Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Econometrics 117 (2003) 123–150
JOURNAL OF Econometrics
www.elsevier.com/locate/econbase

Testing panel data regression models with spatial error correlation

Badi H. Baltagi[a,*], Seuck Heun Song[b], Won Koh[c]

[a] *Department of Economics, Texas A&M University, College Station, TX 77843-4228, USA*
[b] *Department of Statistics, Korea University, Sungbuk-Ku, Seoul, 136-701, South Korea*
[c] *Center for DM&S, Korea Institute for Defense Analyses (KIDA), Cheong Ryang, P.O. Box 250, Seoul, 130-650, South Korea*

Accepted 6 February 2003

Abstract

This paper derives several lagrange multiplier (LM) tests for the panel data regression model with spatial error correlation. These tests draw upon two strands of earlier work. The first is the LM tests for the spatial error correlation model discussed in Anselin (Spatial Econometrics: Methods and Models, Kluwer Academic Publishers, Dordrecht; Rao's score test in spatial econometrics, J. Statist. Plann. Inference 97 (2001) 113) and Anselin et al. (Regional Sci. Urban Econom. 26 (1996) 77), and the second is the LM tests for the error component panel data model discussed in Breusch and Pagan (Rev. Econom. Stud. 47(1980) 239) and Baltagi et al. (J. Econometrics 54 (1992) 95). The idea is to allow for both spatial error correlation as well as random region effects in the panel data regression model and to test for their joint significance. Additionally, this paper derives conditional LM tests, which test for random regional effects given the presence of spatial error correlation. Also, spatial error correlation given the presence of random regional effects. These conditional LM tests are an alternative to the one-directional LM tests that test for random regional effects ignoring the presence of spatial error correlation or the one-directional LM tests for spatial error correlation ignoring the presence of random regional effects. We argue that these joint and conditional LM tests guard against possible misspecification. Extensive Monte Carlo experiments are conducted to study the performance of these LM tests as well as the corresponding likelihood ratio tests.

JEL classification: C23; C12

Keywords: Panel data; Spatial error correlation; Lagrange multiplier tests; Likelihood ratio tests

* Corresponding author. Tel.: 001-979-845-7380; fax: 001-979-847-8757.
E-mail addresses: badi@econmail.tamu.edu (B.H. Baltagi), ssong@mail.korea.ac.kr (S.H. Song), kohwon@kida.re.kr (W. Koh).

0304-4076/03/$ - see front matter
doi:10.1016/S0304-4076(03)00120-9

1. Introduction

Spatial dependence models deal with spatial interaction (spatial autocorrelation) and spatial structure (spatial heterogeneity) primarily in cross-section data, see Anselin (1988, 2001). Spatial dependence models use a metric of economic distance, see Anselin (1988) and Conley (1999) to mention a few. This measure of economic distance provides cross-sectional data with a structure similar to that provided by the time index in time series. There is an extensive literature estimating these spatial models using maximum likelihood methods, see Anselin (1988). More recently, generalized method of moments have been proposed by Kelejian and Prucha (1999) and Conley (1999). Testing for spatial dependence is also extensively studied by Anselin (1988, 2001), Anselin and Bera (1998), Anselin et al. (1996) to mention a few.

With the increasing availability of micro- as well as macro-level panel data, spatial panel data models studied in Anselin (1988) are becoming increasingly attractive in empirical economic research. See Case (1991), Kelejian and Robinson (1992), Case et al. (1993), Holtz-Eakin (1994), Driscoll and Kraay (1998), Baltagi and Li (1999) and Bell and Bockstael (2000) for a few applications. Convergence in growth models that use a pooled set of countries over time could have spatial correlation as well as heterogeneity across countries to contend with, see De Long and Summers (1991) and Islam (1995) to mention a few studies. County level data over time, whether it is expenditures on police, or measuring air pollution levels can be treated with these models. Also, state level expenditures over time on welfare benefits, mass transit, etc. Household level survey data from villages observed over time to study nutrition, female labor participation rates, or the effects of education on wages could exhibit spatial correlation as well as heterogeneity across households and this can be modeled with a spatial error component model.

Estimation and testing using panel data models have also been extensively studied, see Hsiao (1986) and Baltagi (2001), but these models ignore the spatial correlation. Heterogeneity across the cross-sectional units is usually modeled with an error component model. A Lagrange multiplier (LM) test for random effects was derived by Breusch and Pagan (1980), and an extensive Monte Carlo on testing in this error component model was performed by Baltagi et al. (1992). This paper extends the Breusch and Pagan LM test to the spatial error component model. First, a joint LM test is derived which simultaneously tests for the existence of spatial error correlation as well as random region effects. This LM test is based on the estimation of the model under the null hypothesis and its computation is simple requiring only least-squares residuals. This test is important, because ignoring spatial correlation and heterogeneity due to the random region effects will result in inefficient estimates and misleading inference. We emphasize the importance of using the one-sided version of this test since variance components cannot be negative. Next, two conditional LM tests are derived. One for the existence of spatial error correlation assuming the presence of random region effects, and the other for the existence of random region effects assuming the presence of spatial error correlation. These tests guard against misleading inference caused by (i) one-directional LM tests that ignore the presence of random region effects when testing for spatial error correlation, or (ii) one-directional LM

tests that ignore the presence of spatial correlation when testing for random region effects.

Section 2 revisits the spatial error component model considered in Anselin (1988) and provides the joint and conditional LM tests proposed in this paper. Only the final LM test statistics are given in the paper. Their derivations are relegated to the appendices. Section 3 compares the performance of these LM tests as well as the corresponding likelihood ratio LR tests using Monte Carlo experiments. Section 4 gives a summary and conclusion.

2. The model and test statistics

Consider the following panel data regression model, see Baltagi (2001):

$$y_{ti} = X_{ti}'\beta + u_{ti}, \quad i = 1,\ldots,N; \;\; t = 1,\ldots,T, \tag{2.1}$$

where y_{ti} is the observation on the ith region for the tth time period, X_{ti} denotes the $k \times 1$ vector of observations on the non-stochastic regressors and u_{ti} is the regression disturbance. In vector form, the disturbance vector of (2.1) is assumed to have random region effects as well as spatially autocorrelated residual disturbances, see Anselin (1988):

$$u_t = \mu + \varepsilon_t \tag{2.2}$$

with

$$\varepsilon_t = \lambda W \varepsilon_t + v_t, \tag{2.3}$$

where $u_t' = (u_{t1},\ldots,u_{tN})$, $\varepsilon_t' = (\varepsilon_{t1},\ldots,\varepsilon_{tN})$ and $\mu' = (\mu_1,\ldots,\mu_N)$ denote the vector of random region effects which are assumed to be $\mathrm{IIN}(0,\sigma_\mu^2)$. λ is the scalar spatial autoregressive coefficient with $|\lambda| < 1$. W is a known $N \times N$ spatial weight matrix whose diagonal elements are zero. W also satisfies the condition that $(I_N - \lambda W)$ is non-singular for all $|\lambda| < 1$. $v_t' = (v_{t1},\ldots,v_{tN})$, where v_{ti} is i.i.d. over i and t and is assumed to be $\mathrm{N}(0,\sigma_v^2)$. The $\{v_{ti}\}$ process is also independent of the process $\{\mu_i\}$. One can rewrite (2.3) as

$$\varepsilon_t = (I_N - \lambda W)^{-1} v_t = B^{-1} v_t, \tag{2.4}$$

where $B = I_N - \lambda W$ and I_N is an identity matrix of dimension N. The model (2.1) can be rewritten in matrix notation as

$$y = X\beta + u, \tag{2.5}$$

where y is now of dimension $NT \times 1$, X is $NT \times k$, β is $k \times 1$ and u is $NT \times 1$. The observations are ordered with t being the slow running index and i the fast running index, i.e., $y' = (y_{11},\ldots,y_{1N},\ldots,y_{T1},\ldots,y_{TN})$. X is assumed to be of full column rank and its elements are assumed to be asymptotically bounded in absolute value. Eq. (2.2) can be written in vector form as

$$u = (\iota_T \otimes I_N)\mu + (I_T \otimes B^{-1})v, \tag{2.6}$$

where $v' = (v_1', \ldots, v_T')$, $\imath_T$ is a vector of ones of dimension T, I_T is an identity matrix of dimension T and $\otimes$ denotes the Kronecker product. Under these assumptions, the variance–covariance matrix for u can be written as

$$\Omega_u = \sigma_\mu^2 (J_T \otimes I_N) + \sigma_v^2 (I_T \otimes (B'B)^{-1}), \tag{2.7}$$

where J_T is a matrix of ones of dimension T. This variance–covariance matrix can be rewritten as

$$\Omega_u = \sigma_v^2 [\bar{J}_T \otimes (T\phi I_N + (B'B)^{-1}) + E_T \otimes (B'B)^{-1}] = \sigma_v^2 \Sigma_u, \tag{2.8}$$

where $\phi = \sigma_\mu^2/\sigma_v^2$, $\bar{J}_T = J_T/T$, $E_T = I_T - \bar{J}_T$ and $\Sigma_u = [\bar{J}_T \otimes (T\phi I_N + (B'B)^{-1}) + E_T \otimes (B'B)^{-1}]$. Using results in Wansbeek and Kapteyn (1983), Σ_u^{-1} is given by

$$\Sigma_u^{-1} = \bar{J}_T \otimes (T\phi I_N + (B'B)^{-1})^{-1} + E_T \otimes B'B. \tag{2.9}$$

Also, $|\Sigma_u| = |T\phi I_N + (B'B)^{-1}| \cdot |(B'B)^{-1}|^{T-1}$. Under the assumption of normality, the log-likelihood function for this model was derived by Anselin (1988, p. 154) as

$$\begin{aligned} L &= -\frac{NT}{2} \ln 2\pi\sigma_v^2 - \frac{1}{2} \ln |\Sigma_u| - \frac{1}{2\sigma_v^2} u' \Sigma_u^{-1} u \\ &= -\frac{NT}{2} \ln 2\pi\sigma_v^2 - \frac{1}{2} \ln [|T\phi I_N + (B'B)^{-1}|] + \frac{(T-1)}{2} \ln |B'B| \\ &\quad - \frac{1}{2\sigma_v^2} u' \Sigma_u^{-1} u \end{aligned} \tag{2.10}$$

with $u = y - X\beta$. Anselin (1988, p. 154) derived the LM test for $\lambda = 0$ in this model. Here, we extend Anselin's work by deriving the joint test for spatial error correlation as well as random region effects.

The hypotheses under consideration in this paper are the following:

(a) $H_0^a : \lambda = \sigma_\mu^2 = 0$, and the alternative H_1^a is that at least one component is not zero.
(b) $H_0^b : \sigma_\mu^2 = 0$ (assuming no spatial correlation, i.e., $\lambda = 0$), and the one-sided alternative H_1^b is that $\sigma_\mu^2 > 0$ (assuming $\lambda = 0$).
(c) $H_0^c : \lambda = 0$ (assuming no random effects, i.e., $\sigma_\mu^2 = 0$), and the two-sided alternative is $H_1^c : \lambda \neq 0$ (assuming $\sigma_\mu^2 = 0$).
(d) $H_0^d : \lambda = 0$ (assuming the possible existence of random effects, i.e., $\sigma_\mu^2 \geqslant 0$), and the two-sided alternative is $H_1^d : \lambda \neq 0$ (assuming $\sigma_\mu^2 \geqslant 0$).
(e) $H_0^e : \sigma_\mu^2 = 0$ (assuming the possible existence of spatial correlation, i.e., λ may be zero or different from zero), and the one-sided alternative is $H_1^e : \sigma_\mu^2 > 0$ (assuming that λ may be zero or different from zero).

In the next sections, we derive the corresponding LM tests for these hypotheses and we compare their performance with the corresponding LR tests using Monte Carlo experiments.

2.1. Joint LM test for H_0^a: $\lambda = \sigma_\mu^2 = 0$

The joint LM test statistic for testing $H_0^a : \lambda = \sigma_\mu^2 = 0$ vs. H_1^a is given by

$$LM_J = \frac{NT}{2(T-1)} G^2 + \frac{N^2 T}{b} H^2, \tag{2.11}$$

where $G = \tilde{u}'(J_T \otimes I_N)\tilde{u}/\tilde{u}'\tilde{u} - 1$, $H = \tilde{u}'(I_T \otimes (W+W')/2)\tilde{u}/\tilde{u}'\tilde{u}$, $b = tr(W+W')^2/2 = tr(W^2 + W'W)$ and $\tilde{u}$ denotes the OLS residuals. The derivation of this LM test statistic is given in Appendix A.1. It is important to note that the large sample distribution of the LM test statistics derived in this paper are not formally established, but are likely to hold under similar sets of low level assumptions developed in Kelejian and Prucha (2001) for the Moran I test statistic and its close cousins the LM tests for spatial correlation. See also Pinkse (1998, 1999) for general conditions under which Moran flavoured tests for spatial correlation have a limiting normal distribution in the presence of nuisance parameters in six frequently encountered spatial models. Section 2.4 shows that the one-sided version of this joint LM test should be used because variance components cannot be negative.

2.2. Marginal LM test for H_0^b: $\sigma_\mu^2 = 0$ (assuming $\lambda = 0$)

Note that the first term in (2.11), call it $LM_G = (NT/2(T-1))G^2$, is the basis for the LM test statistic for testing H_0^b: $\sigma_\mu^2 = 0$ assuming there are no spatial error dependence effects, i.e., assuming that $\lambda = 0$, see Breusch and Pagan (1980). This LM statistic should be asymptotically distributed as χ_1^2 under H_0^b as $N \to \infty$, for a given T. But this LM test has the problem that the alternative hypothesis is assumed to be two-sided when we know that the variance component cannot be negative. Honda (1985) suggested a uniformly most powerful test for H_0^b based upon the square root of the G^2 term, i.e.,

$$LM_1 = \sqrt{\frac{NT}{2(T-1)}}\, G. \tag{2.12}$$

This should be asymptotically distributed as $\mathrm{N}(0,1)$ under H_0^b as $N \to \infty$, for T fixed. Moulton and Randolph (1989) showed that the asymptotic $\mathrm{N}(0,1)$ approximation for this one sided LM test can be poor even in large samples. This occurs when the number of regressors is large or the intra-class correlation of some of the regressors is high. They suggest an alternative standardized LM (SLM) test statistic whose asymptotic critical values are generally closer to the exact critical values than those of the LM test. This SLM test statistic centers and scales the one-sided LM statistic so that its mean is zero and its variance is one:

$$SLM_1 = \frac{LM_1 - E(LM_1)}{\sqrt{var(LM_1)}} = \frac{d_1 - E(d_1)}{\sqrt{var(d_1)}}, \tag{2.13}$$

where $d_1 = \tilde{u}' D_1 \tilde{u}/\tilde{u}'\tilde{u}$ and $D_1 = (J_T \otimes I_N)$ with $\tilde{u}$ denoting the OLS residuals. Using the normality assumption and results on moments of quadratic forms in regression residuals

(see e.g. Evans and King, 1985), we get

$$E(d_1) = tr(D_1M)/s, \tag{2.14}$$

where $s = NT - k$ and $M = I_{NT} - X(X'X)^{-1}X'$. Also.

$$var(d_1) = 2\{s\,tr(D_1M)^2 - [tr(D_1M)]^2\}/s^2(s+2). \tag{2.15}$$

Under H_0^b, SLM_1 should be asymptotically distributed as $\mathrm{N}(0,1)$.

2.3. Marginal LM test for H_0^c: $\lambda = 0$ (assuming $\sigma_\mu^2 = 0$)

Similarly, the second term in (2.11), call it $LM_H = (N^2T/b)H^2$, is the basis for the LM test statistic for testing H_0^c: λ=0 assuming there are no random regional effects, i.e., assuming that $\sigma_\mu^2 = 0$, see Anselin (1988). This LM statistic should be asymptotically distributed as χ_1^2 under H_0^c. Alternatively, this can be obtained as

$$LM_2 = \sqrt{\frac{N^2T}{b}}\,H. \tag{2.16}$$

This LM_2 test statistic should be asymptotically distributed as $\mathrm{N}(0,1)$ under H_0^c. The corresponding standardized LM (SLM) test statistic is given by

$$SLM_2 = \frac{LM_2 - E(LM_2)}{\sqrt{var(LM_2)}} = \frac{d_2 - E(d_2)}{\sqrt{var(d_2)}}, \tag{2.17}$$

where $d_2 = \tilde{u}'D_2\tilde{u}/\tilde{u}'\tilde{u}$ and $D_2 = (I_T \otimes W)$. Under H_0^c, SLM_2 should be asymptotically distributed as $\mathrm{N}(0,1)$. SLM_2 should have asymptotic critical values that are generally closer to the corresponding exact critical values than those of the unstandardized LM_2 test statistic.

2.4. One-sided joint LM test for H_0^a: $\lambda = \sigma_\mu^2 = 0$

Following Honda (1985) for the two-way error component model, a handy one-sided test statistic for H_0^a: $\lambda = \sigma_\mu^2 = 0$ is given by

$$LM^{\mathrm{H}} = (LM_1 + LM_2)/\sqrt{2}, \tag{2.18}$$

which is asymptotically distributed $\mathrm{N}(0,1)$ under H_0^a.

Note that LM_1 in (2.12) can be negative for a specific application, especially when the true variance component σ_μ^2 is small and close to zero. Similarly, LM_2 in (2.16) can be negative especially when the true λ is small and close to zero. Following Gourieroux et al. (1982), here after GHM, we propose the following test for the joint null hypothesis H_0^a:

$$\chi_m^2 = \begin{cases} LM_1^2 + LM_2^2 & \text{if } LM_1 > 0,\ LM_2 > 0, \\ LM_1^2 & \text{if } LM_1 > 0,\ LM_2 \leqslant 0, \\ LM_2^2 & \text{if } LM_1 \leqslant 0,\ LM_2 > 0, \\ 0 & \text{if } LM_1 \leqslant 0,\ LM_2 \leqslant 0. \end{cases} \tag{2.19}$$

Under the null hypothesis H_0^a, χ_m^2 has a mixed χ^2-distribution:

$$\chi_m^2 \sim (\tfrac{1}{4})\chi^2(0) + (\tfrac{1}{2})\chi^2(1) + (\tfrac{1}{4})\chi^2(2), \tag{2.20}$$

where $\chi^2(0)$ equals zero with probability one. The weights $(\frac{1}{4}),(\frac{1}{2})$ and $(\frac{1}{4})$ follow from the fact that LM_1 and LM_2 are asymptotically independent of each other and the results in Gourieroux et al. (1982). The critical values for the mixed χ_m^2 are 7.289, 4.321 and 2.952 for $\alpha = 0.01$, 0.05 and 0.1, respectively.

2.5. LR Test for H_0^a: $\lambda = \sigma_\mu^2 = 0$

We also compute the likelihood ratio (LR) test for H_0^a: $\lambda = \sigma_\mu^2 = 0$. Estimation of the unrestricted log-likelihood function is obtained using the method of scoring. The details of the estimation procedure are available upon request from the authors. Let $\hat{\sigma}_v^2$, $\hat{\phi}$, $\hat{\lambda}$ and $\hat{\beta}$ denote the unrestricted maximum likelihood estimators (MLEs) and let $\hat{B} = I_N - \hat{\lambda}W$ and $\hat{u} = y - X'\hat{\beta}$, then the unrestricted maximum log-likelihood estimator function is given by

$$\begin{aligned} L_U = &-\frac{NT}{2}\ln 2\pi\hat{\sigma}_v^2 - \frac{1}{2}\ln[|T\hat{\phi}I_N + (\hat{B}'\hat{B})^{-1}|] \\ &+ (T-1)\ln|\hat{B}| - \frac{1}{2\hat{\sigma}_v^2}\hat{u}'\hat{\Sigma}_u^{-1}\hat{u}, \end{aligned} \tag{2.21}$$

see Anselin (1988), where $\hat{\Sigma}$ is obtained from (2.8) with $\hat{B}$ replacing B and $\hat{\phi}$ replacing ϕ. But under the null hypothesis H_0^a, the variance–covariance matrix reduces to $\Omega_u^* = \Omega_u = \sigma_v^2 I_{TN}$ and the restricted MLE of β is $\tilde{\beta}_{\mathrm{OLS}}$, so that $\tilde{u} = y - X'\tilde{\beta}_{\mathrm{OLS}}$ are the OLS residuals and $\tilde{\sigma}_v^2 = \tilde{u}'\tilde{u}/NT$. Therefore, the restricted maximum log-likelihood function under H_0^a is given by

$$L_R = -\frac{NT}{2}\ln 2\pi\tilde{\sigma}_v^2 - \frac{1}{2\tilde{\sigma}_v^2}\tilde{u}'\tilde{u}. \tag{2.22}$$

Hence, the likelihood ratio test statistic for H_0^a: $\lambda = \sigma_\mu^2 = 0$ is given by

$$LR_J^* = 2(L_U - L_R) \tag{2.23}$$

and this should be asymptotically distributed as a mixture of χ^2 given in (2.20) under the null hypothesis.

2.6. Conditional LM test for H_0^d: $\lambda = 0$ (assuming $\sigma_\mu^2 \geqslant 0$)

When one uses LM_2, given by (2.16), to test H_0^c: $\lambda = 0$, one implicitly assumes that the random region effects do not exist. This may lead to incorrect decisions especially when σ_μ^2 is large. To overcome this problem, this section derives a conditional LM test for spatially uncorrelated disturbances assuming the possible existence of random regional effects. The null hypothesis for this model is H_0^d: $\lambda = 0$ (assuming $\sigma_\mu^2 \geqslant 0$).

Under the null hypothesis, the variance–covariance matrix reduces to $\Omega_0 = \sigma_\mu^2 J_T \otimes I_N + \sigma_\nu^2 I_{NT}$. It is the familiar form of the one-way error component model, see Baltagi (1995), with $\Omega_0^{-1} = (\sigma_1^2)^{-1}(\bar{J}_T \otimes I_N) + (\sigma_\nu^2)^{-1}(E_T \otimes I_N)$, where $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\nu^2$, and $E_T = I_T - \bar{J}_T$. Using derivations analogous to those for the joint LM-test, see Appendix A.1, we obtain the following LM test for H_0^d vs. H_1^d.

$$LM_\lambda = \frac{\hat{D}(\lambda)^2}{[(T-1) + \hat{\sigma}_\nu^4/\hat{\sigma}_1^4]b}, \tag{2.24}$$

where

$$\hat{D}(\lambda) = \frac{1}{2}\hat{u}'\left[\frac{\hat{\sigma}_\nu^2}{\hat{\sigma}_1^4}(\bar{J}_T \otimes (W' + W)) + \frac{1}{\hat{\sigma}_\nu^2}(E_T \otimes (W' + W))\right]\hat{u}.$$

Here, $\hat{\sigma}_\nu^2 = \hat{u}'(E_T \otimes I_N)\hat{u}/N(T-1)$ and $\hat{\sigma}_1^2 = \hat{u}'(\bar{J}_T \otimes I_N)\hat{u}/N$ are the MLEs of σ_ν^2 and σ_1^2 under H_0^d, and $\hat{u}$ denotes the maximum likelihood residuals under the null hypothesis H_0^d. See Appendix A.2 for more details.

Therefore, the one-sided test for zero spatial error dependence (assuming $\sigma_\mu^2 \geqslant 0$) against an alternative, say of $\lambda > 0$ is obtained from

$$LM_\lambda^* = \frac{\hat{D}(\lambda)}{\sqrt{[(T-1) + \hat{\sigma}_\nu^4/\hat{\sigma}_1^4]b}} \tag{2.25}$$

and this test statistic should be asymptotically distributed as $\mathrm{N}(0,1)$ under H_0^d for $N \to \infty$ and T fixed.

We can also get the LR test for H_0^d, using the scoring method. Details are available upon request from the authors. Under the null hypothesis, and a similar set of low level assumptions developed in Kelejian and Prucha (2001) the conditional LR test statistic will have the same asymptotic distribution as its LM counterpart.

2.7. Conditional LM test for H_0^e: $\sigma_\mu^2 = 0$ (assuming λ may or may not be $=0$)

Similarly, if one uses LM_1, given by (2.12), to test $H_0^b : \sigma_\mu^2 = 0$, one is implicitly assuming that no spatial error correlation exists. This may lead to incorrect decisions especially when λ is significantly different from zero. To overcome this problem, this section derives a conditional LM test for no random regional effects assuming the possible existence of spatial error correlation. The null hypothesis for this model is $H_0^e : \sigma_\mu^2 = 0$ (assuming λ may or may not be $=0$).

This LM test statistic is derived in Appendix A.3 and is given by

$$LM_\mu = \hat{D}_\mu' \hat{J}_\theta^{-1} \hat{D}_\mu, \tag{2.26}$$

where

$$\hat{D}_\mu = -\frac{T}{2\hat{\sigma}_\nu^2} tr(\hat{B}'\hat{B}) + \frac{1}{2\hat{\sigma}_\nu^4}\hat{u}'[J_T \otimes (\hat{B}'\hat{B})^2]\hat{u} \tag{2.27}$$

and

$$\hat{J}_\theta = \begin{bmatrix} \frac{TN}{2\hat{\sigma}_v^2} & \frac{NT}{2\hat{\sigma}_v^2} tr[(W'\hat{B} + \hat{B}'W) + (\hat{B}'\hat{B})^{-1}] & \frac{T}{2\hat{\sigma}_v^4} tr[\hat{B}'\hat{B}] \\ & \frac{T}{2} tr[((W'\hat{B} + \hat{B}'W) + (\hat{B}'\hat{B})^{-1})^2] & \frac{T}{2\hat{\sigma}_v^2} tr[W'\hat{B} + \hat{B}'W] \\ & & \frac{T^2}{2\hat{\sigma}_v^4} tr[(\hat{B}'\hat{B})^2] \end{bmatrix}, \quad (2.28)$$

$$= \frac{T}{2\hat{\sigma}_v^4} \begin{bmatrix} N & \hat{\sigma}_v^2 g & h \\ \hat{\sigma}_v^2 g & \hat{\sigma}_v^4 c & \hat{\sigma}_v^2 d \\ h & \hat{\sigma}_v^2 d & Te \end{bmatrix}, \quad (2.29)$$

where $g = tr[(W'\hat{B} + \hat{B}'W)(\hat{B}'\hat{B})^{-1}]$, $h = tr[\hat{B}'\hat{B}]$, $c = tr[((W'\hat{B} + \hat{B}'W)(\hat{B}'\hat{B})^{-1})^2]$, $d = tr\,[W'\hat{B} + \hat{B}'W]$ and $e = tr[(\hat{B}'\hat{B})^2]$. Therefore,

$$LM_\mu = (\hat{D}_\mu)^2 \left(\frac{2\hat{\sigma}_v^4}{T}\right) (TN\hat{\sigma}_v^4 ec - N\hat{\sigma}_v^4 d^2 - T\hat{\sigma}_v^4 g^2 e + 2\hat{\sigma}_v^4 ghd - \hat{\sigma}_v^4 h^2 c)^{-1}$$
$$\times (N\hat{\sigma}_v^4 c - \hat{\sigma}_v^4 g^2), \quad (2.30)$$

where $\hat{D}_\mu$ and $\hat{J}_\theta$ are evaluated at the MLEs under the null hypothesis H_0^e. However, LM_μ ignores the fact that the variance component cannot be negative. Therefore, the one-sided version of this LM test is given by

$$LM_\mu^* = \frac{\hat{D}_\mu \sqrt{(2\hat{\sigma}_v^4/T)(N\hat{\sigma}_v^4 c - \hat{\sigma}_v^4 g^2)}}{\sqrt{TN\hat{\sigma}_v^4 ec - N\hat{\sigma}_v^4 d^2 - T\hat{\sigma}_v^4 g^2 e + 2\hat{\sigma}_v^4 ghd - \hat{\sigma}_v^4 h^2 c}} \quad (2.31)$$

and this should be asymptotically distributed as $\mathrm{N}(0,1)$ under H_0^e as $N \to \infty$ for T fixed.

3. Monte Carlo results

The experimental design for the Monte Carlo simulations is based on the format extensively used in earlier studies in the spatial regression model by Anselin and Rey (1991) and Anselin and Florax (1995) and in the panel data model by Nerlove (1971).

The model is set as follows:

$$y_{it} = \alpha + x_{it}'\beta + u_{it}, \qquad i = 1, \ldots N, \; t = 1, \ldots, T, \quad (3.1)$$

where $\alpha = 5$ and $\beta = 0.5$. x_{it} is generated by a similar method to that of Nerlove (1971). In fact, $x_{it} = 0.1t + 0.5x_{i,t-1} + z_{it}$, where z_{it} is uniformly distributed over the interval $[-0.5, 0.5]$. The initial values x_{i0} are chosen as $(5 + 10z_{i0})$. For the disturbances, $u_{it} = \mu_i + \varepsilon_{it}$, $\varepsilon_{it} = \lambda \sum_{j=1}^N w_{ij}\varepsilon_{it} + v_{it}$ with $\mu_i \sim \mathrm{IIN}(0, \sigma_\mu^2)$ and $v_{it} \sim \mathrm{IIN}(0, \sigma_v^2)$. The matrix W is either a rook or a queen type weight matrix, and the rows of this matrix are standardized so that they sum to one.[1] We fix $\sigma_\mu^2 + \sigma_v^2 = 20$ and let $\rho = \sigma_\mu^2/(\sigma_\mu^2 + \sigma_v^2)$

[1] The weight matrix with first-order contiguity according to the rook criterion has the cells immediately above, below, to the right, and to the left, for a total of four neighboring cells. The weight matrix with first order contiguity according to the queen criterion is eight cells immediately surrounding the central cell, see Anselin and Rey (1991).

vary over the set (0, 0.2, 0.5, 0.8). The spatial autocorrelation factor λ is varied over a positive range from 0 to 0.9 by increments of 0.1. Two values for $N=25$ and 49, and two values for $T=3$ and 7 are chosen. In total, this amounts to 320 experiments.[2] For each experiment, the joint, conditional and marginal LM and LR tests are computed and 2000 replications are performed. In a first draft of this paper we reported the two-sided LM and LR test results to show how misleading the results of these tests can be. These results are available upon request from the authors. In this version, we focus on the one-sided version of these tests except for testing H_0^b: $\sigma_\mu^2=0$ where we thought a warning should be given to applied econometricians using packages that still report two-sided versions of this test.

3.1. Joint tests for H_0^a: $\lambda=\sigma_\mu^2=0$

Table 1 gives the frequency of rejections at the 5% level for the handy one-sided Honda-type LM test statistic LM^{H} given in (2.18), the GHM test statistic given in (2.19) and LR_J^* given in (2.23). The results are reported for $N=25,49$ and $T=3,7$ for both the queen and rook weight matrices based on 2000 replications. Table 1 shows that at the 5% level, the size of the joint LR test (LR_J^*) is not significantly different from 0.05 for all values of N and T and choice of the weight matrix W. The same is true for LM^{H} and GHM except for $N=25$ and $T=3$ where they are undersized. The power of all three tests is reasonably high as long as $\lambda>0.3$ or $\rho>0.2$. In fact, for $\rho\geqslant 0.5$ this power is almost one in all cases. For a fixed λ or ρ, this power improves as N or T increase.

3.2. Marginal and conditional tests for $\lambda=0$

Fig. 1 plots the frequency of rejections in 2000 replications for testing $\lambda=0$, i.e., zero spatial error correlation. Fig. 1 reports these frequencies for various values of $N=25,49$ and $T=3,7$, for both rook and queen weight matrices. Marginal tests for H_0^c: $\lambda=0$ (assuming $\sigma_\mu^2=0$) as well as conditional tests for H_0^d: $\lambda=0$ (assuming $\sigma_\mu^2\geqslant 0$) are plotted for various values of λ. As clear from the graphs, marginal tests can have misleading size when ρ is large (0.5 or 0.8). Marginal tests also have lower power than conditional tests for $\rho>0.2$ and $0.2\leqslant\lambda\leqslant 0.8$. This is true whether we use LM or LR type tests. This difference in power is quite substantial for example when $\rho=0.8$ and $\lambda=0.6$. This phenomena persists even when we increase N or T. However, it is important to note that marginal tests still detect that something is wrong when ρ is large.

3.3. Marginal and conditional tests for $\sigma_\mu^2=0$

Table 2 gives the frequency of rejections in 2000 replications for the marginal LR and LM tests for H_0^b: $\sigma_\mu^2=0$ (assuming $\lambda=0$). The results are reported only when

[2] The Monte Carlo experiments were also run for negative λ ranging between -0.1 and -0.9. The results were similar and are not reproduced here to save space.

Table 1
Joint tests for H_0^a, $\lambda = \sigma_\mu^2 = 0$

N, T	λ	$\rho = 0.0$			$\rho = 0.2$			$\rho = 0.5$		
		LM^{H}	GHM	LR_J^*	LM^{H}	GHM	LR_J^*	LM^{H}	GHM	LR_J^*
Weight matrix is rook										
25, 3	0.0	0.021	0.022	0.048	0.241	0.351	0.379	0.825	0.961	0.963
	0.1	0.073	0.074	0.089	0.359	0.386	0.405	0.875	0.967	0.970
	0.2	0.154	0.198	0.219	0.474	0.432	0.473	0.920	0.969	0.973
	0.3	0.277	0.414	0.435	0.605	0.566	0.633	0.944	0.970	0.980
	0.4	0.469	0.688	0.728	0.727	0.711	0.788	0.963	0.976	0.989
	0.5	0.634	0.876	0.900	0.845	0.844	0.909	0.983	0.980	0.996
	0.6	0.832	0.967	0.976	0.934	0.953	0.978	0.989	0.988	0.998
	0.7	0.950	0.999	1.000	0.973	0.984	0.996	0.997	0.996	0.999
	0.8	0.987	1.000	1.000	0.995	0.999	1.000	0.999	0.997	1.000
	0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
25, 7	0.0	0.038	0.039	0.061	0.805	0.895	0.890	1.000	1.000	1.000
	0.1	0.115	0.144	0.151	0.916	0.940	0.916	1.000	1.000	1.000
	0.2	0.299	0.450	0.463	0.962	0.949	0.947	1.000	1.000	1.000
	0.3	0.611	0.827	0.836	0.986	0.984	0.986	1.000	1.000	1.000
	0.4	0.866	0.983	0.984	0.997	0.997	0.997	1.000	1.000	1.000
	0.5	0.975	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
49, 3	0.0	0.035	0.041	0.062	0.440	0.608	0.613	0.979	0.999	0.999
	0.1	0.120	0.133	0.135	0.613	0.638	0.633	0.996	1.000	1.000
	0.2	0.289	0.384	0.403	0.786	0.750	0.767	0.999	1.000	1.000
	0.3	0.513	0.724	0.738	0.902	0.860	0.895	1.000	1.000	1.000
	0.4	0.782	0.948	0.956	0.965	0.946	0.969	1.000	1.000	1.000
	0.5	0.920	0.992	0.992	0.989	0.994	0.997	1.000	1.000	1.000
	0.6	0.993	1.000	1.000	1.000	0.999	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	1.000
49, 7	0.0	0.033	0.040	0.050	0.977	0.996	0.994	1.000	1.000	1.000
	0.1	0.188	0.259	0.263	0.994	0.997	0.996	1.000	1.000	1.000
	0.2	0.563	0.779	0.775	0.999	0.999	1.000	1.000	1.000	1.000
	0.3	0.895	0.984	0.986	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Weight matrix is queen										
25, 3	0.0	0.020	0.029	0.064	0.216	0.344	0.372	0.822	0.955	0.958
	0.1	0.059	0.067	0.066	0.301	0.369	0.380	0.864	0.962	0.967
	0.2	0.112	0.151	0.140	0.375	0.402	0.411	0.902	0.967	0.968
	0.3	0.229	0.299	0.282	0.510	0.505	0.519	0.932	0.977	0.981
	0.4	0.379	0.536	0.514	0.638	0.608	0.650	0.954	0.973	0.984
	0.5	0.576	0.748	0.743	0.786	0.764	0.806	0.970	0.983	0.992
	0.6	0.771	0.895	0.891	0.903	0.899	0.923	0.985	0.990	0.995
	0.7	0.916	0.975	0.974	0.970	0.973	0.980	0.997	0.993	1.000
	0.8	0.972	0.997	0.997	0.992	0.992	0.996	0.999	0.999	0.999
	0.9	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 1 (*continued*)

N, T	λ	$\rho = 0.0$			$\rho = 0.2$			$\rho = 0.5$		
		LM^{H}	GHM	LR_J^*	LM^{H}	GHM	LR_J^*	LM^{H}	GHM	LR_J^*
25, 7	0.0	0.040	0.042	0.056	0.782	0.893	0.876	1.000	1.000	1.000
	0.1	0.109	0.131	0.117	0.870	0.920	0.904	1.000	1.000	1.000
	0.2	0.226	0.337	0.317	0.937	0.936	0.932	1.000	1.000	1.000
	0.3	0.493	0.679	0.646	0.964	0.956	0.951	1.000	1.000	1.000
	0.4	0.772	0.893	0.883	0.995	0.990	0.989	1.000	1.000	1.000
	0.5	0.937	0.987	0.987	0.999	0.999	0.999	1.000	1.000	1.000
	0.6	0.997	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000
49, 3	0.0	0.029	0.032	0.057	0.406	0.6070	0.622	0.990	1.000	1.000
	0.1	0.094	0.108	0.100	0.552	0.6380	0.636	0.992	1.000	1.000
	0.2	0.212	0.259	0.232	0.719	0.7140	0.716	0.998	1.000	1.000
	0.3	0.392	0.546	0.517	0.811	0.7860	0.804	0.999	1.000	1.000
	0.4	0.641	0.797	0.782	0.915	0.8890	0.901	0.999	1.000	1.000
	0.5	0.862	0.959	0.953	0.972	0.9650	0.973	1.000	1.000	1.000
	0.6	0.968	0.991	0.993	0.993	0.9940	0.995	1.000	1.000	1.000
	0.7	1.000	1.000	1.000	1.000	0.9990	1.000	1.000	1.000	1.000
49, 7	0.0	0.052	0.056	0.056	0.961	0.994	0.995	1.000	1.000	1.000
	0.1	0.156	0.167	0.153	0.989	0.993	0.995	1.000	1.000	1.000
	0.2	0.419	0.566	0.536	0.998	0.998	0.996	1.000	1.000	1.000
	0.3	0.775	0.908	0.899	1.000	1.000	0.999	1.000	1.000	1.000
	0.4	0.961	0.995	0.994	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	0.992	1.000	1.000	1.000	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	1.000

Frequency of rejections in 2000 replications.

$\sigma_\mu^2 = 0$ for $N = 25, 49$ and $T = 3, 7$ for both the queen and rook weight matrices. Table 2 shows that at the 5% level, the size of the two-sided LM test (LM_G) for H_0^b (compared to its one sided counterpart LM_1) could be missleading, especially when λ is large. For example, for the queen weight matrix when $N = 49$, $T = 7$ and $\lambda = 0.9$, the frequency of rejection for LM_G is 50.4% whereas the corresponding one-sided LM (LM_1) has a size of 7.6%. The two-sided likelihood ratio (LR_G) test for H_0^b performs better than its two-sided LM counterpart (LM_G). However, in most experiments, LR_G underestimates its size and is outperformed by its one-sided LR alternative (LR_1).

Table 2 also gives the frequency of rejections in 2000 replications for the conditional LR and LM tests (LR_μ^* and LM_μ^*) for H_0^e: $\sigma_\mu^2 = 0$ (assuming $\lambda \neq 0$). These were derived in Section 2.2. The results are reported only when $\sigma_\mu^2 = 0$ for $N = 25, 49$ and $T = 3, 7$ for both the queen and rook weight matrices. For most experiments, the conditional LM and LR tests have size not significantly different from 5%. For cases where λ is large, conditional tests have better size than marginal tests. For example, when the weight matrix is queen, $N = 49, T = 3$ and $\lambda = 0.9$, the frequency of rejections at the 5% significance level, when the null is true, is 11.4% and 12% for LM_1 and LR_1 compared to 4.9% and 4.3% for LM_μ^* and LR_μ^*.

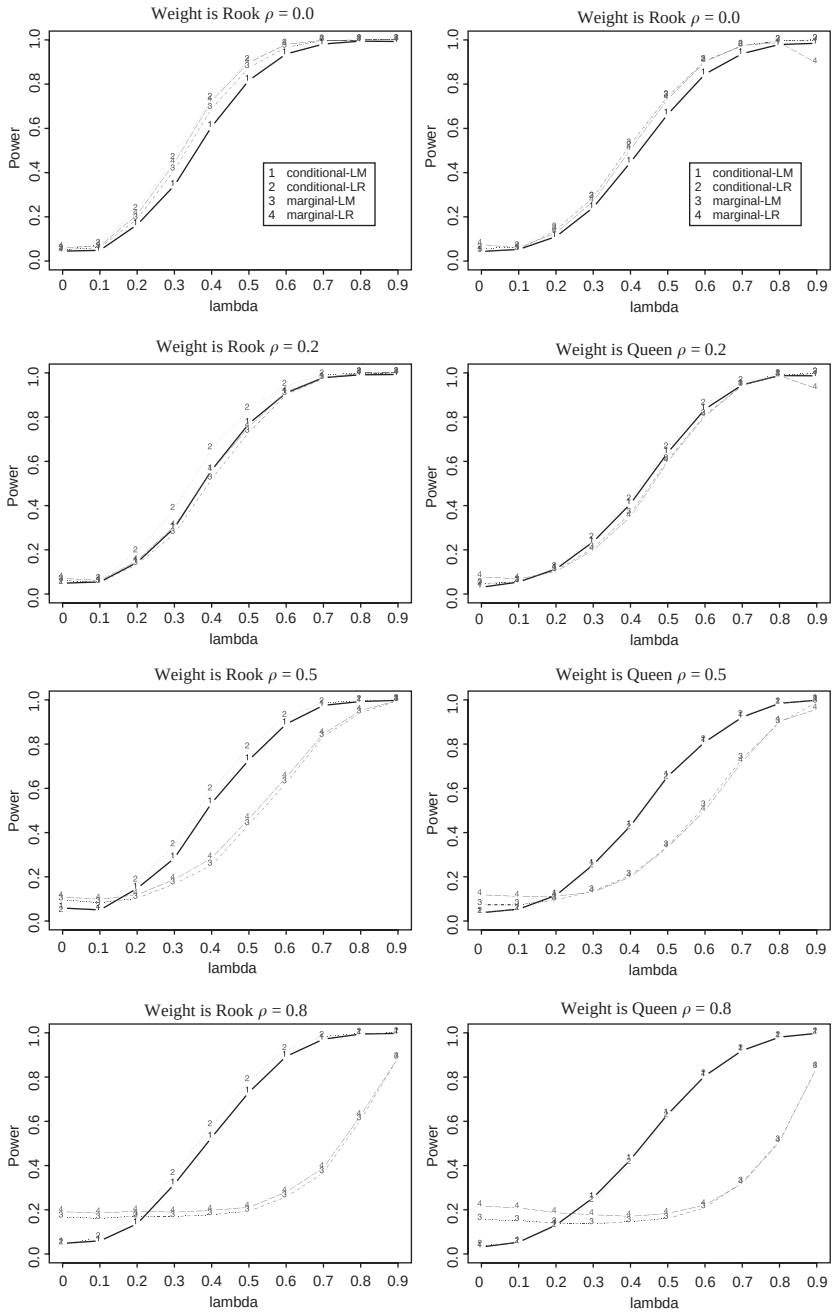

Fig. 1. Frequency of rejections in 2000 replications: (A) $H_0^c = \lambda = 0$ (assuming $\sigma_\mu^2 = 0$) and H_0^d; $\lambda = 0$ (assuming $\sigma_\mu^2 > 0$ $N = 25$, $T = 3$.

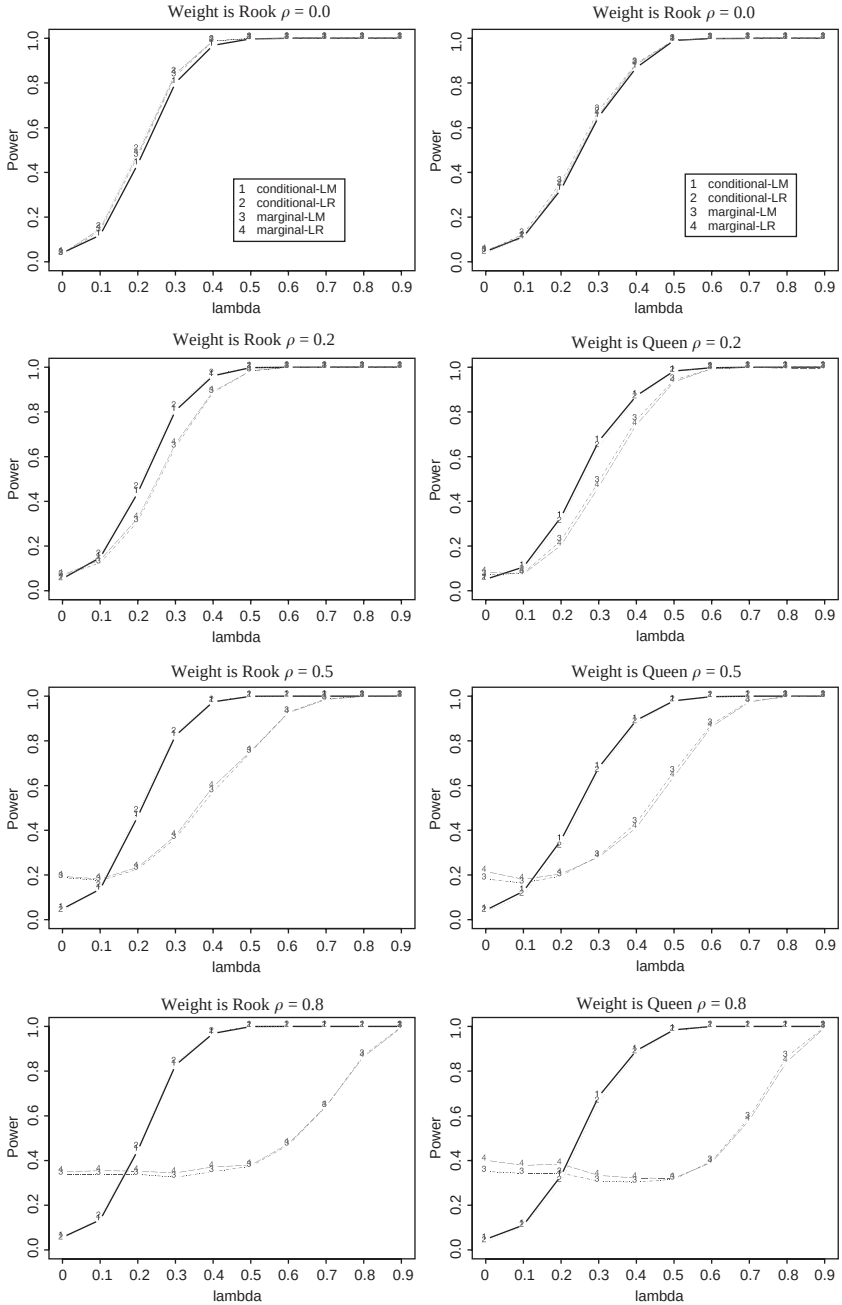

Fig. 1. (B) H_0^c; $\lambda = 0$ (assuming $\sigma_\mu^2 = 0$) and H_0^d; $\lambda = 0$ (assuming $\sigma_\mu^2 > 0$) $N = 25$, $T = 7$.

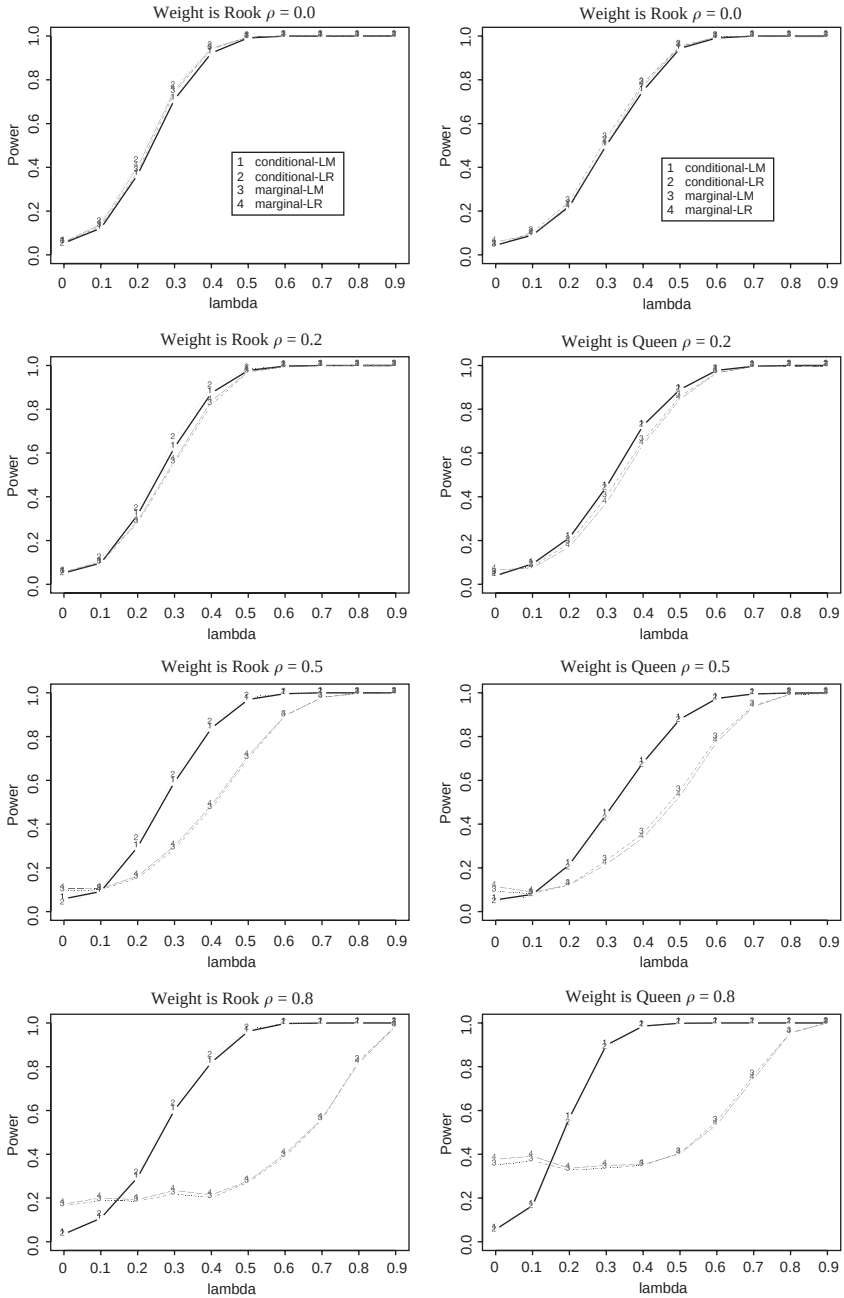

Fig. 1. (C) H_0^c; $\lambda = 0$ (assuming $\sigma_\mu^2 = 0$) and H_0^d; $\lambda = 0$ (assuming $\sigma_\mu^2 > 0$) $N = 49$, $T = 3$.

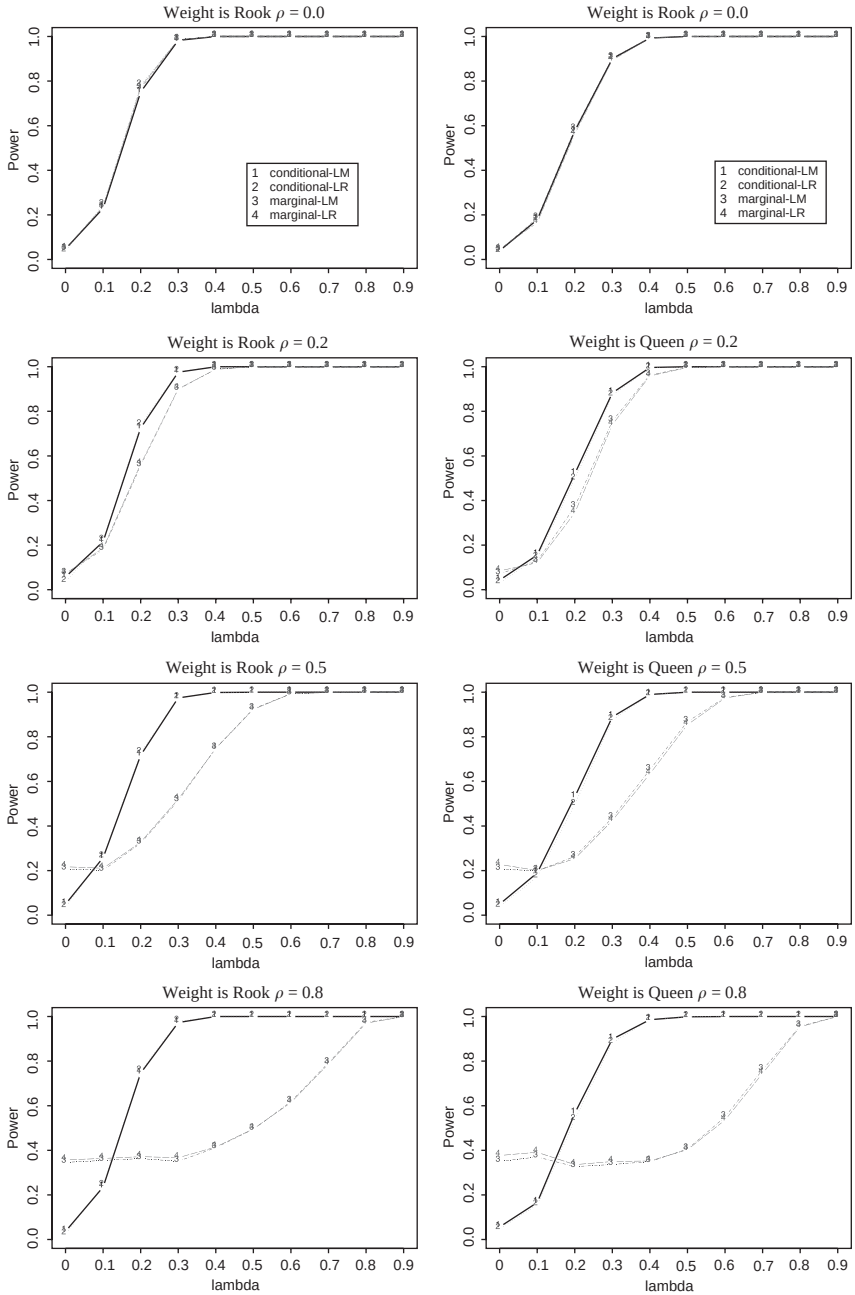

Fig. 1. (D) H_0^c; $\lambda = 0$ (assuming $\sigma_\mu^2 = 0$) and H_0^d; $\lambda = 0$ (assuming $\sigma_\mu^2 > 0$) $N = 49$, $T = 7$.

Table 2
Tests for $\sigma_\mu^2 = 0$

N, T	W	λ	H_0^b, $\sigma_\mu^2 = 0$ (assuming $\lambda = 0$)					H_0^e, $\sigma_\mu^2 = 0$ (assuming $\lambda \neq 0$)	
			Two-sided		One-sided			One-sided	
			LM_G	LR_G	LM_1	SLM_1	LR_1	LM_μ^*	LR_μ^*
25, 3	Rook	0.0	0.037	0.014	0.038	0.038	0.045	0.036	0.027
		0.1	0.052	0.020	0.044	0.047	0.050	0.043	0.041
		0.2	0.044	0.015	0.039	0.045	0.045	0.040	0.039
		0.3	0.055	0.023	0.044	0.049	0.047	0.040	0.036
		0.4	0.058	0.023	0.046	0.049	0.054	0.031	0.039
		0.5	0.069	0.017	0.043	0.046	0.048	0.041	0.038
		0.6	0.100	0.031	0.060	0.067	0.067	0.043	0.038
		0.7	0.143	0.043	0.067	0.070	0.075	0.043	0.035
		0.8	0.215	0.050	0.074	0.080	0.083	0.040	0.034
		0.9	0.341	0.055	0.068	0.074	0.072	0.036	0.035
25, 3	Queen	0.0	0.038	0.016	0.037	0.043	0.042	0.039	0.033
		0.1	0.045	0.020	0.042	0.046	0.050	0.036	0.041
		0.2	0.044	0.020	0.044	0.045	0.053	0.045	0.041
		0.3	0.058	0.020	0.043	0.043	0.051	0.039	0.040
		0.4	0.054	0.017	0.038	0.041	0.043	0.030	0.030
		0.5	0.083	0.033	0.054	0.062	0.062	0.038	0.042
		0.6	0.089	0.032	0.055	0.060	0.058	0.041	0.034
		0.7	0.154	0.041	0.068	0.073	0.072	0.044	0.030
		0.8	0.239	0.038	0.057	0.066	0.063	0.042	0.049
		0.9	0.388	0.033	0.039	0.045	0.051	0.031	0.035
25, 7	Rook	0.0	0.053	0.020	0.052	0.058	0.059	0.043	0.042
		0.1	0.045	0.018	0.045	0.047	0.053	0.039	0.030
		0.2	0.033	0.011	0.035	0.039	0.033	0.051	0.032
		0.3	0.051	0.018	0.049	0.050	0.051	0.045	0.044
		0.4	0.063	0.022	0.046	0.046	0.047	0.051	0.036
		0.5	0.062	0.018	0.044	0.044	0.046	0.052	0.039
		0.6	0.100	0.023	0.051	0.050	0.055	0.066	0.046
		0.7	0.168	0.035	0.056	0.059	0.058	0.050	0.036
		0.8	0.257	0.032	0.050	0.050	0.050	0.061	0.049
		0.9	0.510	0.017	0.024	0.024	0.023	0.050	0.035
25, 7	Queen	0.0	0.059	0.022	0.059	0.057	0.058	0.052	0.044
		0.1	0.043	0.014	0.038	0.041	0.043	0.039	0.031
		0.2	0.036	0.014	0.041	0.039	0.039	0.053	0.037
		0.3	0.050	0.012	0.045	0.044	0.049	0.055	0.041
		0.4	0.050	0.014	0.040	0.044	0.040	0.061	0.046
		0.5	0.069	0.025	0.050	0.054	0.049	0.053	0.039
		0.6	0.089	0.023	0.036	0.036	0.041	0.042	0.030
		0.7	0.151	0.022	0.040	0.038	0.041	0.049	0.045
		0.8	0.307	0.015	0.030	0.030	0.030	0.052	0.051
		0.9	0.589	0.008	0.014	0.014	0.014	0.042	0.057
49, 3	Rook	0.0	0.053	0.024	0.336	0.054	0.059	0.045	0.046
		0.1	0.041	0.017	0.049	0.050	0.052	0.048	0.039

Table 2 (*continued*)

N, T	W	λ	H_0^b, $\sigma_\mu^2 = 0$ (assuming $\lambda = 0$)					H_0^e, $\sigma_\mu^2 = 0$ (assuming $\lambda \neq 0$)	
			Two-sided		One-sided			One-sided	
			LM_G	LR_G	LM_1	SLM_1	LR_1	LM_μ^*	LR_μ^*
		0.2	0.065	0.024	0.048	0.048	0.051	0.041	0.037
		0.3	0.061	0.025	0.050	0.051	0.055	0.052	0.053
		0.4	0.085	0.031	0.056	0.056	0.060	0.042	0.036
		0.5	0.104	0.037	0.065	0.067	0.069	0.048	0.035
		0.6	0.118	0.046	0.074	0.080	0.078	0.042	0.038
		0.7	0.162	0.050	0.075	0.078	0.078	0.040	0.032
		0.8	0.260	0.067	0.107	0.111	0.110	0.050	0.027
		0.9	0.422	0.096	0.123	0.129	0.130	0.042	0.040
49, 3	Queen	0.0	0.048	0.021	0.046	0.048	0.049	0.044	0.043
		0.1	0.045	0.018	0.046	0.047	0.050	0.045	0.040
		0.2	0.062	0.031	0.059	0.066	0.063	0.054	0.059
		0.3	0.055	0.023	0.039	0.045	0.044	0.046	0.039
		0.4	0.086	0.036	0.062	0.065	0.064	0.051	0.047
		0.5	0.085	0.026	0.051	0.053	0.053	0.037	0.038
		0.6	0.126	0.038	0.068	0.072	0.076	0.047	0.044
		0.7	0.168	0.051	0.087	0.093	0.088	0.046	0.043
		0.8	0.282	0.074	0.102	0.105	0.103	0.045	0.051
		0.9	0.463	0.098	0.114	0.121	0.120	0.049	0.043
49, 7	Rook	0.0	0.057	0.023	0.062	0.060	0.061	0.053	0.044
		0.1	0.046	0.016	0.045	0.039	0.042	0.046	0.033
		0.2	0.053	0.022	0.055	0.052	0.057	0.053	0.039
		0.3	0.067	0.019	0.048	0.049	0.052	0.054	0.044
		0.4	0.067	0.030	0.064	0.063	0.067	0.048	0.043
		0.5	0.077	0.021	0.060	0.058	0.058	0.054	0.042
		0.6	0.129	0.036	0.067	0.066	0.067	0.051	0.036
		0.7	0.190	0.046	0.086	0.088	0.083	0.061	0.036
		0.8	0.280	0.062	0.097	0.096	0.093	0.043	0.025
		0.9	0.504	0.055	0.076	0.081	0.077	0.054	0.033
49, 7	Queen	0.0	0.043	0.017	0.043	0.046	0.048	0.047	0.034
		0.1	0.057	0.017	0.054	0.057	0.061	0.058	0.045
		0.2	0.042	0.017	0.048	0.051	0.050	0.060	0.043
		0.3	0.053	0.022	0.054	0.056	0.057	0.048	0.038
		0.4	0.061	0.024	0.048	0.048	0.046	0.046	0.035
		0.5	0.084	0.023	0.065	0.061	0.066	0.056	0.043
		0.6	0.117	0.037	0.071	0.068	0.069	0.060	0.048
		0.7	0.195	0.043	0.071	0.073	0.070	0.059	0.046
		0.8	0.313	0.046	0.075	0.073	0.069	0.056	0.056
		0.9	0.574	0.034	0.050	0.050	0.049	0.041	0.045

Frequency of rejections in 2000 replications, $\rho = 0.0$.

4. Conclusion

It is clear from the extensive Monte Carlo experiments performed that the spatial econometrics literature should not ignore the heterogeneity across cross-sectional units when testing for the presence of spatial error correlation. Similarly, the panel data econometrics literature should not ignore the spatial error correlation when testing for the presence of random regional effects. Both joint and conditional LM tests have been derived in this paper that are easy to implement and that perform better in terms of size and power than the one-directional LM tests. The latter tests ignore the random regional effects when testing for spatial error correlation or ignore spatial error correlation when testing for random regional effects. We emphasize the importance of using the one-sided version of these tests since variance components cannot be negative. This paper does not consider testing for spatial lag dependence and random regional effects in a panel. This should be the subject of future research. Also, the results in the paper should be tempered by the fact that the $N = 25, 49$ used in our Monte Carlo experiments may be small for a typical micro panel. Larger N will probably improve the performance of these tests whose critical values are based on their large sample distributions. However, it will also increase the computation difficulty and accuracy of the eigenvalues of the big weighting matrix W. Finally, it is important to point out that the asymptotic distribution of our test statistics were not explicitly derived in the paper but that they are likely to hold under a similar set of low level assumptions developed by Kelejian and Prucha (2001).

Acknowledgements

We would like to thank the associate editor and two referees for helpful comments. An earlier version of this paper was given at the North American Summer Meeting of the Econometric Society held at the University of Maryland, June, 2001. Also at UC-Berkeley, UC-San Diego and University of Southern California. Baltagi would like to thank the Bush School Program in the Economics of Public Policy. Also, we would like to thank the Korea Research Foundation Grant (KRF-2002–042-C00008) for its financial support.

Appendix A

A.1. Joint LM test

This appendix derives the joint LM test for spatial error correlation and random regional effects. The null hypothesis is given by H_0^a: $\sigma_\mu^2 = \lambda = 0$. Let $\theta = (\sigma_\nu^2, \sigma_\mu^2, \lambda)'$. Note that the part of the information matrix corresponding to β will be ignored in computing the LM 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

MLE will be zero. The LM statistic is given by

$$LM = \tilde{D}'_\theta \tilde{J}_\theta^{-1} \tilde{D}_\theta, \tag{A.1}$$

where $\tilde{D}_\theta = (\partial L/\partial\theta)(\tilde{\theta})$ is a 3×1 vector of partial derivatives with respect to each element of θ, evaluated at the restricted MLE $\tilde{\theta}$. Also, $\tilde{J} = E[-\partial^2 L/\partial\theta\partial\theta'](\tilde{\theta})$ is the information matrix corresponding to θ, evaluated at the restricted MLE $\tilde{\theta}$. Under the null hypothesis, the variance–covariance matrix reduces to $\sigma_\nu^2 I_{TN}$ and the restricted MLE of β is $\tilde{\beta}_{\mathrm{OLS}}$, so that $\tilde{u} = y - X'\tilde{\beta}_{\mathrm{OLS}}$ are the OLS residuals and $\tilde{\sigma}_\nu^2 = \tilde{u}'\tilde{u}/NT$.

Hartley and Rao (1967) or Hemmerle and Hartley(1973) give a useful general formula to obtain $\tilde{D}_\theta$:

$$\partial L/\partial\theta_r = -\tfrac{1}{2}\, tr[\Omega_u^{-1}(\partial\Omega_u/\partial\theta_r)] + \tfrac{1}{2}[u'\Omega_u^{-1}(\partial\Omega_u/\partial\theta_r)\Omega_u^{-1}u] \tag{A.2}$$

for $r = 1, 2, 3$. It is easy to check that $\partial\Omega_u/\partial\sigma_\nu^2 = I_T \otimes (B'B)^{-1}$, $\partial\Omega_u/\partial\sigma_\mu^2 = J_T \otimes I_N$ and $\partial\Omega_u/\partial\lambda = \sigma_\nu^2[I_T \otimes (B'B)^{-1}(W'B + B'W)(B'B)^{-1}]$ using the fact that $\partial(B'B)^{-1}/\partial\lambda = (B'B)^{-1}(W'B + B'W)(B'B)^{-1}$, see Anselin (1988, p. 164).

Under H_0^a, we get

$$\Omega_u^{-1}|_{H_0^a} = \frac{1}{\sigma_\nu^2} I_T \otimes I_N, \tag{A.3}$$

$$\left.\frac{\partial\Omega_u}{\partial\sigma_\nu^2}\right|_{H_0^a} = I_T \otimes I_N,$$

$$\left.\frac{\partial\Omega_u}{\partial\sigma_\mu^2}\right|_{H_0^a} = J_T \otimes I_N,$$

$$\left.\frac{\partial\Omega_u}{\partial\lambda}\right|_{H_0^a} = \sigma_\nu^2 I_T \otimes (W' + W).$$

This uses the fact that $B = I_N$ under H_0^a. Using (A.2), we obtain

$$\left.\frac{\partial L}{\partial\sigma_\nu^2}\right|_{H_0^a} = -\frac{1}{2}\, tr\left[\frac{1}{\tilde{\sigma}_\nu^2}(I_T \otimes (B'B)^{-1})\right] + \frac{1}{2}\left[\tilde{u}'\,\frac{1}{\tilde{\sigma}_\nu^4}\,(I_T \otimes (B'B)^{-1})\tilde{u}\right]$$

$$= -\frac{1}{2}\, tr\left[\frac{1}{\tilde{\sigma}_\nu^2} I_{NT}\right] + \frac{1}{2}\left[\frac{\tilde{u}'\tilde{u}}{\tilde{\sigma}_\nu^4}\right] = 0,$$

$$\left.\frac{\partial L}{\partial\sigma_\mu^2}\right|_{H_0^a} = D(\tilde{\sigma}_\mu^2) = \frac{NT}{2\tilde{\sigma}_\nu^2}\left(\frac{\tilde{u}'(J_T \otimes I_N)\tilde{u}}{\tilde{u}'\tilde{u}} - 1\right),$$

$$\left.\frac{\partial L}{\partial\lambda}\right|_{H_0^a} = D(\tilde{\lambda}) = \frac{NT}{2}\,\frac{\tilde{u}'(I_T \otimes (W + W'))\tilde{u}}{\tilde{u}'\tilde{u}}.$$

Therefore, the score with respect to θ, evaluated at the restricted MLE is given by

$$\tilde{D}_\theta = \begin{bmatrix} 0 \\ D(\tilde{\sigma}_\mu^2) \\ D(\tilde{\lambda}) \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{NT}{2\tilde{\sigma}_v^2}\left(\frac{\tilde{u}'(J_T \otimes I_N)\tilde{u}}{\tilde{u}'\tilde{u}} - 1\right) \\ NT\frac{\tilde{u}'(I_T \otimes (WW'))\tilde{u}}{\tilde{u}'\tilde{u}} \end{bmatrix}. \tag{A.4}$$

For the information matrix, it is useful to use the formula given by Harville (1977):

$$J_{rs} = E[-\partial^2 L/\partial\theta_r\partial\theta_s] = \frac{1}{2}\, tr[\Omega_u^{-1}(\partial\Omega_u/\partial\theta_r)\Omega_u^{-1}(\partial\Omega_u/\partial\theta_s)] \tag{A.5}$$

for $r, s = 1, 2, 3$. The corresponding elements of the information matrix are given by

$$J_{11} = E\left[-\frac{\partial^2 L}{\partial(\sigma_v^2)^2}\right] = \frac{1}{2}\, tr\left[\left(\frac{1}{\sigma_v^2}(I_T \otimes I_N)\right)^2\right] = \frac{NT}{2\sigma_v^4},$$

$$J_{22} = E\left[-\frac{\partial^2 L}{\partial(\sigma_\mu^2)^2}\right] = \frac{1}{2}\, tr\left[\frac{1}{\sigma_v^4}(J_T \otimes I_N)^2\right] = \frac{NT^2}{2\sigma_v^4},$$

$$J_{33} = E\left[-\frac{\partial^2 L}{\partial\lambda^2}\right] = \frac{1}{2}\, tr[I_T \otimes (W + W')^2]$$

$$= \tfrac{1}{2}\, tr[I_T \otimes (2W^2 + 2W'W)] = Tb,$$

$$J_{12} = \frac{1}{2}\, tr\left[\frac{1}{\sigma_v^2}(I_T \otimes I_N)\frac{1}{\sigma_v^2}(J_T \otimes I_N)\right] = \frac{NT}{2\sigma_v^4},$$

$$J_{13} = \frac{1}{2}\, tr\left[\frac{1}{\sigma_v^2}(I_T \otimes I_N)(I_T \otimes (W + W'))\right]$$

$$= \frac{1}{2\sigma_v^2}\, tr[I_T \otimes (W + W')] = 0,$$

$$J_{23} = \frac{1}{2}\, tr\left[\frac{1}{\sigma_v^2}(J_T \otimes I_N)(I_T \otimes (W + W'))\right]$$

$$= \frac{1}{2\sigma_v^2}\, tr\,[J_T \otimes (W + W')] = 0,$$

where the result that $J_{13} = J_{23} = 0$ follows from the fact that the diagonal elements of W are 0 and J_{33} uses the fact that $tr(W^2) = tr(W'^2)$ and $b = tr(W^2 + W'W)$.

Therefore, the information matrix evaluated under H_0^a is given by

$$\tilde{J}_\theta = \frac{NT}{2\tilde{\sigma}_v^4}\begin{bmatrix} 1 & 1 & 0 \\ 1 & T & 0 \\ 0 & 0 & \frac{2b\tilde{\sigma}_v^4}{N} \end{bmatrix}, \tag{A.6}$$

using (A.1), we get

$$LM_J = \left[0, \frac{NT}{2\tilde{\sigma}_v^2} G, NTH\right] \left(\frac{2\tilde{\sigma}_v^4}{NT}\right) \begin{bmatrix} \frac{T}{T-1} & -\frac{1}{T-1} & 0 \\ -\frac{1}{T-1} & \frac{1}{T-1} & 0 \\ 0 & 0 & \frac{N}{2b\tilde{\sigma}_v^4} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{NT}{2\tilde{\sigma}_v^2} G \\ NTH \end{bmatrix}$$

$$= \frac{NT}{2(T-1)} G^2 + \frac{N^2 T}{b} H^2, \tag{A.7}$$

where $G = \tilde{u}'(J_T \otimes I_N)\tilde{u}/\tilde{u}'\tilde{u} - 1$ and $H = \tilde{u}'(I_T \otimes (WW')/2)\tilde{u}/\tilde{u}'\tilde{u}$ as described in (2.11).

A.2. Conditional LM test for $\lambda = 0$ (given $\sigma_\mu^2 > 0$)

In this appendix we derive the conditional LM test which tests for no spatial error correlation given the existence of random regional effects. The null hypothesis is given by H_0^d: $\lambda = 0$ (assuming $\sigma_\mu^2 > 0$). Under the null hypothesis, the variance–covariance matrix reduces to $\Omega_0 = \sigma_\mu^2 J_T \otimes I_N + \sigma_v^2 I_{NT}$. It is the familiar form of the one-way error component model, see Baltagi (1995), with $\Omega_0^{-1} = (\sigma_1^2)^{-1}(\bar{J}_T \otimes I_N) + (\sigma_v^2)^{-1}(E_T \otimes I_N)$, where $\sigma_1^2 = T\sigma_\mu^2 + \sigma_v^2$.

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

$$\Omega_u^{-1}|_{H_0^d} = \left(\frac{1}{\sigma_1^2}\bar{J}_T + \frac{1}{\sigma_v^2} E_T\right) \otimes I_N, \tag{A.8}$$

$$\left.\frac{\partial \Omega_u}{\partial \sigma_v^2}\right|_{H_0^d} = I_T \otimes I_N, \tag{A.9}$$

$$\left.\frac{\partial \Omega_u}{\partial \sigma_\mu^2}\right|_{H_0^d} = J_T \otimes I_N, \tag{A.10}$$

$$\left.\frac{\partial \Omega_u}{\partial \lambda}\right|_{H_0^d} = \sigma_v^2 I_T \otimes (W + W'). \tag{A.11}$$

Using (A.2), one obtains

$$\left.\frac{\partial L}{\partial \sigma_v^2}\right|_{H_0^d} = -\frac{1}{2} tr[\Omega_u^{-1}(I_T \otimes (B'B)^{-1})] + \frac{1}{2}[\hat{u}'\Omega_u^{-1}(I_T \otimes (B'B)^{-1})\Omega_u^{-1}\hat{u}]$$

$$= -\frac{1}{2} tr\left[\left(\frac{1}{\hat{\sigma}_1^2}\bar{J}_T + \frac{1}{\hat{\sigma}_v^2}E_T\right) \otimes I_N\right] + \frac{1}{2}\left[\hat{u}'\left[\left(\frac{1}{\hat{\sigma}_1^4}\bar{J}_T + \frac{1}{\hat{\sigma}_v^4}E_T\right) \otimes I_N\right]\hat{u}\right]$$

$$= -\frac{\mathrm{N}(T-1)}{2\hat{\sigma}_v^2} - \frac{N}{2\hat{\sigma}_1^2} + \frac{1}{2}\hat{u}'\left(\frac{1}{\hat{\sigma}_1^4}\bar{J}_T \otimes I_N\right)\hat{u} + \frac{1}{2}\hat{u}'\left(\frac{1}{\hat{\sigma}_v^4}E_T \otimes I_N\right)\hat{u}$$

$$= 0, \tag{A.12}$$

$$\left.\frac{\partial L}{\partial \sigma_\mu^2}\right|_{H_0^d} = -\frac{1}{2}\, tr[\Omega_u^{-1}(J_T \otimes I_N)] + \frac{1}{2}[\hat{u}'\Omega_u^{-1}(J_T \otimes I_N)\Omega_u^{-1}\hat{u}]$$

$$= -\frac{NT}{2\hat{\sigma}_1^2} + \frac{1}{2\hat{\sigma}_1^4}\,[\hat{u}'(\bar{J}_T \otimes I_N)\hat{u}] = 0, \tag{A.13}$$

$$\left.\frac{\partial L}{\partial \lambda}\right|_{H_0^d} = -\frac{1}{2}\, tr[\Omega_u^{-1}(\sigma_v^2 I_T \otimes (W + W'))] + \frac{1}{2}[\hat{u}'\Omega_u^{-1}(\sigma_v^2 I_T \otimes (W + W'))\Omega_u^{-1}\hat{u}]$$

$$= \frac{1}{2\hat{\sigma}_v^2}\,[\hat{u}'(E_T \otimes (W + W'))\hat{u}] + \frac{\hat{\sigma}_v^2}{2\hat{\sigma}_1^4}[\hat{u}'(\bar{J}_T \otimes (W + W'))\hat{u}] = \hat{D}_\lambda, \tag{A.14}$$

where $\hat{\sigma}_v^2 = \hat{u}'(E_T \otimes I_N)\hat{u}/\mathrm{N}(T-1)$ and $\hat{\sigma}_1^2 = \hat{u}'(\bar{J}_T \otimes I_N)\hat{u}/N$ are the MLEs of σ_v^2 and σ_1^2, and $\hat{u}$ is the maximum likelihood residual under the null hypothesis H_0^d.

Therefore, the score vector under H_0^d is given by

$$\hat{D} = \begin{bmatrix} 0 \\ 0 \\ \hat{D}_\lambda \end{bmatrix}. \tag{A.15}$$

Using (A.5), the elements of the information matrix are given by

$$J_{11} = E\left[-\frac{\partial^2 L}{\partial(\sigma_v^2)^2}\right] = \frac{1}{2}\, tr[((\sigma_1^{-2}\bar{J}_T + \sigma_v^{-2}E_T) \otimes I_N)^2]$$

$$= \frac{N}{2}\left(\frac{1}{\sigma_1^4} + \frac{T-1}{\sigma_v^4}\right), \tag{A.16}$$

$$J_{22} = E\left[-\frac{\partial^2 L}{\partial(\sigma_\mu^2)^2}\right] = \frac{1}{2}\, tr[(((\sigma_1^{-2}\bar{J}_T + \sigma_v^{-2}E_T) \otimes I_N)(J_T \otimes I_N))^2]$$

$$= \frac{NT^2}{2\sigma_1^4}, \tag{A.17}$$

$$J_{33} = E\left[-\frac{\partial^2 L}{\partial\lambda^2}\right] = \frac{1}{2}\, tr[(((\sigma_1^{-2}\bar{J}_T + \sigma_v^{-2}E_T) \otimes I_N)(\sigma_v^2 I_T \otimes (W + W')))^2]$$

$$= \frac{\sigma_v^4}{2}\, tr[(\sigma_1^{-4}\bar{J}_T + \sigma_v^{-4}E_T) \otimes (2W^2 + 2W'W)] = \left(\frac{\sigma_v^4}{\sigma_1^4} + (T-1)\right) b, \tag{A.18}$$

$$J_{12} = E\left[-\frac{\partial^2 L}{\partial\sigma_v^2\partial\sigma_\mu^2}\right] = \frac{1}{2}\, tr[((\sigma_1^{-2}\bar{J}_T + \sigma_v^{-2}E_T) \otimes I_N)\sigma_1^{-2}(J_T \otimes I_N)]$$

$$= \frac{NT}{2\sigma_1^4}, \tag{A.19}$$

$$J_{13} = E\left[-\frac{\partial^2 L}{\partial\sigma_v^2\partial\lambda}\right] = \frac{1}{2}\,tr[((\sigma_1^{-2}\bar{J}_T + \sigma_v^{-2}E_T)\otimes I_N)$$
$$\times((\sigma_1^{-2}\bar{J}_T + \sigma_v^{-2}E_T)\otimes I_N)(\sigma_v^2 I_T\otimes(W+W'))] = 0, \tag{A.20}$$

$$J_{23} = E\left[-\frac{\partial^2 L}{\partial\sigma_\mu^2\partial\lambda}\right]$$
$$= \tfrac{1}{2}\,tr[\sigma_1^{-2}(J_T\otimes I_N)((\sigma_1^{-2}\bar{J}_T + \sigma_v^{-2}E_T)\otimes I_N)(\sigma_v^2 I_T\otimes(W+W'))]$$
$$= \frac{\sigma_v^2}{2}\,tr\left[\frac{1}{\sigma_1^4}J_T\otimes(W+W')\right] = 0, \tag{A.21}$$

where the result that $J_{13} = J_{23} = 0$ follows from the fact that the diagonal elements of W is 0 and J_{33} uses the fact that $tr(W^2) = tr(W'^2)$, and $b = tr(W^2 + W'W)$.

Therefore, the information matrix evaluated under H_0^d is given by

$$\hat{J}_\theta = \begin{bmatrix} \frac{N}{2}(\frac{1}{\hat{\sigma}_1^4} + \frac{T-1}{\hat{\sigma}_v^4}) & \frac{NT}{2\hat{\sigma}_1^4} & 0 \\ \frac{NT}{2\hat{\sigma}_1^4} & \frac{NT^2}{2\hat{\sigma}_1^4} & 0 \\ 0 & 0 & (T-1+\frac{\hat{\sigma}_v^4}{\hat{\sigma}_1^4})b \end{bmatrix}. \tag{A.22}$$

Therefore,

$$LM_\lambda = \hat{D}_\lambda'\hat{J}_\theta^{-1}\hat{D}_\lambda$$
$$= \frac{\hat{D}(\lambda)^2}{[(T-1)+\frac{\hat{\sigma}_v^4}{\hat{\sigma}_1^4}]b},$$

as described in (2.24).

A.3. Conditional LM test for $\sigma_\mu^2 = 0$ (assuming $\lambda \neq 0$)

This appendix derives the conditional LM tests for zero random regional effects assuming that spatial error correlation exists. We give the detailed derivation of the score and information matrix for testing $H_0^e : \sigma_\mu^2 = 0$ (assuming $\lambda \neq 0$).

Under H_0^e,

$$\Omega_0 = \sigma_v^2 I_T \otimes (B'B)^{-1}, \tag{A.23}$$

$$\Omega_0^{-1} = \frac{1}{\sigma_v^2}\,I_T\otimes(B'B) \tag{A.24}$$

and

$$\left.\frac{\partial\Omega_u}{\partial\sigma_\mu^2}\right|_{H_0^e} = J_T\otimes I_N, \tag{A.25}$$

$$\left.\frac{\partial \Omega_u}{\partial \sigma_\nu^2}\right|_{H_0^e} = I_T \otimes (\hat{B}'\hat{B})^{-1}, \tag{A.26}$$

$$\left.\frac{\partial \Omega_u}{\partial \lambda}\right|_{H_0^e} = \hat{\sigma}_\nu^2 I_T \otimes (\hat{B}'\hat{B})^{-1}(W'\hat{B} + \hat{B}'W)(\hat{B}'\hat{B})^{-1}, \tag{A.27}$$

where $\hat{B} = I_N - \hat{\lambda}W$ and $\hat{\lambda}$ is the MLE of λ under H_0^e. Using (A.2), we get

$$\hat{D}_\mu = \left.\frac{\partial L}{\partial \sigma_\mu^2}\right|_{H_0^e} = -\frac{T}{2\hat{\sigma}_\nu^2}\, tr(\hat{B}'\hat{B}) + \frac{1}{2\hat{\sigma}_\nu^4}\, \hat{u}'[J_T \otimes (\hat{B}'\hat{B})^2]\hat{u}. \tag{A.28}$$

This uses the fact that

$$\Omega_u^{-1}\left.\frac{\partial \Omega_u}{\partial \sigma_\mu^2}\right|_{H_0^e} = \left(\frac{1}{\hat{\sigma}_\nu^2} I_T \otimes \hat{B}'\hat{B}\right)(J_T \otimes I_N) = \frac{1}{\hat{\sigma}_\nu^2} J_T \otimes \hat{B}'\hat{B},$$

$$\Omega_u^{-1}\left.\frac{\partial \Omega_u}{\partial \sigma_\mu^2}\Omega_u^{-1}\right|_{H_0^e} = \left(\frac{1}{\hat{\sigma}_\nu^2} J_T \otimes \hat{B}'\hat{B}\right)\left(\frac{1}{\hat{\sigma}_\nu^2} I_T \otimes \hat{B}'\hat{B}\right) = \frac{1}{\hat{\sigma}_\nu^4} J_T \otimes (\hat{B}'\hat{B})^2,$$

$$tr\left[\Omega_u^{-1}\frac{\partial \Omega_u}{\partial \sigma_\mu^2}\right]\Bigg|_{H_0^e} = \frac{1}{\hat{\sigma}_\nu^2}\, tr[J_T \otimes \hat{B}'\hat{B}] = \frac{T}{\hat{\sigma}_\nu^2}\, tr[\hat{B}'\hat{B}].$$

Similarly

$$\left.\frac{\partial L}{\partial \sigma_\nu^2}\right|_{H_0^e} = -\frac{TN}{2\hat{\sigma}_\nu^2} + \frac{1}{2\hat{\sigma}_\nu^4}\, \hat{u}'[I_T \otimes (\hat{B}'\hat{B})]\hat{u} = 0, \tag{A.29}$$

which yields

$$\hat{\sigma}_\nu^2 = \frac{\hat{u}'[I_T \otimes \hat{B}'\hat{B}]\hat{u}}{TN}.$$

This uses the fact that

$$\Omega_u^{-1}\left.\frac{\partial \Omega_u}{\partial \sigma_\nu^2}\right|_{H_0^e} = \left(\frac{1}{\hat{\sigma}_\nu^2} I_T \otimes \hat{B}'\hat{B}\right)(I_T \otimes (\hat{B}'\hat{B})^{-1}) = \frac{1}{\hat{\sigma}_\nu^2} I_T \otimes I_N,$$

$$\Omega_u^{-1}\left.\frac{\partial \Omega_u}{\partial \sigma_\nu^2}\Omega_u^{-1}\right|_{H_0^e} = \left(\frac{1}{\hat{\sigma}_\nu^2} I_T \otimes I_N\right)\left(\frac{1}{\hat{\sigma}_\nu^2} I_T \otimes \hat{B}'\hat{B}\right) = \frac{1}{\hat{\sigma}_\nu^4} I_T \otimes \hat{B}'\hat{B},$$

$$tr\left[\Omega_u^{-1}\frac{\partial \Omega_u}{\partial \sigma_\nu^2}\right]\Bigg|_{H_0^e} = \frac{TN}{\hat{\sigma}_\nu^2}.$$

Also,

$$\left.\frac{\partial L}{\partial \lambda}\right|_{H_0^e} = -\frac{T}{2}\, tr[(W'\hat{B} + \hat{B}'W)(\hat{B}'\hat{B})^{-1}]$$
$$+ \frac{1}{2\hat{\sigma}_\nu^2}\, \hat{u}[I_T \otimes (W'\hat{B} + \hat{B}'W)]\hat{u} = 0,$$

because

$$\Omega_u^{-1}\left.\frac{\partial\Omega_u}{\partial\lambda}\right|_{H_0^e}=\left(\frac{1}{\hat\sigma_\nu^2}I_T\otimes\hat B'\hat B\right)\left(\hat\sigma_\nu^2 I_T\otimes(\hat B'\hat B)^{-1}(W'\hat B+\hat B'W)(\hat B'\hat B)^{-1}\right)$$

$$=I_T\otimes(W'\hat B+\hat B'W)(\hat B'\hat B)^{-1},$$

$$\Omega_u^{-1}\left.\frac{\partial\Omega_u}{\partial\lambda}\,\Omega_u^{-1}\right|_{H_0^e}=\frac{1}{\hat\sigma_\nu^2}I_T\otimes(W'\hat B+\hat B'W).$$

Therefore, the score vector under H_0^e is given by

$$\hat D=\begin{bmatrix}0\\0\\\hat D_\mu\end{bmatrix}. \tag{A.30}$$

The elements of the information matrix for this model using (A.5) are given by

$$J_{11}=E\left[-\frac{\partial^2 L}{\partial(\sigma_\nu^2)^2}\right]\Bigg|_{H_0^e}=\frac{1}{2}\,tr[((\hat\sigma_\nu^2)^{-1}I_T\otimes I_N)^2]=\frac{TN}{2\hat\sigma_\nu^4},$$

$$J_{22}=E\left[-\frac{\partial^2 L}{\partial\lambda^2}\right]\Bigg|_{H_0^e}=\frac{1}{2}\,tr[(I_T\otimes(W'\hat B+\hat B'W)(\hat B'\hat B)^{-1})^2]$$

$$=\frac{T}{2}\,tr[((W'\hat B+\hat B'W)(\hat B'\hat B)^{-1})^2],$$

$$J_{33}=E\left[-\frac{\partial^2 L}{\partial\sigma_\mu^2}\right]\Bigg|_{H_0^e}=\frac{1}{2}\,tr[((\hat\sigma_\nu^2)^{-1}J_T\otimes\hat B'\hat B)^2]=\frac{T^2}{2\hat\sigma_\nu^4}\,tr[(\hat B'\hat B)^2],$$

$$J_{12}=E\left[-\frac{\partial^2 L}{\partial\sigma_\nu^2\partial\lambda}\right]\Bigg|_{H_0^e}=\frac{1}{2}\,tr[(\hat\sigma_\nu^2)^{-1}(I_T\otimes I_N)(I_T\otimes(W'\hat B+\hat B'W)(\hat B'\hat B)^{-1})]$$

$$=\frac{1}{2}\,tr[(\hat\sigma_\nu^2)^{-1}I_T\otimes(W'\hat B+\hat B'W)(\hat B'\hat B)^{-1})]$$

$$=\frac{T}{2\hat\sigma_\nu^2}tr[(W'\hat B+\hat B'W)(\hat B'\hat B)^{-1}],$$

$$J_{13}=E\left[-\frac{\partial^2 L}{\partial\sigma_\nu^2\partial\sigma_\mu^2}\right]\Bigg|_{H_0^e}=\frac{1}{2}\,tr[(\hat\sigma_\nu^2)^{-1}(I_T\otimes I_N)((\hat\sigma_\nu^2)^{-1}J_T\otimes\hat B'\hat B)]$$

$$=\frac{1}{2\hat\sigma_\nu^4}\,tr[J_T\otimes\hat B'\hat B]=\frac{T}{2\hat\sigma_\nu^4}\,tr[\hat B'\hat B],$$

$$J_{23} = E\left[-\frac{\partial^2 L}{\partial\sigma_\mu^2 \partial\lambda}\right]\Bigg|_{H_0^e} = \frac{1}{2}\, tr[(I_T \otimes (W'\hat{B} + \hat{B}'W)(\hat{B}'\hat{B})^{-1})((\hat{\sigma}_v^2)^{-1} J_T \otimes \hat{B}'\hat{B})]$$

$$= \frac{1}{2\hat{\sigma}_v^2}\, tr[J_T \otimes (W'\hat{B} + \hat{B}'W)] = \frac{T}{2\hat{\sigma}_v^2}\, tr[W'\hat{B} + \hat{B}'W],$$

$$\hat{J}_\theta = \begin{bmatrix} \frac{TN}{2\hat{\sigma}_v^4} & \frac{T}{2\hat{\sigma}_v^2}\, tr[(W'\hat{B} + \hat{B}'W)(\hat{B}'\hat{B})^{-1}] & \frac{T}{2\hat{\sigma}_v^4}\, tr[\hat{B}'\hat{B}] \\ & \frac{T}{2}\, tr[((W'\hat{B} + \hat{B}'W)(\hat{B}'\hat{B})^{-1})^2] & \frac{T}{2\hat{\sigma}_v^2}\, tr[W'\hat{B} + \hat{B}'W] \\ & & \frac{T^2}{2\hat{\sigma}_v^4}\, tr[(\hat{B}'\hat{B})^2] \end{bmatrix}. \tag{A.31}$$

References

Anselin, L., 1988. Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Dordrecht.

Anselin, L., 2001. Rao's score tests in spatial econometrics. Journal of Statistical Planning and Inference 97, 113–139.

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., Florax, R., 1995. Small sample properties of tests for spatial dependence in regression models: some further results. In: Anselin, L., Florax, R. (Eds.), New Directions in Spatial Econometrics. Springer, Berlin, pp. 21–74.

Anselin, L., Rey, S., 1991. Properties of tests for spatial dependence in linear regression models. Geographical Analysis 23, 112–131.

Anselin, L., Bera, A.K., Florax, R., Yoon, M.J., 1996. Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics 26, 77–104.

Baltagi, B.H., 2001. Econometrics Analysis of Panel Data. Wiley, Chichester.

Baltagi, B.H., Li, D., 1999. Prediction in the panel data model with spatial correlation. In: Anselin, L., Florax, R.J.G.M. (Eds.), New Advances in Spatial Econometrics, forthcoming.

Baltagi, B.H., Chang, Y.J., Li, Q., 1992. Monte Carlo results on several new and existing tests for the error component model. Journal of Econometrics 54, 95–120.

Bell, K.P., Bockstael, N.R., 2000. Applying the generalized-moments estimation approach to spatial problems involving microlevel data. Review of Economics and Statistics 82, 72–82.

Breusch, T.S., Pagan, A.R., 1980. The Lagrange Multiplier test and its application to model specification in econometrics. Review of Economic Studies 47, 239–254.

Case, A.C., 1991. Spatial patterns in household demand. Econometrica 59, 953–965.

Case, A.C., Hines Jr., J., Rosen, H., 1993. Budget spillovers and fiscal policy independence: evidence from the states. Journal of Public Economics 52, 285–307.

Conley, T.G., 1999. GMM estimation with cross sectional dependence. Journal of Econometrics 92, 1–45.

De Long, J.B., Summers, L.H., 1991. Equipment investment and economic growth. Quarterly Journal of Economics 106, 445–502.

Driscoll, J., Kraay, A., 1998. Consistent covariance matrix estimation with spatially dependent panel data. Review of Economics and Statistics 80, 549–560.

Evans, M.A., King, M.L., 1985. Critical value approximations for tests of linear regression disturbances. Australian Journal of Statistics 27, 68–83.

Gourieroux, C., Holly, A., Monfort, A., 1982. Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters. Econometrica 50, 63–80.

Hartley, H.O., 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., Hartley, H.O., 1973. Computing maximum likelihood estimates for the mixed A.O.V. model using the W-transformation. Technometrics 15, 819–831.

Holtz-Eakin, D., 1994. Public-sector capital and the productivity puzzle. Review of Economics and Statistics 76, 12–21.

Honda, Y., 1985. Testing the error components model with non-normal disturbances. Review of Economic Studies 52, 681–690.

Hsiao, C., 1986. Analysis of Panel Data. Cambridge University Press, Cambridge.

Islam, N., 1995. Growth empirics: a panel data approach. Quarterly Journal of Economics 110, 1127–1170.

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.

Kelejian, H.H., Prucha, I.R., 2001. On the asymptotic distribution of the Moran I test with applications. Journal of Econometrics 104, 219–257.

Kelejian, H.H., Robinson, D.P., 1992. Spatial autocorrelation: a new computationally simple test with an application to per capita county police expenditures. Regional Science and Urban Economics 22, 317–331.

Moulton, B.R., Randolph, W.C., 1989. Alternative tests of the error components model. Econometrica 57, 685–693.

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

Pinkse, J., 1998. Asymptotic properties of Moran and related tests and a test for spatial correlation in probit models. Working paper, Department of Economics, University of British Columbia.

Pinkse, J., 1999. Moran-flavoured tests with nuisance parameters: examples. In: Anselin, L., Florax, R.J.G.M. (Eds.), New Advances in Spatial Econometrics, forthcoming.

Wansbeek, T.J., Kapteyn, A., 1983. A note on spectral decomposition and maximum likelihood estimation of ANOVA models with balanced data. Statistics and Probability Letters 1, 213–215.