Spatial Economic Analysis
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rsea20

Testing for Spatial Effects in Seemingly Unrelated Regressions

Jesús Mur , Fernando López & Marcos Herrera

Published online: 08 Nov 2010.

To cite this article: Jesús Mur , Fernando López & Marcos Herrera (2010) Testing for Spatial Effects in Seemingly Unrelated Regressions, Spatial Economic Analysis, 5:4, 399-440, DOI: 10.1080/17421772.2010.516443

To link to this article: http://dx.doi.org/10.1080/17421772.2010.516443

Spatial Economic Analysis, Vol. 5, No. 4, December 2010

Testing for Spatial Effects in Seemingly Unrelated Regressions

JESÚS MUR, FERNANDO LÓPEZ & MARCOS HERRERA

(Received November 2009; accepted July 2010)

ABSTRACT *The paper focuses on the case of a panel data set, without unobserved individual effects, treated by means of an SUR specification. The problem raised is to test for the presence of spatial effects in these multivariate systems. Various useful tests are developed based on the principle of the Lagrange Multiplier in a maximum-likelihood framework. Also, we address the question of the time stability of the sequence of spatial dependence coefficients, as a maintained hypothesis that is not necessarily true in applied work. The second part of the paper presents the results of a Monte Carlo experiment.*

Essais sur les effets spatiaux dans des régressions apparemment sans rapport

RESUME *Cette communication se concentre sur le cas de l'ensemble de données de panel, sans effets individuels non observés, traitées au moyen d'une spécification SUR. Le problème soulevé concerne l'examen de la présence d'effets spatiaux dans ces systèmes à multi-variables. On développe plusieurs essais utiles, basés sur le principe du multiplicateur d'Euler-Lagrange dans un cadre de probabilité maximale. En outre, nous nous penchons sur la question de la stabilité en fonction du temps des coefficients de dépendance spatiale, en tant qu'hypothèse maintenue qui n'est pas nécessairement vraie dans les applications pratiques. La deuxième partie de la communication présente les résultats d'une expérience Monte Carlo.*

Ensayando los efectos espaciales en ecuaciones aparentemente no relacionadas

Extracto *El trabajo se centra en el caso de un conjunto de datos panel, donde no existen efectos individuales inobservados, tratado mediante una especificación SUR. El problema que se plantea es contrastar la existencia de efectos espaciales en ese tipo de sistemas multivariantes. Se desarrollan varios contrastes basados en el principio del Multiplicador de Lagrange en un contexto de máxima*

Jesús Mur (to whom correspondence should be sent), Department of Economic Analysis, University of Zaragoza, Gran Vía 2–4 (50005), Zaragoza, Spain. Email: jmur@unizar.es. Fernando López, Department of Quantitative Methods and Computing, Technical University of Cartagena, Calle Real 3, 30201 Cartagena, Spain. Email: fernando.lopez@upct.es. Marcos Herrera, Department of Economic Analysis, University of Zaragoza, Gran Vía 2–4 (50005), Zaragoza, Spain. Email: mherreragomez@gmail.com. This work has been carried out with the financial support of project ECO2009-10534/ECON of the Ministerio de Ciencia y Tecnología of the Reino de España and the project 11897/PHCS/09 of the Fundación Séneca—Agencia de Ciencia y Tecnología de la Región de Murcia.

ISSN 1742-1772 print; 1742-1780 online/10/040399-42 DOI: 10.1080/17421772.2010.516443

verosimilitud. Igualmente tratamos la cuestiœn de la estabilidad temporal en la secuencia de coeficientes de dependencia espacial, como una hipótesis mantenida que no es necesariamente cierta en trabajos de tipo aplicados. La parte final del artículo presenta los resultados de un experimento de Monte Carlo.

半相依回归中空间效应的检测

摘要：本文讨论了一个经过SUR处理、没有未观察到的个体效应的面板数据集。本文提出的问题是检验这些多元系统中空间效应的存在。笔者在最大似然法框架下，根据拉格朗日乘数法原理，开发了多种有用的检验方法。此外，笔者还讨论了空间依赖性系数序列的时间稳定性问题，这是一个在应用工作中不一定有效的保留假设。文章的第二部分列出了一项蒙特卡罗实验的结果。

KEYWORDS: *Spatial dependence; seemingly unrelated regressions; Monte Carlo*

JEL CLASSIFICATION: C21*;* C50*;* R15

1. Introduction

The seemingly unrelated regressions (SUR) equations are traditional multi-variant econometric formulations employed in very different fields, obviously including spatial analysis. The basis of the approach is very well-known due to the initial works of Zellner (1962), Malinvaud (1970), Theil (1971), Schmidt (1976) and also Dwivedi & Srivastava (1978). Almost every econometric textbook includes a discussion of the SUR methodology, which is available in most popular econometric software packages. It hardly requires any further justification.

SUR models are discussed in Chapter 10 of the textbook by Anselin (1988a), which introduces the term 'spatial SUR' as that which 'consists of an equation for each time period which is estimated for a cross-section of spatial units' (p. 141). In each of these equations (cross-sections, in fact), some spatial elements may be introduced in either the form of mechanisms of intra-equation spatial error autocorrelation, or a spatially lagged dependent variable or both. Rey & Montouri (1999), Fingleton (2001, 2007), Egger & Pfaffermayr (2004), Moscone *et al.* (2007), LeGallo & Chasco (2008) and Lauridsen *et al.* (2010) have all used this technique on different occasions.

The main characteristics of the spatial SUR approach are the existence of a limited heterogeneity among the individuals (the regression coefficients are assumed to remain constant) and a certain imbalance between the cross-section dimension (where R is the number of individuals in the sample) and the number of cross-sections (where T is the time dimension). Typically, using Anselin's approach in the context of our paper, the former should be greater than the latter (or the ratio T/R should tend towards zero). If we need more flexibility, allowing for some individual heterogeneity, the next step may consist of specifying a spatial panel data model (Elhorst, 2003, 2005, 2008; Anselin *et al.*, 2007; Baltagi, 2008).

If the number of cross-sections is greater than the number of spatial units (the ratio R/T tends towards zero) other approaches may be preferable, specifying for example, one equation for each individual. Arora & Brown (1977), Hordijk & Nijkamp (1977), Hordijk (1979) or White & Hewings (1982) are examples of this line of thought that allow for an absolute heterogeneity. The covariance matrix among the error terms of the different equations is used as a measure of spatial dependence, thus avoiding the necessity of specifying a weighting matrix. This is the method followed by Conley (1999), Chen & Conley (2001), Coakley *et al.* (2002), Pesaran (2005) or Conley & Molinari (2008) among others. Problems appear when R increases at a rate similar to T. In this case, as indicated by Driscoll & Kraay (1998), it is necessary to introduce restrictions on the parameters to keep the dimensionality of the problem more manageable. Some recent applications in this line, referred to as spatial VAR models, are Carlino & DeFina (1999), Di Giacinto (2003, 2006), Badinger *et al.* (2004) or Beenstock & Felsenstein (2007, 2008).

Referring to the present paper, our preferred combination is a finite T and a large R. The case of an SUR model with spatial effects will be solely addressed, taking into account the question of specification and estimation. The problem is, first, to test for the presence of spatial effects in the SUR specification and, second, to confront the question of identifying the type of spatial process most relevant to the data. With regards to the latter question it should be acknowledged that the current situation isn't very satisfactory, as LeGallo & Dall'erba (2006, p. 279) point out: 'For our SUR specification with spatial autocorrelation and spatial regimes, no specification procedure has been formally suggested. Consequently, we apply here a sequential strategy similar to that applied for cross-sections'.

The paper contains five sections. In the second section we develop a set of Lagrange Multipliers to test for the presence of spatial effects in a standard spatial SUR model. The third section contains some additions (robustness, instability and diagonality) to the basic case presented in Section 2. In the fourth section we use a Monte Carlo experiment designed to study the behaviour of the tests previously obtained in a finite sample context. Finally, the fifth section comments on the main conclusion reached in the paper.

2. Testing for Spatial Autocorrelation: A Maximum-Likelihood Approach

Our specification follows the original design of Anselin (1988b): R individuals are observed in T different time periods; these are TR observations. Only one equation is specified with an endogenous variable, called y, and a set of explanatory variables, $\{x_1, x_2, \ldots, x_k\}$. This equation is copied for each individual and time period. If there are unobserved effects, a set of (individual, temporal) dummy variables will also be introduced. The data for each individual show some inertia which results in a structure of temporal dependence in the corresponding error term (Elhorst, 2001).

The model described above corresponds to an SUR specification to which we add the existence of spatial effects. The general case of a model with an autoregressive structure, both in the main equation and in the errors, is presented first (known as an SARAR model). Next, the Spatial Lag Model (SLM) contains a spatial autoregressive term only in the main equation, whereas this autoregression appears only in the equation of the errors in the case of the Spatial Error Model (SEM).

2.1. An SARAR Model

The characteristic of this model is that the spatial effects appear both in the main equations and in the equations of the errors:

$$\left.\begin{array}{ll} y_t = \lambda_t \mathbf{W}_1 y_t + x_t\beta + u_t & \Rightarrow \mathbf{A}_t y_t = x_t\beta + u_t \\ u_t = \rho_t \mathbf{W}_2 u_t + \varepsilon_t & \Rightarrow \mathbf{B}_t u_t = \varepsilon_t \end{array}\right\}$$
$$\mathbf{A}_t = I_R - \lambda_t \mathbf{W}_1 \quad \mathbf{B}_t = I_R - \rho_t \mathbf{W}_2 \tag{1}$$

y_t, u_t and ε_t are vectors of order $(R \times 1)$, x_t is a matrix of order $(R \times k)$, β is a vector of parameters of order $(k \times 1)$, I_R is the identity matrix of order $(R \times R)$ and $\mathbf{W}_1$ and $\mathbf{W}_2$ are weighting matrices also of order $(R \times R)$. For the sake of simplicity, in the following we assume that the two weighting matrices coincide ($\mathbf{W}_1 = \mathbf{W}_2 = \mathbf{W}$). The coefficients of the model and the weighting matrix are allowed to vary among the different cross-sections. However, in order to simplify the discussion, in the following we assume that vector β and matrix $\mathbf{W}$ remain the same, but that the parameters of dependence, ρ_t and λ_t may take different values in different cross-sections. Using a more compact notation:

$$\left.\begin{array}{l} \mathbf{A}y = X\beta + u \\ \mathbf{B}u = \varepsilon \\ \varepsilon \sim N(0, \boldsymbol{\Omega}) \end{array}\right\}$$
$$\mathbf{y} = \underset{TRx1}{\begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_T \end{bmatrix}}; \quad \mathbf{X} = \underset{TRXTk}{\begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & x_T \end{bmatrix}}; \quad \mathbf{u} = \underset{TRx1}{\begin{bmatrix} u_1 \\ u_2 \\ \cdots \\ u_T \end{bmatrix}}; \quad \boldsymbol{\varepsilon} = \underset{TRx1}{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_T \end{bmatrix}} \tag{2}$$

where $\mathbf{A} = I_{TR} - \Lambda \otimes \mathbf{W}$ and $\mathbf{B} = I_{TR} - \Upsilon \otimes \mathbf{W}$; Λ (respectively Υ) is an $R \times R$ diagonal matrix containing the parameters λ_t (respectively ρ_t) and $\otimes$ is the Kronecker product. The temporal dependence, introduced by the SUR mechanism, leads to the matrix $\boldsymbol{\Omega} = \Sigma \otimes I_R$ where Σ is a $T \times T$ matrix, $\Sigma = \{\sigma_{ij}, i, j = 1, 2, \ldots, T\}$. As usual, the error terms are assumed to be normal.

The logarithm of the likelihood function of the model in Equation (2) is:

$$l(y;\theta) = -\frac{RT}{2}\ln(2\pi) - \frac{R}{2}\ln|\boldsymbol{\Sigma}| + \sum\nolimits_{t=1}^{T} \ln|\mathbf{B}_t| + \sum\nolimits_{t=1}^{T} \ln|\mathbf{A}_t|$$
$$- \frac{(\mathbf{A}y - X\beta)'\mathbf{B}'(\boldsymbol{\Sigma} \otimes I_R)^{-1}\mathbf{B}(\mathbf{A}y - X\beta)}{2} \tag{3}$$

where $\theta' = [\beta'; \lambda_1; \cdots; \lambda_T; \rho_1; \cdots; \rho_T; \text{vech}(\Sigma)']$ is the vector of parameters of the model, of order $(k + 2T + T(T+1)/2) \times 1$. The ML estimation of these parameters results in a nonlinear optimization problem which can be solved by applying standard techniques (Wang & Kockelman, 2007). In Section A.I of the Appendix more details are included in relation to the score vector and the information matrix.

In order to test for the existence of spatial effects in the specification of Equation (1), we can obtain the corresponding Lagrange Multiplier. The null hypothesis of the absence of spatial effects implies that:

$$H_0: \lambda_t = \rho_t = 0 \ (\forall \text{t}) \qquad \text{vs} \qquad H_A: \text{No H}_0 \tag{4}$$

After a few calculations, the details of which appear in Section A.II of the Appendix, the final expression of the test is:[1]

$$\mathbf{LM}_{\text{SARAR}}^{\text{SUR}} = [g'_{(\lambda)_{|H_0}} \quad g'_{(\rho)_{|H_0}}] \begin{bmatrix} I_{\lambda\lambda} - I_{\lambda\beta} I_{\beta\beta}^{-1} I_{\beta\lambda} & I_{\lambda\rho} \\ I_{\rho\lambda} & I_{\rho\rho} \end{bmatrix}^{-1} \begin{bmatrix} g_{(\lambda)_{|H_0}} \\ g_{(\rho)_{|H_0}} \end{bmatrix} \underset{as}{\sim} \chi^2(2T) \tag{5}$$

with $g_{(\lambda)_{|H_0}} = [\mathbf{\Sigma}^{-1} \circ (\hat{U}' Y_L)]'\tau$ and $g_{(\rho)_{|H_0}} = [\mathbf{\Sigma}^{-1} \circ (\hat{U}' \hat{U}_L)]'\tau$, where '$\circ$' is the Hadamard product, τ is a $(T \times 1)$ vector of ones, and $\hat{U}$ is a $(R \times T)$ matrix where the columns contain the SUR residuals corresponding to each of the cross-sections: $\hat{U} = [\hat{u}_1 \quad \hat{u}_2 \quad \cdots \quad \hat{u}_T]$. $\hat{U}_L$ is the vector of spatially lagged residuals $\hat{U}_L = [\mathbf{W}\hat{u}_1 \quad \mathbf{W}\hat{u}_2 \quad \cdots \quad \mathbf{W}\hat{u}_T]$; Y_L is specified analogously. The terms of the information matrix appear in expression (A9) of Section A.II of the Appendix.

2.2. *An SLM model*

The Spatial Lag Model is a particular case of the SARAR model:

$$\begin{aligned} y_t &= \lambda_t \mathbf{W} y_t + x_t\beta + \varepsilon_t \Rightarrow \mathbf{A}_t y_t = x_t\beta + \varepsilon_t \\ \mathbf{A}_t &= I_R - \lambda_t \mathbf{W} \end{aligned} \tag{6}$$

All the elements that intervene in this specification have already been defined. Using a more compact matrix notation:

$$\left.\begin{array}{c} \mathbf{A}y = X\beta + \varepsilon \\ \varepsilon \sim N(0, \mathbf{\Omega}) \end{array}\right\} \tag{7}$$

where $\mathbf{A} = I_{TR} - \Lambda \otimes \mathbf{W}$ and ω is the same matrix as that indicated in Equation (2). The logarithm of the likelihood function, introducing the SLM structure of Equation (7), is:

$$\begin{aligned} l(y;\theta) = &-\frac{RT}{2}\ln(2\pi) \\ &-\frac{R}{2}\ln|\mathbf{\Sigma}| + \sum\nolimits_{t=1}^{T} \ln|\mathbf{A}_t| - \frac{(\mathbf{A}y - X\beta)'(\mathbf{\Sigma} \otimes I_R)^{-1}(\mathbf{A}y - X\beta)}{2} \end{aligned} \tag{8}$$

where $\theta' = [\beta'; \lambda_1; \cdots; \lambda_T; vech(\Sigma)']$ is the vector of parameters of the model, of order $(k + T + T(T+1)/2) \times 1$. More details are shown in the Appendix, Section A.III.

The null hypothesis is that the spatial lag of the endogenous variable is, statistically, not relevant:

$$H_0{:}\lambda_t = 0 \ (\forall t) \text{ vs } H_A{:}\text{No } H_0 \tag{9}$$

The Lagrange Multiplier emerges as:

$$\mathbf{LM}_{\text{SLM}}^{\text{SUR}} = g'_{(\lambda)_{|H_0}} [I_{\lambda\lambda} - I_{\lambda\beta} I_{\beta\beta}^{-1} I_{\beta\lambda}]^{-1} g_{(\lambda)_{|H_0}} \underset{\text{as}}{\sim} \chi^2(T) \tag{10}$$

The terms of the score can be expressed as: $g_{(\lambda)_{|H_0}} = [\mathbf{\Sigma}^{-1} \circ (\hat{U}' Y_L)]'\tau$ whereas the elements that intervene in the matrix $I_{\lambda\lambda} - I_{\lambda\beta} I_{\beta\beta}^{-1} I_{\beta\lambda}$ are defined in expression (A21) in Section A.IV of the Appendix.

2.3. The SEM Model

The SUR model with a structure of spatial dependence in the error term, SEM, is well known due to the seminal work of Anselin (1988b). However, to be thorough, we have included this case in our discussion. The structure of the model is simple:

$$\left.\begin{array}{l} y_t = x_t\beta + u_t \\ u_t = \rho_t \mathbf{W} u_t + \varepsilon_t \quad \Rightarrow \mathbf{B}_t u_t = \varepsilon_t \end{array}\right\} \\ \mathbf{B}_t = I_R - \rho_t \mathbf{W} \tag{11}$$

More concisely:

$$y = X\beta + u \;\; ; \; \mathbf{B}u = \varepsilon \text{ with } \varepsilon \sim N(0, \boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes I_R) \text{ and } \mathbf{B} = I_{TR} - \Upsilon \otimes \mathbf{W} \tag{12}$$

The logarithm of the likelihood function:

$$l(y;\theta) = -\frac{RT}{2}\ln(2\pi) \\ -\frac{R}{2}\ln|\boldsymbol{\Sigma}| + \sum\nolimits_{t=1}^{T} \ln|\mathbf{B}_t| - \frac{(y - X\beta)'\mathbf{B}'(\boldsymbol{\Sigma} \otimes I_R)^{-1}\mathbf{B}(y - X\beta)}{2} \tag{13}$$

where $\theta' = [\beta'; \rho_1; \cdots; \rho_T; vech(\Sigma)']$ is a vector of parameters, of order $(k + T + T(T+1)/2) \times 1$. More details appear in Section A.V of the Appendix.

The null hypothesis that there is no SEM structure in the error terms of the SUR means that:

$$H_0\colon \rho_t = 0 \;\; (\forall t) \text{ vs } H_A\colon \text{No } H_0 \tag{14}$$

The obtaining of the corresponding Multiplier is straightforward (Section A.VI of the Appendix):

$$\mathbf{LM}_{\text{SEM}}^{\text{SUR}} = g'_{(\rho)_{|H_0}} [I_{\rho\rho}]^{-1} g_{(\rho)_{|H_0}} \underset{as}{\sim} \chi^2(T) \tag{15}$$

As in previous cases, a more compact expression is proposed for the term of the score vector: $g_{(\rho)_{|H_0}} = [\boldsymbol{\Sigma}^{-1} \circ (\hat{U}'\hat{U}_L)]'\tau$ and $I_{\rho\rho} = tr(\mathbf{W}'\mathbf{W})\, I_R + tr(\mathbf{W}\mathbf{W})\, \boldsymbol{\Sigma}^{-1} \circ \boldsymbol{\Sigma}$.

3. The Robust Multipliers and Some Other Extensions

The Lagrange Multipliers developed in the second section are flexible so they can be adapted to different situations. It is worth mentioning that they depend only on the estimation of the SUR model under the null hypothesis of no spatial effects, which can be solved using standard techniques.

Several extensions of the basic discussion of Section 2 follow. Firstly, we obtain the robust version of the raw Lagrange Multipliers; then we move to the problem of the lack of constancy of the parameters of spatial dependence, considering a temporal perspective. The final part of the section focuses on a simpler, but important, question of SUR modelling—that of testing for the diagonality of the matrix $\boldsymbol{\Sigma}$.

3.1. The Robust Multipliers

The tests of the second section will help us to improve the specification of spatial SUR models. The weakness of this range of Lagrange Multipliers is that they are not robust to misspecification errors in the alternative hypothesis (Davidson, 2000). A consequence is that it will be more difficult to identify the type of misspecification that affects the model (Bera & Yoon, 1993). In this section, we follow the reasoning of Anselin *et al.* (1996) in search of robustness for the raw $\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$ and $\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$ tests. A brief introduction to the problem follows.

The likelihood function of the SARAR model discussed in Section 2.1, $L[\varphi;\lambda;\rho]$, depends on three groups of parameters: those associated with the basic SUR structure, $\varphi' = [\beta';\mathrm{vech}(\Sigma)]$, those related to the spatial lag of the endogenous variable, $\lambda' = [\lambda_1;\cdots;\lambda_T]$, and those that introduce spatial dependence into the error terms, $\rho' = [\rho_1;\cdots;\rho_T]$. If, for example, vector ρ is zero, the likelihood function collapses in $L_1[\varphi;\lambda]$ and the unidirectional test, which null hypothesis is that $H_0{:}\,\lambda = 0$, leads to the $\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$ statistic of Equation (10), distributed as a centred chi-squared $\chi^2(T,0)$. Under alternative hypotheses of the type $H_A{:}\lambda = \xi/\sqrt{R}$, with $\xi \neq 0$, this distribution will have a non-centrality parameter, $\chi^2(T,\pi_1)$, equal to $\pi_1 = \xi' I_{\lambda\cdot\varphi}\xi$ with $I_{\lambda\cdot\varphi} = I_{\lambda\lambda} - I_{\lambda\varphi} I_{\varphi\varphi}^{-1} I_{\varphi\lambda}$. The presence of this parameter is what gives strength to the test.

If the true likelihood function is assumed to be $L_2[\varphi;\rho]$, where λ is zero and ρ is different to zero, the $\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$ statistic will have a non-centrality bias even if the null hypothesis, $H_0{:}\,\lambda = 0$, stands. In particular, for alternatives of the type $H_A{:}\rho = \zeta/\sqrt{R};\zeta \neq 0$, the asymptotic distribution of the test statistic is a $\chi^2(T,\pi_2)$, where $\pi_2 = \zeta' I_{\lambda\rho\cdot\varphi} I_{\lambda\cdot\varphi}^{-1} I_{\lambda\rho\cdot\varphi}\zeta$ with $I_{\lambda\rho\cdot\varphi} = I_{\rho\lambda} - I_{\lambda\varphi} I_{\varphi\varphi}^{-1} I_{\varphi\rho}$. This factor of non-centrality will be positive if the terms of the score associated with λ and ρ ($g_{(\lambda)_{|H_0}}$ and $g_{(\rho)_{|H_0}}$, respectively) are correlated. Under the composite hypothesis that the two sets of parameters are zero, expressions (A9) and (A10) show that $I_{\lambda\rho}$ is a diagonal matrix. Given that matrix $I_{\varphi\lambda}$ is also different to zero but that $I_{\varphi\rho} = 0$, then $I_{\lambda\rho\cdot\varphi} = I_{\rho\lambda} \geq 0$. The consequence is that the parameter of non-centrality π_2 will be positive and the $\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$ statistic will unduly reject the hypothesis that the vector of parameters λ is zero when the data have been generated by an SUR with an SEM structure.

The discussion for the $\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$ statistic of Equation (15), under the assumption that the data have been generated by $L_1[\varphi;\lambda]$, is analogous. To be precise, under alternatives of the type $H_A{:}\lambda = \zeta/\sqrt{R};\zeta \neq 0$, the asymptotic distribution of this statistic is $\chi^2(T,\pi_3)$, where $\pi_3 = \zeta' I_{\lambda\rho\cdot\varphi} I_{\rho\cdot\varphi}^{-1} I_{\lambda\rho\cdot\varphi}\zeta$ with $I_{\rho\cdot\varphi} = I_{\rho\rho} - I_{\rho\varphi} I_{\varphi\varphi}^{-1} I_{\varphi\rho}$. For the same reasons as previously highlighted, it will be true that $I_{\lambda\rho\cdot\varphi} = I_{\rho\lambda} \geq 0$ and $\pi_3 > 0$. This result induces the $\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$ statistic to wrongly reject the hypothesis that the vector of parameters ρ is zero when the data have been generated by an SUR with an SLM structure.

The proposal by Bera & Yoon (1993, p. 652) 'is to construct a size-resistant test … where we adapt the statistic for the nuisance parameters'. The correction consists of adjusting the bias that affects both the score and the covariance matrix that intervene in the raw Lagrange Multiplier. In our case:

$$\left.\begin{array}{l} H_0{:}\lambda_t = 0 \\ H_A{:}No\ H_0 \end{array}\right\}$$

$$\mathbf{LM}^{*\ \mathrm{SUR}}_{\mathrm{SLM}} = [g_{(\lambda)_{|H_0}} - I_{\lambda\rho\cdot\varphi} I^{-1}_{\rho\cdot\varphi} g_{(\rho)_{|H_0}}]'[I_{\lambda\cdot\varphi} - I_{\lambda\rho\cdot\varphi} I^{-1}_{\rho\cdot\varphi} I_{\lambda\rho\cdot\varphi}]^{-1}$$
$$[g_{(\lambda)_{|H_0}} - I_{\lambda\rho\cdot\varphi} I^{-1}_{\rho\cdot\varphi} g_{(\rho)_{|H_0}}] \underset{\mathrm{as}}{\sim} \chi^2(T) \tag{16}$$

where:

$$g_{(\rho)_{|H_0}} = [\mathbf{\Sigma}^{-1} \circ (\hat{U}'\hat{U}_L)]'\tau$$
$$g_{(\lambda)_{|H_0}} = [\mathbf{\Sigma}^{-1} \circ (\hat{U}'Y_L)]'\tau$$
$$I_{\lambda\rho\cdot\varphi} = I_{\rho\lambda} \qquad I_{\rho\cdot\varphi} = I_{\rho\rho}$$
$$I_{\lambda\rho\cdot\varphi} = I_{\rho\lambda} \quad I_{\lambda\cdot\varphi} = I_{\lambda\lambda} - I_{\lambda\varphi} I^{-1}_{\varphi\varphi} I_{\lambda\varphi}$$
$$I^{-1}_{\varphi\varphi} = \begin{bmatrix} I_{\beta\beta} & 0 \\ & I_{\sigma\sigma} \end{bmatrix}^{-1} \tag{17}$$

The information needed to obtain the value of Equation (16) appears in expressions (A9) and (A10) in the Appendix. Likewise:

$$\left.\begin{matrix} H_0{:}\rho_t = 0 \\ H_A{:}No H_0 \end{matrix}\right\}$$

$$\mathbf{LM}^{*\ \mathrm{SUR}}_{\mathrm{SEM}} = [g_{(\rho)_{|H_0}} - I_{\lambda\rho\cdot\varphi} I^{-1}_{\lambda\cdot\varphi} g_{(\lambda)_{|H_0}}]'[I_{\rho\cdot\varphi} - I_{\lambda\rho\cdot\varphi} I^{-1}_{\lambda\cdot\varphi} I_{\lambda\rho\cdot\varphi}]^{-1}[g_{(\rho)_{|H_0}} - I_{\lambda\rho\cdot\varphi} I^{-1}_{\lambda\cdot\varphi} g_{(\lambda)_{|H_0}}]$$
$$\times \underset{\mathrm{as}}{\sim} \chi^2(T) \tag{18}$$

where:

$$g_{(\rho)_{|H_0}} = [\mathbf{\Sigma}^{-1} \circ (\hat{U}'\hat{U}_L)]'\tau$$
$$g_{(\lambda)_{|H_0}} = [\mathbf{\Sigma}^{-1} \circ (\hat{U}'Y_L)]'\tau$$
$$I_{\lambda\rho\cdot\varphi} = I_{\rho\lambda} \quad I_{\lambda\cdot\varphi} = I_{\lambda\lambda} - I_{\lambda\varphi} I^{-1}_{\varphi\varphi} I_{\lambda\varphi}$$

$$I_{\lambda\varphi} = [I_{\lambda\beta} \quad 0] \;\; I^{-1}_{\varphi\varphi} = \begin{bmatrix} I_{\beta\beta} & 0 \\ & I_{\sigma\sigma} \end{bmatrix}^{-1} \tag{19}$$

Anselin *et al.* (1996) show that the corrections suggested by Bera & Yoon (1993) work well in relation to the spatial Lagrange Multipliers obtained from a single cross-section estimation. Our expectancy is that the same will happen in an SUR context.

3.2. Testing for the Time-Constancy of the Parameters of Spatial Dependence

The coefficients of spatial dependence of the SUR specifications of Section 2 are allowed to vary in each cross-section. If these cross-sections refer to consecutive time periods, it is reasonable to question if these coefficients remain the same over time.

In order to further appreciate the testing procedure for this case, let us first of all introduce the null model. Under the restriction of time constancy of the spatial coefficients, the SARAR model of Equation (2) becomes:

$$\left.\begin{array}{l}\mathbf{A}y = X\beta + u \\ \mathbf{B}u = \varepsilon \\ \varepsilon \sim N(0, \boldsymbol{\Omega}) \Rightarrow \boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes I_R\end{array}\right\}$$

$$\mathbf{A} = I_{TR} - \lambda \otimes \mathbf{W} = I_T \otimes (I_R - \lambda\mathbf{W}) = I_T \otimes \hat{\mathbf{A}}$$

$$\mathbf{B} = I_{TR} - \rho \otimes \mathbf{W} = I_T \otimes (I_R - \rho\mathbf{W}) = I_T \otimes \hat{\mathbf{B}} \tag{20}$$

where $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ are two matrices of order $(R \times R)$. The number of parameters to estimate reduces to $(k+2+T(T+1)/2)$. If we adopt an SLM model, the restricted version is:

$$\left.\begin{array}{c}\mathbf{A}y = X\beta + \varepsilon \\ \varepsilon \sim N(0, \boldsymbol{\Omega}) \Rightarrow \boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes I_R\end{array}\right\}$$

$$\mathbf{A} = I_{TR} - \lambda \otimes \mathbf{W} = I_T \otimes (I_R - \lambda\mathbf{W}) = I_T \otimes \hat{\mathbf{A}} \tag{21}$$

The number of parameters to estimate is $(k+1+T(T+1)/2)$, the same as in the SEM case, under the assumption of constancy:

$$\left.\begin{array}{c}y = X\beta + u \\ \mathbf{B}u = \varepsilon \\ \varepsilon \sim N(0, \boldsymbol{\Omega}) \Rightarrow \boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes I_R\end{array}\right\}$$

$$\mathbf{B} = I_{TR} - \rho \otimes \mathbf{W} = I_T \otimes (I_R - \rho\mathbf{W}) = I_T \otimes \hat{\mathbf{B}} \tag{22}$$

It is evident that the hypothesis of temporal homogeneity in the spatial dependence parameters has important consequences, and therefore it should be tested adequately. One simple solution is the likelihood ratio which compares the likelihood of the ample model against that obtained with its restricted version (that is, the model of Equation (2) against the model of Equation (20) in the SARAR case; the model of Equation (7) against that of Equation (21) in the SLM case and the model of Equation (11) against that of Equation (22) in the SEM case). This procedure requires the estimation of the null hypothesis model and the alternative hypothesis model. The Lagrange Multiplier is a much more acceptable procedure because only the estimation of the restricted model is needed. The main results are shown below.

As mentioned, in the SARAR case, we compare the unrestricted specification of expression (2) with the restricted version of Equation (20). These two models are related by a set of $2(T–1)$ restrictions that lead to the composite null hypothesis:

$$H_0\colon \lambda_1 = \lambda_2 = \ldots = \lambda_T \text{ and } \rho_1 = \rho_2 = \ldots = \rho_T \text{ vs } H_A\colon \text{No } H_0 \tag{23}$$

The expression of the Lagrange Multiplier is the quadratic form of the score vector as expected (which appears in expression (A3)) on the inverse of the information matrix (as indicated in expressions (A5) and (A6) in the Appendix), in both cases having been evaluated under the null hypothesis of Equation (23). That is:

$$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SARAR}}(\lambda, \rho) = [g(\theta)_{|H_0}]'[I^{T,R}(\theta)_{|H_0}]^{-1}[g(\theta)_{|H_0}] \underset{\mathrm{as}}{\sim} \chi^2(2(T-1)) \tag{24}$$

with:

$$g(0)_{|H_0} = \begin{bmatrix} \dfrac{\partial l}{\partial \beta} \\ \dfrac{\partial l}{\partial \lambda_t}_{t=1,\dots T} \\ \dfrac{\partial l}{\partial \rho_t}_{t=1,\dots T} \\ \dfrac{\partial l}{\partial \sigma_{ij}}_{i,j=1,\dots T} \end{bmatrix}_{|H_0} = \begin{bmatrix} g_{(\beta)_{|H_0}} \\ g_{(\lambda)_{|H_0}} \\ g_{(\rho)_{|H_0}} \\ g_{(\Sigma)_{|H_0}} \end{bmatrix} = \begin{bmatrix} 0 \\ -\mathrm{tr}(\hat{\mathbf{A}}^{-1}\mathbf{W}) + \hat{\varepsilon}'[(\boldsymbol{\Sigma}^{-1}\mathbf{E}^{tt}) \otimes (\hat{\mathbf{B}}\mathbf{W})]\gamma_{\,t=1,\dots T} \\ -\mathrm{tr}(\hat{\mathbf{B}}^{-1}\mathbf{W}) + \hat{\varepsilon}'[(\boldsymbol{\Sigma}^{-1}\mathbf{E}^{tt}) \otimes \mathbf{W}]\hat{u}_{\,t=1,\dots T} \\ 0 \end{bmatrix} \tag{25}$$

In the case of the SLM of Equation (6), there are (T–1) restrictions in the null hypothesis:

$$H_0\colon \lambda_1 = \lambda_2 = \ \dots = \lambda_T \text{ vs } H_A\colon \text{No } H_0 \tag{26}$$

The expression of the Multiplier does not vary:

$$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SLM}}(\lambda) = [g(\theta)_{|H_0}]'[I^{T,R}(\theta)_{|H_0}]^{-1}[g(\theta)_{|H_0}] \underset{\mathrm{as}}{\sim} \chi^2(T-1) \tag{27}$$

with:

$$g(\theta)_{|H_0} = \begin{bmatrix} \dfrac{\partial l}{\partial \beta} \\ \dfrac{\partial l}{\partial \lambda_t}_{t=1,\dots T} \\ \dfrac{\partial l}{\partial \sigma_{ij}}_{i,j=1,\dots T} \end{bmatrix} = \begin{bmatrix} g_{(\beta)_{|H_0}} \\ g_{(\lambda)_{|H_0}} \\ g_{(\Sigma)_{|H_0}} \end{bmatrix} = \begin{bmatrix} 0 \\ -\mathrm{tr}(\hat{\mathbf{A}}^{-1}\mathbf{W}) + \sum_{s=1}^{R} \sigma^{st}\hat{\varepsilon}_s'\mathbf{W}\gamma_t \;{}_{t=1,\dots T} \\ 0 \end{bmatrix} \tag{28}$$

Lastly, in the SEM of Equation (11) we obtain:

$$H_0\colon \rho_1 = \ \rho_2 = \dots = \rho_T \text{ vs } H_A\colon \text{No } H_0 \tag{29}$$

with:

$$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SEM}}(\rho) = [g(\theta)_{|H_0}]'[I^{T,R}(\theta)_{|H_0}]^{-1}[g(\theta)_{|H_0}] \underset{\mathrm{as}}{\sim} \chi^2(T-1) \tag{30}$$

where:

$$g(\theta)_{|H_0} = \begin{bmatrix} \dfrac{\partial l}{\partial \beta} \\ \dfrac{\partial l}{\partial \rho_t} \\ {\scriptstyle t=1,\ldots T} \\ \dfrac{\partial l}{\partial \sigma_{ij}} \\ {\scriptstyle i,j=1,\ldots T} \end{bmatrix}_{|H_0} = \begin{bmatrix} g_{(\beta)_{|H_0}} \\ g_{(\rho)_{|H_0}} \\ g_{(\Sigma)_{|H_0}} \end{bmatrix} = \begin{bmatrix} 0 \\ -\mathrm{tr}(\hat{\mathbf{B}}^{-1}\mathbf{W}) + \sum_{s=1}^{R} \sigma^{st} \hat{\varepsilon}_s' \mathbf{W} \hat{u}_t \\ {\scriptstyle t=1,\ldots T} \\ 0 \end{bmatrix} \tag{31}$$

More details can be found in Section A.VII of the Appendix.

3.3. Testing for the Diagonality of the Matrix Σ

Up to now, our concern has been analysing the spatial structure of the SUR model. In this section we focus on the assumption of non-diagonality of matrix $\mathbf{\Sigma}$. If the hypothesis of diagonality cannot be rejected, the SUR model simplifies in a set of (unrelated) cross-sections. On the other hand, the non-diagonality of matrix $\mathbf{\Sigma}$ reinforces the need to also consider the temporal evolution of the spatial system.

This problem is often referred to in SUR literature where different proposals can be found. Among them, those that best fit our approach are the Likelihood Ratio and the Lagrange Multiplier tests (Breusch & Pagan, 1980). One aspect to bear in mind is that, in this case, both tests must be used in models where a certain spatial structure has been introduced. For example, in the case of the SARAR of Equation (2):

$$\left.\begin{array}{l} \mathbf{A}y = X\beta + u \\ \mathbf{B}u = \varepsilon \\ \varepsilon \sim N(0,\mathbf{\Omega}) \Rightarrow \mathbf{\Omega} = \Sigma \otimes I_R \end{array}\right\} \tag{32}$$

matrix $\mathbf{\Sigma}$ describes the temporal correlation that exists between the errors of the T cross-sections once the spatial structure that intervenes in the specification has been filtered. Then, the test statistic is obtained in the usual way:

$$\begin{array}{ll} H_0: & \mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 & \ldots & 0 \\ 0 & \sigma_1^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \sigma_T^2 \end{bmatrix} \Bigg\} \to LM_{\Sigma} \\ H_A: & \qquad \text{No } H_0 \end{array}$$

$$= R \sum_{t=1}^{T} \sum_{s=1}^{t-1} r_{ts}^2 \underset{\mathrm{as}}{\sim} \chi^2(T(T-1)/2) \tag{33}$$

with:

$$r_{ts} = \frac{\sum_{r=1}^{R} (\hat{\varepsilon}_{tr} - \hat{\bar{\varepsilon}}_t)(\hat{\varepsilon}_{sr} - \hat{\bar{\varepsilon}}_s)}{\sqrt{\sum_{r=1}^{R} (\hat{\varepsilon}_{tr} - \hat{\bar{\varepsilon}}_t)^2} \sqrt{\sum_{r=1}^{R} (\hat{\varepsilon}_{sr} - \hat{\bar{\varepsilon}}_s)^2}} \tag{34}$$

where r_{ts} is the coefficient of correlation between the residuals of cross-sections t and s. The Likelihood Ratio is defined as usual:

$$LR_{\Sigma} = 2\left[l(y;\theta)|_{H_A} - l(y;\theta)|_{H_0} \right] = R\left[\ln|\tilde{\Sigma}_{H_0}| - \ln|\tilde{\Sigma}_{H_A}| \right] \underset{as}{\sim} \chi^2(T(T-1)/2) \quad (35)$$

with $l(y;\theta)|_{H_0}$ and $l(y;\theta)|_{H_A}$ being the value of the log-likelihood function under the null and alternative hypotheses, respectively, and $\tilde{\Sigma}_{H_0}$ and $\tilde{\Sigma}_{H_A}$ being the corresponding covariance matrices. Under mild conditions, the two statistics converge asymptotically to the χ^2 with $T(T-1)/2$ degrees of freedom (as shown, for example, in Breusch & Pagan, 1980). Breusch & Pagan (1980) use the least squares residuals given that they discuss the case of a static model, but this isn't so in our case. In the framework of Equation (33), and under the null hypothesis of diagonality, there is a panel of ML residuals obtained from the estimation of a spatial model (SARAR, SLM or SEM) for each of the T cross-sections. In other words, if we introduce spatial dependencies in the cross-sections, obtaining the asymptotic distribution under the null of diagonality is not straightforward. As a useful approximation, Lee (2004) shows that the (quasi) maximum likelihood residuals of a correctly specified (spatial) econometric model can be estimated, asymptotically with R, to a Gaussian random variable. Under this setting, the test statistic of Equation (33) will converge to the chi-squared distribution, as assumed.

4. A Monte Carlo Study: Design and Empirical Results

In this section we present the results of an extensive Monte Carlo exercise which focus on the tests developed in Sections 2 and 3. The first sub-section examines the robustness of the diagonality test in the presence of spatial dependence in SURE models. The second sub-section deals with the behaviour of the raw and robust Lagrange Multipliers. The third part of the Monte Carlo study considers the test of instability in the spatial dependence parameters.

The design of the experiment considers the factors that, in our opinion, may have greater impact on the tests: the structure of temporal dependence among the errors of the cross-sections, the intensity of the spatial dependence, its stability (or lack of stability) and the size of the sample.

To evaluate the power of the test, two types of alternatives are considered:

$$\left.\begin{array}{l} \text{SLM:} y_t = (I_R - \rho_t \mathbf{W})^{-1}(\alpha + \beta x_t + e_t);\ t = 1, \ldots, T \\ \text{SEM:} y_t = \alpha + \beta x_t + (I_R - \lambda_t \mathbf{W})^{-1} e_t;\ t = 1, \ldots, T \end{array}\right\} \quad (36)$$

The structure of the equations is simple: they contain an intercept, with a value of 2 ($\alpha_t = 2$), and only one regressor (x_t) drawn from an $N(0,1)$ distribution. The value of the coefficient β is 3, which assures that, in the absence of spatial effects ($\rho_t = 0/\lambda_t = 0$), the R^2 coefficient of the Least Square (LS) regression of the equation will be close to 0.9.

The error terms $e = (e_1, \ldots, e_T)$ have been obtained from a normal multivariant distribution, centred on zero and the covariance matrix has been defined as:

$$\Sigma_k = (\sigma_{k,ij}) \text{ with } \sigma_{k,ij} = 0.95^{s_k|i-j|} \quad (37)$$

The parameter s_k determines the intensity of the dependency between the residuals of the equations i and j. Three values for s_k have been used, (10^6, 20, 0.2), which introduce different levels of time dependence: null ($s_1 = 10^6$; $\mathbf{\Sigma}_1 = I_T$); moderate ($s_2 = 20$; in matrix $\mathbf{\Sigma}_2$) and extreme ($s_3 = 0.2$; in matrix $\mathbf{\Sigma}_3$).

Moreover, we have simulated different values for T, namely $T = 2$, 4 and 10, combined with two different sample sizes corresponding to regular lattices of (7×7), $R = 49$, and (9×9), $R = 81$, respectively. For each cross-section a row-standardized weighting matrix has been defined using a rook scheme. Each experiment has been repeated 2,000 times.

4.1. The Diagonality Test

There are several studies where the hypothesis of non-diagonality of the matrix $\mathbf{\Sigma}$ has been revised (Shiba & Tsurumi, 1988; Kurata, 2001; Dufour & Kahlaf, 2002; Tsay, 2004). This assumption between the errors of the different equations is crucial as it clarifies the SUR specification. In this part of the paper we highlight the size problems of the two most popular tests, LR and LM, when a spatial dependency structure is present in the cross-sections. The design of our experiment considers only one regressor and a balanced relation between R and T.

In spite of this favourable outcome, Table 1 and Figure 1 show that both tests have size problems in the presence of spatial dependence. If there is independence among the SUR equations, the LR_{Σ} test appears to slightly overestimate the 5% nominal size, especially when the number of cross-section ($T = 10$) increases. This result is in accordance with well documented literature on LR-based multivariant tests. Under spatial dependence, both tests, LM_{Σ} and LR_{Σ}, have serious problems with empirical size, even for low to moderate values of the spatial dependence parameter. There is a strong tendency to over-reject high values of the spatial dependence parameter, whatever the type of process (SLM or SEM) used. The number of cross-sections (T) also has clear negative effects.

If the cross-sections repeat a similar pattern of spatial dependence, a false impression emerges of time persistence in the values corresponding to each individual. The tests interpret this situation as a symptom of temporal dependence. In order to correct this weakness, as indicated in Section 3.4, it is first necessary to filter the spatial structure in the different cross-section. By doing this, assumption of diagonality in the $\mathbf{\Sigma}$ matrix can be properly analysed.

4.2. The Tests of Spatial Dependence

Below we present the size and power of the spatial dependence tests developed in Sections 2 and 3. The empirical size of the test, at a 5% nominal level, appears in Table 2. The estimates are very close to the nominal value; any tendency to over or underestimate the size can be seen.

Tables 3, 4 and 5 present the results corresponding to the power of the tests, also at a nominal value of 5%. Case 1 combines time stability in the spatial dependence coefficients and temporal independence among the SUR equations. Case 2 maintains the situation of time stability of the spatial parameters but introduces an extreme temporal correlation in the SUR structure. Cases 3 and 4 are characterized by the time heterogeneity of the spatial dependence parameters. In Case 3, the T cross-sections are independent, as in Case 1, but spatial dependence only exists in the first cross-section (the remaining cross-sections are spatially independent). Case 4 combines a very strong correlation in the SUR equations with the instability in the spatial coefficients of Case 3.

In summary, we present only the results for the case where $R = 81$ $(9 \times 9$ lattice) for a representative range of values of the spatial correlation coefficient.

Table 1. Empirical size of the LM and LR test in spatial dependent processes ($R = 81$)⋆

	SLM model						SEM model					
	$T=2$		$T=4$		$T=10$		$T=2$		$T=4$		$T=10$	
ρ	LM_{Σ}	LR_{Σ}	LM_{Σ}	LR_{Σ}	LM_{Σ}	LR_{Σ}	LM_{Σ}	LR_{Σ}	LM_{Σ}	LR_{Σ}	LM_{Σ}	LR_{Σ}
0.00	0.061	0.065	0.051	0.055	0.058	0.088	0.061	0.065	0.051	0.055	0.058	0.088
0.10	0.051	0.054	0.047	0.060	0.066	0.103	0.054	0.055	0.048	0.059	0.069	0.107
0.20	0.065	0.07	0.073	0.085	0.108	0.161	0.059	0.061	0.073	0.082	0.074	0.128
0.30	0.073	0.079	0.097	0.107	0.207	0.274	0.066	0.069	0.084	0.093	0.142	0.192
0.40	0.076	0.078	0.261	0.286	0.516	0.600	0.088	0.093	0.105	0.120	0.251	0.322
0.50	0.173	0.186	0.315	0.348	0.959	0.975	0.102	0.107	0.164	0.185	0.500	0.582
0.60	0.213	0.235	0.377	0.415	0.978	0.988	0.133	0.142	0.277	0.302	0.799	0.848
0.70	0.289	0.296	0.538	0.578	1.000	1.000	0.191	0.198	0.445	0.473	0.980	0.986
0.80	0.719	0.729	0.811	0.821	1.000	1.000	0.279	0.286	0.728	0.740	1.000	1.000
0.90	0.798	0.788	0.992	0.992	1.000	1.000	0.477	0.485	0.935	0.939	1.000	1.000
0.95	0.842	0.848	0.994	0.995	1.000	1.000	0.592	0.596	0.990	0.991	1.000	1.000

Note: ⋆ Similar results are obtained for other values of R.

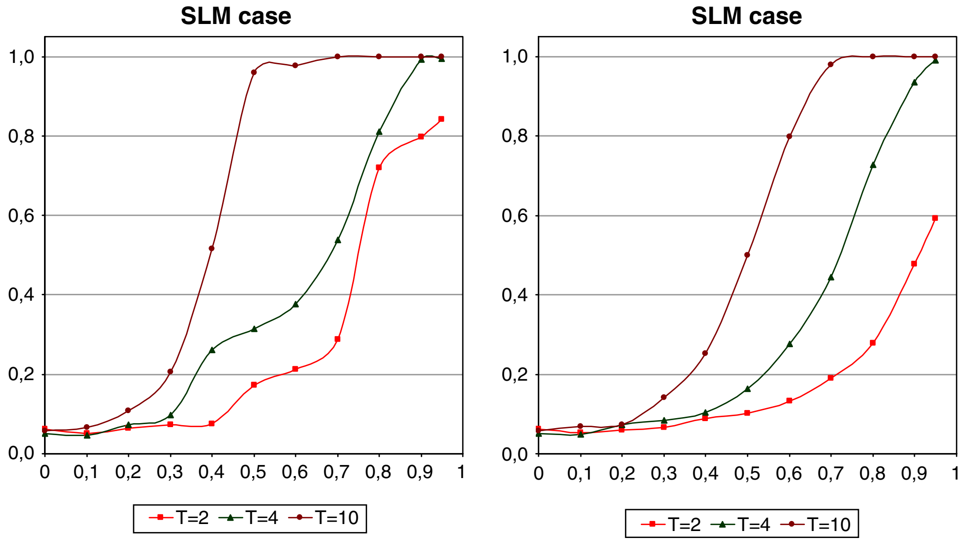

Figure 1. Empirical size of the LM_Σ and LR_Σ tests in spatially dependent processes.

Anselin *et al.* (1996) study the behaviour of the Lagrange Multiplier for an omitted spatial lag of the endogenous variable, using one single cross-section, referred to as LM_ϕ. This test is equivalent to our $\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$. Anselin *et al.* (1996) report an empirical power of 0.463 for the LM_ϕ with $\rho = 0.1$ while in our experiment, and for $T = 10$, Case 1, the empirical power of the equivalent test raises to 0.861 (0.543 with $T = 2$). This increment is due solely to the use of several cross-sections. Moreover, the power of the test grows faster when there is time dependence among the equations. For example, in Case 2 and using 10 cross-sections, the power function attains its maximum value of 1 with a very low value of the parameter, 0.02 when $T = 10$.

Table 2. Empirical size of the spatial dependence tests, 5% nominal level

		R	$\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$
$T = 2$	Σ_1	49	0.045	0.048	0.056	0.055	0.043
		81	0.053	0.049	0.045	0.044	0.046
	Σ_2	49	0.048	0.051	0.056	0.058	0.053
		81	0.050	0.047	0.053	0.052	0.051
	Σ_3	49	0.050	0.051	0.056	0.056	0.052
		81	0.048	0.049	0.052	0.057	0.057
$T = 4$	Σ_1	49	0.045	0.041	0.056	0.050	0.048
		81	0.042	0.042	0.051	0.052	0.047
	Σ_2	49	0.043	0.046	0.053	0.056	0.048
		81	0.050	0.053	0.054	0.056	0.051
	Σ_3	49	0.041	0.044	0.055	0.053	0.051
		81	0.043	0.048	0.053	0.056	0.049
$T = 10$	Σ_1	49	0.036	0.040	0.055	0.060	0.047
		81	0.044	0.040	0.049	0.051	0.047
	Σ_2	49	0.041	0.040	0.062	0.063	0.049
		81	0.044	0.045	0.055	0.052	0.047
	Σ_3	49	0.038	0.040	0.070	0.075	0.056
		81	0.045	0.046	0.067	0.069	0.059

Note: See expression (37) for the meaning of Σ_j ($j = 1,2,3$).

Table 3. Empirical power $T=2$; $R=81$. Percentage of rejections of the null hypothesis, 5% sig. level

SLM model						SEM model					
Case 1: Temporal independence ($\Sigma_1 = I_T$) and stability in spatial dependence $\rho_i = \rho_j$ $(\forall i, j)$											
ρ	$\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$	$\boldsymbol{\lambda}$	$\mathbf{LM}_{SEM}^{SUR}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$
0.02	0.044	0.046	0.065	0.075	0.054	**0.10**	0.075	0.063	0.071	0.056	0.071
0.06	0.058	0.053	0.239	0.228	0.175	**0.20**	0.289	0.236	0.097	0.065	0.215
0.10	0.079	0.049	0.545	0.523	0.424	**0.30**	0.609	0.482	0.190	0.064	0.500
0.14	0.123	0.057	0.732	0.681	0.617	**0.40**	0.897	0.786	0.380	0.063	0.829
0.16	0.181	0.048	0.814	0.735	0.708	**0.50**	0.986	0.961	0.532	0.102	0.975
0.20	0.306	0.040	0.973	0.943	0.939	**0.60**	0.999	0.995	0.725	0.102	0.998
0.30	0.501	0.051	1.000	1.000	1.000	**0.70**	1.000	1.000	0.911	0.109	1.000
0.40	0.775	0.048	1.000	1.000	1.000	**0.80**	1.000	1.000	0.996	0.129	1.000
0.90	1.000	0.059	1.000	1.000	1.000	**0.90**	1.000	1.000	0.997	0.378	1.000
Case 2: Extreme temporal correlation Σ_3 and stability in spatial dependence $\rho_i = \rho_j$ $(\forall i, j)$											
ρ	$\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$	$\boldsymbol{\lambda}$	$\mathbf{LM}_{SEM}^{SUR}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$
0.02	0.046	0.053	0.643	0.652	0.530	**0.10**	0.073	0.078	0.047	0.053	0.067
0.06	0.188	0.086	1.000	1.000	1.000	**0.20**	0.263	0.271	0.068	0.065	0.211
0.10	0.077	0.079	1.000	1.000	1.000	**0.30**	0.626	0.612	0.068	0.061	0.523
0.14	0.195	0.058	1.000	1.000	1.000	**0.40**	0.893	0.889	0.079	0.080	0.833
0.16	0.238	0.066	1.000	1.000	1.000	**0.50**	0.983	0.983	0.127	0.117	0.962
0.20	0.383	0.050	1.000	1.000	1.000	**0.60**	1.000	0.999	0.160	0.106	0.998
0.30	0.551	0.051	1.000	1.000	1.000	**0.70**	1.000	1.000	0.195	0.185	1.000
0.40	0.792	0.049	1.000	1.000	1.000	**0.80**	1.000	1.000	0.284	0.194	1.000
0.90	1.000	0.057	1.000	1.000	1.000	**0.90**	1.000	1.000	0.459	0.265	1.000

Case 3: Temporal independence ($\Sigma_1 = I_T$) and instability in spatial dependence $\rho_1 \neq 0;\ \ \rho_i = 0\ \ (i > 1)$

ρ	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SARAR}}$	λ	$\mathbf{LM}^{SUR}_{SEM}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SARAR}}$
0.02	0.042	0.047	0.059	0.061	0.055	**0.10**	0.068	0.067	0.060	0.052	0.064
0.06	0.041	0.050	0.079	0.075	0.060	**0.20**	0.174	0.153	0.069	0.056	0.137
0.10	0.062	0.047	0.223	0.196	0.161	**0.30**	0.355	0.279	0.131	0.060	0.284
0.14	0.076	0.039	0.620	0.590	0.502	**0.40**	0.617	0.506	0.192	0.075	0.515
0.16	0.108	0.043	0.502	0.438	0.384	**0.50**	0.843	0.669	0.385	0.070	0.751
0.20	0.211	0.040	0.871	0.822	0.793	**0.60**	0.942	0.891	0.440	0.091	0.909
0.30	0.631	0.048	1.000	1.000	1.000	**0.70**	0.993	0.981	0.561	0.104	0.983
0.40	0.883	0.055	1.000	1.000	1.000	**0.80**	1.000	0.995	0.830	0.257	0.999
0.90	1.000	0.057	1.000	1.000	1.000	**0.90**	1.000	1.000	0.985	0.101	1.000

Case 4: Extreme temporal correlation (Σ_3) and instability in spatial dependence $\rho_1 \neq 0;\ \ \rho_i = 0\ \ (i > 1)$

ρ	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SARAR}}$	λ	$\mathbf{LM}^{SUR}_{SEM}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SARAR}}$
0.02	0.074	0.047	0.460	0.429	0.343	**0.10**	0.755	0.637	0.253	0.082	0.665
0.06	0.266	0.064	1.000	1.000	1.000	**0.20**	0.997	0.986	0.632	0.112	0.993
0.10	0.486	0.062	1.000	1.000	1.000	**0.30**	1.000	1.000	0.786	0.161	1.000
0.14	0.463	0.071	1.000	1.000	1.000	**0.40**	1.000	1.000	0.791	0.202	1.000
0.16	0.683	0.052	1.000	1.000	1.000	**0.50**	1.000	1.000	0.945	0.204	1.000
0.20	0.697	0.062	1.000	1.000	1.000	**0.60**	1.000	1.000	0.948	0.217	1.000
0.30	0.775	0.058	1.000	1.000	1.000	**0.70**	1.000	1.000	0.946	0.268	1.000
0.40	0.821	0.059	1.000	1.000	1.000	**0.80**	1.000	1.000	0.921	0.342	1.000
0.90	0.991	0.061	1.000	1.000	1.000	**0.90**	1.000	1.000	0.995	0.260	1.000

Table 4. Empirical power $T=4$; $R=81$. Percentage of rejections of the null hypothesis, 5% sig. level

	SLM model						SEM model				
	Case 1: Temporal independence ($\Sigma_1 = I_T$) and stability in spatial dependence $\rho_i = \rho_j$ $(\forall i, j)$										
ρ	$\mathbf{LM}_{SEM}^{SUR}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{SLM}^{SUR}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$	$\boldsymbol{\lambda}$	$\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$
0.02	0.039	0.045	0.076	0.079	0.065	**0.10**	0.084	0.084	0.065	0.054	0.076
0.06	0.059	0.056	0.228	0.220	0.164	**0.20**	0.373	0.298	0.108	0.060	0.285
0.10	0.084	0.044	0.671	0.624	0.524	**0.30**	0.817	0.707	0.229	0.082	0.715
0.14	0.185	0.041	0.949	0.917	0.899	**0.40**	0.984	0.950	0.451	0.069	0.956
0.16	0.221	0.050	0.972	0.952	0.928	**0.50**	1.000	0.997	0.696	0.098	1.000
0.20	0.394	0.037	0.999	0.996	0.994	**0.60**	1.000	1.000	0.924	0.163	1.000
0.30	0.510	0.047	1.000	1.000	1.000	**0.70**	1.000	1.000	0.991	0.166	1.000
0.40	0.971	0.060	1.000	1.000	1.000	**0.80**	1.000	1.000	1.000	0.234	1.000
0.90	1.000	0.052	1.000	1.000	1.000	**0.90**	1.000	1.000	1.000	0.335	1.000
	Case 2: Extreme temporal correlation (Σ_3) and stability in spatial dependence $\rho_i = \rho_j$ $(\forall i, j)$										
ρ	$\mathbf{LM}_{SEM}^{SUR}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{SLM}^{SUR}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$	$\boldsymbol{\lambda}$	$\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$
0.02	0.049	0.059	0.990	0.988	0.963	**0.10**	0.090	0.099	0.051	0.055	0.083
0.06	0.099	0.084	1.000	1.000	1.000	**0.20**	0.359	0.349	0.070	0.069	0.291
0.10	0.168	0.094	1.000	1.000	1.000	**0.30**	0.817	0.814	0.073	0.065	0.711
0.14	0.139	0.141	1.000	1.000	1.000	**0.40**	0.987	0.987	0.102	0.089	0.967
0.16	0.292	0.088	1.000	1.000	1.000	**0.50**	1.000	1.000	0.122	0.125	0.998
0.20	0.837	0.096	1.000	1.000	1.000	**0.60**	1.000	1.000	0.175	0.146	1.000
0.30	0926	0.099	1.000	1.000	1.000	**0.70**	1.000	1.000	0.235	0.208	1.000
0.40	0.983	0.105	1.000	1.000	1.000	**0.80**	1.000	1.000	0.306	0.268	1.000
0.90	1.000	0.120	1.000	1.000	1.000	**0.90**	1.000	1.000	0.487	0.502	1.000

Case 3: Temporal independence ($\Sigma_1 = I_T$) and instability in spatial dependence $\rho_1 \neq 0; \quad \rho_i = 0 \quad (i > 1)$

ρ	$\mathbf{LM}_{SEM}^{SUR}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{SUR}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$	$\boldsymbol{\lambda}$	$\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$
0.02	0.048	0.050	0.053	0.050	0.047	**0.10**	0.061	0.059	0.059	0.053	0.059
0.06	0.052	0.046	0.112	0.107	0.088	**0.20**	0.127	0.104	0.065	0.051	0.100
0.10	0.055	0.042	0.180	0.161	0.119	**0.30**	0.259	0.200	0.088	0.049	0.185
0.14	0.065	0.042	0.277	0.228	0.186	**0.40**	0.506	0.415	0.126	0.056	0.398
0.16	0.091	0.054	0.268	0.229	0.201	**0.50**	0.770	0.547	0.348	0.062	0.646
0.20	0.124	0.054	0.702	0.621	0.558	**0.60**	0.898	0.775	0.396	0.095	0.850
0.30	0.218	0.055	0.885	0.752	0.722	**0.70**	0.984	0.952	0.477	0.097	0.959
0.40	0.340	0.044	1.000	0.881	0.859	**0.80**	0.998	0.979	0.848	0.099	0.995
0.90	0.577	0.058	1.000	0.991	0.910	**0.90**	1.000	0.986	0.968	0.245	1.000

Case 4: Extreme temporal correlation (Σ_3) and instability in spatial dependence $\rho_1 \neq 0; \quad \rho_i = 0 \quad (i > 1)$

ρ	$\mathbf{LM}_{SEM}^{SUR}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{SUR}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$	$\boldsymbol{\lambda}$	$\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$	$\mathbf{LM}_{\mathrm{SARAR}}^{\mathrm{SUR}}$
0.02	0.066	0.051	0.483	0.458	0.362	**0.10**	0.641	0.554	0.140	0.054	0.514
0.06	0.247	0.051	0.988	0.983	0.960	**0.20**	0.996	0.989	0.275	0.086	0.981
0.10	0.327	0.075	1.000	1.000	1.000	**0.30**	1.000	1.000	0.357	0.091	1.000
0.14	0.577	0.046	1.000	1.000	1.000	**0.40**	1.000	1.000	0.565	0.102	1.000
0.16	0.521	0.060	1.000	1.000	1.000	**0.50**	1.000	1.000	0.578	0.126	1.000
0.20	0.705	0.055	1.000	1.000	1.000	**0.60**	1.000	1.000	0.771	0.103	1.000
0.30	0.888	0.058	1.000	1.000	1.000	**0.70**	1.000	1.000	0.875	0.143	1.000
0.40	0.921	0.049	1.000	1.000	1.000	**0.80**	1.000	1.000	0.976	0.072	1.000
0.90	1.000	0.061	1.000	1.000	1.000	**0.90**	1.000	1.000	0.982	0.111	1.000

Table 5. Empirical power $T=10$; $R=81$. Percentage of rejections of the null hypothesis, 5% sig. level

SLM model						SEM model					
Case 1: Temporal independence ($\Sigma_1 = I_T$) and stability in spatial dependence $\rho_i = \rho_j$ $(\forall i, j)$											
ρ	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{SUR}_{\mathrm{SLM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SARAR}}$	$\boldsymbol{\lambda}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SARAR}}$
0.02	0.033	0.041	0.075	0.082	0.051	**0.10**	0.098	0.083	0.073	0.057	0.079
0.06	0.044	0.044	0.340	0.338	0.241	**0.20**	0.536	0.419	0.120	0.058	0.398
0.10	0.090	0.045	0.861	0.827	0.718	**0.30**	0.963	0.911	0.273	0.079	0.897
0.14	0.194	0.041	1.000	0.997	0.996	**0.40**	1.000	0.999	0.553	0.087	0.998
0.16	0.291	0.052	1.000	1.000	0.999	**0.50**	1.000	1.000	0.857	0.125	1.000
0.20	0.570	0.035	1.000	1.000	1.000	**0.60**	1.000	1.000	0.978	0.138	1.000
0.30	0.681	0.045	1.000	1.000	1.000	**0.70**	1.000	1.000	0.999	0.218	1.000
0.40	0.882	0.051	1.000	1.000	1.000	**0.80**	1.000	1.000	1.000	0.225	1.000
0.90	1.000	0.086	1.000	1.000	1.000	**0.90**	1.000	1.000	1.000	0.320	1.000
Case 2: Extreme temporal correlation (Σ_3) and stability in spatial dependence $\rho_i = \rho_j$ $(\forall i, j)$											
ρ	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{SUR}_{\mathrm{SLM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SARAR}}$	$\boldsymbol{\lambda}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SEM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{*\mathrm{SUR}}_{\mathrm{SLM}}$	$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SARAR}}$
0.02	0.057	0.055	1.000	1.000	1.000	**0.10**	0.087	0.086	0.076	0.068	0.084
0.06	0.149	0.105	1.000	1.000	1.000	**0.20**	0.523	0.519	0.064	0.055	0.397
0.10	0.299	0.159	1.000	1.000	1.000	**0.30**	0.952	0.948	0.094	0.093	0.894
0.14	0.422	0.129	1.000	1.000	1.000	**0.40**	1.000	1.000	0.094	0.079	0.999
0.16	0.530	0.133	1.000	1.000	1.000	**0.50**	1.000	1.000	0.159	0.156	1.000
0.20	0.626	0.148	1.000	1.000	1.000	**0.60**	1.000	1.000	0.199	0.203	1.000
0.30	0.810	0.267	1.000	1.000	1.000	**0.70**	1.000	1.000	0.232	0.209	1.000
0.40	0.926	0.441	1.000	1.000	1.000	**0.80**	1.000	1.000	0.361	0.353	1.000
0.90	1.000	0.908	1.000	1.000	1.000	**0.90**	1.000	1.000	0.388	0.329	1.000

Case 3: Temporal independence ($\Sigma_1 = I_T$) and instability in spatial dependence $\rho_1 \neq 0; \quad \rho_i = 0 \ \ (i > 1)$

ρ	$\mathbf{LM}^{SUR}_{SEM}$	$\mathbf{LM}^{*SUR}_{SEM}$	$\mathbf{LM}^{SUR}_{SLM}$	$\mathbf{LM}^{*SUR}_{SLM}$	$\mathbf{LM}^{SUR}_{SARAR}$	λ	$\mathbf{LM}^{SUR}_{SEM}$	$\mathbf{LM}^{*SUR}_{SEM}$	$\mathbf{LM}^{SUR}_{SLM}$	$\mathbf{LM}^{*SUR}_{SLM}$	$\mathbf{LM}^{SUR}_{SARAR}$
0.02	0.046	0.042	0.062	0.057	0.050	**0.10**	0.059	0.050	0.059	0.055	0.054
0.06	0.042	0.046	0.091	0.098	0.078	**0.20**	0.086	0.075	0.059	0.049	0.077
0.10	0.048	0.038	0.125	0.111	0.090	**0.30**	0.160	0.136	0.077	0.065	0.135
0.14	0.057	0.047	0.196	0.195	0.144	**0.40**	0.305	0.252	0.101	0.066	0.239
0.16	0.065	0.045	0.217	0.193	0.151	**0.50**	0.550	0.445	0.145	0.059	0.422
0.20	0.116	0.046	0.556	0.470	0.394	**0.60**	0.770	0.691	0.160	0.067	0.658
0.30	0.189	0.048	0.622	0.684	0.625	**0.70**	0.944	0.822	0.454	0.069	0.875
0.40	0.263	0.053	0.751	0.754	0.699	**0.80**	0.989	0.932	0.643	0.104	0.973
0.90	0.670	0.049	1.000	1.000	1.000	**0.90**	1.000	0.944	0.947	0.107	0.999

Case 4: Extreme temporal correlation (Σ_3) and instability in spatial dependence $\rho_1 \neq 0; \quad \rho_i = 0 \ \ (i > 1)$

ρ	$\mathbf{LM}^{SUR}_{SEM}$	$\mathbf{LM}^{*SUR}_{SEM}$	$\mathbf{LM}^{SUR}_{SLM}$	$\mathbf{LM}^{*SUR}_{SLM}$	$\mathbf{LM}^{SUR}_{SARAR}$	λ	$\mathbf{LM}^{SUR}_{SEM}$	$\mathbf{LM}^{*SUR}_{SEM}$	$\mathbf{LM}^{SUR}_{SLM}$	$\mathbf{LM}^{*SUR}_{SLM}$	$\mathbf{LM}^{SUR}_{SARAR}$
0.02	0.059	0.050	0.170	0.162	0.136	**0.10**	0.424	0.362	0.137	0.082	0.348
0.06	0.120	0.048	0.985	0.979	0.939	**0.20**	0.950	0.880	0.325	0.079	0.882
0.10	0.243	0.055	1.000	1.000	1.000	**0.30**	0.998	0.994	0.296	0.062	0.992
0.14	0.396	0.038	1.000	1.000	1.000	**0.40**	1.000	1.000	0.476	0.090	1.000
0.16	0.241	0.065	1.000	1.000	1.000	**0.50**	1.000	1.000	0.492	0.112	1.000
0.20	0.440	0.036	1.000	1.000	1.000	**0.60**	1.000	1.000	0.610	0.108	1.000
0.30	0.517	0.066	1.000	1.000	1.000	**0.70**	1.000	1.000	0.762	0.102	1.000
0.40	0.598	0.061	1.000	1.000	1.000	**0.80**	1.000	1.000	0.916	0.077	1.000
0.90	0.872	0.049	1.000	1.000	1.000	**0.90**	1.000	1.000	0.929	0.093	1.000

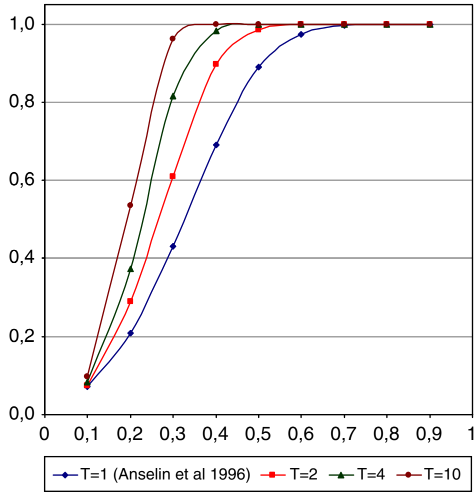

Figure 2. Estimated power of the $\mathbf{LM}_{SEM}^{SUR}$ test, Case 1.

The power of the $\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$ test also improves with the number of equations (T). Figure 2 represents the estimated power curves for the different cross-sections ($T=$ 2, 4, 10; Case 1). We add the results obtained by Anselin *et al.* (1996) for the same test, using only one equation. The improvement is evident, especially for low values of the spatial dependence coefficient. However contrary to what happened in the case of the test $\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$, the use of a stronger temporal correlation between the SUR equations does not help to increase the power of the test.

As highlighted by Anselin *et al.* (1996), the $\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$ and $\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$ tests are not robust to the specification of the alternative hypothesis. This is the reason for developing the robust tests in Section 3.1.

The performance of the two tests, $\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$ and $\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$, are very similar. They are scarcely affected by the number of equations (our results, with different T's are in line with the findings of Anselin *et al.* (1996)). The intensity of the correlation among the SUR equations has a greater impact, especially in the case of a very high temporal dependence (Case 2). $\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$ clearly shows that the 'robustness adjustment' does not work quite as well. For the case of the $\mathbf{LM}_{\mathrm{SLM}}^{*\mathrm{SUR}}$ test we obtain rejection frequencies that are slightly higher than the nominal value of 5%. This tendency is more severe as the SUR structure among the equations becomes stronger. The same comment extends to the $\mathbf{LM}_{\mathrm{SEM}}^{*\mathrm{SUR}}$: good behaviour for moderate to low relations of temporal dependence in the SUR system and a sharp deterioration in the case of strong time dependence.

4.3. Testing for the Instability of the Spatial Dependence Coefficients in Spatial SUR Models

The (possible) time instability of the parameter of spatial dependence is a factor that clearly affects the power of the tests. Cases 3 and 4 presented in Tables 3–5 show that this instability is a potential risk factor that especially affects the $\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$ and $\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$ tests.

It is interesting to note that the two tests react differently, depending on the level of correlation between the SUR equations. As such, for processes of substantive spatial dependence, SLM, the $\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}$ test suffers severe power losses under time instability. On the other hand, for an SEM type process, the time dependence appears to favour the power of the $\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}$ test. Neither of the two robust Multipliers appears to be affected by the (possible) time stability of the spatial parameter and, in some cases, they perform better than their corresponding raw Multipliers.

Given this apparently puzzling situation, we believe that it is really important to check the constancy of the coefficients of spatial dependence in SUR specifications. The tests developed in Section 3.2 may be useful in this aspect. Tables 6 and 7 present the results obtained for the two Multipliers, $\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}(\lambda)$ and $\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}(\rho)$.

It is encouraging that neither of the two tests suffer size problems: as shown in Table 6, the empirical size of both tests is always close to the nominal 5%.

In order to evaluate the power, we have repeated the design of Cases 3 and 4 (that is, there is spatial correlation in the first cross-section, whereas the remaining cross-sections are characterized by spatial independence). The results are shown in Table 7. Both tests react to the value of the spatial dependence parameter of the first cross-section, especially in the case of an SLM process. In the two cases, when the number of cross-sections (T) increases then the power decreases. As expected, the estimated power function improves as we introduce a stronger time correlation structure between the SUR equations.

5. Final Conclusions

Multivariant models are very common, useful techniques to deal simultaneously with temporal and spatial structures. The seemingly unrelated regressions, SUR models, are some of these techniques, which are very popular due to the fact that they require few assumptions and simple computations. As far as we know, since the seminal works of Anselin (1988a,b) very little has been presented about the possibilities of combining SUR models with spatial processes. We think that it is time to fill the gap. Specifically, in this paper we develop a range of tests based on the principle of the Lagrange Multiplier that are, firstly, efficient in order to detect

Table 6. Empirical size of the time instability spatial dependence tests. Percentage of rejections of the null hypothesis, 5% sig. level

			SLM: $\mathrm{LM}_{\mathrm{SLM}}^{\mathrm{SUR}}(\rho)$			SEM: $\mathrm{LM}_{\mathrm{SEM}}^{\mathrm{SUR}}(\lambda)$		
	R	ρ/λ	$T=2$	$T=4$	$T=10$	$T=2$	$T=4$	$T=10$
Σ_1	49	**0.2**	0.048	0.054	0.049	0.038	0.032	0.026
		0.5	0.052	0.047	0.045	0.035	0.034	0.036
		0.9	0.044	0.053	0.041	0.031	0.049	0.044
	81	**0.2**	0.064	0.045	0.051	0.044	0.030	0.044
		0.5	0.064	0.053	0.053	0.040	0.050	0.032
		0.9	0.053	0.049	0.044	0.041	0.037	0.055
Σ_2	49	**0.2**	0.048	0.047	0.047	0.045	0.036	0.034
		0.5	0.059	0.045	0.048	0.031	0.036	0.035
		0.9	0.042	0.050	0.046	0.025	0.036	0.050
	81	**0.2**	0.051	0.048	0.056	0.041	0.042	0.027
		0.5	0.043	0.054	0.051	0.048	0.045	0.060
		0.9	0.036	0.049	0.054	0.039	0.034	0.050
Σ_3	49	**0.2**	0.050	0.055	0.042	0.027	0.044	0.031
		0.5	0.046	0.053	0.044	0.028	0.034	0.039
		0.9	0.025	0.026	0.037	0.025	0.051	0.079
	81	**0.2**	0.060	0.046	0.046	0.044	0.043	0.037
		0.5	0.054	0.047	0.043	0.047	0.041	0.048
		0.9	0.048	0.045	0.027	0.030	0.049	0.076

Table 7. Estimated power of the time instability spatial dependence tests. Percentage of rejections of the null hypothesis, 5% sig. level; $R=81$

		SLM: $LM_{SLM}^{SUR}(\rho)$				SEM: $LM_{SEM}^{SUR}(\lambda)$		
	ρ_1	$T=2$	$T=4$	$T=10$	λ_1	$T=2$	$T=4$	$T=10$
Σ_1	**0.10**	0.190	0.201	0.114	**0.10**	0.082	0.090	0.057
	0.20	0.601	0.663	0.431	**0.20**	0.143	0.136	0.100
	0.30	0.905	0.977	0.845	**0.30**	0.273	0.252	0.192
	0.40	0.998	0.996	0.959	**0.40**	0.468	0.436	0.347
	0.50	1.000	1.000	1.000	**0.50**	0.621	0.693	0.562
	0.60	1.000	1.000	1.000	**0.60**	0.838	0.866	0.777
	0.70	1.000	1.000	1.000	**0.70**	0.955	0.972	0.928
	0.80	1.000	1.000	1.000	**0.80**	0.992	0.995	0.988
	0.90	1.000	1.000	1.000	**0.90**	1.000	0.999	0.999
	ρ_1	$T=2$	$T=4$	$T=10$	λ_1	$T=2$	$T=4$	$T=10$
Σ_3	**0.02**	0.371	0.405	0.141	**0.02**	0.125	0.103	0.070
	0.06	0.963	0.992	0.965	**0.06**	0.279	0.179	0.125
	0.10	1.000	1.000	1.000	**0.10**	0.493	0.324	0.196
	0.14	1.000	1.000	1.000	**0.14**	0.706	0.532	0.319
	0.16	1.000	1.000	1.000	**0.16**	0.859	0.722	0.470
	0.18	1.000	1.000	1.000	**0.18**	0.954	0.907	1.000
	0.20	1.000	1.000	1.000	**0.20**	0.995	0.976	1.000
	0.30	1.000	1.000	1.000	**0.30**	1.000	0.993	1.000
	0.40	1.000	1.000	1.000	**0.40**	1.000	1.000	1.000
	0.90	1.000	1.000	1.000	**0.90**	1.000	0.995	0.960

the presence of spatial elements in SUR models and, secondly, useful to distinguish the type of spatial structure, SLM or SEM, that highlights the data.

The Monte Carlo study contains some results that are worth highlighting. Firstly, the tests of diagonality (specifically, time independence between the different cross-sections) tend to be unreliable under the presence of spatial dependence. In general, they tend to reject the null hypotheses. Secondly, the performance of the robust Multipliers, which have the aim of differentiating between SLM and SEM processes, deteriorates as the symptoms of time instability increase. For these reasons, and because of the difficulties of dealing simultaneously with several cross-sections, it is advisable to follow a conservative model specification strategy when dealing with SUR models. General-to-specific approaches appear to adapt better to this scheme, although these comments pertain to a future research agenda.

Note

1. Henceforth the following notation will be used:

$$\mathbf{LM}=[g(\theta)_{|H_0}]'[I^{T,R}(\theta)_{|H_0}]^{-1}[g(\theta)_{|H_0}]\underset{\text{as}}{\sim}\chi^2(\text{d.f.})$$

where $g(\theta)$ is the score (vector of first derivatives of the likelihood function), $I^{T,R}(\theta)$ the information matrix and d.f. means degrees of freedom.

References

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

Anselin, L. (1988b) A test for spatial autocorrelation in seemingly unrelated regressions, *Economics Letters*, 28, 335–341.

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

Anselin, L., LeGallo, J. & Jayet, H. (2007) Spatial panel econometrics, in: L. Matyas & P. Sevestre (eds) *The Econometrics of Panel Data*, Dordrecht, Kluwer Academic Publishers.

Arora, S. & Brown, M. (1977) Alternative approaches to spatial autocorrelation: an improvement over current practice, *International Regional Science Review*, 2, 67–78.

Badinger, H., Müller, W. & Tondl, G. (2004) Regional convergence in the European Union, 1985–1999: a spatial dynamic panel analysis, *Regional Studies*, 38, 241–253.

Baltagi, B. (2008) *Econometric Analysis of Panel Data*, 4th edn, New York, John Wiley and Sons.

Beenstock, M. & Felsenstein, D. (2007) Spatial vector autoregressions, *Spatial Economic Analysis*, 2, 167–196.

Beenstock, M. & Felsenstein, D. (2008) Regional heterogeneity, conditional convergence and regional inequality, *Regional Studies*, 42, 475–488.

Bera, A. & Yoon, M. (1993) Specification testing with locally misspecified alternatives, *Econometric Theory*, 9, 649–658.

Breusch, T. & Pagan, A. (1980) The Lagrange multiplier test and its applications to model specification in econometrics, *Review of Economic Studies*, 47, 239–254.

Carlino, G. & DeFina, R. (1999) The differential regional effects of monetary policy: evidence from the US states, *Journal of Regional Science*, 39, 339–358.

Chen, X. & Conley, T. (2001) A new semiparametric spatial model for panel time series, *Journal of Econometrics*, 105, 59–83.

Coakley, J., Fuentes, A. & Smith, R. (2002) *A Principal Components Approach to Cross-section Dependence in Panels*, Working Paper, Department of Economics, Birbeck College, London, University of London.

Conley, T. (1999) GMM estimation with cross-sectional dependence, *Journal of Econometrics*, 92, 1–45.

Conley, T. & Molinari, F. (2008) Spatial correlation robust inference with errors in location or distance, *Journal of Econometrics*, 40, 76–96.

Davidson, J. (2000) *Econometric Theory*, Oxford, Blackwell.

Di Giacinto, V. (2003) Differential regional effects of monetary policy: a geographical SVAR approach, *International Regional Science Review*, 26, 313–341.

Di Giacinto, V. (2006) A generalized space–time model with an application to regional unemployment analysis in Italy, *International Regional Science Review*, 29, 159–198.

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

Dufour, J. & Khalaf, L. (2002) Exact tests for contemporaneous correlation of disturbances in seemingly unrelated regressions, *Journal of Econometrics*, 106, 143–170.

Dwivedi, T. & Srivastava, K. (1978) Optimality of least squares in seemingly unrelated regressions models, *Journal of Econometrics*, 7, 391–395.

Egger, P. & Pfaffermayr, M. (2004) Distance, trade and FDI: a Hausman–Taylor SUR approach, *Journal of Applied Econometrics*, 16, 227–246.

Elhorst, P. (2001) Dynamic models in space and time, *Geographical Analysis*, 33, 244–268.

Elhorst, P. (2003) Specification and estimation of spatial panel data models, *International Regional Science Review*, 26, 244–268.

Elhorst, P. (2005) Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels, *Geographical Analysis*, 37, 85–106.

Elhorst, P. (2008) Serial and spatial error correlation, *Economics Letters*, 100, 422–424.

Fingleton, B. (2001) Theoretical economic geography and spatial econometrics: dynamic perspectives, *Journal of Economic Geography*, 1, 201–225.

Fingleton, B. (2007) Multi-equation spatial econometric model, with application to EU manufacturing productivity growth, *Journal of Geographical Systems*, 9, 119–144.

Hordijk, L. (1979) Problems in estimating econometric relations in space, Papers, *Regional Science Association*, 42, 99–115.

Hordijk, L. & Nijkamp, P. (1977) Dynamic models of spatial autocorrelation, *Environment and Planning A*, 9, 505–519.

Kurata, H. (2001) One-sided tests for independence of seemingly unrelated regression equations, *Journal of Multivariate Analysis*, 90, 393–406.

Lauridsen, J., Bech, M., López, F. & Maté, M. (2010) A spatiotemporal analysis of public pharmaceutical expenditures, *Annals of Regional Science*, 44(2), 299–314.

Lee, L. (2004) Asymptotic distributions of maximum likelihood estimators for spatial autoregressive models, *Econometrica*, 72, 1899–1925.

LeGallo, J. & Chasco, C. (2008) Spatial analysis of urban growth in Spain, 1900–2001, *Empirical Economics*, 34, 59–80.

LeGallo, J. & Dall'erba, S. (2006) Evaluating the temporal and spatial heterogeneity of the European convergence process: 1980–1999, *Journal of Regional Science*, 46, 269–288.
Malinvaud, E. (1970) *Statistical Methods of Econometrics*, Amsterdam, North Holland.
Moscone, F., Tosetti, E. & Knapp, M. (2007) SUR model with spatial effects: an application to mental health expenditure, *Health Economics*, 16, 1403–1408.
Pesaran, H. (2005) *Estimation and Inference in Large Heterogeneous Panels with Cross Section Dependence*, Technical Report, Department of Economics, Cambridge, University of Cambridge.
Rey, S. & Montouri, B. (1999) US regional income convergence: a spatial econometrics perspective, *Regional Studies*, 33, 143–156.
Schmidt, P. (1976) *Econometrics*, New York, Marcel Dekker.
Shiba, T. & Tsurumi, H. (1988) Bayesian and non-Bayesian test of independence in seemingly unrelated regressions, *International Economic Review*, 29(2), 377–395.
Theil, H. (1971) *Principles of Econometrics*, New York, John Wiley and Sons.
Tsay, W. (2004) Testing for contemporaneous correlation of disturbances in seemingly unrelated regressions with serial dependence, *Economics Letters*, 83, 69–76.
Wang, X. & Kockelman, K. (2007) Specification and estimation of a spatially and temporally autocorrelated seemingly unrelated regression model: application to crash rates in China, *Transportation*, 34, 281–300.
White, E. & Hewings, G. (1982) Space–time employment modeling: some results using seemingly unrelated regression estimators, *Journal of Regional Science*, 22, 283–302.
Zellner, A. (1962) An efficient method of estimating seemingly unrelated regressions and test of aggregation bias, *Journal of the American Statistical Association*, 57, 348–368.

Appendix: Results of the ML estimation of SUR models with spatial effects

This Appendix contains additional results on the ML estimation of the different SUR models introduced in Sections 2 and 3. Specifically, we focus on the expressions of the score vector and of the information matrix. We include also details in relation to the obtaining of the Lagrange Multipliers.

A.1. Score Vector and Information Matrix of the SUR–SARAR Model

As indicated, the compact expression of this model is:

$$\left.\begin{array}{l}\mathbf{A}y = X\beta + u\\ \mathbf{B}u = \varepsilon\\ \varepsilon \sim N(0,\mathbf{\Omega}) \Rightarrow \mathbf{\Omega} = \mathbf{\Sigma}\otimes I_R\end{array}\right\}$$

$$\mathbf{A} = I_{TR} - \Lambda\otimes\mathbf{W} \quad \mathbf{B} = I_{TR} - \Upsilon\otimes\mathbf{W} \tag{A1}$$

The logarithm of the likelihood function of the model of expression (A1) is the following:

$$l(y;\theta) = -\frac{RT}{2}\ln(2\pi) - \frac{R}{2}\ln|\mathbf{\Sigma}| + \sum\nolimits_{t=1}^{T}\ln|\mathbf{B}_t| + \sum\nolimits_{t=1}^{T}\ln|\mathbf{A}_t|$$

$$-\frac{(\mathbf{A}y - X\beta)'\mathbf{B}'(\mathbf{\Sigma}\otimes I_R)^{-1}\mathbf{B}(\mathbf{A}y - X\beta)}{2} \tag{A2}$$

where $\theta' = [\beta'; \lambda_1; \cdots; \lambda_T; \rho_1; \cdots; \rho_T; \mathrm{vech}(\mathbf{\Sigma})']$ is the vector of parameters of the model, of order $(k+2T+T(T+1)/2)\times 1$. The score vector has the following structure:

$$g(\theta)=\begin{bmatrix}\dfrac{\partial l}{\partial\beta}\\ \dfrac{\partial l}{\partial\lambda_t}\Big|_{t=1,\dots T}\\ \dfrac{\partial l}{\partial\rho_t}\Big|_{t=1,\dots T}\\ \dfrac{\partial l}{\partial\sigma_{ij}}\Big|_{i,j=1,\dots T}\end{bmatrix}=\begin{bmatrix}X'\mathbf{B}'[\mathbf{\Sigma}^{-1}\otimes I_R]\mathbf{B}(\mathbf{A}\gamma-X\beta)\\ -\mathrm{tr}[\mathbf{A}_t^{-1}\mathbf{W}]+(\mathbf{A}\gamma-X\beta)'\mathbf{B}'[\mathbf{\Sigma}^{-1}\otimes I_R]\mathbf{B}(\mathbf{E}^{tt}\otimes\mathbf{W})\gamma\\ -\mathrm{tr}[\mathbf{B}_t^{-1}\mathbf{W}]+(\mathbf{A}\gamma-X\beta)'\mathbf{B}'[\mathbf{\Sigma}^{-1}\otimes I_R](\mathbf{E}^{tt}\otimes\mathbf{W})(\mathbf{A}\gamma-X\beta)\\ -\dfrac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}]+\dfrac{(\mathbf{A}\gamma-X\beta)'\mathbf{B}'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R]\mathbf{B}(\mathbf{A}\gamma-X\beta)}{2}\end{bmatrix}\tag{A3}$$

where vech(Σ) is a $T(T+1)\times 1$ vector of variances and covariances, vech(Σ) $=\{\sigma_{ij}; j\leq i, i=1,2,\dots,T\}$. In Equation (A3) $\dfrac{\partial l}{\partial\sigma_{ij}}$ means the first partial derivative of the function l with respect to the (i,j)-th element of vech(Σ); the same applies in relation to $\dfrac{\partial l}{\partial\lambda_t}$ and $\dfrac{\partial l}{\partial\rho_t}$: both refer to the first derivative of function l with respect to the t-th spatial dependence coefficient. Moreover, $\mathbf{E}^{ij}$ (analogously $\mathbf{E}^{tt}$) is a matrix of order $(T\times T)$ whose elements are all zero except the (i,j) and the (j,i) which have the value 1. The set of second derivatives is:

$$\begin{aligned}
\frac{\partial^2 l}{\partial\beta\partial\beta'}&=-X'\mathbf{B}'[\mathbf{\Sigma}^{-1}\otimes I_R]\mathbf{B}X\\
\frac{\partial^2 l}{\partial\beta\partial\lambda_t}&=-X'\mathbf{B}'[\mathbf{\Sigma}^{-1}\otimes I_R]\mathbf{B}(\mathbf{E}^{tt}\otimes\mathbf{W})\gamma\\
\frac{\partial^2 l}{\partial\beta\partial\rho_t}&=-X'[\mathbf{B}'\{(\mathbf{\Sigma}^{-1}\mathbf{E}^{tt})\otimes\mathbf{W}\}+\{(\mathbf{E}^{tt}\mathbf{\Sigma}^{-1})\otimes\mathbf{W}'\}\mathbf{B}](\mathbf{A}\gamma-X\beta)\\
\frac{\partial^2 l}{\partial\beta\partial\sigma_{ij}}&=-X'\mathbf{B}'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R]\mathbf{B}(\mathbf{A}\gamma-X\beta)
\end{aligned}\tag{A4.a}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial\lambda_t^2}&=-\mathrm{tr}[\mathbf{A}_t^{-1}\mathbf{W}\mathbf{A}_t^{-1}\mathbf{W}]-\gamma'(\mathbf{E}^{tt}\otimes\mathbf{W})'\mathbf{B}'[\mathbf{\Sigma}^{-1}\otimes I_R]\mathbf{B}(\mathbf{E}^{tt}\otimes\mathbf{W})\gamma\\
\frac{\partial^2 l}{\partial\lambda_t\partial\lambda_s}&=-\gamma'(\mathbf{E}^{ss}\otimes\mathbf{W})'\mathbf{B}'[\mathbf{\Sigma}^{-1}\otimes I_R]\mathbf{B}(\mathbf{E}^{tt}\otimes\mathbf{W})\gamma\\
\frac{\partial^2 l}{\partial\lambda_t\partial\rho_t}&=-(\mathbf{A}\gamma-X\beta)'[(\mathbf{E}^{tt}\otimes\mathbf{W})'(\mathbf{\Sigma}^{-1}\otimes I_R)B+B'(\mathbf{\Sigma}^{-1}\otimes I_R)(\mathbf{E}^{tt}\otimes\mathbf{W})](\mathbf{E}^{tt}\otimes\mathbf{W})\gamma\\
\frac{\partial^2 l}{\partial\lambda_t\partial\rho_s}&=-(\mathbf{A}\gamma-X\beta)'[(\mathbf{E}^{ss}\otimes\mathbf{W})'(\mathbf{\Sigma}^{-1}\otimes I_R)B+B'(\mathbf{\Sigma}^{-1}\otimes I_R)(\mathbf{E}^{ss}\otimes\mathbf{W})](\mathbf{E}^{tt}\otimes\mathbf{W})\gamma
\end{aligned}$$

$$\frac{\partial^2 l}{\partial \lambda_t \partial \sigma_{ij}} = -(\mathbf{A}\gamma - X\beta)'\mathbf{B}'[(\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}) \otimes I_R]\mathbf{B}(\mathbf{E}^{tt} \otimes \mathbf{W})\gamma \tag{A4.b}$$

$$\frac{\partial^2 l}{\partial \rho_t^2} = -\mathrm{tr}[\mathbf{B}_t^{-1}\mathbf{W}\mathbf{B}_t^{-1}\mathbf{W}] - (\mathbf{A}\gamma - X\beta)'(\mathbf{E}^{tt} \otimes \mathbf{W})'[\boldsymbol{\Sigma}^{-1} \otimes I_R](\mathbf{E}^{tt} \otimes \mathbf{W})(\mathbf{A}\gamma - X\beta)$$

$$\frac{\partial^2 l}{\partial \rho_t \partial \rho_s} = -(\mathbf{A}\gamma - X\beta)'(\mathbf{E}^{ss} \otimes \mathbf{W}')[\boldsymbol{\Sigma}^{-1} \otimes I_R](\mathbf{E}^{tt} \otimes \mathbf{W})(\mathbf{A}\gamma - X\beta)$$

$$\frac{\partial^2 l}{\partial \rho_t \partial \sigma_{ij}} = -(\mathbf{A}\gamma - X\beta)'(\mathbf{E}^{tt} \otimes \mathbf{W}')[(\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}) \otimes I_R]\mathbf{B}(\mathbf{A}\gamma - X\beta) \tag{A4.c}$$

$$\frac{\partial^2 l}{\partial \sigma_{ij}^2} = \frac{R}{2}\mathrm{tr}[\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}] - (\mathbf{A}\gamma - X\beta)'\mathbf{B}'[(\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}) \otimes I_R]\mathbf{B}(\mathbf{A}\gamma - X\beta)$$

$$\frac{\partial^2 l}{\partial \sigma_{ij} \partial \sigma_{sr}} = \frac{R}{2}\mathrm{tr}[\boldsymbol{\Sigma}^{-1}\mathbf{E}^{sr}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}] - (\mathbf{A}\gamma - X\beta)'\mathbf{B}'[(\boldsymbol{\Sigma}^{-1}\mathbf{E}^{sr}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}) \otimes I_R]\mathbf{B}(\mathbf{A}\gamma - X\beta) \tag{A4.d}$$

The expected value of these terms, changing the sign, is (for brevity's sake, we use the notation $I_{\eta\gamma} = -E\left[\frac{\partial^2 l}{\partial \eta \partial \gamma'}\right]$):

$$\begin{aligned} I_{\beta\beta} &= X'\mathbf{B}'[\boldsymbol{\Sigma}^{-1} \otimes I_R]\mathbf{B}X \\ I_{\beta\lambda_t} &= X'\mathbf{B}'[\boldsymbol{\Sigma}^{-1} \otimes I_R](\mathbf{E}^{tt} \otimes \mathbf{W})\mathbf{B}\mathbf{A}^{-1}X\beta; \quad t = 1,2,\dots,T \\ I_{\beta\rho_t} &= 0; \; t = 1,2,\dots,T \\ I_{\beta\sigma_{ij}} &= 0; \; i,j = 1,2,\dots,T \end{aligned} \tag{A5.a}$$

$$\begin{aligned} I_{\lambda_t\lambda_t} &= \mathrm{tr}[\mathbf{A}_t^{-1}\mathbf{W}\mathbf{A}_t^{-1}\mathbf{W}] + \beta'X'\mathbf{H}^{tt}_{\mathrm{SARAR}}X\beta + \mathrm{tr}\mathbf{H}^{tt}_{\mathrm{SARAR}}\mathbf{B}'^{-1}[\boldsymbol{\Sigma} \otimes I_R]\mathbf{B}^{-1}; \\ &\quad t = 1,2,\dots,T \\ I_{\lambda_t\lambda_s} &= \beta'X'\mathbf{H}^{ts}_{\mathrm{SARAR}}X\beta + \mathrm{tr}\mathbf{H}^{ts}_{\mathrm{SARAR}}\mathbf{B}'^{-1}[\boldsymbol{\Sigma} \otimes I_R]\mathbf{B}^{-1}; \quad t,s = 1,2,\dots,T \\ I_{\lambda_t\rho_t} &= \mathrm{tr}\{(\boldsymbol{\Sigma} \otimes I_R)\mathbf{B}'^{-1}[(\mathbf{E}^{tt}\boldsymbol{\Sigma}^{-1}) \otimes \mathbf{W}']\mathbf{B}[\mathbf{E}^{tt} \otimes \mathbf{W}] + [\mathbf{E}^{tt} \otimes (\mathbf{W}\mathbf{W})]\}(\mathbf{B}\mathbf{A})^{-1}; \\ &\quad t = 1,2,\dots,T \\ I_{\lambda_t\rho_s} &= \mathrm{tr}\{(\boldsymbol{\Sigma} \otimes I_R)\mathbf{B}'^{-1}[(\mathbf{E}^{ss}\boldsymbol{\Sigma}^{-1}) \otimes \mathbf{W}']\mathbf{B}[\mathbf{E}^{tt} \otimes \mathbf{W}](\mathbf{B}\mathbf{A})^{-1}\}; \quad t,s = 1,2,\dots,T \\ I_{\lambda_t\sigma_{ij}} &= \mathrm{tr}[\mathbf{B}(\mathbf{E}^{tt} \otimes \mathbf{W})(\mathbf{B}\mathbf{A})^{-1}][(\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}) \otimes I_R]; \quad t; i,j = 1,2,\dots,T \end{aligned} \tag{A5.b}$$

$$\begin{aligned} I_{\rho_t\rho_t} &= \mathrm{tr}[\mathbf{B}_t^{-1}\mathbf{W}\mathbf{B}_t^{-1}\mathbf{W}] + \sigma^{tt}\mathrm{tr}[\mathbf{E}^{tt} \otimes (\mathbf{W}'\mathbf{W})][\mathbf{B}^{-1}(\boldsymbol{\Sigma} \otimes I_R)\mathbf{B}'^{-1}]; \quad t = 1,2,\dots,T \\ I_{\rho_t\rho_s} &= \mathrm{tr}[(\mathbf{E}^{tt}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ss}) \otimes (\mathbf{W}'\mathbf{W})][\mathbf{B}^{-1}(\boldsymbol{\Sigma} \otimes I_R)\mathbf{B}'^{-1}]; \quad t,s = 1,2,\dots,T \\ I_{\rho_t\sigma_{ij}} &= \mathrm{tr}[(\mathbf{E}^{tt}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}) \otimes \mathbf{W}]\mathbf{B}'^{-1}; \quad t; i,j = 1,2,\dots,T \end{aligned} \tag{A5.c}$$

$$I_{\sigma_{ij}\sigma_{sr}} = \frac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1}\mathbf{E}^{sr}]; \quad i,j,s,r = 1,2,....,T \tag{A5.d}$$

where $\mathbf{H}^{ts}_{\mathrm{SARAR}} = \mathbf{A}'^{-1}(\mathbf{E}^{ss}\otimes\mathbf{W})'\mathbf{B}'[\mathbf{\Sigma}^{-1}\otimes I_R]\mathbf{B}(\mathbf{E}^{tt}\otimes\mathbf{W})\mathbf{A}^{-1}$ and σ^{st} is the element (t,s) of the matrix $\mathbf{\Sigma}^{-1}$. We introduce the following ordering of the information matrix:

$$I^{T,R}(\theta) = \begin{bmatrix} \underset{(kxk)}{I_{\beta\beta}} & \underset{(kxT)}{I_{\beta\lambda}} & \underset{(kxT)}{I_{\beta\rho}} & \underset{(kx(T(T+1)/2))}{I_{\beta\sigma}} \\ & \underset{(TxT)}{I_{\lambda\lambda}} & \underset{(TxT)}{I_{\lambda\rho}} & \underset{(Tx(T(T+1)/2))}{I_{\lambda\sigma}} \\ & & \underset{(TxT)}{I_{\rho\rho}} & \underset{(Tx(T(T+1)/2))}{I_{\rho\sigma}} \\ & & & \underset{\substack{((T(T+1)/2) \\ x(T(T+1)/2))}}{I_{\sigma\sigma}} \end{bmatrix} \tag{A6}$$

A.II. Testing for Spatial Effects in the SUR–SARAR Model

The null hypothesis is that there are no spatial effects in the SUR model of Equation (A1):

$$\left.\begin{aligned} H_0&:\lambda_t = \rho_t = 0 \\ H_A&: \text{No } H_0 \end{aligned}\right\}$$

$$\text{Under} H_0 \Rightarrow \mathbf{A} = \mathbf{B} = I_{TR}$$

$$\left.\begin{aligned} &y = X\beta + u \\ &\Rightarrow u = \varepsilon \\ &\varepsilon \sim N(0,\Omega) \Rightarrow \Omega = \Sigma\otimes I_R \end{aligned}\right\} \tag{A7}$$

The score of Equation (A3) becomes:

$$g(\theta)_{|H_0} = \begin{bmatrix} \frac{\partial l}{\partial\beta} \\ \underset{t=1,...T}{\frac{\partial l}{\partial\lambda_t}} \\ \underset{t=1,...T}{\frac{\partial l}{\partial\rho_t}} \\ \underset{i,j=1,...T}{\frac{\partial l}{\partial\sigma_{ij}}} \end{bmatrix} = \begin{bmatrix} g_{(\beta)|H_0} \\ g_{(\lambda)|H_0} \\ g_{(\rho)|H_0} \\ g_{(\Sigma)|H_0} \end{bmatrix} = \begin{bmatrix} 0 \\ \underset{t=1,...T}{\hat{u}'[\mathbf{\Sigma}^{-1}\mathbf{E}^{tt}\otimes\mathbf{W}]y} \\ \underset{t=1,...T}{\hat{u}'[\mathbf{\Sigma}^{-1}\mathbf{E}^{tt}\otimes\mathbf{W}]\hat{u}} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \underset{t=1,...T}{\sum_{s=1}^{T}\sigma^{st}\hat{u}_s'\mathbf{W}y_t} \\ \underset{t=1,...T}{\sum_{s=1}^{T}\sigma^{st}\hat{u}_s'\mathbf{W}\hat{u}_t} \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ [\mathbf{\Sigma}^{-1}\circ(\hat{U}'Y_L)]'\tau \\ [\mathbf{\Sigma}^{-1}\circ(\hat{U}'\hat{U}_L)]'\tau \\ 0 \end{bmatrix} \tag{A8}$$

$\hat{u}$ is the $(TR \times 1)$ vector of residuals of the SUR model, estimated in the absence of spatial effects. The score is made up of two sub-vectors of zeros, of orders $(k \times 1)$ and $(T(T+1)/2) \times 1$, respectively, and another two non-zero sub-vectors, both of order $(T \times 1)$, as appears in Equation (A8). Using the Hadamard product, $\circ$, the first non-zero subvector can be expressed as: $g_{(\lambda)_{|H_0}} = [\boldsymbol{\Sigma}^{-1} \circ (\hat{U}' Y_L)]'\tau$ and the second as $g_{(\rho)_{|H_0}} = [\boldsymbol{\Sigma}^{-1} \circ (\hat{U}' \hat{U}_L)]'\tau$. In both cases, τ is a $(T \times 1)$ vector of ones, $\hat{U}$ is an $(R \times T)$ matrix whose columns contain the SUR residuals corresponding to the different cross-sections: $\hat{U} = [\hat{u}_1 \quad \hat{u}_2 \quad \cdots \quad \hat{u}_T]$, and $\hat{U}_L$ is its spatial lag $\hat{U}_L = [\mathbf{W}\hat{u}_1 \quad \mathbf{W}\hat{u}_2 \quad \cdots \quad \mathbf{W}\hat{u}_T]$; Y_L is specified analogously.

Furthermore, the elements of the information matrix, also under the null hypothesis, become:

$$
\begin{aligned}
&I_{\beta\beta} = X'[\boldsymbol{\Sigma}^{-1} \otimes I_R]X \\
&I_{\beta\lambda_t} = X'[\boldsymbol{\Sigma}^{-1}\mathbf{E}^{tt} \otimes \mathbf{W}]X\beta; \quad t = 1, 2, ..., T \\
&I_{\beta\rho_t} = 0; \quad t = 1, 2, ..., T \\
&I_{\beta\sigma_{ij}} = 0; \quad i, j = 1, 2, ..., T
\end{aligned} \tag{A9.a}
$$

$$
\begin{aligned}
&I_{\lambda_t\lambda_t} = \sigma^{tt}(\beta'X'[\mathbf{E}^{tt} \otimes (\mathbf{W}'\mathbf{W})]X\beta + \sigma_{tt}\mathrm{tr}(\mathbf{W}'\mathbf{W})) + \mathrm{tr}\mathbf{W}^2; \quad t = 1, 2, ..., T \\
&I_{\lambda_t\lambda_s} = \sigma^{ts}\beta'X'[\mathbf{E}^{ts} \otimes (\mathbf{W}'\mathbf{W})]X\beta + \sigma^{ts}\sigma_{ts}\mathrm{tr}(\mathbf{W}'\mathbf{W}); \quad t, s = 1, 2, ..., T \\
&I_{\lambda_t\rho_t} = \sigma^{tt}\sigma_{tt}\mathrm{tr}(\mathbf{W}'\mathbf{W}) + \mathrm{tr}(\mathbf{W}\mathbf{W}); \quad t = 1, 2, ..., T \\
&I_{\lambda_t\rho_s} = \sigma^{ts}\sigma_{ts}(\mathrm{tr}\mathbf{W}'\mathbf{W}); \quad t, s = 1, 2, ..., T \\
&I_{\lambda_t\sigma_{ij}} = 0; \quad t; i, j = 1, 2, ..., T
\end{aligned} \tag{A9.b}
$$

$$
\begin{aligned}
&I_{\rho_t\rho_t} = \mathrm{tr}(\mathbf{W}'\mathbf{W})(\sigma^{tt}\sigma_{tt} + 1); \quad t = 1, 2, ..., T \\
&I_{\rho_t\rho_s} = \sigma^{ts}\sigma_{ts}\mathrm{tr}(\mathbf{W}'\mathbf{W}); \quad t, s = 1, 2, ..., T \\
&I_{\rho_t\sigma_{ij}} = 0; \quad t; i, j = 1, 2, ..., T
\end{aligned} \tag{A9.c}
$$

$$
I_{\sigma_{ij}\sigma_{sr}} = \frac{R}{2}\mathrm{tr}[\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{sr}]; \; i, j, s, r = 1, 2, ..., T \tag{A9.d}
$$

All these elements have been defined previously except σ_{ts} which refers to the element (t,s) of matrix $\boldsymbol{\Sigma}$. To sum up, the information matrix becomes:

$$
I^{T,R}(\theta)_{|H_0} = \begin{bmatrix} \underset{(kxk)}{I_{\beta\beta}} & \underset{(kxT)}{I_{\beta\lambda}} & \underset{(kxT)}{0} & \underset{(kx(T(T+1)/2))}{0} \\ & \underset{(TxT)}{I_{\lambda\lambda}} & \underset{(TxT)}{I_{\lambda\rho}} & \underset{(Tx(T(T+1)/2))}{0} \\ & & \underset{(TxT)}{I_{\rho\rho}} & \underset{(Tx(T(T+1)/2))}{0} \\ & & & \underset{\substack{((T(T+1)/2) \\ x(T(T+1)/2))}}{I_{\sigma\sigma}} \end{bmatrix} \tag{A10}
$$

This matrix is block-diagonal:

$$I^{T,R}(\theta)_{|H_0} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \Rightarrow \begin{cases} M_{11} = \begin{bmatrix} \underset{(kxk)}{I_{\beta\beta}} & \underset{(kxT)}{I_{\beta\lambda}} & \underset{(kxT)}{0} \\ & \underset{(TxT)}{I_{\lambda\lambda}} & \underset{(TxT)}{I_{\lambda\rho}} \\ & & \underset{(TxT)}{I_{\rho\rho}} \end{bmatrix} \\ M_{22} = \begin{bmatrix} \underset{\substack{((T(T-1)/2) \\ x(T(T-1)/2))}}{I_{\sigma\sigma}} \end{bmatrix} \\ M_{12} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ M_{12} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \end{cases} \tag{A11}$$

Sub-matrix $I_{\lambda\rho}$ is diagonal, which implies that the ML estimators of ρ_t and of λ_t, under the null hypothesis, are correlated only for the same cross-section but they are independent for different cross-sections; that is, $\mathrm{Cov}(\rho_t, \lambda_s) = 0$ if $t \neq s$ and $\mathrm{Cov}(\rho_t, \lambda_s) \neq 0$ if $t = s$. The Lagrange Multiplier, for the hypothesis of Equation (A7), is the quadratic form of the score evaluated in the null hypothesis (as in Equation (A8)), on the inverse of the information matrix, also evaluated in the null hypothesis (as in Equations (A9) and (A10)). The result that we obtain is that:

$$\mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SARAR}} = [g(\theta)_{|H_0}]'[I^{T,R}(\theta)_{|H_0}]^{-1}[g(\theta)_{|H_0}] \underset{\mathrm{as}}{\sim} \chi^2(2T)$$

$$\Rightarrow \mathbf{LM}^{\mathrm{SUR}}_{\mathrm{SARAR}} = [g'_{(\lambda)_{|H_0}} \quad g'_{(\rho)_{|H_0}}] \begin{bmatrix} I_{\lambda\lambda} - I_{\lambda\beta} I_{\beta\beta}^{-1} I_{\beta\lambda} & I_{\lambda\rho} \\ I_{\rho\lambda} & I_{\rho\rho} \end{bmatrix}^{-1} \begin{bmatrix} g_{(\lambda)_{|H_0}} \\ g_{(\rho)_{|H_0}} \end{bmatrix} \underset{\mathrm{as}}{\sim} \chi^2(2T) \tag{A12}$$

A.III. Score Vector and Information Matrix of the SUR–SLM Model

The model that corresponds to this case is:

$$\begin{aligned} y_t &= \lambda_t \mathbf{W} y_t + x_t\beta + \varepsilon_t \Rightarrow \mathbf{A}_t y_t = x_t\beta + \varepsilon_t \\ \mathbf{A}_t &= I_R - \lambda_t \mathbf{W} \end{aligned} \tag{A13}$$

All the elements that intervene in this specification are known. More compactly:

$$\begin{aligned} &\left.\begin{aligned} \mathbf{A}y &= X\beta + \varepsilon \\ \varepsilon &\sim N(0, \mathbf{\Omega}) \Rightarrow \mathbf{\Omega} = \Sigma \otimes I_R \end{aligned}\right\} \\ &\mathbf{A} = I_{TR} - \Lambda \otimes \mathbf{W} \end{aligned} \tag{A14}$$

Matrix $\mathbf{\Omega}$ is the same as that indicated in Equation (A1). The logarithm of the likelihood function, introducing the SLM structure of Equation (A14), is:

$$l(y;\theta) = -\frac{RT}{2}\ln(2\pi) - \frac{R}{2}\ln|\mathbf{\Sigma}| + \sum_{t=1}^{T} \ln|\mathbf{A}_t| - \frac{(\mathbf{A}y - X\beta)'(\mathbf{\Sigma} \otimes I_R)^{-1}(\mathbf{A}y - X\beta)}{2} \tag{A15}$$

where $\theta' = [\beta'; \lambda_1; \cdots; \lambda_T; \mathrm{vech}(\Sigma)']$ is the vector of parameters of the model, of order $(k + T + T(T+1)/2) \times 1$. The score vector has the following structure:

$$g(\theta)=\begin{bmatrix}\dfrac{\partial l}{\partial \beta}\\ \dfrac{\partial l}{\partial \lambda_t}\\ {\scriptstyle t=1,\ldots T}\\ \dfrac{\partial l}{\partial \sigma_{ij}}\\ {\scriptstyle i,j=1,\ldots T}\end{bmatrix}$$

$$=\begin{bmatrix} X'[\mathbf{\Sigma}^{-1}\otimes I_R](\mathbf{A}\gamma-X\beta)\\ -\mathrm{tr}[\mathbf{A}_t^{-1}\mathbf{W}]+(\mathbf{A}\gamma-X\beta)'[\mathbf{\Sigma}^{-1}\otimes I_R](\mathbf{E}^{tt}\otimes\mathbf{W})\gamma\\ -\dfrac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}]+\dfrac{(\mathbf{A}\gamma-X\beta)'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R](\mathbf{A}\gamma-X\beta)}{2}\end{bmatrix} \tag{A16}$$

The second derivatives corresponding to this case are:

$$\begin{aligned}\frac{\partial^2 l}{\partial\beta\partial\beta'}&=-X'[\mathbf{\Sigma}^{-1}\otimes I_R]X\\ \frac{\partial^2 l}{\partial\beta\partial\lambda_t}&=-X'[\mathbf{\Sigma}^{-1}\otimes I_R](\mathbf{E}^{tt}\otimes\mathbf{W})\gamma\\ \frac{\partial^2 l}{\partial\beta\partial\sigma_{ij}}&=-X'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R](\mathbf{A}\gamma-X\beta)\end{aligned} \tag{A17.a}$$

$$\begin{aligned}\frac{\partial^2 l}{\partial\lambda_t^2}&=-\mathrm{tr}[\mathbf{A}_t^{-1}\mathbf{W}\mathbf{A}_t^{-1}\mathbf{W}]-\gamma'(\mathbf{E}^{tt}\otimes\mathbf{W})'[\mathbf{\Sigma}^{-1}\otimes I_R](\mathbf{E}^{tt}\otimes\mathbf{W})\gamma\\ \frac{\partial^2 l}{\partial\lambda_t\partial\lambda_s}&=-\gamma'(\mathbf{E}^{ss}\otimes\mathbf{W})'[\mathbf{\Sigma}^{-1}\otimes I_R](\mathbf{E}^{tt}\otimes\mathbf{W})\gamma\\ \frac{\partial^2 l}{\partial\lambda_t\partial\sigma_{ij}}&=-(\mathbf{A}\gamma-X\beta)'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R](\mathbf{E}^{tt}\otimes\mathbf{W})\gamma\end{aligned} \tag{A17.b}$$

$$\begin{aligned}\frac{\partial^2 l}{\partial\sigma_{ij}^2}&=\frac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}]-(\mathbf{A}\gamma-X\beta)'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R](\mathbf{A}\gamma-X\beta)\\ \frac{\partial^2 l}{\partial\sigma_{ij}\partial\sigma_{sr}}&=\frac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{sr}\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}]-(\mathbf{A}\gamma-X\beta)'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{sr}\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R](\mathbf{A}\gamma-X\beta)\end{aligned} \tag{A17.c}$$

The expected values, changing their sign, are:

$$\begin{aligned}I_{\beta\beta}&=X'[\mathbf{\Sigma}^{-1}\otimes I_R]X\\ I_{\beta\lambda_t}&=X'[\mathbf{\Sigma}^{-1}\otimes I_R](\mathbf{E}^{tt}\otimes\mathbf{W})\mathbf{A}^{-1}X\beta;\quad t=1,2,\ldots.,T\\ I_{\beta\sigma_{ij}}&=0;\quad i,j=1,2,\ldots.,T\end{aligned} \tag{A18.a}$$

$$I_{\lambda_t\lambda_t} = \mathrm{tr}[\mathbf{A}_t^{-1}\mathbf{W}\mathbf{A}_t^{-1}\mathbf{W}] + \beta' X' \mathbf{H}_{\mathrm{SLM}}^{tt} X\beta + tr\mathbf{H}_{\mathrm{SLM}}^{tt}[\boldsymbol{\Sigma} \otimes I_R]; \quad t = 1, 2, \ldots, T$$
$$I_{\lambda_t\lambda_s} = \beta' X' \mathbf{H}_{\mathrm{SLM}}^{ts} X\beta + tr\mathbf{H}_{\mathrm{SLM}}^{ts}[\boldsymbol{\Sigma} \otimes I_R]; \quad t, s = 1, 2, \ldots, T$$
$$I_{\lambda_t\sigma_{ij}} = \mathrm{tr}[(\mathbf{E}^{tt} \otimes \mathbf{W})\mathbf{A}^{-1}][(\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}) \otimes I_R]; \quad t; i, j = 1, 2, \ldots, T \tag{A18.b}$$

$$I_{\sigma_{ij}\sigma_{sr}} = \frac{R}{2}\mathrm{tr}[\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{sr}]; \quad i, j, s, r = 1, 2, \ldots, T \tag{A18.c}$$

where $\mathbf{H}_{\mathrm{SLM}}^{ts} = \mathbf{A}'^{-1}(\mathbf{E}^{ss} \otimes \mathbf{W})'[\boldsymbol{\Sigma}^{-1} \otimes I_R](\mathbf{E}^{tt} \otimes \mathbf{W})\mathbf{A}^{-1}$ and σ^{st}, as before, is the element (t,s) of matrix $\boldsymbol{\Sigma}^{-1}$. Now we use the following ordering of the information matrix:

$$I^{T,R}(\theta) = \begin{bmatrix} \underset{(kxk)}{I_{\beta\beta}} & \underset{(kxT)}{I_{\beta\lambda}} & \underset{(kx(T(T+1)/2))}{I_{\beta\sigma}} \\ & \underset{(TxT)}{I_{\lambda\lambda}} & \underset{(Tx(T(T+1)/2))}{I_{\lambda\sigma}} \\ & & \underset{\substack{((T(T+1)/2) \\ x(T(T+1)/2))}}{I_{\sigma\sigma}} \end{bmatrix}$$

A.IV. Testing for Spatial Effects in the SUR–SLM Model

The null hypothesis in which we are interested is that the spatial lag of the endogenous variable included in the right-hand side of the SUR model of Equation (A13) is not relevant:

$$\left.\begin{array}{r} H_0{:}\lambda_t = 0 \\ H_A{:}\mathrm{No}\ H_0 \end{array}\right\}$$
$$\text{Under}\ \ H_0 \Rightarrow \mathbf{A} = I_{TR} \Rightarrow \left.\begin{array}{l} y = X\beta + \varepsilon \\ \varepsilon \sim N(0, \boldsymbol{\Omega}) \Rightarrow \boldsymbol{\Omega} = \Sigma \otimes I_R \end{array}\right\} \tag{A19}$$

The score of Equation (A16) becomes:

$$g(\theta)_{|H_0} = \begin{bmatrix} \dfrac{\partial l}{\partial \beta} \\ \underset{t=1,\ldots T}{\dfrac{\partial l}{\partial \lambda_t}} \\ \underset{i,j=1,\ldots T}{\dfrac{\partial l}{\partial \sigma_{ij}}} \end{bmatrix} = \begin{bmatrix} g_{(\beta)|H_0} \\ g_{(\lambda)|H_0} \\ g_{(\Sigma)|H_0} \end{bmatrix} = \begin{bmatrix} 0 \\ \underset{t=1,\ldots T}{\hat{u}'[\boldsymbol{\Sigma}^{-1}\mathbf{E}^{tt} \otimes \mathbf{W}]y} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \underset{t=1,\ldots T}{\sum_{s=1}^{T} \sigma^{st}\hat{u}_s'\mathbf{W}y_t} \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 \\ [\boldsymbol{\Sigma}^{-1} \circ (\hat{U}' Y_L)]'\tau \\ 0 \end{bmatrix} \tag{A20}$$

As before, $\hat{u}$ is the $(TR \times 1)$ vector of residuals of the SUR model in the absence of spatial effects. The other terms are known. The elements of the information matrix, also under the null hypothesis, are:

$$
\begin{aligned}
&I_{\beta\beta} = X'[\mathbf{\Sigma}^{-1} \otimes I_R]X \\
&I_{\beta\lambda_t} = X'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{tt}) \otimes \mathbf{W}]X\beta; \quad t = 1, 2, ..., T \\
&I_{\beta\sigma_{ij}} = 0; \quad i, j = 1, 2, ..., T
\end{aligned} \tag{A21.a}
$$

$$
\begin{aligned}
&I_{\lambda_t\lambda_t} = \sigma^{tt}[\beta' X'(\mathbf{E}^{tt} \otimes (\mathbf{W}'\mathbf{W}))X\beta + \sigma_{tt}\mathrm{tr}(\mathbf{W}'\mathbf{W})] + tr\mathbf{W}^2; \quad t = 1, 2, ..., T \\
&I_{\lambda_t\lambda_s} = \sigma^{ts}[\beta' X'(\mathbf{E}^{ts} \otimes (\mathbf{W}'\mathbf{W}))X\beta + \sigma_{ts}\mathrm{tr}(\mathbf{W}'\mathbf{W})]; \quad t, s = 1, 2, ..., T \\
&I_{\lambda_t\sigma_{ij}} = 0; \quad t; i, j = 1, 2, ..., T
\end{aligned} \tag{A21.b}
$$

$$
I_{\sigma_{ij}\sigma_{sr}} = \frac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1}\mathbf{E}^{sr}]; \quad i, j, s, r = 1, 2, ..., T \tag{A21.c}
$$

This matrix has a block-diagonal structure:

$$
I^{T,R}(\theta)_{|H_0} = \begin{bmatrix} \underset{(kxk)}{I_{\beta\beta}} & \underset{(kxT)}{I_{\beta\lambda}} & \underset{(kxT)}{0} \\ & \underset{(TxT)}{I_{\lambda\lambda}} & \underset{(Tx(T(T-1)/2))}{0} \\ & & \underset{\substack{((T(T-1)/2) \\ x(T(T-1)/2))}}{I_{\sigma\sigma}} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{cases} M_{11} = \begin{bmatrix} \underset{(kxk)}{I_{\beta\beta}} & \underset{(kxT)}{I_{\beta\lambda}} \\ & \underset{(TxT)}{I_{\lambda\lambda}} \end{bmatrix} \\ M_{22} = \begin{bmatrix} \underset{\substack{((T(T-1)/2) \\ x(T(T-1)/2))}}{I_{\sigma\sigma}} \end{bmatrix} \\ M_{12} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ M_{12} = [0 \quad 0] \end{cases} \tag{A22}
$$

Finally, the Lagrange Multiplier emerges as:

$$
\begin{aligned}
\mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}} &= [g(\theta)_{|H_0}]'[I^{T,R}(\theta)_{|H_0}]^{-1}[g(\theta)_{|H_0}] \underset{\mathrm{as}}{\sim} \chi^2(T) \\
\Rightarrow \mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR}} &= g'_{(\lambda)_{|H_0}}[I_{\lambda\lambda} - I_{\lambda\beta}I_{\beta\beta}^{-1}I_{\beta\lambda}]^{-1}g_{(\lambda)_{|H_0}} \underset{\mathrm{as}}{\sim} \chi^2(T)
\end{aligned} \tag{A23}
$$

The terms of the score associated with parameters λ can also be expressed as:

$$
g_{(\lambda)_{|H_0}} = [\mathbf{\Sigma}^{-1} \circ (\hat{U}' Y_L)]'\tau \tag{A24}
$$

A.V. Score Vector and Information Matrix of the SUR–SEM Model

The specification corresponding to this model is:

$$
\left.\begin{aligned}
&y = X\beta + u \\
&\mathbf{B}u = \varepsilon \\
&\varepsilon \sim N(0, \mathbf{\Omega}) \Rightarrow \mathbf{\Omega} = \mathbf{\Sigma} \otimes I_R
\end{aligned}\right\} \\
\mathbf{B} = I_{TR} - \Upsilon \otimes \mathbf{W} \tag{A25}
$$

The logarithm of the likelihood function is:

$$l(y;\theta) = -\frac{RT}{2}\ln(2\pi) - \frac{R}{2}\ln|\mathbf{\Sigma}| + \sum_{t=1}^{T}\ln|\mathbf{B}_t| - \frac{(y - X\beta)\mathbf{B}'(\mathbf{\Sigma}\otimes I_R)^{-1}\mathbf{B}(y - X\beta)}{2} \tag{A26}$$

where $\theta' = [\beta'; \rho_1; \cdots; \rho_T; \mathrm{vech}(\Sigma)']$ is a vector of parameters, of order $(k + T + T(T+1)/2) \times 1$. The score vector has the following composition:

$$g(\theta) = \begin{bmatrix} \dfrac{\partial l}{\partial \beta} \\ \dfrac{\partial l}{\partial \rho_t}_{t=1,\ldots T} \\ \dfrac{\partial l}{\partial \sigma_{ij}}_{i,j=1,\ldots T} \end{bmatrix}$$

$$= \begin{bmatrix} X'\mathbf{B}'[\mathbf{\Sigma}^{-1}\otimes I_R]\mathbf{B}(y - X\beta) \\ -\mathrm{tr}[\mathbf{B}_t^{-1}\mathbf{W}] + (y - X\beta)'\mathbf{B}'[\mathbf{\Sigma}^{-1}\otimes I_R](\mathbf{E}^{tt}\otimes\mathbf{W})(y - X\beta) \\ -\dfrac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}] + \dfrac{(y - X\beta)'\mathbf{B}'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R]\mathbf{B}(y - X\beta)}{2} \end{bmatrix} \tag{A27}$$

The second derivatives of the logarithm of the likelihood function are:

$$\frac{\partial^2 l}{\partial\beta\partial\beta'} = -X'\mathbf{B}'[\mathbf{\Sigma}^{-1}\otimes I_R]\mathbf{B}X$$

$$\frac{\partial^2 l}{\partial\beta\partial\rho_t} = -X'[\mathbf{B}'\{(\mathbf{\Sigma}^{-1}\mathbf{E}^{tt})\otimes\mathbf{W}\} + \{(\mathbf{E}^{tt}\mathbf{\Sigma}^{-1})\otimes\mathbf{W}'\}\mathbf{B}](y - X\beta)$$

$$\frac{\partial^2 l}{\partial\beta\partial\sigma_{ij}} = -X'\mathbf{B}'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R]\mathbf{B}(y - X\beta) \tag{A28.a}$$

$$\frac{\partial^2 l}{\partial\rho_t^2} = -\mathrm{tr}[\mathbf{B}_t^{-1}\mathbf{W}\mathbf{B}_t^{-1}\mathbf{W}] - (y - X\beta)'(\mathbf{E}^{tt}\otimes\mathbf{W})'[\mathbf{\Sigma}^{-1}\otimes I_R](\mathbf{E}^{tt}\otimes\mathbf{W})(y - X\beta)$$

$$\frac{\partial^2 l}{\partial\rho_t\partial\rho_s} = -(y - X\beta)'(\mathbf{E}^{ss}\otimes\mathbf{W})'[\mathbf{\Sigma}^{-1}\otimes I_R](\mathbf{E}^{tt}\otimes\mathbf{W})(y - X\beta)$$

$$\frac{\partial^2 l}{\partial\rho_t\partial\sigma_{ij}} = -(y - X\beta)'(\mathbf{E}^{tt}\otimes\mathbf{W})[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R]\mathbf{B}(y - X\beta) \tag{A28.b}$$

$$\frac{\partial^2 l}{\partial\sigma_{ij}^2} = \frac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}] - (y - X\beta)'\mathbf{B}'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R]\mathbf{B}(y - X\beta)$$

$$\frac{\partial^2 l}{\partial\sigma_{ij}\partial\sigma_{sr}} = \frac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{sr}\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}] - (y - X\beta)'\mathbf{B}'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{sr}\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1})\otimes I_R]\mathbf{B}(y - X\beta) \tag{A28.c}$$

Their expected value, after changing the sign is:

$$I_{\beta\beta} = X'\mathbf{B}'[\mathbf{\Sigma}^{-1} \otimes I_R]\mathbf{B}X$$

$$I_{\beta\rho_t} = 0; \quad t = 1, 2, ..., T$$

$$I_{\beta\sigma_{ij}} = 0; \quad i, j = 1, 2, ..., T \tag{A29.a}$$

$$I_{\rho_t\rho_t} = \mathrm{tr}[\mathbf{B}_t^{-1}\mathbf{W}\mathbf{B}_t^{-1}\mathbf{W}] + \sigma^{tt}\mathrm{tr}(\mathbf{E}^{tt} \otimes (\mathbf{W}'\mathbf{W}))[\mathbf{B}^{-1}(\mathbf{\Sigma} \otimes I_R)\mathbf{B}^{'-1}]; \quad t = 1, 2, ..., T$$

$$I_{\rho_t\rho_s} = \mathrm{tr}[(\mathbf{E}^{tt}\mathbf{\Sigma}^{-1}\mathbf{E}^{ss}) \otimes (\mathbf{W}'\mathbf{W})][\mathbf{B}^{-1}(\mathbf{\Sigma} \otimes I_R)\mathbf{B}^{'-1}]; \quad t, s = 1, 2, ..., T$$

$$I_{\rho_t\sigma_{ij}} = \mathrm{tr}[(\mathbf{E}^{tt}\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}) \otimes \mathbf{W}]\mathbf{B}^{'-1}; \quad t; i, j = 1, 2, ..., T \tag{A29.b}$$

$$I_{\sigma_{ij}\sigma_{sr}} = \frac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1}\mathbf{E}^{sr}]; \quad i, j, s, r = 1, 2, ..., T \tag{A29.c}$$

We use the following ordering of the information matrix:

$$I^{T,R}(\theta) = \begin{bmatrix} \underset{(kxk)}{I_{\beta\beta}} & \underset{(kxT)}{I_{\beta\rho}} & \underset{(kx(T(T+1)/2))}{I_{\beta\sigma}} \\ & \underset{(TxT)}{I_{\rho\rho}} & \underset{(Tx(T(T+1)/2))}{I_{\rho\sigma}} \\ & & \underset{\substack{((T(T+1)/2) \\ x(T(T+1)/2))}}{I_{\sigma\sigma}} \end{bmatrix} \tag{A30}$$

A.VI. Testing for Spatial Effects in the SUR–SEM Model

The null hypothesis that we want to test is that there is no SEM structure in the error terms of the SUR, which means that:

$$\left.\begin{array}{l} H_0{:}\rho_t = 0 \\ H_A{:}\text{No } H_0 \end{array}\right\} \Rightarrow \text{Under} \quad H_0 \Rightarrow \mathbf{B} = I_{TR} \Rightarrow \left.\begin{array}{c} y = X\beta + \varepsilon \\ \varepsilon \sim N(0, \mathbf{\Omega}); \mathbf{\Omega} = \Sigma \otimes I_R \end{array}\right\} \tag{A31}$$

The score of Equation (A27) under the null hypothesis of Equation (A31) becomes:

$$g(\theta)_{|H_0} = \begin{bmatrix} \dfrac{\partial l}{\partial \beta} \\ \dfrac{\partial l}{\partial \rho_t} \\ {\scriptstyle t=1,\dots T} \\ \dfrac{\partial l}{\partial \sigma_{ij}} \\ {\scriptstyle i,j=1,\dots T} \end{bmatrix} = \begin{bmatrix} g_{(\beta)_{|H_0}} \\ g_{(\rho)_{|H_0}} \\ g_{(\Sigma)_{|H_0}} \end{bmatrix} = \begin{bmatrix} 0 \\ \underset{t=1,\dots T}{\hat{u}'[\boldsymbol{\Sigma}^{-1}\mathbf{E}^{tt} \otimes \mathbf{W}]\hat{u}} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \underset{t=1,\dots T}{\sum_{s=1}^{T} \sigma^{st}\hat{u}_s'\mathbf{W}\hat{u}_t} \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ [\boldsymbol{\Sigma}^{-1} \circ (\hat{U}'\hat{U}_L)]'\tau \\ 0 \end{bmatrix} \tag{A32}$$

$\hat{u}$ is the $(TR \times 1)$ vector of residuals of the SUR model without spatial effects. The elements of the information matrix, under the null hypothesis, are also simplified:

$$\begin{aligned} I_{\beta\beta} &= X'[\boldsymbol{\Sigma}^{-1} \otimes I_R]X \\ I_{\beta\rho_t} &= 0; \quad t = 1, 2, ..., T \\ I_{\beta\sigma_{ij}} &= 0; \quad i, j = 1, 2, ..., T \end{aligned} \tag{A33.a}$$

$$\begin{aligned} I_{\rho_t\rho_t} &= \mathrm{tr}\mathbf{W}'\mathbf{W}(\sigma^{tt}\sigma_{tt} + 1); \ t = 1, 2, ..., T \\ I_{\rho_t\rho_s} &= \sigma^{ts}\sigma_{ts} tr\mathbf{W}'\mathbf{W}; \ t, s = 1, 2, ..., T \\ I_{\rho_t\sigma_{ij}} &= 0; \ t; i, j = 1, 2, ..., T \end{aligned} \tag{A33.b}$$

$$I_{\sigma_{ij}\sigma_{sr}} = \frac{R}{2}\mathrm{tr}[\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{sr}]; \quad i, j, s, r = 1, 2, ..., T \tag{A33.c}$$

The structure of that matrix is, once again, block-diagonal:

$$I^{T,R}(\theta)_{|H_0} = \begin{bmatrix} \underset{(kxk)}{I_{\beta\beta}} & \underset{(kxT)}{0} & \underset{(kxT)}{0} \\ & \underset{(TxT)}{I_{\rho\rho}} & \underset{(Tx(T(T-1)/2))}{0} \\ & & \underset{\substack{((T(T-1)/2) \\ x(T(T-1)/2))}}{I_{\sigma\sigma}} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{cases} M_{11} = \begin{bmatrix} \underset{(kxk)}{I_{\beta\beta}} & 0 \\ & \underset{\substack{((T(T-1)/2) \\ x(T(T-1)/2))}}{I_{\sigma\sigma}} \end{bmatrix} \\ M_{22} = \begin{bmatrix} \underset{(TxT)}{I_{\rho\rho}} \end{bmatrix} \\ M_{12} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ M_{12} = \begin{bmatrix} 0 & 0 \end{bmatrix} \end{cases} \tag{A34}$$

This result facilitates the obtaining of the corresponding Multiplier:

$$\begin{aligned} \mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}} &= [g(\theta)_{|H_0}]'[I^{T,R}(\theta)_{|H_0}]^{-1}[g(\theta)_{|H_0}] \underset{\mathrm{as}}{\sim} \chi^2(T) \\ \Rightarrow \mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR}} &= g'_{(\rho)_{|H_0}}[I_{\rho\rho}]^{-1}g_{(\rho)_{|H_0}} = \end{aligned}$$

$$= \left[\sum_{s=1}^{T} \sigma^{st}\hat{u}_s'\mathbf{W}\hat{u}_t\right]' [I_{\rho\rho}]^{-1} \left[\sum_{s=1}^{T} \sigma^{st}\hat{u}_s'\mathbf{W}\hat{u}_t\right] = \underset{\text{as}}{\sim} \chi^2(T) \tag{A35}$$

As in the previous cases, we can propose a more compact expression for the term of the score vector associated with the parameters ?:

$$g_{(\rho)_{|H_0}} = [\boldsymbol{\Sigma}^{-1} \circ (\hat{U}'\hat{U}_L)]'\tau \tag{A36}$$

A.VII. Testing for the Time-Constancy of the Spatial Dependence Coefficients

The log-likelihood function of the SARAR model specified in expression (20) is similar to that of expression (2). The first corresponds to the restricted specification and the second to the unrestricted version of the model. The likelihood function of the first one is:

$$\begin{aligned} l(y;\theta) = &-\frac{RT}{2}\ln(2\pi) \\ &-\frac{R}{2}\ln|\boldsymbol{\Sigma}|+T\ln|\hat{\mathbf{B}}|+T\ln|\hat{\mathbf{A}}|-\frac{(\mathbf{A}y - X\beta)'[\boldsymbol{\Sigma}^{-1} \otimes (\hat{\mathbf{B}}'\hat{\mathbf{B}})](\mathbf{A}y - X\beta)}{2} \end{aligned} \tag{A37}$$

The associated score vector is more compact in the restricted model:

$$g(\theta) = \begin{bmatrix} \dfrac{\partial l}{\partial \beta} \\ \dfrac{\partial l}{\partial \lambda} \\ \dfrac{\partial l}{\partial \rho} \\ \dfrac{\partial l}{\partial \sigma_{ij}} \end{bmatrix}_{i,j=1,\dots T} = \begin{bmatrix} X'[\boldsymbol{\Sigma}^{-1} \otimes (\hat{\mathbf{B}}'\hat{\mathbf{B}})](\mathbf{A}y - X\beta) \\ -T\mathrm{tr}[\hat{\mathbf{A}}^{-1}\mathbf{W}] + (\mathbf{A}y - X\beta)'[\boldsymbol{\Sigma}^{-1} \otimes (\hat{\mathbf{B}}'\hat{\mathbf{B}})](I_T \otimes \mathbf{W})y \\ -T\mathrm{tr}[\hat{\mathbf{B}}^{-1}\mathbf{W}] + (\mathbf{A}y - X\beta)'[\boldsymbol{\Sigma}^{-1} \otimes (\hat{\mathbf{B}}'\mathbf{W})](\mathbf{A}y - X\beta) \\ -\dfrac{R}{2}\mathrm{tr}[\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}] + \dfrac{(\mathbf{A}y - X\beta)'[(\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}) \otimes (\hat{\mathbf{B}}'\hat{\mathbf{B}})](\mathbf{A}y - X\beta)}{2} \end{bmatrix} \tag{A38}$$

It is easy to test for the absence of spatial effects in the specification of (20). Now the null hypothesis only affects two parameters:

$$\left.\begin{array}{r} H_0{:}\lambda = \rho = 0 \\ H_A{:}\ \text{No } H_0 \end{array}\right\} \quad \Rightarrow \text{Under } H_0 \Rightarrow \mathbf{A} = \mathbf{B} = I_{TR} \tag{A39}$$

The associated Multiplier maintains a relatively complex structure:

$$\Rightarrow \mathbf{LM}_{\text{SARMA}}^{\text{SUR(cons)}} = [g'_{(\lambda)_{|H_0}} \quad g'_{(\rho)_{|H_0}}] \begin{bmatrix} I_{\lambda\lambda} - I_{\lambda\beta}I_{\beta\beta}^{-1}I_{\beta\lambda} & I_{\lambda\rho} \\ I_{\rho\lambda} & I_{\rho\rho} \end{bmatrix}^{-1} \begin{bmatrix} g_{(\lambda)_{|H_0}} \\ g_{(\rho)_{|H_0}} \end{bmatrix} \underset{\text{as}}{\sim} \chi^2(2) \tag{A40}$$

with:

$$I_{\beta\beta} = X'[\mathbf{\Sigma}^{-1} \otimes I_R]X$$
$$I_{\beta\lambda} = X'[\mathbf{\Sigma}^{-1} \otimes \mathbf{W}]X\beta$$
$$I_{\lambda\lambda} = T[\mathrm{tr}(\mathbf{W}'\mathbf{W}) + \mathrm{tr}(\mathbf{W}\mathbf{W})] + \beta' X'[\mathbf{\Sigma}^{-1} \otimes (\mathbf{W}'\mathbf{W})]X\beta \quad g_{(\lambda)_{|H_0}} = \tau'[\mathbf{\Sigma}^{-1} \circ (\hat{U}' Y_L)]\tau$$
$$I_{\rho\rho} = T[\mathrm{tr}(\mathbf{W}'\mathbf{W}) + \mathrm{tr}(\mathbf{W}\mathbf{W})] \qquad g_{(\rho)_{|H_0}} = \tau'[\mathbf{\Sigma}^{-1} \circ (\hat{U}' \hat{U}_L)]$$
$$I_{\rho\lambda} = T[\mathrm{tr}(\mathbf{W}'\mathbf{W}) + \mathrm{tr}(\mathbf{W}\mathbf{W})] \tag{A41}$$

The discussion for the case of the Spatial Lag Model of expression (21) is very similar. Now, the log-likelihood function and the score vector are as follows:

$$l(\gamma;\theta) = -\frac{RT}{2}\ln(2\pi) - \frac{R}{2}\ln|\mathbf{\Sigma}| + T\ln|\hat{\mathbf{A}}| - \frac{(\mathbf{A}\gamma - X\beta)'[\mathbf{\Sigma}^{-1} \otimes I_R](\mathbf{A}\gamma - X\beta)}{2} \tag{A42}$$

$$g(\theta) = \begin{bmatrix} \dfrac{\partial l}{\partial \beta} \\ \dfrac{\partial l}{\partial \lambda} \\ \dfrac{\partial l}{\partial \sigma_{ij}} \end{bmatrix}_{i,j=1,\ldots T} = \begin{bmatrix} X'[\mathbf{\Sigma}^{-1} \otimes I_R](\mathbf{A}\gamma - X\beta) \\ -T\mathrm{tr}[\hat{\mathbf{A}}^{-1}\mathbf{W}] + (\mathbf{A}\gamma - X\beta)'[\mathbf{\Sigma}^{-1} \otimes I_R](I_T \otimes \mathbf{W})\gamma \\ -\dfrac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}] + \dfrac{(\mathbf{A}\gamma - X\beta)'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1}) \otimes I_R](\mathbf{A}\gamma - X\beta)}{2} \end{bmatrix} \tag{A43}$$

The hypothesis that parameter λ is zero leads us back to the SUR without spatial effects:

$$\left.\begin{matrix} H_0: \lambda = 0 \\ H_A: \text{No } H_0 \end{matrix}\right\} \qquad \Rightarrow \text{Under} \quad H_0 \Rightarrow \mathbf{A} = I_{TR} \tag{A44}$$

and the Multiplier result is:

$$\Rightarrow \mathbf{LM}_{\mathrm{SLM}}^{\mathrm{SUR(cons)}} = g'_{(\lambda)_{|H_0}}[I_{\lambda\lambda} - I_{\lambda\beta}I_{\beta\beta}^{-1}I_{\beta\lambda}]^{-1}g_{(\lambda)_{|H_0}} \underset{\mathrm{as}}{\sim} \chi^2(1) \tag{A45}$$

with:

$$g_{(\lambda)_{|H_0}} = \tau'[\mathbf{\Sigma}^{-1} \circ (\hat{U}' Y_L)]\tau$$
$$\begin{cases} I_{\beta\beta} = X'[\mathbf{\Sigma}^{-1} \otimes I_R]X \\ I_{\beta\lambda} = X'[\mathbf{\Sigma}^{-1} \otimes \mathbf{W}]X\beta \\ I_{\lambda\lambda} = T[tr(\mathbf{W}'\mathbf{W}) + tr(\mathbf{W}\mathbf{W})] + \beta' X'[\mathbf{\Sigma}^{-1} \otimes (\mathbf{W}'\mathbf{W})]X\beta \end{cases} \tag{A46}$$

Finally, the log-likelihood function and the score vector for the Spatial Error Model of Equation (22) are:

$$l(y;\theta) = -\frac{RT}{2}\ln(2\pi) - \frac{R}{2}\ln|\mathbf{\Sigma}| + T\ln|\hat{\mathbf{B}}| - \frac{(y - X\beta)'[\mathbf{\Sigma}^{-1} \otimes (\hat{\mathbf{B}}'\hat{\mathbf{B}})](y - X\beta)}{2} \tag{A47}$$

$$g(\theta) = \begin{bmatrix} \dfrac{\partial l}{\partial \beta} \\ \dfrac{\partial l}{\partial \rho_t} \\ \dfrac{\partial l}{\partial \sigma_{ij}} \end{bmatrix} = \begin{bmatrix} X'[\mathbf{\Sigma}^{-1} \otimes (\hat{\mathbf{B}}'\hat{\mathbf{B}})](y - X\beta) \\ -T\mathrm{tr}[\hat{\mathbf{B}}^{-1}\mathbf{W}] + (y - X\beta)'[\mathbf{\Sigma}^{-1} \otimes (\hat{\mathbf{B}}'\mathbf{W})](y - X\beta) \\ -\dfrac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}] + \dfrac{(y - X\beta)'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1}) \otimes (\hat{\mathbf{B}}'\hat{\mathbf{B}})](y - X\beta)}{2} \end{bmatrix} \tag{A48}$$

The null hypothesis of absence of spatial effects implies that parameter ρ is zero:

$$\left.\begin{array}{c} H_0{:}\rho = 0 \\ H_A{:}\ \text{No } H_0 \end{array}\right\} \qquad \Rightarrow \text{Under} \quad H_0 \Rightarrow \mathbf{B} = I_{TR} \tag{A49}$$

The expression of the Multiplier is relatively simple:

$$\mathbf{LM}_{\mathrm{SEM}}^{\mathrm{SUR(cons)}} = g'_{(\rho)_{|H_0}} I_{\rho\rho}^{-1} g_{(\rho)_{|H_0}} \underset{\mathrm{as}}{\sim} \chi^2(1)$$

$$g_{(\lambda)_{|H_0}} = \tau'[\mathbf{\Sigma}^{-1} \circ (\hat{U}'\hat{U}_L)]\tau$$

$$I_{\rho\rho} = T[\mathrm{tr}(\mathbf{W}'\mathbf{W}) + \mathrm{tr}(\mathbf{W}\mathbf{W})] \tag{A50}$$

The problem addressed in Section 3.2 refers to testing the time-constancy of the parameters of spatial dependence introduced in the SUR specifications. In the case of the SARAR model associated with the hypothesis of Equation (23), the Lagrange Multiplier appears in expression (24) and the information matrix that should be introduced in this expression is:

$$I^{T,R}(\theta) = \begin{bmatrix} I_{\beta\beta} & I_{\beta\lambda} & I_{\beta\rho} & I_{\beta\sigma} \\ & I_{\lambda\lambda} & I_{\lambda\rho} & I_{\lambda\sigma} \\ & & I_{\rho\rho} & I_{\rho\sigma} \\ & & & I_{\sigma\sigma} \end{bmatrix} \tag{A51}$$

where $\hat{u} = [I_T \otimes (I_R - \hat{\lambda}\mathbf{W})]y - X\hat{\beta}$ and $\hat{\varepsilon} = [I_T \otimes (I_R - \hat{\rho}\mathbf{W})]\hat{u}$. Furthermore:

$$
\begin{cases}
I_{\beta\beta} = X'[\mathbf{\Sigma}^{-1} \otimes (\hat{\mathbf{B}}'\hat{\mathbf{B}})]X \\
I_{\beta\lambda_t} = X'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{tt}) \otimes (\hat{\mathbf{B}}'\mathbf{W}\hat{\mathbf{B}}\hat{\mathbf{A}}^{-1})]X\beta; \quad t = 1, 2, \ldots., T \\
I_{\beta\rho_t} = 0; \quad t = 1, 2, \ldots., T \\
I_{\beta\sigma_{ij}} = 0; \quad i, j = 1, 2, \ldots., T \\
I_{\lambda_t\lambda_t} = \mathrm{tr}(\hat{\mathbf{A}}'^{-1}\mathbf{W}'\hat{\mathbf{A}}^{-1}\mathbf{W}) + \sigma^{tt}\beta' X'[\hat{\mathbf{A}}'^{-1}\mathbf{W}'\hat{\mathbf{B}}'\hat{\mathbf{B}}\mathbf{W}\hat{\mathbf{A}}^{-1}]X\beta + \mathrm{tr}[\hat{\mathbf{A}}'^{-1}\mathbf{W}'\hat{\mathbf{B}}'\hat{\mathbf{B}}\hat{\mathbf{A}}^{-1}\mathbf{W}(\hat{\mathbf{B}}'\hat{\mathbf{B}})^{-1}] \\
I_{\lambda_t\lambda_s} = \sigma^{ts}\beta' X'[\mathbf{E}^{ts} \otimes (\hat{\mathbf{A}}'^{-1}\mathbf{W}'\hat{\mathbf{B}}'\hat{\mathbf{B}}\mathbf{W}\hat{\mathbf{A}}^{-1})]X\beta; \quad t, s = 1, 2, \ldots., T \\
I_{\lambda_t\rho_t} = \sigma^{tt}\sigma_{tt}\mathrm{tr}[\hat{\mathbf{B}}'^{-1}\mathbf{W}\hat{\mathbf{B}}\mathbf{W}\hat{\mathbf{A}}^{-1}\hat{\mathbf{B}}^{-1}] + \mathrm{tr}(\mathbf{W}\mathbf{W}\hat{\mathbf{A}}^{-1}\hat{\mathbf{B}}^{-1}); \quad t = 1, 2, \ldots., T \\
I_{\lambda_t\rho_s} = \sigma^{ts}\sigma_{ss}\mathrm{tr}[\hat{\mathbf{B}}'^{-1}\mathbf{W}'\hat{\mathbf{B}}\mathbf{W}\hat{\mathbf{A}}^{-1}\hat{\mathbf{B}}^{-1}]; \quad t, s = 1, 2, \ldots., T \\
I_{\lambda_t\sigma_{ij}} = \begin{cases} 0 & \text{if } t \neq i \text{ or } t \neq j \\ \sigma^{tt}\mathrm{tr}(\hat{\mathbf{A}}^{-1}\mathbf{W}) & \text{if } t = i = j \end{cases} \\
I_{\rho_t\rho_t} = \mathrm{tr}[\hat{\mathbf{B}}'^{-1}\mathbf{W}\hat{\mathbf{B}}^{-1}\mathbf{W}] + \sigma^{tt}\sigma_{tt}\mathrm{tr}[\hat{\mathbf{B}}'^{-1}\mathbf{W}\mathbf{W}'\hat{\mathbf{B}}^{-1}]; \quad t = 1, 2, \ldots., T \\
I_{\rho_t\rho_s} = \sigma^{ts}\sigma_{ts}\mathrm{tr}[\hat{\mathbf{B}}'^{-1}\mathbf{W}\mathbf{W}'\hat{\mathbf{B}}^{-1}]; \quad t, s = 1, 2, \ldots., T \\
I_{\rho_t\sigma_{ij}} = \begin{cases} 0 & \text{if } t \neq i \text{ and } t \neq j \\ \sigma^{ti}\mathrm{tr}[\hat{\mathbf{B}}'^{-1}\mathbf{W}] & \text{if } t = i \text{ or } t = j \end{cases} \\
I_{\sigma_{ij}\sigma_{sr}} = \dfrac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1}\mathbf{E}^{sr}]; \quad i, j, s, r = 1, 2, \ldots., T
\end{cases}
\tag{A52}
$$

In the case of the SLM model to which the hypothesis of Equation (26) refers, we need the following information matrix in order to solve for the expression of the Lagrange Multiplier that appears in Equation (27):

$$
I^{T,R}(\theta) = \begin{bmatrix} I_{\beta\beta} & I_{\beta\lambda} & I_{\beta\sigma} \\ & I_{\lambda\lambda} & I_{\lambda\sigma} \\ & & I_{\sigma\sigma} \end{bmatrix}
$$

$$
\begin{cases}
I_{\beta\beta} = X'[\mathbf{\Sigma}^{-1} \otimes I_R]X \\
I_{\beta\lambda_t} = X'[(\mathbf{\Sigma}^{-1}\mathbf{E}^{tt}) \otimes (\mathbf{W}\hat{\mathbf{A}}^{-1})]X\beta; \quad t = 1, 2, \ldots., T \\
I_{\beta\sigma_{ij}} = 0; \quad i, j = 1, 2, \ldots., T \\
I_{\lambda_t\lambda_t} = \mathrm{tr}(\hat{\mathbf{A}}'^{-1}\mathbf{W}'\hat{\mathbf{A}}^{-1}\mathbf{W}) + \sigma^{tt}\beta_t' x_t'[\hat{\mathbf{A}}'^{-1}\mathbf{W}'\mathbf{W}\hat{\mathbf{A}}^{-1}]x_t\beta_t + \sigma^{tt}\sigma_{tt}\mathrm{tr}[\hat{\mathbf{A}}'^{-1}\mathbf{W}'\mathbf{W}\hat{\mathbf{A}}^{-1}]; t = 1, 2, \ldots., T \\
I_{\lambda_t\lambda_s} = \sigma^{ts}\beta_t' x_t'[\hat{\mathbf{A}}'^{-1}\mathbf{W}'\mathbf{W}\hat{\mathbf{A}}^{-1}]x_s\beta_s + \sigma^{ts}\sigma_{ts}\mathrm{tr}[\hat{\mathbf{A}}'^{-1}\mathbf{W}'\mathbf{W}\hat{\mathbf{A}}^{-1}]; t, s = 1, 2, \ldots., T \\
I_{\lambda_t\sigma_{ij}} = \begin{cases} 0 & \text{if } t \neq i \text{ or } t \neq j \\ \sigma^{tt}\mathrm{tr}(\hat{\mathbf{A}}^{-1}\mathbf{W}) & \text{if } t = i = j \end{cases} \\
I_{\sigma_{ij}\sigma_{sr}} = \dfrac{R}{2}\mathrm{tr}[\mathbf{\Sigma}^{-1}\mathbf{E}^{ij}\mathbf{\Sigma}^{-1}\mathbf{E}^{sr}]; \quad i, j, s, r = 1, 2, \ldots., T
\end{cases}
\tag{A53}
$$

Finally, the same can be said with respect to the SEM model of Equation (11). The null hypothesis of time-constancy appears in Equation (29), the Lagrange Multiplier is that of Equation (30) which intervenes in the following information matrix:

$$I^{T,R}(\theta) = \begin{bmatrix} I_{\beta\beta} & I_{\beta\rho} & I_{\beta\sigma} \\ & I_{\rho\rho} & I_{\rho\sigma} \\ & & I_{\sigma\sigma} \end{bmatrix}$$

$$\begin{cases} I_{\beta\beta} = X'[\boldsymbol{\Sigma}^{-1} \otimes (\hat{\mathbf{B}}'\hat{\mathbf{B}})]X \\ I_{\beta\rho_t} = 0;\, t = 1, 2, ..., T \\ I_{\beta\sigma_{ij}} = 0;\, i, j = 1, 2, ..., T \\ I_{\rho_t\rho_t} = \mathrm{tr}[\hat{\mathbf{B}}'^{-1}\mathbf{W}\hat{\mathbf{B}}^{-1}\mathbf{W}] + \sigma^{tt}\sigma_{tt}\mathrm{tr}[\hat{\mathbf{B}}'^{-1}\mathbf{W}'\mathbf{W}\hat{\mathbf{B}}^{-1}]; \quad t = 1, 2, ..., T \\ I_{\rho_t\rho_s} = \sigma^{ts}\sigma_{ts}\mathrm{tr}[\hat{\mathbf{B}}'^{-1}\mathbf{W}\mathbf{W}'\hat{\mathbf{B}}^{-1}]; \quad t, s = 1, 2, ..., T \\ I_{\rho_t\sigma_{ij}} = \begin{cases} 0 & \text{if } \mathrm{t} \neq \mathrm{i},\ \mathrm{t} \neq \mathrm{j} \text{ and } \mathrm{i} \neq \mathrm{j} \\ \sigma^{ti}\mathrm{tr}[\hat{\mathbf{B}}'^{-1}\mathbf{W}] & \text{if } \mathrm{i} \neq \mathrm{j} \text{ and } \mathrm{t} = \mathrm{i} \text{ or } \mathrm{t} = \mathrm{j} \\ \sigma^{tt}\mathrm{tr}[\hat{\mathbf{B}}'^{-1}\mathbf{W}] & \text{if } \mathrm{t} = \mathrm{i} = \mathrm{j} \end{cases} \\ I_{\sigma_{ij}\sigma_{sr}} = \dfrac{R}{2}\mathrm{tr}[\boldsymbol{\Sigma}^{-1}\mathbf{E}^{ij}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{sr}]; \quad i, j, s, r = 1, 2, ..., T \end{cases} \tag{A54}$$