# QUADERNI DEL DIPARTIMENTO DI SCIENZE ECONOMICHE E SOCIALI

## A NOTE ON TWO-WAY ECM ESTIMATION OF SUR SYSTEMS ON UNBALANCED PANEL DATA

*Silvia Platoni, Paolo Sckokai e Daniele Moro*

*Serie Rossa: Economia – Quaderno N. 67 luglio 2010*

**UNIVERSITÀ CATTOLICA DEL SACRO CUORE**
**PIACENZA**

# A NOTE ON TWO-WAY ECM ESTIMATION OF SUR SYSTEMS ON UNBALANCED PANEL DATA[♣]


SILVIA PLATONI[1♠], PAOLO SCKOKAI[2] AND DANIELE MORO[2]

[1]*Dipartimento di Scienze Economiche e Sociali, Università Cattolica del Sacro Cuore, Piacenza, Italy*
[2]*Istituto di Economia Agro-alimentare, Università Cattolica del Sacro Cuore, Piacenza, Italy*



This paper considers the two-way error components model (*ECM*) estimation of seemingly unrelated regressions (*SUR*) on unbalanced panel by generalized least squares (*GLS*). As suggested by Biørn (2004) for the one-way case, in order to use the standard results for the balanced case the individuals are arranged in groups according to the number of times they are observed. Thus, the *GLS* estimator can be interpreted as a matrix weighted average of the group specific *GLS* estimators with weights equal to the inverse of their respective covariance matrices.



KEYWORDS Unbalanced panels; Error Components Model; Seemingly Unrelated Regressions.
JEL CLASSIFICATION C13; C23; C33.


## 1. INTRODUCTION

The error components model (*ECM*) is the most frequently used approach to estimate models on panel data. The phenomenon of missing observations—not all cross-sectional units are observed during all time periods—is a problem that often occurs in practice: the unbalanced panel is the rule rather than the exception when the data come from large-scale surveys. Biørn (1981) and Baltagi (1985) discuss the single-equation *ECM* with missing observations, but they focus on the one-way case, where only the individual-specific effects are considered. Wansbeek and Kapteyn (1989) and Davis (2002) extend this estimation method to the two-way case, where both the individual-specific and the time-specific effects are taken into account, as well as to the multi-way case, where a third specific effect can be considered, for example the location effects[1].

---


♣ This research has been carried out as part of the WEMAC (World Econometric Modelling of Arable Crops) research project (Scientific coordinator: Catherine Benjamin), funded by the European Commission under the 6th Framework programme. The FADN data have been provided in the context of that project.



♠ Dipartimento di Scienze Economiche e Sociali, Università Cattolica del Sacro Cuore, via Emilia Parmense 84, 29122 Piacenza, Italy (silvia.platoni@unicatt.it).


[1] Among the recent empirical applications, Boumahdi, Chaaban and Thomas (2004) estimate the Lebanon import demand elasticities using a three-way *ECM*—nested and non-nested—in which product, country and time effects are introduced.

All these papers apply the *ECM* to the single-equation case. Baltagi (1980) and Magnus (1982) extend the estimation procedure of the single-equation model for balanced panels to the case of seemingly unrelated regressions (*SUR*), while Biørn (2004) proposes a parsimonious technique to estimate one-way *SUR* systems on unbalanced panel data.

The purpose of this paper is to extend the Generalized Least Squares (*GLS*) estimation of the *SUR* system suggested by Biørn (2004) to the case of the two-way *ECM* for unbalanced panels, considering not only the individual-specific effect, but also the time-specific effect. This extension is rather important, since the estimation of the time-specific effects is likely to play an important role in many practical situations, especially when the time period is sufficiently long. For example a parametric trend is often used to parameterize the effect (on the response variable) either of technical improvements in case of panels of firms or of the change in tastes over time in case of panels of households.

In order to use the standard results for the balanced case, the key element of the Biørn (2004)'s technique is arranging the data such that individuals are grouped according to the number of times they are observed. Extending this approach to the two-way *SUR* allows to estimate systems of equations also on large unbalanced panel databases with a relevant time dimension.

The structure of the paper is the following: in section 2 we introduce the logic and the notation of the single-equation case, while in section 3 we develop the corresponding *SUR* system. Finally, in section 4 some simulation results are provided for illustrative purpose.

## 2. Single-Equation Two-Way *ECM* for Unbalanced Panels

We analyze an unbalanced panel characterized by a total of $n$ observations, with $N$ individuals (indexed $i=1,\ldots,N$) observed over $T$ periods (indexed $t=1,\ldots,T$). Let $T_i$ denote the number of times the individual $i$ is observed and $N_t$ the number of individuals observed in period $t$. Hence $\sum_i T_i = \sum_t N_t = n$.

In the following we consider the regression model[2]

$$y_{it} = \underset{1\times k}{\mathbf{x}'_{it}}\ \underset{k\times 1}{\boldsymbol{\beta}} + \mu_i + \nu_t + u_{it}, \tag{1}$$

[2] Throughout the paper, all vectors and matrices are in boldface.

where $\mathbf{x}_{it}$ is a $1 \times k$ vector of explanatory variables and $\boldsymbol{\beta}$ a $k \times 1$ vector of parameters, $\mu_i$ is the individual-specific effect, $\nu_t$ the time-specific effect and $u_{it}$ the remainder error term. Since the panel is unbalanced, the standard projection and transformation results no longer hold. Thus, Wansbeek and Kapteyn (1989) propose to order the data on the $N$ individuals in $T$ consecutive sets, one for each period. Let $\mathbf{D}_t$ be the $N_t \times N$ matrix obtained from the $N \times N$ identity matrix $\mathbf{I}_N$ by omitting the rows corresponding to individuals not observed in period $t$. Using the matrices $\underset{n\times N}{\boldsymbol{\Delta}_\mu} \equiv (\underset{N\times N_1}{\mathbf{D}'_1}, \ldots, \underset{N\times N_T}{\mathbf{D}'_T})'$ and $\underset{n\times T}{\boldsymbol{\Delta}_\nu} \equiv \text{blockdiag}[\underset{N_t\times N}{\mathbf{D}_t} \underset{N\times 1}{\boldsymbol{\iota}_N}] = \text{blockdiag}[\underset{N_t\times 1}{\boldsymbol{\iota}_{N_t}}]$, where $\boldsymbol{\iota}_N$ and $\boldsymbol{\iota}_{N_t}$ are vectors of ones of dimension $N$ and $N_t$ respectively, we can define the diagonal matrices $\underset{N\times N}{\boldsymbol{\Delta}_N} \equiv \underset{N\times n}{\boldsymbol{\Delta}'_\mu} \underset{n\times N}{\boldsymbol{\Delta}_\mu}$ and $\underset{T\times T}{\boldsymbol{\Delta}_T} \equiv \underset{T\times n}{\boldsymbol{\Delta}'_\nu} \underset{n\times T}{\boldsymbol{\Delta}_\nu}$, as well as the matrix of zeros and ones $\underset{T\times N}{\boldsymbol{\Delta}_{TN}} \equiv \underset{T\times n}{\boldsymbol{\Delta}'_\nu} \underset{n\times N}{\boldsymbol{\Delta}_\mu}$, indicating the absence or presence of an individual in a certain time period. Moreover, we can consider the matrix $\underset{n\times(N+T)}{\boldsymbol{\Delta}} \equiv (\underset{n\times N}{\boldsymbol{\Delta}_\mu}, \underset{n\times T}{\boldsymbol{\Delta}_\nu})$, which gives the dummy-variable structure for the unbalanced panel model (see Baltagi, 2005). Hence, using matrix notation, we can write

$$\underset{n\times 1}{\mathbf{y}} = \underset{n\times k}{\mathbf{X}} \underset{k\times 1}{\boldsymbol{\beta}} + \underset{n\times N}{\boldsymbol{\Delta}_\mu} \underset{N\times 1}{\boldsymbol{\mu}} + \underset{n\times T}{\boldsymbol{\Delta}_\nu} \underset{T\times 1}{\boldsymbol{\nu}} + \underset{n\times 1}{\mathbf{u}} = \underset{n\times k}{\mathbf{X}} \underset{k\times 1}{\boldsymbol{\beta}} + \underset{n\times 1}{\boldsymbol{\varepsilon}}, \tag{2}$$

where $\mathbf{X}$ is a $n \times k$ matrix of explanatory variables and $\varepsilon_{it} = \mu_i + \nu_t + u_{it}$ the composite error term.

In the fixed effects (FE) case, where the error components $\mu_i$ and $\nu_t$ are parameters to be estimated, we assume the following (see Appendix A for details).

- (FE.1) Strict exogeneity: the set of explanatory variables in each time period $\underset{N_t k}{\mathbf{x}_{t\{T\}}}$ is uncorrelated with the idiosyncratic error $u_{it}$ and the set of explanatory variables for each individual $\underset{T_i k}{\mathbf{x}_{i\{N\}}}$ is also uncorrelated with the same idiosyncratic error $u_{it}$.
- (FE.2) Consistency: the within estimator is asymptotically well behaved, in the sense that the "adjusted" outer product matrix has the appropriate rank.
- (FE.3) Homoscedasticity and no serial correlation: the conditional variance-covariance matrix of the idiosyncratic error terms $u_{it}$ coincides with the unconditional one, and it is characterized by constant variances and zero covariances.

The assumptions FE.1 and FE.3 guarantee the efficiency of the within estimator.

In the random effect (RE) case, where all error components are random variables, we assume the following (see Appendix B for details).

- (RE.1.a) Strict exogeneity: same definition as assumption FE.1.
- (RE.1.b) and (RE.1.c) Orthogonality conditions: both $\mu_i$ and $\nu_t$ are orthogonal to the corresponding sets of explanatory variables $\underset{T_i k}{\mathbf{x}_{i\{N\}}}$ and $\underset{N_t k}{\mathbf{x}_{t\{T\}}}$.
- (RE.2) Consistency: the RE *GLS* estimator is consistent, in the sense that the weighted outer product matrix has the appropriate rank.
- (RE.3a), (RE.3b) and (RE.3c) Homoscedasticity and no serial correlation: the conditional variance-covariance matrix of the idiosyncratic error terms $u_{it}$ is characterized by constant variances and zero covariances; in addition, the variance of both the individual specific effects $\mu_i$ and the time-specific effects $\nu_t$ is constant.

## 3. TWO-WAY *SUR* SYSTEMS FOR UNBALANCED PANELS

Biørn (2004) estimates a one-way *SUR* system of equations on unbalanced panel data. Thus, he considers only the individual-specific effect, while in this paper we extend his analysis considering also the time-specific effect.

### 3.1. MODEL AND NOTATION

Grouping individuals according to the number of times they are observed, as suggested by Biørn (2004) for the one-way case, is essential also in our two-way *SUR* systems, that would not be manageable adopting traditional estimation techniques. This can be done as follows.

Let $\widetilde{N}_p$ denote the number of individuals observed exactly in $p$ periods, with $p = 1,\ldots,T$. Hence $\sum_p \widetilde{N}_p = N$ and $\sum_p (\widetilde{N}_p p) = n$. We assume that the $T$ groups of individuals are ordered such that the $\widetilde{N}_1$ individuals observed once come first, the $\widetilde{N}_2$ individuals observed twice come second, etc. Hence with $C_p = \sum_{h=1}^{p} \widetilde{N}_h$ being the cumulated number of individuals observed at most $p$ times, the index sets of the individuals observed exactly $p$ times can be written as $I_p = \{C_{p-1}+1,\ldots,C_p\}$. Note that $I_1$ may be considered as a pure cross section and $I_p$, with $p \geq 2$, as a pseudo-balanced panel with $p$ observations for each individual. This structure allows to use a number of results derived for the two-way *SUR* in the balanced case.

If $k_m$ is the number of regressors for equation $m$, the total number of regressors for the system is $K=\sum_{m=1}^{M}k_m$. Stacking the $M$ equations, indexed by $m=1,\dots,M$, for the observation $(i,t)$ we have

$$\underset{M\times 1}{\mathbf{y}_{it}}=\underset{M\times K}{\mathbf{X}_{it}}\underset{K\times 1}{\boldsymbol{\beta}}+\underset{M\times 1}{\boldsymbol{\mu}_{i\{N\}}}+\underset{M\times 1}{\mathbf{v}_{t\{T\}}}+\underset{M\times 1}{\mathbf{u}_{it}}=\underset{M\times K}{\mathbf{X}_{it}}\underset{K\times 1}{\boldsymbol{\beta}}+\underset{M\times 1}{\boldsymbol{\varepsilon}_{it}}, \tag{3}$$

where $\underset{M\times K}{\mathbf{X}_{it}}=\operatorname{diag}[\underset{1\times k_1}{\mathbf{x}_{1it}},\dots,\underset{1\times k_M}{\mathbf{x}_{Mit}}]$ and $\boldsymbol{\beta}=(\underset{1\times k_1}{\boldsymbol{\beta}'_1},\dots,\underset{1\times k_M}{\boldsymbol{\beta}'_M})'$ [3]. If we do not have cross-equation restrictions, we can assume $\mathrm{E}(u_{mit}\mid\mathbf{x}_{1it},\mathbf{x}_{2it},\dots,\mathbf{x}_{Mit})=0$ and then $\mathrm{E}(y_{mit}\mid\mathbf{x}_{1it},\mathbf{x}_{2it},\dots,\mathbf{x}_{Mit})=\mathrm{E}(y_{mit}\mid\mathbf{x}_{mit})=\mathbf{x}_{mit}\boldsymbol{\beta}_m$. On the contrary, if we have cross-equation restrictions we can only assume $\mathrm{E}(\mathbf{u}_{it}\mid\mathbf{x}_{it})=0$ where $\underset{M\times 1}{\mathbf{u}_{it}}\equiv(u_{1it},\dots,u_{Mit})'$ and $\underset{1\times K}{\mathbf{x}_{it}}\equiv(\mathbf{x}_{1it},\mathbf{x}_{2it},\dots,\mathbf{x}_{Mit})$. With $\underset{M\times 1}{\boldsymbol{\mu}_{i\{N\}}}=(\mu_{1i},\dots,\mu_{Mi})'$ and $\underset{M\times 1}{\mathbf{v}_{t\{T\}}}=(v_{1t},\dots,v_{Mt})'$ we assume

$$\begin{aligned}
&\mathrm{E}\left(\mu_{mi},\mu_{ji'}\right) && \begin{cases}=\ \sigma^2_{\mu_{mj}} & i=i'\\ =\ 0 & i\neq i',\end{cases}\\
&\mathrm{E}\left(v_{mt},v_{jt'}\right) && \begin{cases}=\ \sigma^2_{v_{mj}} & t=t'\\ =\ 0 & t\neq t',\end{cases}\\
&\mathrm{E}\left(u_{mit},u_{ji't'}\right) && \begin{cases}=\ \sigma^2_{u_{mj}} & i=i' \text{ and } t=t'\\ =\ 0 & i\neq i' \text{ and/or } t\neq t',\end{cases}
\end{aligned} \tag{4}$$

and then $\underset{N\times 1}{\boldsymbol{\mu}_m}\equiv(\mu_{m1},\dots,\mu_{mN})'$, $\underset{T\times 1}{\mathbf{v}_m}\equiv(v_{m1},\dots,v_{mT})'$ and $\underset{n\times 1}{\mathbf{u}_m}\equiv(u_{m11},u_{m12},\dots,u_{m1T_1},u_{m21},\dots,u_{mNT_N})'$ are random vectors with zero means and covariance matrix

$$\mathrm{E}\left(\begin{pmatrix}\boldsymbol{\mu}_m\\ \mathbf{v}_m\\ \mathbf{u}_m\end{pmatrix}\begin{pmatrix}\boldsymbol{\mu}'_j & \mathbf{v}'_j & \mathbf{u}'_j\end{pmatrix}\right)=\begin{bmatrix}\sigma^2_{\mu_{mj}} & 0 & 0\\ 0 & \sigma^2_{v_{mj}} & 0\\ 0 & 0 & \sigma^2_{u_{mj}}\end{bmatrix}. \tag{5}$$

---

[3] As Biørn (2004) suggests, if the coefficient vectors are not disjointed across equations, we can redefine $\underset{K\times 1}{\boldsymbol{\beta}}$ as the complete coefficient vector (without duplication) and the regression matrix as $\underset{M\times K}{\mathbf{X}_{it}}=(\underset{k_1\times 1}{\mathbf{x}'_{1it}},\underset{k_2\times 1}{\mathbf{x}'_{2it}},\dots,\underset{k_{M\times 1}}{\mathbf{x}'_{Mit}})'$ where the $k^{\text{th}}$ element of $\underset{1\times k_m}{\mathbf{x}_{mit}}$ (i) contains the observation on the variable in the $m^{\text{th}}$ equation which corresponds to the $k^{\text{th}}$ coefficient in $\underset{K\times 1}{\boldsymbol{\beta}}$ or (ii) is zero if the $k^{\text{th}}$ coefficient does not occur in the $m^{\text{th}}$ equation.

With $\underset{NM\times 1}{\boldsymbol{\mu}} \equiv (\boldsymbol{\mu}'_{1\{N\}},\ldots,\boldsymbol{\mu}'_{N\{N\}})'$, $\underset{TM\times 1}{\mathbf{v}} \equiv (\mathbf{v}'_{1\{T\}},\ldots,\mathbf{v}'_{T\{T\}})'$ and $\underset{nM\times 1}{\mathbf{u}} \equiv (\mathbf{u}'_{11},\mathbf{u}'_{12},\ldots,\mathbf{u}'_{1T_1},\mathbf{u}'_{21},\ldots,\mathbf{u}'_{NT_N})'$, since we have $\underset{NM\times 1}{\boldsymbol{\mu}} \sim \left(0, \underset{M\times M}{\boldsymbol{\Sigma}_\mu}\right)$, $\underset{TM\times 1}{\mathbf{v}} \sim \left(0, \underset{M\times M}{\boldsymbol{\Sigma}_v}\right)$ and $\underset{nM\times 1}{\mathbf{u}} \sim \left(0, \underset{M\times M}{\boldsymbol{\Sigma}_u}\right)$ with $\boldsymbol{\Sigma}_\mu = \left[\sigma^2_{\mu_{mj}}\right]$, $\boldsymbol{\Sigma}_v = \left[\sigma^2_{v_{mj}}\right]$ and $\boldsymbol{\Sigma}_u = \left[\sigma^2_{u_{mj}}\right]$, we can assume that the expected values of $\underset{M\times 1}{\boldsymbol{\mu}_{i\{N\}}}$, $\underset{M\times 1}{\mathbf{v}_{t\{T\}}}$ and $\underset{M\times 1}{\mathbf{u}_{it}}$ are zero and their covariance matrices are equal to $\underset{M\times M}{\boldsymbol{\Sigma}_\mu}$, $\underset{M\times M}{\boldsymbol{\Sigma}_v}$ and $\underset{M\times M}{\boldsymbol{\Sigma}_u}$. It follows that $\mathrm{E}(\underset{M\times 1}{\boldsymbol{\varepsilon}_{it}}\,\underset{1\times M}{\boldsymbol{\varepsilon}'_{i't'}}) = \delta_{ii'}\underset{M\times M}{\boldsymbol{\Sigma}_\mu} + \delta_{tt'}\underset{M\times M}{\boldsymbol{\Sigma}_v} + \delta_{ii'}\delta_{tt'}\underset{M\times M}{\boldsymbol{\Sigma}_u}$ with $\delta_{ii'} = 1$ for $i = i'$ and $\delta_{ii'} = 0$ for $i \neq i'$, $\delta_{tt'} = 1$ for $t = t'$ and $\delta_{tt'} = 0$ for $t \neq t'$.

Let us consider $\underset{pM\times 1}{\mathbf{y}_{i(p)}} \equiv (\underset{1\times M}{\mathbf{y}'_{i1}},\ldots,\underset{1\times M}{\mathbf{y}'_{ip}})'$, $\underset{pM\times K}{\mathbf{X}_{i(p)}} \equiv (\underset{K\times M}{\mathbf{X}'_{i1}},\ldots,\underset{K\times M}{\mathbf{X}'_{ip}})'$ and $\underset{pM\times 1}{\boldsymbol{\varepsilon}_{i(p)}} \equiv (\underset{1\times M}{\boldsymbol{\varepsilon}'_{i1}},\ldots,\underset{1\times M}{\boldsymbol{\varepsilon}'_{ip}})'$ for $i \in I_p$ (and then for $i = 1,\ldots,C_1,C_1+1,\ldots,C_2,\ldots,C_{T-1}+1,\ldots,C_T$ with $C_T = N$).

We define the matrix $\underset{pM\times TM}{\boldsymbol{\Delta}_{i(p)}}$ indicating in which period $t$ the individual $i$ of the group $p$ is observed. For example, with $T = 4$, if the individual $i$ is observed in the periods $t = 2$ and $t = 4$ (the individual $i$ belongs to group $p = 2$) we have

$$\underset{2M\times 4M}{\boldsymbol{\Delta}_{i(2)}} \equiv \begin{bmatrix} \underset{M\times M}{\mathbf{0}} & \underset{M\times M}{\mathbf{I}_M} & \underset{M\times M}{\mathbf{0}} & \underset{M\times M}{\mathbf{0}} \\ \underset{M\times M}{\mathbf{0}} & \underset{M\times M}{\mathbf{0}} & \underset{M\times M}{\mathbf{0}} & \underset{M\times M}{\mathbf{I}_M} \end{bmatrix},$$

where $\mathbf{I}_M$ is an identity matrix of dimension $M$. This is a convenient way of structuring the data in order to obtain vectors of time-specific errors of appropriate dimension. In fact, considering $\underset{TM\times 1}{\mathbf{v}}$, for the individual $i \in I_p$ we can define the vector $\underset{pM\times 1}{\mathbf{v}_{i(p)}} \equiv \underset{pM\times TM}{\boldsymbol{\Delta}_{i(p)}}\underset{TM\times 1}{\mathbf{v}}$ and write the model

$$\underset{pM\times 1}{\mathbf{y}_{i(p)}} = \underset{pM\times K}{\mathbf{X}_{i(p)}}\underset{K\times 1}{\boldsymbol{\beta}} + \left(\underset{p\times 1}{\boldsymbol{\iota}_p} \otimes \underset{M\times 1}{\boldsymbol{\mu}_i}\right) + \underset{pM\times 1}{\mathbf{v}_{i(p)}} + \underset{pM\times 1}{\mathbf{u}_{i(p)}} = \underset{pM\times K}{\mathbf{X}_{i(p)}}\underset{K\times 1}{\boldsymbol{\beta}} + \underset{pM\times 1}{\boldsymbol{\varepsilon}_{i(p)}}, \tag{6}$$

where $\boldsymbol{\iota}_p$ is a vector of ones of dimension $p$. The variance-covariance matrix of the composite error term $\varepsilon_{i(p)}$ is given by

$$\begin{aligned} \underset{pM\times pM}{\boldsymbol{\Omega}_p} = \mathrm{E}\left(\underset{pM\times 1}{\varepsilon_{i(p)}}\,\underset{1\times pM}{\varepsilon'_{i(p)}}\right) &= \underset{p\times p}{\mathbf{I}_p} \otimes \underset{M\times M}{\left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_v\right)} + \underset{p\times p}{\mathbf{J}_p} \otimes \underset{M\times M}{\boldsymbol{\Sigma}_\mu} = \\ &= \mathbf{E}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_v\right) + \bar{\mathbf{J}}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_v\right) + \mathbf{J}_p \otimes \boldsymbol{\Sigma}_\mu = \\ &= \mathbf{E}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_v\right) + \bar{\mathbf{J}}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_v\right) + \bar{\mathbf{J}}_p \otimes p\boldsymbol{\Sigma}_\mu = \\ &= \underset{p\times p}{\mathbf{E}_p} \otimes \underset{M\times M}{\left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_v\right)} + \underset{p\times p}{\bar{\mathbf{J}}_p} \otimes \underset{M\times M}{\left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_v + p\boldsymbol{\Sigma}_\mu\right)}, \end{aligned} \tag{7}$$

where $\mathbf{I}_p$ is an identity matrix of dimension $p$, $\mathbf{J}_p$ a matrix of ones of dimension $p$, $\mathbf{E}_p = \mathbf{I}_p - \overline{\mathbf{J}}_p$ and $\overline{\mathbf{J}}_p = \frac{\mathbf{J}_p}{p}$. Since $\mathbf{E}_p$ and $\overline{\mathbf{J}}_p$ are symmetric, idempotent and have orthogonal columns, the inverse of the variance-covariance matrix is

$$\underset{pM\times pM}{\boldsymbol{\Omega}_p^{-1}} = \underset{p\times p}{\mathbf{E}_p} \otimes \underset{M\times M}{\left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu\right)^{-1}} + \underset{p\times p}{\overline{\mathbf{J}}_p} \otimes \underset{M\times M}{\left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu + p\boldsymbol{\Sigma}_\mu\right)^{-1}}. \tag{8}$$

## 3.2. *GLS* ESTIMATION

If we assume that $\boldsymbol{\Sigma}_\mu$, $\boldsymbol{\Sigma}_\nu$ and $\boldsymbol{\Sigma}_u$ are known, we can write the *GLS* estimator for $\boldsymbol{\beta}$ as the problem of minimizing

$$\begin{aligned} &\sum_{p=1}^{T}\sum_{i\in I_p} \underset{1\times pM}{\boldsymbol{\varepsilon}'_{i(p)}} \underset{pM\times pM}{\boldsymbol{\Omega}_p^{-1}} \underset{pM\times 1}{\boldsymbol{\varepsilon}_{i(p)}} = \\ &= \sum_{p=1}^{T}\sum_{i\in I_p} \underset{1\times pM}{\boldsymbol{\varepsilon}'_{i(p)}} \underset{pM\times pM}{\left[\mathbf{E}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu\right)^{-1}\right]} \underset{pM\times 1}{\boldsymbol{\varepsilon}_{i(p)}} + \sum_{p=1}^{T}\sum_{i\in I_p} \underset{1\times pM}{\boldsymbol{\varepsilon}'_{i(p)}} \underset{pM\times pM}{\left[\overline{\mathbf{J}}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu + p\boldsymbol{\Sigma}_\mu\right)^{-1}\right]} \underset{pM\times 1}{\boldsymbol{\varepsilon}_{i(p)}}. \end{aligned} \tag{9}$$

If we apply *GLS* on the observations for the individuals observed $p$ times we obtain

$$\begin{aligned} \underset{K\times 1}{\boldsymbol{\beta}_p^{GLS}} &= \left[\sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\boldsymbol{\Omega}_p^{-1}} \underset{pM\times K}{\mathbf{X}_{i(p)}}\right]^{-1} \times \left[\sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\boldsymbol{\Omega}_p^{-1}} \underset{pM\times 1}{\mathbf{y}_{i(p)}}\right] = \\ &= \left[\sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\left[\mathbf{E}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu\right)^{-1}\right]} \underset{pM\times K}{\mathbf{X}_{i(p)}} + \sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\left[\overline{\mathbf{J}}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu + p\boldsymbol{\Sigma}_\mu\right)^{-1}\right]} \underset{pM\times K}{\mathbf{X}_{i(p)}}\right]^{-1} \times \\ &\times \left[\sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\left[\mathbf{E}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu\right)^{-1}\right]} \underset{pM\times 1}{\mathbf{y}_{i(p)}} + \sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\left[\overline{\mathbf{J}}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu + p\boldsymbol{\Sigma}_\mu\right)^{-1}\right]} \underset{pM\times 1}{\mathbf{y}_{i(p)}}\right], \end{aligned} \tag{10}$$

while the full *GLS* estimator is

$$\begin{aligned} \underset{K\times 1}{\boldsymbol{\beta}^{GLS}} &= \left[\sum_{p=1}^{T}\sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\boldsymbol{\Omega}_p^{-1}} \underset{pM\times K}{\mathbf{X}_{i(p)}}\right]^{-1} \times \left[\sum_{p=1}^{T}\sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\boldsymbol{\Omega}_p^{-1}} \underset{pM\times 1}{\mathbf{y}_{i(p)}}\right] = \\ &= \left[\sum_{p=1}^{T}\sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\left[\mathbf{E}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu\right)^{-1}\right]} \underset{pM\times K}{\mathbf{X}_{i(p)}} + \sum_{p=1}^{T}\sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\left[\overline{\mathbf{J}}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu + p\boldsymbol{\Sigma}_\mu\right)^{-1}\right]} \underset{pM\times K}{\mathbf{X}_{i(p)}}\right]^{-1} \times \\ &\times \left[\sum_{p=1}^{T}\sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\left[\mathbf{E}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu\right)^{-1}\right]} \underset{pM\times 1}{\mathbf{y}_{i(p)}} + \sum_{p=1}^{T}\sum_{i\in I_p} \underset{K\times pM}{\mathbf{X}'_{i(p)}} \underset{pM\times pM}{\left[\overline{\mathbf{J}}_p \otimes \left(\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_\nu + p\boldsymbol{\Sigma}_\mu\right)^{-1}\right]} \underset{pM\times 1}{\mathbf{y}_{i(p)}}\right]. \end{aligned} \tag{11}$$

## 3.3. ESTIMATION OF THE COVARIANCE MATRICES

The next step is to find an appropriate technique to estimate the three error component variance-covariance matrices of the two-way *SUR* system $\underset{M\times M}{\boldsymbol{\Sigma}_\mu}$, $\underset{M\times M}{\boldsymbol{\Sigma}_\nu}$ and $\underset{M\times M}{\boldsymbol{\Sigma}_u}$. This can be achieved adopting either the within-between procedure suggested by Biørn (2004) for the one-way *SUR* or the Quadratic Unbiased Estimator (*QUE)* procedure suggested by

Wansbeek and Kapteyn (1989) for the single equation case. In the following sub-sections we modify both procedures making them suitable for the two-way *SUR* system.

### 3.3.2. The QUE Procedure

The QUE procedure considers the FE residuals $\underset{n\times 1}{\mathbf{e}_m} \equiv \underset{n\times 1}{\mathbf{y}_m} - \underset{n\times k_m}{\mathbf{X}_m}\underset{k_m\times 1}{\boldsymbol{\beta}_m^{WT}}$ for the equation $m=1,\ldots,M$. If we assume that the matrix $\underset{n\times k_m}{\mathbf{X}_m}$ contains a vector of ones, we have to define the centered residuals $\underset{n\times 1}{\mathbf{f}_m} \equiv \underset{n\times n}{\mathbf{E}_n}\cdot\underset{n\times 1}{\mathbf{e}_m} = \underset{n\times 1}{\mathbf{e}_m - \overline{e}_m}$ , where $\mathbf{E}_n = \mathbf{I}_n - \overline{\mathbf{J}}_n$, with $\mathbf{I}_n$ being an identity matrix of dimension *n*, $\overline{\mathbf{J}}_n = \frac{\mathbf{J}_n}{n}$ and $\mathbf{J}_n$ a matrix of ones of dimension *n* (see Wansbeek and Kapteyn, 1989). Thus, we can obtain the adapted *QUE's* for $\sigma^2_{u_{mj}}$, $\sigma^2_{\mu_{mj}}$ and $\sigma^2_{\nu_{mj}}$ by equating

$$
\begin{aligned}
q_{n_{mj}} &\equiv \underset{1\times n}{\mathbf{f}'_j}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times 1}{\mathbf{f}_m}, \\
q_{N_{mj}} &\equiv \underset{1\times n}{\mathbf{f}'_j}\underset{n\times T}{\boldsymbol{\Delta}_\nu}\underset{T\times T}{\boldsymbol{\Delta}_T^{-1}}\underset{T\times n}{\boldsymbol{\Delta}'_\nu}\underset{n\times 1}{\mathbf{f}_m}, \\
q_{T_{mj}} &\equiv \underset{1\times n}{\mathbf{f}'_j}\underset{n\times N}{\boldsymbol{\Delta}_\mu}\underset{N\times N}{\boldsymbol{\Delta}_N^{-1}}\underset{N\times n}{\boldsymbol{\Delta}'_\mu}\underset{n\times 1}{\mathbf{f}_m}
\end{aligned}
\tag{12}
$$

where $\underset{n\times n}{\mathbf{Q}_{[\Delta]}}$ is the projection matrix onto the null-space of $\underset{n\times(N+T)}{\boldsymbol{\Delta}}$ (see Appendix A), to their expected values (see Appendix C)

$$
\begin{aligned}
\mathrm{E}\left(q_{n_{mj}}\right) &= \left(n-T-N+1+k_{mj}-k_m-k_j\right)\cdot\sigma^2_{u_{mj}}, \\
\mathrm{E}\left(q_{N_{mj}}\right) &= \left(T+k_{N_{mj}}-k_{0_{mj}}-1\right)\cdot\sigma^2_{u_{mj}}+\left(T-\lambda_\mu\right)\cdot\sigma^2_{\mu_{mj}}+\left(n-\lambda_\nu\right)\cdot\sigma^2_{\nu_{mj}}, \\
\mathrm{E}\left(q_{T_{mj}}\right) &= \left(N+k_{T_{mj}}-k_{0_{mj}}-1\right)\cdot\sigma^2_{u_{mj}}+\left(n-\lambda_\mu\right)\cdot\sigma^2_{\mu_{mj}}+\left(N-\lambda_\nu\right)\cdot\sigma^2_{\nu_{mj}},
\end{aligned}
\tag{13}
$$

where we have defined $k_{mj} \equiv \mathrm{tr}((\mathbf{X}'_m\mathbf{Q}_{[\Delta]}\mathbf{X}_m)^{-1}\mathbf{X}'_m\mathbf{Q}_{[\Delta]}\mathbf{X}_j(\mathbf{X}'_j\mathbf{Q}_{[\Delta]}\mathbf{X}_j)^{-1}\mathbf{X}'_j\mathbf{Q}_{[\Delta]}\mathbf{X}_m)$, $k_{N_{mj}} \equiv \mathrm{tr}((\mathbf{X}'_m\mathbf{Q}_{[\Delta]}\mathbf{X}_m)^{-1}\mathbf{X}'_m\mathbf{Q}_{[\Delta]}\mathbf{X}_j(\mathbf{X}'_j\mathbf{Q}_{[\Delta]}\mathbf{X}_j)^{-1}\mathbf{X}'_j\boldsymbol{\Delta}_\nu\boldsymbol{\Delta}_T^{-1}\boldsymbol{\Delta}'_\nu\mathbf{X}_m)$, $\lambda_\mu \equiv \frac{\boldsymbol{\iota}'_n\boldsymbol{\Delta}_\mu\boldsymbol{\Delta}'_\mu\boldsymbol{\iota}_n}{n} = \frac{\sum_{i=1}^N T_i^2}{n}$, $k_{T_{mj}} \equiv \mathrm{tr}((\mathbf{X}'_m\mathbf{Q}_{[\Delta]}\mathbf{X}_m)^{-1}\mathbf{X}'_m\mathbf{Q}_{[\Delta]}\mathbf{X}_j(\mathbf{X}'_j\mathbf{Q}_{[\Delta]}\mathbf{X}_j)^{-1}\mathbf{X}'_j\boldsymbol{\Delta}_\mu\boldsymbol{\Delta}_N^{-1}\boldsymbol{\Delta}'_\mu\mathbf{X}_m)$, $\lambda_\nu \equiv \frac{\boldsymbol{\iota}'_n\boldsymbol{\Delta}_\nu\boldsymbol{\Delta}'_\nu\boldsymbol{\iota}_n}{n} = \frac{\sum_{t=1}^T N_t^2}{n}$ and $k_{0_{mj}} \equiv \frac{\boldsymbol{\iota}'_n\mathbf{X}_m(\mathbf{X}'_m\mathbf{Q}_{[\Delta]}\mathbf{X}_m)^{-1}\mathbf{X}'_m\mathbf{Q}_{[\Delta]}\mathbf{X}_j(\mathbf{X}'_j\mathbf{Q}_{[\Delta]}\mathbf{X}_j)^{-1}\mathbf{X}'_j\boldsymbol{\iota}_n}{n}$.

The difference with respect to the single equation case is that the centered residuals $\mathbf{f}_j$ and $\mathbf{f}_m$ may refer to different equations. Since $k_{jm} = k_{mj}$, $k_{N_{jm}} = k_{N_{mj}}$, $k_{T_{jm}} = k_{T_{mj}}$ and $k_{0_{jm}} = k_{0_{mj}}$ obviously we have $\sigma^2_{u_{jm}} = \sigma^2_{u_{mj}}$, $\sigma^2_{\mu_{jm}} = \sigma^2_{\mu_{mj}}$ and $\sigma^2_{\nu_{jm}} = \sigma^2_{\nu_{mj}}$.

### 3.3.2. The Within-between Procedure

As the *QUE* procedure, the within-between procedure considers the FE residuals $\underset{M\times 1}{\mathbf{e}_{it}} \equiv \underset{M\times 1}{\mathbf{y}_{it}} - \underset{M\times K}{\mathbf{X}_{it}} \underset{K\times 1}{\boldsymbol{\beta}^{WT}}$ for the individual $i$ in period $t$. As before, if we assume that the matrix $\underset{M\times K}{\mathbf{X}_{it}}$ always contains $M$ vectors of ones (a vector of ones for each equation $m$), we have to define the consistent centered residuals $\underset{M\times 1}{\mathbf{f}_{it}} = \underset{M\times 1}{\mathbf{e}_{it}} - \underset{M\times 1}{\overline{\mathbf{e}}}$, where $\overline{e}_m = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T_i} e_{mit}}{n} = \frac{\sum_{t=1}^{T}\sum_{i=1}^{N_t} e_{mit}}{n}$[4]. Therefore the $M\times M$ matrices of within individuals, between individuals and between times[5] (co)variations in the $\mathbf{f}$'s of the different equations are the following:

$$
\begin{aligned}
\underset{M\times M}{\mathbf{W}_f} &= \sum_{i=1}^{N}\sum_{t=1}^{T_i} \underset{M\times 1}{\left(\mathbf{f}_{it} - \overline{\mathbf{f}}_{i\bullet} - \overline{\mathbf{f}}_{\bullet t}\right)} \underset{1\times M}{\left(\mathbf{f}_{it} - \overline{\mathbf{f}}_{i\bullet} - \overline{\mathbf{f}}_{\bullet t}\right)'},\\
\underset{M\times M}{\mathbf{B}_f^C} &= \sum_{i=1}^{N} T_i \underset{M\times 1}{\left(\overline{\mathbf{f}}_{i\bullet} - \overline{\mathbf{f}}\right)} \underset{1\times M}{\left(\overline{\mathbf{f}}_{i\bullet} - \overline{\mathbf{f}}\right)'},\\
\underset{M\times M}{\mathbf{B}_f^T} &= \sum_{t=1}^{T} N_t \underset{M\times 1}{\left(\overline{\mathbf{f}}_{\bullet t} - \overline{\mathbf{f}}\right)} \underset{1\times M}{\left(\overline{\mathbf{f}}_{\bullet t} - \overline{\mathbf{f}}\right)'},
\end{aligned}
\tag{14}
$$

where for each equation $m$ we have $\overline{f}_{mi\cdot} = \frac{\sum_{t=1}^{T_i} f_{mit}}{T_i}$, $\overline{f}_{m\cdot t} = \frac{\sum_{i=1}^{N_t} f_{mit}}{N_t}$ and $\overline{f}_m = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T_i} f_{mit}}{n} = \frac{\sum_{i=1}^{N}(T_i \cdot \overline{f}_{mi\cdot})}{n}$ or $\overline{f}_m = \frac{\sum_{t=1}^{T}\sum_{i=1}^{N_t} f_{mit}}{n} = \frac{\sum_{t=1}^{T}(N_t \cdot \overline{f}_{m\cdot t})}{n}$. The between times (co)variation $\mathbf{B}_f^T$ is needed to adapt the Biørn's (2004) procedure to the two-way *ECM*. In Appendix D equation A(21) allows us to conclude that

$$
\begin{aligned}
\underset{M\times M}{\hat{\boldsymbol{\Sigma}}_u} &= \frac{\mathbf{W}_f}{n-N-T},\\
\underset{M\times M}{\hat{\boldsymbol{\Sigma}}_\mu} &= \frac{\mathbf{B}_f^C - (N-1)\cdot\hat{\boldsymbol{\Sigma}}_u}{n - \sum_{i=1}^{N}\frac{T_i^2}{n}},\\
\underset{M\times M}{\hat{\boldsymbol{\Sigma}}_\nu} &= \frac{\mathbf{B}_f^T - (T-1)\cdot\hat{\boldsymbol{\Sigma}}_u}{n - \sum_{t=1}^{T}\frac{N_t^2}{n}}
\end{aligned}
\tag{15}
$$

are consistent and unbiased estimators of $\underset{M\times M}{\boldsymbol{\Sigma}_\mu}$, $\underset{M\times M}{\boldsymbol{\Sigma}_\nu}$ and $\underset{M\times M}{\boldsymbol{\Sigma}_u}$.

---

[4] To obtain consistent estimates of the variance-covariance matrices, we need consistent residuals (Biørn, 2004). In the two-way case, since the *QUE* procedure is based on the FE residuals, for coherence we use the same FE residuals, and then the corresponding $M\times 1$ consistent centered residuals $\mathbf{f}_{it}$, also in the within-between procedure.

[5] Kang (1985) uses the between time periods estimator to build the equivalent tests for the two-way error components model.

## 4. Simulation Results

In order to analyze the performances of the proposed techniques, in this section we develop a simple simulation on a three equations system ($M = 3$). We assume an unbalanced panel with a large number of individuals ($N = 4000$) extended over a rather long time period ($T = 8$). This should mimic a real world situation of a large unbalanced panel for which the two-way *SUR* system is the appropriate model. The simulated model is

$$
\begin{aligned}
y_1 &= \beta_{10} + \beta_{11} \cdot x_1 + \beta_{12} \cdot x_2 + \varepsilon_1, \\
y_2 &= \beta_{20} + \beta_{21} \cdot x_1 + \beta_{22} \cdot x_2 + \beta_{23} \cdot x_3 + \varepsilon_2, \\
y_3 &= \beta_{30} + \beta_{32} \cdot x_2 + \beta_{33} \cdot x_3 + \varepsilon_3,
\end{aligned}
$$

where $\boldsymbol{\beta}_1 = (15, 6, -3)'$, $\boldsymbol{\beta}_2 = (10, -3, 8, -2)'$ and $\boldsymbol{\beta}_3 = (20, -2, 5)'$, which imply the cross equations restrictions $\beta_{12} = \beta_{21}$ and $\beta_{23} = \beta_{32}$. We consider the following variance-covariance matrices[6]

$$
\boldsymbol{\Sigma}_\mu = \begin{bmatrix} 968.5 & -88.2 & 21.5 \\ & 725.2 & -55.0 \\ & & 513.4 \end{bmatrix}, \boldsymbol{\Sigma}_\nu = \begin{bmatrix} 87.52 & 15.81 & -4.65 \\ & 79.97 & 5.89 \\ & & 53.22 \end{bmatrix} \text{ and } \boldsymbol{\Sigma}_u = \begin{bmatrix} 86.28 & 17.39 & -5.94 \\ & 77.98 & 7.53 \\ & & 56.46 \end{bmatrix}.
$$

Finally, the scalars $x_{kit}$ have been generated according to the scheme introduced by Nerlove (1971) and used, among others, by Baltagi (1981) and Wansbeek and Kapteyn (1989)

$$
x_{kit} = 0.1 \cdot t + 0.5 \cdot x_{kit-1} + \omega_{kit},
$$

with $\omega_{kit}$ following the uniform distribution $[-\frac{1}{2}, \frac{1}{2}]$ and $x_{ki0} = 5 + 10 \cdot \omega_{ki0}$.

In order to construct the unbalanced panel, we have adopted the procedure currently used for rotating panels, in which we have approximately the same number of individuals every year: a fixed percentage of individuals (20% in our case[7]) is replaced each year, but they can re-enter the sample in the following years. Thus, for each group $p$ we have the following number of individuals: $\widetilde{N}_1 = 962$, $\widetilde{N}_2 = 769$, $\widetilde{N}_3 = 615$, $\widetilde{N}_4 = 492$, $\widetilde{N}_5 = 394$, $\widetilde{N}_6 = 315$, $\widetilde{N}_7 = 252$ and $\widetilde{N}_8 = 201$ (and then $n = 13545$).

The results of a 150-run simulation are shown in table 1.

The covariance matrices estimated through the one-way within-between procedure are

[6] The three variance-covariance matrices have been randomly generated using the *sprandsym* command in *MatLab*, that produces positive-definite symmetric matrices with all non-zero entries.
[7] Also in Wansbeek and Kapteyn (1989) each period 20% of the households in the panels is removed randomly.

**TABLE 1** Simulation Results: Means of the Estimated Parameters and Average Variances of the Error Components

| | True value | FE one-way | FE two-way | RE one-way (GLS) | RE one-way (ML) | RE two-way (GLS) | RE two-way (ML) | SUR WB one-way (GLS) | SUR QUE two-way (GLS) | SUR WB two-way (GLS) |
|---|---|---|---|---|---|---|---|---|---|---|
| $\beta_{10}$ | 15 | | | 15.0114 (12.1377) | 15.0114 (12.1371) | 15.0304 (12.0983) | 15.0304 (12.0981) | 15.0031 (12.2501) | 15.0033 (12.2478) | 15.0023 (12.2584) |
| | | | | (0.5883) | (0.5883) | (3.2396) | (3.0494) | (0.5800) | (0.5825) | (0.5932) |
| $\beta_{11}$ | 6 | 6.0057 (0.0280) | 6.0048 (0.0157) | 6.0081 (0.0221) | 6.0082 (0.0221) | 6.0046 (0.0138) | 6.0047 (0.0137) | 6.0047 (0.0142) | 6.0051 (0.0141) | 6.0059 (0.0142) |
| | | (0.1786) | (0.1270) | (0.1652) | (0.1652) | (0.1219) | (0.1219) | (0.1417) | (0.1416) | (0.1486) |
| $\beta_{12}$ | -3 | -3.0085 (0.0341) | -3.0158 (0.0174) | -3.0078 (0.0251) | -3.0078 (0.0251) | -3.0143 (0.0147) | -3.0144 (0.0147) | -3.0000 (0.0099) | -3.0004 (0.0099) | -3.0006 (0.0097) |
| | | (0.1787) | (0.1272) | (0.1654) | (0.1654) | (0.1220) | (0.1220) | (0.1076) | (0.1076) | (0.1127) |
| $\sigma^2_{\mu_{11}}$ | 968.5 | | | 956.6491 | 956.8983 | 967.8692 | 968.0000 | 957.0824 | 967.8692 | 976.2906 |
| $\sigma^2_{v_{11}}$ | 87.52 | | | | | 86.9819 | 76.6019 | | 86.9819 | 87.2374 |
| $\sigma^2_{u_{11}}$ | 86.28 | 173.2588 | 86.3215 | 173.2407 | 173.2091 | 86.3215 | 86.2966 | 173.2044 | 86.3215 | 108.0565 |
| $\beta_{20}$ | 10 | | | 10.3448 (10.5953) | 10.3448 (10.5957) | 10.3406 (10.5691) | 10.3406 (10.5694) | 10.3433 (10.5928) | 10.3441 (10.5898) | 10.3434 (10.5934) |
| | | | | (0.5316) | (0.5317) | (3.0779) | (2.8912) | (0.5213) | (0.5239) | (0.5352) |
| $\beta_{21}$ | -3 | -2.9913 (0.0339) | 2.9860 (0.0173) | -2.9905 (0.0281) | -2.9905 (0.0281) | -2.9882 (0.0160) | -2.9882 (0.0160) | -3.0000 (0.0099) | -3.0004 (0.0099) | -3.0006 (0.0097) |
| | | (0.1824) | (0.1296) | (0.1653) | (0.1653) | (0.1229) | (0.1229) | (0.1076) | (0.1076) | (0.1127) |
| $\beta_{22}$ | 8 | 7.9877 (0.0347) | 7.9847 (0.0129) | 7.9903 (0.0238) | 7.9904 (0.0238) | 7.9878 (0.0113) | 7.9878 (0.0113) | 7.9973 (0.0139) | 7.9974 (0.0138) | 7.9977 (0.0136) |
| | | (0.1826) | (0.1297) | (0.1655) | (0.1655) | (0.1230) | (0.1230) | (0.1393) | (0.1392) | (0.1461) |
| $\beta_{23}$ | -2 | -2.0094 (0.0344) | -2.0061 (0.0167) | -2.0080 (0.0261) | -2.0080 (0.0261) | -2.0051 (0.0153) | -2.0051 (0.0153) | -2.0046 (0.0093) | -2.0047 (0.0092) | -2.0044 (0.0091) |
| | | (0.1824) | (0.1296) | (0.1653) | (0.1653) | (0.1229) | (0.1229) | (0.0970) | (0.0970) | (0.1013) |
| $\sigma^2_{\mu_{22}}$ | 725.2 | | | 713.3330 | 713.8177 | 723.5082 | 723.6495 | 713.7265 | 723.5082 | 731.2457 |
| $\sigma^2_{v_{22}}$ | 79.97 | | | | | 79.4254 | 69.8026 | | 79.4254 | 79.6144 |
| $\sigma^2_{u_{22}}$ | 77.98 | 157.2736 | 77.9915 | 157.2571 | 157.2128 | 77.9915 | 77.9612 | 157.2077 | 77.9915 | 97.7202 |
| $\beta_{30}$ | 20 | | | 19.8231 (8.5881) | 19.8230 (8.5881) | 19.8121 (8.5321) | 19.8120 (8.5323) | 19.8189 (8.5543) | 19.8189 (8.5509) | 19.8188 (8.5488) |
| | | | | (0.4405) | (0.4404) | (2.5527) | (2.4030) | (0.4342) | (0.4361) | (0.4446) |
| $\beta_{32}$ | -2 | -2.0131 (0.0221) | -2.0049 (0.0115) | -2.0087 (0.0187) | -2.0087 (0.0187) | -2.0029 (0.0101) | -2.0029 (0.0101) | -2.0046 (0.0093) | -2.0047 (0.0092) | -2.0044 (0.0091) |
| | | (0.1431) | (0.1029) | (0.1308) | (0.1308) | (0.0979) | (0.0979) | (0.0970) | (0.0970) | (0.1013) |
| $\beta_{33}$ | 5 | 5.0117 (0.0168) | 5.0089 (0.0096) | 5.0052 (0.0131) | 5.0053 (0.0131) | 5.0050 (0.0084) | 5.0050 (0.0084) | 5.0031 (0.0100) | 5.0032 (0.0100) | 5.0028 (0.0101) |
| | | (0.1430) | (0.1028) | (0.1306) | (0.1306) | (0.0978) | (0.0978) | (0.1172) | (0.1171) | (0.1224) |
| $\sigma^2_{\mu_{33}}$ | 513.4 | | | 506.5589 | 506.3825 | 513.6266 | 513.5890 | 506.7917 | 513.6266 | 518.9794 |
| $\sigma^2_{v_{33}}$ | 53.22 | | | | | 54.0729 | 47.6716 | | 54.0729 | 54.2076 |
| $\sigma^2_{u_{33}}$ | 56.46 | 110.9766 | 56.5285 | 110.9650 | 110.9453 | 56.5285 | 56.5108 | 110.9417 | 56.5285 | 69.6443 |

The numbers in brackets at the right in the first line are the variances of the estimated parameters and the numbers in brackets in the second line are the average standard errors.

$$\boldsymbol{\Sigma}_{\mu}=\begin{bmatrix}957.0824 & -87.8098 & 22.7150\\ & 713.7265 & -55.4356\\ & & 506.7917\end{bmatrix} \text{ and } \boldsymbol{\Sigma}_{u}=\begin{bmatrix}173.2044 & 32.7289 & -10.4171\\ & 157.2077 & 13.3276\\ & & 110.9417\end{bmatrix},$$

while those estimated through the two-way *QUE* procedure are

$$\boldsymbol{\Sigma}_{\mu}=\begin{bmatrix}967.8692 & -85.9968 & 22.2379\\ & 723.5082 & -54.7695\\ & & 513.6266\end{bmatrix}, \boldsymbol{\Sigma}_{\nu}=\begin{bmatrix}86.9819 & 15.8670 & -5.0449\\ & 79.4254 & 5.9641\\ & & 54.0729\end{bmatrix} \text{ and } \boldsymbol{\Sigma}_{u}=\begin{bmatrix}86.3215 & 17.3926 & -5.8825\\ & 77.9915 & 7.5897\\ & & 56.5285\end{bmatrix}$$

and those estimated through the two-way within-between procedure are

$$\boldsymbol{\Sigma}_{\mu}=\begin{bmatrix}976.2906 & -84.4980 & 21.8235\\ & 731.2457 & -54.2465\\ & & 518.9794\end{bmatrix}, \boldsymbol{\Sigma}_{\nu}=\begin{bmatrix}87.2374 & 15.8408 & -5.0379\\ & 79.6144 & 5.9480\\ & & 54.2076\end{bmatrix} \text{ and } \boldsymbol{\Sigma}_{u}=\begin{bmatrix}108.0565 & 21.4853 & -7.3919\\ & 97.7202 & 9.2659\\ & & 69.6443\end{bmatrix}.$$

In table 1 we present the means and the variances of the estimated parameters, the average standard errors and the average variances of the error term components.

While to estimate all the single-equation version of the model (FE one-way, RE *GLS* one-way, RE *ML* one-way and RE *ML* two-way) we used the commands built in the econometric software *TSP* version 5.0, we computed the FE two-way, the RE *GLS* two-way, the *SUR GLS* one-way and the two versions of the *SUR GLS* two-way–adopting either the *QUE* procedure or the within-between procedure–through a specific routine written in *TSP* version 5.0.

The advantages of adopting a two-way specification for analyzing our unbalanced panel (through either single-equations or a system of equations) are clear when we analyze the estimated variance-covariance matrices. For example, all the one-way techniques produce biased estimates for the variances of the idiosyncratic error term $\sigma^2_{u_{mm}}$ : when the time dimension in the data is relevant, two-way techniques produce better estimates since they allow to disentangle the time component from the remainder error term.

This is true also comparing the *SUR GLS* one-way with both versions of the *SUR GLS* two-way, even though, between the two procedures we have proposed in the previous sections, the *QUE* turns out to be more precise than the within-between.

In terms of parameter estimates, it is clear the system of equation techniques perform better than the single equation ones, although the gain in efficiency of the *SUR GLS* two-way may become more relevant for a panel with a longer time dimension.

In general, we can conclude that all the estimates obtained are consistent, but the *SUR GLS* two-way procedures guarantee a gain in efficiency. Moreover the parameters and the variances estimated with the *SUR GLS* two-way adopting the *QUE* procedure tend to be closer to the true values.

## APPENDIX A: FIXED EFFECTS ESTIMATION ASSUMPTIONS

Adapting the formulation proposed by Wooldridge (2002) for the one-way case to the two-way case, the assumptions related to the FE estimation are the following.

**FE.1** **Strict exogeneity**: the set of explanatory variables in each time period $\underset{N_t k}{\mathbf{x}_{t\{T\}}}$ is uncorrelated with the idiosyncratic error $u_{it}$ and the set of explanatory variables for each individual $\underset{T_i k}{\mathbf{x}_{i\{N\}}}$ is also uncorrelated with the same idiosyncratic error $u_{it}$.

This means that

$$\mathrm{E}\left(u_{it}\left|\underset{1\times nk}{\mathbf{x}}, \mu_i, \nu_t\right.\right) = \mathrm{E}\left(u_{it}\left|\underset{1\times N_t k}{\mathbf{x}_{t\{T\}}}, \mu_i, \nu_t\right.\right) = \mathrm{E}\left(u_{it}\left|\underset{1\times T_i k}{\mathbf{x}_{i\{N\}}}, \mu_i, \nu_t\right.\right) = 0,$$

where $\underset{1\times kn}{\mathbf{x}} \equiv (\mathbf{x}_{11}, \ldots, \mathbf{x}_{1T_1}, \mathbf{x}_{21}, \ldots, \mathbf{x}_{2T_2}, \ldots, \mathbf{x}_{NT_N})$ or $\underset{1\times kn}{\mathbf{x}} \equiv (\mathbf{x}_{11}, \ldots, \mathbf{x}_{N_1 1}, \mathbf{x}_{12}, \ldots, \mathbf{x}_{N_2 2}, \ldots, \mathbf{x}_{N_T T})$, $\underset{1\times kN_t}{\mathbf{x}_{t\{T\}}} \equiv (\mathbf{x}_{1t}, \mathbf{x}_{2t}, \ldots, \mathbf{x}_{N_t t})$ and $\underset{1\times kT_i}{\mathbf{x}_{i\{N\}}} \equiv (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \ldots, \mathbf{x}_{iT_i})$.

Therefore we have strict exogeneity of $\{\underset{1\times k}{\mathbf{x}_{it}} : i = 1, \ldots, N; t = 1, \ldots, T_i\}$ or $\{\underset{1\times k}{\mathbf{x}_{it}} : t = 1, \ldots, T; i = 1, \ldots, N_t\}$ conditional on the unobserved effects.

**FE.2** **Consistency**: the within estimator is asymptotically well behaved, in the sense that the "adjusted" outer product matrix has the appropriate rank.

The idea of estimating $\underset{k\times 1}{\boldsymbol{\beta}}$ under assumption FE.1 is to transform the equation to eliminate the unobserved effects $\mu_i$ and $\nu_t$. When we have an unbalanced panel, the simple projection and transformation results no longer hold. Therefore, following Wansbeek and Kapteyn (1989), we order the data on the $N$ individuals in $T$ consecutive sets and we define the following matrices:

$$\underset{n\times T}{\overline{\boldsymbol{\Delta}}} \equiv \underset{n\times T}{\boldsymbol{\Delta}_\nu} - \underset{n\times N}{\boldsymbol{\Delta}_\mu}\, \underset{N\times N}{\boldsymbol{\Delta}_N^{-1}}\, \underset{N\times T}{\boldsymbol{\Delta}'_{TN}} = \left(\underset{n\times n}{\mathbf{I}_n} - \underset{n\times N}{\boldsymbol{\Delta}_\mu}\, \underset{N\times N}{\boldsymbol{\Delta}_N^{-1}}\, \underset{N\times n}{\boldsymbol{\Delta}'_\mu}\right) \cdot \underset{n\times T}{\boldsymbol{\Delta}_\nu} = \left(\underset{n\times n}{\mathbf{I}_n} - \underset{n\times n}{\mathbf{P}_{[\Delta_\mu]}}\right) \cdot \underset{n\times T}{\boldsymbol{\Delta}_\nu} = \underset{n\times n}{\mathbf{Q}_{[\Delta_\mu]}}\, \underset{n\times T}{\boldsymbol{\Delta}_\nu},$$

$$\underset{T\times T}{\mathbf{Q}} \equiv \underset{T\times T}{\boldsymbol{\Delta}_T} - \underset{T\times N}{\boldsymbol{\Delta}_{TN}}\, \underset{N\times N}{\boldsymbol{\Delta}_N^{-1}}\, \underset{N\times T}{\boldsymbol{\Delta}'_{TN}} = \underset{T\times n}{\boldsymbol{\Delta}'_\nu} \cdot \left(\underset{n\times T}{\boldsymbol{\Delta}_\nu} - \underset{n\times N}{\boldsymbol{\Delta}_\mu}\, \underset{N\times N}{\boldsymbol{\Delta}_N^{-1}}\, \underset{N\times T}{\boldsymbol{\Delta}'_{TN}}\right) = \underset{T\times n}{\boldsymbol{\Delta}'_\nu} \cdot \underset{n\times T}{\overline{\boldsymbol{\Delta}}} = \underset{T\times n}{\boldsymbol{\Delta}'_\nu}\, \underset{n\times n}{\mathbf{Q}_{[\Delta_1]}}\, \underset{n\times T}{\boldsymbol{\Delta}_\nu}.$$

Hence the projection matrix onto the null-space of $\boldsymbol{\Delta}$ is:

$$\underset{n\times n}{\mathbf{Q}_{[\Delta]}} = \left( \underset{n\times n}{\mathbf{I}_n} - \underset{n\times N}{\boldsymbol{\Delta}_\mu} \underset{N\times N}{\boldsymbol{\Delta}_N^{-1}} \underset{N\times n}{\boldsymbol{\Delta}'_\mu} \right) - \underset{n\times T}{\overline{\boldsymbol{\Delta}}} \underset{T\times T}{\mathbf{Q}^-} \underset{T\times n}{\overline{\boldsymbol{\Delta}}'} = \underset{n\times n}{\mathbf{Q}_{[\Delta_\mu]}} - \underset{n\times n}{\mathbf{Q}_{[\Delta_\mu]}} \underset{n\times T}{\boldsymbol{\Delta}_\nu} \underset{T\times T}{\mathbf{Q}^-} \underset{T\times n}{\boldsymbol{\Delta}'_\nu} \underset{n\times n}{\mathbf{Q}_{[\Delta_\mu]}}.$$

Given the previous within transformation, we can perform the regression

$$\underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times 1}{\mathbf{y}} = \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times k}{\mathbf{X}} \underset{k\times 1}{\boldsymbol{\beta}} + \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times 1}{\mathbf{u}}. \tag{A.1}$$

Finally, in order to ensure that the FE estimator is well behaved asymptotically, we need the following standard rank condition:

$$\text{rank}\left( \underset{k\times n}{\mathbf{X}'} \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times k}{\mathbf{X}} \right) = k. \tag{A.2}$$

If $\underset{1\times k}{\mathbf{x}_{it}}$ contains an element that does not vary over time for any $i$, then the corresponding element in $\underset{1\times k}{\tilde{\mathbf{x}}_{it}}$ from the matrix $\underset{n\times k}{\tilde{\mathbf{X}}} = \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times k}{\mathbf{X}}$ is identically zero for all $t$ and for any draw from the cross section. Since $\underset{n\times k}{\tilde{\mathbf{X}}}$ contains a column of zeros, assumption FE.2 cannot be true. Thus, assumption FE.2 shows explicitly why time-constant variables are not allowed in FE.

**FE.3** **Homoscedasticity and no serial correlation**: the conditional variance-covariance matrix of the idiosyncratic error terms $u_{it}$ coincides with the unconditional one, and it is characterized by constant variances and zero covariances

$$\mathrm{E}\left( \underset{n\times 1}{\mathbf{u}} \underset{1\times n}{\mathbf{u}'} \middle| \underset{1\times nk}{\mathbf{x}}, \mu_i, \nu_t \right) = \sigma_u^2 \underset{n\times n}{\mathbf{I}_n}.$$

The FE estimator can be expressed as

$$\underset{k\times 1}{\boldsymbol{\beta}^{WT}} = \left( \underset{n\times k}{\mathbf{X}'} \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times k}{\mathbf{X}} \right)^{-1} \left( \underset{n\times k}{\mathbf{X}'} \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times 1}{\mathbf{y}} \right). \tag{A.3}$$

Without further assumptions, the FE estimator is not necessarily the most efficient estimator based on assumption FE.1. Since $\mathrm{E}(\mathbf{u}|\mathbf{x}, \mu_i, \nu_t) = \mathbf{0}$ by assumption FE.1, assumption FE.3 is the same as saying $\mathrm{Var}(\mathbf{u}|\mathbf{x}, \mu_i, \nu_t) = \sigma_u^2 \mathbf{I}_n$ if assumption FE.1 also holds. It is useful to think of assumption FE.3 as having two parts. The first is that $\mathrm{E}(\mathbf{uu}'|\mathbf{x}, \mu_i, \nu_t) = \mathrm{E}(\mathbf{uu}')$, which is standard in system estimation contexts. The second is that the unconditional variance matrix $\mathrm{E}(\mathbf{uu}')$ has the special form $\sigma_u^2 \mathbf{I}_n$. This implies that the idiosyncratic errors $u_{it}$ have a constant variance across $i$ and across $t$ $\mathrm{E}(u_{it}^2) = \sigma_u^2$ and are serially uncorrelated $\mathrm{E}(u_{it}u_{js}) = 0$ for all $i = j, t \neq s$, all $i \neq j, t = s$ and all $i \neq j, t \neq s$.

## APPENDIX B: RANDOM EFFECTS ESTIMATION ASSUMPTIONS

Adapting again the formulation proposed by Wooldridge (2002) to the two-way case, the assumptions related to the RE estimation are the following.

**RE.1.a Strict exogeneity**: the set of explanatory variables in each time period $\underset{N_t k}{\mathbf{x}_{t\{T\}}}$ is uncorrelated with the idiosyncratic error $u_{it}$ and the set of explanatory variables for each individual $\underset{T_i k}{\mathbf{x}_{i\{N\}}}$ is also uncorrelated with the same idiosyncratic error $u_{it}$.

This means that

$$\mathrm{E}\left(u_{it}\left|\underset{1\times nk}{\mathbf{x}}, \mu_i, \nu_t\right.\right) = \mathrm{E}\left(u_{it}\left|\underset{1\times N_t k}{\mathbf{x}_{t\{T\}}}, \mu_i, \nu_t\right.\right) = \mathrm{E}\left(u_{it}\left|\underset{1\times T_i k}{\mathbf{x}_{i\{N\}}}, \mu_i, \nu_t\right.\right) = 0,$$

where $\underset{1\times kn}{\mathbf{x}}$, $\underset{1\times kN_t}{\mathbf{x}_{t\{T\}}}$ and $\underset{1\times kT_i}{\mathbf{x}_{i\{N\}}}$ have been already defined in Appendix A.

**RE.1.b Orthogonality** between $\mu_i$ and each $\underset{1\times T_i k}{\mathbf{x}_{i\{N\}}}$:

$$\mathrm{E}\left(\mu_i\left|\underset{1\times T_i k}{\mathbf{x}_{i\{N\}}}\right.\right) = \mathrm{E}\left(\mu_i\right) = 0.$$

**RE.1.c Orthogonality** between $\nu_t$ and each $\underset{1\times kN_t}{\mathbf{x}_{t\{T\}}} \equiv \left(\mathbf{x}_{1t}, \mathbf{x}_{2t}, \ldots, \mathbf{x}_{N_t t}\right)$:

$$\mathrm{E}\left(\nu_t\left|\underset{1\times N_t k}{\mathbf{x}_{t\{T\}}}\right.\right) = \mathrm{E}\left(\nu_t\right) = 0.$$

While assumption RE.1.a is identical to assumption FE.1, the key difference with respect to the FE case is that we assume also RE.1.b and RE.1.c. In other words, while for FE analysis $\mathrm{E}(\mu_i|\underset{1\times T_i k}{\mathbf{x}_{i\{N\}}})$ is allowed to be any function of $\mathbf{x}_{i\{N\}}$ and $\mathrm{E}(\nu_t|\underset{1\times kN_t}{\mathbf{x}_{t\{T\}}})$ any function of $\mathbf{x}_{i\{T\}}$, for the RE analysis this is not allowed. In fact assumptions RE.1.b and RE.1.c allow to derive the traditional asymptotic variance for the RE estimator.

**RE.2** **Consistency**: the RE *GLS* estimator is consistent, in the sense that the weighted outer product matrix has the appropriate rank.

A standard RE analysis adds assumptions on the idiosyncratic errors that give $\underset{n\times n}{\mathbf{\Omega}}$ a special form. The first assumption is that the idiosyncratic errors $u_{it}$ have a constant unconditional variance across $i$ and across $t$

$$\mathrm{E}\left(u_{it}^2\right)=\sigma_u^2. \tag{A.4}$$

The second assumption is that the idiosyncratic errors are serially uncorrelated

$$\mathrm{E}\left(u_{it}u_{js}\right)=0 \;\left(\text{all } i=j,t\neq s \text{ and all } i\neq j,t=s \text{ and all } i\neq j,t\neq s\right). \tag{A.5}$$

Under these two assumptions (special to RE), we can derive the variances and covariances of the elements of $\underset{n\times 1}{\boldsymbol{\varepsilon}}$. Under assumption RE.1.a we have $\mathrm{E}\left(\mu_i u_{it}\right)=0$ and $\mathrm{E}\left(\nu_t u_{it}\right)=0$ and given $\mathrm{E}\left(\mu_i\nu_t\right)=0$ we have

$$\mathrm{E}\left(\varepsilon_{it}^2\right)=\sigma_\mu^2+\sigma_\nu^2+\sigma_u^2.$$

Also for $t\neq s$ we have

$$\mathrm{E}\left(\varepsilon_{it}\varepsilon_{is}\right)=\mathrm{E}\left[\left(\mu_i+\nu_t+u_{it}\right)\left(\mu_i+\nu_s+u_{is}\right)\right]=\sigma_\mu^2,$$

and for $i\neq j$ we have

$$\mathrm{E}\left(\varepsilon_{it}\varepsilon_{jt}\right)=\mathrm{E}\left[\left(\mu_i+\nu_t+u_{it}\right)\left(\mu_j+\nu_t+u_{jt}\right)\right]=\sigma_\nu^2.$$

In fact the covariances are characterized by

$$\begin{aligned}\mathrm{Cov}\left(\varepsilon_{it},\varepsilon_{js}\right) &=\sigma_\mu^2+\sigma_\nu^2+\sigma_u^2 && \text{for } i=j,t=s,\\ &=\sigma_\mu^2 && \text{for } i=j,t\neq s,\\ &=\sigma_\nu^2 && \text{for } i\neq j,t=s,\\ &=0 && \text{for } i\neq j,t\neq s.\end{aligned}$$

If we order the data on the $N$ individuals in $T$ consecutive sets (see Wansbeek and Kapteyn 1989 and Baltagi 2005), under assumption RE.1.a, (A.4) and (A.5), $\underset{n\times n}{\mathbf{\Omega}}$ takes the special form

$$\underset{n\times n}{\mathbf{\Omega}}=\mathrm{E}\left(\underset{n\times 1}{\boldsymbol{\varepsilon}}\,\underset{1\times n}{\boldsymbol{\varepsilon}'}\right)=\sigma_u^2\underset{n\times n}{\mathbf{I}_n}+\sigma_\mu^2\underset{n\times N}{\boldsymbol{\Delta}_\mu}\,\underset{N\times n}{\boldsymbol{\Delta}'_\mu}+\sigma_\nu^2\underset{n\times T}{\boldsymbol{\Delta}_\nu}\,\underset{T\times n}{\boldsymbol{\Delta}'_\nu}. \tag{A.6}$$

Therefore, the appropriate rank of the weighted outer matrix is

$$\operatorname{rank}\left(\underset{k\times n}{\mathbf{X}'}\underset{n\times n}{\mathbf{\Omega}^{-1}}\underset{n\times k}{\mathbf{X}}\right)=k. \tag{A.7}$$

For efficiency of feasible *GLS*, we assume that the variance matrix of $\boldsymbol{\varepsilon}$ conditional on $\underset{1\times nk}{\mathbf{x}}$ is constant

$$\left.\begin{aligned}
&\mathrm{E}\left(\varepsilon_{it}^2\,\middle|\,\underset{1\times k}{\mathbf{x}_{it}}\right)=\mathrm{E}\left(\varepsilon_{it}^2\right)=\sigma_\mu^2+\sigma_\nu^2+\sigma_u^2\\
&\mathrm{E}\left(\varepsilon_{it}\varepsilon_{is}\,\middle|\,\underset{1\times T_i k}{\mathbf{x}_{i\{N\}}}\right)=\mathrm{E}\left(\varepsilon_{it}\varepsilon_{is}\right)=\sigma_\mu^2\\
&\mathrm{E}\left(\varepsilon_{it}\varepsilon_{jt}\,\middle|\,\underset{1\times N_t k}{\mathbf{x}_{t\{T\}}}\right)=\mathrm{E}\left(\varepsilon_{it}\varepsilon_{jt}\right)=\sigma_\nu^2\\
&\mathrm{E}\left(\varepsilon_{it}\varepsilon_{js}\,\middle|\,\underset{1\times nk}{\mathbf{x}}\right)=\mathrm{E}\left(\varepsilon_{it}\varepsilon_{js}\right)=0
\end{aligned}\right\}\mathrm{E}\left(\underset{n\times 1}{\boldsymbol{\varepsilon}}\underset{1\times n}{\boldsymbol{\varepsilon}'}\,\middle|\,\underset{1\times nk}{\mathbf{x}}\right)=\mathrm{E}\left(\underset{n\times 1}{\boldsymbol{\varepsilon}}\underset{1\times n}{\boldsymbol{\varepsilon}'}\right). \tag{A.8}$$

Assumptions (A.4), (A.5) and (A.8) are implied by the third RE assumption.

**RE.3** **Homoscedasticity and no serial correlation**: the conditional variance-covariance matrix of the idiosyncratic error terms $u_{it}$ is characterized by constant variances and zero covariances; in addition, the variance of both the individual specific effects $\mu_i$ and the time-specific effects $\nu_t$ is constant. This means that

**a.** $\mathrm{E}\left(\underset{n\times 1}{\mathbf{u}}\underset{1\times n}{\mathbf{u}'}\,\middle|\,\underset{1\times nk}{\mathbf{x}},\mu_i,\nu_t\right)=\sigma_u^2\underset{n\times n}{\mathbf{I}_n}$,

**b.** $\mathrm{E}\left(\mu_i^2\,\middle|\,\underset{1\times T_i k}{\mathbf{x}_{i\{N\}}}\right)=\sigma_\mu^2$,

**c.** $\mathrm{E}\left(\nu_t^2\,\middle|\,\underset{1\times N_t k}{\mathbf{x}_{t\{T\}}}\right)=\sigma_\nu^2$.

The assumption RE.3.a is identical to assumption FE.3. Under assumption RE.3.a, we have (i) $\mathrm{E}(u_{it}^2\,|\,\underset{1\times nk}{\mathbf{x}},\mu_i,\nu_t)=\sigma_u^2$, which implies assumption (A.4), and (ii) $\mathrm{E}(u_{it}u_{jt}|\underset{1\times N_t k}{\mathbf{x}_{t\{T\}}},\mu_i,\nu_t)=0\ (i\neq j)$, $\mathrm{E}(u_{it}u_{is}|\underset{1\times T_i k}{\mathbf{x}_{i\{N\}}},\mu_i,\nu_t)=0\ (t\neq s)$ and $\mathrm{E}(u_{it}u_{js}|\underset{1\times nk}{\mathbf{x}},\mu_i,\nu_t)=0\ (i\neq j,t\neq s)$, which imply assumption (A.5). But assumption

RE.3.a is stronger because it assumes that the conditional variances are constant and the conditional covariances are zero. Along with assumption RE.1.b, assumption RE.3.b is the same as $\mathrm{Var}(\mu_i|\underset{1\times T_i k}{\mathbf{x}_{i\{N\}}}) = \mathrm{Var}(\mu_i) = \sigma_\mu^2$ and along with assumption RE.1.c, assumption RE.3.c is the same as $\mathrm{Var}(\nu_t|\underset{1\times N_t k}{\mathbf{x}_{t\{T\}}}) = \mathrm{Var}(\nu_t) = \sigma_\nu^2$. Moreover, under assumption RE.3, assumption (A.8) holds and $\underset{n\times n}{\mathbf{\Omega}}$ has the form (A.6).

To implement a *FGLS* procedure, assume that we have consistent estimator of $\sigma_u^2$, $\sigma_\mu^2$ and $\sigma_\nu^2$. Then we can form

$$\underset{n\times n}{\hat{\mathbf{\Omega}}} = \hat{\sigma}_u^2 \underset{n\times n}{\mathbf{I}_n} + \hat{\sigma}_\mu^2 \underset{n\times N}{\mathbf{\Delta}_\mu} \underset{N\times n}{\mathbf{\Delta}'_\mu} + \hat{\sigma}_\nu^2 \underset{n\times T}{\mathbf{\Delta}_\nu} \underset{T\times n}{\mathbf{\Delta}'_\nu} \tag{A.9}$$

and the RE estimator is

$$\underset{k\times 1}{\mathbf{\beta}^{GLS}} = \left(\underset{k\times n}{\mathbf{X}'} \underset{n\times n}{\hat{\mathbf{\Omega}}} \underset{n\times k}{\mathbf{X}}\right)^{-1} \left(\underset{k\times n}{\mathbf{X}'} \underset{n\times n}{\hat{\mathbf{\Omega}}} \underset{n\times 1}{\mathbf{y}}\right). \tag{A.10}$$

## APPENDIX C: PROOF OF (13)

Since $\mathbf{e}_m \equiv \mathbf{y}_m - \mathbf{X}_m \mathbf{\beta}_m^{WT} = \left[\mathbf{I}_n - \mathbf{X}_m \left(\mathbf{X}'_m \mathbf{Q}_{[\Delta]} \mathbf{X}_m\right)^{-1} \mathbf{X}'_m \mathbf{Q}_{[\Delta]}\right] \mathbf{y}_m$ by definition we can write $\mathbf{e}_m = \mathbf{M}_m \cdot \mathbf{y}_m = \mathbf{M}_m \cdot \mathbf{\varepsilon}_m$. We assume that there is a constant term and then we consider the centered residuals $\mathbf{f}_m = \mathbf{E}_n \cdot \mathbf{e}_m = \mathbf{E}_n \cdot \mathbf{M}_m \cdot \mathbf{\varepsilon}_m$. With $\mathbf{Q}_{[\Delta]}$ idempotent, $\mathbf{Q}_{[\Delta]} \cdot \mathbf{\Delta}_\mu = \mathbf{Q}_{[\Delta]} \cdot \mathbf{\Delta}_\nu = \mathbf{0}$ and $\mathbf{Q}_{[\Delta]} \cdot \mathbf{\Omega} = \sigma_u^2 \cdot \mathbf{Q}_{[\Delta]}$ we have

$$\begin{aligned}
\underset{T\times n}{\mathbf{\Delta}'_\nu} \underset{n\times n}{\mathbf{M}_m} \underset{n\times n}{\mathbf{\Omega}_{mj}} \underset{n\times n}{\mathbf{M}'_j} \underset{n\times T}{\mathbf{\Delta}_\nu} &= \sigma_{\mu_{mj}}^2 \underset{T\times N}{\mathbf{\Delta}_{TN}} \underset{N\times T}{\mathbf{\Delta}'_{TN}} + \sigma_{\nu_{mj}}^2 \underset{T\times T}{\mathbf{\Delta}_T^2} + \sigma_{u_{mj}}^2 \underset{T\times T}{\mathbf{\Delta}_T} + \\
&+ \sigma_{u_{mj}}^2 \underset{T\times n}{\mathbf{\Delta}'_\nu} \underset{n\times k_m}{\mathbf{X}_m} \left(\underset{k_m\times n}{\mathbf{X}'_m} \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times k_m}{\mathbf{X}_m}\right)^{-1} \underset{k_m\times n}{\mathbf{X}'_m} \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times k_j}{\mathbf{X}_j} \left(\underset{k_j\times n}{\mathbf{X}'_j} \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times k_j}{\mathbf{X}_j}\right)^{-1} \underset{k_j\times n}{\mathbf{X}'_j} \underset{n\times T}{\mathbf{\Delta}_\nu}
\end{aligned} \tag{A.11}$$

and

$$\begin{aligned}
\underset{N\times n}{\mathbf{\Delta}'_\mu} \underset{n\times n}{\mathbf{M}_m} \underset{n\times n}{\mathbf{\Omega}_{mj}} \underset{n\times n}{\mathbf{M}'_j} \underset{n\times N}{\mathbf{\Delta}_\mu} &= \sigma_{\mu_{mj}}^2 \underset{N\times N}{\mathbf{\Delta}_N^2} + \sigma_{\nu_{mj}}^2 \underset{N\times T}{\mathbf{\Delta}'_{TN}} \underset{T\times N}{\mathbf{\Delta}_{TN}} + \sigma_{u_{mj}}^2 \underset{N\times N}{\mathbf{\Delta}_N} + \\
&+ \sigma_{u_{mj}}^2 \underset{N\times n}{\mathbf{\Delta}'_\mu} \underset{n\times k_m}{\mathbf{X}_m} \left(\underset{k_m\times n}{\mathbf{X}'_m} \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times k_m}{\mathbf{X}_m}\right)^{-1} \underset{k_m\times n}{\mathbf{X}'_m} \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times k_j}{\mathbf{X}_j} \left(\underset{k_j\times n}{\mathbf{X}'_j} \underset{n\times n}{\mathbf{Q}_{[\Delta]}} \underset{n\times k_j}{\mathbf{X}_j}\right)^{-1} \underset{k_j\times n}{\mathbf{X}'_j} \underset{n\times N}{\mathbf{\Delta}_\mu},
\end{aligned} \tag{A.12}$$

and we obtain

$$
\begin{aligned}
\mathrm{E}\left(q_{n_{mj}}\right) &= \mathrm{E}\left(\underset{1\times n}{\mathbf{u}'_j}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\mathbf{E}_n}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times n}{\mathbf{E}_n}\underset{n\times n}{\mathbf{M}_m}\underset{n\times 1}{\mathbf{u}_m}\right) = \mathrm{tr}\,\mathrm{E}\left(\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times n}{\mathbf{E}_n}\underset{n\times n}{\mathbf{M}_m}\underset{n\times 1}{\mathbf{u}_m}\underset{1\times n}{\mathbf{u}'_j}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\mathbf{E}_n}\right) = \\
&= \mathrm{tr}\left(\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times n}{\mathbf{E}_n}\underset{n\times n}{\mathbf{M}_m}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\mathbf{E}_n}\right) = \mathrm{tr}\left(\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times n}{\left(\mathbf{I}_n - \frac{\boldsymbol{\iota}_n\boldsymbol{\iota}'_n}{n}\right)}\underset{n\times n}{\mathbf{M}_m}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\left(\mathbf{I}_n - \frac{\boldsymbol{\iota}_n\boldsymbol{\iota}'_n}{n}\right)}\right) = \\
&= \mathrm{tr}\left(\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\right)\cdot\sigma^2_{u_{mj}} + \mathrm{tr}\left(\left(\underset{k_m\times n}{\mathbf{X}'_m}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_m}{\mathbf{X}_m}\right)^{-1}\underset{k_m\times n}{\mathbf{X}'_m}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_j}{\mathbf{X}_j}\left(\underset{k_j\times n}{\mathbf{X}'_j}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_j}{\mathbf{X}_j}\right)^{-1}\underset{k_j\times n}{\mathbf{X}'_j}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_m}{\mathbf{X}_m}\right)\cdot\sigma^2_{u_{mj}} + \\
&- \mathrm{tr}\left(\left(\underset{k_m\times n}{\mathbf{X}'_m}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_m}{\mathbf{X}_m}\right)^{-1}\underset{k_m\times n}{\mathbf{X}'_m}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_m}{\mathbf{X}_m}\right)\cdot\sigma^2_{u_{mj}} - \mathrm{tr}\left(\left(\underset{k_j\times n}{\mathbf{X}'_j}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_j}{\mathbf{X}_j}\right)^{-1}\underset{k_j\times n}{\mathbf{X}'_j}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_j}{\mathbf{X}_j}\right)\cdot\sigma^2_{u_{mj}}
\end{aligned}
\tag{A.13}
$$

and

$$
\begin{aligned}
\mathrm{E}\left(q_{N_{mj}}\right) &= \mathrm{E}\left(\underset{1\times n}{\mathbf{u}'_j}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\mathbf{E}_n}\underset{n\times T}{\mathbf{\Delta}_\nu}\underset{T\times T}{\mathbf{\Delta}_T^{-1}}\underset{T\times n}{\mathbf{\Delta}'_\nu}\underset{n\times n}{\mathbf{E}_n}\underset{n\times n}{\mathbf{M}_m}\underset{n\times 1}{\mathbf{u}_m}\right) = \mathrm{tr}\,\mathrm{E}\left(\underset{T\times T}{\mathbf{\Delta}_T^{-1}}\underset{T\times n}{\mathbf{\Delta}'_\nu}\underset{n\times n}{\mathbf{E}_n}\underset{n\times n}{\mathbf{M}_m}\underset{n\times 1}{\mathbf{u}_m}\underset{1\times n}{\mathbf{u}'_j}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\mathbf{E}_n}\underset{n\times T}{\mathbf{\Delta}_\nu}\right) \\
&= \mathrm{tr}\left(\underset{T\times T}{\mathbf{\Delta}_T^{-1}}\underset{T\times n}{\mathbf{\Delta}'_\nu}\underset{n\times n}{\mathbf{E}_n}\underset{n\times n}{\mathbf{M}_m}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\mathbf{E}_n}\underset{n\times T}{\mathbf{\Delta}_\nu}\right) = \mathrm{tr}\left(\underset{T\times T}{\mathbf{\Delta}_T^{-1}}\underset{T\times n}{\mathbf{\Delta}'_\nu}\underset{n\times n}{\left(\mathbf{I}_n - \frac{\boldsymbol{\iota}_n\boldsymbol{\iota}'_n}{n}\right)}\underset{n\times n}{\mathbf{M}_m}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\left(\mathbf{I}_n - \frac{\boldsymbol{\iota}_n\boldsymbol{\iota}'_n}{n}\right)}\underset{n\times T}{\mathbf{\Delta}_\nu}\right) = \\
&= \mathrm{tr}\left(\underset{T\times T}{\mathbf{\Delta}_T^{-1}}\underset{T\times n}{\mathbf{\Delta}'_\nu}\underset{n\times n}{\mathbf{M}_m}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times T}{\mathbf{\Delta}_\nu}\right) - \frac{\underset{1\times n}{\boldsymbol{\iota}'_n}\underset{n\times n}{\mathbf{M}_m}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times 1}{\boldsymbol{\iota}_n}}{n} = \\
&= \left(T + k_{N_{mj}}\right)\cdot\sigma^2_{u_{mj}} + T\cdot\sigma^2_{\mu_{mj}} + n\cdot\sigma^2_{\nu_{mj}} + \\
&- \frac{\underset{1\times n}{\boldsymbol{\iota}'_n}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times 1}{\boldsymbol{\iota}_n} + \sigma^2_{u_{mj}}\cdot\underset{1\times n}{\boldsymbol{\iota}'_n}\underset{n\times k_m}{\mathbf{X}_m}\left(\underset{k_m\times n}{\mathbf{X}'_m}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_m}{\mathbf{X}_m}\right)^{-1}\underset{k_m\times n}{\mathbf{X}'_m}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_j}{\mathbf{X}_j}\left(\underset{k_j\times n}{\mathbf{X}'_j}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_j}{\mathbf{X}_j}\right)^{-1}\underset{k_j\times n}{\mathbf{X}'_j}\underset{n\times 1}{\boldsymbol{\iota}_n}}{n}
\end{aligned}
\tag{A.14}
$$

and

$$
\begin{aligned}
\mathrm{E}\left(q_{T_{mj}}\right) &= \mathrm{E}\left(\underset{1\times n}{\mathbf{u}'_j}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\mathbf{E}_n}\underset{n\times N}{\mathbf{\Delta}_\mu}\underset{N\times N}{\mathbf{\Delta}_N^{-1}}\underset{N\times n}{\mathbf{\Delta}'_\mu}\underset{n\times n}{\mathbf{E}_n}\underset{n\times n}{\mathbf{M}_m}\underset{n\times 1}{\mathbf{u}_m}\right) = \mathrm{tr}\,\mathrm{E}\left(\underset{N\times N}{\mathbf{\Delta}_N^{-1}}\underset{N\times n}{\mathbf{\Delta}'_\mu}\underset{n\times n}{\mathbf{E}_n}\underset{n\times n}{\mathbf{M}_m}\underset{n\times 1}{\mathbf{u}_m}\underset{1\times n}{\mathbf{u}'_j}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\mathbf{E}_n}\underset{n\times N}{\mathbf{\Delta}_\mu}\right) \\
&= \mathrm{tr}\left(\underset{N\times N}{\mathbf{\Delta}_N^{-1}}\underset{N\times n}{\mathbf{\Delta}'_\mu}\underset{n\times n}{\mathbf{E}_n}\underset{n\times n}{\mathbf{M}_m}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\mathbf{E}_n}\underset{n\times N}{\mathbf{\Delta}_\mu}\right) = \mathrm{tr}\left(\underset{N\times N}{\mathbf{\Delta}_N^{-1}}\underset{N\times n}{\mathbf{\Delta}'_\mu}\underset{n\times n}{\left(\mathbf{I}_n - \frac{\boldsymbol{\iota}_n\boldsymbol{\iota}'_n}{n}\right)}\underset{n\times n}{\mathbf{M}_m}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times n}{\left(\mathbf{I}_n - \frac{\boldsymbol{\iota}_n\boldsymbol{\iota}'_n}{n}\right)}\underset{n\times N}{\mathbf{\Delta}_\mu}\right) = \\
&= \mathrm{tr}\left(\underset{N\times N}{\mathbf{\Delta}_N^{-1}}\underset{N\times n}{\mathbf{\Delta}'_\mu}\underset{n\times n}{\mathbf{M}_m}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times N}{\mathbf{\Delta}_\mu}\right) - \frac{\underset{1\times n}{\boldsymbol{\iota}'_n}\underset{n\times n}{\mathbf{M}_m}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times n}{\mathbf{M}'_j}\underset{n\times 1}{\boldsymbol{\iota}_n}}{n} = \\
&= \left(N + k_{T_{mj}}\right)\cdot\sigma^2_{u_{mj}} + n\cdot\sigma^2_{\mu_{mj}} + N\cdot\sigma^2_{\nu_{mj}} + \\
&\quad \frac{\underset{1\times n}{\boldsymbol{\iota}'_n}\underset{n\times n}{\mathbf{\Omega}_{mj}}\underset{n\times 1}{\boldsymbol{\iota}_n} + \sigma^2_{u_{mj}}\cdot\underset{1\times n}{\boldsymbol{\iota}'_n}\underset{n\times k_m}{\mathbf{X}_m}\left(\underset{k_m\times n}{\mathbf{X}'_m}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_m}{\mathbf{X}_m}\right)^{-1}\underset{k_m\times n}{\mathbf{X}'_m}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_j}{\mathbf{X}_j}\left(\underset{k_j\times n}{\mathbf{X}'_j}\underset{n\times n}{\mathbf{Q}_{[\Delta]}}\underset{n\times k_j}{\mathbf{X}_j}\right)^{-1}\underset{k_j\times n}{\mathbf{X}'_j}\underset{n\times 1}{\boldsymbol{\iota}_n}}{n},
\end{aligned}
\tag{A.15}
$$

where $\mathbf{J}_n = \boldsymbol{\iota}_n\boldsymbol{\iota}'_n$ and $\boldsymbol{\iota}_n$ is a vector of ones of dimension $n$.

## APPENDIX D: PROOF OF (15)

Since the $\underset{M\times 1}{\boldsymbol{\mu}_{i\{N\}}}$'s, the $\underset{M\times 1}{\mathbf{v}_{t\{T\}}}$'s and $\underset{M\times 1}{\mathbf{u}_{it}}$'s are independent, from the equations in (14) we can write

$$\begin{aligned}
\mathrm{E}\Big(\underset{M\times M}{\mathbf{W}_f}\Big) &= \mathrm{E}\Big(\underset{M\times M}{\mathbf{W}_u}\Big),\\
\mathrm{E}\Big(\underset{M\times M}{\mathbf{B}_f^C}\Big) &= \mathrm{E}\Big(\underset{M\times M}{\mathbf{B}_\mu^C}\Big)+\mathrm{E}\Big(\underset{M\times M}{\mathbf{B}_u^C}\Big),\\
\mathrm{E}\Big(\underset{M\times M}{\mathbf{B}_f^T}\Big) &= \mathrm{E}\Big(\underset{M\times M}{\mathbf{B}_\nu^T}\Big)+\mathrm{E}\Big(\underset{M\times M}{\mathbf{B}_u^T}\Big),
\end{aligned} \tag{A.16}$$

where the within individuals (co)variation is

$$\begin{aligned}
\underset{M\times M}{\mathbf{W}_u} &= \sum_{i=1}^{N}\sum_{t=1}^{T_i}\underset{M\times 1}{\left(\mathbf{u}_{it}-\bar{\mathbf{u}}_{i\bullet}-\bar{\mathbf{u}}_{\bullet t}\right)}\underset{1\times M}{\left(\mathbf{u}_{it}-\bar{\mathbf{u}}_{i\bullet}-\bar{\mathbf{u}}_{\bullet t}\right)'} =\\
&= \sum_{i=1}^{N}\sum_{t=1}^{T_i}\underset{M\times 1}{\mathbf{u}_{it}}\,\underset{1\times M}{\mathbf{u}'_{it}}-\sum_{i=1}^{N}T_i\,\underset{M\times 1}{\bar{\mathbf{u}}_{i\bullet}}\,\underset{1\times M}{\bar{\mathbf{u}}'_{i\bullet}}-\sum_{t=1}^{T}N_t\,\underset{M\times 1}{\bar{\mathbf{u}}_{\bullet t}}\,\underset{1\times M}{\bar{\mathbf{u}}'_{\bullet t}},
\end{aligned} \tag{A.17}$$

the between individuals (co)variations are

$$\begin{aligned}
\underset{M\times M}{\mathbf{B}_\mu^C} &= \sum_{i=1}^{N}T_i\underset{M\times 1}{\left(\boldsymbol{\mu}_{i\{N\}}-\bar{\boldsymbol{\mu}}\right)}\underset{1\times M}{\left(\boldsymbol{\mu}_{i\{N\}}-\bar{\boldsymbol{\mu}}\right)'}=\sum_{i=1}^{N}T_i\,\underset{M\times 1}{\boldsymbol{\mu}_{i\{N\}}}\,\underset{1\times M}{\boldsymbol{\mu}'_{i\{N\}}}-n\,\underset{M\times 1}{\bar{\boldsymbol{\mu}}}\,\underset{1\times M}{\bar{\boldsymbol{\mu}}'},\\
\underset{M\times M}{\mathbf{B}_u^C} &= \sum_{i=1}^{N}T_i\underset{M\times 1}{\left(\bar{\mathbf{u}}_{i\bullet}-\bar{\mathbf{u}}\right)}\underset{1\times M}{\left(\bar{\mathbf{u}}_{i\bullet}-\bar{\mathbf{u}}\right)'}=\sum_{i=1}^{N}T_i\,\underset{M\times 1}{\bar{\mathbf{u}}_{i\bullet}}\,\underset{1\times M}{\bar{\mathbf{u}}'_{i\bullet}}-n\,\underset{M\times 1}{\bar{\mathbf{u}}}\,\underset{1\times M}{\bar{\mathbf{u}}'}
\end{aligned} \tag{A.18}$$

and the between times (co)variations are

$$\begin{aligned}
\underset{M\times M}{\mathbf{B}_\nu^T} &= \sum_{t=1}^{T}N_t\underset{M\times 1}{\left(\boldsymbol{\nu}_{t\{T\}}-\bar{\boldsymbol{\nu}}\right)}\underset{1\times M}{\left(\boldsymbol{\nu}_{t\{T\}}-\bar{\boldsymbol{\nu}}\right)'}=\sum_{t=1}^{T}N_t\,\underset{M\times 1}{\boldsymbol{\nu}_{t\{T\}}}\,\underset{1\times M}{\boldsymbol{\nu}'_{t\{T\}}}-n\,\underset{M\times 1}{\bar{\boldsymbol{\nu}}}\,\underset{1\times M}{\bar{\boldsymbol{\nu}}'},\\
\underset{M\times M}{\mathbf{B}_u^T} &= \sum_{t=1}^{T}N_t\underset{M\times 1}{\left(\bar{\mathbf{u}}_{\bullet t}-\bar{\mathbf{u}}\right)}\underset{1\times M}{\left(\bar{\mathbf{u}}_{\bullet t}-\bar{\mathbf{u}}\right)'}=\sum_{t=1}^{T}N_t\,\underset{M\times 1}{\bar{\mathbf{u}}_{\bullet t}}\,\underset{1\times M}{\bar{\mathbf{u}}'_{\bullet t}}-n\,\underset{M\times 1}{\bar{\mathbf{u}}}\,\underset{1\times M}{\bar{\mathbf{u}}'},
\end{aligned} \tag{A.19}$$

where $\bar{u}_{mi\bullet}=\frac{\sum_{t=1}^{T_i}u_{mit}}{T_i}$, $\bar{u}_{m\bullet t}=\frac{\sum_{i=1}^{N_t}u_{mit}}{N_t}$, $\bar{u}_m=\frac{\sum_{i=1}^{N}\sum_{t=1}^{T_i}u_{mit}}{n}=\frac{\sum_{i=1}^{N}(T_i\cdot\bar{u}_{mi\bullet})}{n}$ or $\bar{u}_m=\frac{\sum_{t=1}^{T}\sum_{i=1}^{N_t}u_{mit}}{n}=\frac{\sum_{t=1}^{T}(N_t\cdot\bar{u}_{m\bullet t})}{n}$, $\bar{\mu}_m=\frac{\sum_{i=1}^{N}(T_i\cdot\mu_{mi})}{n}$ and $\bar{\nu}_m=\frac{\sum_{t=1}^{T}(N_t\cdot\nu_{mt})}{n}$.

Since $\mathrm{E}(\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}'_{i't'})=\delta_{ii'}\boldsymbol{\Sigma}_\mu+\delta_{tt'}\boldsymbol{\Sigma}_\nu+\delta_{ii'}\delta_{tt'}\boldsymbol{\Sigma}_u$, where $\mathrm{E}(\boldsymbol{\mu}_{i\{N\}}\boldsymbol{\mu}'_{i'\{N\}})=\delta_{ii'}\boldsymbol{\Sigma}_\mu$, $\mathrm{E}(\boldsymbol{\nu}_{t\{T\}}\boldsymbol{\nu}'_{t'\{T\}})=\delta_{tt'}\boldsymbol{\Sigma}_\nu$ and $\mathrm{E}(\mathbf{u}_{it}\mathbf{u}'_{i't'})=\delta_{ii'}\delta_{tt'}\boldsymbol{\Sigma}_u$, it follows that $\mathrm{E}(\bar{\mathbf{u}}_{i\bullet}\bar{\mathbf{u}}'_{i\bullet})=\frac{\boldsymbol{\Sigma}_u}{T_i}$, $\mathrm{E}(\bar{\mathbf{u}}_{\bullet t}\bar{\mathbf{u}}'_{\bullet t})=\frac{\boldsymbol{\Sigma}_u}{N_t}$, $\mathrm{E}(\bar{\boldsymbol{\mu}}\bar{\boldsymbol{\mu}}')=\frac{(\sum_{i=1}^{N}T_i^2)\cdot\boldsymbol{\Sigma}_\mu}{n^2}$, $\mathrm{E}(\bar{\boldsymbol{\nu}}\bar{\boldsymbol{\nu}}')=\frac{(\sum_{t=1}^{T}N_t^2)\cdot\boldsymbol{\Sigma}_\nu}{n^2}$ and $\mathrm{E}(\bar{\mathbf{u}}\bar{\mathbf{u}}')=\frac{\boldsymbol{\Sigma}_u}{n}$.

Therefore from the equations in (A.17), (A.18) and (A.19) we have

$$
\begin{aligned}
\mathrm{E}\left(\underset{M\times M}{\mathbf{W}_u}\right) &= \sum_{i=1}^{N}\sum_{t=1}^{T_i}\boldsymbol{\Sigma}_u - \sum_{i=1}^{N}\left(T_i\cdot\frac{\boldsymbol{\Sigma}_u}{T_i}\right) - \sum_{t=1}^{T}\left(N_t\cdot\frac{\boldsymbol{\Sigma}_u}{N_t}\right) = \left(n-N-T\right)\cdot\underset{M\times M}{\boldsymbol{\Sigma}_u}\,,\\
\mathrm{E}\left(\underset{M\times M}{\mathbf{B}_\mu^C}\right) &= \sum_{i=1}^{N}T_i\cdot\boldsymbol{\Sigma}_\mu - n\cdot\left(\sum_{i=1}^{N}T_i^2\right)\cdot\boldsymbol{\Sigma}_\mu\Big/n^2 = \left(n-\sum_{i=1}^{N}T_i^2\Big/n\right)\cdot\underset{M\times M}{\boldsymbol{\Sigma}_\mu}\,,\\
\mathrm{E}\left(\underset{M\times M}{\mathbf{B}_u^C}\right) &= \sum_{i=1}^{N}\left(T_i\cdot\frac{\boldsymbol{\Sigma}_u}{T_i}\right) - n\cdot\frac{\boldsymbol{\Sigma}_u}{n} = \left(N-1\right)\cdot\underset{M\times M}{\boldsymbol{\Sigma}_u}\,,\\
\mathrm{E}\left(\underset{M\times M}{\mathbf{B}_\nu^T}\right) &= \sum_{t=1}^{T}N_t\cdot\boldsymbol{\Sigma}_\nu - n\cdot\left(\sum_{t=1}^{T}N_t^2\right)\cdot\boldsymbol{\Sigma}_\nu\Big/n^2 = \left(n-\sum_{t=1}^{T}N_t^2\Big/n\right)\cdot\underset{M\times M}{\boldsymbol{\Sigma}_\nu}\,,\\
\mathrm{E}\left(\underset{M\times M}{\mathbf{B}_u^T}\right) &= \sum_{t=1}^{T}\left(N_t\cdot\frac{\boldsymbol{\Sigma}_u}{N_t}\right) - n\cdot\frac{\boldsymbol{\Sigma}_u}{n} = \left(T-1\right)\cdot\underset{M\times M}{\boldsymbol{\Sigma}_u}
\end{aligned}
\tag{A.20}
$$

and combining the equations in (A.20) with the equations in (A.16) we obtain

$$
\begin{aligned}
\mathrm{E}\left(\underset{M\times M}{\mathbf{W}_f}\right) &= \left(n-N-T\right)\cdot\underset{M\times M}{\boldsymbol{\Sigma}_u}\,,\\
\mathrm{E}\left(\underset{M\times M}{\mathbf{B}_f^C}\right) &= \left(n-\sum_{i=1}^{N}T_i^2\Big/n\right)\cdot\underset{M\times M}{\boldsymbol{\Sigma}_\mu} + \left(N-1\right)\cdot\underset{M\times M}{\boldsymbol{\Sigma}_u}\,,\\
\mathrm{E}\left(\underset{M\times M}{\mathbf{B}_f^T}\right) &= \left(n-\sum_{t=1}^{T}N_t^2\Big/n\right)\cdot\underset{M\times M}{\boldsymbol{\Sigma}_\nu} + \left(T-1\right)\cdot\underset{M\times M}{\boldsymbol{\Sigma}_u}\,.
\end{aligned}
\tag{A.21}
$$

## REFERENCES

BALTAGI, B. H. (1980). On seemingly unrelated regressions with error components. *Econometrica* 48, 1547–1551.

BALTAGI, B. H. (1981). Pooling: an experimental study of alternative testing and estimation procedures in a two-way error component model. *Journal of Econometrics* 17, 21–49.

BALTAGI, B. H. (1985). Pooling cross-sections with unequal time series lengths. *Economics Letters* 18, 133–136.

BALTAGI, B. H. (2005). *Econometric analysis of panel data*. 3rd ed. Chichester: Wiley and Sons.

BIØRN, E. (1981). Estimating economic relations from incomplete cross-section/time-series data. *Journal of Econometrics* 16, 221-236.

BIØRN, E. (2004). Regression systems for unbalanced panel data: a stepwise maximum likelihood procedure. *Journal of Econometrics* 122, 281-291.

BOUMAHDI, R., CHAABAN, J., THOMAS, A. (2004). Import demand estimation with country and products effects: application of multi-way unbalanced panel data models to Lebanese imports. *Cahier de Recherche INRA de Toulouse* 2004-17.

DAVIS, P. (2002). Estimating multi-way error components models with unbalanced data structures. *Journal of Econometrics* 106, 67-95.

KANG, S. (1985). A note on the equivalence of specification tests in the two-factor multivariate variance components model. *Journal of Econometrics* 28, 193-203.

MAGNUS, J. R. (1982). Multivariate error components analysis of linear and nonlinear regression models by maximum likelihood. *Journal of Econometrics* 19, 239–285.

NERLOVE, M. (1971). Further evidence on the estimation of dynamic economic relations from a time series of cross sections. *Econometrica* 39, 359-382.

WANSBEEK, T., KAPTEYN, A. (1989). Estimation of the error-components model with incomplete panels. *Journal of Econometrics* 41, 341–361.

WOOLDRIDGE J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA and London, UK: The MIT Press.