Prediction in a Spatial Nested Error Components Panel Data Model

Badi H. Baltagi[a,*] and Alain Pirotte[b,c]

[a] *Syracuse University, Department of Economics and Center for Policy Research, USA*

[b] *CRED-TEPP, University Panthéon-Assas Paris II, France*

[c] *IFSTTAR-DEST, French Institute of Science and Technology for Transport, Development and Networks, France*

(forthcoming in ***Journal of Forecasting***, 2014)

Abstract

This paper derives the Best Linear Unbiased Predictor (BLUP) for a spatial nested error components panel data model. This predictor is useful for panel data applications that exhibit spatial dependence and a nested (hierarchical) structure. The predictor allows for unbalancedness in the number of observations in the nested groups. One application includes forecasting average housing prices located in a county nested in a state. We derive the BLUP accounting for the spatial correlation across counties as well as the unbalancedness due to observing different number of counties nested in each state. Ignoring the nested spatial structure leads to inefficiency and inferior forecasts. Using Monte Carlo simulations, we show that our feasible predictor is better in root mean square error performance than the usual fixed and random effects panel predictors ignoring the spatial nested structure of the data.

Keywords: Spatial Nested Error Components; Unbalanced Panels; Forecasting; Linear Predictor; BLUP.

JEL classification: C33.

[*] Corresponding author. Tel.: + 1 315 443 1630; Fax: + 1 315 443 1081.

E-mail address: bbaltagi@maxwell.syr.edu (B.H. Baltagi).

1 Introduction

Baltagi and Pirotte (2013) derive the Best Linear Unbiased Predictor (BLUP) for a nested error components panel data model that ignores the spatial correlation along the cross-sectional units. They show that forecasting a nested panel data model with a non-nested error structure leads to higher root mean square error (RMSE) forecasts. This emphasizes the need to account for the *nested structure* of the data in forecasting. However, Baltagi and Pirotte (2013) did not consider possible *spatial autocorrelation* in the data. This is done in this paper. In fact, Baltagi and Pirotte (2010) emphasized that if the spatial dimension is neglected, test of hypotheses using the usual panel data estimators like random effects (RE) and fixed effects (FE) estimators perform badly and can lead to misleading inference. Accounting for spatial dependence in forecasting using panel data has been considered by Baltagi and Li (2004, 2006) who forecasted sales of cigarette and liquor per capita for U.S. states over time. However, these applications were for balanced panels and had no nested structure for the data. Spatial correlation arises in many examples, see Anselin (1988) and LeSage and Pace (2009) for several examples and a nice introduction to this literature. The structure of the spatial dependence can be related to location and distance, both in a geographical space as well as a more general economic or social network space (see Anselin, Le Gallo and Jayet, 2008). One application includes forecasting average housing prices located in a county nested in a state. For this application, one has to account for the spatial correlation across counties as well as the unbalancedness due to observing different number of counties nested in each state.[1] For a survey of panel data forecasting that does not include spatial dependence, see Baltagi (2008) and for spatial panel data forecasting that does *not* account for the nested structure in the data, see Baltagi, Bresson and Pirotte (2012). The latter study considered the case where the true Data Generating Process (DGP) is random effects with a spatial auto-regressive (SAR) or a spatial moving average (SMA) remainder error. Using Monte Carlo experiments, Baltagi, Bresson and Pirotte (2012) find that estimators

[1] It is important to note that this paper does not allow for unbalancedness in the time dimension and assumes that there are no missing observations across the sample period for all counties and states. This is the likely case when forecasting the average price in a county but not when forecasting individual house prices. The latter most likely exhibits unbalancedness in the time dimension as not all house prices are observed over the sample period.

that ignore heterogeneity/spatial autocorrelation perform badly in RMSE forecasts. Their results also show that accounting for heterogeneity improves the RMSE forecast performance by a big margin, while accounting for spatial autocorrelation improves the forecast performance by a smaller margin. Ignoring both lead to the worst RMSE forecasting performance. Other applications include Longhi and Nijkamp (2007) who obtain short-term forecasts of employment in a panel of 326 West German regional labor markets observed over the period 1987-2002. The authors find that taking into account spatial autocorrelation by means of spatial error models leads to forecasts that are, on average, more reliable than those models neglecting regional spatial autocorrelation. Girardin and Kholodilin (2011) obtain multi-step forecasts of the annual growth rates of the real gross regional product (GRP) for a panel of 31 Chinese regions over the period 1979–2007. This is done using a dynamic spatial panel model. They argue that using panel data and accounting for spatial effects substantially improve forecasting performance compared to the benchmark models estimated for each of the provinces separately. They also find that accounting for spatial dependence is even more pronounced at longer forecasting horizons where the root mean squared forecast error (RMSFE) improves from 8% at the 1-year horizon to over 25% at the 13- and 14-year horizons. They recommend incorporating a spatial dependence structure into regional forecasting models, especially when long-run forecasts are made. Also, Kholodilin et al. (2008) who consider a dynamic spatial panel model to forecast the GDP of 16 German Länder (states) over the period 1991-2006, at horizons varying from 1 to 5 years. Using root mean squared forecast error, they show that accounting for spatial effects helps to improve the forecast performance especially at longer horizons. In fact, they find that this gain in RMSFE is about 9% at the 1-year horizon and exceeds 40% at the 5-year horizon.

This paper focuses on prediction and derives the Goldberger (1962) BLUP for a spatial nested error components panel data model.[2] Using Monte Carlo experiments, this paper shows that this predictor performs well in terms of out of sample root mean square error. The predictions are based on the Maximum Likelihood estimator which takes into account the special *unbalanced* aspect of the data, the *spatial autocorrelation* and the *nested structure*

[2] While Baltagi and Li (2004) extend the BLUP to spatial panel models, Song and Jung (2002) extend the BLUP to the case of spatially *and* serially correlated error components models.

of the disturbances. The paper is organized as follows: In section 2, we derive the BLUP for the spatial nested random effects model with the special *unbalanced* aspect of the data. Section 3 describes the Monte Carlo design, while Section 4 describes the Monte Carlo results. Section 5 concludes with suggestions for further work.

2 The Spatial Nested Error Components Model

Consider the unbalanced panel data regression model:

$$y_{ijt} = \mathbf{x}_{ijt}\boldsymbol{\beta} + \varepsilon_{ijt}, \tag{1}$$

where $i = 1, \ldots, N$, $j = 1, \ldots, M_i$ and $t = 1, \ldots, T$. The dependent variable y_{ijt} could denote the average house price in county j located in state i at time period t. $\mathbf{x}_{ijt}$ is a $(1 \times K)$ vector of explanatory (exogenous) variables, while $\boldsymbol{\beta}$ represents a $(K \times 1)$ vector of parameters to be estimated. N denotes the number of states. M_i denotes the number of counties in each state i. This model allows for an unequal number of counties in each state i. However, it does not allow for missing observations across time. Moreover, in contrast to the usual panel data framework, we allow ε_{ijt} to be contemporaneously correlated. A simple and widely used approach to modelling spatial error dependence is to assume a SAR process:

$$\varepsilon_{ijt} = \rho \sum_{g=1}^{N} \sum_{h=1}^{M_g} w_{ij,gh}\varepsilon_{ght} + u_{ijt}, \tag{2}$$

where ρ is the autoregressive parameter to be estimated. The weight $w_{ij,gh} = w_{k,l}$ is the (k,l) element of the matrix W_S with ij denoting county j within state i, and similarly for gh. Thus $k, l = 1, ..., S$ where $S = \sum_{i=1}^{N} M_i$ and W_S is an $(S \times S)$ known spatial weights matrix which has zero diagonal elements and is usually row-normalized so that for row k, $\sum_{g=1}^{N} \sum_{h=1}^{M_g} w_{k,gh} = 1$. Typically, $\mathbf{W}_S$ is defined as first order contiguity, such elements consist of location pairs that have a common border but there is no higher order contiguity or could be based on distances between counties. The error component structure of the disturbances u_{ijt} contain an unobserved permanent unit-specific error component α_i, a nested permanent unit-specific error component μ_{ij} together with a remainder error component v_{ijt}. More formally,

$$u_{ijt} = \alpha_i + \mu_{ij} + v_{ijt}, \tag{3}$$

where α_i denotes an unobservable state specific time-invariant effect which is assumed to be i.i.d.$N\left(0,\sigma_\alpha^2\right)$, μ_{ij} denotes the nested effect of county j within the ith state which is assumed to be i.i.d.$N\left(0,\sigma_\mu^2\right)$, and v_{ijt} is a remainder disturbance term which is also assumed to be i.i.d.$N\left(0,\sigma_v^2\right)$. The α_i's, μ_{ij}'s and v_{ijt}'s are independent of each other and among themselves. In contrast to the classical literature on panel data, grouping the data by periods rather than units is more convenient when we consider the spatial autocorrelation due to (2). For a cross-section t, the model (1) can be written as:

$$\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t, \tag{4}$$

where $\mathbf{y}_t$ is of dimension $(S \times 1)$, $\mathbf{X}_t$ is an $(S \times K)$ matrix of explanatory variables and the expression of the first order spatial autoregressive disturbances process $\boldsymbol{\varepsilon}_t$ is given by

$$\boldsymbol{\varepsilon}_t = \rho\mathbf{W}_S\boldsymbol{\varepsilon}_t + \mathbf{u}_t, \tag{5}$$

and

$$\mathbf{u}_t = diag\left(\boldsymbol{\iota}_{M_i}\right)\boldsymbol{\alpha} + \boldsymbol{\mu} + \mathbf{v}_t, \tag{6}$$

where $\mathbf{u}_t$ is $(S \times 1)$, $\boldsymbol{\alpha}$ is the vector of state effects of dimension $(N \times 1)$, $\boldsymbol{\mu}^\top = \left(\boldsymbol{\mu}_1^\top, \ldots, \boldsymbol{\mu}_N^\top\right)$, a vector of dimension $(1 \times S)$, $\boldsymbol{\mu}_i^\top = \left(\mu_{i1}, \ldots, \mu_{iM_i}\right)$, a vector of dimension $(1 \times M_i)$, $\boldsymbol{\iota}_{M_i}$ is a vector of ones of dimension $(M_i \times 1)$. By $diag\left(\boldsymbol{\iota}_{M_i}\right)$, we mean $diag\left(\boldsymbol{\iota}_{M_1}, \ldots, \boldsymbol{\iota}_{M_N}\right)$. $\mathbf{v}_t$ is of dimension $(S \times 1)$. The covariance matrix of $\mathbf{u}_t$ is

$$E\left[\mathbf{u}_t\mathbf{u}_t^\top\right] = \sigma_\alpha^2 diag\left(\mathbf{J}_{M_i}\right) + \left(\sigma_\mu^2 + \sigma_v^2\right)\mathbf{I}_S, \tag{7}$$

where $\mathbf{I}_S$ $\left(= diag\left(\mathbf{I}_{M_i}\right)\right)$ is an identity matrix of dimension S. $\mathbf{J}_{M_i} = \left(\boldsymbol{\iota}_{M_i}\boldsymbol{\iota}_{M_i}^\top\right)$ is a matrix of ones of dimension $(M_i \times M_i)$. For the full $(TS \times 1)$ vector of disturbances $\mathbf{u}$, we have

$$\mathbf{u} = \left(\boldsymbol{\iota}_T \otimes diag\left(\boldsymbol{\iota}_{M_i}\right)\right)\boldsymbol{\alpha} + \left(\boldsymbol{\iota}_T \otimes \mathbf{I}_S\right)\boldsymbol{\mu} + \mathbf{v}, \tag{8}$$

and the covariance matrix of $\mathbf{u}$ corresponds to

$$\begin{aligned}\boldsymbol{\Omega}_u &= \sigma_\alpha^2\left(\mathbf{Z}_\alpha\mathbf{Z}_\alpha^\top\right) + \sigma_\mu^2\left(\mathbf{Z}_\mu\mathbf{Z}_\mu^\top\right) + \sigma_v^2\left(\mathbf{I}_T \otimes \mathbf{I}_S\right) \\ &= \sigma_\alpha^2\left(\mathbf{J}_T \otimes diag\left(\mathbf{J}_{M_i}\right)\right) + \left(\sigma_\mu^2\mathbf{J}_T + \sigma_v^2\mathbf{I}_T\right) \otimes \mathbf{I}_S,\end{aligned} \tag{9}$$

where $\mathbf{Z}_\alpha = \boldsymbol{\iota}_T \otimes diag\left(\boldsymbol{\iota}_{M_i}\right)$, $\mathbf{Z}_\mu = \boldsymbol{\iota}_T \otimes \mathbf{I}_S$ and $\mathbf{J}_T = \left(\boldsymbol{\iota}_T\boldsymbol{\iota}_T^\top\right)$ is a matrix of ones of dimension $(T \times T)$. Replace $\mathbf{J}_T$ by its idempotent counterpart $T\overline{\mathbf{J}}_T$, $\mathbf{J}_{M_i}$

by $M_i\overline{\mathbf{J}}_{M_i}$. Also, define $\mathbf{E}_T = \mathbf{I}_T - \overline{\mathbf{J}}_T$, and $\mathbf{E}_{M_i} = \mathbf{I}_{M_i} - \overline{\mathbf{J}}_{M_i}$, and replace $\mathbf{I}_T$ by $\left(\mathbf{E}_T + \overline{\mathbf{J}}_T\right)$, $\mathbf{I}_{M_i}$ by $\left(\mathbf{E}_{M_i} + \overline{\mathbf{J}}_{M_i}\right)$. Collecting terms with the same matrices, one gets the spectral decomposition of $\boldsymbol{\Omega}_u$:

$$\boldsymbol{\Omega}_u = \lambda_1 \mathbf{Q}_1 + \lambda_2 \mathbf{Q}_2 + \left(\mathbf{I}_T \otimes diag\left(\lambda_{3i}\mathbf{I}_{M_i}\right)\right)\mathbf{Q}_3, \tag{10}$$

with

$$\lambda_1 = \sigma_v^2,\ \lambda_2 = T\sigma_\mu^2 + \sigma_v^2,\ \lambda_{3i} = M_i T\sigma_\alpha^2 + T\sigma_\mu^2 + \sigma_v^2, \tag{11}$$

$$\mathbf{Q}_1 = \mathbf{E}_T \otimes \mathbf{I}_S,\ \mathbf{Q}_2 = \overline{\mathbf{J}}_T \otimes diag\left(\mathbf{E}_{M_i}\right), \tag{12}$$

$$\mathbf{Q}_3 = \overline{\mathbf{J}}_T \otimes diag\left(\overline{\mathbf{J}}_{M_i}\right), \tag{13}$$

and $\overline{\mathbf{J}}_T = \mathbf{J}_T/T$, $\overline{\mathbf{J}}_{M_i} = \mathbf{J}_{M_i}/M_i$. The operators $\mathbf{Q}_1$, $\mathbf{Q}_2$ and $\mathbf{Q}_3$ are symmetric and idempotent, with their rank equal to their trace. Moreover, they are pairwise orthogonal and sum to the identity matrix. From (10), we can easily obtain $\boldsymbol{\Omega}_u^{-1}$ as

$$\boldsymbol{\Omega}_u^{-1} = \lambda_1^{-1}\mathbf{Q}_1 + \lambda_2^{-1}\mathbf{Q}_2 + \left(\mathbf{I}_T \otimes diag\left(\lambda_{3i}^{-1}\mathbf{I}_{M_i}\right)\right)\mathbf{Q}_3. \tag{14}$$

For the full $(TS \times 1)$ vector disturbances $\boldsymbol{\varepsilon}$, we have:

$$\boldsymbol{\varepsilon} = \left(\mathbf{I}_T \otimes \mathbf{B}_S^{-1}\right)\mathbf{u}, \tag{15}$$

where $\mathbf{B}_S = \mathbf{I}_S - \rho\mathbf{W}_S$. The corresponding $(TS \times TS)$ covariance matrix is given by:

$$\boldsymbol{\Omega}_\varepsilon = \mathbf{A}\boldsymbol{\Omega}_u\mathbf{A}^\top, \tag{16}$$

where $\mathbf{A}$ is a block-diagonal matrix equal to $\mathbf{I}_T \otimes \mathbf{B}_S^{-1}$. Following the properties of the matrices $\boldsymbol{\Omega}_u$ and $\mathbf{A}$, we obtain the inverse covariance matrix of $\boldsymbol{\varepsilon}$ defined as:

$$\boldsymbol{\Omega}_\varepsilon^{-1} = \left(\mathbf{A}^\top\right)^{-1}\boldsymbol{\Omega}_u^{-1}\mathbf{A}^{-1}. \tag{17}$$

These expressions are useful for deriving the Maximum Likelihood (ML) estimator, see Anselin (1988) and Elhorst (2003, 2010) and LeSage and Pace (2009) for ML estimation of spatial models. Under normality of the disturbances, the log-likelihood function is given by:

$$\ln L = -\frac{1}{2}\ln\left|\boldsymbol{\Omega}_\varepsilon\right| - \frac{1}{2}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)^\top \boldsymbol{\Omega}_\varepsilon^{-1}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right). \tag{18}$$

After some basic mathematical manipulations, we obtain:

$$\ln L = -\frac{1}{2}\ln\left|\boldsymbol{\Omega}_u\right| + T\ln\left|\mathbf{B}_S\right| - \frac{1}{2}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)^\top \boldsymbol{\Omega}_\varepsilon^{-1}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right). \tag{19}$$

Let $\delta_1 = \sigma_\alpha^2/\sigma_v^2$, $\delta_2 = \sigma_\mu^2/\sigma_v^2$ and $\mathbf{\Omega}_\varepsilon = \sigma_v^2\mathbf{\Sigma}$, then the log-likelihood function (19) can be written as[3]

$$\begin{aligned}\ln L &= -\frac{TS}{2}\ln\sigma_v^2 - \frac{1}{2}\sum_{i=1}^{N}\ln\left(T\left(M_i\delta_1+\delta_2\right)+1\right) \\ &\quad -\frac{1}{2}\sum_{i=1}^{N}\left(M_i-1\right)\ln\left(T\delta_2+1\right)+T\ln\left|\mathbf{B}_S\right| \\ &\quad -\frac{1}{2\sigma_v^2}\left(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\right)^\top\mathbf{\Sigma}^{-1}\left(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\right). \end{aligned} \tag{20}$$

The first-order conditions give closed form solutions for $\boldsymbol{\beta}$ and σ_v^2 conditional on δ_1, δ_2 and ρ:

$$\widehat{\boldsymbol{\beta}}=\left(\mathbf{X}^\top\widehat{\mathbf{\Sigma}}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^\top\widehat{\mathbf{\Sigma}}^{-1}\mathbf{y}, \tag{21}$$

$$\widehat{\sigma}_v^2=\frac{\left(\mathbf{y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)^\top\widehat{\mathbf{\Sigma}}^{-1}\left(\mathbf{y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)}{TS}. \tag{22}$$

However, the first-order conditions for the other parameters in (20) are intertwined which mean that they are non-linear, i.e. the equations cannot be solved analytically. Therefore, a numerical solution by means of iteration procedure is needed. In a first stage, we fix initial values for δ_1, δ_2 and ρ to obtain $\widehat{\boldsymbol{\beta}}$ and $\widehat{\sigma}_v^2$. In a second step, we find δ_1, δ_2 and ρ that maximize the concentrated log-likelihood.

3 Prediction

Following Goldberger (1962), the BLUP for $y_{ij,T+\tau}$, the average house price in county j located in state i at time period $T+\tau$, denoted by $\widehat{y}_{ij,T+\tau}$, can be written in general form as

$$\widehat{y}_{ij,T+\tau}=\mathbf{x}_{ij,T+\tau}\widehat{\boldsymbol{\beta}}_{GLS}+\boldsymbol{\omega}^\top\mathbf{\Omega}_\varepsilon^{-1}\widehat{\boldsymbol{\varepsilon}}_{GLS}, \tag{23}$$

[3] Grouping the data by units rather than periods, Baltagi et al. (2001) have shown that the covariance matrix of $\mathbf{u}$ is given by: $\mathbf{\Omega}_u^p = diag\left(\mathbf{\Lambda}_i^p\right) = diag\left(\lambda_{1i}^p\mathbf{Q}_{1i}+\lambda_{2i}^p\mathbf{Q}_{2i}+\lambda_{3i}^p\mathbf{Q}_{3i}\right)$ where $\mathbf{Q}_{1i}=\mathbf{I}_{M_i}\otimes\mathbf{E}_T$, $\mathbf{Q}_{2i}=\mathbf{E}_{M_i}\otimes\overline{\mathbf{J}}_T$, $\mathbf{Q}_{3i}=\overline{\mathbf{J}}_{M_i}\otimes\overline{\mathbf{J}}_T$ and $\lambda_{pi,}$, for $p=1,2,3$, are the distinct characteristic roots of $\mathbf{\Lambda}_i$ then $\left|\mathbf{\Lambda}_i\right|=\left(\lambda_{3i}\right)\left(\lambda_{2i}^{M_i-1}\right)\left(\lambda_{1i}^{M_i(T-1)}\right)$.

where $\boldsymbol{\omega} = E\left[\varepsilon_{ij,T+\tau}\boldsymbol{\varepsilon}\right]$ is the covariance between the future disturbance $\varepsilon_{ij,T+\tau}$ and the sample disturbances $\boldsymbol{\varepsilon}$. $\widehat{\boldsymbol{\beta}}_{GLS}$ is the GLS estimator of $\boldsymbol{\beta}$ from equation (1) based on true $\boldsymbol{\Omega}_{\varepsilon}$. While $\widehat{\boldsymbol{\varepsilon}}_{GLS}$ denotes the corresponding GLS residual vector. For the nested error components model *without* spatial autocorrelation ($\rho = 0$), Baltagi and Pirotte (2013) derive the BLUP which reduces to:

$$\widehat{y}_{ij,T+\tau} = \mathbf{x}_{ij,T+\tau}\widehat{\boldsymbol{\beta}}_{GLS} + \left(\frac{T\sigma_{\mu}^{2}}{\lambda_{2}}\right)\widehat{\boldsymbol{\varepsilon}}_{ij.,GLS} + \left[\frac{M_{i}T\sigma_{\alpha}^{2}\sigma_{v}^{2}}{\lambda_{2}\lambda_{3i}}\right]\widehat{\boldsymbol{\varepsilon}}_{i..,GLS}, \qquad (24)$$

where $\widehat{\boldsymbol{\varepsilon}}_{ij.,GLS} = \sum_{t=1}^{T}\widehat{\varepsilon}_{ijt,GLS}/T$ and $\widehat{\boldsymbol{\varepsilon}}_{i..,GLS} = \sum_{j=1}^{M_i}\sum_{t=1}^{T}\widehat{\varepsilon}_{ijt,GLS}/M_iT$. Therefore, the BLUP of $y_{ij,T+\tau}$ for the nested error components model modifies the usual GLS forecast by adding two terms. The first is a fraction of the average of the GLS residuals (over time) corresponding to the average house price in county j located in state i. The second term adds a fraction of the average GLS residual (over time as well county j) corresponding to state i. In order to make (24) operational, $\widehat{\boldsymbol{\beta}}_{GLS}$ and the variance components are replaced by their feasible estimates proposed by Baltagi, Song and Jung (2001).

Here, we derive the BLUP when *both* spatial autocorrelation (2) and nested error components (3) are present in the model. After some algebra provided in the appendix, one can show that

$$\begin{aligned}\boldsymbol{\omega}^{\top} &= \sigma_{\alpha}^{2}\mathbf{b}_{ij}diag\left(\boldsymbol{\iota}_{M_i}\right)\left(\boldsymbol{\iota}_{T}^{\top}\otimes diag\left(\boldsymbol{\iota}_{M_i}\right)^{\top}\left(\mathbf{B}_{S}^{-1}\right)^{\top}\right)\\ &\quad+\sigma_{\mu}^{2}\mathbf{b}_{ij}\left(\boldsymbol{\iota}_{T}^{\top}\otimes\left(\mathbf{B}_{S}^{-1}\right)^{\top}\right), \qquad (25)\end{aligned}$$

where $\mathbf{b}_{ij}$ is the ijth row of the matrix $\mathbf{B}_{S}^{-1}$, and the second term of (23) is given by

$$\begin{aligned}\boldsymbol{\omega}^{\top}\boldsymbol{\Omega}_{\varepsilon}^{-1}\widehat{\boldsymbol{\varepsilon}}_{GLS} &= \frac{\sigma_{\alpha}^{2}}{\sigma_{v}^{2}}\mathbf{b}_{ij}diag\left(\boldsymbol{\iota}_{M_i}\right)\left[\boldsymbol{\iota}_{T}^{\top}\otimes diag\left(\theta_{2i}^{-1}\boldsymbol{\iota}_{M_i}^{\top}\right)\right]\left[\mathbf{I}_{T}\otimes\mathbf{B}_{S}\right]\widehat{\boldsymbol{\varepsilon}}_{GLS}\\ &\quad+\frac{\sigma_{\mu}^{2}}{\sigma_{v}^{2}}\mathbf{b}_{ij}\left[\boldsymbol{\iota}_{T}^{\top}\otimes\left[\theta_{1}^{-1}diag\left(\mathbf{E}_{M_i}\right)+diag\left(\theta_{2i}^{-1}\overline{\mathbf{J}}_{M_i}\right)\right]\right]\\ &\quad\left[\mathbf{I}_{T}\otimes\mathbf{B}_{S}\right]\widehat{\boldsymbol{\varepsilon}}_{GLS}, \qquad (26)\end{aligned}$$

where $\theta_1 = (T\delta_2 + 1)$, $\theta_{2i} = (M_iT\delta_1 + T\delta_2 + 1)$, $\delta_1 = \sigma_{\alpha}^{2}/\sigma_{v}^{2}$ and $\delta_2 = \sigma_{\mu}^{2}/\sigma_{v}^{2}$.

Thus, if we replace (26) in (23), the BLUP of $y_{ij,T+\tau}$ is given by

$$\begin{aligned}\widehat{y}_{ij,T+\tau} &= \mathbf{x}_{ij,T+\tau}\widehat{\boldsymbol{\beta}}_{GLS} \\ &+\frac{\sigma_\alpha^2}{\sigma_v^2}\mathbf{b}_{ij}diag\left(\boldsymbol{\iota}_{M_i}\right)\left[\boldsymbol{\iota}_T^\top\otimes diag\left(\theta_{2i}^{-1}\boldsymbol{\iota}_{M_i}^\top\right)\right]\left[\mathbf{I}_T\otimes\mathbf{B}_S\right]\widehat{\boldsymbol{\varepsilon}}_{GLS} \\ &+\frac{\sigma_\mu^2}{\sigma_v^2}\mathbf{b}_{ij}\left[\boldsymbol{\iota}_T^\top\otimes\left[\theta_1^{-1}diag\left(\mathbf{E}_{M_i}\right)+diag\left(\theta_{2i}^{-1}\overline{\mathbf{J}}_{M_i}\right)\right]\right] \\ &\left[\mathbf{I}_T\otimes\mathbf{B}_S\right]\widehat{\boldsymbol{\varepsilon}}_{GLS}. \end{aligned} \tag{27}$$

To compute (27), we use the ML estimates obtained from the iterative procedure described in section 2.

4 Monte Carlo Design

We consider the simple regression model:

$$y_{ijt}=\beta_0+\beta_1x_{ijt}+\varepsilon_{ijt} \qquad i=1,\ldots,N,\ j=1,\ldots,M_i,\ t=1,\ldots,T, \tag{28}$$

where the disturbance ε_{ijt} has a spatial autocorrelation structure

$$\varepsilon_{ijt}=\rho\sum_{g=1}^{N}\sum_{h=1}^{M_g}w_{ij,gh}\varepsilon_{ght}+u_{ijt},$$

and u_{ijt} a nested error components structure given by

$$u_{ijt}=\alpha_i+\mu_{ij}+v_{ijt}. \tag{29}$$

Throughout the experiment the parameters of (28) were set at $\beta_0=3$, $\beta_1=1$ and $\rho=0.2$ and 0.8. We have taken the spatial matrix $\mathbf{W}_S$ proposed by Kelejian and Prucha (1999). The weight matrices are labelled as "j ahead and j behind" with the non-zero elements being $1/2j$. Note that, as j increase, the value of non-zero elements $1/2j$ decreases and, this in turn may reduce the amount of spatial correlation. Here, we have considered $j=1$, namely $\mathbf{W}(1,1)$. The x_{ijt} explanatory variable is generated by a similar method to that of Nerlove (1971). More precisely, we have:

$$x_{ijt}=0.3t+0.8x_{ijt-1}+\omega_{ijt}, \tag{30}$$

where ω_{ijt} is a random variable uniformly distributed on the interval $[-0.5, 0.5]$ and $x_{ij0} = 40 + 20\omega_{ij0}$, see Antweiler (2001). The first 10 period observations were discarded to minimize the effect of initial values. For the data generating processes of the disturbances, we assume $\alpha_i \sim iid.N(0, \sigma_\alpha^2)$, $\mu_{ij} \sim iid.N\left(0, \sigma_\mu^2\right)$ and $v_{ijt} \sim iid.N(0, \sigma_v^2)$. We fix $\sigma_u^2 = \sigma_\alpha^2 + \sigma_\mu^2 + \sigma_v^2 = 5$ and define $\gamma_1 = \sigma_\alpha^2/\sigma_u^2$ and $\gamma_2 = \sigma_\mu^2/\sigma_u^2$. These two ratios vary over the set $(0, 0.2, 0.4, 0.6, 0.8)$ such that $(1 - \gamma_1 - \gamma_2)$ is always positive. For all experiments, $(N, T) = (10, 5)$, the sample size (i.e. TS) is fixed at 300. We consider 3 patterns describing the unbalanced patterns $(M_1, ..., M_N)$, see Table 1. For example, Pattern P_2 has 2 counties in the first and second states; 3 counties for the third to sixth states; 5 counties in the seventh state; and 13 counties in the eighth to tenth states, all observed over five periods. We also report summary measure of unbalancedness described in Baltagi, Song and Jung (2001). This is defined by

$$c = \frac{N}{\overline{M} \sum_{i=1}^{N} (1/M_i)}, \tag{31}$$

where $\overline{M} = \frac{1}{N} \sum_{i=1}^{N} M_i$. The coefficient c denotes the measure of subgroup unbalancedness due to each group size. This measure takes the value one when the data are balanced. Smaller values than one indicate that the data are more unbalanced.

[**INSERT TABLE 1**]

For each experiment, we compute the post sample RMSE forecasts of OLS, Fixed Effects (FE), Random Effects (RE), Nested random Effects (NE) and Spatial Nested random Effects (SNE) predictors. In order to do this, we generate 5 more time periods for each county (i.e. $T + \tau, \tau = 1, \dots, 5$) which are not used in the estimation. Note that the usual OLS, FE, RE and NE estimators ignore the spatial structure in the data. The FE estimator accounts for state fixed effects. For the RE estimator, we use the unbalanced Swamy and Arora (1972) FGLS estimator proposed by Baltagi and Chang (1994) and the predictor described in Taub (1979). For the nested random effects (NE) predictor, we use the unbalanced Swamy and Arora (1972) FGLS estimator proposed by Baltagi, Song and Jung (2001) and the predictor proposed in (24). For the Spatial Nested random Effects (SNE) predictor, we use the ML iterative procedure described in Section 2 and the BLUP proposed in

(27). An average RMSE is also calculated across all counties (i.e. $S = 60$) at different forecast horizon. For each experiment, $1,000$ replications are performed.

5 Monte Carlo Results

Table 2 gives the post sample RMSE for one year ahead, five year ahead and the average over 5 years for the various estimators considered over 3 patterns of unbalancedness and for $\rho = 0.2$ and various values of $\gamma_1 = \sigma_\alpha^2/\sigma_u^2$ and $\gamma_2 = \sigma_\mu^2/\sigma_u^2$. One thing is clear, since $\rho \neq 0$, whatever the values of (γ_1, γ_2), the OLS estimator is not BLUE. This estimator ignores the random effects and the spatial nested error structure of the data. If $\gamma_1 = 0$ and $\gamma_2 = 0$, the disturbances follow a SAR process. Nevertheless, the performance of OLS in this case performs second best in RMSE right after the SNE estimator. For all other cases, OLS gives the worse RMSE prediction over all horizons considered and for all unbalanced patterns. FE which controls for fixed state effects gives a much improved forecast RMSE over OLS but is still not as good as RE, NE and SNE across different (γ_1,γ_2) patterns and different degrees of unbalancedness and nested error structure. OLS, FE, RE and NE estimators do not consider the spatial dimension of the data, i.e. the fact that $\rho \neq 0$. The SNE estimator performs well whatever case is considered. The spatial autocorrelation has a significant impact in terms of RMSE prediction. Table 3 generates results that are similar to Table 2, only now for a higher spatial autocorrelation coefficient $\rho = 0.8$. The results are the same, but the magnitudes of the RMSEs are different. For all cases considered, the RMSE forecasts of SNE performs best and by a bigger margin over estimators that do not account for the spatial correlation.

[**INSERT TABLES 2 AND 3**]

6 Conclusion

This paper describes a Maximum Likelihood iterative procedure to estimate the spatial nested error components model. It also derives the Best Linear Unbiased Predictor for a spatial nested error components panel data model with a special *unbalanced* aspect of the data. Our Monte Carlo simulations show that this predictor is better in root mean square error performance than the usual random, fixed or nested effects predictors ignoring the spatial nested structure of the data. An important extension of our forecasting exercise is to allow for unbalancedness across time as this will allow the forecast of housing prices based on repeated transaction sales for each house allowing for the spatial nested random effects structure of the panel.

References

[1] Anselin, L., 1988, *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht.

[2] Anselin, L., J. Le Gallo and H. Jayet, 2008, Spatial panel econometrics, Chapter 19 in L. Mátyás and P. Sevestre (eds), The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice, Springer-Verlag, Berlin, 625-660.

[3] Antweiler, W., 2001, Nested random effects estimation in unbalanced panel data. *Journal of Econometrics* **101**, 295-313.

[4] Baltagi, B.H., 2008, Forecasting with panel data, *Journal of Forecasting* **27**, 155-173.

[5] Baltagi, B.H., G. Bresson and A. Pirotte, 2012, Forecasting with spatial panel data. *Computational Statistics and Data Analysis* **56**, 3381-3397.

[6] Baltagi, B.H. and Y.-J. Chang, 1994, Incomplete panels. A comparative study of alternative estimators for the unbalanced one-way error component regression model. *Journal of Econometrics* **62**, 67-89.

[7] Baltagi, B.H. and D. Li, 2004, Prediction in the panel data model with spatial correlation. In *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Anselin L., R.J.G.M. Florax, S.J. Rey (eds). Springer, Berlin, 283-295.

[8] Baltagi, B.H. and D. Li, 2006, Prediction in the panel data model with spatial correlation: The case of liquor. *Spatial Economic Analysis* **1**, 175-185.

[9] Baltagi, B.H. and A. Pirotte, 2010, Panel data inference under spatial dependence. *Economic Modelling* **27**, 1368-1381.

[10] Baltagi, B.H. and A. Pirotte, 2013, Prediction in an unbalanced nested error components panel data model. Forthcoming in *Journal of Forecasting.*

[11] Baltagi, B.H., S.H. Song and B.C. Jung, 2001, The unbalanced nested error component regression model. *Journal of Econometrics* **101**, 357-381.

[12] Elhorst, J.P., 2003, Specification and Estimation of Spatial Panel Data Models. *International Regional Science Review* **26**, 244-268.

[13] Elhorst, J.P., 2010, *Spatial Panel Data Models*, Chapter 2 in M.M. Fischer and A. Getis (eds), Handbook of Applied Spatial Analysis, Springer-Verlag, Berlin, 377-405.

[14] Girardin, E. and K.A. Kholodilin, 2011, How helpful are spatial effects in forecasting the growth of Chinese provinces?. *Journal of Forecasting* **30**, 622-643.

[15] Goldberger, A.S., 1962, Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association* **57**, 369-375.

[16] Kelejian, H.H. and I.R. Prucha, 1999, A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review* **40**, 509-533.

[17] Kholodilin, K.A., B. Siliverstovs and S. Kooths, 2008, A dynamic panel data approach to the forecasting of the GDP of German Länder. *Spatial Economic Analysis* **3**, 195–207.

[18] LeSage, J.P. and R.K. Pace, 2009, *Introduction to Spatial Econometrics*, Chapman & Hall/CRC Press, Boca Raton.

[19] Longhi, S. and P. Nijkamp, 2007, Forecasting regional labor market developments under spatial heterogeneity and spatial correlation. *International Regional Science Review* **30**, 100–119.

[20] Nerlove, M., 1971, Further evidence on the estimation of dynamic economic relations from a time-series of cross-sections. *Econometrica* **39**, 359-383.

[21] Song, S.H. and B.C. Jung, 2002, BLUP in the panel regression model with spatially and serially correlated error components. *Statistical Papers* **43**, 551-566.

[22] Swamy, P.A.V.B. and S.S. Arora, 1972, The exact finite sample properties of the estimators of coefficients in the error components regression models. *Econometrica* **40**, 261-275.

Appendix

This appendix derives the BLUP for the unbalanced spatial nested error components model described in (1), (2) and (3). From (5) and (6), we have

$$\boldsymbol{\varepsilon}_{T+\tau} = \mathbf{B}_S^{-1}\mathbf{u}_{T+\tau} = \mathbf{B}_S^{-1}\left[diag\left(\boldsymbol{\iota}_{M_i}\right)\boldsymbol{\alpha} + \boldsymbol{\mu} + \mathbf{v}_{T+\tau}\right], \tag{32}$$

where $\mathbf{B}_S = \mathbf{I}_S - \rho\mathbf{W}_S$ and $\mathbf{I}_S = diag\left(\mathbf{I}_{M_i}\right)$ and $S = \sum_{i=1}^{N} M_i$. Moreover, using (8) and (15) the full vector disturbances $\boldsymbol{\varepsilon}$ is given by:

$$\begin{aligned}\boldsymbol{\varepsilon} &= \left(\mathbf{I}_T \otimes \mathbf{B}_S^{-1}\right)\left[\left(\boldsymbol{\iota}_T \otimes diag\left(\boldsymbol{\iota}_{M_i}\right)\right)\boldsymbol{\alpha} + \left(\boldsymbol{\iota}_T \otimes \mathbf{I}_S\right)\boldsymbol{\mu} + \mathbf{v}\right] \\ &= \left(\boldsymbol{\iota}_T \otimes \mathbf{B}_S^{-1} diag\left(\boldsymbol{\iota}_{M_i}\right)\right)\boldsymbol{\alpha} + \left(\boldsymbol{\iota}_T \otimes \mathbf{B}_S^{-1}\right)\boldsymbol{\mu} + \left(\mathbf{I}_T \otimes \mathbf{B}_S^{-1}\right)\mathbf{v}. \end{aligned} \tag{33}$$

Replacing (32) and (33) in $E\left[\boldsymbol{\varepsilon}_{T+\tau}\boldsymbol{\varepsilon}^\top\right]$ gives

$$\begin{aligned}E\left[\boldsymbol{\varepsilon}_{T+\tau}\boldsymbol{\varepsilon}^\top\right] &= E\Big[\mathbf{B}_S^{-1}\left(diag\left(\boldsymbol{\iota}_{M_i}\right)\boldsymbol{\alpha} + \boldsymbol{\mu} + \mathbf{v}_{T+\tau}\right) \\ &\quad \Big[\left(\boldsymbol{\iota}_T \otimes \mathbf{B}_S^{-1} diag\left(\boldsymbol{\iota}_{M_i}\right)\right)\boldsymbol{\alpha} + \left(\boldsymbol{\iota}_T \otimes \mathbf{B}_S^{-1}\right)\boldsymbol{\mu} \\ &\quad + \left(\mathbf{I}_T \otimes \mathbf{B}_S^{-1}\right)\mathbf{v}\Big]^\top\Big]. \end{aligned} \tag{34}$$

Using $E\left[\boldsymbol{\alpha}\boldsymbol{\alpha}^\top\right] = \sigma_\alpha^2\mathbf{I}_N$, $E\left[\boldsymbol{\mu}\boldsymbol{\mu}^\top\right] = \sigma_\mu^2\mathbf{I}_S$, $E\left[\boldsymbol{\alpha}\mathbf{v}^\top\right] = \mathbf{0}$, $E\left[\mathbf{v}_{T+\tau}\boldsymbol{\alpha}^\top\right] = \mathbf{0}$, $E\left[\boldsymbol{\mu}\boldsymbol{\alpha}^\top\right] = \mathbf{0}$, $E\left[\boldsymbol{\mu}\mathbf{v}^\top\right] = \mathbf{0}$, $E\left[\mathbf{v}_{T+\tau}\boldsymbol{\mu}^\top\right] = \mathbf{0}$ and $E\left[\mathbf{v}_{T+\tau}\mathbf{v}^\top\right] = \mathbf{0}$, the formula (34) can be simplified as

$$\begin{aligned}E\left[\boldsymbol{\varepsilon}_{T+\tau}\boldsymbol{\varepsilon}^\top\right] &= \sigma_\alpha^2\mathbf{B}_S^{-1} diag\left(\boldsymbol{\iota}_{M_i}\right)\left(\boldsymbol{\iota}_T \otimes \mathbf{B}_S^{-1} diag\left(\boldsymbol{\iota}_{M_i}\right)\right)^\top \\ &\quad + \sigma_\mu^2\mathbf{B}_S^{-1}\left(\boldsymbol{\iota}_T \otimes \mathbf{B}_S^{-1}\right)^\top \\ &= \sigma_\alpha^2\mathbf{B}_S^{-1} diag\left(\boldsymbol{\iota}_{M_i}\right)\left(\boldsymbol{\iota}_T^\top \otimes diag\left(\boldsymbol{\iota}_{M_i}\right)^\top\left(\mathbf{B}_S^{-1}\right)^\top\right) \\ &\quad + \sigma_\mu^2\mathbf{B}_S^{-1}\left(\boldsymbol{\iota}_T^\top \otimes \left(\mathbf{B}_S^{-1}\right)^\top\right), \end{aligned} \tag{35}$$

and the expression of this covariance for the disturbance of county j in state i at time period $T+\tau$ is given by

$$\begin{aligned}\boldsymbol{\omega}^\top &= E\left[\varepsilon_{ij,T+\tau}\boldsymbol{\varepsilon}^\top\right] = \sigma_\alpha^2\mathbf{b}_{ij} diag\left(\boldsymbol{\iota}_{M_i}\right)\left(\boldsymbol{\iota}_T^\top \otimes diag\left(\boldsymbol{\iota}_{M_i}\right)^\top\left(\mathbf{B}_S^{-1}\right)^\top\right) \\ &\quad + \sigma_\mu^2\mathbf{b}_{ij}\left(\boldsymbol{\iota}_T^\top \otimes \left(\mathbf{B}_S^{-1}\right)^\top\right), \end{aligned} \tag{36}$$

where $\mathbf{b}_{ij}$ is the ijth row of the matrix $\mathbf{B}_S^{-1}$. From (17), we have

$$\begin{aligned}\boldsymbol{\Omega}_\varepsilon^{-1} &= \sigma_v^{-2}\left(\mathbf{I}_T\otimes\mathbf{B}_S\right)^\top\left[\mathbf{Q}_1+\theta_1^{-1}\mathbf{Q}_2\right.\\&\quad\left.+\left(\mathbf{I}_T\otimes diag\left(\theta_{2i}^{-1}\mathbf{I}_{M_i}\right)\right)\mathbf{Q}_3\right]\left(\mathbf{I}_T\otimes\mathbf{B}_S\right)\\&= \sigma_v^{-2}\left[\left(\mathbf{E}_T\otimes\mathbf{B}_S^\top\mathbf{B}_S\right)+\theta_1^{-1}\left(\overline{\mathbf{J}}_T\otimes\mathbf{B}_S^\top diag\left(\mathbf{E}_{M_i}\right)\mathbf{B}_S\right)\right.\\&\quad\left.+\left(\overline{\mathbf{J}}_T\otimes\mathbf{B}_S^\top diag\left(\theta_{2i}^{-1}\overline{\mathbf{J}}_{M_i}\right)\mathbf{B}_S\right)\right]\end{aligned}\tag{37}$$

with $\theta_1=(T\delta_2+1)$, $\theta_{2i}=(M_iT\delta_1+T\delta_2+1)$, $\delta_1=\sigma_\alpha^2/\sigma_v^2$ and $\delta_2=\sigma_\mu^2/\sigma_v^2$. The product $\boldsymbol{\omega}^\top\boldsymbol{\Omega}_\varepsilon^{-1}$ is a function of the following products

$$\left(\boldsymbol{\iota}_T^\top\otimes diag\left(\boldsymbol{\iota}_{M_i}\right)^\top\left(\mathbf{B}_S^{-1}\right)^\top\right)\left(\mathbf{E}_T\otimes\mathbf{B}_S^\top\mathbf{B}_S\right)=\mathbf{0}\tag{38}$$

because $\boldsymbol{\iota}_T^\top\mathbf{E}_T=\mathbf{0}$,

$$\left(\boldsymbol{\iota}_T^\top\otimes diag\left(\boldsymbol{\iota}_{M_i}\right)^\top\left(\mathbf{B}_S^{-1}\right)^\top\right)\left(\overline{\mathbf{J}}_T\otimes\mathbf{B}_S^\top diag\left(\mathbf{E}_{M_i}\right)\mathbf{B}_S\right)=\mathbf{0}\tag{39}$$

because $diag\left(\boldsymbol{\iota}_{M_i}^\top\mathbf{E}_{M_i}\right)=\mathbf{0}$,

$$\left(\boldsymbol{\iota}_T^\top\otimes diag\left(\boldsymbol{\iota}_{M_i}\right)^\top\left(\mathbf{B}_S^{-1}\right)^\top\right)\left(\overline{\mathbf{J}}_T\otimes\mathbf{B}_S^\top diag\left(\theta_{2i}^{-1}\overline{\mathbf{J}}_{M_i}\right)\mathbf{B}_S\right)=\boldsymbol{\iota}_T^\top\otimes diag\left(\theta_{2i}^{-1}\boldsymbol{\iota}_{M_i}^\top\right)\mathbf{B}_S,\tag{40}$$

$$\left(\boldsymbol{\iota}_T^\top\otimes\left(\mathbf{B}_S^{-1}\right)^\top\right)\left(\mathbf{E}_T\otimes\mathbf{B}_S^\top\mathbf{B}_S\right)=\mathbf{0}\tag{41}$$

because $\boldsymbol{\iota}_T^\top\mathbf{E}_T=\mathbf{0}$,

$$\left(\boldsymbol{\iota}_T^\top\otimes\left(\mathbf{B}_S^{-1}\right)^\top\right)\left(\overline{\mathbf{J}}_T\otimes\mathbf{B}_S^\top diag\left(\mathbf{E}_{M_i}\right)\mathbf{B}_S\right)=\boldsymbol{\iota}_T^\top\otimes diag\left(\mathbf{E}_{M_i}\right)\mathbf{B}_S,\tag{42}$$

$$\left(\boldsymbol{\iota}_T^\top\otimes\left(\mathbf{B}_S^{-1}\right)^\top\right)\left(\overline{\mathbf{J}}_T\otimes\mathbf{B}_S^\top diag\left(\theta_{2i}^{-1}\overline{\mathbf{J}}_{M_i}\right)\mathbf{B}_S\right)=\boldsymbol{\iota}_T^\top\otimes diag\left(\theta_{2i}^{-1}\overline{\mathbf{J}}_{M_i}\right)\mathbf{B}_S.\tag{43}$$

Thus, using (38), (39), (40), (41), (42) and (43), the product $\boldsymbol{\omega}^\top\boldsymbol{\Omega}_\varepsilon^{-1}$ is reduced to

$$\begin{aligned}\boldsymbol{\omega}^\top\boldsymbol{\Omega}_\varepsilon^{-1} &= \frac{\sigma_\alpha^2}{\sigma_v^2}\mathbf{b}_{ij}diag\left(\boldsymbol{\iota}_{M_i}\right)\left[\boldsymbol{\iota}_T^\top\otimes diag\left(\theta_{2i}^{-1}\boldsymbol{\iota}_{M_i}^\top\right)\mathbf{B}_S\right]\\&\quad+\frac{\sigma_\mu^2}{\sigma_v^2}\mathbf{b}_{ij}\left[\theta_1^{-1}\left[\boldsymbol{\iota}_T^\top\otimes diag\left(\mathbf{E}_{M_i}\right)\mathbf{B}_S\right]+\left[\boldsymbol{\iota}_T^\top\otimes diag\left(\theta_{2i}^{-1}\overline{\mathbf{J}}_{M_i}\right)\mathbf{B}_S\right]\right]\\&= \frac{\sigma_\alpha^2}{\sigma_v^2}\mathbf{b}_{ij}diag\left(\boldsymbol{\iota}_{M_i}\right)\left[\boldsymbol{\iota}_T^\top\otimes diag\left(\theta_{2i}^{-1}\boldsymbol{\iota}_{M_i}^\top\right)\right]\left[\mathbf{I}_T\otimes\mathbf{B}_S\right]\\&\quad+\frac{\sigma_\mu^2}{\sigma_v^2}\mathbf{b}_{ij}\left[\boldsymbol{\iota}_T^\top\otimes\left(\theta_1^{-1}diag\left(\mathbf{E}_{M_i}\right)+diag\left(\theta_{2i}^{-1}\overline{\mathbf{J}}_{M_i}\right)\right)\right]\left[\mathbf{I}_T\otimes\mathbf{B}_S\right].\end{aligned}\tag{44}$$

The BLUP of the model (1) is given by:

$$\begin{aligned}\widehat{y}_{ij,T+\tau} &= \mathbf{x}_{ij,T+\tau}\widehat{\boldsymbol{\beta}}_{GLS} \\ &+\frac{\sigma_\alpha^2}{\sigma_v^2}\mathbf{b}_{ij} diag\left(\boldsymbol{\iota}_{M_i}\right)\left[\boldsymbol{\iota}_T^\top \otimes diag\left(\theta_{2i}^{-1}\boldsymbol{\iota}_{M_i}^\top\right)\right]\left[\mathbf{I}_T \otimes \mathbf{B}_S\right]\widehat{\boldsymbol{\varepsilon}}_{GLS} \\ &+\frac{\sigma_\mu^2}{\sigma_v^2}\mathbf{b}_{ij}\left[\boldsymbol{\iota}_T^\top \otimes \left[\theta_1^{-1} diag\left(\mathbf{E}_{M_i}\right) + diag\left(\theta_{2i}^{-1}\overline{\mathbf{J}}_{M_i}\right)\right]\right]\left[\mathbf{I}_T \otimes \mathbf{B}_S\right]\widehat{\boldsymbol{\varepsilon}}_{GLS}.\end{aligned} \tag{45}$$

If $\rho = 0$ (no spatial autocorrelation), we obtain

$$\frac{\sigma_\alpha^2}{\sigma_v^2}\mathbf{b}_{ij} diag\left(\boldsymbol{\iota}_{M_i}\right)\left[\boldsymbol{\iota}_T^\top \otimes diag\left(\theta_{2i}^{-1}\boldsymbol{\iota}_{M_i}^\top\right)\right]\left[\mathbf{I}_T \otimes \mathbf{B}_S\right]\widehat{\boldsymbol{\varepsilon}}_{GLS} = \frac{M_i T\sigma_\alpha^2}{\lambda_{3i}}\widehat{\boldsymbol{\varepsilon}}_{i..,GLS}, \tag{46}$$

$$\begin{aligned}\frac{\sigma_\mu^2}{\sigma_v^2}\mathbf{b}_{ij}\left[\boldsymbol{\iota}_T^\top \otimes \left[\theta_1^{-1} diag\left(\mathbf{E}_{M_i}\right) + diag\left(\theta_{2i}^{-1}\overline{\mathbf{J}}_{M_i}\right)\right]\right]\left[\mathbf{I}_T \otimes \mathbf{B}_S\right]\widehat{\boldsymbol{\varepsilon}}_{GLS} &= \frac{T\sigma_\mu^2}{\lambda_2}\left[\widehat{\boldsymbol{\varepsilon}}_{ij..,GLS} - \widehat{\boldsymbol{\varepsilon}}_{i..,GLS}\right] \\ &+\frac{T\sigma_\mu^2}{\lambda_{3i}}\widehat{\boldsymbol{\varepsilon}}_{i..,GLS}.\end{aligned} \tag{47}$$

The sum of (46) and (47) gives

$$\boldsymbol{\omega}^\top \boldsymbol{\Omega}_\varepsilon^{-1}\widehat{\boldsymbol{\varepsilon}}_{GLS} = \left(\frac{T\sigma_\mu^2}{\lambda_2}\right)\widehat{\boldsymbol{\varepsilon}}_{ij.,GLS} + \left[\frac{M_i T\sigma_\alpha^2\sigma_v^2}{\lambda_2\lambda_{3i}}\right]\widehat{\boldsymbol{\varepsilon}}_{i..,GLS}. \tag{48}$$

This term corresponds to the second term of the BLUP for the unbalanced nested error components model derived by Baltagi and Pirotte (2013).

Table 1 - $(M_1, \ldots, M_N)$ patterns considering $(N, T) = (10, 5)$ and their unbalancedness measures

Pattern	$(M_1, \ldots, M_{10})$	c
P_1	$(6, 6, 6, 6, 6, 6, 6, 6, 6, 6)$	1
P_2	$(2, 2, 3, 3, 3, 3, 5, 13, 13, 13)$	0.603
P_3	$(2, 2, 2, 2, 2, 10, 10, 10, 10, 10)$	0.555

Table 2 - RMSE Forecasts for the Spatial Nested Error Components Model, $\rho = 0.2$, (N,T)=(10,5), W(1,1), 1,000 replications

			1st year					5th year					Average				
	γ_1	γ_2	OLS	FE	RE	NE	SNE	OLS	FE	RE	NE	SNE	OLS	FE	RE	NE	SNE
P_1	0.0	0.0	1.8571	2.0361	1.8597	1.8656	1.8549	2.2795	2.4917	2.2832	2.2907	2.2548	2.1238	2.3228	2.1270	2.1338	2.1111
		0.2	1.8500	1.8212	1.7533	1.7561	1.7468	2.2474	2.2286	2.1477	2.1500	2.1130	2.1014	2.0775	2.0016	2.0041	1.9817
		0.4	1.8399	1.5772	1.5458	1.5471	1.5421	2.1997	1.9300	1.8931	1.8942	1.8742	2.0679	1.7992	1.7644	1.7655	1.7534
		0.6	1.8303	1.2878	1.2743	1.2747	1.2712	2.1322	1.5759	1.5602	1.5606	1.5448	2.0212	1.4690	1.4541	1.4546	1.4452
		0.8	1.8196	0.9106	0.9065	0.9066	0.9055	2.0359	1.1143	1.1097	1.1098	1.1052	1.9545	1.0388	1.0343	1.0344	1.0316
	0.2	0.0	1.8977	1.8212	1.7602	1.6951	1.6900	2.3088	2.2286	2.1572	2.0815	2.0567	2.1570	2.0775	2.0099	1.9384	1.9239
		0.2	1.8897	1.5772	1.5496	1.5369	1.5310	2.2588	1.9300	1.8968	1.8819	1.8549	2.1226	1.7992	1.7677	1.7539	1.7381
		0.4	1.8805	1.2878	1.2758	1.2722	1.2699	2.1898	1.5759	1.5616	1.5577	1.5469	2.0751	1.4690	1.4554	1.4518	1.4454
		0.6	1.8717	0.9106	0.9070	0.9061	0.9055	2.0927	1.1143	1.1101	1.1092	1.1060	2.0086	1.0388	1.0347	1.0339	1.0320
	0.4	0.0	1.9336	1.5772	1.5506	1.4712	1.4648	2.3093	1.9300	1.8984	1.8064	1.7809	2.1706	1.7992	1.7691	1.6822	1.6668
		0.2	1.9275	1.2878	1.2765	1.2630	1.2604	2.2398	1.5759	1.5623	1.5463	1.5356	2.1233	1.4690	1.4560	1.4412	1.4347
		0.4	1.9189	0.9106	0.9072	0.9045	0.9037	2.1421	1.1143	1.1103	1.1073	1.1039	2.0565	1.0388	1.0349	1.0320	1.0300
	0.6	0.0	1.9691	1.2878	1.2771	1.2042	1.1978	2.2860	1.5759	1.5628	1.4784	1.4599	2.1675	1.4690	1.4566	1.3768	1.3646
		0.2	1.9640	0.9106	0.9074	0.9001	0.8992	2.1891	1.1143	1.1104	1.1021	1.0979	2.1019	1.0388	1.0350	1.0271	1.0246
	0.8	0.0	2.0030	0.9106	0.9077	0.8566	0.8485	2.2324	1.1143	1.1106	1.0514	1.0382	2.1430	1.0388	1.0352	0.9792	0.9684
P_2	0.0	0.0	1.8571	2.0361	1.8597	1.8647	1.8513	2.2795	2.4917	2.2832	2.2893	2.2417	2.1238	2.3228	2.1270	2.1327	2.1033
		0.2	1.8500	1.8212	1.7533	1.7545	1.7457	2.2474	2.2286	2.1477	2.1492	2.1126	2.1014	2.0775	2.0016	2.0032	1.9813
		0.4	1.8399	1.5772	1.5458	1.5463	1.5401	2.1997	1.9300	1.8931	1.8938	1.8650	2.0679	1.7992	1.7644	1.7651	1.7488
		0.6	1.8303	1.2878	1.2743	1.2745	1.2707	2.1322	1.5759	1.5602	1.5604	1.5443	2.0212	1.4690	1.4541	1.4544	1.4448
		0.8	1.8196	0.9106	0.9065	0.9065	0.9054	2.0359	1.1143	1.1097	1.1097	1.1049	1.9545	1.0388	1.0343	1.0344	1.0314
	0.2	0.0	1.8811	1.8212	1.7560	1.6915	1.6868	2.2790	2.2286	2.1495	2.0775	2.0535	2.1343	2.0775	2.0043	1.9353	1.9215
		0.2	1.8711	1.5777	1.5479	1.5355	1.5296	2.2275	1.9300	1.8941	1.8798	1.8532	2.0980	1.7992	1.7657	1.7526	1.7369
		0.4	1.8595	1.2878	1.2754	1.2718	1.2662	2.1585	1.5759	1.5608	1.5567	1.5306	2.0498	1.4690	1.4549	1.4512	1.4360
		0.6	1.8489	0.9106	0.9069	0.9061	0.9050	2.0612	1.1143	1.1099	1.1089	1.1038	1.9822	1.0388	1.0346	1.0337	1.0307
	0.4	0.0	1.9018	1.5772	1.5478	1.4680	1.4616	2.2557	1.9300	1.8929	1.8033	1.7786	2.1281	1.7992	1.7653	1.6798	1.6648
		0.2	1.8880	1.2878	1.2755	1.2623	1.2584	2.1837	1.5759	1.5603	1.5454	1.5291	2.0768	1.4690	1.4547	1.4407	1.4310
		0.4	1.8772	0.9106	0.9071	0.9045	0.9032	2.0856	1.1143	1.1098	1.1069	1.1014	2.0085	1.0388	1.0346	1.0319	1.0286
	0.6	0.0	1.9191	1.2878	1.2755	1.2011	1.1958	2.2104	1.5759	1.5598	1.4756	1.4622	2.1057	1.4690	1.4545	1.3745	1.3653
		0.2	1.9049	0.9106	0.9072	0.9001	0.8989	2.1097	1.1143	1.1096	1.1017	1.0977	2.0342	1.0388	1.0345	1.0271	1.0246
	0.8	0.0	1.9337	0.9106	0.9073	0.8535	0.8465	2.1352	1.1143	1.1096	1.0483	1.0371	2.0617	1.0388	1.0346	0.9766	0.9673
P_3	0.0	0.0	1.8571	2.0361	1.8597	1.8646	1.8535	2.2795	2.4917	2.2832	2.2886	2.2517	2.1238	2.3228	2.1270	2.1321	2.1089
		0.2	1.8500	1.8212	1.7533	1.7550	1.7475	2.2474	2.2286	2.1477	2.1495	2.1211	2.1014	2.0775	2.0016	2.0034	1.9861
		0.4	1.8399	1.5772	1.5458	1.5465	1.5413	2.1997	1.9300	1.8931	1.8940	1.8733	2.0679	1.7992	1.7644	1.7652	1.7526
		0.6	1.8303	1.2878	1.2743	1.2746	1.2718	2.1322	1.5759	1.5602	1.5606	1.5495	2.0212	1.4690	1.4541	1.4545	1.4478
		0.8	1.8196	0.9106	0.9065	0.9066	0.9058	2.0359	1.1143	1.1097	1.1098	1.1066	1.9545	1.0388	1.0343	1.0344	1.0324
	0.2	0.0	1.8889	1.8212	1.7593	1.6924	1.6872	2.2834	2.2286	2.1509	2.0753	2.0485	2.1396	2.0775	2.0065	1.9341	1.9289
		0.2	1.8776	1.5772	1.5494	1.5365	1.5314	2.2327	1.9300	1.8947	1.8799	1.8580	2.1034	1.7992	1.7666	1.7528	1.7399
		0.4	1.8664	1.2878	1.2759	1.2724	1.2695	2.1647	1.5759	1.5609	1.5569	1.5442	2.0558	1.4690	1.4552	1.4515	1.4440
		0.6	1.8571	0.9106	0.9071	0.9063	0.9055	2.0689	1.1143	1.1099	1.1090	1.1054	1.9898	1.0388	1.0346	1.0338	1.0317
	0.4	0.0	1.9136	1.5772	1.5496	1.4686	1.4635	2.2663	1.9300	1.8940	1.8012	1.7809	2.1388	1.7992	1.7667	1.6786	1.6663
		0.2	1.9017	1.2878	1.2764	1.2630	1.2598	2.1955	1.5759	1.5608	1.5453	1.5325	2.0887	1.4690	1.4553	1.4408	1.4331
		0.4	1.8908	0.9106	0.9073	0.9047	0.9039	2.0988	1.1143	1.1099	1.1070	1.1030	2.0214	1.0388	1.0347	1.0320	1.0297
	0.6	0.0	1.9368	1.2878	1.2765	1.2013	1.1967	2.2277	1.5759	1.5604	1.4736	1.4608	2.1228	1.4690	1.4553	1.3732	1.3646
		0.2	1.9245	0.9106	0.9074	0.9003	0.8992	2.1293	1.1143	1.1099	1.1017	1.0973	2.0534	1.0388	1.0348	1.0271	1.0244
	0.8	0.0	1.9588	0.9106	0.9076	0.8532	0.8472	2.1598	1.1143	1.1098	1.0464	1.0362	2.0862	1.0388	1.0349	0.9752	0.9669

Table 3 - RMSE Forecasts for the Spatial Nested Error Components Model, $\rho = 0.8$, (N,T)=(10,5) , W(1,1), 1,000 replications

			1st year					5th year					Average				
	γ_1	γ_2	OLS	FE	RE	NE	SNE	OLS	FE	RE	NE	SNE	OLS	FE	RE	NE	SNE
P_1	0.0	0.0	3.9797	4.3623	3.9931	4.0900	3.8760	5.1368	5.6786	5.1608	5.2654	4.6588	4.6711	5.1374	4.6900	4.7925	4.3895
		0.2	3.9031	3.9018	3.7539	3.8121	3.6261	4.9555	5.0791	4.8570	4.8877	4.3643	4.5371	4.5951	4.4091	4.4548	4.1084
		0.4	3.8291	3.3790	3.3079	3.3469	3.1950	4.7500	4.3987	4.2952	4.3138	3.8514	4.3886	3.9794	3.8919	3.9215	3.6222
		0.6	3.7594	2.7590	2.7267	2.7435	2.6361	4.5087	3.5915	3.5473	3.5541	3.1857	4.2173	3.2492	3.2110	3.2231	2.9920
		0.8	3.6958	1.9509	1.9405	1.9453	1.8799	4.2096	2.5396	2.5266	2.5283	2.2870	4.0086	2.2975	2.2859	2.2892	2.1397
	0.2	0.0	4.6538	3.9018	3.8390	3.7850	3.6108	5.7283	5.0791	4.9775	4.8542	4.3406	5.3059	4.5951	4.5116	4.4240	4.0875
		0.2	4.5998	3.7090	3.3439	3.3512	3.2050	5.5297	4.3987	4.3425	4.3061	3.8603	5.1656	3.9794	3.9328	3.9197	3.6313
		0.4	4.5462	2.7590	2.7420	2.7472	2.6423	5.2942	3.5915	3.5648	3.5525	3.1942	5.0011	3.2492	3.2268	3.2239	2.9989
		0.6	4.4969	1.9509	1.9455	1.9466	1.8823	5.0065	2.5396	2.5312	2.5279	2.2902	4.8025	2.2975	2.2903	2.2894	2.1427
	0.4	0.0	5.2369	3.3790	3.3562	3.3334	3.1357	6.1565	4.3987	4.3600	4.2751	3.7733	5.7956	3.9794	3.9477	3.8953	3.5510
		0.2	5.1979	2.7590	2.7476	2.7472	2.6325	5.9355	3.5915	3.5724	3.5446	3.1823	5.6425	3.2492	3.2333	3.2198	2.9877
		0.4	5.1554	1.9509	1.9473	1.9467	1.8814	5.6599	2.5396	2.5335	2.5260	2.2903	5.4536	2.2975	2.2923	2.2885	2.1421
	0.6	0.0	5.7586	2.7590	2.7505	2.7456	2.5637	6.4897	3.5915	3.5768	3.5326	3.0881	6.1976	3.2492	3.2370	3.2132	2.9046
		0.2	5.7260	1.9509	1.9483	1.9466	1.8755	6.2333	2.5396	2.5348	2.5230	2.2815	6.0223	2.2975	2.2935	2.2869	2.1347
	0.8	0.0	6.2321	1.9509	1.9488	1.9470	1.8146	6.7411	2.5396	2.5357	2.5195	2.1882	6.5267	2.2975	2.2942	2.2854	2.0570
P_2	0.0	0.0	3.9797	4.3623	3.9931	4.0590	3.8747	5.1368	5.6786	5.1608	5.2299	4.6566	4.6711	5.1374	4.6900	4.7604	4.3881
		0.2	3.9031	3.9018	3.7539	3.7899	3.6241	4.9555	5.0791	4.8570	4.8751	4.3630	4.5371	4.5951	4.4091	4.4389	4.1072
		0.4	3.8291	3.3790	3.3079	3.3326	3.1952	4.7500	4.3987	4.2952	4.3059	3.8566	4.3886	3.9794	3.8919	3.9111	3.6244
		0.6	3.7594	2.7590	2.7267	2.7376	2.6351	4.5087	3.5915	3.5473	3.5514	3.1861	4.2173	3.2492	3.2110	3.2191	3.9912
		0.8	3.6958	1.9509	1.9405	1.9438	1.8804	4.2096	2.5396	2.5266	2.5278	2.2916	4.0086	2.2975	2.2859	2.2883	2.1421
	0.2	0.0	4.6241	3.9018	3.8349	3.7445	3.5826	5.6513	5.0791	4.9657	4.8100	4.3131	5.2528	4.5951	4.5068	4.3827	4.0625
		0.2	4.5487	3.3790	3.3408	3.3324	3.1967	5.4387	4.3987	4.3376	4.2925	3.8554	5.0961	3.9794	3.9310	3.9052	3.6262
		0.4	4.4861	2.7590	2.7404	2.7400	2.6397	5.1991	3.5915	3.5637	3.5483	3.1926	4.9254	3.2492	3.2266	3.2191	3.9978
		0.6	4.4243	1.9509	1.9448	1.9448	1.8829	4.9067	2.5396	2.5313	2.5273	2.2947	4.7188	2.2975	2.2906	2.2887	2.1453
	0.4	0.0	5.1692	3.3790	3.3534	3.2967	3.1085	6.0232	4.3987	4.3535	4.2302	3.7494	5.6965	3.9794	3.9454	3.8547	3.5280
		0.2	5.1011	2.7590	2.7453	2.7369	2.6265	5.7842	3.5915	3.5693	3.5359	3.1779	5.5227	3.2492	3.2320	3.2116	2.9838
		0.4	5.0423	1.9509	1.9463	1.9444	1.8805	5.5019	2.5396	2.5327	2.5249	2.2923	5.3222	2.2975	2.2921	2.2873	2.1431
	0.6	0.0	5.6539	2.7590	2.7491	2.7238	2.5405	6.3096	3.5915	3.5738	3.5016	3.0675	6.0590	3.2492	3.2362	3.1871	2.8848
		0.2	5.5921	1.9509	1.9474	1.9432	1.8731	6.0326	2.5396	2.5337	2.5200	2.2795	5.8597	2.2975	2.2931	2.2843	2.1332
	0.8	0.0	6.0931	1.9509	1.9484	1.9417	1.7982	6.5196	2.5396	2.5348	2.5082	2.1737	6.3521	2.2975	2.2942	2.2774	2.0428
P_3	0.0	0.0	3.9797	4.3623	3.9931	4.0636	3.8768	5.1368	5.6786	5.1608	5.2350	4.6577	4.6711	5.1374	4.6900	4.7643	4.3883
		0.2	3.9031	3.9018	3.7539	3.7960	3.6235	4.9555	5.0791	4.8570	4.8810	4.3669	4.5371	4.5951	4.4091	4.4434	4.1080
		0.4	3.8291	3.3790	3.3079	3.3337	3.1941	4.7500	4.3987	4.2952	4.3081	3.8559	4.3886	3.9794	3.8919	3.9122	3.6233
		0.6	3.7594	2.7590	2.7267	2.7374	2.6363	4.5087	3.5915	3.5473	3.5519	3.1914	4.2173	3.2492	3.2110	3.2191	2.9938
		0.8	3.6958	1.9509	1.9405	1.9435	1.8806	4.2096	2.5396	2.5266	2.5277	2.2902	4.0086	2.2975	2.2859	2.2880	2.1413
	0.2	0.0	4.6685	3.9018	3.8403	3.7462	3.5882	5.6778	5.0791	4.9715	4.8102	4.3128	5.2867	4.5951	4.5116	4.3831	4.0627
		0.2	4.5922	3.3790	3.3419	3.3323	3.1979	5.4666	4.3987	4.3400	4.2937	3.8547	5.1315	3.9794	3.9322	3.9050	3.6252
		0.4	4.5310	2.7590	2.7402	2.7382	2.6400	5.2275	3.5915	3.5640	3.5488	3.1954	4.9613	3.2492	3.2265	3.2185	2.9984
		0.6	4.4728	1.9509	1.9444	1.9437	1.8826	4.9383	2.5396	2.5310	2.5271	2.2941	4.7578	2.2975	2.2902	2.2881	2.1447
	0.4	0.0	5.2449	3.3790	3.3555	3.2931	3.1117	6.0759	4.3987	4.3559	4.2277	3.7463	5.7604	3.9794	3.9470	3.8522	3.5259
		0.2	5.1797	2.7590	2.7456	2.7346	2.6268	5.8396	3.5915	3.5706	3.5358	3.1787	5.5901	3.2492	3.2325	3.2102	2.9835
		0.4	5.1218	1.9509	1.9462	1.9428	1.8805	5.5604	2.5396	2.5329	2.5246	2.2923	5.3918	2.2975	2.2920	2.2866	2.1429
	0.6	0.0	5.7558	2.7590	2.7495	2.7178	2.5431	6.3853	3.5915	3.5745	3.4988	3.0656	6.1486	3.2492	3.2365	3.1838	2.8833
		0.2	5.6952	1.9509	1.9472	1.9411	1.8731	6.1131	2.5396	2.5340	2.5196	2.2809	5.9534	2.2975	2.2931	2.2833	2.1336
	0.8	0.0	6.2206	1.9509	1.9482	1.9370	1.7997	6.6170	2.5396	2.5348	2.5055	2.1721	6.4660	2.2975	2.2940	2.2745	2.0416