Computational Statistics and Data Analysis 56 (2012) 3381–3397

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Forecasting with spatial panel data

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

[a] Department of Economics and Center for Policy Research, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020, USA
[b] ERMES (CNRS) and TEPP (CNRS), Université Panthéon-Assas Paris II, France
[c] INRETS-DEST, National Institute of Research on Transports and Safety, France

ARTICLE INFO

Article history:
Received 8 December 2009
Received in revised form 29 July 2010
Accepted 11 August 2010
Available online 30 August 2010

Keywords:
Forecasting
BLUP
Panel data
Spatial dependence
Heterogeneity

ABSTRACT

Various forecasts using panel data with spatial error correlation are compared using Monte Carlo experiments. The true data generating process is assumed to be a simple error component regression model with spatial remainder disturbances of the autoregressive or moving average type. The best linear unbiased predictor is compared with other forecasts ignoring spatial correlation, or ignoring heterogeneity due to the individual effects. In addition, the root mean squared error performance of these forecasts is examined under misspecification of the spatial error process, various spatial weight matrices, and heterogeneous rather than homogeneous panel data models.

1. Introduction

The literature on forecasting is rich with time series applications, but this is not the case for spatial panel data applications. Exceptions are Baltagi and Li (2004, 2006) with applications to forecasting sales of cigarettes and liquor per capita for US states over time. In order to explain how spatial autocorrelation may arise in the demand for cigarettes, we note that cigarette prices vary among states primarily due to variation in state taxes on cigarettes. Border effect purchases not included in the cigarette demand equation can cause spatial autocorrelation among the disturbances. In forecasting sales of cigarettes, the spatial autocorrelation due to neighboring states and the individual heterogeneity across states is taken explicitly into account. Best linear unbiased prediction (BLUP) in panel data using an error component model have been considered by Taub (1979), Baltagi and Li (1992), and Baillie and Baltagi (1999) to mention a few. Applications include Baltagi and Griffin (1997), Hsiao and Tahmiscioglu (1997), Schmalensee et al. (1998), Baltagi et al. (2000), Hoogstrate et al. (2000), Baltagi et al. (2002, 2004), Frees and Miller (2004), Rapach and Wohar (2004), and Brucker and Siliverstovs (2006), see Baltagi (2008) for a recent survey. However, these panel forecasting applications do not deal with spatial dependence across the panel units. Spatial dependence models – popular in regional science and urban economics – deal with spatial interaction and spatial heterogeneity (see Anselin (1988) and Anselin and Bera (1998)). The structure of the dependence can be related to location and distance, both in a geographic space as well as a more general economic or social network space. Some commonly used spatial error processes include the spatial autoregressive (SAR) and the spatial moving average (SMA) error processes. Two different variants of these models for spatial panels are considered, one discussed in Anselin (1988) and another in Kapoor et al. (2007) and Fingleton (2008a). The best linear unbiased predictors for the Anselin type model was derived by Baltagi and Li (2004). This paper derives the best linear unbiased predictors for the Kapoor et al. (2007) and Fingleton (2008a) variants. More importantly, it compares the performance of sixteen various forecasts of the spatial panel data using Monte

* Corresponding author. Tel.: +1 315 443 1630; fax: +1 315 443 1081.
E-mail addresses: bbaltagi@maxwell.syr.edu (B.H. Baltagi), bresson-georges@orange.fr (G. Bresson), apirotte@aol.com (A. Pirotte).

0167-9473/$ – see front matter
doi:10.1016/j.csda.2010.08.006

Carlo experiments. These include homogeneous as well as heterogeneous estimators of the spatial panel model and their corresponding forecasts. The true data generating process is assumed to be a simple error component regression model with spatial remainder disturbances of the autoregressive or moving average type. The best linear unbiased predictor is compared with other forecasts ignoring spatial correlation, or ignoring heterogeneity due to the individual effects. In addition, we check the performance of these forecasts under misspecification of the spatial error process, different spatial weight matrices, and various sample sizes. Section 2 introduces the error component model with spatially autocorrelated residuals of the SAR and SMA type. Section 3 describes the forecasts using the estimators considered in Section 2, while Section 4 gives the Monte Carlo design. Section 5 reports the results of the Monte Carlo simulations and Section 6 gives our summary and conclusion.

2. The error component model with spatially autocorrelated residuals

Consider a linear panel data regression model:

$$y_{it} = X_{it}\beta + \varepsilon_{it}, \quad i = 1, \ldots, N, \; t = 1, \ldots, T, \tag{1}$$

where the disturbance term follows an error component model with spatially autocorrelated residuals. The disturbance vector for time t is given by:

$$\varepsilon_t = \mu + \phi_t, \tag{2}$$

where $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$, $\mu = (\mu_1, \ldots, \mu_N)'$ denotes the vector of specific effects assumed to be $iid(0, \sigma_\mu^2)$ and $\phi_t = (\phi_{1t}, \ldots, \phi_{Nt})'$ are the remainder disturbances which are independent of μ. We let the ϕ_t's follow a spatial autoregressive (SAR) or a spatial moving average (SMA) error model. The SAR process is known to transmit the shocks globally while the SMA process transmits these shocks locally, see Anselin et al. (2008).

The SAR specification for the $(N \times 1)$ error vector ϕ_t at time t can be expressed as:

$$\phi_t = \rho W_N \phi_t + v_t = (I_N - \rho W_N)^{-1} v_t = B_N^{-1} v_t, \tag{3}$$

where W_N is an $(N \times N)$ known spatial weights matrix. In the simplest case, the weight matrix is binary, with $w_{ij} = 1$ when i and j are neighbors and $w_{ij} = 0$ when they are not. By convention, diagonal elements are null: $w_{ii} = 0$ and the weights are almost always standardized such that the elements of each row sum to 1. ρ is the spatial autoregressive parameter, $|\rho| < 1$, and v_t is an $(N \times 1)$ error vector assumed to be distributed independently across a cross-sectional dimension with constant variance $\sigma_v^2 I_N$. $B_N = (I_N - \rho W_N)$ and is assumed to be non-singular. The error covariance matrix for the cross-section at time t becomes:

$$\Omega_t = E[\varepsilon_t \varepsilon_t'] = \sigma_\mu^2 I_N + \sigma_v^2 (B_N' B_N)^{-1}. \tag{4}$$

For the full $(NT \times 1)$ vector of disturbances:

$$\varepsilon = (\iota_T \otimes I_N)\mu + (I_T \otimes B_N^{-1})v, \tag{5}$$

the corresponding $(NT \times NT)$ covariance matrix is given by:

$$\Omega = \sigma_\mu^2 (J_T \otimes I_N) + \sigma_v^2 [I_T \otimes (B_N' B_N)^{-1}], \tag{6}$$

where ι_T is a $(T \times 1)$ vector of ones and $J_T = \iota_T \iota_T'$ is a $(T \times T)$ matrix of ones.

The spatial moving average (SMA) specification for the $(N \times 1)$ error vector ϕ_t at time t can be expressed as:

$$\phi_t = \lambda W_N v_t + v_t = (I_N + \lambda W_N) v_t = D_N v_t, \tag{7}$$

where $D_N = (I_N + \lambda W_N)$, λ is the spatial moving average parameter, $|\lambda| < 1$. The error covariance matrix for the cross-section at time t becomes:

$$\Omega_t = E[\varepsilon_t \varepsilon_t'] = \sigma_\mu^2 I_N + \sigma_v^2 (D_N D_N'). \tag{8}$$

For the full $(NT \times 1)$ vector of disturbances:

$$\varepsilon = (\iota_T \otimes I_N)\mu + (I_T \otimes D_N)v, \tag{9}$$

the corresponding $(NT \times NT)$ covariance matrix is given by:

$$\Omega = \sigma_\mu^2 (J_T \otimes I_N) + \sigma_v^2 [I_T \otimes (D_N D_N')]. \tag{10}$$

MLE under normality of the disturbances using these error component models with spatial autocorrelation have been derived by Anselin (1988). The log-likelihood is given by:

$$L \propto -\frac{NT}{2} \ln(2\pi\sigma_v^2) - \frac{1}{2} \ln|\Sigma| - \frac{1}{2\sigma_v^2} \varepsilon' \Sigma^{-1} \varepsilon, \tag{11}$$

where

$$\varepsilon = y - X\beta, \qquad \Omega = \sigma_v^2 \Sigma$$

$$\Sigma = \begin{cases} (J_T \otimes \theta I_N) + [I_T \otimes (B_N' B_N)^{-1}] & \text{for SAR} \\ (J_T \otimes \theta I_N) + [I_T \otimes (D_N D_N')] & \text{for SMA} \end{cases} \tag{12}$$

with $\theta = \sigma_\mu^2/\sigma_v^2$.

Regression models containing spatially correlated disturbance terms based on the SAR or SMA models are typically estimated using MLE, where the likelihood function corresponds to the normal distribution. However, this can be computationally demanding for large N. Kelejian and Prucha (1999) suggested a generalized moments (GM) estimation method for the SAR model in a cross-section setting, and Fingleton (2008b) extended this generalized moments estimator to the SMA model. Kapoor et al. (2007) generalized this GM procedure from cross-section to panel data and derived its large sample properties when T is fixed and $N \to \infty$. However, their SAR random effects model (SAR-RE) differs from that described in (2) which we will call (RE-SAR). In fact, in their specification, the disturbance term ε_t itself follows a SAR process and the remainder term follows an error component structure. This allows the individual effects, i.e., the μ's themselves to be spatially correlated but with the same ρ. In particular, the disturbance vector for time t is given by:

$$\varepsilon_t = \rho W_N \varepsilon_t + u_t, \tag{13}$$

where u_t follows an error component structure:

$$u_t = \mu + v_t. \tag{14}$$

The SAR-RE specification for the $(N \times 1)$ error vector ε_t at time t can be expressed as:

$$\varepsilon_t = (I_N - \rho W_N)^{-1} u_t = B_N^{-1} u_t. \tag{15}$$

For the full $(NT \times 1)$ vector of disturbances:

$$\varepsilon = (\iota_T \otimes B_N^{-1})\mu + (I_T \otimes B_N^{-1})v, \tag{16}$$

and the corresponding $(NT \times NT)$ covariance matrix is given by:

$$\Omega = \sigma_\mu^2 (J_T \otimes (B_N' B_N)^{-1}) + \sigma_v^2 [I_T \otimes (B_N' B_N)^{-1}]. \tag{17}$$

Kapoor et al. (2007) proposed three generalized moments (GM) estimators of ρ, σ_v^2 and σ_1^2 $(= \sigma_v^2 + T\sigma_\mu^2)$ based on the following six moment conditions:

$$E \begin{bmatrix} \frac{1}{N(T-1)} u_N' Q_{0,N} u_N \\ \frac{1}{N(T-1)} \overline{u}_N' Q_{0,N} \overline{u}_N \\ \frac{1}{N(T-1)} \overline{u}_N' Q_{0,N} u_N \\ \frac{1}{N} u_N' Q_{1,N} u_N \\ \frac{1}{N} \overline{u}_N' Q_{1,N} \overline{u}_N \\ \frac{1}{N} \overline{u}_N' Q_{1,N} u_N \end{bmatrix} = \begin{bmatrix} \sigma_v^2 \\ \sigma_v^2 \frac{1}{N} tr(W_N' W_N) \\ 0 \\ \sigma_1^2 \\ \sigma_1^2 \frac{1}{N} tr(W_N' W_N) \\ 0 \end{bmatrix}, \tag{18}$$

where

$$u_N = \varepsilon_N - \rho \overline{\varepsilon}_N \tag{19}$$

$$\overline{u}_N = \overline{\varepsilon}_N - \rho \overline{\overline{\varepsilon}}_N \tag{20}$$

$$\overline{\varepsilon}_N = (I_T \otimes W_N)\varepsilon_N \tag{21}$$

$$\overline{\overline{\varepsilon}}_N = (I_T \otimes W_N)\overline{\varepsilon}_N \tag{22}$$

$$Q_{0,N} = \left(I_T - \frac{J_T}{T}\right) \otimes I_N \tag{23}$$

$$Q_{1,N} = \frac{J_T}{T} \otimes I_N. \tag{24}$$

Under the random effects specification considered, the OLS estimator of β is consistent. Using $\widehat{\beta}_{OLS}$ one gets a consistent estimator of the disturbances $\widehat{\varepsilon} = y - X\widehat{\beta}_{OLS}$. The GM estimators of σ_1^2, σ_v^2 and ρ are the solution of the sample counterpart of the six equations given above. Kapoor et al. (2007) suggest three GM estimators. The first involves only the first three moments which do not involve σ_1^2 and yield estimates of ρ and σ_v^2. The fourth moment condition is then used to solve for σ_1^2 given estimates of ρ and σ_v^2. The second GM estimator is based upon weighing the moment equations by the inverse of a properly normalized variance-covariance matrix of the sample moments evaluated at the true parameter values. A simple version of this weighting matrix is derived under normality of the disturbances. The third GM estimator is motivated by computational considerations and replaces a component of the weighting matrix for the second GM estimator by an identity matrix. Kapoor et al. (2007) perform Monte Carlo experiments comparing MLE and these three GM estimation methods. They find that on average, the RMSE of MLE and their weighted GM estimators are quite similar. The feasible GLS estimator of β is then obtained by replacing ρ, σ_v^2 and σ_1^2 by their GM estimators. Later, in our Monte Carlo experiments, we computed the predictors for all three GM estimators suggested by Kapoor et al. (2007). However, the differences in root mean squared error performance were minor. To save space, we only report the second GM estimator, called the weighted GM estimator by Kapoor et al. (2007).

Recently, Fingleton (2008a) extended this GM estimator for the SMA panel data model with random effects. We call this SMA-RE to distinguish it from the RE-SMA procedure described in Anselin et al. (2008). In fact, for the Fingleton (2008a) SMA-RE, the disturbance term ε_t in (2) follows a SMA process and the remainder term follows an error component structure. Unlike the Anselin et al. (2008) RE-SMA, the individual effects, i.e., the μ's themselves are allowed to be spatially correlated but with the same λ. In particular, the disturbance vector for time t is given by:

$$\varepsilon_t = (I_N + \lambda W_N)u_t = D_N u_t, \tag{25}$$

where $D_N = (I_N + \lambda W_N)$, and u_t follows an error component structure (14). So, the full SMA-RE ($NT \times 1$) vector of disturbances is given by:

$$\varepsilon = (\iota_T \otimes D_N)\mu + (I_T \otimes D_N)v, \tag{26}$$

and the corresponding ($NT \times NT$) covariance matrix is given by:

$$\Omega = \sigma_\mu^2(J_T \otimes (D_N D_N')) + \sigma_v^2[I_T \otimes (D_N D_N')]. \tag{27}$$

The moment conditions for SMA-RE are similar to those derived by Kapoor et al. (2007), see Fingleton (2008a).

3. Prediction

Goldberger (1962) has shown that, for a given Ω, the best linear unbiased predictor (BLUP) for the ith individual at a future period $T + \tau$ is given by:

$$\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{GLS} + \omega'\Omega^{-1}\widehat{\varepsilon}_{GLS}, \tag{28}$$

where $\omega = E[\varepsilon_{i,T+\tau}\varepsilon]$ is the covariance between the future disturbance $\varepsilon_{i,T+\tau}$ and the sample disturbances ε. $\widehat{\beta}_{GLS}$ is the GLS estimator of β from Eq. (1) based on Ω and $\widehat{\varepsilon}_{GLS}$ denotes the corresponding GLS residual vector.

For the error component without spatial autocorrelation ($\lambda = 0$), this BLUP reduces to:

$$\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{GLS} + \frac{\sigma_\mu^2}{\sigma_1^2}(\iota_T' \otimes l_i')\widehat{\varepsilon}_{GLS}, \tag{29}$$

where $\sigma_1^2 = T\sigma_\mu^2 + \sigma_v^2$ and l_i is the ith column of I_N. This predictor was considered by Wansbeek and Kapteyn (1978), Lee and Griffiths (1979) and Taub (1979). The typical element of the last term of Eq. (29) is $(T\sigma_\mu^2/\sigma_1^2)\overline{\varepsilon}_{i.,GLS}$ where $\overline{\varepsilon}_{i.,GLS} = \sum_{t=1}^T \widehat{\varepsilon}_{ti,GLS}/T$. Therefore, the BLUP of $y_{i,T+\tau}$ for the RE model modifies the usual GLS forecasts by adding a fraction of the mean of the GLS residuals corresponding to the ith individual. In order to make this forecast operational, $\widehat{\beta}_{GLS}$ is replaced by its feasible GLS estimate and the variance components are replaced by their feasible estimates.

Baltagi and Li (2004, 2006) derived the BLUP correction term when both error components and spatial autocorrelation are present and ϕ_t follows a SAR process. So, the predictors for the SAR and the SMA are given by:

$$\widehat{y}_{i,T+\tau} = \begin{cases} X_{i,T+\tau}\widehat{\beta}_{MLE} + \theta(\iota_T' \otimes l_i'C_1^{-1})\widehat{\varepsilon}_{MLE} = X_{i,T+\tau}\widehat{\beta}_{MLE} + T\theta\displaystyle\sum_{j=1}^N c_{1,j}\overline{\varepsilon}_{j.,MLE} & \text{for SAR} \\ X_{i,T+\tau}\widehat{\beta}_{MLE} + \theta(\iota_T' \otimes l_i'C_2^{-1})\widehat{\varepsilon}_{MLE} = X_{i,T+\tau}\widehat{\beta}_{MLE} + T\theta\displaystyle\sum_{j=1}^N c_{2,j}\overline{\varepsilon}_{j.,MLE} & \text{for SMA,} \end{cases} \tag{30}$$

where c_{1j} (resp. $c_{2,j}$) is the jth element of the ith row of C_1^{-1} (resp. C_2^{-1}) with $C_1 = [T\theta I_N + (B_N'B_N)^{-1}]$ (resp. $C_2 = [T\theta I_N + (D_N D_N')]$) and $\overline{\varepsilon}_{j.,MLE} = \sum_{t=1}^T \widehat{\varepsilon}_{tj,MLE}/T$. In other words, the BLUP of $y_{i,T+\tau}$ adds to $X_{i,T+\tau}\widehat{\beta}_{MLE}$ a weighted average of the MLE residuals for the N individuals averaged over time. The weights depend upon the spatial matrix W_N and the spatial autoregressive (or moving average) coefficients ρ and λ. To make these predictors operational, we replace θ, ρ and λ by their estimates from the RE-spatial MLE with SAR or SMA. When there are no random individual effects, so that $\sigma_\mu^2 = 0$, then $\theta = 0$ and the BLUP prediction terms drop out completely from Eq. (30). In these cases, Ω in Eq. (12) reduces to $\sigma_v^2[I_T \otimes (B_N' B_N)^{-1}]$ for SAR and $\sigma_v^2[I_T \otimes (D_N D_N')]$ for SMA, and the corresponding MLE for these models yield the pooled spatial MLE with SAR or SMA remainder disturbances.

For the Kapoor et al. (2007) model, the BLUP of $y_{i,T+\tau}$ for the SAR-RE also modifies the usual GLS forecasts by adding a fraction of the mean of the GLS residuals corresponding to the ith individual. More specifically, the predictor is given by:

$$\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{GLS} + \left(\frac{\sigma_\mu^2}{\sigma_1^2}\right) b_i(\iota_T' \otimes B_N)\widehat{\varepsilon}_{GLS}, \tag{31}$$

where b_i is the ith row of the matrix B_N^{-1}. This is derived in the Appendix of this paper which also shows that the resulting predictor has the same form as that of the RE model (29). This proof applies to both the Kapoor et al. (2007) SAR-RE specification and the Fingleton (2008a) SMA-RE specification. Therefore, the BLUP of $y_{i,T+\tau}$ for the SAR-RE and the SMA-RE, like the usual RE model with no spatial effects, modifies the usual GLS forecasts by adding a fraction of the mean of the GLS residuals corresponding to the ith individual. While the predictor formula is the same, the MLEs for these specifications yield different estimates which in turn yield different residuals and hence different forecasts.

4. Monte Carlo design

In this section, we consider the small sample performance of several predictors for an error component model with spatially autocorrelated residuals. The data generating process (DGP) considers two specifications on the remainder errors, namely SAR and SMA:

$$y_{it} = \beta_0 + \beta_1 x_{it} + \varepsilon_{it}, \qquad \varepsilon_{it} = \mu_i + \phi_{it}, \quad i = 1, \ldots, N;\ t = 1, \ldots, T, \tag{32}$$

where

$$x_{it} = \delta_i + \xi_{it},$$

with

$$\mu_i \sim iid.\, N(0, \sigma_\mu^2), \qquad \delta_i \sim iid.\, U(-7.5, 7.5),$$
$$\xi_{it} \sim iid.\, U(-5, 5), \qquad \beta_0 = 5, \qquad \beta_1 = 0.5$$
$$\phi_t = \begin{cases} \rho W_N \phi_t + v_t & \text{for SAR} \\ \lambda W_N v_t + v_t & \text{for SMA,} \end{cases} \quad \text{with } \rho, \lambda = \begin{cases} 0.8 \\ 0.4, \end{cases} \tag{33}$$

and

$$v_{it} \sim iid.\, N(0, \sigma_v^2). \tag{34}$$

The μ_i's and ϕ_{it}'s are assumed to be independent of the regressor x_{it}. In the spirit of Nerlove (1971), we have tried another DGP for x_{it}. We obtain the same ranking as those which appear in the reported tables. The only difference is that the gap between the average heterogeneous estimators and the homogeneous estimators widens with a Nerlove (1971) type design. In other words, the forecast performance of the heterogeneous estimators becomes worse. Alternative experiments can be performed where one can allow, say for correlation between the individual effects and the regressor x_{it}. In these experiments, the FE estimator will be consistent, whereas the RE estimator is not. We consider the simple regressions (32) and (33) with $N = (50, 100)$, $T = (10, 20)$ and two cases for the residuals variances:

$$\begin{cases} \sigma_\mu^2 = 4, & \sigma_v^2 = 16, \\ \sigma_\mu^2 = 16, & \sigma_v^2 = 4. \end{cases} \tag{35}$$

Following Kelejian and Prucha (1999), we use two weight matrices which essentially differ in their degree of sparseness. The weight matrices are labelled as "j ahead and j behind" with the non-zero elements being $(2j)^{-1}$, $j = 1$ and 5. Even with this modest design we have 64 experiments.

For each experiment, we obtain the following 16 estimators:

1. Pooled OLS which ignores the individual heterogeneity and the spatial autocorrelation.
2. The average heterogeneous OLS which estimates the cross-sectional equation using OLS for each time period and averages these heterogeneous estimates to obtain a pooled estimator, see Pesaran and Smith (1995).
3. The fixed-effects (FE) estimator which accounts for fixed individual effects but does not take into account the spatial autocorrelation.
4. The random effects (RE) estimator which assumes that the μ_i's are $iid(0, \sigma_\mu^2)$, and independent of the remainder disturbances ϕ_{it}'s. This estimator accounts for random individual effects but does not take into account the spatial autocorrelation.

5. The RE-spatial MLE assuming a SAR specification (RE-SAR) on the remainder disturbances. In this case, the μ_i's are $iid(0, \sigma_\mu^2)$ and are independent of the ϕ_{it}'s which follow a SAR process, see Anselin et al. (2008).
6. The RE-spatial MLE assuming a SMA specification (RE-SMA) on the remainder disturbances. In this case, the μ_i's are $iid(0, \sigma_\mu^2)$ and are independent of the ϕ_{it}'s which follow a SMA process, see Anselin et al. (2008).
7. The pooled spatial MLE assuming a SAR specification (Pooled SAR) on the remainder disturbances. This estimator ignores the individual heterogeneity but takes into account the spatial autocorrelation of the SAR type.
8. The pooled spatial MLE assuming a SMA specification (Pooled SMA) on the remainder disturbances. This estimator ignores the individual heterogeneity but takes into account the spatial autocorrelation of the SMA type.
9. The average heterogeneous spatial MLE assuming a SAR specification on the remainder disturbances. This estimates cross-sectional MLE with SAR disturbances for each time period and averages the estimates over time.
10. The average heterogeneous spatial GM estimator assuming a SAR specification on the remainder disturbances proposed by Kelejian and Prucha (1999). This estimates the cross-sectional GM estimator with SAR disturbances for each time period and averages the estimates over time.
11. The average heterogeneous spatial MLE assuming a SMA specification on the remainder disturbances. This estimates the cross-sectional MLE with SMA disturbances for each time period and averages the estimates over time.
12. The average heterogeneous spatial GM estimator assuming a SMA specification on the remainder disturbances proposed by Fingleton (2008b). This estimates the cross-sectional GM estimator with SMA disturbances for each time period and averages the estimates over time.
13. The FE-spatial MLE assuming a SAR specification (FE-SAR) on the remainder disturbances.
14. The FE-spatial MLE assuming a SMA specification (FE-SMA) on the remainder disturbances.
15. The (SAR-RE) model following Kapoor et al. (2007). This utilizes a panel data GM estimator where the disturbance term itself follows a SAR process and the remainder term follows an error component structure.
16. The (SMA-RE) model following Fingleton (2008a). This utilizes a panel data GM estimator where the disturbance term itself follows a SMA process and the remainder term follows an error component structure.

Next, we compute the following predictors for the ith individual at a future period $T+\tau$ for $\tau = 1, 2, \ldots, 5$:

OLS	$\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{OLS}$
Average hetero. OLS	$\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{av.OLS}$
FE	$\begin{cases}\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{FE} + \widehat{\mu}_i \\ \text{with } \widehat{\mu}_i = \overline{y}_i - \overline{X}_i\widehat{\beta}_{FE}, \overline{y}_i = \sum_{t=1}^{T} y_{it}/T\end{cases}$
RE	$\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{RE} + \frac{\sigma_\mu^2}{\sigma_1^2}(\iota_T' \otimes l_i')\widehat{\varepsilon}_{RE}$
RE-SAR	$\begin{cases}\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{MLE,RE\text{-}SAR} + \theta(\iota_T' \otimes l_i'C_1^{-1})\widehat{\varepsilon}_{MLE,RE\text{-}SAR} \\ \text{with } C_1 = [T\theta I_N + (B_N'B_N)^{-1}] \text{ and } \theta = \sigma_\mu^2/\sigma_v^2\end{cases}$
RE-SMA	$\begin{cases}\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{MLE,RE\text{-}SMA} + \theta(\iota_T' \otimes l_i'C_2^{-1})\widehat{\varepsilon}_{MLE,RE-SMA} \\ \text{with } C_2 = [T\theta I_N + (D_N D_N')] \text{ and } \theta = \sigma_\mu^2/\sigma_v^2\end{cases}$
Pooled SAR	$\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{MLE,SAR}$
Pooled SMA	$\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{MLE,SMA}$
Average hetero. SAR	$\widehat{y}_{i,T+\tau} = \begin{cases}X_{i,T+\tau}\widehat{\beta}_{av.MLE,SAR} \\ X_{i,T+\tau}\widehat{\beta}_{av.GM,SAR}\end{cases}$
Average hetero. SMA	$\widehat{y}_{i,T+\tau} = \begin{cases}X_{i,T+\tau}\widehat{\beta}_{av.MLE,SMA} \\ X_{i,T+\tau}\widehat{\beta}_{av.GM,SMA}\end{cases}$
FE-SAR	$\begin{cases}\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{MLE,FE\text{-}SAR} + \widehat{\mu}_i \\ \text{with } \widehat{\mu}_i = \overline{y}_i - \overline{X}_i\widehat{\beta}_{MLE,FE\text{-}SAR}, \overline{y}_i = \sum_{t=1}^{T} y_{it}/T\end{cases}$
FE-SMA	$\begin{cases}\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{MLE,FE\text{-}SMA} + \widehat{\mu}_i \\ \text{with } \widehat{\mu}_i = \overline{y}_i - \overline{X}_i\widehat{\beta}_{MLE,FE\text{-}SMA}, \overline{y}_i = \sum_{t=1}^{T} y_{it}/T\end{cases}$
SAR-RE	$\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{MLE,SAR-RE} + \left(\frac{\sigma_\mu^2}{\sigma_1^2}\right)(\iota_T' \otimes l_i')\widehat{\varepsilon}_{MLE,SAR\text{-}RE}$
SMA-RE	$\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{MLE,SMA\text{-}RE} + \left(\frac{\sigma_\mu^2}{\sigma_1^2}\right)(\iota_T' \otimes l_i')\widehat{\varepsilon}_{MLE,SMA\text{-}RE}$

For all experiments, 1000 replications are performed and the RMSE for one step to five step ahead forecasts are reported.

5. Monte Carlo results

5.1. The spatial dependence specification effect

Table 1 gives the RMSE for the one year and five year ahead forecasts along with the average RMSE for all 5 years. These are out of sample forecasts when the true DGP is a RE panel model with SAR remainder disturbances, and when the true DGP is a RE panel model with SMA remainder disturbances. The sample size is $N = 50$ and $T = 10$, the weight matrix is $W(1, 1)$, i.e., one neighbor behind and one neighbor ahead.

If we consider the case where the true DGP is a RE panel model with SAR remainder disturbances, for $\rho = 0.4$, 0.8 and $\sigma_\mu^2 = 4$, 16, the lowest RMSE is that of RE-SAR. This is followed closely by SAR-RE and SMA-RE. The results also confirm the findings of Kapoor et al. (2007) that, on average, RMSE of MLE and their GM estimators are quite similar. It also seems like misspecifying the SAR by an SMA in an error component model does not affect the forecast performance as long as it is taken into account. The misspecification here is only in the variance-covariance matrix, and not in the regression or in the spatial lag say of the dependent variable. The latter could be more serious and should be the subject of future study. As the spatial autoregressive parameter ρ doubles from 0.4 to 0.8, the RMSE also doubles. The RMSE improves as σ_μ^2 gets large, i.e., 16 rather than 4, for estimators that take heterogeneity into account. Pooled OLS, average heterogeneous OLS, pooled SAR, pooled SMA, average heterogeneous SAR (MLE and GM) and average heterogeneous SMA (MLE and GM) perform worse in terms of RMSE than spatial/panel homogeneous estimators. This forecast comparison is robust whether we are predicting one period or five periods ahead and is also reflected in the average over the five years. The gain in forecast performance is substantial once we account for RE or FE and is only slightly improved by additionally accounting for spatial autocorrelation, i.e., FE-SAR or RE-SAR, FE-SMA, or RE-SMA.

If we consider the case where the true DGP is a RE panel model with SMA remainder disturbances, for $\lambda = 0.4$, 0.8 and $\sigma_\mu^2 = 4$, 16, the lowest RMSE is that of RE-SMA. This is followed closely by RE-SAR. Misspecifying the SMA by an SAR in an error component model does not seem to affect the forecast performance as long as it is taken into account. However, the magnitudes of the RMSE (where the true DGP is a RE-SMA process) are much lower than those where the true DGP is a RE-SAR process. Once again, the forecast RMSE based on MLE and their GM counterparts are quite similar, compare SAR-RE and SMA-RE with RE-SAR and RE-SMA. The RMSE improves as σ_μ^2 gets large, i.e., 16 rather than 4, for estimators that take heterogeneity into account. As the spatial autoregressive parameter λ increases from 0.4 to 0.8, the RMSE also increases but not as much as it did for the SAR process. Pooled OLS, average heterogeneous OLS, pooled SAR, pooled SMA, average heterogeneous SAR (MLE and GM) and average heterogeneous SMA (MLE and GM) perform worse in terms of RMSE than spatial/panel homogeneous estimators. This forecast performance is robust whether we are predicting one period or 5 periods ahead and is also reflected in the average over the five years. Once again, the gain in forecast performance is substantial once we account for RE or FE and is only slightly improved by additionally accounting for spatial autocorrelation, i.e., FE-SMA, or RE-SMA, FE-SAR or RE-SAR.

Table 2 reports the RMSE results as Table 1 except that the weight matrix is changed from a $W(1, 1)$ to $W(5, 5)$, i.e., five neighbors behind and five neighbors ahead. Except for the magnitudes of the RMSE, the same rankings in terms of RMSE performance are exhibited as before.

Table 3 reports the RMSE results as Table 1 except that T is now doubled from 10 to 20 holding N fixed at 50. Except for the magnitudes of the RMSE, the same rankings in terms of RMSE performance are exhibited as before. Table 4 reports the RMSE results when $\rho = \lambda = 0.8$, the weight matrices are $W(1, 1)$ and $W(5, 5)$, and N is doubled from 50 to 100 holding T fixed at 10. Except for the magnitudes of the RMSE, the same rankings in terms of RMSE performance are exhibited as before. Other Tables for $W(5, 5)$ and $(N, T) = (100, 20)$ show the same rankings in terms of RMSE forecast performance and are not shown here to save space. These are available upon request from the authors.

5.2. Sensitivity analysis

5.2.1. Sensitivity to irregular lattice structures

The spatial weights matrices considered in the paper are regular lattice structures. Using real irregular lattices structures, as in Anselin and Moreno (2003) and in Kelejian and Prucha (1999), does not change the conclusions of the Monte Carlo study. We used real-world matrices by taking spatial groupings of French administrative communes for dimension $N = 50$. Other Tables for $N = 100$ are available upon request from the authors. Those spatial matrices have been used by Baltagi et al. (2007). Spatial weight matrices may represent high-order contiguity relationships. We use a k-order contiguity matrix containing $N-1$ potential neighborhoods in French municipalities. We have patterns of 0 and 1 values in an $(N-1)$ by $(N-1)$ grid for the k-nearest neighborhoods and we use the 1-nearest neighborhood ($k = 1$) and the 5-nearest neighborhoods ($k = 5$). Note that a non-zero entry in row i, column j denotes that neighborhoods i and j have borders that touch and are therefore considered “neighbors”. For $N = 50$ and for $k = 5$, and for the 2401 possible elements in the 49 by 49 matrix, there are only 250 non-zero elements. So, the sparseness value is 10% (=250/2500). These non-zero entries reflect the contiguity relations between the 5-nearest neighborhoods. Results of Tables 5–6 are very similar to those of Tables 1 and 2. Using irregular lattice structures do not change the main conclusions in terms of the RMSE forecast performance of the various estimators considered. These are similar to the rankings obtained when regular lattice structures are used, only the magnitudes of the RMSE differ.

Table 1
Forecasts RMSE–SAR and SMA data generating processes for ϕ, $(N, T) = (50, 10)$, $W(1, 1)$, 1000 replications.

		Estimators																	
		ρ, λ	σ_μ^2	Pooled OLS	Av. hetero. OLS	FE	RE	Pooled SAR	Av. hetero. SAR		FE-SAR	RE-SAR	Pooled SMA	Av. hetero. SMA		FE-SMA	RE-SMA	SAR-RE	SMA-RE
								MLE	MLE	GM	MLE	MLE	MLE	MLE	GM	MLE	MLE	GM	GM
1st year	SAR	0.4	4	3.978	3.978	3.810	3.765	3.978	3.978	3.961	3.809	3.756	3.978	3.978	3.946	3.809	3.756	3.761	3.777
			16	3.629	3.629	1.902	1.899	3.630	3.630	3.652	1.901	1.897	3.630	3.630	3.657	1.901	1.897	1.898	1.913
		0.8	4	7.056	7.055	7.196	7.022	7.053	7.039	7.056	7.192	6.956	7.054	7.040	7.080	7.192	6.967	7.019	7.038
			16	4.653	4.653	3.591	3.576	4.658	4.657	4.670	3.586	3.552	4.659	4.658	4.664	3.587	3.560	3.605	3.591
	SMA	0.4	4	3.670	3.670	3.472	3.426	3.670	3.670	3.671	3.471	3.419	3.670	3.670	3.667	3.471	3.419	3.433	3.438
			16	3.558	3.558	1.748	1.746	3.558	3.558	3.561	1.748	1.745	3.558	3.559	3.557	1.748	1.745	1.744	1.732
		0.8	4	3.987	3.987	3.851	3.791	3.986	3.986	3.988	3.848	3.765	3.986	3.986	3.977	3.848	3.763	3.792	3.821
			16	3.636	3.636	1.912	1.910	3.638	3.638	3.624	1.910	1.907	3.638	3.638	3.632	1.910	1.906	1.927	1.908
5th year	SAR	0.4	4	4.724	4.724	4.544	4.487	4.724	4.724	4.727	4.543	4.478	4.724	4.724	4.720	4.543	4.478	4.484	4.491
			16	4.028	4.028	2.273	2.270	4.029	4.029	4.036	2.272	2.268	4.029	4.029	4.037	2.272	2.268	2.269	2.272
		0.8	4	8.420	8.420	8.568	8.376	8.416	8.399	8.417	8.561	8.300	8.416	8.400	8.433	8.561	8.310	8.430	8.426
			16	5.428	5.428	4.296	4.281	5.431	5.430	5.416	4.291	4.253	5.432	5.431	5.413	4.291	4.262	4.281	4.271
	SMA	0.4	4	4.364	4.364	4.137	4.088	4.364	4.364	4.365	4.136	4.079	4.364	4.364	4.370	4.136	4.079	4.109	4.097
			16	3.912	3.912	2.075	2.072	3.913	3.913	3.915	2.074	2.071	3.913	3.913	3.915	2.074	2.071	2.073	2.075
		0.8	4	4.770	4.770	4.600	4.539	4.770	4.770	4.746	4.597	4.516	4.771	4.771	4.743	4.596	4.512	4.522	4.537
			16	4.040	4.041	2.296	2.293	4.043	4.042	4.040	2.294	2.290	4.043	4.043	4.041	2.293	2.289	2.289	2.290
Average	SAR	0.4	4	4.474	4.474	4.296	4.243	4.474	4.474	4.460	4.295	4.234	4.474	4.474	4.451	4.295	4.234	4.239	4.244
			16	3.898	3.898	2.149	2.145	3.899	3.899	3.910	2.147	2.144	3.899	3.899	3.913	2.147	2.144	2.145	2.151
		0.8	4	7.947	7.947	8.089	7.904	7.943	7.928	7.946	8.083	7.833	7.944	7.929	7.966	8.083	7.842	7.930	7.949
			16	5.173	5.174	4.064	4.048	5.178	5.177	5.170	4.059	4.021	5.178	5.177	5.163	4.059	4.030	4.049	4.041
	SMA	0.4	4	4.132	4.132	3.913	3.865	4.132	4.132	4.133	3.912	3.857	4.132	4.132	4.135	3.912	3.857	3.882	3.872
			16	3.798	3.798	1.964	1.962	3.798	3.798	3.799	1.964	1.961	3.798	3.798	3.799	1.964	1.961	1.960	1.957
		0.8	4	4.505	4.505	4.347	4.286	4.505	4.505	4.488	4.343	4.262	4.506	4.506	4.481	4.343	4.259	4.275	4.296
			16	3.904	3.904	2.161	2.158	3.906	3.905	3.905	2.159	2.156	3.906	3.906	3.908	2.159	2.155	2.167	2.161

Table 2
Forecasts RMSE–SAR and SMA data generating processes for ϕ, $(N, T) = (50, 10)$, $W(5, 5)$, 1000 replications.

		Estimators ρ, λ	σ_μ^2	Pooled OLS	Av. hetero. OLS	FE	RE	Pooled SAR MLE	Av. hetero. SAR MLE	Av. hetero. SAR GM	FE-SAR MLE	RE-SAR MLE	Pooled SMA MLE	Av. hetero. SMA MLE	Av. hetero. SMA GM	FE-SMA MLE	RE-SMA MLE	SAR-RE GM	SMA-RE GM
1st year	SAR	0.4	4	3.660	3.660	3.454	3.412	3.660	3.660	3.645	3.454	3.410	3.660	3.660	3.634	3.454	3.410	3.442	3.442
			16	3.574	3.574	1.741	1.741	3.574	3.574	3.570	1.741	1.740	3.574	3.574	3.567	1.741	1.740	1.735	1.734
		0.8	4	4.915	4.915	4.805	4.773	4.913	4.893	4.898	4.804	4.736	4.913	4.893	4.908	4.804	4.736	4.711	4.782
			16	3.928	3.928	2.415	2.414	3.928	3.927	3.898	2.413	2.406	3.928	3.927	3.904	2.413	2.408	2.395	2.397
	SMA	0.4	4	3.591	3.591	3.376	3.336	3.591	3.591	3.582	3.376	3.336	3.591	3.591	3.586	3.376	3.335	3.342	3.334
			16	3.511	3.512	1.687	1.685	3.511	3.512	3.525	1.687	1.684	3.511	3.512	3.518	1.687	1.684	1.691	1.685
		0.8	4	3.676	3.676	3.473	3.431	3.676	3.676	3.659	3.472	3.428	3.676	3.676	3.655	3.472	3.428	3.404	3.394
			16	3.530	3.530	1.729	1.727	3.530	3.530	3.555	1.729	1.727	3.530	3.531	3.566	1.729	1.727	1.723	1.725
5th year	SAR	0.4	4	4.362	4.362	4.140	4.090	4.362	4.362	4.361	4.140	4.087	4.362	4.362	4.359	4.140	4.087	4.091	4.102
			16	3.913	3.913	2.071	2.070	3.914	3.914	3.925	2.071	2.069	3.914	3.914	3.920	2.071	2.069	2.067	2.067
		0.8	4	5.838	5.838	5.731	5.681	5.837	5.816	5.862	5.729	5.638	5.837	5.816	5.864	5.729	5.641	5.651	5.707
			16	4.419	4.419	2.867	2.864	4.420	4.418	4.402	2.865	2.855	4.420	4.418	4.399	2.865	2.858	2.870	2.876
	SMA	0.4	4	4.256	4.256	4.020	3.975	4.256	4.256	4.266	4.020	3.973	4.256	4.256	4.266	4.020	3.973	3.976	3.979
			16	3.869	3.870	2.016	2.014	3.870	3.870	3.863	2.016	2.014	3.870	3.870	3.861	2.016	2.014	2.014	2.016
		0.8	4	4.347	4.347	4.123	4.074	4.347	4.347	4.345	4.122	4.068	4.347	4.347	4.340	4.122	4.068	4.072	4.063
			16	3.877	3.877	2.058	2.055	3.877	3.877	3.900	2.057	2.055	3.877	3.877	3.907	2.057	2.055	2.055	2.055
Average	SAR	0.4	4	4.119	4.119	3.903	3.855	4.119	4.119	4.118	3.903	3.852	4.119	4.119	4.115	3.903	3.852	3.874	3.878
			16	3.801	3.801	1.959	1.957	3.801	3.801	3.807	1.959	1.957	3.801	3.802	3.804	1.959	1.957	1.954	1.954
		0.8	4	5.523	5.523	5.412	5.369	5.522	5.501	5.539	5.410	5.329	5.522	5.501	5.541	5.410	5.331	5.332	5.390
			16	4.254	4.254	2.713	2.710	4.255	4.254	4.238	2.711	2.702	4.255	4.254	4.238	2.711	2.704	2.706	2.712
	SMA	0.4	4	4.030	4.030	3.801	3.757	4.030	4.030	4.037	3.801	3.756	4.030	4.030	4.036	3.801	3.756	3.764	3.758
			16	3.752	3.752	1.906	1.904	3.753	3.753	3.752	1.905	1.903	3.753	3.753	3.749	1.905	1.903	1.905	1.902
		0.8	4	4.123	4.123	3.903	3.857	4.122	4.122	4.112	3.903	3.852	4.122	4.122	4.106	3.902	3.852	3.847	3.837
			16	3.761	3.761	1.945	1.942	3.761	3.761	3.789	1.944	1.942	3.761	3.762	3.799	1.944	1.942	1.942	1.942

Table 3
Forecasts RMSE–SAR and SMA data generating processes for ϕ, $W(1, 1)$, $(N, T) = (50, 20)$, 1000 replications.

		Estimators																	
		ρ, λ	σ_μ^2	Pooled OLS	Av. hetero. OLS	FE	RE	Pooled SAR	Av. hetero. SAR		FE-SAR	RE-SAR	Pooled SMA	Av. hetero. SMA		FE-SMA	RE-SMA	SAR-RE	SMA-RE
								MLE	MLE	GM	MLE	MLE	MLE	MLE	GM	MLE	MLE	GM	GM
1st year	SAR	0.4	4	3.951	3.951	3.736	3.716	3.952	3.952	3.963	3.736	3.710	3.952	3.952	3.967	3.736	3.710	3.729	3.712
			16	3.616	3.616	1.863	1.862	3.618	3.618	3.643	1.863	1.862	3.618	3.618	3.644	1.863	1.862	1.865	1.865
		0.8	4	7.064	7.064	7.040	6.980	7.063	7.054	7.129	7.036	6.944	7.064	7.055	7.139	7.036	6.944	6.987	6.975
			16	4.660	4.661	3.508	3.505	4.661	4.661	4.728	3.507	3.498	4.660	4.660	4.719	3.507	3.500	3.491	3.498
	SMA	0.4	4	3.670	3.670	3.400	3.386	3.670	3.670	3.687	3.399	3.383	3.671	3.671	3.680	3.399	3.383	3.389	3.352
			16	3.557	3.557	1.692	1.691	3.557	3.558	3.550	1.692	1.691	3.557	3.558	3.538	1.692	1.691	1.704	1.687
		0.8	4	3.986	3.986	3.752	3.737	3.985	3.985	3.982	3.750	3.729	3.986	3.986	3.988	3.750	3.728	3.742	3.727
			16	3.638	3.638	1.880	1.879	3.639	3.639	3.618	1.879	1.878	3.639	3.639	3.623	1.879	1.878	1.899	1.869
5th year	SAR	0.4	4	4.730	4.730	4.464	4.444	4.730	4.730	4.714	4.464	4.439	4.730	4.730	4.716	4.464	4.439	4.443	4.438
			16	4.017	4.017	2.223	2.222	4.018	4.018	4.028	2.222	2.222	4.019	4.019	4.032	2.222	2.222	2.223	2.219
		0.8	4	8.446	8.446	8.404	8.340	8.445	8.437	8.443	8.401	8.304	8.447	8.439	8.445	8.401	8.304	8.364	8.329
			16	5.426	5.426	4.208	4.203	5.428	5.427	5.443	4.206	4.192	5.428	5.428	5.443	4.206	4.196	4.170	4.191
	SMA	0.4	4	4.366	4.366	4.053	4.038	4.367	4.367	4.360	4.052	4.036	4.367	4.367	4.354	4.052	4.035	4.041	4.036
			16	3.911	3.911	2.025	2.024	3.911	3.911	3.907	2.025	2.024	3.911	3.911	3.899	2.025	2.024	2.027	2.025
		0.8	4	4.736	4.736	4.471	4.452	4.737	4.737	4.737	4.469	4.443	4.738	4.738	4.740	4.469	4.441	4.462	4.459
			16	4.042	4.042	2.243	2.242	4.043	4.043	4.031	2.242	2.241	4.044	4.044	4.032	2.242	2.241	2.250	2.232
Average	SAR	0.4	4	4.470	4.470	4.219	4.199	4.470	4.470	4.458	4.218	4.194	4.470	4.471	4.460	4.218	4.194	4.201	4.195
			16	3.884	3.884	2.100	2.099	3.885	3.885	3.900	2.100	2.099	3.885	3.886	3.904	2.100	2.099	2.102	2.097
		0.8	4	7.979	7.979	7.942	7.880	7.978	7.970	7.999	7.938	7.844	7.979	7.972	8.008	7.938	7.844	7.898	7.870
			16	5.173	5.174	3.975	3.971	5.174	5.174	5.199	3.973	3.961	5.174	5.174	5.194	3.974	3.965	3.942	3.959
	SMA	0.4	4	4.128	4.128	3.832	3.817	4.128	4.128	4.132	3.831	3.814	4.128	4.128	4.126	3.831	3.813	3.820	3.800
			16	3.794	3.794	1.913	1.912	3.794	3.794	3.792	1.912	1.911	3.794	3.794	3.782	1.912	1.911	1.918	1.909
		0.8	4	4.483	4.483	4.228	4.210	4.484	4.484	4.484	4.226	4.201	4.485	4.485	4.487	4.226	4.200	4.215	4.214
			16	3.906	3.906	2.119	2.118	3.907	3.907	3.894	2.118	2.117	3.908	3.908	3.897	2.118	2.116	2.128	2.108

Table 4
Forecasts RMSE - $\rho = \lambda = 0.8$ for SAR and SMA data generating processes, $(N, T) = (100, 10)$, 1000 replications.

		Estimators																	
		True DGP	σ_μ^2	Pooled OLS	Av. hetero. OLS	FE	RE	Pooled SAR	Av. hetero. SAR		FE-SAR	RE-SAR	Pooled SMA	Av. hetero. SMA		FE-SMA	RE-SMA	SAR-RE	SMA-RE
								MLE	MLE	GM	MLE	MLE	MLE	MLE	GM	MLE	MLE	GM	GM
1st year	$W(1, 1)$	SAR	4	6.999	6.998	7.123	6.943	6.998	6.994	7.111	7.119	6.890	6.999	6.995	7.093	7.119	6.894	7.074	7.034
			16	4.650	4.650	3.589	3.570	4.651	4.650	4.677	3.587	3.542	4.651	4.650	4.445	3.587	3.553	3.590	3.577
		SMA	4	4.010	4.010	3.859	3.800	4.010	4.010	4.006	3.858	3.780	4.011	4.011	3.985	3.858	3.778	3.792	3.781
			16	3.658	3.658	1.925	1.922	3.658	3.658	3.669	1.924	1.919	3.658	3.658	3.657	1.924	1.918	1.924	1.919
	$W(5, 5)$	SAR	4	4.890	4.890	4.810	4.742	4.889	4.884	4.876	4.809	4.711	4.889	4.884	4.863	4.809	4.710	4.718	4.738
			16	3.914	3.914	2.391	2.388	3.914	3.914	3.934	2.391	2.373	3.914	3.914	3.907	2.391	2.380	2.394	2.377
		SMA	4	3.648	3.648	3.453	3.406	3.648	3.648	3.657	3.452	3.400	3.648	3.648	3.653	3.452	3.400	3.399	3.403
			16	3.564	3.564	1.722	1.720	3.564	3.564	3.564	1.722	1.719	3.564	3.564	3.568	1.722	1.719	1.724	1.718
5th year	$W(1, 1)$	SAR	4	8.403	8.403	8.553	8.340	8.401	8.397	8.434	8.549	8.278	8.402	8.398	8.420	8.549	8.285	8.416	8.374
			16	5.423	5.423	4.281	4.264	5.424	5.424	5.420	4.279	4.236	5.424	5.424	5.433	4.279	4.246	4.283	4.274
		SMA	4	4.758	4.758	4.582	4.518	4.758	4.758	4.751	4.581	4.496	4.759	4.759	4.748	4.580	4.493	4.533	4.520
			16	4.062	4.063	2.295	2.292	4.063	4.063	4.061	2.294	2.289	4.063	4.063	4.052	2.294	2.289	2.288	2.283
	$W(5, 5)$	SAR	4	5.835	5.835	5.750	5.672	5.834	5.829	5.816	5.749	5.631	5.834	5.829	5.816	5.749	5.633	5.639	5.677
			16	4.424	4.424	2.876	2.871	4.424	4.424	4.437	2.875	2.854	4.424	4.424	4.416	2.875	2.863	2.854	2.842
		SMA	4	4.329	4.329	4.107	4.054	4.329	4.329	4.344	4.107	4.048	4.329	4.329	4.351	4.107	4.047	4.070	4.066
			16	3.921	3.921	2.057	2.055	3.921	3.921	3.913	2.057	2.055	3.921	3.921	3.914	2.057	2.055	2.056	2.055
Average	$W(1, 1)$	SAR	4	7.922	7.922	8.064	7.862	7.920	7.917	7.987	8.060	7.803	7.921	7.918	7.971	8.061	7.808	7.965	7.920
			16	5.167	5.167	4.053	4.035	5.168	5.168	5.172	4.051	4.006	5.168	5.168	5.182	4.051	4.017	4.049	4.037
		SMA	4	4.507	4.507	4.339	4.276	4.507	4.507	4.501	4.338	4.255	4.508	4.508	4.490	4.337	4.253	4.282	4.272
			16	3.928	3.928	2.169	2.166	3.928	3.928	3.932	2.168	2.163	3.928	3.928	3.924	2.167	2.162	2.163	2.159
	$W(5, 5)$	SAR	4	5.514	5.514	5.430	5.355	5.513	5.509	5.503	5.428	5.319	5.513	5.509	5.495	5.428	5.319	5.325	5.356
			16	4.254	4.254	2.714	2.709	4.254	4.254	4.274	2.713	2.694	4.254	4.254	4.250	2.713	2.702	2.699	2.685
		SMA	4	4.096	4.096	3.884	3.832	4.096	4.096	4.109	3.883	3.826	4.096	4.096	4.115	3.883	3.826	3.844	3.841
			16	3.805	3.805	1.945	1.943	3.805	3.805	3.798	1.944	1.942	3.805	3.805	3.802	1.944	1.942	1.944	1.941

Table 5
Forecasts RMSE–SAR and SMA data generating processes for ϕ, $(N, T) = (50, 10)$, $W(1, 1)$ asymmetric weight matrix of French administrative communes, 1000 replications.

		Estimators ρ, λ	σ^2_μ	Pooled OLS	Av. hetero. OLS	FE	RE	Pooled SAR	Av. hetero. SAR		FE-SAR	RE-SAR	Pooled SMA	Av. hetero. SMA		FE-SMA	RE-SMA	SAR-RE	SMA-RE
								MLE	MLE	GM	MLE	MLE	MLE	MLE	GM	MLE	MLE	GM	GM
1st year	SAR	0.4	4	4.155	4.155	4.036	3.971	4.156	4.156	4.208	4.034	3.955	4.156	4.156	4.177	4.034	3.959	3.972	4.041
			16	3.674	3.674	2.031	2.027	3.675	3.675	3.710	2.030	2.024	3.674	3.674	3.687	2.030	2.025	2.019	1.777
		0.8	4	9.981	9.986	9.381	9.934	10.093	10.100	9.517	9.351	9.992	10.080	10.049	11.007	9.356	9.577	14.106	11.922
			16	6.060	6.060	5.440	5.395	6.062	6.059	6.035	5.432	5.318	6.062	6.061	5.994	5.432	5.350	5.344	5.914
	SMA	0.4	4	3.781	3.781	3.581	3.538	3.782	3.782	3.804	3.579	3.526	3.783	3.783	3.775	3.579	3.528	3.566	3.551
			16	3.589	3.589	1.802	1.800	3.590	3.589	3.603	1.801	1.799	3.590	3.590	3.579	1.801	1.799	1.809	1.809
		0.8	4	4.367	4.368	4.279	4.198	4.366	4.367	4.399	4.274	4.171	4.366	4.366	4.377	4.273	4.166	4.232	4.221
			16	3.747	3.747	2.148	2.144	3.749	3.749	3.749	2.147	2.141	3.749	3.748	3.732	2.146	2.140	2.136	2.133
5th year	SAR	0.4	4	4.961	4.962	4.817	4.747	4.962	4.962	4.972	4.815	4.731	4.962	4.962	4.967	4.815	4.735	4.750	4.816
			16	4.086	4.087	2.413	2.408	4.088	4.087	4.123	2.411	2.406	4.087	4.087	4.115	2.411	2.407	2.402	2.522
		0.8	4	13.489	13.495	13.751	13.477	13.476	13.477	13.267	13.722	13.583	13.473	13.468	14.185	13.724	13.416	13.952	12.277
			16	7.154	7.154	6.442	6.395	7.162	7.158	7.158	6.434	6.315	7.157	7.157	7.151	6.434	6.346	6.419	6.372
	SMA	0.4	4	4.498	4.498	4.296	4.241	4.499	4.499	4.505	4.294	4.230	4.499	4.499	4.501	4.294	4.231	4.246	4.244
			16	3.947	3.947	2.145	2.143	3.948	3.948	3.965	2.144	2.141	3.948	3.948	3.956	2.144	2.141	2.146	2.149
		0.8	4	5.217	5.217	5.110	5.027	5.218	5.218	5.241	5.106	4.998	5.218	5.218	5.213	5.105	4.992	5.039	5.042
			16	4.191	4.191	2.559	2.555	4.193	4.193	4.201	2.557	2.550	4.193	4.193	4.179	2.557	2.549	2.558	2.553
Average	SAR	0.4	4	4.688	4.688	4.552	4.484	4.689	4.689	4.712	4.550	4.468	4.689	4.689	4.699	4.550	4.471	4.482	4.575
			16	3.949	3.949	2.282	2.278	3.950	3.949	3.989	2.281	2.275	3.949	3.949	3.975	2.281	2.277	2.272	2.212
		0.8	4	12.050	12.055	12.190	12.028	12.071	12.071	11.345	12.145	12.198	12.065	12.056	13.154	12.149	11.932	14.199	12.078
			16	6.778	6.778	6.093	6.047	6.785	6.781	6.779	6.085	5.969	6.781	6.780	6.766	6.085	5.999	6.049	6.423
	SMA	0.4	4	4.257	4.257	4.057	4.005	4.258	4.258	4.268	4.055	3.994	4.259	4.259	4.252	4.055	3.995	4.013	4.010
			16	3.831	3.831	2.027	2.025	3.831	3.832	3.847	2.026	2.024	3.832	3.832	3.834	2.026	2.024	2.031	2.034
		0.8	4	4.936	4.936	4.835	4.753	4.936	4.936	4.954	4.831	4.725	4.936	4.936	4.927	4.829	4.719	4.767	4.760
			16	4.047	4.047	2.421	2.417	4.049	4.049	4.054	2.419	2.413	4.049	4.049	4.031	2.418	2.411	2.414	2.409

Table 6
Forecasts RMSE–SAR and SMA data generating processes for ϕ, $(N, T) = (50, 10)$, $W(5, 5)$ asymmetric weight matrix of French administrative communes, 1000 replications.

		Estimators																	
		ρ, λ	σ_μ^2	Pooled OLS	Av. hetero. OLS	FE	RE	Pooled SAR	Av. hetero. SAR		FE-SAR	RE-SAR	Pooled SMA	Av. hetero. SMA		FE-SMA	RE-SMA	SAR-RE	SMA-RE
								MLE	MLE	GM	MLE	MLE	MLE	MLE	GM	MLE	MLE	GM	GM
1st year	SAR	0.4	4	3.738	3.738	3.550	3.508	3.738	3.738	3.734	3.549	3.502	3.737	3.738	3.735	3.549	3.501	3.500	3.800
			16	3.558	3.558	1.773	1.770	3.560	3.559	3.573	1.773	1.769	3.559	3.559	3.562	1.773	1.769	1.778	1.907
		0.8	4	5.721	5.721	5.699	5.632	5.712	5.691	5.773	5.694	5.669	5.717	5.719	5.671	5.694	5.584	5.726	6.582
			16	4.182	4.182	2.836	2.832	4.186	4.183	4.200	2.834	2.807	4.183	4.182	4.201	2.834	2.820	2.832	3.351
	SMA	0.4	4	3.599	3.599	3.417	3.372	3.600	3.600	3.606	3.416	3.375	3.600	3.600	3.601	3.416	3.369	3.350	3.427
			16	3.540	3.540	1.695	1.694	3.541	3.541	3.518	1.694	1.693	3.541	3.542	3.547	1.694	1.694	1.705	1.667
		0.8	4	3.740	3.741	3.562	3.517	3.742	3.740	3.739	3.562	3.508	3.740	3.740	3.739	3.561	3.505	3.498	3.702
			16	3.576	3.576	1.774	1.772	3.578	3.577	3.576	1.773	1.772	3.576	3.577	3.552	1.773	1.772	1.767	1.986
5th year	SAR	0.4	4	4.433	4.434	4.228	4.175	4.434	4.434	4.433	4.227	4.170	4.433	4.434	4.443	4.227	4.169	4.166	4.234
			16	3.920	3.920	2.112	2.109	3.922	3.921	3.940	2.112	2.108	3.921	3.921	3.934	2.112	2.109	2.110	2.269
		0.8	4	6.808	6.809	6.785	6.706	6.796	6.775	6.868	6.780	6.755	6.802	6.797	6.807	6.781	6.651	6.712	6.711
			16	4.789	4.789	3.394	3.389	4.794	4.789	4.782	3.392	3.363	4.791	4.789	4.788	3.392	3.376	3.395	3.310
	SMA	0.4	4	4.274	4.274	4.048	3.999	4.274	4.274	4.281	4.048	4.004	4.274	4.274	4.279	4.048	3.997	3.997	4.031
			16	3.870	3.870	2.025	2.022	3.871	3.871	3.862	2.024	2.022	3.870	3.871	3.888	2.024	2.023	2.026	1.992
		0.8	4	4.441	4.441	4.234	4.183	4.443	4.442	4.449	4.234	4.175	4.441	4.441	4.458	4.234	4.172	4.190	4.330
			16	3.934	3.934	2.119	2.117	3.937	3.935	3.937	2.119	2.116	3.935	3.935	3.918	2.119	2.116	2.112	2.256
Average	SAR	0.4	4	4.194	4.194	3.994	3.944	4.194	4.194	4.191	3.993	3.940	4.194	4.194	4.204	3.993	3.938	3.939	4.042
			16	3.805	3.805	2.000	1.998	3.807	3.807	3.821	2.000	1.997	3.806	3.807	3.811	2.000	1.998	1.998	2.116
		0.8	4	6.452	6.452	6.429	6.353	6.440	6.419	6.499	6.425	6.403	6.446	6.444	6.418	6.425	6.301	6.386	6.424
			16	4.589	4.589	3.207	3.203	4.593	4.589	4.591	3.205	3.177	4.590	4.589	4.590	3.205	3.190	3.204	3.388
	SMA	0.4	4	4.046	4.046	3.834	3.786	4.046	4.046	4.050	3.834	3.790	4.046	4.046	4.045	3.834	3.784	3.776	3.849
			16	3.760	3.760	1.913	1.911	3.761	3.761	3.750	1.913	1.911	3.761	3.761	3.775	1.913	1.912	1.918	1.903
		0.8	4	4.205	4.205	4.005	3.957	4.208	4.206	4.211	4.005	3.949	4.205	4.205	4.216	4.005	3.945	3.958	4.115
			16	3.816	3.817	1.999	1.997	3.819	3.817	3.819	1.998	1.997	3.817	3.817	3.796	1.998	1.997	1.994	2.147

5.2.2. Robustness to non-normality

So far, we have been assuming that the error components have been generated by the normal distribution. In this section, we check the sensitivity of our results to non-normal disturbances. In particular, we generate the μ_i's from a χ^2 distribution and we let the remainder disturbances follow the normal distribution. One can also study the sensitivity of these results to non-normality of the remainder disturbances. Table 7 gives similar results as those of Table 1 (when the individual effects follow a normal distribution). So, the results seem to be robust to non-normality of the disturbances of the χ^2 type.

6. Summary and conclusion

Our Monte Carlo study finds that when the true DGP is RE with a SAR or SMA remainder disturbances, estimators that ignore heterogeneity/spatial correlation perform badly in RMSE forecasts. For our experiments, accounting for heterogeneity improves the forecast performance by a big margin and accounting for spatial correlation improves the forecast but by a smaller margin. Ignoring both leads to the worst forecasting performance. Heterogeneous estimators based on averaging perform worse than homogeneous estimators in forecasting performance. This performance improves with a larger sample size and seems robust to the type of spatial error structure imposed on the remainder disturbances. These Monte Carlo experiments confirm earlier empirical studies that report similar findings. Having said that, our results should be tempered by the fact that our model is not dynamic nor does it deal with endogenous regressors. Two important characteristics that are surely relevant in empirical studies. This should be the subject of future research.

Acknowledgements

This paper was presented at a conference in honor of Phoebus Dhrymes in Paphos, Cyprus, June 1–3, 2007. Also, at the 14th International Conference on Panel Data at the Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, China, July 16–18, 2007, and the 63rd European Meeting of the Econometric Society (ESEM) held at the University of Bocconi in Milan, Italy, August 27–31, 2008. We would like to thank the editor, three anonymous referees and the participants of those conferences for their useful comments and suggestions.

Appendix

This Appendix first derives the BLUP for the KKP model which we call the (SAR-RE) model described in (13) and (14). The variance-covariance matrix Ω is given in (17). The inverse of Ω is given by:

$$\Omega^{-1} = \frac{1}{\sigma_v^2}\left[\left(I_T - \frac{T\sigma_\mu^2}{\sigma_1^2}\bar{J}_T\right) \otimes (B_N' B_N)\right], \tag{36}$$

where $\bar{J}_T = J_T/T$ and $\sigma_1^2 = T\sigma_\mu^2 + \sigma_v^2$ and $B_N = (I_N - \rho W_N)$. From (13) and (14), we have:

$$\varepsilon_{T+\tau} = B_N^{-1} u_{T+\tau} = B_N^{-1}(\mu + v_{T+\tau}) \tag{37}$$

so that,

$$\begin{aligned} E[\varepsilon_{T+\tau}\varepsilon'] &= E[B_N^{-1}(\mu + v_{T+\tau})((\iota_T \otimes B_N^{-1})\mu + (I_T \otimes B_N^{-1})v)'] \\ &= B_N^{-1}E[\mu\mu'](\iota_T \otimes B_N^{-1})' + B_N^{-1}E[\mu v'](I_T \otimes B_N^{-1})' \\ &\quad + B_N^{-1}E[v_{T+\tau}\mu'](\iota_T \otimes B_N^{-1})' + B_N^{-1}E[v_{T+\tau}v'](I_T \otimes B_N^{-1})'. \end{aligned} \tag{38}$$

Using $E[\mu\mu'] = \sigma_\mu^2 I_N$, $E[\mu v'] = 0$, $E[v_{T+\tau}\mu'] = 0$ and $E[v_{T+\tau}v'] = 0$, the formula (38) can be simplified as

$$E[\varepsilon_{T+\tau}\varepsilon'] = \sigma_\mu^2 B_N^{-1}(\iota_T' \otimes B_N^{-1'}), \tag{39}$$

and the expression of this covariance for an individual i at time $T+\tau$ is given by

$$\omega' = E[\varepsilon_{i,T+\tau}\varepsilon'] = \sigma_\mu^2 b_i(\iota_T' \otimes B_N^{-1'}), \tag{40}$$

where b_i is the ith row of the matrix B_N^{-1}. In this case,

$$\begin{aligned} \omega'\Omega^{-1} &= \frac{\sigma_\mu^2}{\sigma_v^2} b_i(\iota_T' \otimes B_N^{-1'})\left[\left(I_T - \frac{T\sigma_\mu^2}{\sigma_1^2}\bar{J}_T\right) \otimes (B_N' B_N)\right] \\ &= \frac{\sigma_\mu^2}{\sigma_v^2} b_i\left[(\iota_T' \otimes B_N) - \frac{T\sigma_\mu^2}{\sigma_1^2}(\iota_T' \otimes B_N)\right] \\ &= \frac{\sigma_\mu^2}{\sigma_1^2} b_i(\iota_T' \otimes B_N). \end{aligned} \tag{41}$$

Table 7
Forecasts RMSE–SAR and SMA data generating processes for ϕ, $(N, T) = (50, 10)$, $W(1, 1)$, 1000 replications under non-normality of individual effects.

		Estimators																
		ρ, λ	σ_μ^2	Pooled OLS	Av. hetero. OLS	FE	RE	Pooled SAR	Av. hetero. SAR		FE-SAR	RE-SAR	Pooled SMA	Av. hetero. SMA		FE-SMA	RE-SMA	SAR-RE	SMA-RE
								MLE	MLE	GM	MLE	MLE	MLE	MLE	GM	MLE	MLE	GM	GM
1st year	SAR	0.4	4	3.984	3.984	3.838	3.787	3.985	3.985	3.938	3.837	3.781	3.986	3.985	3.965	3.837	3.779	3.766	3.775
			16	3.605	3.605	1.908	1.906	3.606	3.606	3.598	1.907	1.905	3.606	3.606	3.552	1.907	1.905	1.905	1.889
		0.8	4	7.113	7.114	7.255	7.079	7.108	7.100	7.092	7.251	7.058	7.109	7.105	7.071	7.251	7.028	7.027	7.065
			16	4.628	4.628	3.594	3.578	4.629	4.628	4.614	3.591	3.556	4.629	4.628	4.627	3.591	3.564	3.607	3.570
	SMA	0.4	4	3.673	3.673	3.500	3.455	3.675	3.675	3.627	3.499	3.447	3.675	3.675	3.657	3.499	3.448	3.436	3.454
			16	3.493	3.493	1.756	1.753	3.494	3.494	3.480	1.755	1.751	3.494	3.494	3.483	1.755	1.751	1.734	1.735
		0.8	4	3.997	3.997	3.862	3.808	3.997	3.997	3.982	3.858	3.788	4.000	3.999	3.989	3.858	3.786	3.798	3.802
			16	3.550	3.549	1.927	1.924	3.552	3.551	3.566	1.925	1.920	3.552	3.552	3.582	1.924	1.919	1.922	1.920
5th year	SAR	0.4	4	4.693	4.693	4.546	4.482	4.693	4.693	4.687	4.544	4.475	4.694	4.693	4.699	4.544	4.475	4.473	4.495
			16	3.996	3.996	2.272	2.270	3.996	3.996	3.997	2.272	2.269	3.996	3.997	3.953	2.272	2.269	2.267	2.269
		0.8	4	8.480	8.481	8.635	8.436	8.477	8.462	8.440	8.627	8.426	8.478	8.474	8.422	8.627	8.379	8.407	8.381
			16	5.388	5.388	4.297	4.280	5.390	5.388	5.372	4.293	4.252	5.390	5.390	5.385	4.293	4.262	4.295	4.289
	SMA	0.4	4	4.334	4.334	4.150	4.096	4.334	4.334	4.328	4.149	4.090	4.334	4.334	4.338	4.149	4.091	4.084	4.099
			16	3.849	3.849	2.072	2.069	3.849	3.849	3.837	2.071	2.068	3.849	3.849	3.842	2.071	2.068	2.074	2.070
		0.8	4	4.741	4.741	4.600	4.534	4.742	4.741	4.727	4.596	4.512	4.744	4.743	4.722	4.595	4.507	4.525	4.512
			16	3.954	3.954	2.295	2.292	3.955	3.955	3.957	2.293	2.289	3.955	3.955	3.990	2.292	2.287	2.289	2.293
Average	SAR	0.4	4	4.451	4.451	4.304	4.245	4.452	4.452	4.432	4.303	4.237	4.453	4.452	4.450	4.303	4.237	4.234	4.248
			16	3.867	3.867	2.149	2.147	3.868	3.868	3.866	2.148	2.146	3.868	3.868	3.822	2.149	2.146	2.147	2.142
		0.8	4	8.008	8.008	8.161	7.968	8.004	7.990	7.975	8.154	7.952	8.005	8.001	7.972	8.154	7.913	7.932	7.923
			16	5.140	5.140	4.067	4.051	5.142	5.140	5.121	4.064	4.024	5.142	5.142	5.134	4.064	4.033	4.062	4.043
	SMA	0.4	4	4.110	4.110	3.927	3.877	4.110	4.110	4.089	3.926	3.870	4.111	4.111	4.107	3.926	3.871	3.866	3.880
			16	3.733	3.733	1.964	1.961	3.733	3.734	3.723	1.963	1.960	3.733	3.734	3.725	1.963	1.960	1.957	1.958
		0.8	4	4.485	4.485	4.348	4.285	4.485	4.484	4.468	4.344	4.263	4.487	4.487	4.474	4.343	4.259	4.279	4.279
			16	3.820	3.821	2.172	2.169	3.822	3.822	3.830	2.170	2.165	3.822	3.822	3.854	2.169	2.164	2.165	2.165

But $b_i(\iota_T' \otimes B_N) = (1 \otimes b_i)(\iota_T' \otimes B_N) = (\iota_T' \otimes l_i')$, where l_i' is the ith row of I_N. This holds because $B_N^{-1}B_N = I_N$ and therefore $b_iB_N = l_i'$. This means that the predictor of the KKP model from (28) is given by:

$$\widehat{y}_{i,T+\tau} = X_{i,T+\tau}\widehat{\beta}_{GLS} + \frac{\sigma_\mu^2}{\sigma_1^2}(\iota_T' \otimes l_i')\widehat{\varepsilon}_{GLS}, \tag{42}$$

which is the same as that of the RE model with no spatial correlation. While the predictor formula is the same, the MLEs for these specifications yield different estimates which in turn yield different residuals and hence different forecasts.

The proof is the similar for the Fingleton (2008a) specification which we are calling the (SMA-RE) model described in (25) and (14). The variance-covariance matrix Ω is given in (27). The inverse of Ω is given by:

$$\Omega^{-1} = \frac{1}{\sigma_v^2}\left[\left(I_T - \frac{T\sigma_\mu^2}{\sigma_1^2}\bar{J}_T\right) \otimes (D_N D_N')^{-1}\right], \tag{43}$$

where $D_N = (I_N + \lambda W_N)$. From (25) and (14), we have:

$$\varepsilon_{T+\tau} = D_N u_{T+\tau} = D_N(\mu + v_{T+\tau}), \tag{44}$$

so that,

$$\begin{aligned} E[\varepsilon_{T+\tau}\varepsilon'] &= E[D_N(\mu + v_{T+\tau})((\iota_T \otimes D_N)\mu + (I_T \otimes D_N)v)'] \\ &= D_N E[\mu\mu'](\iota_T \otimes D_N)' + D_N E[\mu v'](I_T \otimes D_N)' + D_N E[v_{T+\tau}\mu'](\iota_T \otimes D_N)' + D_N E[v_{T+\tau}v'](I_T \otimes D_N)'. \end{aligned} \tag{45}$$

Using $E[\mu\mu'] = \sigma_\mu^2 I_N$, $E[\mu v'] = 0$, $E[v_{T+\tau}\mu'] = 0$ and $E[v_{T+\tau}v'] = 0$, the formula (45) can be simplified as

$$E[\varepsilon_{T+\tau}\varepsilon'] = \sigma_\mu^2 D_N(\iota_T' \otimes D_N'), \tag{46}$$

and the expression of this covariance for an individual i at time $T + \tau$ is given by

$$\omega' = E[\varepsilon_{i,T+\tau}\varepsilon'] = \sigma_\mu^2 d_i(\iota_T' \otimes D_N'), \tag{47}$$

where d_i is the ith row of the matrix D_N. In this case,

$$\begin{aligned} \omega'\Omega^{-1} &= \frac{\sigma_\mu^2}{\sigma_v^2} d_i(\iota_T' \otimes D_N')\left[\left(I_T - \frac{T\sigma_\mu^2}{\sigma_1^2}\bar{J}_T\right) \otimes (D_N D_N')^{-1}\right] \\ &= \frac{\sigma_\mu^2}{\sigma_v^2} d_i\left[(\iota_T' \otimes D_N^{-1}) - \frac{T\sigma_\mu^2}{\sigma_1^2}(\iota_T' \otimes D_N^{-1})\right] \\ &= \frac{\sigma_\mu^2}{\sigma_1^2} d_i(\iota_T' \otimes D_N^{-1}). \end{aligned} \tag{48}$$

But $d_i(\iota_T' \otimes D_N^{-1}) = (1 \otimes d_i)(\iota_T' \otimes D_N^{-1}) = (\iota_T' \otimes l_i')$, where l_i' is the ith row of I_N. This holds because $D_N D_N^{-1} = I_N$ and therefore $d_i D_N^{-1} = l_i'$. This means that the predictor of the Fingleton (2008a) model is again the same as that of the RE model with no spatial correlation. While the predictor formula is the same, the MLEs for these specifications yield different estimates which in turn yield different residuals and hence different forecasts.

References

Anselin, L., 1988. Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Dordrecht.
Anselin, L., Bera, A.K., 1998. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah, A., Giles, D.E.A. (Eds.), Handbook of Applied Economic Statistics. Marcel Dekker, New York.
Anselin, L., Le Gallo, J., Jayet, H., 2008. Spatial panel econometrics. In: Mátyás, L., Sevestre, P. (Eds.), The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice. Springer-Verlag, Berlin, pp. 625–660 (Chapter 19).
Anselin, L., Moreno, R., 2003. Properties of tests for spatial error components. Regional Science and Urban Economics 33, 595–618.
Baillie, R.T., Baltagi, B.H., 1999. Prediction from the regression model with one-way error components. In: Hsiao, C., Lahiri, K., Lee, L.F., Pesaran, H. (Eds.), Analysis of Panels and Limited Dependent Variable Models. Cambridge University Press, Cambridge, pp. 255–267 (Chapter 10).
Baltagi, B.H., 2008. Forecasting with panel data. Journal of Forecasting 27, 153–173.
Baltagi, B.H., Bresson, G., Pirotte, A., 2002. Comparison of forecast performance for homogeneous, heterogeneous and shrinkage estimators: some empirical evidence from US electricity and natural-gas consumption. Economics Letters 76, 375–382.
Baltagi, B.H., Bresson, G., Pirotte, A., 2004. Tobin q: forecast performance for hierarchical Bayes, shrinkage, heterogeneous and homogeneous panel data estimators. Empirical Economics 29, 107–113.
Baltagi, B.H., Bresson, G., Pirotte, A., 2007. Panel unit root tests and spatial dependence. Journal of Applied Econometrics 22, 339–360.
Baltagi, B.H., Griffin, J.M., 1997. Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline. Journal of Econometrics 77, 303–327.
Baltagi, B.H., Griffin, J.M., Xiong, W., 2000. To pool or not to pool: homogeneous versus heterogeneous estimators applied to cigarette demand. Review of Economics and Statistics 82, 117–126.

Baltagi, B.H., Li, D., 2004. Prediction in the panel data model with spatial correlation. In: Anselin, L., Florax, R.J.G.M., Rey, S.J. (Eds.), Advances in Spatial Econometrics: Methodology, Tools and Applications. Springer, Berlin, pp. 283–295 (Chapter 13).
Baltagi, B.H., Li, D., 2006. Prediction in the panel data model with spatial correlation: the case of liquor. Spatial Economic Analysis 1, 175–185.
Baltagi, B.H., Li, Q., 1992. Prediction in the one-way error component model with serial correlation. Journal of Forecasting 11, 561–567.
Brucker, H., Siliverstovs, B., 2006. On the estimation and forecasting of international migration: how relevant is heterogeneity across countries. Empirical Economics 31, 735–754.
Fingleton, B., 2008a. A generalized method of moments estimation for a spatial panel model with an endogenous spatial lag and spatial moving average errors. Spatial Economic Analysis 3, 27–44.
Fingleton, B., 2008b. A generalized method of moments estimation for a spatial model with moving average errors with application to real estate prices. Empirical Economics 34, 35–57.
Frees, E.W., Miller, T.W., 2004. Sales forecasting using longitudinal data models. International Journal of Forecasting 20, 99–114.
Goldberger, A.S., 1962. Best linear unbiased prediction in the generalized linear regression model. Journal of the American Statistical Association 57, 369–375.
Hoogstrate, A.J., Palm, F.C., Pfann, G.A., 2000. Pooling in dynamic panel-data models: an application to forecasting GDP growth rates. Journal of Business and Economic Statistics 18, 274–283.
Hsiao, C., Tahmiscioglu, A.K., 1997. A panel analysis of liquidity constraints and firm investment. Journal of the American Statistical Association 92, 455–465.
Kapoor, M., Kelejian, H.H., Prucha, I.R., 2007. Panel data models with spatially correlated error components. Journal of Econometrics 140, 97–130.
Kelejian, H.H., Prucha, I.R., 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. International Economic Review 40, 509–533.
Lee, L.F., Griffiths, W.E., 1979. The prior likelihood and best linear unbiased prediction in stochastic coefficient linear models. Working paper. Department of Economics, University of Minnesota.
Nerlove, M., 1971. Futher evidence on the estimation of dynamic economic relations from a time-series of cross-sections. Econometrica 39, 359–382.
Pesaran, M.H., Smith, R., 1995. Estimating long-run relationships from dynamic heterogenous panels. Journal of Econometrics 68, 79–113.
Rapach, D.E., Wohar, M.E., 2004. Testing the monetary model of exchange rate determination: a closer look at panels. Journal of International Money and Finance 23, 867–895.
Schmalensee, R., Stoker, T.M., Judson, R.A., 1998. World carbon dioxide emissions: 1950–2050. Review of Economics and Statistics 80, 15–27.
Taub, A.J., 1979. Prediction in the context of the variance-components model. Journal of Econometrics 10, 103–108.
Wansbeek, T.J., Kapteyn, A., 1978. The seperation of individual variation and systematic change in the analysis of panel data. Annales de l'INSEE 30–31, 659–680.