsilon^2 = \frac{N}{N-1}\hat{u}^{\star'}_{within}\hat{u}^{\star}_{within}$$

$$\hat{\sigma}_\mu^2 = \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\hat{u}^{\star 2}_{ijt} - \hat{\sigma}_\epsilon^2$$

Now for model (6) the Within transformation is

$$u^*_{ijt,within} = (u^*_{ijt} - 1/N\sum_i u^*_{ijt} - 1/N\sum_j u^*_{ijt} + 1/N^2\sum_i\sum_j u^*_{ijt}) \qquad (15)$$

so we get

$$\begin{aligned}
&E\left[\left(u^\star_{ijt} - \bar{u^\star}_{jt} - \bar{u^\star}_{it} + \bar{u^\star}_{t}\right)^2\right] = E\left[\left(\epsilon_{ijt} - \bar{\epsilon}_{jt} - \bar{\epsilon}_{it} + \bar{\epsilon}_t\right)^2\right] \\
&= E\left[\epsilon^2_{ijt}\right] + E\left[\frac{1}{N^2}\left(\sum_{i=1}^{N}\epsilon_{ijt}\right)^2\right] + E\left[\frac{1}{N^2}\left(\sum_{j=1}^{N}\epsilon_{ijt}\right)^2\right] + E\left[\frac{1}{N^4}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\epsilon_{ijt}\right)^2\right] - \\
&- 2E\left[\epsilon_{ijt}\frac{1}{N}\sum_{i=1}^{N}\epsilon_{ijt}\right] - 2E\left[\epsilon_{ijt}\frac{1}{N}\sum_{j=1}^{N}\epsilon_{ijt}\right] + 2E\left[\epsilon_{ijt}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\epsilon_{ijt}\right] + \\
&+ 2E\left[\frac{1}{N^2}\sum_{i=1}^{N}\epsilon_{ijt}\sum_{j=1}^{N}\epsilon_{ijt}\right] - 2E\left[\frac{1}{N^3}\sum_{i=1}^{N}\epsilon_{ijt}\sum_{i=1}^{N}\sum_{j=1}^{N}\epsilon_{ijt}\right] - 2E\left[\frac{1}{N^3}\sum_{j=1}^{N}\epsilon_{ijt}\sum_{i=1}^{N}\sum_{j=1}^{N}\epsilon_{ijt}\right] = \\
&= \sigma^2_\epsilon + \frac{1}{N}\sigma^2_\epsilon + \frac{1}{N}\sigma^2_\epsilon + \frac{1}{N^2}\sigma^2_\epsilon - \frac{2}{N}\sigma^2_\epsilon - \frac{2}{N}\sigma^2_\epsilon + \frac{2}{N^2}\sigma^2_\epsilon + \frac{2}{N^2}\sigma^2_\epsilon - \frac{2}{N^2}\sigma^2_\epsilon - \frac{2}{N^2}\sigma^2_\epsilon = \\
&= \sigma^2_\epsilon\left(1 - \frac{2}{N} + \frac{1}{N^2}\right) = \sigma^2_\epsilon\left(\frac{N^2 - 2N + 1}{N^2}\right) = \sigma^2_\epsilon\frac{(N-1)^2}{N^2}
\end{aligned} \qquad (16)$$

And, also,

$$\begin{aligned}
E\left[u^{\star 2}_{ijt}\right] &= E\left[(u_{jt} + v_{it} + \epsilon_{ijt})^2\right] = \sigma^2_u + \sigma^2_v + \sigma^2_\epsilon \\
E\left[\left(\frac{1}{N}\sum_{i=1}^{N}u^\star_{ijt}\right)^2\right] &= E\left[\left(\frac{1}{N}\sum_{i=1}^{N}(u_{jt} + v_{it} + \epsilon_{ijt})\right)^2\right] \\
&= E\left[u^2_{jt}\right] + \frac{1}{N^2}E\left[\sum_{i=1}^{N}v^2_{it}\right] + \frac{1}{N^2}E\left[\sum_{i=1}^{N}\epsilon^2_{ijt}\right] \\
&= \sigma^2_u + \frac{1}{N}\sigma^2_v + \frac{1}{N}\sigma^2_\epsilon
\end{aligned} \qquad (17)$$

The estimators of the variance components therefore are

$$\hat{\sigma}^2_\epsilon = \frac{N^2}{(N-1)^2}\hat{u}^{\star\prime}_{within}\hat{u}^{\star}_{within}$$

$$\hat{\sigma}^2_u = \frac{1}{N^2T(N-1)}\left(\sum_{j=1}^{N}\sum_{t=1}^{T}\left(\left(\sum_{i=1}^{N}\hat{u}_{ijt}{}^{\star}\right)^2 - \sum_{i=1}^{N}\hat{u}_{ijt}{}^{\star 2}\right)\right)$$

$$\hat{\sigma}^2_v = \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\hat{u}^{\star 2}_{ijt} - \hat{\sigma}^2_\epsilon - \hat{\sigma}^2_u$$

Finally, to derive the estimators of the variance components for model (7), we need first the appropriate Within transformation

$$\begin{aligned} u^*_{ijt,within} = (u^*_{ijt} - 1/T\sum_t u^*_{ijt} - 1/N\sum_i u^*_{ijt} - 1/N\sum_j u^*_{ijt} + 1/N^2\sum_i\sum_j u^*_{ijt} \\ + 1/(NT)\sum_i\sum_t u^*_{ijt} + 1/(NT)\sum_j\sum_t u^*_{ijt} - 1/(N^2T)\sum_i\sum_j\sum_t u^*_{ijt}) \end{aligned}$$

Carrying out the derivation as earlier, we get to the following estimators

$$\begin{aligned} \hat{\sigma}^2_\epsilon &= \frac{N^2T}{(N-1)^2(T-1)}\hat{u}^{\star\prime}_{within}\hat{u}^{\star}_{within} \\ \hat{\sigma}^2_v &= \frac{1}{N^2T(N-1)}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\left(\sum_{j=1}^{N}\hat{u}^{\star}\right)^2 - \sum_{j=1}^{N}\hat{u}^{\star 2}\right)\right) \\ \hat{\sigma}^2_u &= \frac{1}{N^2T(N-1)}\left(\sum_{j=1}^{N}\sum_{t=1}^{T}\left(\left(\sum_{i=1}^{N}\hat{u}^{\star}\right)^2 - \sum_{i=1}^{N}\hat{u}^{\star 2}\right)\right) \\ \hat{\sigma}^2_\mu &= \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\hat{u}^{\star 2}_{ijt} - \hat{\sigma}^2_\epsilon - \hat{\sigma}^2_v - \hat{\sigma}^2_u \end{aligned} \tag{18}$$

Now we have all the tools to properly use the FGLS estimators.

5. Some Data Issues

Like in the case of the usual panel data models, just more frequently, one may be faced with a situation when the data at hand is unbalanced. In our framework of analysis this means that for all models considered, in general $t = 1,\ldots,T_{ij}$, $\sum_i\sum_j T_{ij} = T$ and T_{ij} often is not equal to $T_{i'j'}$. For this unbalanced data case, as we did when the data was balanced, we need to derive the covariance matrices of the models and the appropriate estimators for the variance components.

For model (1), using decomposition (8) we get

$$
\begin{aligned}
u^{\star}_{ij} &= \mu_{ij} \otimes l_{T_{ij}} + \epsilon_{ij} \\
E\left[u^{\star}_{ij} {u^{\star}}'_{ij}\right] &= E\left[\left(\mu_{ij} \otimes l_{T_{ij}}\right)\left(\mu_{ij} \otimes l_{T_{ij}}\right)'\right] + E\left[\epsilon_{ij}\epsilon'_{ij}\right] = \\
&= \sigma^2_{\mu} J_{T_{ij}} + \sigma^2_{\epsilon} I_{T_{ij}} \\
\text{and} \qquad u^{\star}_{i} &= \tilde{\mu}_i + \epsilon_i \\
E\left[u^{\star}_{i} {u^{\star}}'_{i}\right] &= E\left[\tilde{\mu}_i \tilde{\mu}_i{}'\right] + E\left[\epsilon_i \epsilon'_i\right] \\
&= \sigma^2_{\mu} A + \sigma^2_{\epsilon} I_{\sum_{j=1}^{N} T_{ij}}
\end{aligned}
$$

$$
\text{where} \qquad \tilde{\mu}_i = \begin{pmatrix} \mu_{i1} \\ \vdots \\ \mu_{i1} \\ \mu_{i2} \\ \vdots \\ \mu_{i2} \\ \vdots \\ \mu_{iN} \\ \vdots \\ \mu_{iN} \end{pmatrix}, \quad A = \begin{pmatrix} J_{T_{i1}} & 0 & \dots & 0 \\ 0 & J_{T_{i2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{T_{iN}} \end{pmatrix} \quad \text{of size} \sum_{j=1}^{N} T_{ij} \times \sum_{j=1}^{N} T_{ij}
$$

and finally for the complete model

$$
\begin{aligned}
u^{\star} &= \tilde{\mu} + \epsilon \\
E\left[u^{\star} {u^{\star}}'\right] &= E\left[\tilde{\mu}\tilde{\mu}'\right] + E\left[\epsilon\epsilon'\right] \\
&= \sigma^2_{\mu} B + \sigma^2_{\epsilon} I_T
\end{aligned}
$$

$$
\text{where} \qquad \tilde{\mu} = \begin{pmatrix} \mu_{11} \\ \vdots \\ \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{12} \\ \vdots \\ \mu_{ij} \\ \vdots \\ \mu_{ij} \\ \vdots \\ \mu_{NN} \\ \vdots \\ \mu_{NN} \end{pmatrix}, \qquad B = \begin{pmatrix} J_{T_{11}} & 0 & \dots & 0 \\ 0 & J_{T_{12}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{T_{NN}} \end{pmatrix} \quad \text{of size} \quad (T \times T)
$$

Continuing with model (3)

$$
\begin{aligned}
u^\star_{ij} &= \mu_{ij} \otimes l_{T_{ij}} + \lambda + \epsilon_{ij} \\
E\left[u^\star_{ij} {u^\star}'_{ij}\right] &= E\left[\left(\mu_{ij} \otimes l_{T_{ij}}\right)\left(\mu_{ij} \otimes l_{T_{ij}}\right)'\right] + E\left[\lambda\lambda'\right] + E\left[\epsilon_{ij}\epsilon'_{ij}\right] \\
&= \sigma^2_\mu J_{T_{ij}} + \sigma^2_\lambda I_{T_{ij}} + \sigma^2_\epsilon I_{T_{ij}} \\
u^\star_i &= \tilde{\mu}_i + \tilde{\lambda}_i + \epsilon_i
\end{aligned}
$$

where

$$
\begin{aligned}
\tilde{\lambda}'_i &= (\lambda_1, \lambda_2, \dots, \lambda_{T_{i1}}, \dots, \lambda_1, \lambda_2, \dots, \lambda_{T_{iN}}) \\
E\left[u^\star_i {u^\star}'_i\right] &= E\left[\tilde{\mu}_i \tilde{\mu}_i{}'\right] + E\left[\tilde{\lambda}_i \tilde{\lambda}_i{}'\right] + E\left[\epsilon_i \epsilon'_i\right] \\
&= \sigma^2_\mu A + \sigma^2_\lambda D_i + \sigma^2_\epsilon I_{\sum_{j=1}^N T_{ij}} \\
u^\star &= \tilde{\mu} + \tilde{\lambda} + \epsilon \\
E\left[u^\star {u^\star}'\right] &= E\left[\tilde{\mu}\tilde{\mu}'\right] + E\left[\tilde{\lambda}\tilde{\lambda}'\right] + E\left[\epsilon\epsilon'\right] \\
&= \sigma^2_\mu B + \sigma^2_\lambda E + \sigma^2_\epsilon I_T
\end{aligned}
$$

with

$$
E(E_{11}, E_{12}, \dots, E_{1N}, \dots, E_{N1}, E_{N2}, \dots, E_{NN})
$$

$$E_{ij} = \begin{pmatrix} M_{T_{11} \times T_{ij}} \\ M_{T_{12} \times T_{ij}} \\ \vdots \\ M_{T_{NN} \times T_{ij}} \end{pmatrix} \quad \text{and} \quad D_i = \begin{pmatrix} I_{T_{i1}} & M_{T_{i1} \times T_{i2}} & \dots & M_{T_{i1} \times T_{iN}} \\ M_{T_{i2} \times T_{i1}} & I_{T_{i1}} & \dots & M_{T_{i2} \times T_{iN}} \\ \vdots & \vdots & \ddots & \vdots \\ M_{T_{iN} \times T_{i1}} & M_{T_{iN} \times T_{i2}} & \dots & I_{T_{iN}} \end{pmatrix}$$

where

$$M_{T_{ij} \times T_{lj}} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \end{pmatrix} \quad \text{if} \quad T_{lj} > T_{ij}$$

and

$$M_{T_{ij} \times T_{lj}} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \quad \text{if} \quad T_{lj} < T_{ij}$$

Doing the same exercise for model (4) using decomposition (9) we end up with

$$\begin{aligned} u^\star_{ij} &= u_j + \epsilon_{ij} \\ E\left(u^\star_{ij} {u^\star}'_{ij}\right) &= E\left[u_j u_j'\right] + E\left[\epsilon_{ij}\epsilon_{ij}'\right] = \sigma_u^2 I_{T_{ij}} + \sigma_\epsilon^2 I_{T_{ij}} \\ u^\star_i &= u + \epsilon_i \\ E\left(u^\star_i {u^\star}'_i\right) &= E\left[uu'\right] + E\left[\epsilon_i \epsilon_i'\right] = \sigma_u^2 I_{\sum_{j=1}^N T_{ij}} + \sigma_\epsilon^2 I_{\sum_{j=1}^N T_{ij}} \\ u^\star &= \tilde{u} + \epsilon \end{aligned}$$

and so for the complete model we get

$$E\left(u^\star {u^\star}'\right) = E\left[\tilde{u}\tilde{u}'\right] + E\left[\epsilon\epsilon'\right] = \sigma_u^2 C + \sigma_\epsilon^2 I_T$$

where

$$\tilde{u}' = (u_{11}, \dots, u_{1T_{11}}, \dots, u_{N1}, \dots, u_{NT_{1N}}, \dots, u_{11}, \dots, u_{1T_{N1}}, \dots, u_{N1}, \dots, u_{NT_{NN}})$$

$C = (C_1, C_2, C_3)$

$$C_1 = \begin{pmatrix} I_{T_{11}} & 0 & \dots & 0 \\ 0 & I_{T_{12}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I_{T_{1N}} \\ M_{T_{21}\times T_{11}} & 0 & \dots & 0 \\ 0 & M_{T_{22}\times T_{12}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_{T_{2N}\times T_{1N}} \\ \vdots & \vdots & \ddots & \vdots \\ M_{T_{N1}\times T_{11}} & 0 & \dots & 0 \\ 0 & M_{T_{N2}\times T_{12}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_{T_{NN}\times T_{1N}} \end{pmatrix}$$

$$C_2 = \begin{pmatrix} M_{T_{11}\times T_{21}} & 0 & \dots & 0 & \dots \\ 0 & M_{T_{12}\times T_{22}} & \dots & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \dots \\ 0 & 0 & \dots & M_{T_{1N}\times T_{2N}} & \dots \\ I_{T_{21}} & 0 & \dots & 0 & \dots \\ 0 & I_{T_{22}} & \dots & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \dots \\ 0 & 0 & \dots & I_{T_{2N}} & \dots \\ \vdots & \vdots & \ddots & \vdots & \dots \\ M_{T_{N1}\times T_{21}} & 0 & \dots & 0 & \dots \\ 0 & M_{T_{N2}\times T_{22}} & \dots & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \dots \\ 0 & 0 & \dots & M_{T_{NN}\times T_{1N}} & \dots \end{pmatrix}$$

$$C_3 = \begin{pmatrix} M_{T_{11} \times T_{N1}} & 0 & \dots & 0 \\ 0 & M_{T_{12} \times T_{N2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_{T_{1N} \times T_{NN}} \\ M_{T_{21} \times T_{N1}} & 0 & \dots & 0 \\ 0 & M_{T_{22} \times T_{N2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_{T_{2N} \times T_{NN}} \\ \vdots & \vdots & \ddots & \vdots \\ I_{T_{N1}} & 0 & \dots & 0 \\ 0 & I_{T_{N2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I_{T_{NN}} \end{pmatrix}$$

Let us now turn to model (5). Following the same steps as above, we get for the covariance matrix $(\sigma_v^2 D + \sigma_\epsilon^2 I_T)$ where

$$D = \begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ 0 & 0 & \dots & D_N \end{pmatrix}$$

Models (6) and (7) can be dealt with together using decomposition (10)

$$\begin{aligned} u_{ij}^\star &= \mu_{ij} \otimes l_T + u_j + v_i + \epsilon_{ij} \\ E\left(u_{ij}^\star {u^\star}'_{ij}\right) &= E\left[\left(\mu_{ij} \otimes l_{T_{ij}}\right)\left(\mu_{ij} \otimes l'_{T_{ij}}\right)\right] + E\left[u_j u_j'\right] + E\left[v_i v_i'\right] + E\left[\epsilon_{ij}\epsilon_{ij}'\right] \\ &= \sigma_\mu^2 J_{T_{ij}} + \sigma_u^2 I_{T_{ij}} + \sigma_v^2 I_{T_{ij}} + \sigma_\epsilon^2 I_{T_{ij}} \\ u_i^\star &= \tilde{\mu_i} + \tilde{v_i} + u + \epsilon_i \\ E\left(u_i^\star {u^\star}'_i\right) &= E\left[\tilde{\mu_i}\tilde{\mu_i}'\right] + E\left[\tilde{v_i}\tilde{v_i}'\right] + E\left[uu'\right] + E\left[\epsilon_i\epsilon_i'\right] \\ &= \sigma_\mu^2 A + \sigma_u^2 I_{\sum_{j=1}^N T_{ij}} + \sigma_v^2 D_i + \sigma_\epsilon^2 I_{\sum_{j=1}^N T_{ij}} \\ u^\star &= \tilde{\mu} + \tilde{v} + \tilde{u} + \epsilon \\ \text{where} \quad \tilde{v_i}' &= (v_{i1}, v_{i2}, \dots, v_{iT_{i1}}, v_{i1}, v_{i2}, \dots, v_{iT_{i2}}, \dots, v_{i1}, v_{i2}, \dots, v_{iT_{iN}}) \\ \tilde{v}' &= (\tilde{v_1}, \tilde{v_2}, \dots, \tilde{v_N},) \\ E\left(u^\star {u^\star}'\right) &= E\left[\tilde{\mu}\tilde{\mu}'\right] + E\left[\tilde{v}\tilde{v}'\right] + E\left[\tilde{u}\tilde{u}'\right] + E\left[\epsilon\epsilon'\right] = \\ &= \sigma_\mu^2 B + \sigma_u^2 C + \sigma_v^2 D + \sigma_\epsilon^2 I_T \end{aligned}$$

For model (6) the appropriate covariance matrix is the same with $B = 0$.

Now that we derived the covariance matrices for unbalanced data it is time to turn to the estimation of the variance components. Using (11) and (12) for model (1)

$$E\left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(u^{\star}_{ijt} - \bar{u^{\star}}_{ij}\right)^2\right] = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}E\left[\left(\epsilon_{ijt} - \bar{\epsilon}_{ij}\right)^2\right]$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}E\left[\epsilon^2_{ijt} - 2\epsilon_{ijt}\frac{1}{T_{ij}}\sum_{t=1}^{T_{ij}}\epsilon_{ijt} + \left(\frac{1}{T_{ij}}\sum_{t=1}^{T_{ij}}\epsilon_{ijt}\right)^2\right]$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(E\left[\epsilon^2_{ijt}\right] - 2E\left[\epsilon_{ijt}\frac{1}{T_{ij}}\sum_{t=1}^{T_{ij}}\epsilon_{ijt}\right] + E\left[\left(\frac{1}{T_{ij}}\sum_{t=1}^{T_{ij}}\epsilon_{ijt}\right)^2\right]\right)$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\sigma^2_{\epsilon} - \frac{2}{T_{ij}}\sigma^2_{\epsilon} + \frac{1}{T_{ij}}\sigma^2_{\epsilon}\right) = \sigma^2_{\epsilon}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{T_{ij}-1}{T_{ij}}$$

so for the variance components we get the following estimators

$$\hat{\sigma}^2_{\epsilon} = \frac{N^2}{\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{T_{ij}-1}{T_{ij}}}\hat{u}^{\star'}_{within}\hat{u}^{\star}_{within}$$

$$\hat{\sigma}^2_{\mu} = \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}\hat{u^{\star}}^2_{ijt} - \hat{\sigma}^2_{\epsilon}$$

For model (3) we have

$$E\left[{u^{\star}}^2_{ijt}\right] = \sigma^2_{\mu} + \sigma^2_{\lambda} + \sigma^2_{\epsilon}$$

$$E\left[\left(\frac{1}{N}\sum_{i=1}^{N}u^{\star}_{ijt}\right)^2\right] = \frac{1}{N}\sigma^2_{\mu} + \sigma^2_{\lambda} + \frac{1}{N}\sigma^2_{\epsilon}$$

$$E\left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{1}{T_{ij}}\sum_{j=1}^{T_{ij}}u^{\star}_{ijt}\right)^2\right] = \sigma^2_{\mu} + \left(\sigma^2_{\lambda} + \sigma^2_{\epsilon}\right)\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{T_{ij}}$$

and therefore for the variance components we get the following estimators

$$\hat{\sigma}^2_\mu = \frac{1}{N^2 - \sum_{i=1}^N \sum_{j=1}^N \frac{1}{T_{ij}}} \left(\sum_{i=1}^N \sum_{j=1}^N \left(\frac{1}{T_{ij}} \sum_{t=1}^{T_{ij}} \hat{u^\star}_{ijt} \right)^2 - \left(\sum_{i=1}^N \sum_{j=1}^N \frac{1}{T_{ij}} \right) \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^{T_{ij}} \hat{u^\star}^2_{ijt} \right)$$

$$\hat{\sigma}^2_\lambda = \frac{1}{N-1} \left(\sum_{i=1}^N \frac{1}{\sum_{j=1}^N T_{ij}} \sum_{j=1}^N \sum_{t=1}^{T_{ij}} \left(\frac{1}{N} \sum_{i=1}^N \hat{u^\star}_{ijt} \right)^2 - \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^{T_{ij}} \hat{u^\star}^2_{ijt} \right)$$

$$\hat{\sigma}_{\epsilon^2} = \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^{T_{ij}} \hat{u^\star}^2_{ijt} - \hat{\sigma}^2_\mu - \hat{\sigma^2_\lambda}$$

For model (4) (and similarly for model (5)), using (13) and (14), and using the same derivations as there, we get

$$\hat{\sigma}^2_\epsilon = \frac{N}{N-1} \hat{u}^{\star'}_{within} \hat{u}^\star_{within}$$

$$\hat{\sigma}^2_u = \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^{T_{ij}} \hat{u}^{\star 2}_{ijt} - \hat{\sigma}^2_\epsilon$$

Turning now to model (6), as (15) and (16) are the same in the unbalanced case we get

$$\hat{\sigma}^2_\epsilon = \frac{N^2}{(N-1)^2} \hat{u}^{\star'}_{within} \hat{u}^\star_{within}$$

$$\hat{\sigma}^2_u = \frac{1}{N-1} \left(\sum_{i=1}^N \frac{1}{\sum_{j=1}^N T_{ij}} \sum_{j=1}^N \sum_{t=1}^{T_{ij}} \left(\frac{1}{N} \sum_{i=1}^N \hat{u^\star}_{ijt} \right)^2 - \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^{T_{ij}} \hat{u}^{\star 2}_{ijt} \right)$$

$$\hat{\sigma^2_v} = \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^{T_{ij}} \hat{u}^{\star 2}_{ijt} - \hat{\sigma}^2_\epsilon - \hat{\sigma}^2_u$$

And finally for model (7) we get

$$E\left[{u^{\star}}_{ijt}^{2}\right] = E\left[(\mu_{ij} + u_{jt} + v_{it} + \epsilon_{ijt})^2\right] = \sigma_\mu^2 + \sigma_u^2 + \sigma_v^2 + \sigma_\epsilon^2$$

$$E\left[\left(\frac{1}{N}\sum_{i=1}^{N} u^{\star}_{ijt}\right)^2\right] = E\left[\left(\frac{1}{N}\sum_{i=1}^{N}(\mu_{ij} + u_{jt} + v_{it} + \epsilon_{ijt})\right)^2\right]$$
$$= \frac{1}{N}\sigma_\mu^2 + \sigma_u^2 + \frac{1}{N}\sigma_v^2 + \frac{1}{N}\sigma_\epsilon^2$$

$$E\left[\left(\frac{1}{N}\sum_{j=1}^{N} u^{\star}_{ijt}\right)^2\right] = E\left[\left(\frac{1}{N}\sum_{j=1}^{N}(\mu_{ij} + u_{jt} + v_{it} + \epsilon_{ijt})\right)^2\right]$$
$$= \frac{1}{N}\sigma_\mu^2 + \frac{1}{N}\sigma_u^2 + \sigma_v^2 + \frac{1}{N}\sigma_\epsilon^2$$

$$E\left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{1}{T_{ij}}\sum_{j=1}^{T_{ij}} u^{\star}_{ijt}\right)^2\right] = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} E\left[\left(\frac{1}{T_{ij}}\sum_{j=1}^{T_{ij}} u^{\star}_{ijt}\right)^2\right]$$
$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{T_{ij}^2} E\left[\left(\sum_{t=1}^{T_{ij}}(\mu_{ij} + u_{jt} + v_{it} + \epsilon_{ijt})\right)^2\right]$$
$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{T_{ij}^2}\left(E\left[\sum_{t=1}^{T_{ij}}\mu_{ij}^2\right] + E\left[\sum_{t=1}^{T_{ij}}u_{jt}^2\right] + E\left[\sum_{t=1}^{T_{ij}}v_{it}^2\right] + E\left[\sum_{t=1}^{T_{ij}}\epsilon_{ijt}^2\right]\right)$$
$$= \sigma_\mu^2 + \left(\sigma_u^2 + \sigma_v^2 + \sigma_\epsilon^2\right)\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{T_{ij}}$$

Putting all these together, the estimators of the variance components are

$$\hat{\sigma}_\mu^2 = \frac{1}{N^2 - \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{T_{ij}}}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{1}{T_{ij}}\sum_{t=1}^{T_{ij}}{\hat{u}^{\star}}_{ijt}\right)^2 - \left(\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{T_{ij}}\right)\frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}{\hat{u}^{\star}}_{ijt}^{2}\right)$$

$$\hat{\sigma}_u^2 = \frac{1}{N-1}\left(\sum_{i=1}^{N}\frac{1}{\sum_{j=1}^{N}T_{ij}}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}\left(\frac{1}{N}\sum_{i=1}^{N}{\hat{u}^{\star}}_{ijt}\right)^2 - \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}{\hat{u}^{\star}}_{ijt}^{2}\right)$$

$$\hat{\sigma}_v^2 = \frac{1}{N-1}\left(\sum_{j=1}^{N}\frac{1}{\sum_{i=1}^{N}T_{ij}}\sum_{i=1}^{N}\sum_{t=1}^{T_{ij}}\left(\frac{1}{N}\sum_{j=1}^{N}{\hat{u}^{\star}}_{ijt}\right)^2 - \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}{\hat{u}^{\star}}_{ijt}^{2}\right)$$

$$\hat{\sigma}_\epsilon^2 = \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}{\hat{u}^{\star}}_{ijt}^{2} - \hat{\sigma}_\mu^2 - \hat{\sigma}_u^2 - \hat{\sigma}_v^2$$

Now let us have a look at another potential data issue. As in fact we are mostly dealing here with flow types of data like trade, capital movements (FDI), etc., it is important to have a closer look at the case when, by nature, we do not observe self-flow. This means that from the (ijt) indexes we do not have observations for the dependent variable of the model when $i = j$ for any t. This implies that we relax our initial implicit assumption that the observation sets i and j are equivalent.

The no-self-flow case has surprisingly little effect vis-a-vis to what has been said so far. In fact only models (6) and (7) are affected. As in this case the Within transformation does not cancel out all the effects, for these models the estimation of the variance components needs to be modified slightly, such as

$$\hat{\sigma}_u^2 = \frac{1}{N(N-1)(N-2)T}\left(\sum_{j=1,j\neq i}^{N}\sum_{t=1}^{T}\left(\left(\sum_{i=1}^{N}\hat{u}_{ijt}^{\star}\right)^2 - \sum_{i=1}^{N}\left(\hat{u}_{ijt}^{\star}\right)^2\right)\right)$$

$$\hat{\sigma}_v^2 = \frac{1}{N(N-1)(N-2)T}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\left(\sum_{j=1,j\neq i}^{N}\hat{u}_{ijt}^{\star}\right)^2 - \sum_{j=1,j\neq i}^{N}\left(\hat{u}_{ijt}^{\star}\right)^2\right)\right)$$

$$\hat{\sigma}_\epsilon^2 = \frac{1}{N(N-1)T}\sum_{i=1,i\neq j}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\hat{u}_{ijt}^{\star 2} - \hat{\sigma}_v^2 - \hat{\sigma}_u^2$$

for model (6), and

$$\hat{\sigma}_\mu^2 = \frac{1}{N(N-1)T(T-1)}\left(\sum_{i=1,i\neq j}^{N}\sum_{j=1}^{N}\left(\left(\sum_{t=1}^{T}\hat{u}_{ijt}^{\star}\right)^2 - \sum_{t=1}^{T}\left(\hat{u}_{ijt}^{\star}\right)^2\right)\right)$$

$$\hat{\sigma}_v^2 = \frac{1}{N(N-1)(N-2)T}\left(\sum_{i=1,i\neq j}^{N}\sum_{t=1}^{T}\left(\left(\sum_{j=1}^{N}\hat{u}_{ijt}^{\star}\right)^2 - \sum_{j=1}^{N}\left(\hat{u}_{ijt}^{\star}\right)^2\right)\right)$$

$$\hat{\sigma}_u^2 = \frac{1}{N(N-1)(N-2)T}\left(\sum_{j=1,j\neq i}^{N}\sum_{t=1}^{T}\left(\left(\sum_{i=1}^{N}\hat{u}_{ijt}^{\star}\right)^2 - \sum_{i=1}^{N}\left(\hat{u}_{ijt}^{\star}\right)^2\right)\right)$$

$$\hat{\sigma}_\epsilon^2 = \frac{1}{N(N-1)T}\sum_{i=1,i\neq j}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\hat{u}_{ijt}^{\star 2} - \hat{\sigma}_\mu^2 - \hat{\sigma_u^2} - \hat{\sigma_v^2}$$

for model (7).

6. An Extension

In the case of trade models, and in general for most of the flow type data, Assumption (2) may well be too restrictive as it does not allow for any type of cross correlation. We can replace this assumption by introducing a simple cross-correlation such as

$$E(\mu_{ij}\mu_{ks}) = \begin{cases} \sigma_\mu^2 & \text{if } i=k, j=s \\ \rho_{(1)} & \text{if } i=k, j\neq s \\ \rho_{(2)} & \text{if } i\neq k, j=s \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

This will affect models (1), (3) and (7). The respective covariance matrices are now

$$\begin{aligned} E[u^\star_{ij}u^{\star\prime}_{ij}] &= \sigma_\mu^2 J_T + \sigma_\epsilon^2 I_T \\ E[u^\star_{i}u^{\star\prime}_{i}] &= \sigma_\mu^2 I_N \otimes J_T + \rho_{(1)}\left(J_{NT} - I_N \otimes J_T\right) + \sigma_\epsilon^2 I_{NT} \\ E[u^\star u^{\star\prime}] &= \sigma_\mu^2 I_{N^2} \otimes J_T + \rho_{(1)}\left(I_N \otimes J_{NT} - I_{N^2} \otimes J_T\right) + \\ &\quad + \rho_{(2)}\left((J_N - I_N)\otimes(I_N \otimes J_T)\right) + \sigma_\epsilon^2 I_{N^2T} \end{aligned}$$

for model (1),

$$\begin{aligned} E[u^\star u^{\star\prime}] = {} & \sigma_\mu^2 I_{N^2} \otimes J_T + \rho_{(1)}\left(I_N \otimes J_{NT} - I_{N^2} \otimes J_T\right) + \rho_{(2)}\left((J_N - I_N)\otimes(I_N \otimes J_T)\right) + \\ & + \sigma_\lambda^2 J_{N^2} \otimes I_T + \sigma_\epsilon^2 I_{N^2T} \end{aligned}$$

for model (3), and finally

$$\begin{aligned} E[u^\star u^{\star\prime}] = {} & \sigma_\mu^2\left(I_{N^2} \otimes J_T\right) + \rho_{(1)}\left(I_N \otimes J_{NT} - I_{N^2} \otimes J_T\right) + \rho_{(2)}\left((J_N - I_N)\otimes(I_N \otimes J_T)\right) + \\ & + \sigma_u^2\left(J_N \otimes I_{NT}\right) + \sigma_v^2\left(I_N \otimes J_N \otimes I_T\right) + \sigma_\epsilon^2 I_{N^2T} \end{aligned}$$

for model (7). The estimation of the variance components for models (1) and (3) will not change, but we need, of course, to estimate the cross-correlations. For model (1) this can be carried out such as

$$\begin{aligned} E\left[\left(\frac{1}{N}\sum_{i=1}^N u^\star_{ijt}\right)^2\right] &= E\left[\left(\frac{1}{N}\sum_{i=1}^N \mu_{ij} + \epsilon_{ijt}\right)^2\right] \\ &= \frac{1}{N^2}E\left[\sum_{i=1}^N \mu_{ij}^2 + 2\sum_{i,s}\mu_{ij}\mu_{sj} + \sum_{i=1}^N \epsilon_{ijt}^2\right] \\ &= \frac{1}{N}\sigma_\mu^2 + \frac{1}{N}\sigma_\epsilon^2 + \frac{N-1}{N}\rho_{(2)} \\ E\left[\left(\frac{1}{N}\sum_{j=1}^N u^\star_{ijt}\right)^2\right] &= \frac{1}{N}\sigma_\mu^2 + \frac{1}{N}\sigma_\epsilon^2 + \frac{N-1}{N}\rho_{(1)} \end{aligned}$$

So we get

$$\hat{\rho}_{(1)} = \frac{1}{N^2(N-1)T}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\sum_{j=1}^{N}\hat{u}^{\star}_{ijt}\right)^2 - \frac{1}{N-1}\hat{\sigma}^2_{\mu} - \frac{1}{N-1}\hat{\sigma}^2_{\epsilon}$$

$$\hat{\rho}_{(2)} = \frac{1}{N^2(N-1)T}\sum_{j=1}^{N}\sum_{t=1}^{T}\left(\sum_{i=1}^{N}\hat{u}^{\star}_{ijt}\right)^2 - \frac{1}{N-1}\hat{\sigma}^2_{\mu} - \frac{1}{N-1}\hat{\sigma}^2_{\epsilon}$$

For model (3)

$$E\left[\left(\frac{1}{N}\sum_{i=1}^{N}u^{\star}_{ijt}\right)^2\right] = \frac{1}{N}\sigma^2_{\mu} + \sigma^2_{\lambda} + \frac{1}{N}\sigma^2_{\epsilon} + \frac{N-1}{N}\rho_{(2)}$$

$$E\left[\left(\frac{1}{N}\sum_{j=1}^{N}u^{\star}_{ijt}\right)^2\right] = \frac{1}{N}\sigma^2_{\mu} + \sigma^2_{\lambda} + \frac{1}{N}\sigma^2_{\epsilon} + \frac{N-1}{N}\rho_{(1)}$$

and so

$$\hat{\rho}_{(1)} = \frac{1}{N^2(N-1)T}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\sum_{j=1}^{N}\hat{u}^{\star}_{ijt}\right)^2 - \frac{1}{N-1}\hat{\sigma}^2_{\mu} - \frac{N}{N-1}\hat{\sigma}^2_{\lambda} - \frac{1}{N-1}\hat{\sigma}^2_{\epsilon}$$

$$\hat{\rho}_{(2)} = \frac{1}{N^2(N-1)T}\sum_{j=1}^{N}\sum_{t=1}^{T}\left(\sum_{i=1}^{N}\hat{u}^{\star}_{ijt}\right)^2 - \frac{1}{N-1}\hat{\sigma}^2_{\mu} - \frac{N}{N-1}\hat{\sigma}^2_{\lambda} - \frac{1}{N-1}\hat{\sigma}^2_{\epsilon}$$

Finally, in the case of model (7), the Within transformation remains unchanged so estimation of the variance component of ϵ is the same as in (18), otherwise

$$E\left[{u^{\star}}^2_{ijt}\right] = \sigma^2_{\mu} + \sigma^2_{v} + \sigma^2_{u} + \sigma^2_{\epsilon}$$

$$E\left[\left(\frac{1}{N}\sum_{i=1}^{N}u^{\star}_{ijt}\right)^2\right] = \frac{1}{N}\sigma^2_{\mu} + \frac{1}{N}\sigma^2_{v} + \sigma^2_{u} + \frac{1}{N}\sigma^2_{\epsilon} + \frac{N-1}{N}\rho_{(2)}$$

$$E\left[\left(\frac{1}{N}\sum_{j=1}^{N}u^{\star}_{ijt}\right)^2\right] = \frac{1}{N}\sigma^2_{\mu} + \sigma^2_{v} + \frac{1}{N}\sigma^2_{u} + \frac{1}{N}\sigma^2_{\epsilon} + \frac{N-1}{N}\rho_{(1)}$$

$$E\left[\left(\frac{1}{T}\sum_{t=1}^{T}u^{\star}_{ijt}\right)^2\right] = \sigma^2_{\mu} + \frac{1}{T}\sigma^2_{v} + \frac{1}{T}\sigma^2_{u} + \frac{1}{T}\sigma^2_{\epsilon}$$

$$E\left[\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}u^{\star}_{ijt}\right)^2\right] = \frac{1}{N}\sigma^2_{\mu} + \frac{1}{NT}\sigma^2_{v} + \frac{1}{T}\sigma^2_{u} + \frac{1}{NT}\sigma^2_{\epsilon} + \frac{N-1}{N}\rho_{(2)}$$

And by solving this system we get

$$\hat{\sigma}_{\mu}^{2}=\frac{1}{N^{2}T(T-1)}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\left(\sum_{t=1}^{T}\hat{u}_{ijt}^{\star}\right)^{2}-\sum_{t=1}^{T}\left(\hat{u}_{ijt}^{\star}\right)^{2}\right)\right)$$

$$\hat{\rho}_{2}=\frac{1}{N^{2}(N-1)T(T-1)}\times$$
$$\times\sum_{j=1}^{N}\left(\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{u}_{ijt}^{\star}\right)^{2}-\sum_{i=1}^{N}\left(\sum_{t=1}^{T}\hat{u}_{ijt}^{\star}\right)^{2}-\sum_{t=1}^{T}\left(\sum_{i=1}^{N}\hat{u}_{ijt}^{\star}\right)^{2}+\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat{u}_{ijt}^{\star}\right)^{2}\right)$$

$$\hat{\sigma}_{u}^{2}=\frac{1}{N^{2}(N-1)T}\sum_{j=1}^{N}\sum_{t=1}^{T}\left(\left(\sum_{i=1}^{N}\hat{u}_{ijt}^{\star}\right)^{2}-\sum_{i=1}^{N}\left(\hat{u}_{ijt}^{\star}\right)^{2}\right)-\hat{\rho}_{2}$$

$$\hat{\sigma}_{v}^{2}=\frac{1}{N^{2}T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\left(\hat{u}_{ijt}^{\star}\right)^{2}-\hat{\sigma}_{\mu}^{2}-\hat{\sigma}_{u}^{2}-\hat{\sigma}_{\epsilon}^{2}$$

$$\hat{\rho}_{1}=\frac{1}{N^{2}(N-1)T}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\left(\sum_{j=1}^{N}\hat{u}_{ijt}^{\star}\right)^{2}-\sum_{j=1}^{N}\left(\hat{u}_{ijt}^{\star}\right)^{2}\right)-\hat{\sigma}_{v}^{2}$$

Now doing the same exercise as above, but for the unbalanced data case, leads to slightly more complicated terms. Using some definitions introduced in Section 5 above, the covariance matrix of model (1) is

$$\begin{aligned}E[u_{ij}^{\star}u^{\star\prime}_{ij}]&=\sigma_{\mu}^{2}J_{T_{ij}}+\sigma_{\epsilon}^{2}I_{T_{ij}}\\E[u_{i}^{\star}u^{\star\prime}_{i}]&=\sigma_{\mu}^{2}A_{i}+\rho_{(1)}\left(J_{\sum_{j}T_{ij}}-A_{i}\right)+\sigma_{\epsilon}^{2}I_{\sum_{j}T_{ij}}\\E[u^{\star}u^{\star\prime}]&=\sigma_{\mu}^{2}B+\rho_{(1)}\left(P-B\right)+\rho_{(2)}S+\sigma_{\epsilon}^{2}I_{T}\end{aligned}$$

$$S=\begin{pmatrix}O_{1}&R_{12}&\dots&R_{1N}\\R_{21}&O_{2}&\dots&R_{2N}\\\vdots&\vdots&\ddots&\vdots\\R_{N1}&R_{N2}&\dots&O_{N}\end{pmatrix}$$

where O_i is the matrix of zeros of size $\sum_j T_{ij}\times\sum_j T_{ij}$

$$R_{ij}=\begin{pmatrix}J_{T_{i1}\times T_{j1}}&0&\dots&0\\0&J_{T_{i2}\times T_{j2}}&\dots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\dots&J_{T_{iN}\times T_{jN}}\end{pmatrix}$$

where $J_{T_{is}\times T_{js}}$ is the matrix of ones of size $T_{is}\times T_{js}$, and

$$P=\begin{pmatrix} J_{\sum_j T_{1j}} & 0 & \dots & 0 \\ 0 & J_{\sum_j T_{2j}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{\sum_j T_{Nj}} \end{pmatrix}$$

The covariance matrix of model (3) is

$$E[u^\star u^{\star\prime}]=\sigma_\mu^2 B+\rho_{(1)}\left(P-B\right)+\rho_{(2)}S+\sigma_\lambda^2 E+\sigma_\epsilon^2 I_T$$

and that of model (7) is

$$E[u^\star u^{\star\prime}]=\sigma_\mu^2 B+\rho_{(1)}\left(P-B\right)+\rho_{(2)}S+\sigma_u^2 C+\sigma_v^2 D+\sigma_\epsilon^2 I_T$$

Let us turn now our attention to the estimation of the variance components. In the case of model (1) the estimation of the variance components remains the same as for the "usual" unbalanced case and the estimation of the cross-correlations remains the same as in the balanced cross-correlation case. For model (3) the estimation σ_μ^2 is, again, the same as for "usual" unbalanced case. For the other terms

$$E\left[{u^\star}^2_{ijt}\right]=\sigma_\mu^2+\sigma_\lambda^2+\sigma_\epsilon^2$$

$$E\left[\left(\frac{1}{N}\sum_{i=1}^{N}u^\star_{ijt}\right)^2\right]=\frac{1}{N}\sigma_\mu^2+\sigma_\lambda^2+\frac{1}{N}\sigma_\epsilon^2+\frac{N-1}{N}\rho_{(2)}$$

$$E\left[\left(\frac{1}{N}\sum_{j=1}^{N}u^\star_{ijt}\right)^2\right]=\frac{1}{N}\sigma_\mu^2+\sigma_\lambda^2+\frac{1}{N}\sigma_\epsilon^2+\frac{N-1}{N}\rho_{(1)}$$

$$E\left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{T_{ij}}\sum_{t=1}^{T_{ij}}\left(\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}u^\star_{ijt}\right)^2\right]=\frac{1}{N^2}\sigma_\mu^2+\sigma_\lambda^2+\frac{1}{N^2}\sigma_\epsilon^2+\frac{N-1}{N^2}\rho_{(1)}+$$

$$+\frac{N-1}{N^2}\rho_{(2)}$$

$$\hat{\rho}_{(2)} = \frac{1}{(N-1)^2}\times$$

$$\times\left(\sum_{i=1}^{N}\frac{1}{\sum_{j=1}^{N}T_{ij}}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}\left(\sum_{i=1}^{N}\hat{u^\star}_{ijt}\right)^2 - \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{T_{ij}}\sum_{t=1}^{T_{ij}}\left(\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\hat{u^\star}_{ijt}\right)^2\right)+$$

$$+\frac{N}{(N-1)^2}\left(\sum_{j=1}^{N}\frac{1}{\sum_{i=1}^{N}T_{ij}}\sum_{i=1}^{N}\sum_{t=1}^{T_{ij}}\left(\frac{1}{N}\sum_{j=1}^{N}\hat{u^\star}_{ijt}\right)^2 - \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}\left(\hat{u^\star}_{ijt}\right)^2\right)$$

$$\hat{\sigma}^2_\lambda = \frac{1}{N-1}\left(\sum_{i=1}^{N}\frac{1}{\sum_{j=1}^{N}T_{ij}}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}\left(\frac{1}{N}\sum_{i=1}^{N}\hat{u^\star}_{ijt}\right)^2 - \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}\hat{u^\star}^2_{ijt}\right) - \hat{\rho}_{(2)}$$

$$\hat{\rho}_{(1)} = \frac{1}{N-1}\left(\sum_{j=1}^{N}\frac{1}{\sum_{i=1}^{N}T_{ij}}\sum_{i=1}^{N}\sum_{t=1}^{T_{ij}}\left(\frac{1}{N}\sum_{j=1}^{N}\hat{u^\star}_{ijt}\right)^2 - \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}\hat{u^\star}^2_{ijt}\right) - \hat{\sigma}^2_\lambda$$

$$\hat{\sigma^2_\epsilon} = \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}\hat{u^\star}^2_{ijt} - \hat{\sigma}^2_\mu - \hat{\sigma}^2_\lambda$$

Finally, for model (7) the estimation of σ^2_μ is the same as for usual unbalanced case, and

$$E\left[{u^\star}^2_{ijt}\right] = \sigma^2_\mu + \sigma^2_v + \sigma^2_u + \sigma^2_\epsilon$$

$$E\left[\left(\frac{1}{N}\sum_{i=1}^{N}u^\star_{ijt}\right)^2\right] = \frac{1}{N}\sigma^2_\mu + \frac{1}{N}\sigma^2_v + \sigma^2_u + \frac{1}{N}\sigma^2_\epsilon + \frac{N-1}{N}\rho_{(2)}$$

$$E\left[\left(\frac{1}{N}\sum_{j=1}^{N}u^\star_{ijt}\right)^2\right] = \frac{1}{N}\sigma^2_\mu + \sigma^2_v + \frac{1}{N}\sigma^2_u + \frac{1}{N}\sigma^2_\epsilon + \frac{N-1}{N}\rho_{(1)}$$

$$E\left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{\sum_j T_{ij}}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}u^\star_{ijt}\right)^2\right] =$$

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\sum_j T^2_{ij}}{\left(\sum_j T_{ij}\right)^2}\sigma^2_\mu + \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\left(\sum_j T_{ij}\right)^2}\sum_{t=1}^{max_j T_{ij}}\left(\sum_{j=1}^{N}\delta_{ijt}\right)^2\sigma^2_v +$$

$$+\frac{1}{N}\sum_{i=1}^{N}\frac{\sum_j T_{ij}}{\left(\sum_j T_{ij}\right)^2}\left(\sigma^2_u+\sigma^2_\epsilon\right) + \frac{2}{N}\sum_{i=1}^{N}\frac{1}{\left(\sum_j T_{ij}\right)^2}\sum_{j=1}^{N-1}\sum_{s=j+1}^{N}T_{ij}T_{is}\rho_{(1)}$$

$$E\left[\frac{1}{N}\sum_{j=1}^{N}\left(\frac{1}{\sum_i T_{ij}}\sum_{i=1}^{N}\sum_{t=1}^{T_{ij}}u^{\star}_{ijt}\right)^2\right]=$$

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\sum_i T_{ij}^2}{\left(\sum_i T_{ij}\right)^2}\sigma_\mu^2+\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\left(\sum_i T_{ij}\right)^2}\sum_{t=1}^{max_i T_{ij}}\left(\sum_{i=1}^{N}\delta_{ijt}\right)^2\sigma_u^2+$$

$$+\frac{1}{N}\sum_{j=1}^{N}\frac{\sum_i T_{ij}}{\left(\sum_i T_{ij}\right)^2}\left(\sigma_v^2+\sigma_\epsilon^2\right)+\frac{2}{N}\sum_{j=1}^{N}\frac{1}{\left(\sum_i T_{ij}\right)^2}\sum_{i=1}^{N-1}\sum_{k=i+1}^{N}T_{ij}T_{kj}\rho_{(2)}$$

where

$$\delta_{ijt}=\begin{cases}1, & \text{if } t\le T_{ij}\\ 0, & \text{if } t> T_{ij}\end{cases}$$

Solving the system yields the following estimators

$$\hat{\rho_{(1)}}=\frac{N}{\sum_{i=1}^{N}\frac{1}{\left(\sum_j T_{ij}\right)^2}\left(\sum_j T_{ij}+2\sum_{j=1}^{N-1}\sum_{s=j+1}^{N}T_{ij}T_{is}-\sum_{t=1}^{max_j T_{ij}}\left(\sum_{j=1}^{N}\delta_{ijt}\right)^2\right)}\times$$

$$\times\left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{\sum_j T_{ij}}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}u^{\star}_{ijt}\right)^2-\frac{1}{N(N-1)}\sum_{i=1}^{N}\frac{1}{\left(\sum_j T_{ij}\right)^2}\sum_{t=1}^{max_j T_{ij}}\left(\sum_{j=1}^{N}\delta_{ijt}\right)^2\times\right.$$

$$\times\left(\sum_{j=1}^{N}\frac{1}{\sum_{i=1}^{N}T_{ij}}\sum_{i=1}^{N}\sum_{t=1}^{T_{ij}}\left(\frac{1}{N}\sum_{j=1}^{N}\hat{u^{\star}}_{ijt}\right)^2--\frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}\left(\hat{u^{\star}}_{ijt}\right)^2\right)-$$

$$-\frac{1}{N-1}\sum_{i=1}^{N}\frac{\sum_j T_{ij}}{\left(\sum_j T_{ij}\right)^2}\times$$

$$\times\left(\frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T_{ij}}\left(\hat{u^{\star}}_{ijt}\right)^2-\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\sum_{i=1}^{N}T_{ij}}\sum_{i=1}^{N}\sum_{t=1}^{T_{ij}}\left(\frac{1}{N}\sum_{j=1}^{N}\hat{u^{\star}}_{ijt}\right)^2\right)+$$

$$\left.+\hat{\sigma}_\mu^2\frac{1}{N}\sum_{i=1}^{N}\frac{\sum_j\left(T_{ij}-T_{ij}^2\right)}{\left(\sum_j T_{ij}\right)^2}\right]$$

$$\hat{\rho_{(2)}} = \frac{N}{\sum_{j=1}^{N} \frac{1}{\left(\sum_i T_{ij}\right)^2} \left(\sum_i T_{ij} + 2\sum_{i=1}^{N-1} \sum_{k=i+1}^{N} T_{ij}T_{kj} - \sum_{t=1}^{max_i T_{ij}} \left(\sum_{i=1}^{N} \delta_{ijt}\right)^2\right)} \times$$

$$\times [\frac{1}{N} \sum_{j=1}^{N} \left(\frac{1}{\sum_i T_{ij}} \sum_{i=1}^{N} \sum_{t=1}^{T_{ij}} u^{\star}_{ijt} \right)^2 - \frac{1}{N(N-1)} \sum_{j=1}^{N} \frac{1}{\left(\sum_i T_{ij}\right)^2} \sum_{t=1}^{max_i T_{ij}} \left(\sum_{i=1}^{N} \delta_{ijt} \right)^2 \times$$

$$\times \left(\sum_{i=1}^{N} \frac{1}{\sum_{j=1}^{N} T_{ij}} \sum_{j=1}^{N} \sum_{t=1}^{T_{ij}} \left(\frac{1}{N} \sum_{i=1}^{N} \hat{u^\star}_{ijt} \right)^2 - \frac{1}{T} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t=1}^{T_{ij}} \left(\hat{u^\star}_{ijt}\right)^2 \right) -$$

$$- \frac{1}{N-1} \sum_{j=1}^{N} \frac{\sum_i T_{ij}}{\left(\sum_i T_{ij}\right)^2} \left(\frac{1}{T} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t=1}^{T_{ij}} \left(\hat{u^\star}_{ijt}\right)^2 - \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\sum_{j=1}^{N} T_{ij}} \sum_{j=1}^{N} \sum_{t=1}^{T_{ij}} \left(\frac{1}{N} \sum_{i=1}^{N} \hat{u^\star}_{ijt} \right)^2 \right) +$$

$$+ \hat{\sigma}_{\mu}^2 \frac{1}{N} \sum_{j=1}^{N} \frac{\sum_i \left(T_{ij} - T_{ij}^2\right)}{\left(\sum_i T_{ij}\right)^2}]$$

$$\hat{\sigma}_u^2 = \frac{1}{N-1} \left(\sum_{i=1}^{N} \frac{1}{\sum_{j=1}^{N} T_{ij}} \sum_{j=1}^{N} \sum_{t=1}^{T_{ij}} \left(\frac{1}{N} \sum_{i=1}^{N} \hat{u^\star}_{ijt} \right)^2 - \frac{1}{T} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t=1}^{T_{ij}} \hat{u^\star}^2_{ijt} \right) - \hat{\rho}_{(2)}$$

$$\hat{\sigma}_v^2 = \frac{1}{N-1} \left(\sum_{j=1}^{N} \frac{1}{\sum_{i=1}^{N} T_{ij}} \sum_{i=1}^{N} \sum_{t=1}^{T_{ij}} \left(\frac{1}{N} \sum_{j=1}^{N} \hat{u^\star}_{ijt} \right)^2 - \frac{1}{T} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t=1}^{T_{ij}} \hat{u^\star}^2_{ijt} \right) - \hat{\rho}_{(1)}$$

$$\hat{\sigma^2_\epsilon} = \frac{1}{T} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t=1}^{T_{ij}} \hat{u^\star}^2_{ijt} - \hat{\sigma}_\mu^2 - \hat{\sigma}_u^2 - \hat{\sigma}_v^2$$

Let us note here that the cross-correlations (19) can be generalized without too much trouble, for example, in the following way:

$$E(\mu_{ij}\mu_{ks}) = \begin{cases} \sigma_\mu^2 & \text{if } i = k, j = s \\ \rho_{(1)i} & \text{if } i = k, j \neq s \\ \rho_{(2)j} & \text{if } i \neq k, j = s \\ 0 & \text{otherwise} \end{cases}$$

The covariance matrix of this model and the estimation of the cross-correlations and variance components can be derived in a similar, but much more complicated way, as above.

7. An Application: Modelling Within EU Trade

In order to highlight the differences between the usual fixed effects (FE) and the proposed random effects (RE) approach, let us use a typical empirical trade problem.

In a gravity-like panel estimation exercise we explore the effects of geographical distance and membership in the European Union (EU) on bilateral trade flows. We compare the RE estimates to FE and Pooled OLS estimates.

We take a balanced panel data set of bilateral trade flows for all pairs formed by 20 EU member countries for years 2001-2006. Hence, the total number of country pairs is $N^2 = 400$ and $T = 6$. Twelve of the countries were members of the EU in the whole sample period (group A)[1], the remaining eight entered the EU in 2004 (group B)[2]. Apart from foreign trade, the database also includes self-trade, i.e. trade of a country within its own borders.[3] Self-trade for a given year is generated as gross output minus total exports of a country in that year. All data is in current euros.

A first look at the data suggests that countries in group A trade more with each other than countries in group B, and trade of group B countries increased much faster after 2004 than trade of group A countries (*Table 1*). The first fact can simply reflect that larger, more advanced and more strongly integrated economies trade more. The second may be evidence for the trade creating effect of entering the EU.

We fit a simple gravity-type model that explains bilateral trade with country incomes (GDP), bilateral geographical distance and a dummy for EU membership. For better tractability we restrict the elasticities of trade to income to unity and use income adjusted trade, denoted as y, as dependent variable. We take all variables (except the EU dummy) in logarithms: $y_{ijt} = \ln \text{trade}_{ijt} - \ln \text{income}_{it} - \ln \text{income}_{jt}$, and the explanatory variables are simply $[\ln \text{dist}_{ij}, \text{EU}_{ijt}]$. The dummy for EU membership is 1, if both the exporter and the importer countries are EU members, and 0 otherwise. Formally,

$$\text{EU}_{ijt} = \begin{cases} 1 & (i \in \text{A and } \; j \in \text{A}) \text{ or } \; t \geq 2004 \\ 0 & \text{otherwise} \end{cases}$$

Notice that, while distance is time-invariant, the EU dummy is time-varying, changing from 0 to 1 from 2003 to 2004 for country pairs with at least one type-B country.

The (composite) error term, can take the form of any of the error structures (1) to (7) discussed in the previous sections. Assuming an error structure, we estimate

[1] Group A: Austria, Germany, Denmark, Spain, Finland, France, Greece, Ireland, Italy, Portugal, Sweden, United Kingdom

[2] Group B: Czech Republic, Estonia, Hungary, Lithuania, Latvia, Poland, Slovenia, Slovakia

[3] We include self-trade to avoid bias of the Fixed Effects estimates. As it is stressed in *Hornok* [2011] and *Mátyás and Balázsi* [2011], the Fixed Effects within transformation formulas for some of the error structures considered here give biased estimates if self-trade is not included in the database.

the model using FGLS. Estimated coefficients for each error structure (models (1) to (7)) are reported in *Table 2*. We report Pooled OLS, FE and RE estimates, as well as the estimated variances for the error components. Pooled OLS estimates are identical for each model. The parameter estimates for the distance coefficient are very stable across all models and methods as this coefficient is always identified from variation in the country pair dimension. As a consequence, in the FE models with country pair fixed effects (models (1), (3) and (7)) the distance coefficient is not identified. This highlights one important drawback of the FE approach: due to the large number of fixed effects related dummies, other important dummy-like variable often are not identified.

In contrast, the coefficient for the EU dummy is always identified, but its estimates vary considerably across models. In models (1), (3) and (7) it is identified mostly (in the case of FE, only) from the time dimension, i.e., from the change in trade of type-B countries. In the other three models identification is based more on the cross-sectional dimension, i.e., comparing EU pairs to non-EU pairs before 2004. Apart from that, in this case, the RE parameter estimates happen to be quite close to the FE estimates (except for model (6)).

As a next step we change the estimating equation so that the EU dummy is broken up into three separate dummies. One for pairs of two type-A countries, one for pairs of two type-B countries and one for pairs with one type-A and one type-B country:

$$\mathrm{EU}_{AA} = \begin{cases} 1 & \mathrm{EU} = 1 \;\text{ and }\; i \in A \;\text{ and }\; j \in A \\ 0 & \text{otherwise} \end{cases}$$

$$\mathrm{EU}_{BB} = \begin{cases} 1 & \mathrm{EU} = 1 \;\text{ and }\; i \in B \;\text{ and }\; j \in B \\ 0 & \text{otherwise} \end{cases}$$

$$\mathrm{EU}_{AB} = \begin{cases} 1 & \mathrm{EU} = 1 \;\text{ and }\; \mathrm{EU}_{AA} \neq 1 \text{ and } \mathrm{EU}_{BB} \neq 1 \\ 0 & \text{otherwise} \end{cases}$$

This modification enables us to identify separate effects of the EU on the different groups of country pairs. Besides, it helps to bring to light again the already mentioned important disadvantage of the FE approach. When the within transformation nets out fixed effects in it and jt (and ij) dimensions, identification of other regressors (especially dummy variables) may not be possible, even if these regressors vary in the ijt dimension. This is due to perfect collinearity among the fixed effects and the regressors.[4] As reported in *Table 3*, the FE estimator is not able to identify the

[4] This identification problem is also addressed in *Hornok* [2011].

coefficients of EU_{BB} and EU_{AB} separately under models (6) and (7). In contrast, the RE method identifies all coefficients and reveals that there are indeed large and significant differences among the effects of EU on different groups of country pairs. These differences are in line with the raw data evidence in *Table 1.* The coefficient estimates for EU_{BB} are significantly larger than those for EU_{AB}, which are larger than the coefficient estimates for EU_{AA}. The estimates are in some cases negative, depending on whether the model identifies them mostly from the time series or from the cross section dimension. In an empirical research, this has of course important implications on which model to choose and how to interpret the estimates, an issue we do not deal with here.

Finally, let us note that the introduction of cross-correlations (see *Table 4*) does not change substantially the value and significance of the parameters of interest in this application, although the estimates of the cross-correlation coefficients are non-negligible in magnitude.

8. Conclusion

In this paper we presented an alternative random effects approach to the usual fixed effects gravity models of trade, in a three-dimensional panel data setup. We showed that the random effects and fixed effects specifications, just like in the usual panel data cases, may lead to substantially different parameter estimates and inference, although in both cases the corresponding estimators are in fact consistent.

At the end of the day, the main question for an applied researcher, as in any panel data setup, is whether to use a fixed effects or random effects specification. In three (or multi-) dimensional models the fixed effects specification (due to the very large number of dummies to estimate) will result in a massive over-specification, which implies that much less data information will be available for the estimation of the main/focus variables. Also, again due to the fixed effects dummy variables, frequently other (say, for example policy-type, potentially important) dummy variables cannot be identified. On the other hand, in a random effects specification the data is not "burdened" by the massive estimation of the fixed effects parameters. In addition any reasonable covariance structure can be imposed on the disturbance terms, still the model can be estimated without too much trouble. The down side is, of course, that one has to keep an eye on the endogeneity problem. The choice unfortunately not obvious.

References

James E. Anderson [2010]: The Gravity Model, *NBER Working Paper* 16576, (http://www.nber.org/papers/w16576).

Scott L. Baier and Jeffrey H. Bergstrand [2007]: Do Free Trade Agreements Actually Increase Members International Trade? *Journal of International Economics*, 71, 72-95.

Badi H. Baltagi, Peter Egger and Michael Pfaffermayr [2003]: A generalized Design for Bilateral Trade Flow Models, *Economic Letters*, 80, pp. 391-397.

Badi H. Baltagi, Laszlo Matyas and Patrick Sevestre [2008]: *Error Components Models*, in Matyas and Sevetre (eds.) The Econometrics of Panel Data, Springer Verlag, 2008, pp. 49-88.

Peter A. G. van Bergeijk and Steven Brakman, Rijksuniversiteit Groningen(eds) [2010]: *The Gravity Model in International Trade; Advances and Applications*, Cambridge University Press.

Peter Davis [2002]: Estimating multi-way error components models with unbalanced data structures, *Journal of Econometrics*, 106, 67-95.

Richard Baldwin and Daria Taglioni [2006]: Gravity for Dummies and Dummies for the Gravity Equations, NBER Working Paper 12516.

Peter Egger and Michael Pfaffermayr [2003]: The Proper Econometric Specification of the Gravity Equation: 3-Way Model with Bilateral Interaction Effects, *Empirical Economics*, 28, 571-580.

Cecilia Hornok [2011]: Gravity or Dummies? The Limits of Identification in Gravity Estimations, *CeFiG Working Papers* No. 15. (September 2011)

Laszlo Matyas and Laszlo Balazsi [2011]: The Estimation of Three-dimensional Fixed Effects Panel Data Models, *CEU Economics Working Papers*, 2011/12

Radhakrishna Rao and Jurgen Kleffe [1980]: Estimation of Variance Components, in P.R. Krishnaiah (ed.) *Handbook of Statistics 1*, North-Holland Publisher, pp. 1-40.

Table 1: Trade of EU countries before and after 2004

	2001-2003	2004-2006	% change
Foreign trade[1]			
A with A	7,659	8,561	11.8
A with B	1,064	1,471	38.3
B with B	379	712	87.9
Self-trade[2]			
A	268,348	285,006	6.2
B	29,102	32,803	12.7

Notes: Source of trade data (in millions of euros) is Eurostat. Self-trade is authors' calculation based on Eurostat and OECD data. [1] Average of annual pair-specific flows within group. [2] Average of annual country self-trade within group.

Table 2: Comparison of estimators

	(1)	(3)	(4)	(5)	(6)	(7)
Pooled OLS						
EU			0.190631			
			(0.036527)			
ln dist			-1.645467			
			(0.019266)			
Fixed Effects						
EU	-0.030791	0.126275	-0.064553	-0.257496	0.593089	0.006564
	(0.011727)	(0.018922)	(0.054318)	(0.057025)	(0.072609)	(0.033426)
ln dist	–	–	-1.681144	-1.609525	-1.608665	–
	–	–	(0.017238)	(0.018097)	(0.015311)	–
Random Effects						
EU	-0.034560	0.110319	-0.114367	-0.250177	0.262012	-0.003155
	(0.011698)	(0.018636)	(0.048620)	(0.048929)	(0.059702)	(0.032281)
ln dist	-1.647862	-1.650085	-1.677159	-1.616094	-1.627702	-1.645149
	(0.045509)	(0.045181)	(0.017131)	(0.017937)	(0.014976)	(0.035253)
Variance Components						
σ^2_ε	0.052808	0.049494	0.508946	0.560954	0.345769	0.040670
σ^2_μ	0.639168	0.630406				0.342357
σ^2_λ		0.012077				
σ^2_u			0.183030		0.179277	0.179277
σ^2_v				0.131022	0.166930	0.129673

Notes: Dependent variable is log of income-adjusted bilateral trade. Standard errors in parenthesis. Estimation on a balanced panel of pairs of 20 EU countries for years 2001-2006.

Table 3: Comparison of estimators with three EU dummies

	(1)	(3)	(4)	(5)	(6)	(7)
Pooled OLS						
EU_{AA}			-0.425864			
			(0.039215)			
EU_{BB}			0.628679			
			(0.065105)			
EU_{AB}			-0.250615			
			(0.043757)			
ln dist			-1.572309			
			(0.018913)			
Fixed Effects						
EU_{AA}	–	–	-0.058001	-0.250686	0.626499	–
	–	–	(0.052119)	(0.054984)	(0.071516)	–
EU_{BB}	0.041035	0.198101	1.078016	0.843686	–	–
	(0.023386)	(0.027204)	(0.098732)	(0.104160)	–	–
EU_{AB}	-0.054733	0.102333	0.228817	-0.024152	–	–
	(0.013502)	(0.019963)	(0.073639)	(0.077688)	–	–
ln dist	–	–	-1.631179	-1.557583	-1.577510	–
	–	–	(0.017071)	(0.018009)	(0.015482)	–
Random Effects						
EU_{AA}	-0.420115	-0.342579	-0.176215	-0.328905	0.216663	-0.090768
	(0.079668)	(0.078380)	(0.046842)	(0.047293)	(0.058454)	(0.086442)
EU_{BB}	0.057348	0.212035	0.870278	0.716491	1.002332	0.428087
	(0.023273)	(0.027062)	(0.083976)	(0.083104)	(0.135971)	(0.133660)
EU_{AB}	-0.060171	0.095017	0.041347	-0.141667	0.295330	0.202146
	(0.013480)	(0.019876)	(0.061452)	(0.060241)	(0.082291)	(0.068057)
ln dist	-1.627487	-1.627523	-1.625481	-1.562311	-1.596464	-1.637796
	(0.043166)	(0.042274)	(0.016964)	(0.017839)	(0.015113)	(0.035425)
Variance Components						
σ^2_ε	0.052504	0.049187	0.468526	0.521461	0.334464	0.040692
σ^2_μ	0.571179	0.548059				0.343107
σ^2_λ		0.026436				
σ^2_u			0.155157		0.140710	0.140710
σ^2_v				0.102222	0.148509	0.099174

Notes: Dependent variable is log of income-adjusted bilateral trade. Standard errors in parenthesis. Estimation on a balanced panel of pairs of 20 EU countries for years 2001-2006.

Table 4: Random effects with cross correlation

	(1)	(3)	(7)
Random Effects			
EU	-0.029134	0.124569	-0.006853
	(0.011710)	(0.018686)	(0.022627)
ln dist	-1.638312	-1.632182	-1.640335
	(0.035809)	(0.036543)	(0.036307)
Variance Components			
σ_{ε}^2	0.052808	0.049494	0.040670
σ_{μ}^2	0.639168	0.630406	0.630406
σ_{λ}^2		0.012077	
σ_{u}^2			0.004325
σ_{v}^2			0.016576
ρ_1	0.179234	0.167157	0.113097
ρ_2	0.129630	0.117553	0.174952

Notes: Dependent variable is log of income-adjusted bilateral trade. Standard errors in parenthesis. Estimation on a balanced panel of pairs of 20 EU countries for years 2001-2006.