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

A quiet data revolution has been taking place in economics the last decade or so, which has been profoundly affecting the way applied economics and economic analysis is carried out. This also had a direct influence on econometrics, as new types of data frequently require new econometric tools. One such area, where new large data sets have been emerging, relates to trade where the availability and use of multidimensional panel data sets has received momentum the last few years. Especially, three dimensional data bases are becoming very popular, for a recent reviews of the subject see, for example, *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, but observable, 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, where it is highlighted how the this proposed random effects approach can lead to different inference and model results.

2. Different Heterogeneity Formulations

More than two decade ago *Moulton* [1990] draw the attention to the fact that when dealing with disaggregated observations to answer macro type questions the covariance structure of the model is of paramount importance. When no repeated observations are available (i.e., when not working with panel data) heavy assumptions are needed to get treatable models (see, for example, *Gelman* [2006], amongst others). On the other hand, when working with panel data sets we can take advantage of these repeated observations, and generalize them to this higher dimensional setup, in fact deriving higher dimensional random effects panel data models.

Staring from the most widely used fixed effects model specifications, which have been proposed by *Baltagi et al.* [2003], *Egger and Pfanffermayr* [2003], *Baldwin and Taglioni* [2006], and *Baier and Bergstrand* [2007], and highlighted the importance of different types of interactions effects, we propose below similar, but random interactions amongst the individual and time specific factors. Let us start with simple straightforward direct generalization of the standard panel data model

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

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

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

This model can be considered in fact as a straight panel data model, where the individuals are now the (ij) pairs, so essentially it does not take into account the three-dimensional nature of the data. Let us note, that if the μ_{ij} interaction effects were not random, but fixed, as it has been the practice in most applied studies, this would mean the explicit or implicit estimation of N^2 parameters. For example, in the case of a macro trade model, with say 150 countries involved, we are talking about $150 \times 150 = 22,500$ parameters. This looks very much like a textbook over-specification case. Moreover, none of these empirical studies (see, for example, *Silva and Nelson* [2012], *Cheng and Wall* [2005] or *Karemera, Oguledo and Davis* [2000], amongst large number of others) were bothered with testing the significance of these parameters. If most of these parameters turned out to be non-significant, it would imply that a random effects specification may be more suited for the given data. But then, from an empirical point of view, it is crucial to know whether the explanatory variables and the disturbance terms of the model are uncorrelated, otherwise we may end up with inconsistent parameter estimates. The most widely know test for this is the Hausman's specification test, which relies on both the fixed effects and the above random effects model estimation (see more about this problem in our context in *Clark and Linzer* [2012]). So it is fair to say that for empirical analysis it is of paramount importance to be able to properly estimate the random effects models.

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. This in fact is the generalization of the approach used in multilevel modeling, see for example, *Ebbes, Bockenholt and Wedel* [2004] or *Hubler* [2006]. 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 Feasible 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

First, we need to derive the covariance matrix of each of the models introduced in Section 2, then the unknown variance components of these matrices have to be estimated.

For model (1) let us denote

$$u^\star_{ijt} = \mu_{ij} + \epsilon_{ijt} \tag{8}$$

So for all t observations

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

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

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

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

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

Using matrix notation and decomposing the covariance matrix Ω

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

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)

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

and

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

so we obtain

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

Now turning to the inverse of the covariance matrix

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

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

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

with

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

This is amounts to

$$
\sigma_\epsilon^2 \Omega^{-1} = I_{N^2T} - (1-\theta_2) I_{N^2} \otimes \frac{J_T}{T} - (1-\theta_3) \frac{J_{N^2}}{N^2} \otimes I_T + (1-\theta_2-\theta_3+\theta_1) \frac{J_{N^2T}}{N^2T}
$$

Thus, the FGLS estimator is equivalent to the OLS estimator on the transformed variables like

$$
\begin{aligned}
\tilde{y} = \Bigg(& y_{ijt} - (1-\theta_2) \sum_t \frac{1}{T} y_{ijt} - (1-\theta_3) \sum_i \sum_j \frac{1}{N^2} y_{ijt} + \\
& + (1-\theta_2-\theta_3+\theta_1) \sum_i \sum_j \sum_t \frac{1}{N^2T} y_{ijt} \Bigg)
\end{aligned}
$$

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

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

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

$$\begin{aligned}
u^{\star}_{i} &= u + \epsilon_{i} \\
E\left(u^{\star}_{i} {u^{\star}}'_{i}\right) &= E\left[u u'\right] + E\left[\epsilon_{i}\epsilon'_{i}\right] = \sigma^2_u I_{NT} + \sigma^2_\epsilon I_{NT} \\
u^{\star} &= l_N \otimes u + \epsilon \\
E\left(u^{\star} {u^{\star}}'\right) &= E\left[\left(l_N \otimes u\right)\left(l'_N \otimes u'\right)\right] + E\left[\epsilon\epsilon'\right] = \sigma^2_u J_N \otimes I_{NT} + \sigma^2_\epsilon I_{N^2T} = \Omega
\end{aligned}$$

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

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

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

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

which is equivalent to

$$\sigma^2_\epsilon \Omega^{-1} = I_{N^2T} - (1-\theta)\frac{J_N}{N} \otimes I_{NT}$$

and the FGLS estimator is OLS estimator on $\tilde{y}_{ijt} = \left(y_{ijt} - (1-\theta)\sum_i \frac{1}{N} y_{ijt}\right)$ type transformed variables. 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

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

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

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

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

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

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

with

$$\theta_1 = \frac{\sigma_\epsilon^2}{N\sigma_v^2 + N\sigma_u^2 + \sigma_\epsilon^2}$$

$$\theta_2 = \frac{\sigma_\epsilon^2}{N\sigma_v^2 + \sigma_\epsilon^2}$$

$$\theta_3 = \frac{\sigma_\epsilon^2}{N\sigma_u^2 + \sigma_\epsilon^2}$$

This is equivalent to

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

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 be 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

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

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

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

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

rank : $((\mathrm{N}-1)^2(\mathrm{T}-1)+1)$

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

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

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

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

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

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

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

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

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

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

Continuing with model (3)

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

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

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

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

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

and

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

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

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

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

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

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

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

Now for model (6) the Within transformation is

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

so we get

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

And, also,

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

The estimators of the variance components therefore are

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

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

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

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

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

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

$$\begin{aligned}
\hat{\sigma}^2_\epsilon &= \frac{N^2T}{(N-1)^2(T-1)}\hat{u}^{\star'}_{within}\hat{u}^{\star}_{within}\\
\hat{\sigma}^2_v &= \frac{1}{N^2T(N-1)}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\left(\sum_{j=1}^{N}\hat{u}^{\star}\right)^2 - \sum_{j=1}^{N}\hat{u}^{\star 2}\right)\right)\\
\hat{\sigma}^2_u &= \frac{1}{N^2T(N-1)}\left(\sum_{j=1}^{N}\sum_{t=1}^{T}\left(\left(\sum_{i=1}^{N}\hat{u}^{\star}\right)^2 - \sum_{i=1}^{N}\hat{u}^{\star 2}\right)\right)\\
\hat{\sigma}^2_\mu &= \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\hat{u}^{\star 2}_{ijt} - \hat{\sigma}^2_\epsilon - \hat{\sigma}^2_v - \hat{\sigma}^2_u
\end{aligned} \tag{18}$$

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, for all models considered, in general, let T_{ij} be the number of observations available for a country pair (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. As the structure of the data now is quite complex, we need to introduce some new notations and definitions.

Let us define
Z_{ij} as the set of time periods when country i exports to j, with T_{ij} being the number of elements in Z_{ij}, $Z_{ij} = \left\{ z_{ij}^{1}, ..., z_{ij}^{T_{ij}} \right\}$;
Z_i^{**} as the set of time periods when i export to any country, with T_i^{**} being the number of elements in Z_i^{**};
Z_j^{*} as the set of time periods in which j gets export from anywhere, with T_j^{*} being the number of elements in Z_j^{*};
$Q_{jt}^{(1)}$ as the set of countries that export to country j at period of time t, with $N_{jt}^{(1)}$ being the number of elements in Q_{jt};
$Q_{it}^{(2)}$ as the set of countries to which i exports at period t, with $N_{it}^{(2)}$ being the number of elements in $Q_{it}^{(2)}$;
$Q_i^{(2)}$ as the set of countries i exports to at least once, with $N_i^{(2)}$ being the number of elements in $Q_i^{(2)}$;
$Q^{(1)}$ as the set of countries that export, with $N^{(1)}$ being number of elements in $Q^{(1)}$;
$Q^{(2)}$ as the set of countries that get export, with $N^{(2)}$ being number of elements in $Q^{(2)}$.
And, also,

$$Si = \frac{1}{1 - \frac{1}{N^{(2)}} \sum_j \frac{1}{T_j^*} \sum_t \frac{1}{N_{jt}^{(1)}}} \left(\frac{1}{N^{(2)}} \sum_{j \in Q^{(2)}} \frac{1}{T_j^*} \sum_{t \in Z_j^*} \left(\frac{1}{N_{jt}^{(1)}} \sum_{i \in Q_{jt}^{(1)}} \hat{u}_{ijt}^{*} \right)^2 - \right.$$
$$\left. - \frac{1}{N^{(2)}T} \sum_{j \in Q^{(2)}} \frac{1}{T_j^*} \sum_{t \in Z_j^*} \frac{1}{N_{jt}^{(1)}} \sum_{i \in Q^{(1)}} \sum_{j \in Q_i^{(2)}} \sum_{t \in Z_{ij}} \hat{u}_{ijt}^{*^2} \right)$$

$$Sj = \frac{1}{1 - \frac{1}{N^{(1)}} \sum_i \frac{1}{T_i^{**}} \sum_t \frac{1}{N_{it}^{(2)}}} \left(\frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{T_i^{**}} \sum_{t \in Z_i^{**}} \left(\frac{1}{N_{it}^{(2)}} \sum_{j \in Q_{it}^{(2)}} \hat{u}_{ijt}^{*} \right)^2 - \right.$$
$$\left. - \frac{1}{N^{(1)}T} \sum_{i \in Q^{(1)}} \frac{1}{T_i^{**}} \sum_{t \in Z_i^{**}} \frac{1}{N_{it}^{(2)}} \sum_{i \in Q^{(1)}} \sum_{j \in Q_i^{(2)}} \sum_{t \in Z_{ij}} \hat{u}_{ijt}^{*^2} \right)$$

$$St = \frac{1}{1 - \frac{1}{N^{(1)}} \sum_i \frac{1}{N_i^{(2)}} \sum_j \frac{1}{T_{ij}}} \Bigg(\frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{N_i^{(2)}} \sum_{j \in Q_i^{(2)}} \left(\frac{1}{T_{ij}} \sum_{t \in Z_{ij}} u^*_{ijt} \right)^2 -$$

$$- \frac{1}{N^{(1)} T} \sum_{i \in Q^{(1)}} \frac{1}{N_i^{(2)}} \sum_{j \in Q_i^{(2)}} \frac{1}{T_{ij}} \sum_{i \in Q^{(1)}} \sum_{j \in Q_i^{(2)}} \sum_{t \in Z_{ij}} \hat{u}^{*^2}_{ijt} \Bigg)$$

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 the elements of the s-th row and p-th column of matrix $M_{T_{ij}\times T_{lk}}$ (of size $(T_{ij} \times T_{lk})$) are defined as

$$\{m\}_{sp} = \begin{cases} 1 & \text{if } z_{ij}^s \neq z_{lk}^p \\ 0 & \text{otherwise} \end{cases}$$

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

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

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)

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

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

Now that we have 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) the identifying equations now are

$$
\begin{aligned}
& E\left[u^{*^2}_{ijt}\right] = \sigma^2_\mu + \sigma^2_\epsilon \\
& E\left[\frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{N^{(2)}_i} \sum_{j \in Q^{(2)}_i} \left(\frac{1}{T_{ij}} \sum_{t \in Z_{ij}} u^*_{ijt}\right)^2\right] = \frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{N^{(2)}_i} \sum_{j \in Q^{(2)}_i} \frac{1}{T_{ij}} \sigma^2_\epsilon + \sigma^2_\mu
\end{aligned}
$$

and so for the variance components we get the following estimators

$$
\begin{aligned}
\hat{\sigma}^2_\mu &= St \\
\hat{\sigma}^2_\epsilon &= \frac{1}{T} \sum_{i \in Q^{(1)}} \sum_{j \in Q^{(2)}_i} \sum_{t \in Z_{ij}} \hat{u}^{*^2}_{ijt} - \hat{\sigma}^2_\mu
\end{aligned}
$$

For model (3) we have

$$E\left[u_{ijt}^{*^2}\right]=\sigma_\mu^2+\sigma_\lambda^2+\sigma_\epsilon^2$$

$$E\left[\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\left(\frac{1}{N_{jt}^{(1)}}\sum_{i\in Q_{jt}^{(1)}}u_{ijt}^*\right)^2\right]=\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\frac{1}{N_{jt}^{(1)}}\left(\sigma_\mu^2+\sigma_\epsilon^2\right)+\sigma_\lambda^2$$

$$E\left[\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{N_i^{(2)}}\sum_{j\in Q_i^{(2)}}\left(\frac{1}{T_{ij}}\sum_{t\in Z_{ij}}u_{ijt}^*\right)^2\right]=\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{N_i^{(2)}}\sum_{j\in Q_i^{(2)}}\frac{1}{T_{ij}}\left(\sigma_\lambda^2+\sigma_\epsilon^2\right)+\sigma_\mu^2$$

and therefore for the variance components we get the following estimators

$$\begin{aligned}\hat\sigma_\mu^2&=St\\ \hat\sigma_\lambda^2&=Si\\ \hat\sigma_\epsilon^2&=\frac{1}{T}\sum_{i\in Q^{(1)}}\sum_{j\in Q_i^{(2)}}\sum_{t\in Z_{ij}}\hat u_{ijt}^{*^2}-\hat\sigma_\mu^2-\hat\sigma_\lambda^2\end{aligned}$$

For model (4) (and similarly for model (5)) we get

$$E\left[u_{ijt}^{*^2}\right]=\sigma_u^2+\sigma_\epsilon^2$$

$$E\left[\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\left(\frac{1}{N_{jt}^{(1)}}\sum_{i\in Q_{jt}^{(1)}}u_{ijt}^*\right)^2\right]=\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\frac{1}{N_{jt}^{(1)}}\sigma_\epsilon^2+\sigma_u^2$$

And

$$\begin{aligned}\hat\sigma_u^2&=Si\\ \hat\sigma_\epsilon^2&=\frac{1}{T}\sum_{i\in Q^{(1)}}\sum_{j\in Q_i^{(2)}}\sum_{t\in Z_{ij}}\hat u_{ijt}^{*^2}-\hat\sigma_u^2\end{aligned}$$

Turning now to model (6)

$$E\left[u_{ijt}^{*^2}\right]=\sigma_u^2+\sigma_v^2+\sigma_\epsilon^2$$

$$E\left[\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\left(\frac{1}{N_{jt}^{(1)}}\sum_{i\in Q_{jt}^{(1)}}u_{ijt}^*\right)^2\right]=\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\frac{1}{N_{jt}^{(1)}}\left(\sigma_v^2+\sigma_\epsilon^2\right)+\sigma_u^2$$

$$E\left[\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\left(\frac{1}{N_{it}^{(2)}}\sum_{j\in Q_{it}^{(2)}}u_{ijt}^*\right)^2\right]=\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\frac{1}{N_{it}^{(2)}}\left(\sigma_u^2+\sigma_\epsilon^2\right)+\sigma_v^2$$

and the estimators of variance components are

$$\hat{\sigma}_u^2 = Si$$
$$\hat{\sigma}_v^2 = Sj$$
$$\hat{\sigma}_\epsilon^2 = \frac{1}{T} \sum_{i \in Q^{(1)}} \sum_{j \in Q_i^{(2)}} \sum_{t \in Z_{ij}} \hat{u}_{ijt}^{*^2} - \hat{\sigma}_u^2 - \hat{\sigma}_v^2$$

And finally, for model (7) we get

$$E\left[u_{ijt}^{*^2}\right] = \sigma_\mu^2 + \sigma_u^2 + \sigma_v^2 + \sigma_\epsilon^2$$

$$E\left[\frac{1}{N^{(2)}} \sum_{j \in Q^{(2)}} \frac{1}{T_j^*} \sum_{t \in Z_j^*} \left(\frac{1}{N_{jt}^{(1)}} \sum_{i \in Q_{jt}^{(1)}} u_{ijt}^*\right)^2\right] =$$
$$= \frac{1}{N^{(2)}} \sum_{j \in Q^{(2)}} \frac{1}{T_j^*} \sum_{t \in Z_j^*} \frac{1}{N_{jt}^{(1)}} \left(\sigma_\mu^2 + \sigma_v^2 + \sigma_\epsilon^2\right) + \sigma_u^2$$

$$E\left[\frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{T_i^{**}} \sum_{t \in Z_i^{**}} \left(\frac{1}{N_{it}^{(2)}} \sum_{j \in Q_{it}^{(2)}} u_{ijt}^*\right)^2\right] =$$
$$= \frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{T_i^{**}} \sum_{t \in Z_i^{**}} \frac{1}{N_{it}^{(2)}} \left(\sigma_\mu^2 + \sigma_u^2 + \sigma_\epsilon^2\right) + \sigma_v^2$$

$$E\left[\frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{N_i^{(2)}} \sum_{j \in Q_i^{(2)}} \left(\frac{1}{T_{ij}} \sum_{t \in Z_{ij}} u_{ijt}^*\right)^2\right] =$$
$$= \frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{N_i^{(2)}} \sum_{j \in Q_i^{(2)}} \frac{1}{T_{ij}} \left(\sigma_u^2 + \sigma_v^2 + \sigma_\epsilon^2\right) + \sigma_\mu^2$$

And so the estimators of the variance components are

$$\hat{\sigma}_u^2 = Si$$
$$\hat{\sigma}_v^2 = Sj$$
$$\hat{\sigma}_\mu^2 = St$$
$$\hat{\sigma}_\epsilon^2 = \frac{1}{T} \sum_{i \in Q^{(1)}} \sum_{j \in Q_i^{(2)}} \sum_{t \in Z_{ij}} \hat{u}_{ijt}^{*^2} - \hat{\sigma}_\mu^2 - \hat{\sigma}_u^2 - \hat{\sigma}_v^2$$

Now, let us have a look at another potential data problem important in the case of trade, as by nature, usually, 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. Let us note that in many applications this problem is circumvent by substituting the missing observations by "domestic consumption" or similar variables, which is in fact highly inappropriate. In the unbalanced data case, as seen above, this has been dealt with in a natural way.

On the other hand, the no-self-flow case has surprisingly little effect vis-a-vis to what has been said so far, in balanced data. 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). Let us note here that the no-self-flow case can be extended without to much difficulty to the case when $i = 1, \ldots, N_{(1)}$, $j = 1, \ldots, N_{(2)}$ and $N_{(1)} \neq N_{(2)}$, with very similar results to those obtained above.

We must mention here that, as no easy "ready to use" decompositions of the covariances matrixes in the unbalanced case are available, "brute force" should be used to implement the FGLS estimator, i.e., to invert the covariance matrixes involved. For very large data set (usually when the number of observations is larger than 10^6) we must step out of the comfort zone of the usual statistics/econometrics software packages. There are, however, several methods and algorithms available to make this task manageable (see, for example, *Healy* [1968], *Bientinesi, Gunter and van de Geijn* [2000] or *van de Geijn* [2008]).

6. An Extension

In the case of trade models (and in general in fact 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 between the effects. 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

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

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

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

So we get

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

For model (3)

$$
\begin{aligned}
E\left[\left(\frac{1}{N}\sum_{i=1}^{N} u_{ijt}^{\star}\right)^2\right] &= \frac{1}{N}\sigma_\mu^2 + \sigma_\lambda^2 + \frac{1}{N}\sigma_\epsilon^2 + \frac{N-1}{N}\rho_{(2)} \\
E\left[\left(\frac{1}{N}\sum_{j=1}^{N} u_{ijt}^{\star}\right)^2\right] &= \frac{1}{N}\sigma_\mu^2 + \sigma_\lambda^2 + \frac{1}{N}\sigma_\epsilon^2 + \frac{N-1}{N}\rho_{(1)}
\end{aligned}
$$

and so

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

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

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

$$E\left[\left(\frac{1}{N}\sum_{i=1}^{N} u_{ijt}^{\star}\right)^2\right] = \frac{1}{N}\sigma_\mu^2 + \frac{1}{N}\sigma_v^2 + \sigma_u^2 + \frac{1}{N}\sigma_\epsilon^2 + \frac{N-1}{N}\rho_{(2)}$$

$$E\left[\left(\frac{1}{N}\sum_{j=1}^{N} u_{ijt}^{\star}\right)^2\right] = \frac{1}{N}\sigma_\mu^2 + \sigma_v^2 + \frac{1}{N}\sigma_u^2 + \frac{1}{N}\sigma_\epsilon^2 + \frac{N-1}{N}\rho_{(1)}$$

$$E\left[\left(\frac{1}{T}\sum_{t=1}^{T} u_{ijt}^{\star}\right)^2\right] = \sigma_\mu^2 + \frac{1}{T}\sigma_v^2 + \frac{1}{T}\sigma_u^2 + \frac{1}{T}\sigma_\epsilon^2$$

$$E\left[\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} u_{ijt}^{\star}\right)^2\right] = \frac{1}{N}\sigma_\mu^2 + \frac{1}{NT}\sigma_v^2 + \frac{1}{T}\sigma_u^2 + \frac{1}{NT}\sigma_\epsilon^2 + \frac{N-1}{N}\rho_{(2)}$$

And by solving this system we get

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

$$\hat\rho_2 = \frac{1}{N^2(N-1)T(T-1)}\times$$
$$\times\sum_{j=1}^{N}\left(\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat u_{ijt}^{\star}\right)^2 - \sum_{i=1}^{N}\left(\sum_{t=1}^{T}\hat u_{ijt}^{\star}\right)^2 - \sum_{t=1}^{T}\left(\sum_{i=1}^{N}\hat u_{ijt}^{\star}\right)^2 + \sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat u_{ijt}^{\star}\right)^2\right)$$

$$\hat\sigma_u^2 = \frac{1}{N^2(N-1)T}\sum_{j=1}^{N}\sum_{t=1}^{T}\left(\left(\sum_{i=1}^{N}\hat u_{ijt}^{\star}\right)^2 - \sum_{i=1}^{N}\left(\hat u_{ijt}^{\star}\right)^2\right) - \hat\rho_2$$

$$\hat\sigma_v^2 = \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\left(\hat u_{ijt}^{\star}\right)^2 - \hat\sigma_\mu^2 - \hat\sigma_u^2 - \hat\sigma_\epsilon^2$$

$$\hat\rho_1 = \frac{1}{N^2(N-1)T}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\left(\sum_{j=1}^{N}\hat u_{ijt}^{\star}\right)^2 - \sum_{j=1}^{N}\left(\hat u_{ijt}^{\star}\right)^2\right) - \hat\sigma_v^2$$

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

$$
\begin{aligned}
E[u^\star_{ij} u^{\star\prime}_{ij}] &= \sigma^2_\mu J_{T_{ij}} + \sigma^2_\epsilon I_{T_{ij}} \\
E[u^\star_{i} u^{\star\prime}_{i}] &= \sigma^2_\mu A_i + \rho_{(1)} \left(J_{\sum_j T_{ij}} - A_i \right) + \sigma^2_\epsilon I_{\sum_j T_{ij}} \\
E[u^\star u^{\star\prime}] &= \sigma^2_\mu B + \rho_{(1)} \left(P - B \right) + \rho_{(2)} S + \sigma^2_\epsilon I_T
\end{aligned}
$$

$$
S = \begin{pmatrix} O_1 & R_{12} & \dots & R_{1N} \\ R_{21} & O_2 & \dots & R_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ R_{N1} & R_{N2} & \dots & O_N \end{pmatrix}
$$

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

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

and that of model (7) is

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

Let us turn now our attention again 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 unbalanced case, however, cross-correlations need to be estimated in a different way. The identification equations now are

$$E\left[\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\left(\frac{1}{N_{jt}^{(1)}}\sum_{i\in Q_{jt}^{(1)}}u_{ijt}^*\right)^2\right]=$$

$$=\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\frac{1}{N_{jt}^{(1)}}\left(\sigma_\mu^2+\sigma_\epsilon^2\right)+\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\frac{N_{jt}^{(1)}-1}{N_{jt}^{(1)}}\rho_{(2)}$$

$$E\left[\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\left(\frac{1}{N_{it}^{(2)}}\sum_{j\in Q_{it}^{(2)}}u_{ijt}^*\right)^2\right]=$$

$$=\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\frac{1}{N_{it}^{(2)}}\left(\sigma_\mu^2+\sigma_\epsilon^2\right)+\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\frac{N_{it}^{(2)}-1}{N_{it}^{(2)}}\rho_{(1)}$$

And

$$\hat{\rho}_{(1)}=\frac{1}{\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\frac{N_{it}^{(2)}-1}{N_{it}^{(2)}}}\left(\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\left(\frac{1}{N_{it}^{(2)}}\sum_{j\in Q_{it}^{(2)}}\hat{u}_{ijt}^*\right)^2-\right.$$
$$\left.-\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\frac{1}{N_{it}^{(2)}}\left(\sigma_\mu^2+\sigma_\epsilon^2\right)\right)$$

$$\hat{\rho}_{(2)}=\frac{1}{\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\frac{N_{jt}^{(1)}-1}{N_{jt}^{(1)}}}\left(\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\left(\frac{1}{N_{jt}^{(1)}}\sum_{i\in Q_{jt}^{(1)}}\hat{u}_{ijt}^*\right)^2-\right.$$
$$\left.-\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\frac{1}{N_{jt}^{(1)}}\left(\hat{\sigma}_\mu^2+\hat{\sigma}_\epsilon^2\right)\right)$$

In case of model (3) the identification equation become more complicated. The estimation of σ_μ^2 remains the same as in the unbalanced case, however, we need to estimate σ_λ^2, σ_ϵ^2 and the cross-correlations. We also need to introduce some additional definitions. Let

$Q_t^{(1)}$ be the set of countries that export at time t at least to one country, with $N_t^{(1)}$ the number of elements in the set;

$Q_t^{(2)}$ be the set of countries that get export at time t at least from one country, with $N_t^{(2)}$ the number of elements in the set; and

Z_O be the set of time periods when export occurred from at least one country, with T_O the number of elements in the set.
The identifying equations now are

$$E\left[u_{ijt}^{*^2}\right] = \sigma_\mu^2 + \sigma_\lambda^2 + \sigma_\epsilon^2$$

$$E\left[\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\left(\frac{1}{N_{jt}^{(1)}}\sum_{i\in Q_{jt}^{(1)}}u_{ijt}^*\right)^2\right] = \frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\frac{1}{N_{jt}^{(1)}}\left(\sigma_\mu^2+\sigma_\epsilon^2\right)+$$
$$+\sigma_\lambda^2 + \frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\frac{N_{jt}^{(1)}-1}{N_{jt}^{(1)}}\rho_{(2)}$$

$$E\left[\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\left(\frac{1}{N_{it}^{(2)}}\sum_{j\in Q_{it}^{(2)}}u_{ijt}^*\right)^2\right] = \frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\frac{1}{N_{it}^{(2)}}\left(\sigma_\mu^2+\sigma_\epsilon^2\right)+$$
$$+\sigma_\lambda^2 + \frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\frac{N_{it}^{(2)}-1}{N_{it}^{(2)}}\rho_{(1)}$$

$$E\left[\frac{1}{T_O}\sum_{t\in Z_O}\left(\frac{1}{N_t^{(1)}}\sum_{i\in Q_t^{(1)}}\frac{1}{N_{it}^{(2)}}\sum_{j\in Q_{it}^{(2)}}u_{ijt}^*\right)^2\right] = \frac{1}{T_O}\sum_{t\in Z_O}\frac{1}{N_t^{(1)^2}}\sum_{i\in Q_t^{(1)}}\frac{1}{N_{it}^{(2)}}\left(\sigma_\mu^2+\sigma_\epsilon^2\right)+$$
$$+\sigma_\lambda^2 + \frac{1}{T_O}\sum_{t\in Z_O}\frac{1}{N_t^{(1)^2}}\sum_{i\in Q_t^{(1)}}\frac{N_{it}^{(2)}-1}{N_{it}^{(2)}}\rho_{(1)} + \frac{2}{T_O}\sum_{t\in Z_O}\sum_{j\in Q_t^{(2)}}\sum_{i,s\in Q_{jt}^{(1)},s\neq i}\frac{1}{N_{it}^{(2)}N_{st}^{(2)}}\rho_{(2)}$$

For simplicity let denote

$$a = \frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\frac{1}{N_{jt}^{(1)}}$$

$$b = \frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T_j^*}\sum_{t\in Z_j^*}\frac{N_{jt}^{(1)}-1}{N_{jt}^{(1)}}$$

$$c = \frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T_i^{**}}\sum_{t\in Z_i^{**}}\frac{1}{N_{it}^{(2)}}$$

$$d = \frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{T_i^{**}} \sum_{t \in Z_i^{**}} \frac{N_{it}^{(2)} - 1}{N_{it}^{(2)}}$$

$$e = \frac{1}{T_O} \sum_{t \in Z_O} \frac{1}{N_t^{(1)^2}} \sum_{i \in Q_t^{(1)}} \frac{1}{N_{it}^{(2)}}$$

$$f = \frac{1}{T_O} \sum_{t \in Z_O} \frac{1}{N_t^{(1)^2}} \sum_{i \in Q_t^{(1)}} \frac{N_{it}^{(2)} - 1}{N_{it}^{(2)}}$$

$$g = \frac{2}{T_O} \sum_{t \in Z_O} \sum_{j \in Q_t^{(2)}} \sum_{i,s \in Q_{jt}^{(1)}, s \neq i} \frac{1}{N_{it}^{(2)} N_{st}^{(2)}}$$

$$A = \frac{1}{T} \sum_{i \in Q^{(1)}} \sum_{j \in Q_i^{(2)}} \sum_{t \in Z_{ij}} \hat{u}_{ijt}^{*^2} - \hat{\sigma}_\mu^2$$

$$B = \frac{1}{N^{(2)}} \sum_{j \in Q^{(2)}} \frac{1}{T_j^*} \sum_{t \in Z_j^*} \left(\frac{1}{N_{jt}^{(1)}} \sum_{i \in Q_{jt}^{(1)}} \hat{u}_{ijt}^{*^2} \right)^2 - a\hat{\sigma}_\mu^2$$

$$C = \frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{T_i^{**}} \sum_{t \in Z_i^{**}} \left(\frac{1}{N_{it}^{(2)}} \sum_{j \in Q_{it}^{(2)}} \hat{u}_{ijt}^{*} \right)^2 - c\hat{\sigma}_\mu^2$$

$$D = \frac{1}{T_O} \sum_{t \in Z_O} \left(\frac{1}{N_t^{(1)}} \sum_{i \in Q_t^{(1)}} \frac{1}{N_{it}^{(2)}} \sum_{j \in Q_{it}^{(2)}} \hat{u}_{ijt}^{*} \right)^2 - e\hat{\sigma}_\mu^2$$

So we get

$$\hat{\sigma}_\epsilon^2 = \frac{dbD - dB - bC - (bd - b - d)A}{db(e-1) - bf(c-1) - dg(a-1)}$$

$$\hat{\sigma}_\lambda^2 = A - \hat{\sigma}_\epsilon^2$$

$$\hat{\rho}_{(1)} = \frac{1}{d}(C - A) - \frac{c-1}{d}\hat{\sigma}_\epsilon^2$$

$$\hat{\rho}_{(2)} = \frac{1}{b}(B - A) - \frac{a-1}{b}\hat{\sigma}_\epsilon^2$$

Finally, for the model (7), the estimation σ_μ^2 is, again, the same as for the balanced case. For the other terms the identification equations are now becoming, unfortunately, slightly more complicated

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

$$E\left[\frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T^*_j}\sum_{t\in Z^*_j}\left(\frac{1}{N^{(1)}_{jt}}\sum_{i\in Q^{(1)}_{jt}}u^*_{ijt}\right)^2\right] = \frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T^*_j}\sum_{t\in Z^*_j}\frac{1}{N^{(1)}_{jt}}\left(\sigma^2_\mu+\sigma^2_v+\sigma^2_\epsilon\right)+$$

$$+\sigma^2_u + \frac{1}{N^{(2)}}\sum_{j\in Q^{(2)}}\frac{1}{T^*_j}\sum_{t\in Z^*_j}\frac{N^{(1)}_{jt}-1}{N^{(1)}_{jt}}\rho_{(2)}$$

$$E\left[\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T^{**}_i}\sum_{t\in Z^{**}_i}\left(\frac{1}{N^{(2)}_{it}}\sum_{j\in Q^{(2)}_{it}}u^*_{ijt}\right)^2\right] = \frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T^{**}_i}\sum_{t\in Z^{**}_i}\frac{1}{N^{(2)}_{it}}\left(\sigma^2_\mu+\sigma^2_u+\sigma^2_\epsilon\right)+$$

$$+\sigma^2_v + \frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{T^{**}_i}\sum_{t\in Z^{**}_i}\frac{N^{(2)}_{it}-1}{N^{(2)}_{it}}\rho_{(1)}$$

$$E\left[\frac{1}{T_O}\sum_{t\in Z_O}\left(\frac{1}{N^{(1)}_t}\sum_{i\in Q^{(1)}_t}\frac{1}{N^{(2)}_{it}}\sum_{j\in Q^{(2)}_{it}}u^*_{ijt}\right)^2\right] = \frac{1}{T_O}\sum_{t\in Z_O}\frac{1}{N^{(1)^2}_t}\sum_{i\in Q^{(1)}_t}\frac{1}{N^{(2)}_{it}}\left(\sigma^2_\mu+\sigma^2_\epsilon\right)+$$

$$+\frac{1}{T_O}\sum_{t\in Z_O}\frac{1}{N^{(1)}_t}\sigma^2_v + \frac{1}{T_O}\sum_{t\in Z_O}\frac{1}{N^{(1)}_t}\sum_{i\in Q^{(1)}_t}\frac{1}{N^{(2)}_{it}}\sigma^2_u +$$

$$+\frac{1}{T_O}\sum_{t\in Z_O}\frac{1}{N^{(1)^2}_t}\sum_{i\in Q^{(1)}_t}\frac{N^{(2)}_{it}-1}{N^{(2)}_{it}}\rho_{(1)}+$$

$$+\frac{2}{T_O}\sum_{t\in Z_O}\sum_{j\in Q^{(2)}_t}\sum_{i,s\in Q^{(1)}_{jt},s\neq i}\frac{1}{N^{(2)}_{it}N^{(2)}_{st}}\rho_{(2)}$$

$$E\left[\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\left(\frac{1}{N^{(2)}_i}\sum_{j\in Q^{(2)}_i}\frac{1}{T_{ij}}\sum_{t\in Z_{ij}}u^*_{ijt}\right)^2\right] = \frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{N^{(2)}_i}\sigma^2_\mu +$$

$$+\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{N^{(2)}_i}\sum_{j\in Q^{(2)}_i}\frac{1}{T_{ij}}\sigma^2_v + \frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{1}{N^{(2)^2}_i}\sum_{j\in Q^{(2)}_i}\frac{1}{T_{ij}}\left(\sigma^2_u+\sigma^2_\epsilon\right)+$$

$$+\frac{1}{N^{(1)}}\sum_{i\in Q^{(1)}}\frac{N^{(2)}_i-1}{N^{(2)}_i}\rho_{(1)}$$

Using previous notations and

$$
\begin{aligned}
k &= \frac{1}{T_O} \sum_{t \in Z_O} \frac{1}{N_t^{(1)}} \\
l &= \frac{1}{T_O} \sum_{t \in Z_O} \frac{1}{N_t^{(1)}} \sum_{i \in Q_t^{(1)}} \frac{1}{N_{it}^{(2)}} \\
m &= \frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{N_i^{(2)}} \sum_{j \in Q_i^{(2)}} \frac{1}{T_{ij}} \\
p &= \frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{N_i^{(2)2}} \sum_{j \in Q_i^{(2)}} \frac{1}{T_{ij}} \\
r &= \frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{N_i^{(2)} - 1}{N_i^{(2)}} \\
E &= \frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \left(\frac{1}{N_i^{(2)}} \sum_{j \in Q_i^{(2)}} \frac{1}{T_{ij}} \sum_{t \in Z_{ij}} \hat{u}_{ijt}^* \right)^2 - \frac{1}{N^{(1)}} \sum_{i \in Q^{(1)}} \frac{1}{N_i^{(2)}} \hat{\sigma}_\mu^2
\end{aligned}
$$

So finally we get

$$
\begin{aligned}
\hat{\sigma}_v^2 &= \frac{d(E - pA) - r(C - cA)}{d(m - p) - r(1 - c)} \\
\hat{\rho}_{(1)} &= \frac{C - cA}{d} - \frac{1 - c}{d} \hat{\sigma}_v^2 \\
\hat{\sigma}_\epsilon^2 &= \frac{b(D^* - lA^*) - g(B^* - A^*)}{b(e - l) - g(a - 1)} \\
\hat{\rho}_{(2)} &= \frac{B^* - A^*}{b} - \frac{a - 1}{b} \hat{\sigma}_\epsilon^2 \\
\hat{\sigma}_u^2 &= A^* - \hat{\sigma}_\epsilon
\end{aligned}
$$

where $A^* = A - \hat{\sigma}_v^2$, $B^* = B - a\hat{\sigma}_v^2$ and $D^* = D - k\hat{\sigma}_v^2 - f\hat{\rho}_{(1)}$.

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 of European countries. We capture joint EU membership with a dummy variable, being 1 if both the exporter and the importer countries are members of the EU. We assume either of the six types of error structures considered in this paper and compare the RE estimates to the FE and Pooled OLS estimates.

The data consist of a balanced panel of bilateral trade flows (in current euros) for all pairs formed by 20 European countries for years 2001-2006. Hence, the total number of country pairs is $N^2 = 400$ and the number of years is $T = 6$. Twelve of the countries are members of the EU in the whole sample period (group A): Austria, Germany, Denmark, Spain, Finland, France, Greece, Ireland, Italy, Portugal, Sweden, United Kingdom. The remaining eight entered the EU in 2004 (group B): Czech Republic, Estonia, Hungary, Lithuania, Latvia, Poland, Slovenia, Slovakia.

We augment our foreign trade database with self-trade flows, i.e. the trade flow of a country within its own borders. We include self-trade to avoid the bias of the Fixed Effects within transformation formulas under models (6) and (7).[2] Following *Wei* [1996], self-trade of a country for a given year is generated as gross output minus total exports of the country in that year.

Bilateral distance is based on the weighted distances between the biggest cities in the two countries and is taken from CEPII. In the case of self-trade, we use the internal distance variable from the same source.

A first look at the trade flows 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 both for faster economic growth of group B countries and for the trade creating effect of their entering the EU. Our estimation exercise tries to disentangle these forces, where we assume for simplicity that geographical distance and joint EU membership are sufficient to describe the level of bilateral trade barriers for the sampled countries.

Turning now to the model specification, for better tractability, we restrict the elasticities of trade to income to unity and use income-adjusted trade as dependent variable. We measure country incomes with nominal GDP (in current euros). We take all variables except the EU dummy in logarithms. Our left-hand side variable is then $y_{ijt} = \ln \text{trade}_{ijt} - \ln \text{GDP}_{it} - \ln \text{GDP}_{jt}$. The explanatory variables are the time-invariant distance variable, $\ln \text{dist}_{ij}$, and the dummy for joint EU membership,

[2] As it is stressed in *Hornok* [2011] and *Mátyás and Balázsi* [2011], the usual Fixed Effects within transformation formulas for these error structures give biased estimates if self-trade is not included in the database.

EU_{ijt}. The EU dummy has some time variation due to the enlargement of the EU with type-B countries in 2004, which falls in the middle of our sample period.

We estimate two specifications: one with a single EU dummy and another with three EU dummies. Formally, the EU dummy in the first specification is

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

It is 1 for country pairs of two type-A countries in all years and for all pairs from 2004 onwards, and zero otherwise.

Our aim with the second specification is to estimate heterogeneous trade effects per country group. We define separate EU dummies for pairs of two type-A countries ($\mathrm{EU_{AA}}$), for pairs of two type-B countries ($\mathrm{EU_{BB}}$) and for pairs with one type-A and one type-B country ($\mathrm{EU_{AB}}$) as below.

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

$$\mathrm{EU}_{BB} = \begin{cases} 1 & (i \in \mathrm{B} \text{ and } \ j \in \mathrm{B}) \text{ and } \ t \geq 2004 \\ 0 & \text{otherwise} \end{cases}$$

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

The specification with the three EU dummies implicitly assumes that EU membership can have heterogeneous effects across groups of country pairs. Indeed, data in *Table 1* shows that type-B countries' mutual trade grew more strongly around EU enlargement than trade between type-A and type-B countries. This pattern remains after controlling for differences in GDP growth rates. We consider this specification also because it highlights some of the advantages of the RE estimator over the FE estimator, when dummy regressors are to capture heterogenous effects of trade policies.

We estimate separate panel gravity equations by assuming either of the error structures (1) to (7) discussed in the previous sections. We report both FE, RE and pooled OLS estimates, as well as the estimated variances for the random error components. The RE estimation is done by Feasible GLS. Pooled OLS estimates are naturally identical for each error structure.

The estimates for the specification with a single EU dummy are reported in *Table 2*. Clearly, the FE estimator cannot identify the distance elasticity parameter under models (1), (3) and (7), for the time-invariant distance variable is perfectly collinear with the country pair fixed effects. Apart from that, the distance coefficient estimates are very stable across models and methods.

The coefficient for joint EU membership is identified under all the six models, thanks to its variation both across country pairs and in time. However, the estimates vary considerably across models. The large differences are likely to be driven by whether identification is mostly from variation of the EU dummy in the country pair dimension (model (6)) or in time (models (1), (3) and (7)) and by the fact that the EU dummy has limited variation in both dimensions. Apart from that, the RE parameter estimates happen to be quite close to the FE estimates (except for model (6)).

Table 3 reports the estimates for the specification with heterogenous EU effects. In most of the cases, we get significantly different RE estimates for the effect of EU membership across the three country groups, which supports the choice of this specification. As expected, the EU parameter estimate is the highest for pairs of two type-B countries and the lowest for pairs of two type-A countries.

Our preferred RE estimates are from model (7). This model is consistent with the recent developments in the structural gravity model (*Anderson and van Wincoop* [2003]), since it accounts for the time-varying trade barriers with third countries via country-time random effects. The EU estimates under this model show that the entry of type-B countries in the EU lead to a 20% trade growth between type-B and type-A countries and a 50% trade growth in the mutual trade of type-B countries.

The FE estimation results in *Table 3* help to bring to light again an important disadvantage of the FE over the RE estimator. The FE estimator is not able to identify the coefficients of EU_{BB} and EU_{AB} separately under models (6) and (7), although the two dummy variables vary in the ijt dimension, while the within transformations net out fixed effects only in the it and jt (and ij) dimensions. The unidentification is due to perfect or near-perfect collinearity among the fixed effects and the dummy regressors and the fact that the variation of the EU dummies in the ijt dimension is limited.[3]

Let us turn now our attention to the estimation results with no self-trade flows and cross-correlation. We report estimates for a modified database that lacks self-trade flows, a case discussed in Section 5, in *Table 4* and *Table 5*. Working out this case empirically is especially relevant, since international trade databases never

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

include self flows. Moreover, statistical measurement of self-trade (i.e. gross domestic sales) is non-existent, and deriving self-trade from other statistics is ultimately subject to measurement error.

The RE estimates on the database without self-trade are qualitatively similar to our baseline RE estimates. Joint EU membership seems to affect trade at a somewhat more moderate extent though. As it is estimated under model (7), mutual trade of type-B countries went up by 40%, trade between type-B and type-A countries by 13% following EU enlargement.

Finally, we report estimates for the extension in Section 6, where we assume away the uncorrelatedness of country pair effects with the same exporter or importer. The RE estimation results for the affected models (models (1), (3) and (7)) are reported in *Table 6*. We can conclude that the introduction of cross-correlations 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 (around 0.15).

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) will result in a massive over-specification, which implies that much less data information will be available for the estimation of the main/focus parameters. 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, 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, 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.

Paolo Bientinesi, Brian Gunter and Robert A. van de Geijn [2000]: Families of Algorithms Related to the Inversion of a Symmetric Positive Definite Matrix, mimeo.

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.

I-Hui Cheng and Howard J. Wall [2005]: Controlling for Heterogeneity in Gravity Models of Trade and Integration, The Federal Reserve Bank of St. Louis Review, January/February 2005, 87(1), 49-63.

Peter Ebbes, Ulf Bockenholt and Michel Wedel [2004]: Regressor and Random-effects Dependencies in Multilevel Models, *Statistica Neerlandica*, 58(2), 161178.

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.

Robert A. van de Geijn (ed.) [2008]: A Systematic Approach to Matrix Computations, The University of Texas at Austin, manuscript.

Andrew Gelman [2006]: Multilevel (Hierarchical) Modeling: What It Can and Cannot Do, *Technometrics*, 432-435.

Michael Healy [1968]: Algorithm AS 7: Inversion of a Positive Semi-Definite Symmetric Matrix, *Applied Statistics*, 17(2), 198-199.

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

Olaf Hubler [2006]: Multilevel and Nonlinear Panel Data Models, *Allgemeines Statistisches Archiv*, 90, 121-136.

David Karemera, Victor Iwuagwu Oguledo and Bobby Davis [2000]: A Gravity Model Analysis of International Migration to North America, *Applied Economics*, 32(13), 1745-1755.

Brent Moulton [1990]: An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units, *The Review of Economics and Statistics*, 72, 334-338.

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.

Simone Juhasz Silva and Douglas Nelson [2012]: Does Aid Cause Trade? Evidence from an Asymmetric Gravity Model, *The World Economy*, 545-577.

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			0.191			
			(0.037)			
ln dist			-1.645			
			(0.019)			
Fixed Effects						
EU	-0.031	0.126	-0.065	-0.257	0.593	0.007
	(0.012)	(0.019)	(0.054)	(0.057)	(0.073)	(0.033)
ln dist	–	–	-1.681	-1.610	-1.609	–
	–	–	(0.017)	(0.018)	(0.015)	–
Random Effects						
EU	-0.035	0.110	-0.114	-0.250	0.262	-0.003
	(0.012)	(0.019)	(0.049)	(0.049)	(0.060)	(0.032)
ln dist	-1.648	-1.650	-1.677	-1.616	-1.628	-1.645
	(0.046)	(0.045)	(0.017)	(0.018)	(0.015)	(0.035)
Variance Components						
σ_ε^2	0.053	0.049	0.509	0.561	0.346	0.041
σ_μ^2	0.639	0.630				0.342
σ_λ^2		0.012				
σ_u^2			0.183		0.179	0.179
σ_v^2				0.131	0.167	0.130

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}			-0.426			
			(0.039)			
EU_{BB}			0.629			
			(0.065)			
EU_{AB}			-0.251			
			(0.044)			
ln dist			-1.572			
			(0.019)			
Fixed Effects						
EU_{AA}	–	–	-0.058	-0.251	0.626	–
	–	–	(0.052)	(0.055)	(0.072)	–
EU_{BB}	0.041	0.198	1.078	0.844	–	–
	(0.023)	(0.027)	(0.099)	(0.104)	–	–
EU_{AB}	-0.055	0.102	0.229	-0.024	–	–
	(0.014)	(0.020)	(0.074)	(0.078)	–	–
ln dist	–	–	-1.631	-1.558	-1.578	–
	–	–	(0.017)	(0.018)	(0.015)	–
Random Effects						
EU_{AA}	-0.420	-0.343	-0.176	-0.329	0.217	-0.091
	(0.080)	(0.078)	(0.047)	(0.047)	(0.058)	(0.086)
EU_{BB}	0.057	0.212	0.870	0.716	1.002	0.428
	(0.023)	(0.027)	(0.084)	(0.083)	(0.136)	(0.134)
EU_{AB}	-0.060	0.095	0.041	-0.142	0.295	0.202
	(0.013)	(0.020)	(0.061)	(0.060)	(0.082)	(0.068)
ln dist	-1.627	-1.628	-1.625	-1.562	-1.596	-1.638
	(0.043)	(0.042)	(0.017)	(0.018)	(0.015)	(0.035)
Variance Components						
σ^2_ε	0.053	0.049	0.469	0.521	0.334	0.041
σ^2_μ	0.571	0.548				0.343
σ^2_λ		0.026				
σ^2_u			0.155		0.141	0.141
σ^2_v				0.102	0.149	0.099

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: No self-trade: Random Effects estimates

	(1)	(3)	(4)	(5)	(6)	(7)
Random Effects						
EU	-0.019	0.122	-0.029	-0.191	0.156	-0.031
	(0.012)	(0.019)	(0.046)	(0.046)	(0.052)	(0.030)
ln dist	-1.631	-1.641	-1.732	-1.567	-1.673	-1.664
	(0.059)	(0.058)	(0.023)	(0.024)	(0.021)	(0.044)
Variance Components						
σ^2_ε	0.053	0.050	0.424	0.485	0.269	0.037
σ^2_μ	0.546	0.541				0.228
σ^2_λ	0.008					
σ^2_u			0.175		0.167	0.167
σ^2_v				0.114	0.163	0.167

Notes: Dependent variable is log of income-adjusted bilateral trade. Standard errors in parenthesis. Estimation on a balanced panel with no self-trade of pairs of 20 EU countries for years 2001-2006.

Table 5: No self-trade: Random Effects estimates with three EU dummies

	(1)	(3)	(4)	(5)	(6)	(7)
Random Effects						
EU_{AA}	-0.309	-0.233	-0.084	-0.267	0.129	-0.025
	(0.079)	(0.078)	(0.045)	(0.045)	(0.052)	(0.076)
EU_{BB}	0.131	0.282	0.759	0.656	0.849	0.347
	(0.025)	(0.029)	(0.081)	(0.081)	(0.120)	(0.085)
EU_{AB}	-0.059	0.093	0.077	-0.111	0.229	0.126
	(0.013)	(0.020)	(0.058)	(0.058)	(0.074)	(0.044)
ln dist	-1.573	-1.573	-1.660	-1.487	-1.631	-1.641
	(0.057)	(0.056)	(0.023)	(0.024)	(0.021)	(0.045)
Variance Components						
σ^2_ε	0.052	0.049	0.399	0.455	0.262	0.037
σ^2_μ	0.502	0.489				0.243
σ^2_λ		0.016				
σ^2_u			0.155		0.137	0.137
σ^2_v				0.099	0.155	0.137

Notes: Dependent variable is log of income-adjusted bilateral trade. Standard errors in parenthesis. Estimation on a balanced panel with no self-trade of pairs of 20 EU countries for years 2001-2006.

Table 6: Random Effects with cross correlation

	(1)	(3)	(7)
Random Effects			
EU	-0.029	0.125	-0.007
	(0.012)	(0.019)	(0.023)
ln dist	-1.638	-1.632	-1.641
	(0.036)	(0.037)	(0.036)
Variance Components			
σ^2_ε	0.053	0.049	0.041
σ^2_μ	0.639	0.630	0.630
σ^2_λ		0.012	
σ^2_u			0.004
σ^2_v			0.017
ρ_1	0.179	0.167	0.113
ρ_2	0.130	0.118	0.175

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.