Journal of the Korean Statistical Society (2000), **29**: *2, pp 231–245*

Analysis of Linear Regression Model with Two Way Correlated Errors †

Seuck Heun Song[1]

ABSTRACT

This paper considers a linear regression model with space and time data in where the disturbances follow spatially correlated error components. We provide the best linear unbiased predictor for the one way error component model with spatial autocorrelation. Further, we derive two diagnostic test statistics for the assessment of model specification due to spatial dependence and random effects as an application of the Lagrange Multiplier principle.

Keywords: Space and time data, Error components model, Spatial autocorrelation, BLUP, LM tests.

1. INTRODUCTION

In studies of regional science, urban economics and enviroment, the analysis with spatially autocorrelated error models is widely applied(Anselin and Florax (1995) and Anselin (1988)). Panel data adds an error components dimension due to space(region) effects, see Anselin (1988). In particular, we consider the panel data regession model with spatial autocorrelation. In this paper we will present an illustration of how spatial dependence can be incorporated and how tests for its presence can be developed.

First, we consider also the question of a proper prediction in the context of the spatial panel data models. Most researchers have been applied Goldberger's (1962) result in the various error components models (Taub (1979), Koning (1989) and Baltagi and Li (1992)). We will provide the best linear unbiased predictor(BLUP) for the error components model with spatial autocorrelation.

Breusch and Pagan (1980), Engle (1984) and Godfrey (1989) demonstrated the wide applicability of Lagrange Multiplier(LM) test to various model specifications in regression models. One important application considered is that of the error components model, where a LM test was developed to test for zero random

†This work was supported by Korea Research Foundation Grant(KRF-98-003-D00059)
[1]Department of Statistics, Korea University, Seoul 136-701, Korea

effects. The LM test computation requires only ordinary least squares residuals and this test is much simpler to compute than the Likelihood Ratio(LR) test. This paper extends the Breusch and Pagan's LM test to the spatial panel data model case.

The outline of the paper is as follows. Section 2 gives a linear regression model for space and time data with a spatially correlated error components. In section 3 the BLUP for this model is provided. Section 4 derives first a LM test statistic which jointly tests for the presence of random space(region) effects and spatial correlation and the second LM test statistic for existence of spatial correlation given random space(region) effects. Section 5 compares the performance of LM tests with LR tests using Monte Carlo experiments.

2. THE MODEL

Consider the following spatial linear regression model

$$y_{it} = x_{it}'\beta + u_{it}, \qquad i = 1, \cdots, N \text{ and } t = 1, \cdots, T, \tag{2.1}$$

where y_{it} is an observation on a dependent variable for the tth time period and for the ith spatial unit (for example, country, state and census track). x_{it} denotes a nonstochastic regressor vector of k independent variables for spatial unit i at time period t. β is a vector of regression coefficients to be estimated and u_{it} is a disturbance. The disturbance is assumed to incorporate unobserved effects due to space μ_i as well as the usual disturbance ε_{it}. For example, the unobservable space specific effects could be due to aspects of regional structure and this effect is assumed to be constant over time. Formally, the overall disturbances consist of two error components, as:

$$u_{it} = \mu_i + \varepsilon_{it}, \qquad i = 1, \cdots, N \text{ and } t = 1, \cdots, T, \tag{2.2}$$

where μ_i denote the ith space specific effects which are assumed to be $i.i.N(0, \sigma_\mu^2)$ and the ε_{it} are the remainder disturbances.

A more general specification for the disturbances, see Anseln (1988), is that the remainder disturbance in (2.2) be generated by a spatial autoregressive process

$$\varepsilon_{it} = \lambda \sum_{k=1}^{N} w_{ik}\varepsilon_{kt} + \nu_{it}, \qquad i = 1, \cdots, N \text{ and } t = 1, \cdots, T, \tag{2.3}$$

where λ is a spatial autoregressive coefficient, w_{ik} is the (i,k) th elements of W which is a $N \times N$ weight matrix(Cliff and Ord (1981)) whose diagonal elements are zero and ν_{it} are the remainder disturbances which are also assumed to be $i.i.N(0, \sigma_\nu^2)$. In vector form, the model (2.1) can be written as

$$y = X\beta + u, \tag{2.4}$$

where y is an $NT \times 1$ observation vector, X is an $NT \times k$ design matrix, β is a $k \times 1$ vector of regression coefficients to be estimated and u is an $NT \times 1$ disturbance vector. Both N and T are assumed to be larger than k. Also, the equation (2.3) can be rewritten as

$$\varepsilon = \Big[(I_N - \lambda W)^{-1} \otimes I_T\Big]\nu. = (B^{-1} \otimes I_T)\nu. \tag{2.5}$$

Therefore, the overall disturbance vector of (2.2) can be presented as

$$\begin{aligned} u &= (I_N \otimes \iota_T)\mu + \varepsilon \\ &= (I_N \otimes \iota_T)\mu + \Big(B^{-1} \otimes I_T\Big)\nu, \end{aligned} \tag{2.6}$$

where $\mu' = (\mu_1, \cdots, \mu_N)$, $\nu' = (\nu_{11}, \cdots, \nu_{N1}, \cdots, \nu_{NT})$, ι_T is a vector of ones of dimension T, I_N and I_T are the identity matrix of dimension N and T, respectively, and $\otimes$ denotes the Kronecker product, and $B = I_N - \lambda W$. Under these assumptions, the variance-covariance matrix of u is given by

$$E(uu') = \Omega = \sigma_\mu^2(I_N \otimes J_T) + \sigma_\nu^2((B'B)^{-1} \otimes I_T), \tag{2.7}$$

where J_T is a matrix of ones of dimension T.

3. BEST LINEAR UNBIASED PREDICTOR

Now let us focus on prediction. Goldberger (1962) showed that, for any Ω, the best linear unbiased predictor(BLUP) for $y_{i,T+s}$ is given by

$$\widehat{y}_{i,T+s} = x'_{i,T+s}\widehat{\beta}_{GLS} + w'\Omega^{-1}\widehat{u}_{GLS}, \quad i = 1, \cdots, N, \tag{3.1}$$

where $w' = E\big(u_{i,T+s}u'\big)$, $u_{i,T+s} = \mu_i + \varepsilon_{i,T+s} = \mu_i + b'_i\nu_{T+s}$, where ν_{T+s} is the $N \times 1$ vector of $(T+s)$ th time periods and b'_i is the ith row of B^{-1}, and $u = (I_N \otimes \iota_T)\nu + (B^{-1} \otimes I_T)\nu$ and $\widehat{u}_{GLS} = y - X\widehat{\beta}_{GLS}$.

To derive the predictor we need the inverse matrix Ω, but the inverse of Ω could be obtained by the "method of tearing", we suggested a simple and useful tool to get the inverse of Ω.

Lemma 1: Consider an $Mn \times Mn$ matrix Q which can be expressed as $Q = \sum_{i=1}^{r} D_i \otimes Q_i$ where r is a finite integer, Q_i is an $(M \times M)$ nonsingular matrix for

$i = 1, \cdots, r$, and D_i is a symmetric, idempotent and orthogonal $(n \times n)$ matrix such that $\sum_{i=1}^{r} D_i = I_n$. Then, we have

$$Q^{-1} = \sum_{i=1}^{r} D_i \otimes Q_i^{-1}. \tag{3.2}$$

Proof : Since Q is positive definite matrix, the Q^{-1} is uniquely existence. To this show, direct multiplication of Q and Q^{-1} reveals that

$$\begin{aligned} QQ^{-1} &= \Big(\sum_{i=1}^{r} D_i \otimes Q_i\Big)\Big(\sum_{i=1}^{r} D_i \otimes Q_i^{-1}\Big) \\ &= \sum_{i=1}^{r} \Big(D_i D_i \otimes Q_i Q_i^{-1}\Big) + \sum_{i \neq j}^{r} \Big(D_i D_j \otimes Q_i \Omega_j^{-1}\Big) \\ &= \sum_{i=1}^{r} \Big(D_i \otimes I_M\Big) \\ &= \sum_{i=1}^{r} D_i \otimes I_M = I_n \otimes I_M, \end{aligned}$$

where the third equality follows from the fact that D_i is an idempotent and orthogonal matrix and the fourth equality uses the property of Kronecker products over addition of matrices. Therefore, Q^{-1} given by (3.2) is the unique inverse of Q.

In order to get Ω^{-1}, in (2.7) we replace J_T by $T\bar{J}_T$ and I_T by $E_T + \bar{J}_T$, where $E_T = I_T - \bar{J}_T$ and collect terms with the same matrices for the variance-covariance matrix. This gives

$$\Omega = \Big\{T\sigma_\mu^2 I_N + \sigma_\nu^2 (B'B)^{-1}\Big\} \otimes \bar{J}_T + \Big\{\sigma_\nu^2 (B'B)^{-1}\Big\} \otimes E_T. \tag{3.3}$$

It is easy to verify that $\bar{J}_T$ and E_T are symmetric, idempotent, and sum to the identity matrix I_T. Next, we assume $T\sigma_\mu^2 I_N + \sigma_\nu^2 (B'B)^{-1}$ is a non-singular matrix so that all the conditions of Lemma 1 are satisfied. Therefore, using the Lemma 1, the inverse matrix of Ω is given by

$$\begin{aligned} \Omega^{-1} &= \Big\{T\sigma_\mu^2 I_N + \sigma_\nu^2 (B'B)^{-1}\Big\}^{-1} \otimes \bar{J}_T + \Big\{\sigma_\nu^2 (B'B)^{-1}\Big\}^{-1} \otimes E_T \\ &= Z \otimes \bar{J}_T + \frac{1}{\sigma_\nu^2}(B'B) \otimes E_T, \end{aligned} \tag{3.4}$$

where $Z = [T\sigma_\mu^2 I_N + \sigma_\nu^2 (B'B)^{-1}]^{-1}$.

Since Ω^{-1} is given in (3.4) and the last term of equation (3.1) is given by

$$\begin{aligned} w'\Omega^{-1}\widehat{u}_{GLS} &= E\big[u_{i,T+s}u'\big]\Omega^{-1}\widehat{u}_{GLS} \\ &= E\Big[(\mu_i + b_i'\nu_{T+s})\Big\{(I_N \otimes \iota_T)\mu + (B^{-1}\otimes I_T)\nu\Big\}\Big]\Omega^{-1}\widehat{u}_{GLS} \\ &= \sigma_\mu^2(l_i' \otimes \iota_T')\Big[Z \otimes \bar{J}_T + \frac{1}{\sigma_\nu^2}(B'B)\otimes E_T\Big]\widehat{u}_{GLS} \\ &= \sigma_\mu^2(l_i'Z \otimes \iota_T')\widehat{u}_{GLS} = \sigma_\mu^2\sum_{k=1}^{N} z_{ik}\sum_{t=1}^{T}\widehat{u}_{kt}, \end{aligned} \tag{3.5}$$

where l_i in the third equation is the ith column of I_N, and the second equality follows from the fact that μ_i is the independent of ν and ν_{T+s} is independent of ν_t, $t = 1, \cdots, T$. Also, the fourth equality follows from the fact that $\iota_T'\bar{J}_T = \iota_T'$ and $\iota_T'E_T = 0$, and z_{ik} in last equality is the (i,k)th elements of Z. Therefore the BLUP for $\widehat{y}_{i,T+s}$ is given by

$$\widehat{y}_{i,T+s} = x_{i,T+s}'\widehat{\beta}_{GLS} + \sigma_\mu^2\sum_{k=1}^{N} z_{ik}\sum_{t=1}^{T}\widehat{u}_{kt}. \tag{3.6}$$

4. LM TEST STATISTICS

The log-likelihood function under normality of the disturbances is given by

$$L(\beta, \sigma_\nu^2, \sigma_\mu^2, \lambda) = Constant - \frac{1}{2}log|\Omega| - \frac{1}{2}u'\Omega^{-1}u.$$

Using the result of Magnus (1982), we obtain $|\Omega| = |T\sigma_\mu^2 I_N + \sigma_\nu^2(B'B)^{-1}| \cdot |\sigma_\nu^2(B'B)^{-1}|^{T-1}$. Therefore, the log likelihood function can be rewritten as

$$\begin{aligned} L(\beta, \sigma_\nu^2, \sigma_\mu^2, \lambda) &= Costant - \frac{1}{2}log\Big|T\sigma_\mu^2 I_N + \sigma_\nu^2(B'B)^{-1}\Big| \\ &\quad - \frac{N(T-1)}{2}log(\sigma_\nu^2) + (T-1)log|B| - \frac{1}{2}u'\Omega^{-1}u, \end{aligned} \tag{4.1}$$

where Ω^{-1} is given by (3.4).

4.1 A joint LM test for $\lambda = \sigma_\mu^2 = 0$

Let us first consider the joint hypothesis for testing $H_0 : \sigma_\mu^2 = 0$ and $\lambda = 0$. Let $\theta = (\sigma_\nu^2, \sigma_\mu^2, \lambda)'$, then the part of the information matrix corresponding to β will be ignored in computing the LM test statistic, since the information matrix between the θ and β parameters will be block diagonal and the first derivatives with respect to β evaluated at the restricted maximum likelihood estimation(mle) will be zero. Therefore the LM test statistic is given by

$$LM = \widetilde{D}_\theta'\widetilde{J}_\theta^{-1}\widetilde{D}_\theta, \tag{4.2}$$

where $\widetilde{D}_\theta = (\partial L/\partial\theta)(\widetilde{\theta})$ is a 3×1 vector of partial derivatives with respect to each element of θ, evaluated at the restricted mle $\widetilde{\theta}$ and $\widetilde{J}_\theta = E[-\partial^2 L/\partial\theta\partial\theta'](\widetilde{\theta})$ is the information matrix corresponding to θ, evaluated at the restricted mle $\widetilde{\theta}$. Under the null hypothesis, the variance-covariance matrix reduces to $\Omega^* = \Omega = \sigma_\nu^2 I_{NT}$ and the restricted mle of β is $\widetilde{\beta}_{OLS}$, so that $\widetilde{u} = y - X'\widetilde{\beta}_{OLS}$ are the OLS residuals and $\widetilde{\sigma}_\nu^2 = \widetilde{u}'\widetilde{u}/NT$. Hartley and Rao (1967) and Hemmerle and Hartley (1973) give a useful general formula to obtain $\widetilde{D}_\theta$:

$$\partial L/\partial\theta_r = -\frac{1}{2}tr[\Omega^{-1}(\partial\Omega/\partial\theta_r)] + \frac{1}{2}[u'\Omega^{-1}(\partial\Omega/\partial\theta_r)\Omega^{-1}u], \ for\ r = 1,2,3. \quad (4.3)$$

Using the formula of (4.3), we obtain

$$\frac{\partial L}{\partial\sigma_\nu^2}|_{H_0} = 0, \qquad \frac{\partial L}{\partial\sigma_\mu^2}|_{H_0} = D(\widetilde{\sigma}_\mu^2) = \frac{NT}{2\widetilde{\sigma}_\nu^2}(\frac{\widetilde{u}'(I_N\otimes J_T)\widetilde{u}}{\widetilde{u}'\widetilde{u}} - 1),$$
$$\frac{\partial L}{\partial\lambda}|_{H_0} = D(\widetilde{\lambda}) = \frac{NT}{2}\frac{\widetilde{u}'(W + W'\otimes I_T)\widetilde{u}}{\widetilde{u}'\widetilde{u}} = NT\frac{\widetilde{u}'(W\otimes I_T)\widetilde{u}}{\widetilde{u}'\widetilde{u}}.$$

Therefore, the partial derivatives with respect to each element of θ, evaluated at the restricted mle is given by

$$\widetilde{D}_\theta = \begin{bmatrix} 0 \\ D(\widetilde{\sigma}_\mu^2) \\ D(\widetilde{\lambda}) \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{NT}{2\widetilde{\sigma}_\nu^2}(\frac{\widetilde{u}'(I_N\otimes J_T)\widetilde{u}}{\widetilde{u}'\widetilde{u}} - 1) \\ NT\frac{\widetilde{u}'(W\otimes I_T)\widetilde{u}}{\widetilde{u}'\widetilde{u}} \end{bmatrix}. \quad (4.4)$$

The information matrix for this model using the formula of Harville (1977) is (see Appendix (1))

$$\widetilde{J}_\theta = \frac{NT}{2\widetilde{\sigma}_\nu^4}\begin{bmatrix} 1 & 1 & 0 \\ 1 & T & 0 \\ 0 & 0 & \frac{2b\widetilde{\sigma}_\nu^4}{N} \end{bmatrix}, \quad (4.5)$$

where $b = tr(W^2 + W'W)$. Therefore, the resulting LM statistic of (4.2) is given by

$$LM_1 = \widetilde{D}_\theta'\widetilde{J}_\theta^{-1}\widetilde{D}_\theta = \frac{NT}{2(T-1)}A^2 + \frac{N^2T}{b}B^2, \quad (4.6)$$

where $A = \frac{\widetilde{u}'(I_N\otimes J_T)\widetilde{u}}{\widetilde{u}'\widetilde{u}} - 1$ and $B = \frac{\widetilde{u}'(W\otimes I_T)\widetilde{u}}{\widetilde{u}'\widetilde{u}}$. Note that the A^2 is the basis for the LM test statistic for $H_0 : \sigma_\mu^2 = 0$, assuming there is no spatial correlation(see Breusch and Pagan (1980)), while B^2 is the basis for the LM_1

test statistic for $H_0 : \lambda = 0$, assuming there is no individual effects. Under the null hypothesis, the LM_1 test statistic of (4.6) is asymptotically distributed as χ^2_2.

4.2 An LM test for $\lambda = 0$ given $\sigma^2_\mu > 0$

Furthermore, we consider the LM test for spatially uncorrelated given the existence of random individual effects. The null hypothesis for this model is $H_0 : \lambda = 0$(given $\sigma^2_\mu > 0$). Under the null hypothesis, the variance-covariance matrix reduces to $\Omega_0 = \sigma^2_\mu(I_N \otimes J_T) + \sigma^2_\nu(I_N \otimes I_T)$. It is the familiar form of the one-way error component model, see Baltagi (1995), and $\Omega_0^{-1} = (\sigma_1^2)^{-1}(I_N \otimes \bar{J}_T) + (\sigma^2_\nu)^{-1}(I_N \otimes E_T)$, where $\sigma_1^2 = T\sigma^2_\mu + \sigma^2_\nu$. Using the analogous derivation of equation (4.3), we obtain

$$\frac{\partial L}{\partial \sigma^2_\nu}|_{H_0} = \frac{\partial L}{\partial \sigma^2_\mu}|_{H_0} = 0$$

$$\frac{\partial L}{\partial \lambda}|_{H_0} = D(\widehat{\lambda}) = \widehat{u}'[\frac{\widehat{\sigma}^2_\nu}{\widehat{\sigma}^4_1}(W \otimes \bar{J}_T) + \frac{1}{\widehat{\sigma}^2_\nu}(W \otimes E_T)]\widehat{u},$$

where $\widehat{\sigma}^2_\nu = \widehat{u}'(I_N \otimes E_T)\widehat{u}/N(T-1)$ and $\widehat{\sigma}^2_1 = \widehat{u}'(I_N \otimes \bar{J}_T)\widehat{u}/N$ are the mle of σ^2_ν and σ^2_1, and $\widehat{u}$ is the maximum likelihood residuals under the null hypothesis. Therefore, we have

$$\widehat{D}_\theta = \begin{bmatrix} 0 \\ 0 \\ D_{\widehat{\lambda}} \end{bmatrix}. \tag{4.7}$$

Next, the information matrix for this model, when evaluated under the null hypotesis ($\sigma^2_\mu = 0$) is given by(see Appendix (2))

$$\widehat{J}_\theta = \begin{bmatrix} \frac{N}{2}(\frac{1}{\widehat{\sigma}^4_1} + \frac{T-1}{\widehat{\sigma}^4_\nu}) & \frac{NT}{2\widehat{\sigma}^4_1} & 0 \\ \frac{NT}{2\widehat{\sigma}^4_1} & \frac{NT^2}{2\widehat{\sigma}^4_1} & 0 \\ 0 & 0 & (T-1)b + \frac{\widehat{\sigma}^4_\nu}{\widehat{\sigma}^4_1}\, b \end{bmatrix}. \tag{4.8}$$

Therefore, the resulting LM statistic of (4.2) is given by

$$LM_2 = \widehat{D}'_\theta \widehat{J}_\theta^{-1} \widehat{D}_\theta = \frac{\widehat{D}(\lambda)^2}{[(T-1) + \frac{\widehat{\sigma}^4_\nu}{\widehat{\sigma}^4_1}]\, b}. \tag{4.9}$$

Under the null hypothesis, the LM_2 statistic of (4.9) is asymptotically distributed as χ^2_1.

5. MONTE CARLO RESULTS

Consider the following simple regression model

$$y_{it} = \alpha + x'_{it}\beta + u_{it}, \qquad i = 1, \cdots, N \text{ and } t = 1, \cdots, T, \tag{5.1}$$

where $\alpha = 5$ and $\beta = 0.5$, x_{it} was generated by a similar method to that of Nerlove (1971). In fact $x_{it} = 0.1t + 0.5x_{i,t-1} + w_{it}$, where w_{it} is the uniformly distributed on the interval $(-0.5, 0.5)$. The initial values x_{i0} was chosen as $(5 + 10w_{i0})$. For the disturbances, $u_{it} = \mu_i + \varepsilon_{it}, \varepsilon_{it} = \lambda\sum_{k=1}^{N} w_{ik}\varepsilon_{kt} + \nu_{it}$ with $\mu_i \sim IIN(0, \sigma_\mu^2)$, $\nu_i \sim IIN(0, \sigma_\nu^2)$ and W matrix is rook or queen type weight matrix (see Cliff and Ord (1981, pp.243)). We fix $\sigma_\mu^2 + \sigma_\nu^2 = 20$ and let $\rho = \sigma_\mu^2/(\sigma_\mu^2 + \sigma_\nu^2)$ were varied over the set $(0, 0.2, 0.5, 0.8)$. Spatial autocorrelation factor λ is varied over positive set which was from 0 to 0.9 by 0.1. For sample size combinations the following (N, T) are used: (N, T) : $(25, 3)(25, 7)(49, 3)(49.7)$. For each experimemt, 1000 replication are performed. For each replication, we calculate the LM_1, LM_2 test and corresponding LR test, respectively. Finally, the significance level is set as 0.05.

Table 1 gives the power of the LM_1 and LR_1 test employed for $H_0 : \sigma_\mu^2 = \lambda = 0$ when $(N, T) = (25, 3)$ and $(N, T) = (25, 7)$. (Similar tables for other combinations of (M, N, T) are not produced here to save space. These results are available upon request from the authors.) The LM_1 test has estimated size that is not significantly different from the nominal size. However, the LR_1 test significantly underestimate the nominal size for all cases. The power of all the tests increase as ρ or λ increases. The power of LM_1 test is similar to that of LR_1 test and the power of tests using the Rook type weight matrix has higher than that of tests using the Queen type weight matrix.

Table 2 gives the power of the LM_2 and LR_2 test for testing $H_0 : \lambda = 0$ (given $\sigma_\mu^2 > 0$) when $(N, T) = (25, 3)$ and $(N, T) = (25, 7)$. The LR_2 test also significantly underestimate the nominal size for all cases, and therefore LR_2 test is not recommended. But, the estimated size of the LM_2 test is not different from the nominal size. The power of all the tests increase as T increase and ρ and(or) λ increases. In this case, the power of tests using the Rook type also is higher than that of tests using the Queen type.

6. CONCLUSION

This paper provided the BLUP for the spatially correlated error component model. In the context of this model two types of LM test are derived. The

LM_1 test statistic which jointly tests for the presence of random space effects and spatial correlation and the second LM_2 test statistic for the existence of spatial correlation given random space effects. Using the Monte-Carlo studies the following results are obtained. (*i*) The estimated size of the LM tests is not significantly different from the nominal size. (*ii*) The LR test significantly underestimate the nominal size, and it is not recommended. (*iii*) The power of tests using the Rook type weight matrix is higher than that of tests using the Queen type weight matrix.

APPENDIX

In this Appendix, we give the detail derivation of the information matrix for testing $H_0 : \sigma_\mu^2 = \lambda = 0$ and $H_0 : \lambda = 0 (\sigma_\mu^2 > 0)$. The variance-covariance matrix for this model is given by (2.7) and the likelihood function is given by (4.1). Let $\theta' = (\sigma_\nu^2, \sigma_\mu^2, \lambda)$, then using the formula of Harville (1977), we have

$$J_{rs} = E\Big[- \partial^2 L / \partial\theta_r \partial\theta_s \Big] = \frac{1}{2} tr \Big[\Omega^{-1} \Big(\partial\Omega / \partial\theta_r \Big) \Omega^{-1} \Big(\partial\Omega / \partial\theta_s \Big) \Big], \qquad (A.1)$$

for $r, s = 1, 2, 3$. It is easily checked that $\partial\Omega / \partial\sigma_\nu^2 = I_N \otimes I_T$, $\partial\Omega / \partial\sigma_\mu^2 = I_N \otimes J_T$ and $\partial\Omega / \partial\lambda = \sigma_\nu^2[(B'B)^{-1}(W'B + B'W)(B'B)^{-1} \otimes I_T]$.

(1) Under the null hypothesis($H_0 : \sigma_\mu^2 = 0$ and $\lambda = 0$), we obtain

$$\Omega^{-1}|_{H_0} = \frac{1}{\sigma_\nu^2} I_N \otimes I_T,$$

$$\frac{\partial\Omega}{\partial\sigma_\nu^2}|_{H_0} = I_N \otimes I_T, \quad \frac{\partial\Omega}{\partial\sigma_\mu^2}|_{H_0} = I_N \otimes J_T, \quad \frac{\partial\Omega}{\partial\lambda}|_{H_0} = \sigma_\nu^2 (W + W') \otimes I_T. \quad (A.2)$$

Therefore, the elements of the information matrix for this model using (A.1) and (A.2) are given by

$$\begin{aligned} J_{11} = E\Big[- \frac{\partial^2 L}{\partial(\sigma_\nu^2)^2} \Big] &= \frac{1}{2} tr \Big[(\sigma_\nu^2)^{-2} (I_N \otimes I_T)(I_N \otimes I_T)(I_N \otimes I_T)(I_N \otimes I_T) \Big] \\ &= \frac{NT}{2\sigma_\nu^4} \\ J_{12} = E\Big[- \frac{\partial^2 L}{\partial\sigma_\nu^2 \partial\sigma_\mu^2} \Big] &= \frac{1}{2} tr \Big[(\sigma_\nu^2)^{-2} (I_N \otimes I_T)(I_N \otimes I_T)(I_N \otimes I_T)(I_N \otimes J_T) \Big] \\ &= \frac{NT}{2\sigma_\nu^4} \end{aligned}$$

$$
\begin{aligned}
J_{13} = E\Big[-\frac{\partial^2 L}{\partial\sigma_\nu^2\partial\lambda}\Big] &= \frac{1}{2}\Big[(\sigma_\nu^2)^{-2}(I_N\otimes I_T)(I_N\otimes I_T)(I_N\otimes I_T)\sigma_\nu^2((W+W')\otimes I_T)\Big] \\
&= \frac{1}{2\sigma_\nu^2}tr[(W+W')\otimes I_T] = 0 \\
J_{22} = E\Big[-\frac{\partial^2 L}{\partial(\sigma_\mu^2)^2}\Big] &= \frac{1}{2}tr\Big[(\sigma_\nu^2)^{-2}(I_N\otimes I_T)(I_N\otimes J_T)(I_N\otimes I_T)(I_N\otimes J_T)\Big] \\
&= \frac{NT^2}{2\sigma_\nu^4} \\
J_{23} = E\Big[-\frac{\partial^2 L}{\partial\sigma_\mu^2\partial\lambda}\Big] &= \frac{1}{2}tr\Big[(\sigma_\nu^2)^{-1}(I_N\otimes I_T)(I_N\otimes J_T)(I_N\otimes I_T)((W+W')\otimes I_T)\Big] \\
&= \frac{1}{2\sigma_\nu^2}tr[W+W'\otimes J_T] = 0 \\
J_{33} = E\Big[-\frac{\partial^2 L}{\partial\lambda^2}\Big] &= \frac{1}{2}tr[(I_N\otimes I_T)((W+W')\otimes I_T)(I_N\otimes I_T)((W+W')\otimes I_T)\Big] \\
&= \frac{1}{2}tr[(2W^2+2W'W)\otimes I_T] = Tb,
\end{aligned}
$$

where the result of $J_{13}=0$ and $J_{23}=0$ follows from the fact that the diagonal elements of W is 0 and J_{33} uses that $tr(W^2)=tr(W'^2)$, and $b=tr(W^2+W'W)$. Since $J_{12}=J_{21}$, $J_{13}=J_{31}$ and $J_{23}=J_{32}$, the information matrix is given by

$$
\widetilde{J}_\theta = \frac{NT}{2\widetilde{\sigma}_\nu^4}\begin{bmatrix} 1 & 1 & 0 \\ 1 & T & 0 \\ 0 & 0 & \dfrac{2b\widetilde{\sigma}_\nu^4}{N} \end{bmatrix}. \tag{A.3}
$$

(2) Under the null hypothesis($H_0:\lambda=0(\sigma_\mu^2>0)$), we have

$$
\Omega^{-1}|_{H_0} = I_N\otimes(\frac{1}{\sigma_1^2}\bar{J}_T+\frac{1}{\sigma_\nu^2}E_T),\quad \text{where } \sigma_1^2 = T\sigma_\mu^2+\sigma_\nu^2,
$$

$$
\frac{\partial\Omega}{\partial\sigma_\nu^2}|_{H_0} = I_N\otimes I_T,\ \frac{\partial\Omega}{\partial\sigma_\mu^2}|_{H_0} = I_N\otimes J_T,\ \frac{\partial\Omega}{\partial\lambda}|_{H_0} = \sigma_\nu^2(W+W')\otimes I_T. \tag{A.4}
$$

Using (A.1) and (A.4) the elements of the information matrix are given by

$$
\begin{aligned}
J_{11} = E\Big[-\frac{\partial^2 L}{\partial(\sigma_\nu^2)^2}\Big] &= \frac{1}{2}tr\Big[\{(I_N\otimes I_T)(I_N\otimes(\sigma_1^{-2}\bar{J}_T+\sigma_\nu^{-2}E_T)\}^2\Big] \\
&= \frac{N}{2}(\frac{1}{\sigma_1^4}+\frac{T-1}{\sigma_\nu^4})
\end{aligned}
$$

$$
\begin{aligned}
J_{22} = E\Big[-\frac{\partial^2 L}{\partial(\sigma_\mu^2)^2}\Big] &= \frac{1}{2}tr\Big[\{(I_N\otimes J_T)(I_N\otimes(\sigma_1^{-2}\bar{J}_T+\sigma_\nu^{-2}E_T))\}^2\Big] \\
&= \frac{NT^2}{2\sigma_1^4} \\
J_{33} = E\Big[-\frac{\partial^2 L}{\partial\lambda^2}\Big] &= \frac{1}{2}tr\Big[\{\sigma_\nu^2((W+W')\otimes I_T)(I_N\otimes(\sigma_1^{-2}\bar{J}_T+\sigma_\nu^{-2}E_T))\}^2\Big] \\
&= \frac{\sigma_\nu^4}{2}tr\Big[(2W^2+2W'W)\otimes(\sigma_1^{-4}\bar{J}_T+\sigma_\nu^{-4}E_T))\Big] \\
&= \Big(\frac{\sigma_\nu^4}{\sigma_1^4}+T-1\Big)b \\
J_{12} = E\Big[-\frac{\partial^2 L}{\partial\sigma_\nu^2\partial\sigma_\mu^2}\Big] &= \frac{1}{2}tr\Big[(I_N\otimes I_T)(I_N\otimes(\sigma_1^{-2}\bar{J}_T+\sigma_\nu^{-2}E_T)) \\
&\quad (I_N\otimes J_T)(I_N\otimes(\sigma_1^{-2}\bar{J}_T+\sigma_\nu^{-2}E_T))\Big] = \frac{NT}{2\sigma_1^4} \\
J_{13} = E\Big[-\frac{\partial^2 L}{\partial\sigma_\nu^2\partial\lambda}\Big] &= \frac{1}{2}tr\Big[(I_N\otimes I_T)(I_N\otimes(\sigma_1^{-2}\bar{J}_T+\sigma_\nu^{-2}E_T))(\sigma_\nu^2(W+W')\otimes I_T) \\
&\quad (I_N\otimes(\sigma_1^{-2}\bar{J}_T+\sigma_\nu^{-2}E_T))\Big] \\
&= \frac{\sigma_\nu^2}{2}tr\Big[(W+W')\otimes(\sigma_1^{-4}\bar{J}_T+\sigma_\nu^{-4}E_T\Big] = 0 \\
J_{23} = E\Big[-\frac{\partial^2 L}{\partial\sigma_\mu^2\partial\lambda}\Big] &= \frac{1}{2}tr\Big[(I_N\otimes J_T)(I_N\otimes(\sigma_1^{-2}\bar{J}_T+\sigma_\nu^{-2}E_T))(\sigma_\nu^2(W+W')\otimes I_T) \\
&\quad (I_N\otimes(\sigma_1^{-2}\bar{J}_T+\sigma_\nu^{-2}E_T))\Big] \\
&= \frac{\sigma_\nu^2}{2}tr\Big[(W+W')\otimes\sigma_1^4 J_T\Big] = 0,
\end{aligned}
$$

where the result of $J_{13} = 0$ and $J_{23} = 0$ follows from the fact that the diagonal elements of W is 0 and J_{33} uses that $tr(W^2) = tr(W'^2)$, and $b = tr(W^2 + W'W)$. Since $J_{12} = J_{21}$, $J_{13} = J_{31}$ and $J_{23} = J_{32}$, the information matrix is given by

$$
\widehat{J}_\theta = \begin{bmatrix} \frac{N}{2}(\frac{1}{\widehat{\sigma}_1^4}+\frac{T-1}{\widehat{\sigma}_\nu^4}) & \frac{NT}{2\widehat{\sigma}_1^4} & 0 \\ \frac{NT}{2\widehat{\sigma}_1^4} & \frac{NT^2}{2\widehat{\sigma}_1^4} & 0 \\ 0 & 0 & (T-1)b+\frac{\widehat{\sigma}_\nu^4}{\widehat{\sigma}_1^4}\,b \end{bmatrix}. \qquad (A.5)
$$

Table 1 : Power of the test: $H_0 = \sigma_\mu^2 = \lambda$

			$\rho = 0.0$		$\rho = 0.2$		$\rho = 0.5$		$\rho = 0.8$	
N, T	W	λ	LM_1	LR_1	LM_1	LR_1	LM_1	LR_1	LM_1	LR_1
25, 3	Rook	0.0	0.035	0.013	0.280	0.208	0.943	0.921	1.000	1.000
		0.1	0.060	0.042	0.269	0.243	0.935	0.929	1.000	1.000
		0.2	0.149	0.128	0.324	0.314	0.942	0.945	1.000	1.000
		0.3	0.350	0.306	0.422	0.483	0.939	0.946	1.000	1.000
		0.4	0.593	0.591	0.556	0.673	0.933	0.970	1.000	1.000
		0.5	0.832	0.839	0.764	0.838	0.948	0.989	1.000	1.000
		0.6	0.941	0.948	0.908	0.956	0.949	0.991	0.999	1.000
		0.7	0.997	0.996	0.971	0.992	0.980	0.998	1.000	1.000
		0.8	0.999	0.999	1.000	1,000	0.995	1.000	1.000	1.000
		0.9	1.000	1.000	1.000	1,000	1.000	1.000	1.000	1.000
25, 3	Queen	0.0	0.039	0.007	0.287	0.219	0.932	0.912	1.000	1.000
		0.1	0.049	0.021	0.269	0.222	0.937	0.929	1.000	1.000
		0.2	0.123	0.087	0.298	0.272	0.945	0.942	1.000	1.000
		0.3	0.259	0.182	0.368	0.373	0.950	0.952	1.000	1.000
		0.4	0.484	0.401	0.492	0.508	0.947	0.961	0.999	1.000
		0.5	0.716	0.667	0.668	0.698	0.962	0.985	1.000	1.000
		0.6	0.874	0.846	0.832	0.856	0.969	0.986	1.000	1.000
		0.7	0.975	0.968	0.955	0.962	0.978	0.993	1.000	1.000
		0.8	0.998	0.996	0.984	0.993	0.994	0.999	1.000	1.000
		0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
25, 7	Rook	0.0	0.048	0.022	0.874	0.818	1.000	1.000	1.000	1.000
		0.1	0.105	0.078	0.903	0.858	1.000	1.000	1.000	1.000
		0.2	0.375	0.352	0.910	0.912	1.000	1.000	1.000	1.000
		0.3	0.761	0.740	0.956	0.970	1.000	1.000	1.000	1.000
		0.4	0.971	0.969	0.993	0.997	1.000	1.000	1.000	1.000
		0.5	0.997	0.997	0.998	1.000	1.000	1.000	1.000	1.000
		0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
25, 7	Queen	0.0	0.045	0.014	0.851	0.792	1.000	1.000	1.000	1.000
		0.1	0.111	0.062	0.875	0.832	1.000	1.000	1.000	1.000
		0.2	0.289	0.198	0.903	0.875	1.000	1.000	1.000	1.000
		0.3	0.601	0.518	0.935	0.925	1.000	1.000	1.000	1.000
		0.4	0.880	0.843	0.974	0.977	1.000	1.000	1.000	1.000
		0.5	0.982	0.974	0.995	0.997	1.000	1.000	1.000	1.000
		0.6	0.998	0.998	0.999	0.999	1.000	1.000	1.000	1.000
		0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 2 : Power of the test: $H_0 : \lambda = 0(\sigma_\mu^2 > 0)$

			$\rho = 0.0$		$\rho = 0.2$		$\rho = 0.5$		$\rho = 0.8$	
N, T	W	λ	LM_2	LR_2	LM_2	LR_2	LM_2	LR_2	LM_2	LR_2
25, 3	Rook	0.0	0.050	0.034	0.042	0.020	0.063	0.027	0.040	0.031
		0.1	0.062	0.084	0.062	0.077	0.061	0.074	0.066	0.080
		0.2	0.170	0.241	0.142	0.213	0.161	0.210	0.143	0.195
		0.3	0.364	0.493	0.315	0.419	0.303	0.378	0.331	0.391
		0.4	0.630	0.752	0.588	0.707	0.535	0.617	0.526	0.613
		0.5	0.827	0.922	0.782	0.849	0.732	0.808	0.741	0.816
		0.6	0.942	0.984	0.922	0.957	0.905	0.943	0.900	0.942
		0.7	0.988	1.000	0.978	0.990	0.974	0.990	0.976	0.990
		0.8	0.995	0.999	0.994	1.000	0.992	0.994	0.995	1.000
		0.9	0.994	1.000	0.993	1.000	0.998	1.000	1.000	1.000
25, 3	Queen	0.0	0.042	0.016	0.029	0.015	0.043	0.023	0.026	0.014
		0.1	0.052	0.053	0.063	0.060	0.051	0.046	0.060	0.051
		0.2	0.126	0.163	0.125	0.138	0.116	0.106	0.146	0.147
		0.3	0.248	0.295	0.251	0.276	0.273	0.273	0.273	0.270
		0.4	0.484	0.565	0.415	0.448	0.453	0.457	0.432	0.436
		0.5	0.708	0.789	0.663	0.699	0.663	0.667	0.646	0.647
		0.6	0.853	0.907	0.845	0.869	0.829	0.849	0.832	0.843
		0.7	0.949	0.981	0.956	0.970	0.935	0.948	0.930	0.937
		0.8	0.978	0.999	0.987	0.994	0.988	0.988	0.989	0.989
		0.9	0.981	1.000	0.993	1.000	0.999	0.999	0.998	0.998
25, 7	Rook	0.0	0.039	0.030	0.054	0.041	0.054	0.028	0.052	0.026
		0.1	0.121	0.159	0.146	0.164	0.131	0.147	0.140	0.157
		0.2	0.451	0.520	0.439	0.471	0.455	0.482	0.457	0.472
		0.3	0.805	0.864	0.802	0.826	0.822	0.838	0.825	0.834
		0.4	0.973	0.984	0.970	0.976	0.973	0.975	0.979	0.983
		0.5	0.998	0.999	0.997	0.998	0.999	0.999	0.999	0.999
		0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
25, 7	Queen	0.0	0.046	0.022	0.061	0.052	0.048	0.027	0.043	0.015
		0.1	0.128	0.116	0.122	0.106	0.131	0.117	0.110	0.100
		0.2	0.338	0.358	0.340	0.318	0.348	0.324	0.341	0.319
		0.3	0.656	0.675	0.668	0.647	0.689	0.673	0.680	0.661
		0.4	0.889	0.907	0.895	0.889	0.902	0.895	0.883	0.876
		0.5	0.986	0.987	0.980	0.980	0.983	0.984	0.984	0.984
		0.6	0.998	0.998	0.998	0.998	1.000	1.000	1.000	1.000
		0.7	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.9	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000

References

Anselin, L. and Florax, R.J.G.M. (1995). *Econometrics Analysis of Panel Data.* Wiley, New York.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models.* Kluwer Academic publishers, Dordrecht.

Baltagi, B.H. (1995). *Econometrics Analysis of Panel Data.* Wiley, New York.

Baltagi, B.H. and Li, Q. (1992). Prediction in the one-way error component model with serial correlation. *Journal of Forecasting,* **11**, 561-567.

Breusch, T.S. and Pagan, A.R. (1980). The Lagrange Multiplier test and its application to model specification in econometrics. *Review of Econonic Studies,* **47**, 239-254.

Cliff, A.D. and Ord, J.K. (1981). *Spatial Processes Models and Application.* Pion, London.

Engle, R. (1984). Wald, Likelihood–Ratio, Lagrange Multiplier tests in econometrics, in: Griliches, Z., M.D. Intriligator, eds., *Handbook of Econometrics.* North Holland, Amsterdam.

Godfrey, L.G. (1989). *Misspecification tests in econometrics.* Cambridge University Press, Cambridge.

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

Hartley, H. O. and Rao, J. N. K. (1967). Maximum likelihood estimation for the mixed analysis of variance model. *Biometrika,* **54**, 93-108.

Harville, D.A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association,* **72**, 320-338.

Hemmerle, W.J. and Hartley, H.O. (1973). Computing maximum likelihood estimates for the mixed A.O.V. model using the W-transformation. *Technometrics,* **15**, 819-831.

Koning, R.H. (1989). Prediction with a two-way error component regression model. *Econometric Theory,* **5**, 175.

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

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

Taub, A.J. (1979). Prediction in the context of the variance components model. *Journal of Econometrics,* **10**, 103-108.