



**UvA-DARE (Digital Academic Repository)**

**Power properties of invariant tests for spatial autocorrelation in linear regression**

Martellosio, F.

[Link to publication](#)

*Citation for published version (APA):*

Martellosio, F. (2006). Power properties of invariant tests for spatial autocorrelation in linear regression. (UvA-Econometrics Working Paper; No. 2006/01). Amsterdam: Faculteit Economie en Bedrijfskunde.

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <http://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Discussion Paper: 2006/01

# Power properties of invariant tests for spatial autocorrelation in linear regression

Federico Martellosio

[www.fee.uva.nl/ke/UvA-Econometrics](http://www.fee.uva.nl/ke/UvA-Econometrics)

**Amsterdam School of Economics**

Department of Quantitative Economics

Roetersstraat 11

1018 WB AMSTERDAM

The Netherlands

UvA  UNIVERSITEIT VAN AMSTERDAM



# Power Properties of Invariant Tests for Spatial Autocorrelation in Linear Regression

Federico Martellosio\*

University of Amsterdam

March 15, 2006

---

## Abstract

Many popular tests for residual spatial autocorrelation in the context of the linear regression model belong to the class of invariant tests. This paper derives a number of exact properties of the power function of such tests. In particular, we extend the work of Krämer (2005, *Journal of Statistical Planning and Inference* 128, 489-496) by characterizing the circumstances under which the limiting power, as the autocorrelation increases, vanishes. More generally, the analysis in the paper sheds new light on how the power of invariant tests for spatial autocorrelation is affected by the matrix of regressors and by the spatial structure. A numerical study aimed at assessing the practical relevance of the theoretical results is included.

*Keywords:* Cliff-Ord test; invariant tests; linear regression model; point optimal tests; power; similar tests; spatial autocorrelation.

*JEL Classification:* C12, C21.

---

## 1 Introduction

Testing for residual spatial autocorrelation in the context of the linear regression model (e.g., Cliff and Ord, 1981, Anselin, 1988, Cressie, 1993) is now recognized as a crucial step in much empirical work in economics, geography and regional science. The present paper is concerned with finite sample power properties of tests used for this purpose. More specifically, our main objective is to understand how power is affected by the regressors and by the spatial structure.

So far, the power properties of tests for residual spatial autocorrelation have received much less attention than those of tests for residual serial correlation, and have mainly been studied by Monte Carlo simulation (see Florax and de Graaff, 2004, and references therein). Very few

---

\*Department of Quantitative Economics, University of Amsterdam, Roetersstraat 11, 1018 WB, Amsterdam, The Netherlands. E-mail: [F.Martellosio@uva.nl](mailto:F.Martellosio@uva.nl). I am grateful to Grant Hillier and Kees Jan van Garderen for very helpful comments and to Anurag Banerjee for providing GAUSS routines for the Imhof's procedure.

attempts have been made to derive exact properties of such tests, two notable exceptions being King (1981) and Krämer (2005). The former paper has established that the most popular test for spatial autocorrelation in regression residuals, the Cliff-Ord test, is locally best invariant for an important class of alternatives. The latter paper has generalized some results previously available for tests of serial autocorrelation (see Krämer, 1985, and Zeisel, 1989); in particular, Krämer (2005) has shown analytically that there are cases in which the power of some tests for spatial autocorrelation (namely, those whose associated test statistics can be expressed as ratios of quadratic forms in the regression residuals) can vanish as the spatial autocorrelation in the data increases. In general, it is fair to state that, while there is some evidence in the literature that the properties of tests for spatial autocorrelation can be very sensitive to the regressors and to the spatial structure, little is known about which combinations of regressors and spatial structures lead to low or high power.

Of course—given the popularity of the linear regression model and the pervasiveness of the issue of spatial autocorrelation in many empirical investigations—a large number of procedures are available for the purpose of testing for residual spatial autocorrelation, and one can choose among them on the basis of the suspected form of autocorrelation or on the basis of the desired properties of the testing procedure. In this paper, we confine ourselves to a rather simple, but extremely popular, framework. We assume that the regression errors follow a (first-order) *spatial autoregressive process* (e.g., Cressie, 1993) and we focus on *invariant tests* (e.g., Lehmann and Romano, 2005). Even in this simple setup the analytical investigation of exact power properties of tests is complicated. Because of the availability of many approximating techniques for power functions, this is not necessarily a problem when interest lies in the properties of a test in the context of a given model, i.e., when both the matrix of regressors and the spatial structure are fixed. However, when interest is, as in this paper, in how the properties of a test depend on the regressors and on the spatial structure, none of the available numerical or analytical approximations is likely to yield conclusive results. One feature of our approach is that some new properties of the power function of invariant tests are deduced directly from the density of the pertinent *maximal invariant* avoiding the need for complicated expressions for power functions or approximations to them.

Our contributions are as follows. Firstly, we extend the results of Krämer (2005) in several directions: we formulate conditions, that are in general very simple to check, for the limiting (as the autocorrelation increases) power of any invariant test to be 0, 1, or in  $(0, 1)$ ; we prove that, for any given spatial structure and irrespective of the size of the tests, there exists an infinite number of subspaces spanned by the regressors such that the limiting power of a locally best or point

optimal invariant test vanishes; we characterize such “hostile” subspaces. Secondly, we discuss some conditions that are sufficient for unbiasedness of invariant tests for spatial autocorrelation and for monotonicity of their power function. Such conditions are not necessary, but they help to understand the causes of undesirable properties of the tests.

These results call for caution in interpreting the outcome of tests for residual spatial autocorrelation, especially when the number of degrees of freedom is low and large autocorrelation is suspected. The results have also implications outside a formal hypothesis testing framework, because they imply that there are circumstances in which the practice of interpreting the Cliff-Ord statistic, or even the Moran statistic when the model does not contain regressors, as an autocorrelation coefficient (e.g., Cliff and Ord, 1981, Anselin, 1988) cannot be justified.

The remainder of the paper is organized as follows. Section 2 presents the theoretical framework of the paper, i.e., the testing problem and the tests considered. Section 3 analyzes the limiting power of invariant tests for spatial autocorrelation. This is done by first considering the general case of a model with arbitrary regressors, and then specializing the results to zero-mean models. A numerical study of the practical relevance of the results is included. Further insights into the role played by the regressors and the spatial structure in determining the power of invariant tests of autocorrelation are gained in Section 4, by analyzing the conditions for unbiasedness of the tests and monotonicity of their power functions. Section 5 concludes. All proofs are collected in the Appendix.

## 2 The Setup

### 2.1 The Testing Problem

Consider a *fixed* and *finite* set of  $n$  observational units, such as the regions of a country, and let  $y = (y_1, \dots, y_n)'$ , where  $y_i$  denotes the random variable observed at the  $i$ -th (according to some arbitrary ordering) unit. We assume that  $y$  is modelled according to a Gaussian linear regression model, i.e.,

$$y \sim N(X\beta, \sigma^2\Sigma(\rho)), \tag{1}$$

where  $X$  is a non-stochastic  $n \times k$  matrix of rank  $k < n$ ,  $\beta$  is a  $k \times 1$  vector of unknown parameters,  $\sigma^2$  is an unknown positive parameter and  $\rho$  is a scalar unknown parameter belonging to some connected subset  $\Psi$  of the set of values of  $\rho$  such that  $\Sigma(\rho)$  is positive definite. We assume that the function  $\rho \rightarrow \Sigma(\rho)$  is differentiable, and that the parameters of the model are identified (in the sense that the parameter space of the model does not contain two distinct points indexing

the same distribution) and functionally independent. In this paper, we will often refer to the case of a general  $\Sigma(\rho)$ , but will be mostly concerned with the specific covariance structures implied by spatial autoregressive process.

There are two distinct classes of Gaussian spatial autoregressive processes: conditional autoregressive (CAR) processes and simultaneous autoregressive (SAR) processes. They are both discussed extensively in many books and articles in the statistics and econometrics literature (e.g., Whittle, 1954, Besag, 1974, Cliff and Ord, 1981, Anselin, 1988, Cressie, 1993), to which we refer for details concerning the construction and interpretation of the models. Here, we only briefly define the covariance matrices implied by the models. As in most of the theoretical and empirical literature on spatial autoregressive processes, we confine ourselves to *first-order* (or one-parameter) processes. Such processes are specified on the basis of a fixed  $n \times n$  (*spatial weights matrix*)  $W$ , chosen to reflect a priori information on relations among the  $n$  observations. Typically, the  $(i, j)$ -th entry of  $W$  is set to zero if  $i$  and  $j$  are not neighbors according to some metric that is deemed to be relevant for the phenomenon under analysis, and is set to some non-zero number, possibly reflecting the degree of interaction, otherwise. For instance, if the observational units are the regions of a country, one may set  $W(i, j) = 1$  if  $i$  and  $j$  share a common boundary,  $W(i, j) = 0$  otherwise.

Let  $I$  denote the  $n \times n$  identity matrix. Provided it is symmetric and positive definite, the matrix  $\Sigma(\rho)$  implied by a CAR specification is

$$\Sigma(\rho) = (I - \rho W)^{-1} L, \tag{2}$$

where  $L$  is a fixed  $n \times n$  diagonal matrix such that  $L^{-1}W$  is symmetric ( $\sigma^2 L(i, i)$  represents the variance of  $y_i$  conditional on all the remaining random variables in  $y$ ). We remark that, even without reference to CAR models, structure (2) constitutes a very natural framework in which to study tests for autocorrelation; see, e.g., Anderson (1948), Kadiyala (1970), Kariya (1980), King (1980).

On the other hand, provided that  $I - \rho W$  is nonsingular, a SAR process implies

$$\Sigma(\rho) = (I - \rho W)^{-1} V (I - \rho W')^{-1}, \tag{3}$$

where  $V$  is a fixed  $n \times n$  symmetric and positive definite matrix.

For both CAR and SAR models, we assume:

- (i)  $W(i, i) = 0$ , for  $i = 1, \dots, n$ ;
- (ii)  $W(i, j) \geq 0$ , for  $i, j = 1, \dots, n$ ;

(iii)  $W$  is an *irreducible* matrix (e.g., Gantmacher, 1974, Ch. 13).

Condition (i) is required by the way CAR models are constructed, and is assumed for SAR models merely for convenience. Condition (ii) is not required by the definition of the models, but is virtually always satisfied in empirical applications. Condition (iii) is a natural assumption in a spatial context; in graph theoretic terms, it amounts to requiring that the graph with adjacency matrix  $W$  (that is, the graph with the  $n$  observational units as vertices and an edge from  $i$  to  $j$  if and only if  $W(i, j) \neq 0$ ) is *strongly connected*, i.e., has a path between any two distinct vertices (e.g., Cvetković *et al.*, 1980, p. 18). Observe that (non-circular) AR(1) models are not in our class of SAR processes, because the matrix  $W$  necessary to write the covariance matrix of an AR(1) process as in equation (3) would be triangular and hence reducible. Also note that, as a consequence of the symmetry of  $L^{-1}W$ , in CAR models  $W(i, j) = 0$  if and only if  $W(j, i) = 0$ , for  $i, j = 1, \dots, n$ . This implies that, in CAR models,  $W$  can be assumed to be irreducible without loss of generality, because otherwise the model could be decomposed into the product of (at least) two processes.

In the context of model (1), we wish to test the null hypothesis  $\rho = 0$  versus the one-sided alternative  $\rho > 0$  (here and throughout,  $\rho > 0$  stands for  $\mathbb{R}^+ \cap \Psi$ , i.e., we leave it implicit that  $\rho$  must belong to the parameter space of the model). The choice of a one-sided alternative rather than a two-sided one is dictated by the fact that the former is more relevant in the context of many popular specification of  $\Sigma(\rho)$ , including those implied by CAR and SAR processes, due to the interpretation of  $\rho$  in such processes (see below).

We assume—and this is an important point for the reading of the present paper—that  $\Sigma(0) = I$ . Since, as far as our testing purposes are concerned, this does not involve any loss of generality, from now on and unless otherwise specified we reserve the term “CAR model” to refer to the family of distributions

$$N(X\beta, \sigma^2(I - \rho W)^{-1}), \tag{4}$$

(for a fixed  $W$ ) and the term “SAR model” to refer to the family of distributions

$$N(X\beta, \sigma^2 [(I - \rho W')(I - \rho W)]^{-1}), \tag{5}$$

(again, for a fixed  $W$ ). The normalization to  $\Sigma(0) = I$  emphasizes a crucial difference between CAR and SAR models, with regards to our testing problem: for CAR models there is no loss of generality in assuming that  $W$  is a symmetric matrix, whereas for SAR models we have to allow explicitly the possibility of a nonsymmetric  $W$ . In fact, we shall see that there are substantial differences between SAR models with a symmetric weights matrix, henceforth referred

to as *symmetric SAR models*, and SAR models with a nonsymmetric weights matrix—henceforth referred to as *asymmetric SAR models*. The most popular nonsymmetric weights matrices in applications are those obtained by row-standardizing a preliminary matrix (e.g., Anselin, 1988). In the rest of the paper, a *row-standardized*  $W$  is one that can be written as  $W = D^{-1}A$ , where  $A$  is a symmetric  $(0 - 1)$  matrix and  $D$  is the diagonal matrix with  $D(i, i) = \sum_{j=1}^n A(i, j)$ ,  $i = 1, \dots, n$ .

By the Perron-Frobenius theorem,  $W$  admits a positive eigenvalue that is (algebraically and geometrically) simple and non-smaller in modulus than any other eigenvalue. We denote such an eigenvalue by  $\lambda$ . For both CAR and SAR models we take the set  $\mathbb{R}^+ \cap \Psi$  to be the interval  $(0, \lambda^{-1})$ . Such a restriction is necessary for positive definiteness of the covariance matrix of a CAR model. For a SAR model, it is not necessary, but has the advantage of guaranteeing connectedness of the parameter space and of avoiding some undesirable properties of the covariance structure implied by the model.

When  $0 < \rho < \lambda^{-1}$ , it is easily established (e.g., Gantmacher, 1974, p. 69) that conditions (ii) and (iii) imply that the covariance between any two variables  $y_i$  and  $y_j$  in both CAR and SAR models is strictly positive (similarly, it can be shown that when  $\rho < 0$  the covariances may be positive or negative, but not all of them are positive in a left neighborhood of  $\lambda^{-1}$ ). Thus, the hypothesis  $\rho > 0$  represents positive spatial autocorrelation, a much more common phenomenon in practice than negative spatial autocorrelation.

## 2.2 The Tests

This paper is concerned with *invariant* tests (see, e.g., Lehmann and Romano, 2005). Roughly speaking, these are the tests that preserve the symmetries satisfied by the testing problem in question. More precisely, a test is said to be invariant with respect to a certain group of transformations of the sample space if it is based on a test statistic that is constant on each orbit of that group. A necessary and sufficient condition for this type of invariance is that the test statistic is a function of a *maximal invariant* under that group.

Our problem of testing  $\rho = 0$  against  $\rho > 0$  in model (1) is invariant with respect to the group of transformations  $y \rightarrow ay + Xb$ , where  $a \in \mathbb{R}^+$  and  $b \in \mathbb{R}^k$  (e.g., Kariya, 1980, or King, 1980). A maximal invariant under this group is  $v = Cy / \|Cy\|$ , where  $C$  is an  $(n - k) \times n$  matrix such that  $CC'$  is the identity matrix of order  $n - k$  and  $C'C$  is  $M = I - X(X'X)^{-1}X'$ , and  $\|\cdot\|$  denotes the Euclidean norm. For some positive integer  $m$ , let  $S_m = \{v \in \mathbb{R}^m : \|v\| = 1\}$  denote the unit  $m$ -dimensional sphere. The distribution of  $v$  depends on the single parameter  $\rho$ , and

has density, with respect to the normalized Haar measure on  $S_{n-k}$ ,

$$pdf(v; \rho) = |C\Sigma(\rho)C'|^{-\frac{1}{2}} \left[ v' (C\Sigma(\rho)C')^{-1} v \right]^{-\frac{n-k}{2}} \quad (6)$$

(see Kariya, 1980, equation (3.7)). Since  $pdf(v; \rho)$  does not depend on  $v$  when  $\rho$  vanishes, testing  $\rho = 0$  in  $N(X\beta, \sigma^2\Sigma(\rho))$  by means of an invariant test reduces to testing uniformity of  $v$  on  $S_{n-k}$ .

Besides the “principle of invariance”, there are at least two other reasons why invariant tests are particularly appealing for our testing problem. Firstly, invariant tests can be implemented easily. Since an invariant test statistic must depend on  $y$  only through  $v$ , its distribution under the null (and also under the alternative) is free of nuisance parameters, and hence critical values can, in general, be obtained accurately by Monte Carlo or other numerical methods. In fact, the class of *similar* tests for  $\rho = 0$  coincides with that of invariant tests (Hillier, 1987). Secondly, expression (6) turns out to be proportional to the *marginal likelihood* of  $\rho$ , which has often been found to provide a better basis for inference about  $\rho$  than the full likelihood of model (1) (particularly when  $k$  is large with respect to  $n$ ); see, for instance, Diggle (1988), Tunnicliffe Wilson (1989) and Rahman and King (1997).<sup>1</sup>

Despite the elimination of the nuisance parameters achieved in (6), it is well known that, in general, a uniformly most powerful invariant (UMPI) test does not exist for the testing problem under consideration. In such a situation, one can resort to a test that is optimal according to some exact criterion (see, for instance, Cox and Hinkley, 1974, p. 102), or to a test that has less clear-cut optimality properties but performs well in general, such as a likelihood ratio test (which is an invariant test, as proved for instance in Cox and Hinkley, 1974, p. 173) or its restricted version based on  $pdf(v; \rho)$ . The present paper is particularly concerned with the tests—named point optimal invariant (POI) tests by King (1988)—that are the most powerful amongst all invariant tests against a specific alternative  $\rho = \bar{\rho} > 0$ , and with the locally best invariant (LBI) test, which is obtained as the limiting case for  $\bar{\rho} \rightarrow 0$ . In general, and certainly for our testing problem, the locally most powerful test coincides with the test maximizing the slope of the power function at  $\rho = 0$  (see Lehmann and Romano, 2005, p. 339). The choice of POI and LBI tests is mainly motivated by the fact that POI tests define the upper bound (the so-called power envelope, see below) to the power attainable by any invariant test of a fixed size, and by the popularity of LBI tests, especially in the context of the spatial models we are concerned with in this paper.

The size of a critical region (henceforth c.r.) is denoted by  $\alpha$  and, to avoid trivial cases and

---

<sup>1</sup>The literature on the comparison between maximum likelihood and restricted maximum likelihood, REML, is also relevant here, although REML usually refers to the marginal likelihood of all the covariance parameters, i.e., both  $\rho$  and  $\sigma^2$  in our case.

unless otherwise specified, is assumed to be in  $(0, 1)$ . The critical value for a size- $\alpha$  test will be denoted by  $c_\alpha$ . The POI (or best invariant) c.r. at the point  $\bar{\rho}$ , obtained by application of the Neyman-Pearson Lemma to the density (6), is defined by

$$v' (C\Sigma(\bar{\rho})C')^{-1} v < c_\alpha. \quad (7)$$

Denoting by  $\pi_{\bar{\rho}}(\rho)$  the power of such a c.r., the power envelope of size- $\alpha$  invariant tests is the function that associates the value  $\pi_\rho(\rho)$  to each  $\rho \geq 0$ . When needed, we will emphasize the dependence of  $\pi_\rho(\rho)$  on  $X$  by writing  $\pi_\rho(\rho, X)$ . The LBI c.r. for testing  $\rho = 0$  against  $\rho > 0$  is defined by

$$v' CA_0C'v < c_\alpha, \quad (8)$$

where  $A_0 = d\Sigma^{-1}(\rho)/d\rho|_{\rho=0}$  (King and Hillier, 1985). When  $-A_0$  is equal to some spatial weights matrix  $W$  (or to  $W + W'$ ), as it is in the case of CAR or SAR models, the LBI test is known in the literature as Cliff-Ord test (see Cliff and Ord, 1981, and King, 1981). The Cliff-Ord test represents the generalization to regression residuals of the Moran test (Moran, 1950), and is, by far, the most popular test for spatial autocorrelation in regression models.<sup>2</sup>

Before we continue, some notation is in order. For a  $q \times q$  symmetric matrix  $Q$ , we denote by  $\lambda_1(Q), \dots, \lambda_q(Q)$  its eigenvalues, labeled in non-decreasing order of magnitude; by  $m_i(Q)$  the multiplicity of  $\lambda_i(Q)$ , for  $i = 1, \dots, q$ ; by  $f_1(Q), \dots, f_q(Q)$  a set of orthonormal (with respect to the Euclidean norm) eigenvectors of  $Q$ , with the eigenvector  $f_i(Q)$  being pertinent to the eigenvalue  $\lambda_i(Q)$ ; by  $E_i(Q)$  the eigenspace associated to  $\lambda_i(Q)$ , for  $i = 1, \dots, q$ . In all of the above quantities, we suppress the reference to  $Q$  when  $Q$  is a (symmetric, with  $q = n$ ) weights matrix. Thus, for a symmetric  $W$ ,  $\lambda_n = \lambda$ .

### 3 Limiting Power

This section contains the main results of the paper. Broadly speaking, they concern the role of the regressors in determining power properties of invariant tests for autocorrelation. In Section 3.1, we discuss some preliminary results in the context of the general model (1). Then, we focus on the limiting behavior of the power function, as the autocorrelation increases, in CAR and SAR models, with general regressors in Section 3.2, and without regressors in Section 3.3. Finally,

---

<sup>2</sup>Being based on only the first derivative of  $\Sigma(\rho)$ , the LBI tests for  $\Sigma(\rho) = I$  have generally non-trivial power against a large class of alternative specifications of  $\Sigma(\rho)$ , which is at the same time a strength and a weakness of such tests. In any case, this does not detract from the fact that it is important to study their performance against particular alternatives, spatial autoregressive models in our case.

in Section 3.4 we report results from numerical experiments aimed at assessing the practical relevance of the theoretical results.

### 3.1 Preliminaries

Consider the issue of how  $X$  in  $N(X\beta, \sigma^2\Sigma(\rho))$  affects the the power properties of invariant tests of  $\rho = 0$  versus  $\rho > 0$ , for some covariance structure  $\Sigma(\rho)$ . The following proposition sets the scene for the analysis to follow. It is concerned with comparing the envelope  $\pi_\rho(\rho, X)$ , for an  $X \neq 0$ , with the envelope when  $X = 0$  (here and throughout a zero matrix is simply denoted by 0).

**Proposition 3.1** *Consider testing  $\rho = 0$  versus  $\rho > 0$  in model  $N(X\beta, \sigma^2\Sigma(\rho))$ . For any  $X \neq 0$ , any  $\rho > 0$ , and any  $\alpha$ ,*

$$\pi_\rho(\rho, X) \leq \pi_\rho(\rho, 0). \tag{9}$$

*In (9) equality is attained if and only if, for some  $i = 2, \dots, n - 1$ ,  $\text{col}(X) \subseteq E_i(\Sigma(\rho))$  and  $\alpha = \Pr(v'\Sigma^{-1}(\rho)v < \lambda_i^{-1}(\Sigma(\rho)))$ .*

The conditions for equality in (9) are extremely restrictive, because they pose very severe constraints on  $X$ ,  $\alpha$ , and  $\Sigma(\rho)$ . Proposition 3.1 asserts that, except when these conditions are met, the presence of any  $X \neq 0$  in  $N(X\beta, \sigma^2\Sigma(\rho))$  has a detrimental effect (with respect to the case  $X = 0$ , and as long as  $\beta$  is unknown) on the maximum power achievable by an invariant test for testing  $\rho = 0$  versus  $\rho > 0$ .

Two comments arise naturally from Proposition 3.1. The first comment is that the comparison in the proposition involves models (the one with  $X = 0$  and the one with an  $X \neq 0$ ) with different degrees of freedom. An interesting question is which matrices  $X$  of a fixed dimension  $n \times k$  are favorable, and which are less favorable, to our testing purposes (from the point of view of the maximum power achievable by invariant tests). Such a question is a difficult one, because, for a given matrix  $\Sigma(\rho)$ , in general the answer *depends on  $\rho$  and  $\alpha$* . Some partial answers are available in the literature for regression models with AR(1) errors; see, e.g., Tillman (1975). The second comment is that, in practice, one is usually more concerned with the power of a specific test than with the power envelope. For a general  $\Sigma(\rho)$ , Proposition 3.1 does certainly not imply that the power function of a particular test when  $X = 0$  is uniformly (over  $\rho > 0$ ) non-smaller than when  $X \neq 0$  (it is interesting, however, that such an implication does hold when  $\Sigma(\rho)$  is that of a CAR model and the test in question is a POI or LBI test, because for a zero-mean CAR model the

POI c.r. (7) does not depend on  $\bar{\rho}$ , i.e., there exists a UMPI test, and hence the power function of any POI or LBI test coincides with  $\pi_\rho(\rho, 0)$ .

To deal with the issues raised in the previous paragraph, we will focus on large values of  $\rho$  in CAR and SAR models. Exact power properties of invariant tests will be deduced directly from the density of the maximal invariant  $v$ . For convenience, we now list some fundamental properties of  $pdf(v; \rho)$ , valid for any fixed  $\rho$ . Let  $\Omega_\rho = C\Sigma(\rho)C'$ , and let  $\tilde{E}_i(\Omega_\rho) = S_{n-k} \cap E_i(\Omega_\rho)$ ,  $1 \leq i \leq n - k$ . The density  $pdf(v; \rho)$  is antipodally symmetric (that is,  $pdf(v; \rho) = pdf(-v; \rho)$ ) and, more specifically, is constant on the regions of constant  $v'\Omega_\rho^{-1}v$  (geometrically, such regions are the intersection of the surfaces of a sphere and of an ellipsoid in  $\mathbb{R}^{n-k}$ ). It follows immediately that:

- (i)  $pdf(v; \rho)$  is maximized over  $S_{n-k}$  when  $v'\Omega_\rho^{-1}v$  is minimized, that is, when  $v \in \tilde{E}_{n-k}(\Omega_\rho)$ .  
Note that  $\tilde{E}_{n-k}(\Omega_\rho)$  consists of two antipodal points if and only if  $m_{n-k}(\Omega_\rho) = 1$ ;
- (ii)  $pdf(v; \rho)$  is strictly decreasing as  $v$  moves from  $\tilde{E}_i(\Omega_\rho)$  to  $\tilde{E}_j(\Omega_\rho)$  along any geodesic of  $S_{n-k}$ , for any  $1 \leq j < i \leq n - k$ ;
- (iii) upon rotation to a coordinate system provided by a set of  $n - k$  orthogonal eigenvectors of  $\Omega_\rho$ ,  $pdf(v; \rho)$  is invariant to the sign of each component of the vector  $v$ .

Property (i) will be used to derive some of the results to follow. Property (ii) implies that any invariant c.r. that is not centrally symmetric (we say that an invariant c.r.  $\Phi \in S_{n-k}$  is centrally symmetric if  $t \in \Phi$  implies  $-t \in \Phi$ ) is dominated uniformly over  $\rho$  (in terms of power) by a centrally symmetric c.r. of the same size. Because of this reason, from now on we assume that an invariant c.r. is centrally symmetric. This corresponds to enlarging the group of transformations with respect to which we require invariance to include also the transformation  $y \rightarrow -y$ . Property (iii) is not exploited directly in this paper, but is very useful when thinking geometrically about our testing problem, for it implies that the study of  $pdf(v; \rho)$  can be limited to a single orthant of  $S_{n-k}$ .

The following preliminary result links the limit, as  $\rho$  approaches some positive value  $a$  (from the left, and with  $a$  an accumulation point of  $\Psi$ ), of the power function of an arbitrary invariant c.r. to the limiting eigenstructure of  $\Omega_\rho$ . We denote  $\lim_{\rho \rightarrow a} \Omega_\rho$  by  $\Omega$ , the limit being taken entrywise.

**Lemma 3.2** *Suppose that  $\Sigma(\rho)$  is positive definite for  $\rho \in (0, a)$  and for  $\rho \rightarrow a$ . If  $\lambda_{n-k}(\Omega)$  is finite, then the power of any invariant c.r. for testing  $\rho = 0$  against  $\rho > 0$  in model  $N(X\beta, \sigma^2\Sigma(\rho))$  tends, as  $\rho \rightarrow a$ , to a number strictly between 0 and 1. If  $\lambda_{n-k}(\Omega)$  is infinite and simple, then the*

power of an invariant c.r. for the same testing problem tends, as  $\rho \rightarrow a$ , to 1 if the c.r. contains  $f_{n-k}(\Omega)$ , to 0 otherwise.

Lemma 3.2 holds for a very general class of matrices  $\Sigma(\rho)$ , including CAR and SAR models, and the (time-series) stationary AR(1) model. For the latter model, the power of the Durbin-Watson and some related tests as  $\rho \rightarrow 1$  has been investigated extensively; see Krämer (1985), Zeisel (1989) and Bartels (1992). Lemma 3.2 shows how some results on the power of such tests can be extended to any invariant (similar) test for serial correlation.

Of course, when  $a \in \Psi$  the power of any c.r. must be in  $(0, 1)$ , and this is reflected in Lemma 3.2 by the fact that  $\lambda_{n-k}(\Omega)$  is finite when  $a \in \Psi$ . The possibility that, in the setting of Lemma 3.2, the power of a certain c.r. vanishes as  $\rho$  goes to the boundary of  $\Psi$  should be regarded as a problem of the statistical model, rather than of a particular test. A simple geometric argument clarifies this point. Let  $\nu = n - \lim_{\rho \rightarrow a} \{\text{rank}(\Sigma^{-1}(\rho))\}$ , with  $\Sigma(\rho)$  as in Lemma 3.2. Note that for  $\lambda_{n-k}(\Omega) = \infty$  it is necessary that  $\nu > 0$ . When  $\nu > 0$ , the model  $N(X\beta, \sigma^2\Sigma(\rho))$  tends, as  $\rho \rightarrow a$ , to a family of (improper) distributions defined on a  $\nu$ -dimensional subspace, say  $S_\nu$ , of  $\mathbb{R}^n$ . This is easily seen by observing that the contours of  $N(X\beta, \sigma^2\Sigma(\rho))$  are the ellipses  $(y - X\beta)' \Sigma^{-1}(\rho)(y - X\beta) = k$ , which also shows that, for any fixed  $\beta$ ,  $S_\nu$  is the translation of  $\lim_{\rho \rightarrow a} \{E_n(\Sigma(\rho))\}$  by  $X\beta$ . It is then clear that the limiting power, as  $\rho \rightarrow a$ , of a certain test—not necessarily invariant—depends on the position of the c.r. in  $\mathbb{R}^n$  relative to  $S_\nu$  (if the test is invariant, the relative position does not depend on  $\beta$ ). In particular, the power of a test vanishes whenever the intersection between  $S_\nu$  and the c.r. has 1-dimensional Lebesgue measure zero. In CAR and SAR models (when  $a = \lambda^{-1}$ ) and in stationary AR(1) models (when  $a = 1$ ),  $\nu = 1$ .<sup>3</sup> Note that the stationarity assumption on the AR(1) model is not superfluous, in that generally  $\nu = 0$  otherwise. On the contrary, for the CAR and SAR models considered in this paper  $\nu$  is always 1. This represents an important difference between time-series and spatial autoregressive models, from the point of view of testing for residual autocorrelation.

Clearly, the above geometric argument does not depend on the normality assumption, but holds for any elliptically symmetric distribution. It also holds for any c.r.; when the c.r. is invariant, the conditions in Lemma 3.2 can be exploited. Moreover, further progress can be made by focusing on a specific class of matrices  $\Sigma(\rho)$ , those implied by CAR and SAR models in

---

<sup>3</sup>A very similar situation occurs in regression models such that  $\Sigma(\rho)$ , rather than  $\Sigma^{-1}(\rho)$ , tends to a singular matrix as  $\rho$  tends to some value  $a$ . In this case, the distributions are defined, as  $\rho \rightarrow a$ , in a subspace of  $\mathbb{R}^n$  of dimension  $\lim_{\rho \rightarrow a} \{\text{rank}(\Sigma(\rho))\}$ . Examples are a spatial moving average model (i.e., a model with covariance matrix equal to the inverse of that of a SAR model), and a fractionally integrated white noise, with  $\rho$  being the differencing parameter and  $a = 1/2$  (see Kleiber and Krämer, 2005).

the rest of this section.

### 3.2 Main Results

In this subsection we focus on the power of invariant tests in CAR and SAR models when  $\rho \rightarrow \lambda^{-1}$  (from the left). Accordingly, from now on, by “limiting power” we mean the limit of the power function as  $\rho \rightarrow \lambda^{-1}$ . The restriction to the case  $\rho \rightarrow \lambda^{-1}$  is of practical relevance, because (a) it corresponds to studying power when it is most needed, i.e., when the autocorrelation in the data, and hence the inefficiency of the OLS estimator of  $\beta$ , is large; (b) often, in order to fit real data, spatial autoregressive models require a large value of  $\rho$  (e.g., Besag and Kooperberg, 1995).

In order to state the key result of this section some new notation is needed. Henceforth, an invariant critical region defined as a subset of  $S_{n-k}$  is denoted by  $\Phi_v$ , whereas its image on the sample space  $\mathbb{R}^n$  is denoted by  $\Phi_y$ . The column space of the matrix  $X$ , often referred to as the “regression space”, is denoted by  $\text{col}(X)$ . The entrywise positive and normalized eigenvector of  $W$  pertaining to  $\lambda$  is denoted by  $f$ . Existence and uniqueness of  $f$  are guaranteed by the Perron-Frobenius theorem.

**Theorem 3.3** *In CAR and SAR models, the limiting power of an invariant c.r.  $\Phi_y$  for testing  $\rho = 0$  against  $\rho > 0$  is:*

- in  $(0, 1)$  if  $f \in \text{col}(X)$ ;
- 1 if  $f \in \Phi_y \setminus \text{col}(X)$ ;
- 0 otherwise.

The theorem asserts that, to some degree, the limiting power of  $\Phi_y$  is determined by which of three mutually disjoint subsets of the sample space  $f$  belongs to. Of course, the result can be restated on the space  $S_{n-k}$  of the maximal invariant, in which case  $f$  must be replaced by  $Cf/\|Cf\|$  and the three subsets become  $\{0\}$ ,  $\Phi_v \setminus \{0\}$  and  $\overline{\Phi_v \cup \{0\}}$ .

Theorem 3.3 is strongly related to Theorems 1 and 2 in Krämer (2005), the most important differences being: (a) the class of tests considered there (i.e., tests that can be expressed as ratios of quadratic forms in the regression residuals) and the class considered in the present paper (i.e., invariant tests) are different, although they certainly intersect; (b) our result does not require symmetry of  $W$ . We stress that Theorem 3.3 holds for *any* invariant (similar) c.r., regardless of the analytical form of the associated test statistic. Thus, it also holds for tests whose test statistics are analytically complicated, or, as in the case of a likelihood ratio test based on the full or the marginal likelihood, unavailable in closed form.

The practical usefulness of Theorem 3.3 is in providing simple conditions for the limiting power of any invariant c.r. to vanish, given any matrices  $X$  and  $W$ . Consider an invariant c.r. that rejects  $\rho = 0$  for small values of some statistic  $T(y)$ , i.e.,

$$\Phi_y = \{y \in \mathbb{R}^n : T(y) < c_\alpha\}. \quad (10)$$

Theorem 3.3 asserts that the limiting power of such a c.r. is 0 if  $T(f) < c_\alpha$ , 1 if  $T(f) \geq c_\alpha$ , in  $(0,1)$  if  $f \in \text{col}(X)$ . These conditions are typically very simple to check because, in most cases, (i)  $f$  is either known (e.g., it is a vector of identical entries when  $W$  is row-standardized) or can be computed efficiently (e.g., by the power method); (ii) since  $\Phi_y$  is similar,  $c_\alpha$  can be obtained accurately by simulation methods. For instance, it is readily verified that, for CAR or SAR models, the limiting power of a POI test is 0, 1, or strictly between 0 and 1, depending on whether

$$f'(\Sigma(\bar{\rho})M_{\bar{\rho}} - c_\alpha M) f \quad (11)$$

is respectively positive, negative, or zero (zero occurring if and only if  $f \in \text{col}(X)$ ). Analogously, the limiting power of a LBI test is 0, 1, or strictly between 0 and 1, depending on whether

$$f'(MA_0M - c_\alpha M) f \quad (12)$$

is respectively positive, negative or zero. Note that, since they refer to test statistics that are ratios of quadratic forms in  $y$ , the conditions based on (11) and (12) reduce, in the case of a symmetric SAR model, to conditions given in Krämer (2005). The specific form of the test statistics also implies that, for POI or LBI tests,  $c_\alpha$  can also be obtained by exploiting one of the many approximations available for the distribution of a quadratic form in a vector uniformly distributed on a hypersphere.

Further remarks concerning Theorem 3.3 follow.

**Remark 1** The condition  $f \in \text{col}(X)$ , under which the limiting power of an invariant test is in  $(0,1)$ , is satisfied whenever  $W$  in a CAR or SAR model is row-standardized and an intercept is included in the regression, because row-standardization implies that  $f$  has identical entries. Note that here whether  $W$  refers to a model before or after normalization to  $\Sigma(0) = I$  is irrelevant, because the condition  $f \in \text{col}(X)$  is invariant under any invertible linear transformation of  $y \sim N(X\beta, \sigma^2\Sigma(\rho))$ , where  $\Sigma(\rho)$  is that of a CAR or SAR model. For any weights matrix that is not row-standardized, in general  $f \notin \text{col}(X)$ , with the consequence that the limiting power of an invariant test is either 0 or 1.

**Remark 2** An important and immediate consequence of Theorem 3.3 is that, in CAR and SAR models, the limit of the envelope  $\pi_\rho(\rho)$  as  $\rho \rightarrow \lambda^{-1}$  is 1 if  $f \notin \text{col}(X)$ , and is in  $(\alpha, 1)$  otherwise. Hence, as  $\rho \rightarrow \lambda^{-1}$  in CAR and SAR models, the null hypothesis  $\rho = 0$  can be distinguished (by means of an invariant tests) from the alternative hypothesis  $\rho > 0$  with zero type II error probability only if  $f \notin \text{col}(X)$ .

**Remark 3** Consider an invariant test, constructed on the basis of some assumed weights matrix. By Theorem 3.3, whether its limiting power is 0, 1, or in  $(0, 1)$  depends on the “true”  $W$  appearing in the CAR or SAR model only through  $f$  (which, for instance, is the same for any row-standardized  $W$ ), and does not depend on whether the model is a CAR or a SAR model. This property implies some robustness of invariant tests to model misspecification, when the spatial autocorrelation is large.

In the rest of this subsection we take a close look at the case in which the limiting power vanishes, and, consequently, we restrict attention to the case  $f \notin \text{col}(X)$ .

Suppose that, for a given CAR or a SAR model, one finds, by application of Theorem 3.3, that the limiting power of a certain c.r.  $\Phi_y$  vanishes. Theorem 3.3 itself guarantees that if  $\Phi_y$  is enlarged so as to include  $f$ , then its limiting power jumps to 1. From a practical point of view, a question of concern is how large  $\Phi_y$  must be in order to avoid the vanishing of the limiting power. We define (allowing, for convenience and contrary to what is done elsewhere in the paper,  $\alpha$  to take the value 1):

**Definition 1** For an invariant c.r. for testing  $\rho = 0$  against  $\rho > 0$  in a CAR or SAR model,  $\alpha^*$  is the infimum of the set of values of  $\alpha \in (0, 1]$  such that the limiting power does not vanish.

When  $f \notin \text{col}(X)$ ,  $\alpha^*$  is a measure of the distinguishability between the null hypothesis  $\rho = 0$  and the alternative  $\rho \rightarrow \lambda^{-1}$ . A large  $\alpha^*$  indicates that a large size is necessary to avoid the zero limiting power problem;  $\alpha^* = 1$  means that the limiting power is 0 for any  $\alpha$ ;  $\alpha^* = 0$  indicates that the limiting power is 1 for any  $\alpha$ .<sup>4</sup>

When an invariant c.r. is in form (10) (and  $f \notin \text{col}(X)$ ),  $\alpha^*$  is the probability that  $T(y) < T(f)$  under the null hypothesis  $y \sim N(X\beta, \sigma^2 I)$ , or, by invariance,

$$\alpha^* = \Pr(T(y) < T(f); y \sim N(0, I)). \quad (13)$$

---

<sup>4</sup>Recall that we are here focusing on the case  $f \notin \text{col}(X)$ . If  $f \in \text{col}(X)$ ,  $\alpha^*$  is always zero, by Theorem 1, and hence uninformative. In order to study the power of invariant tests when  $f \in \text{col}(X)$ , one could define  $\alpha^*$  as the infimum of the set of values such that the limiting power is greater than some positive value, but this is not pursued in the present paper.

Thus,  $\alpha^*$  can be computed accurately by simulation or other numerical methods. We stress that  $\alpha^*$  depends on  $X$  (through  $\text{col}(X)$ , because of the invariance property of the tests),  $W$ , the choice of test and the choice between a CAR and a SAR specification. In particular, for a given error process, a given test, and a given  $k$ ,  $\alpha^*$  may depend to a very large extent on  $\text{col}(X)$ . Numerical examples will be given in the next subsection. In the following, we will explore the dependence of  $\alpha^*$  on  $\text{col}(X)$  by studying the circumstances in which  $\alpha^* = 0$  and those in which  $\alpha^* = 1$ . We will first give a lemma that holds for any test based on a quadratic form in the maximal invariant, and then we will apply the lemma to POI and LBI tests. Extensions of the analysis below to more general tests (a likelihood ratio test, say) are possible, but may be more involved.

Consider an invariant c.r. of the form

$$\Phi_v(B) = \{v \in S_{n-k} : v' B v < c_\alpha\}, \quad (14)$$

where  $B$  is an  $(n-k) \times (n-k)$  known symmetric matrix independent of  $\alpha$ . For instance, any c.r. based on a ratio of quadratic forms in the OLS residuals can be written in this form. Typically  $B$  will depend on  $X$  and  $W$  (but could depend on a weights matrix different from the one appearing in CAR models, thus allowing for the possibility of misspecification of  $W$ ). We have:

**Lemma 3.4** *Consider, in the context of CAR or symmetric SAR models, testing  $\rho = 0$  against  $\rho > 0$  by means of a c.r.  $\Phi_v(B)$ . Provided that  $f \notin \text{col}(X)$ ,  $\alpha^* = 0$  if and only if  $Cf \in E_1(B)$ , and  $\alpha^* = 1$  if and only if  $Cf \in E_{n-k}(B)$ .*

Lemma 3.4 implies that, in a CAR or SAR model and for a c.r.  $\Phi_v(B)$ ,  $\alpha^* \in (0, 1)$  as long as  $Cf$  is not an eigenvector of  $B$  associated to the smallest or the largest eigenvalue of  $B$ . The important question remains of whether the extremes  $\alpha^* = 0$  and  $\alpha^* = 1$  are attainable, and, if so, in which circumstances. In particular, it is of interest to understand whether for a fixed  $W$  in a CAR or SAR model,  $\alpha^*$  has a non-trivial (i.e., smaller than 1) upper bound as  $\text{col}(X)$  ranges over the set of all subspaces of  $\mathbb{R}^n$  of low (with respect to  $n$ ) dimension. Obviously, given a certain model and a certain c.r., one would hope that  $\alpha^*$  is small, since the limiting power vanishes whenever  $\alpha^* > \alpha$ .

To answer the above question we focus on POI tests ( $B = \Omega_\rho^{-1}$ ) and LBI tests ( $B = CA_0C'$ ). First, we consider the case  $\alpha^* = 0$ . Two conditions that are easily seen to lead to  $\alpha^* = 0$  (i.e., to  $Cf \in E_1(B)$ ) for POI and LBI tests are (i)  $W$  symmetric and  $X = 0$  and (ii)  $W$  symmetric and  $f \perp \text{col}(X)$ . More generally, the following sufficient condition can be established.

**Proposition 3.5** *Consider, in the context of CAR or symmetric SAR models, testing  $\rho = 0$  against  $\rho > 0$  by means of a POI or LBI c.r. Provided that  $f \notin \text{col}(X)$ ,  $\alpha^* = 0$  if  $E_{n-k}(\Omega_\rho)$  does*

not depend on  $\rho$  for  $\rho > 0$ .

**Remark 4** Although the condition in Proposition 3.5 is not necessary, a simple geometric argument suggests that when the condition is not met  $\alpha^*$  is 0 only in very special circumstances. Let the center of the c.r.  $\Phi_v$  based on a certain test statistic be the set of points of  $S_{n-k}$  that are in  $\Phi_v$  for any  $\alpha$ . For instance, the center of  $\Phi_v(B)$  is  $E_1(B) \cap S_{n-k}$ . Clearly, for a certain c.r.,  $\alpha^* = 0$  if and only if, as  $\rho \rightarrow \lambda^{-1}$ ,  $pdf(v; \rho)$  vanishes anywhere outside the center of that c.r. Now, as long as  $f \notin \text{col}(X)$ ,  $pdf(v; \rho)$  tends, as  $\rho \rightarrow \lambda^{-1}$ , to be concentrated in the direction of  $v = Cf$  (see Section 3.1). Thus, for a POI test ( $B = \Omega_{\bar{\rho}}^{-1}$ ),  $\alpha^* = 0$  if and only if  $Cf \in E_1(\Omega_{\bar{\rho}}^{-1}) = E_{n-k}(\Omega_{\bar{\rho}})$ . This is the case if  $E_{n-k}(\Omega_{\rho})$  does not depend on  $\rho$  for  $\rho > 0$  (by Proposition 3.5), but, otherwise, poses a strong restriction on the trajectories described on  $S_{n-k}$  by the eigenvectors in  $E_{n-k}(\Omega_{\bar{\rho}})$ .

We now turn to characterize the case  $\alpha^* = 1$  for POI and LBI tests. Theorem 1 of Krämer (2005), contains the crucial statement that, in symmetric SAR models, “given any matrix  $W$  of weights, and independently of sample size, there is always some regressor  $X$  such that for the Cliff–Ord test the limiting power disappears” (note that here “some regressor  $X$ ” means  $k = 1$ ). Now, from Theorem 3.3 it is clear that whether or not a particular  $X$  (with  $k \geq 1$ ) causes the limiting power to disappear depends on  $\alpha$ . Thus, if interpreted as holding for any  $\alpha$  (less than 1), the above statement would imply that for any  $W$  there exist some particularly hostile regressors that cause a zero limiting power even when the size of the Cliff–Ord c.r. (i.e., the LBI c.r.) is very large. This is clearly an extremely strong property, in a negative sense, of a c.r. Unfortunately, whether it holds or not for the Cliff–Ord test in the context of a symmetric SAR model remains to be established, because the proof of Krämer’s theorem holds only when  $\alpha \rightarrow 0$ .<sup>5</sup> The next theorem settles the issue and places it in a more general context. Recall that  $m_1$  denotes the multiplicity of  $\lambda_1$ , for a symmetric  $W$ . Unless  $W$  satisfies particular symmetries, generally  $m_1 = 1$  (see, for instance, Biggs, 1993).

**Theorem 3.6** *Consider, in the context of CAR or symmetric SAR models, testing  $\rho = 0$  against  $\rho > 0$  by means of a POI or LBI c.r. For any fixed  $W$ , there exist  $m_1$ -dimensional regression spaces such that the limiting power of the selected c.r. vanishes, irrespective of  $\alpha$ . For instance, when  $m_1 = 1$ , let  $X$  be a vector proportional to  $f_1 + bf$ , for some  $b \in \mathbb{R}$ . Then, the limiting power of POI and LBI tests vanishes, irrespective of  $\alpha$ , if  $|b| \geq b^*$ , where  $b^*$  is a threshold that depends*

<sup>5</sup>This is because  $d_1$  in equation (12) of Krämer (2005) is not necessarily positive for any  $W$ , unless  $\alpha \rightarrow 0$ . As a consequence, the regressors that Krämer constructs in his proof do not need to cause the limiting power to vanish when  $d_1 < 0$ .

on the model and on the c.r. Namely, letting

$$b_1 = \left( \frac{\lambda - \lambda_2}{\lambda_2 - \lambda_1} \right)^{\frac{1}{2}}, \quad b_2 = \frac{1 - \bar{\rho}\lambda_1}{1 - \bar{\rho}\lambda}, \quad b_3 = \frac{2 - \bar{\rho}(\lambda + \lambda_2)}{2 - \bar{\rho}(\lambda_2 + \lambda_1)},$$

$b^*$  is equal to  $b_1\sqrt{b_2}$  for a POI c.r. in a CAR model,  $b_1b_2b_3$  for a POI c.r. in a symmetric SAR model,  $b_1$  for a LBI c.r. in both models.

Theorem 3.6 establishes that, for any fixed  $W$  in a CAR or symmetric SAR model, there are  $m_1$ -dimensional regression spaces such that  $\alpha^* = 1$ . In the presence of such regression spaces, the zero limiting power problem *cannot* be solved by increasing  $\alpha$ . Note that if an  $m_1$ -dimensional regression space causes a zero limiting power (of a POI or LBI test in a CAR or symmetric SAR model), then also all the  $k$ -dimensional regression spaces, with  $k \geq m_1$ , that contain it but do not contain  $f$  will yield a zero limiting power, as an obvious consequence of the fact that the power of an invariant test does not depend on  $\beta$ .

We now aim to show that, in the context of Theorem 3.6, the vanishing of the limiting power is not an event of measure zero (in a sense to be specified). In order to do so, it is convenient to introduce some new notation. Let  $G_{k,n}$  denote the set, known as a Grassmann manifold, of all  $k$ -dimensional subspaces of  $\mathbb{R}^n$ , and let  $H_k(\alpha) \subseteq G_{k,n}$ , for  $0 < \alpha < 1$ , be the set of  $k$ -dimensional  $\text{col}(X)$  such that the limiting power of a POI or LBI c.r. of size less than  $\alpha$  vanishes (for some CAR or symmetric SAR model). Clearly,  $H_k(\alpha_1) \subseteq H_k(\alpha_2)$  for any  $\alpha_1 \geq \alpha_2$ .

A natural measure of the size of  $H_k(\alpha)$  is the probability that  $\text{col}(X) \in H_k(\alpha)$ , as  $\text{col}(X)$  ranges over  $G_{k,n}$  according to some probability distribution (with respect to the invariant measure on  $G_{k,n}$ , as given in James, 1954). Such a probability, which we denote by  $z_\alpha$ , can be interpreted as the *probability of a zero limiting power* of a size- $\alpha$  POI or LBI test, in a CAR or symmetric SAR model (we stress that, for each realization of  $X$ , POI and LBI tests are derived by treating  $X$  as fixed). We have:

**Proposition 3.7** *Consider, in the context of CAR or symmetric SAR models, testing  $\rho = 0$  against  $\rho > 0$  by means of a POI or LBI c.r. If  $\text{col}(X)$  has density that is almost everywhere positive on  $G_{k,n}$ ,  $k \geq m_1$ , then  $z_\alpha > 0$ , for any  $W$  and regardless of how large  $\alpha$  or  $n - k$  is.*

Clearly, in some circumstances  $z_\alpha$  can be very small (e.g.,  $z_\alpha$  is usually small when  $n - k$  or  $\alpha$  are large). The important point made by Proposition 3.7 is that, under the stated conditions,  $z_\alpha$  is *never* zero. In Section 3.4 we will compute  $z_{0.05}$  numerically for some choice of  $W$  and of the probability distribution of  $\text{col}(X)$ .

We now provide an interpretation of the  $m_1$ -dimensional regression spaces  $\text{col}(X)$  that, according to Theorem 3.6, are particularly hostile for testing  $\rho = 0$  versus  $\rho > 0$  when  $\rho$  is large.

Starting from  $m_1 = 1$ , Theorem 3.6 asserts that the set of such regression spaces is a region, defined by  $b^*$ , of the plane spanned by  $f_1$  and  $f$ ; for a POI test, it is easily seen that, as  $\bar{\rho}$  increases, this set becomes smaller and more concentrated in the direction of  $f$ . Generalizing to  $m_1 \geq 1$ , the set of the hostile regression spaces is a certain region of the  $(m_1 + 1)$ -dimensional subspace of  $\mathbb{R}^n$  spanned by the vectors  $f_1, \dots, f_{m_1}, f$  (see the proof of the theorem). Consider the Moran statistic  $x'Wx/x'x$  associated to a vector  $x \in \mathbb{R}^n$  and a symmetric  $W$  (the standard version of the Moran statistic would include a normalizing factor and a correction for the sample mean of  $x$  that are not relevant here). By the Rayleigh-Ritz theorem (e.g., Horn and Johnson, 1985),  $f$  represents a vector that is most autocorrelated according to the Moran statistic, and  $f_1, \dots, f_{m_1}$  represent vectors that are least autocorrelated. Note that  $\lambda_n$ , the value of the Moran static when  $x = f$ , is positive by the Perron-Frobenius theorem, and  $\lambda_1$ , the value of the Moran static when  $x = f_1, \dots, f_{m_1}$ , is negative for  $\text{tr}[W] = \sum_{i=1}^n \lambda_i = 0$  by assumption. Thus, Theorem 3.6 asserts that in CAR and symmetric SAR models it is difficult, or even impossible, to detect large positive spatial autocorrelation in the presence of regressors that can be expressed as the sum of a strongly positively autocorrelated component and a strongly negatively autocorrelated component, with the former component being the dominant one.

We mention an extension of Theorem 3.6 that is directly related to the interpretation just given, and can be proved similarly to Theorem 3.6. If  $f_j + bf \in \text{col}(X)$ , with  $f_j \notin E_{n-1}$  and  $f_j \notin E_n$ , the limiting power of a LBI test in a CAR or symmetric SAR model is 1 for any  $\alpha$  (i.e.,  $\alpha^* = 0$ ) provided that

$$|b| \leq \left( \frac{\lambda - \lambda_{n-1}}{\lambda_{n-1} - \lambda_j} \right)^{\frac{1}{2}}.$$

Expressions of this sort can be used to infer how  $W$  affects (through its spectrum, under Gaussianity) the power properties of tests of  $\rho = 0$ . For instance, if  $W$  is such that  $\lambda - \lambda_{n-1}$  is large, then any vector  $X$  in a large region of the plane spanned by  $f_j$  and  $f$  yields  $\alpha^* = 0$ .

Some further remarks concerning Theorem 3.6 end this section.

**Remark 5** With regards to the statement from Krämer (2005) reported above, Theorem 3.6 (i) establishes that the statement is correct when  $m_1 = 1$ ; (ii) provides a generalization to the case  $m_1 > 1$ ; (iii) provides a generalization to POI tests and to CAR models.

**Remark 6** The strongest implication of Theorem 3.6 is perhaps that regression spaces such that the limiting power of a POI test vanishes exist even when  $\bar{\rho}$  is large (i.e., close to  $\lambda^{-1}$ ) and  $\alpha$  is large. This is surprising because, by Proposition 4.1 below, the power at  $\bar{\rho}$  of a POI test must be larger than  $\alpha$ . Since, if  $f \notin \text{col}(X)$ ,  $\pi_\rho(\rho) \rightarrow 1$  as  $\rho \rightarrow \lambda^{-1}$  (see Remark 2), it also holds that

the supremum—as  $\text{col}(X)$  ranges over the set of all  $k$ -dimensional,  $k \geq m_1$ , subspaces of  $\mathbb{R}^n$ —of the maximum shortcoming (e.g., Lehmann and Romano, 2005, p. 337) of any POI or LBI c.r. is always one, for any  $W$  and any  $\alpha$ .

**Remark 7** For POI and LBI tests and for any  $W$ ,  $T(f)$ , regarded as a function from  $G_{k,n}$  to  $\mathbb{R}$ , is continuous. Thus, by (13),  $\alpha^*$  is itself a continuous, and generally smooth, function of  $\text{col}(X)$ , which implies, in particular, that the regression spaces that are sufficiently close (according to some distance on  $G_{k,n}$ ) to regression spaces yielding a large (resp. small)  $\alpha^*$  yield a large (resp. small)  $\alpha^*$ .

**Remark 8** We have not attempted to generalize Theorem 3.6 to asymmetric SAR models, for two reasons. Firstly, such models generally do not satisfy the condition  $f \notin \text{col}(X)$  necessary for the zero limiting problem. This is because the nonsymmetric weights matrices generally used in SAR models are row-stochastic, implying, as already noted above, that  $f \in \text{col}(X)$  as long as an intercept is included in the regression. Secondly, although the proof of Theorem 3.6 suggests that regression spaces (of low dimension) such that the limiting power of a POI or LBI c.r. vanishes for any  $\alpha$  always exist also in the context of asymmetric SAR models, the exact characterization of such regression spaces appears to be more involved. It should be noted, however, that an approximated characterization can be obtained from Theorem 3.6, by approximating an asymmetric SAR model by a CAR model with  $\Sigma^{-1}(\rho) = I - \rho(W + W')$  (i.e., omitting terms in  $\rho^2$ ).

### 3.3 Zero-Mean Models

In this subsection we specialize some of the above results to zero-mean (or constant-mean, by obvious extension) CAR and SAR models. For our purposes, setting  $X = 0$  in the models analyzed above has two main advantages. Firstly, it clarifies—by direct comparison with the regression case—the role played by the regressors in determining power. Secondly, it allows to focus on the effect of the specification of  $W$  on power.

Let us start from the following corollary of Theorem 3.3.

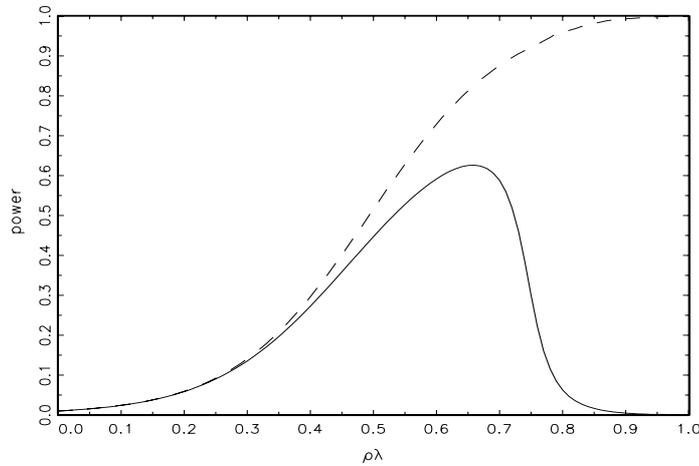
**Corollary 3.8** *In zero-mean CAR and SAR models, the limiting power of an invariant c.r.  $\Phi_y$  for testing  $\rho = 0$  against  $\rho > 0$  is 1 for any  $\alpha$  if  $f \in \Phi_y$ , 0 otherwise.*

It is instructive to relate Corollary 3.8 to the Moran statistic  $y'Wy/y'y$ . In the context of CAR and SAR models, the Moran statistic is usually interpreted as an autocorrelation coefficient. In view of this interpretation, the result in Corollary 3.8 is precisely what one would expect when

$W$  is symmetric, since in that case  $y = f$  maximizes the Moran statistic. The same cannot be said when  $W$  is nonsymmetric, because in that case the Moran statistic is not, in general, maximized by  $f$ .

In fact, the differences between models with symmetric  $W$  (CAR and symmetric SAR models) and models with nonsymmetric  $W$  (asymmetric SAR models) are not only a matter of interpretation. We provide an example in the context of possibly the simplest SAR model; the same model was used by Whittle (1954) in his seminal paper on spatial autoregressions.

**Example 1** A random variable is observed at  $n$  units placed along a line and, in the context of a zero-mean SAR process, it is to be tested whether  $\rho = 0$  or  $\rho > 0$ . Suppose that it is believed that there is only first-order interaction and that the interaction amongst first-order neighbors is stronger in one direction than in the other. Accordingly,  $W$  is chosen so that  $W(i, j)$  is equal to some fixed positive scalar  $w \neq 1$  if  $i - j = 1$ , to 1 if  $j - i = 1$ , and to 0 otherwise, for  $i, j = 1, \dots, n$ . In Figure 1, we plot the power function of the LBI test, i.e., the Moran test, and the envelope  $\pi_\rho(\rho)$  for  $n = 6$ ,  $w = 10$  and  $\alpha = 0.01$ . The power has been computed numerically, via the Imhof method (Imhof, 1961), and is plotted against  $\rho\lambda$ , which ranges between 0 and 1.



**Figure 1:** The power function of the Moran test (solid line) and the envelope  $\pi_\rho(\rho)$  (dashed line) for the zero-mean asymmetric SAR model described in Example 1.

Although it is based on a model with an artificial  $W$  (for more practically relevant models, see Section 3.4), Figure 1 illustrates the theoretically important point that in a SAR model with nonsymmetric  $W$ , the limiting power of the Moran test may vanish even when the model is not contaminated by regressors. On the contrary, when  $W$  is symmetric, the power function of the Moran test always goes to 1 as  $\rho \rightarrow \lambda^{-1}$  (by Lemma 3.4) and—as we shall see in Proposition 4.3

below—is monotonic. Note that this feature of the power function of the Moran test entails that there are zero-mean asymmetric SAR models in which the interpretation of the Moran statistic  $y'Wy/y'y$  as an autocorrelation coefficient cannot be justified, because for such models there exist values  $0 < k < \lambda^{-1}$  such that  $\Pr(y'Wy/y'y > k)$  is not increasing over  $0 < \rho < \lambda^{-1}$ . The next result gives further insights into the problem.

**Proposition 3.9** *In zero-mean SAR models, the limiting power of a POI or LBI c.r. for testing  $\rho = 0$  against  $\rho > 0$  is 1 for any  $\alpha$  if and only if  $f$  is an eigenvector of  $W'$ .*

The weights matrices  $W$  satisfying the condition in Proposition 3.9 are those such that  $\lambda$  is perfectly well-conditioned (e.g., Golub and Van Loan, 1996, p. 323). In practice, it turns out that the condition is very restrictive when  $W$  is nonsymmetric (whereas it is trivially satisfied when  $W$  is symmetric), and hence that in asymmetric SAR models typically  $\alpha^* > 0$  even when  $X = 0$ . For row-standardized  $W$ 's—the most popular, by far, nonsymmetric weights matrices in SAR models—the restrictiveness of the condition is emphasized by the following result.

**Corollary 3.10** *In zero-mean asymmetric SAR models with row-stochastic  $W$ , the limiting power of a POI or LBI c.r. for testing  $\rho = 0$  against  $\rho > 0$  is 1 for any  $\alpha$  if and only if  $W$  is doubly stochastic.*

Clearly, a nonsymmetric row-stochastic weights matrix  $W$  is doubly stochastic, i.e., has not only all rows but also all columns summing to 1, only in very special cases.<sup>6</sup> The condition in Proposition 3.9 remains very unlikely to be satisfied also for nonsymmetric  $W$ 's that are not row-stochastic. This is essentially because, given any choice of the neighborhood structure of a set of observational units (i.e., any choice of the pairs of units deemed to be neighbors) the choice of weights yielding a well-conditioned  $\lambda$  is typically a very particular one,<sup>7</sup> and corresponds to some relevant notion of distance amongst the units only in exceptional cases.

Having argued that the condition in Proposition 3.9 is generally not satisfied, the interesting issue becomes to understand which (nonsymmetric) matrices  $W$  are associated to large values of  $\alpha^*$ . Let us return to our example of a SAR model defined on a line.

---

<sup>6</sup>Formally, this can be deduced from Birkhoff's theorem on doubly stochastic matrices, which states that any such matrix must be a convex combination of permutation matrices; e.g., Horn and Johnson, 1985. We remark that the doubly stochastic weights matrices used in SAR models by Pace and LeSage (2002) are symmetric.

<sup>7</sup>That such a choice exists can be seen by starting from a (nonsymmetric) matrix  $W$ , and transforming it to  $S^{-1}WS$ , where  $S$  is a diagonal matrix with  $S(i, i) = (f_i/l_i)^{1/2}$ , with  $l$  denoting the left eigenvector of  $W$  associated to  $\lambda$ .

**Example 2** For the case of Example 1 above, the Imhof method (or some other numerical approximation to the null distribution of the Moran statistic) can be used to verify that  $\alpha^*$  is decreasing in  $n$  and increasing in  $|w - 1|$ . For the particular case of Figure 1,  $\alpha^*$  is about 0.056. Note that if one closes the line to form a circle (by setting  $W(1, n) = w$  and  $W(n, 1) = 1$ ), then  $W$  becomes a scalar multiple of a doubly stochastic matrix, and consequently  $\alpha^* = 0$  by Proposition 3.9.

Numerical investigations not reported here show that, typically, for a fixed  $n$ , large values of  $\alpha^*$  are associated to matrices  $W$  such that  $W(i, j)/W(j, i)$  is large for at least one pair  $(i, j)$  (we note that this type of asymmetry yields large values of  $\alpha^*$  even when  $X \neq 0$ ). This suggests that the asymmetry introduced by using row-standardized weights matrices  $W = D^{-1}A$  (see Section 2.1) does not yield very large values of  $\alpha^*$  in zero-mean SAR models, because for such matrices  $W(i, j)/W(j, i) \leq u(A)$ ,  $i, j = 1, \dots, n$ , where  $u(A)$  denotes the ratio of the largest to the smallest row-sum of  $A$ . Note that the largest possible value of  $u(A)$  over all  $n \times n$  matrices  $A$  is  $n - 1$ , obtained for the adjacency matrix of a star graph (i.e., a graph with one vertex having  $n - 1$  neighbors, and all other vertices having 1 neighbor). One can check that, even in this case, the value of  $\alpha^*$  associated to the corresponding row-standardized  $W$  is very small, and decreasing in  $n$ .<sup>8</sup> For instance, for the Moran test, when  $W$  is the row-standardized version of the adjacency matrix of a star graph,  $\alpha^* > 0.01$  only when  $n < 6$ . Thus, to summarize, in SAR models asymmetry of  $W$  may cause the limiting power of POI and LBI tests to disappear even when  $X = 0$ ; for row-standardized  $W$ 's, however, this typically occurs only for very small values of  $\alpha$  or  $n$ .

### 3.4 Numerical Examples

In this subsection we report numerical results aimed at illustrating how  $X$  and  $W$  affect the exact power of tests for residual spatial autocorrelation. More specifically, the objective is to show how sensitive power can be to  $X$  when  $\rho$  is large but not necessarily in a small neighborhood of  $\lambda^{-1}$ , in some situations of practical interest. For simplicity, we restrict attention to the power, which we denote by  $\pi_{LBI}(\rho)$ , of the Cliff-Ord test in the context of a SAR model. Related numerical investigations are contained in Krämer (2005).

For some selected specifications of  $W$ , we conduct Monte Carlo experiments where  $X$  is drawn

---

<sup>8</sup>Interestingly, the effect of the asymmetry of a row-standardized weight matrix  $D^{-1}A$  (or any other non-symmetric matrix that is similar to a symmetric matrix) can always be eliminated by suitably selecting  $V$  in (3). In fact, model (3) with  $W = D^{-1}A$  and  $V = D^{-1}$  is reduced, upon normalization to  $\Sigma(0) = I$ , to a SAR model with symmetric weight matrix  $D^{-1/2}AD^{-1/2}$ .

from some probability distribution, and the power is computed by the Imhof method. Because of its invariance property, the power of the Cliff-Ord test depends on  $X$  only through  $\text{col}(X)$ . A natural choice for the distribution of  $X$  would then be to take  $\text{vec}(X) \sim N(0, I_{nk})$ , because this would imply that  $\text{col}(X)$  is uniformly distributed on the Grassmann manifold  $G_{k,n}$  (see James, 1954, for the definition of uniform distribution on  $G_{k,n}$ ). Since, however, an intercept is in practice always included in the regression, we prefer to take  $X = [\iota \mid X_1]$ , with  $\text{vec}(X_1) \sim N(0, I_{n(k-1)})$  (the effect on power of including an intercept will be discussed below). In the results reported below,  $k = 2$ , i.e., the regression includes just an intercept and an i.i.d. standard normal variate. The simulation is based on  $10^6$  replications of  $X$ . All computations are done in GAUSS v7. We set  $\alpha = 0.05$ .

We construct weights matrices from the maps of the  $n = 17$  counties of Nevada and the  $n = 23$  counties of Wyoming. We consider both a binary  $W$ , specified according to the queen criterion (i.e.,  $W[i, j] = 1$  if counties  $i$  and  $j$  share a common boundary or a common point,  $W[i, j] = 0$  otherwise), and its row-standardized version. The average number of neighbors of a county is 4.35 in Nevada, 4.52 in Wyoming, whereas the sparseness of  $W$  (as measured by the percentage of zero entries) is 74.40 for Nevada and 80.34 for Wyoming. We shall see that, despite their similarities, these two spatial configurations are very different with respect to our testing purposes.

Firstly, in order to show how sensitive  $\pi_{LBI}(\rho)$  is to  $X$ , in Table 1 we display the percentage frequency distribution of  $\pi_{LBI}(\rho)$ , with  $W$  as described above. We report values for  $\rho = 0.9\lambda^{-1}$  and  $\rho = 0.95\lambda^{-1}$ , which represent points at which low power is particularly troublesome (because of the large inefficiency of the ordinary least squares estimator of  $\beta$ ), but that are not too close to  $\lambda^{-1}$ . Note that, by Theorem 3.3, in our experiment  $\lim \pi_{LBI}(\rho)$  (as  $\rho \rightarrow \lambda^{-1}$ ) is either 0 or 1 when  $W$  is binary (as in that case  $f \notin \text{col}(X)$  almost surely), whereas it is in  $(0, 1)$  when  $W$  is row-standardized (as in that case  $f = \iota \in \text{col}(X)$ ). It appears from Table 1 that in the case of Nevada  $\pi_{LBI}(\rho)$  depends to a very large extent on  $X$ , even at points that are relatively far from  $\lambda^{-1}$ . The dependence is less pronounced in the case of Wyoming.

Next, we consider the zero limiting power problem more closely, which requires restricting attention to binary weights matrices (so that  $f \notin \text{col}(X)$  almost surely). In Table 2 we display  $z_{0.05}$  (see Section 3.2), obtained as the frequency of times that (12) (with  $c_\alpha$  computed by the Imhof method) is positive in our experiment. Note that  $z_{0.05}$  is very large in the case of Nevada, whereas it is very small in the case of Wyoming. The table also displays the average shortcoming (i.e.,  $\pi_\rho(\rho) - \pi_{LBI}(\rho)$ ) of the Cliff-Ord test at  $\rho = 0.9\lambda^{-1}$  and  $\rho = 0.95\lambda^{-1}$ , when  $\lim \pi_{LBI}(\rho) = 0$  and when  $\lim \pi_{LBI}(\rho) = 1$ . It appears that the impact of the zero limiting power problem is

**Table 1:** Percentage frequency distribution of the power  $\pi_{LBI}(\rho)$  of the Cliff-Ord test, in model  $y = X\beta + \varepsilon$ , where  $\varepsilon$  is a SAR process and  $X$  contains an intercept and a standard normal variate. The power is computed by the Imhof method over  $10^6$  replications of  $X$ .

		$\pi_{LBI}(\rho)$							
		$\rho\lambda$	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1
Nevada									
binary W	0.90	0.11	0.25	28.42	71.05	0.17	.	.	
	0.95	0.29	5.75	36.29	53.43	4.11	0.13	.	
row-st W	0.90	.	.	0.02	0.16	41.47	58.35	.	
	0.95	.	.	0.01	0.05	1.56	98.38	.	
Wyoming									
binary W	0.90	.	.	.	0.02	0.69	99.29	.	
	0.95	.	.	.	0.02	0.10	1.76	98.12	
row-st W	0.90	.	.	.	.	.	0.50	99.50	
	0.95	.	.	.	.	.	.	100	

not localized only in a very small neighborhood of  $\lambda^{-1}$ , because, on average, an  $X$  causing  $\lim \pi_{LBI}(\rho) = 0$  causes shortcomings at  $\rho = 0.9\lambda^{-1}$  and  $\rho = 0.95\lambda^{-1}$  that are significantly larger than the corresponding shortcomings associated to an  $X$  such that  $\lim \pi_{LBI}(\rho) = 1$ .<sup>9</sup>

We remark that the probability  $z_\alpha$  is generally very sensitive to  $W$ ,  $n$ ,  $k$ , the choice of a test,  $\alpha$ , and the distribution of  $X$ . In most situations,  $z_\alpha$  is small (but positive under the condition in Corollary 3.7) when  $n - k$  is large (although it is possible to construct matrices  $W$ , e.g., the adjacency matrix of a star graph or a very dense matrix, such that this is not the case). This suggests that, from a practical point of view, the zero limiting power problem is mainly a small sample problem. In general, and interestingly,  $z_\alpha$  is significantly larger when the regression includes an intercept. This is because, due to the nonnegativity of  $W$ ,  $\iota$  usually (and especially if the row sums of  $W$  are all of similar magnitude) yields a large value of the Moran statistic, and therefore its presence tends to put more probability mass on the regression spaces close to the hostile ones defined by Theorem 3.6. When  $W$  is defined on a regular grid, one can study how  $z_\alpha$  depends on  $n$  explicitly (see Table 1 of Krämer, 2005). Note that  $z_\alpha$  is related to the measure  $\alpha^*$  by the relation  $z_\alpha = \Pr(\alpha^* > \alpha)$  (where the randomness of  $\alpha^*$  is due to that of  $X$ ). In our

<sup>9</sup>Note that when  $m_1 = 1$ , as in the examples we are considering, and  $\text{col}(X)$  contains the span of a vector  $f_1 + bf$  with large  $b$ , the power function goes to zero (by Theorem 3.6), but it does so very rapidly, because the condition  $f \in \text{col}(X)$  is nearly satisfied and therefore the power function tends to be close to that when  $f \in \text{col}(X)$ , which goes to a positive number as  $\rho \rightarrow \lambda^{-1}$ .

**Table 2:** Probability of zero limiting power ( $z_{0.05}$ ) and average shortcoming of the Cliff-Ord test, in the case of a binary  $W$ .

	av. shortc. at $\rho\lambda = 0.90$			av. shortc. at $\rho\lambda = 0.95$	
	$z_{0.05}$	$\pi_{LBI}(\rho) \rightarrow 0$	$\pi_{LBI}(\rho) \rightarrow 1$	$\pi_{LBI}(\rho) \rightarrow 0$	$\pi_{LBI}(\rho) \rightarrow 1$
Nevada	0.77	0.20	0.16	0.32	0.24
Wyoming	$5.2 \cdot 10^{-4}$	0.15	0.03	0.26	0.02

experiment,  $\alpha^*$  varied between  $2.8 \cdot 10^{-4}$  and 0.994 for the case of Nevada, between  $2.1 \cdot 10^{-7}$  and 0.430 for the case of Wyoming.

The main conclusion of our numerical study is that, in some cases of practical interest, the probability that the limiting power of the Cliff-Ord test vanishes may well be non-negligible. This obviously induces a large dependence of the power of the Cliff-Ord test on  $X$  when  $\rho \rightarrow \lambda^{-1}$ , but the numerical results indicate that both the power and the shortcoming may still depend to a large extent on  $X$  for values of  $\rho$  in a rather large neighborhood of  $\lambda^{-1}$ .

## 4 Unbiasedness and Monotonicity

In this section we discuss some conditions on model  $N(X\beta, \sigma^2\Sigma(\rho))$  that are sufficient for POI and LBI tests to be unbiased (for a general  $\Sigma(\rho)$ ) and to have power functions monotonic in  $\rho$  (for CAR or symmetric SAR models). The conditions are by no means necessary, but (i) are important to understand the structure of the testing problem under analysis; (ii) in the case of spatial autoregressive models, admit a simple interpretation.

We start from the following known, and fundamental, fact: any POI test is strictly unbiased for testing  $\rho = 0$  against the specific alternative  $\rho = \bar{\rho}$  for which it is constructed to be optimal. This property, for the general regression model (1), was derived in Theorem 1 of Kadiyala (1970) by an astute, but somewhat indirect, argument. For convenience, we restate the result in terms of the power envelope  $\pi_\rho(\rho)$ , and we point out (see the proof) that the result is a straightforward consequence of the Neyman-Pearson lemma.

**Proposition 4.1** *In model  $N(X\beta, \sigma^2\Sigma(\rho))$ , the inequality  $\pi_\rho(\rho) > \alpha$  holds for any  $\rho > 0$ .*

Proposition 4.1 is a very general result, since it holds for any  $X$  and any  $\Sigma(\rho)$ . However, it cannot be used to establish unbiasedness of a particular invariant test for  $\rho = 0$  against  $\rho > 0$ , except of course when a UMPI test exists (which is a very restrictive condition, because it requires

the c.r. defined by (7) to be independent of  $\bar{\rho}$ ). Next we formulate two conditions that, when taken together, lead to unbiasedness of POI and LBI tests.

By a commuting family of matrices it is meant a finite or infinite set of matrices that are pairwise commutative under standard multiplication.

**Condition A** *The matrices  $\Sigma(\rho)$ , for  $\rho > 0$ , form a commuting family.*

Condition A is particularly relevant in the present paper because it is satisfied by CAR and symmetric SAR models. Except for very special cases, it is not satisfied by asymmetric SAR models. A well-known characterization of a commuting family of symmetric matrices is that all its members share the same eigenvectors. This explains, in view of Proposition 3.5, why  $\alpha^* = 0$  in zero-mean CAR and symmetric SAR models, whereas generally  $\alpha^* > 0$  in zero-mean asymmetric SAR models. An important advantage of Condition A is that it allows a natural extension of many properties of the models  $N(0, \sigma^2 \Sigma(\rho))$  to the models  $N(X\beta, \sigma^2 \Sigma(\rho))$  that satisfy the next condition.

**Condition B** *For a fixed  $\tilde{\rho} > 0$ ,  $\text{col}(X)$  is spanned by  $k$  linearly independent eigenvectors of  $\Sigma(\tilde{\rho})$ .*

An interpretation of Condition B in CAR and SAR models will be given at the end of this section. Because of the characterization mentioned above, if Condition A holds, Condition B does not depend on  $\tilde{\rho}$ . Condition B, in any of its many equivalent formulations, has played a crucial role in the theoretical analysis of regression models with non-spherical errors since Anderson (1948). Although Condition B is unlikely to be met in practice, in some circumstances one may expect it to hold approximately (see the end of this section for CAR and symmetric SAR models, and Durbin, 1970, for the case of serial correlation). There is evidence in the literature that the power properties of tests for  $\rho = 0$  when Condition B holds exactly are similar to those when Condition B holds approximately (e.g., Tillman, 1975, p. 971).

Letting  $\text{col}^\perp(X)$  denote the orthogonal complement of  $\text{col}(X)$ , we have:

**Proposition 4.2** *Assume that Conditions A and B hold. Then, in model  $N(X\beta, \sigma^2 \Sigma(\rho))$ , any POI or LBI c.r. for testing  $\rho = 0$  against  $\rho > 0$  is unbiased. The unbiasedness is strict except when  $\text{col}^\perp(X)$  is a subset of an eigenspace of  $\Sigma(\rho)$ , in which case the power is  $\alpha$  for any  $\rho > 0$ .*

In Proposition 4.2, as in Propositions 4.3 and 4.4 below, “any” means for any  $\alpha$  and any  $\bar{\rho}$ . It is worth pointing out that, in general, Conditions A and B are not sufficient for the existence of a UMPI test for the stated testing problem, and therefore Proposition 4.2 is not a consequence of Proposition 4.1. An important counterexample in which a UMPI test exists is a

CAR model satisfying Condition B (the reason why Condition B combines particularly well with a CAR specification is that the resulting model is an exponential family with number of sufficient statistics equal to the number of parameters,  $k + 2$ ). It should also be noted that Conditions A and B are not sufficient for the monotonicity in  $\rho$  of the power functions of the tests in Proposition 4.2, not even when  $X = 0$ , because, given a  $\Sigma(\rho)$  satisfying Condition A, a reparametrization  $\rho \rightarrow f(\rho)$  may destroy the monotonicity of the power function without causing Condition A to fail. Note that while unbiasedness is a vital property of any c.r., monotonicity of the power function in  $\rho$  is a much stronger property and may or may not be desirable depending on the specification of  $\Sigma(\rho)$ . In general, it is desirable whenever  $\rho$  is interpreted as an autocorrelation parameter. This is the case for CAR and SAR models. We can prove:

**Proposition 4.3** *Assume that Condition B holds. Then, in CAR and symmetric SAR models the power function of any POI and LBI c.r. for testing  $\rho = 0$  against  $\rho > 0$  is non-decreasing. It is strictly increasing except when  $\text{col}^\perp(X)$  is a subset of an eigenspace of  $W$ , in which case the power is  $\alpha$  for any  $\rho > 0$ .*

Proposition 4.3 implies that in CAR and symmetric SAR models having zero mean or, more generally, satisfying Condition B, the LBI and POI test statistics can be regarded as indexes of (residual) autocorrelation, in that they are non-decreasing (as any correlation between pairs of variables in CAR and SAR models) in  $\rho$ , over  $(0, \lambda^{-1})$ . Another important consequence of Proposition 4.3 is the monotonicity of the envelope  $\pi_\rho(\rho)$ , for CAR and symmetric SAR models satisfying Condition B. One would expect the same property to hold for zero-mean asymmetric SAR models, but, so far, we have found neither a proof nor a counterexample (by numerical analysis).

**Remark 9** The power functions in Proposition 4.3 are, in fact, typically strictly increasing, because, unless  $n - k$  is small or an eigenspace of  $W$  has large dimension (see Example 3 below), the chances of  $\text{col}^\perp(X)$  falling into an eigenspace of  $W$  are very low. In the special case  $X = 0$ , the power functions must be strictly increasing, for  $\text{col}^\perp(X) = \mathbb{R}^n$  cannot be an eigenspace of  $W$ .

**Remark 10** For CAR models, Proposition 4.3 can alternatively be proved by showing that the density  $pdf(v; \rho)$  has a monotone likelihood ratio under Condition B, and then by using Theorem 3.4.1 of Lehmann and Romano (2005). Such an argument, however, does not extend to symmetric SAR models.

As it provides a link to the analysis in Section 3, the following result is also of interest.

**Proposition 4.4** *Assume that Condition B holds. Then, in CAR and symmetric SAR models the limiting power of any POI and LBI c.r. for testing  $\rho = 0$  against  $\rho > 0$  is 1 if  $f \notin \text{col}(X)$ ; strictly between  $\alpha$  and 1 if  $f \in \text{col}(X)$  and  $\text{col}^\perp(X)$  is not a subset of an eigenspace of  $W$ ;  $\alpha$  otherwise.*

We now provide a discussion of CAR and symmetric SAR models satisfying Condition B. From a practical perspective, the discussion is helpful to understand in which circumstances Condition B can be expected to hold approximately. We start from some examples. The most obvious case of a CAR model that satisfies Condition B is a model with mean assumed to be unknown but constant across observations and with a row-standardized  $W$  (see Section 2.1). On setting  $L = D^{-1}$  and normalizing to  $\Sigma(0) = I$ , the mean of the model becomes proportional to  $D^{\frac{1}{2}}\iota$ , where  $\iota$  is the  $n$ -dimensional vector of all ones, and the covariance matrix becomes  $\Sigma(\rho) = \sigma^2(I - \rho D^{-\frac{1}{2}}AD^{-\frac{1}{2}})^{-1}$ . Condition B is then satisfied because  $D^{\frac{1}{2}}\iota$  is an eigenvector of  $D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$  (since  $\iota$  is an eigenvector of  $D^{-1}A$ ) and hence of  $\Sigma(\rho)$ . When other regressors are included in the model, a case in which Condition B has some chances of being met in practice is when the number of eigenspaces of  $W$  (and hence of  $\Sigma(\rho)$ , for CAR and symmetric SAR models) is small relative to  $n$ . This typically occurs when  $W$  satisfies a large number of symmetries, in the sense of being invariant under a large group of permutations of its index set (e.g., Biggs, 1993). The extreme case of equicorrelation serves as an illustration.

**Example 3** In the context of CAR and SAR models, all regression errors are equicorrelated when  $W$  has constant off-diagonal entries (and zero diagonal entries). In that case,  $W$  is invariant with respect to the whole symmetric group on  $n$  elements and has only two eigenspaces, the one spanned by  $\iota$  and the hyperplane orthogonal to it. Thus, in the case of equicorrelation, in order for Condition B to be met it suffices that every regressor in the model satisfies a single linear constraint, namely, that its entries sum to zero. Interestingly, if  $X$  contains an intercept, then  $\text{col}^\perp(X)$  is a subset of an eigenspace of  $W$ , and thus the power function of a POI or LBI c.r. is flat by Proposition 4.3.

The cases discussed above, albeit theoretically important, are of limited practical relevance in non-experimental contexts, since other regressors are typically used along with an intercept and the matrices  $W$  are usually not highly regular. We therefore take a more general view. Call two units  $i$  and  $j$  neighbors if  $W(i, j) > 0$ . Consider, for simplicity, the case when there is only one regressor,  $x = (x_1, \dots, x_n)'$  say, and let  $\bar{x}_i$  be the weighted average  $\sum_{j \neq i} W(i, j)x_j$  of the values of  $x$  observed at units that are neighbors of  $i$  (the extension to  $k > 1$  is obvious). For CAR and symmetric SAR models, the eigenvectors of  $\Sigma(\rho)$  are the same as those of  $W$ . Hence, in such

models Condition B is met if and only if the ratio  $x_i/\bar{x}_i$  does not depend on  $i$ , because, by the definition of an eigenvector of  $W$ , such a ratio must, for any  $i$ , be equal to the corresponding eigenvalue. Now, the ratio  $x_i/\bar{x}_i$  may be regarded as a measure of “similarity” (as far as  $x$  is concerned) between  $i$  and its neighbors. This suggests that Condition B is approximately met (and hence the power of optimal invariant tests has desirable properties) when  $x$  is such that the degree of similarity between  $i$  and its neighbors does not change substantially with  $i$ .

## 5 Conclusion

The paper has investigated a number of properties of invariant/similar tests for autocorrelation in the context of a linear regression model with errors following a first-order conditional or simultaneous spatial autoregressive process. The main message of our analysis is that the power properties of exact tests for residual spatial autocorrelation may depend to a very large extent on the regressors, especially when the number of degrees of freedom is small and the autocorrelation is large. Intuitively, this is largely due to the fact that CAR and SAR models tend, as the autocorrelation increases, to a family of (improper) distributions on a 1-dimensional subspace of the sample space. If, in the context of a CAR or SAR model, the regressors are such that the intersection between such a subspace and a critical region has 1-dimensional Lebesgue measure zero, then the power of that critical region vanishes in the limit.

More formally, we have characterized the cases when the limiting power of invariant tests vanishes and we have shown that the minimum size  $\alpha^*$  such that the limiting power of a POI or LBI test does not vanish may, for some spatial structures, depend on  $\text{col}(X)$  to a very large extent. Furthermore, we have established that the sets of regression spaces  $\text{col}(X)$  causing a zero limiting power of a size- $\alpha$  POI or LBI test have non-zero (invariant) measure on the set  $G_{k,n}$  of all  $k$ -dimensional subspaces of  $\mathbb{R}^n$ , for any  $\alpha$ , any spatial structure and any  $k > m_1$ . In fact, in some circumstances, the probability content of these subsets (according to some distribution on  $G_{k,n}$ ) may be far from negligible.

A remark concerning the distributional assumptions underlying our results is in order. As is well known (e.g., Kariya, 1980), the density of the maximal invariant (6) remains the same for any elliptically symmetric distributions of  $y$ , so the assumption of Gaussianity is much more than what is required to study the properties of the test considered in this paper. It should be noted, however, that while the generalization of SAR models to non-Gaussian distributions is straightforward, that is not so for CAR models; see Besag (1974).

Two possible extensions of our work are as follows. Firstly, although in this paper we have

mostly focused on the power as  $\rho \rightarrow \lambda^{-1}$ , the techniques we have used should also prove useful to study local power, namely by studying the right derivative of the power function at  $\rho = 0$ . Secondly, an extension to *mixed regressive, spatial autoregressive* models (e.g., Ord, 1975, and Lee, 2002), which are not in the class of regression models (1) and therefore have not been considered in this paper, would be of interest.

## Appendix: Proofs

**Proof of Proposition 3.1** For some matrix  $X$ , denote by  $G_X$  the group of transformations  $y \rightarrow ay + Xb$ , with  $a \in \mathbb{R}^+$  and  $b \in \mathbb{R}^k$ , and by  $\Upsilon_X(\bar{\rho})$  the size- $\alpha$  POI c.r., defined on the sample space. By definition,  $\Upsilon_X(\bar{\rho})$  is the size- $\alpha$  c.r. that is invariant under  $G_X$  and has maximum probability content under  $N(X\beta, \sigma^2\Sigma(\bar{\rho}))$ . Observe that, for any  $X$ , the probability content  $\pi_{\bar{\rho}}(\bar{\rho}, X)$  of  $\Upsilon_X(\bar{\rho})$  under  $N(X\beta, \sigma^2\Sigma(\bar{\rho}))$  is the same as under  $N(0, \Sigma(\bar{\rho}))$ , by invariance under  $G_X$ . It immediately follows that, for any  $X \neq 0$ , any  $\bar{\rho} > 0$ , and any  $\alpha$ ,  $\pi_{\bar{\rho}}(\bar{\rho}, X) \leq \pi_{\bar{\rho}}(\bar{\rho}, 0)$ , because  $G_X$  is strictly larger than  $G_0$  (as all transformations in  $G_0$ , i.e.  $y \rightarrow ay$ , are in  $G_X$ , and there are transformations in  $G_X$ , i.e., those with  $b \neq 0$ , that are not in  $G_0$ ). Since, by the Neyman-Pearson Lemma applied to  $pdf(v; \bar{\rho})$ ,  $\Upsilon_0(\bar{\rho})$  is unique (up to a set of measure zero), a necessary and sufficient condition for  $\pi_{\bar{\rho}}(\bar{\rho}, X) = \pi_{\bar{\rho}}(\bar{\rho}, 0)$ ,  $X \neq 0$ , is that  $\Upsilon_X(\bar{\rho}) = \Upsilon_0(\bar{\rho})$ , i.e.,  $y'C'[(C\Sigma(\bar{\rho})C')^{-1} - c_\alpha I]Cy < 0$  if and only if  $y'[\Sigma^{-1}(\bar{\rho}) - c_\alpha I]y < 0$ . Since  $rank(C'RC) \leq n - k$  for any  $(n - k) \times (n - k)$  matrix  $R$ ,  $\Upsilon_X(\bar{\rho}) = \Upsilon_0(\bar{\rho})$ ,  $X \neq 0$ , requires  $rank(\Sigma^{-1}(\bar{\rho}) - c_\alpha I) \leq n - k$ , and hence  $c_\alpha = \lambda_i^{-1}(\Sigma(\bar{\rho}))$ ,  $i = 2, \dots, n - 1$ , which is equivalent to  $\alpha = \Pr(y'\Sigma^{-1}(\bar{\rho})y/y'y < \lambda_i^{-1}(\Sigma(\bar{\rho})))$ ,  $i = 2, \dots, n - 2$  (the cases  $i = 1, n$  are excluded because  $\alpha$  is assumed to be in  $(0, 1)$ ). It is easily seen that if  $c_\alpha = \lambda_i^{-1}(\Sigma(\bar{\rho}))$ ,  $i = 2, \dots, n - 1$ ,  $\Upsilon_0(\bar{\rho})$  is invariant under  $y \rightarrow ay + Xb$ , and hence is equal to  $\Upsilon_X(\bar{\rho})$ , if and only if  $\text{col}(X) \subseteq E_i(\Sigma(\rho))$ . This completes the proof of the proposition.

**Proof of Lemma 3.2** Let  $p = \lim_{\rho \rightarrow a} pdf(v; \rho) = |\Omega|^{-\frac{1}{2}} (v'\Omega^{-1}v)^{-\frac{n-k}{2}}$ . If all the eigenvalues of  $\Sigma(\rho)$  tend to a positive value as  $\rho \rightarrow a$ , then, by the Poincaré separation theorem (e.g., Rao, 1973, p. 64), all the eigenvalues of  $\Omega$  are positive. It follows that the term  $|\Omega|^{-\frac{1}{2}}$  of  $p$  is positive and finite if  $\lambda_{n-k}(\Omega) < \infty$ , and it vanishes otherwise. As for the term  $(v'\Omega^{-1}v)^{-\frac{n-k}{2}}$ , this is infinite if  $\lambda_{n-k}(\Omega) = \infty$  and  $v \in \tilde{E}_{n-k}(\Omega)$ , positive and finite in any other case. Combining the results, we have that if  $\lambda_{n-k}(\Omega) = \infty$ , then  $p = 0$  when  $v \notin \tilde{E}_{n-k}(\Omega)$ . Hence, by property (i) of  $pdf(v; \rho)$ , when  $\lambda_{n-k}(\Omega)$  is infinite and simple,  $p$  must be infinite when  $v \in \tilde{E}_{n-k}(\Omega)$ . Also, we have that  $0 < p < \infty$  if  $\lambda_{n-k}(\Omega) < \infty$ . The lemma now follows straightforwardly, on recalling that we are assuming that any invariant c.r. is centrally symmetric, so that it contains either both or neither of two antipodal points.

**Proof of Theorem 3.3** Nonnegativity and irreducibility of  $W$  imply that  $(I - \rho W)^{-1}$  is entrywise

positive, for any  $\rho > 0$  (see, e.g., Gantmacher 1974, p. 69, and recall that when we write  $\rho > 0$  we implicitly assume  $\rho < \lambda^{-1}$ ). It follows that, for both CAR and SAR models and for any  $\rho > 0$ ,  $\Sigma(\rho)$  is positive and hence, by Perron's theorem (e.g., Horn and Johnson, 1985, Theorem 8.2.11), that  $\lambda_n[\Sigma(\rho)]$  is simple. Also, observe that, for both CAR and SAR models, as  $\rho \rightarrow \lambda^{-1}$ ,  $\lambda_n[\Sigma(\rho)] \rightarrow \infty$  and all of the other eigenvalues of  $\Sigma(\rho)$  tend to a finite value, because, as it is easily verified,  $\text{rank}[(I - \lambda^{-1}W')(I - \lambda^{-1}W)] = n - 1$ . For CAR and symmetric SAR models and for any  $\rho > 0$ ,  $f_n[\Sigma(\rho)] = f$  and thus the spectral decomposition  $\Sigma(\rho) = \sum_{i=1}^n \{\lambda_i[\Sigma(\rho)]f_i[\Sigma(\rho)]f_i'[\Sigma(\rho)]\}$  shows that the matrix  $\lambda_n^{-1}[\Sigma(\rho)]\Omega_\rho$  tends to  $Cff'C'$  as  $\rho \rightarrow \lambda^{-1}$ . The same limit result holds also for asymmetric SAR models, since in that case  $f_n[\Sigma(\rho)] \rightarrow f$  as  $\rho \rightarrow \lambda^{-1}$  (because  $(I - \lambda^{-1}W')(I - \lambda^{-1}W)$  has an eigenvector  $f$  corresponding to its smallest eigenvalue 0). Now, since  $\text{rank}(Cff'C') \leq \text{rank}(ff') = 1$ , all eigenvalues of  $Cff'C'$  are zero except possibly one, which must then be equal to  $\bar{\lambda} = \text{tr}[Cff'C'] = f'Mf$ . If  $f \notin \text{col}(X)$ , then  $\bar{\lambda}$  is a simple positive eigenvalue of  $Cff'C'$  and has an associated eigenvector equal to  $Cf$ , for

$$Cff'C'Cf = Cff'Mf = \bar{\lambda}Cf.$$

It is easily seen that for any CAR or SAR model with  $f \notin \text{col}(X)$ ,  $Cf$  is also an eigenvector of  $\lim_{\rho \rightarrow \lambda^{-1}} \Omega_\rho$ , with (simple) eigenvalue equal to  $\lim_{\rho \rightarrow \lambda^{-1}} \{\lambda_n[\Sigma(\rho)]\bar{\lambda}\} = \infty$ . If  $f \in \text{col}(X)$ , then  $Cf = 0$  and thus  $\Omega_\rho = \sum_{i=1}^{n-1} \{\lambda_i^{-1}[\Sigma(\rho)]Cf_i f_i' C'\}$ , which tends to a matrix whose entries are all finite. Hence, when  $f \in \text{col}(X)$ ,  $\lim_{\rho \rightarrow \lambda^{-1}} \lambda_{n-k}(\Omega_\rho)$  must be finite. The theorem now follows by applying Lemma 3.2 with  $a = \lambda^{-1}$ .

**Proof of Lemma 3.4** From (13), we have that, provided that  $Cf \neq 0$ ,  $\alpha^* = 0$  if and only if  $Cf/\|Cf\| = \arg \min_{v \in S_{n-k}} \{v'Bv\}$ , and  $\alpha^* = 1$  if and only if  $Cf/\|Cf\| = \arg \max_{v \in S_{n-k}} \{v'Bv\}$ . The proposition follows by application of the Rayleigh-Ritz theorem (e.g., Horn and Johnson, 1985).

**Proof of Proposition 3.5** It can be deduced from the proof of Theorem 3.3 that, for a CAR or SAR model with  $f \notin \text{col}(X)$ ,  $E_{n-k}(\Omega_\rho)$  tends, as  $\rho \rightarrow \lambda^{-1}$ , to a 1-dimensional subspace containing  $Cf$ . It follows that if  $E_{n-k}(\Omega_\rho)$  does not depend on  $\rho$  for  $\rho > 0$ , it must be spanned by  $Cf$  for any  $\rho > 0$ , and hence, by Lemma 3.4 with  $B = \Omega_\rho^{-1}$ ,  $\alpha^* = 0$  for any POI test. Since this property holds for any  $\bar{\rho} > 0$ , it also holds for the LBI test.

**Proof of Theorem 3.6** We start from the case of the LBI test, which is notationally simpler than that of POI tests. By Lemma 3.4, for CAR and symmetric SAR models the limiting power of a LBI test vanishes for any  $\alpha$  (less than 1) if and only if  $f \notin \text{col}(X)$  and  $Cf \in E_1(CWC')$ . For a fixed  $W$ , consider the  $m_1$ -dimensional subspaces belonging to the span of  $f_1, \dots, f_{m_1}, f$ , and denote by  $\Theta$  the set of all such subspaces that do not contain  $f$  and are not  $E_1$ . It is easily shown that if  $\text{col}(X) \in \Theta$ ,  $CWC'$  admits the eigenpairs  $(\lambda_i, Cf_i)$ ,  $i = m_1 + 1, \dots, n - 1$ . But then, by the symmetry of  $CWC'$  and the fact that the vectors  $Cf_i$ ,  $i = m_1 + 1, \dots, n - 1$  are pairwise orthogonal (because the  $f_i$  are),  $CWC'$  must also admit an eigenvector in the subspace spanned by  $Cf_1, \dots, Cf_{m_1}, Cf$ . Since when  $\text{col}(X) \in \Theta$  such a subspace

is 1-dimensional, it follows that  $Cf$  is an eigenvector of  $CWC'$ , i.e.,

$$CWMf = \tilde{\lambda}Cf \tag{15}$$

for some eigenvalue  $\tilde{\lambda}$ . Thus, a  $\text{col}(X) \in \Theta$  causing the limiting power of the LBI test to disappear for any  $\alpha$  exists if and only if  $\tilde{\lambda} \leq \lambda_{m_1+1}$ . Observe that as  $\text{col}(X) \in \Theta$  approaches a subspace orthogonal to  $E_1$ ,  $Mf/\|Mf\|$  tends to a vector in  $E_1$ , which implies that  $\tilde{\lambda} \rightarrow \lambda_1$  (note that, by the definition of  $\Theta$ , no  $\text{col}(X) \in \Theta$  is orthogonal to  $E_1$ ). Thus, by the continuity of the eigenvalues of a matrix ( $CWC'$  here) in the entries of the matrix itself plus the fact that  $\lambda_1 < \lambda_{m_1+1}$ , a  $\text{col}(X) \in \Theta$  such that  $\tilde{\lambda} \leq \lambda_{m_1+1}$  always exists. The extension to POI tests, for any  $\bar{\rho} > 0$ , is straightforward and is obtained by replacing  $W$  with  $\Sigma(\bar{\rho})$  and  $\lambda_i$  by  $(1 - \rho\lambda_i)^{-r}$ ,  $i = 1, \dots, n$ , in the arguments used above, for any CAR ( $r = 1$ ) or symmetric SAR model ( $r = 2$ ).

The second part of the theorem considers explicitly the case  $m_1 = 1$ . Take  $X$  to be a vector proportional to  $f_1 + bf$ , so that  $\text{col}(X) \in \Theta$  as long as  $b \neq 0$ . For the LBI test, we just need to establish which values of  $b$  yield  $\tilde{\lambda} \leq \lambda_2$  (existence of an infinite number of such values of  $b$  follows from the first part of the theorem). Recalling that the vectors  $f_1$  and  $f$  are normalized, it is easily seen that  $Mf = (f - bf_1)/(1 + b^2)$ . Plugging such an expression in (15), and using the fact that  $Cf_1 = -bCf$  (since  $CX = 0$ ), we obtain  $\tilde{\lambda} = (\lambda + b^2\lambda_1)/(1 + b^2)$ . Hence,  $\tilde{\lambda} \leq \lambda_2$  requires  $|b| \geq [(\lambda - \lambda_2)/(\lambda_2 - \lambda_1)]^{1/2}$ , proving the part of the theorem relative to the LBI test when  $m_1 = 1$  (note that the non-uniqueness of  $f_1$  does not affect this result). By obvious extension, the limiting power of a POI test disappears for any  $\alpha$  if  $|b| \geq \{(\lambda_n[\Gamma(\bar{\rho})] - \lambda_2[\Gamma(\bar{\rho})])/(\lambda_2[\Gamma(\bar{\rho})] - \lambda_1[\Gamma(\bar{\rho})])\}^{1/2}$ . The proof of the theorem is then completed on substituting  $\lambda_i[\Gamma(\bar{\rho})] = (1 - \bar{\rho}\lambda_i)^{-p}$  in the last inequality, with  $p = 1$  for a CAR model and  $p = 2$  for a symmetric SAR model.

**Proof of Proposition 3.7** By Theorem 3.6, for POI or LBI tests in CAR or symmetric SAR models with any  $W$ , there exist  $\text{col}(X) \in G_{k,n}$  such that  $\alpha^* = 1$ . Let  $T(y)$  represents the test statistic associated to a POI or LBI test. Then, by equation (13), the subspaces  $\text{col}(X)$  yielding  $\alpha^* = 1$  are those that maximize  $T(f)$ , regarded as a function from  $G_{k,n}$  to  $\mathbb{R}$ . Next, observe that  $T(f)$  is continuous at its points of maximum, which implies that, for any  $\alpha$ , it is possible to find a neighborhood (defined according to some distance on  $G_{k,n}$ ) of the points of maximum such that any  $\text{col}(X)$  in this neighborhood causes the limiting power of size- $\alpha$  tests to disappear. This implies immediately that  $H_k(\alpha)$  has non-zero invariant measure on  $G_{k,n}$  (see James, 1954), for any  $0 < \alpha < 1$  and for  $k = m_1$ , and for any POI or LBI c.r. in any CAR or symmetric SAR model. Since the power of an invariant test does not depend on  $\beta$ , the proposition also holds for  $k > m_1$ .

**Proof of Corollary 3.8** The result follows immediately by taking  $C = I$  in Theorem 3.3.

**Proof of Proposition 3.9** Observe that if  $f$  is an eigenvector of  $W'$ , it must be associated to  $\lambda$ .

To see this, call  $\phi$  the eigenvalue of  $W'$  associated to  $f$ . Transposing both left and right hand sides of the equation  $W'f = \phi f$  and post-multiplying them by  $f$  yield  $f'Wf = \phi$ . But then  $\phi = \lambda$ , because it must also hold that  $f'Wf = \lambda$ . Let  $\Gamma(\rho) = [(I - \rho W')(I - \rho W)]^{-1}$ . By Lemma 3.4 with  $B = \Gamma^{-1}(\bar{\rho})$ , in order to prove the statement of the proposition regarding POI tests, we need to show that  $W'f = \lambda f$  is necessary and sufficient for  $f \in E_n(\Gamma(\bar{\rho}))$ , for any  $\bar{\rho} > 0$ . Clearly, if this holds for any  $\bar{\rho} > 0$ , it also holds for  $\bar{\rho} \rightarrow 0$ , establishing the part of the proposition regarding the LBI test. The necessity is straightforward, because if  $\Gamma(\bar{\rho})f = \lambda_n(\Gamma(\bar{\rho}))f$ , then  $\Gamma^{-1}(\bar{\rho})f = \lambda_n^{-1}(\Gamma(\bar{\rho}))f$ . From the latter equation we have  $(1 - \bar{\rho}\lambda)(I - \rho W')f = \lambda_n^{-1}(\Gamma(\bar{\rho}))f$ , which requires  $f$  to be an eigenvector of  $I - \rho W'$  and hence of  $W'$  (associated to  $\lambda$  by the above argument). As for the sufficiency, note that if  $W'f = \lambda f$ , then  $f$  is clearly an eigenvector of  $\Gamma(\bar{\rho})$ , for any  $\bar{\rho} > 0$ . By Perron's theorem (e.g., Horn and Johnson, 1985, Theorem 8.2.11), a vector in  $E_n(\Gamma(\bar{\rho}))$  is entrywise nonnegative (or nonpositive), for any  $\bar{\rho} > 0$ . But  $f$  is entrywise positive (by the Perron-Frobenius theorem applied to  $W$ ), and hence it must be in  $E_n(P_{\bar{\rho}}P'_{\bar{\rho}})$ , for any  $\bar{\rho} > 0$ , because if it were not, then, by the symmetry of  $P_{\bar{\rho}}P'_{\bar{\rho}}$ , it should be orthogonal to an entrywise nonnegative vector, which is impossible. This completes the proof of the proposition.

**Proof of Corollary 3.10** If  $W$  is a row-stochastic matrix, then  $f$  has identical entries, and therefore the condition in Proposition 3.9 is satisfied if and only if the columns of  $W$ , as its rows, sum to 1.

**Proof of Proposition 4.1** The assertion follows by applying Corollary 3.2.1 of Lehmann and Romano (2005) to the density (6), plus the assumption of identification of model (1).

**Proof of Proposition 4.2** For a POI test, we need to prove that  $\pi_{\bar{\rho}}(\rho) \geq \alpha$  for any positive  $\rho$  and  $\bar{\rho}$  and any size  $\alpha$ . If unbiasedness holds for any  $\bar{\rho} > 0$ , then it also holds for the LBI test. Letting

$$M_{\bar{\rho}} = I - X[X'\Sigma^{-1}(\bar{\rho})X]^{-1}X'\Sigma^{-1}(\bar{\rho}),$$

the matrix  $C'(C\Sigma(\bar{\rho})C')^{-1}C$  can be rewritten as  $\Sigma^{-1}(\bar{\rho})M_{\bar{\rho}}$  (e.g., Lemma 2 of King, 1980). Thus, for  $0 \leq \rho < \lambda^{-1}$ ,

$$\pi_{\bar{\rho}}(\rho) = \Pr\left(\frac{y'\Sigma^{-1}(\bar{\rho})M_{\bar{\rho}}y}{y'My} < c_{\alpha}; y \sim N(0, \Sigma(\rho))\right). \quad (16)$$

Under Conditions A and B,  $M_{\bar{\rho}} = M$  and, as is easily seen by exploiting the fact that Condition B is equivalent to the existence of an invertible matrix  $A$  such that  $\Sigma(\bar{\rho})X = XA$ , the matrices  $\Sigma^{-1}(\bar{\rho})$  and  $M$  commute for any  $\bar{\rho} > 0$ . Hence,

$$\pi_{\bar{\rho}}(\rho) = \Pr\left(\frac{z'\Sigma(\rho)\Sigma^{-1}(\bar{\rho})Mz}{z'\Sigma(\rho)Mz} < c_{\alpha}\right),$$

where  $z \sim N(0, I)$ . Moreover, under Conditions A and B, the matrix  $M$  has an eigenvalue 0 with eigenspace spanned by the  $k$  eigenvectors of  $\Sigma(\rho)$  that are in  $\text{col}(X)$ , and an eigenvalue 1 with eigenspace spanned by the remaining eigenvectors of  $\Sigma(\rho)$ . Let  $H$  be the set of indexes  $i$  of the  $n - k$  eigenvalues  $\lambda_i[\Sigma(\rho)]$  associated to a set of linearly independent eigenvectors of  $\Sigma(\rho)$  that are not in  $\text{col}(X)$ . Note that, when Condition A holds,  $H$  does not depend on  $\rho$ . Under Conditions A and B, the power of a POI

c.r. can then be expressed as

$$\pi_{\bar{\rho}}(\rho) = \Pr \left( \frac{\sum_{i \in H} \lambda_i[\Sigma(\rho)] \lambda_i^{-1}[\Sigma(\bar{\rho})] z_i^2}{\sum_{i \in H} \lambda_i[\Sigma(\rho)] z_i^2} < c_{\alpha} \right), \quad (17)$$

and its size as

$$\alpha = \Pr \left( \frac{z' \Sigma^{-1}(\bar{\rho}) M z}{z' M z} < c_{\alpha} \right) = \Pr \left( \frac{\sum_{i \in H} \lambda_i^{-1}[\Sigma(\bar{\rho})] z_i^2}{\sum_{i \in H} z_i^2} < c_{\alpha} \right). \quad (18)$$

Observe now that the sequences  $\lambda_i[\Sigma(\rho)]$ ,  $i \in H$ , and  $\lambda_i^{-1}[\Sigma(\bar{\rho})]$ ,  $i \in H$ , are oppositely ordered in the sense of Hardy *et al.*, 1952, p. 43. Then, the application of Tchebychef's inequality (Hardy *et al.*, 1952, Theorem 43) to the weighted arithmetic means (with weights  $z_i^2 / \sum_{i \in H} z_i^2$ ) of the  $\lambda_i[\Sigma(\rho)]$ ,  $i \in H$ , and of the  $\lambda_i^{-1}[\Sigma(\bar{\rho})]$ ,  $i \in H$ , yields that

$$\sum_{i \in H} \lambda_i[\Sigma(\rho)] z_i^2 \sum_{i \in H} \lambda_i^{-1}[\Sigma(\bar{\rho})] z_i^2 \geq \sum_{i \in H} z_i^2 \sum_{i \in H} \lambda_i[\Sigma(\rho)] \lambda_i^{-1}[\Sigma(\bar{\rho})] z_i^2,$$

for any vector  $z \in \mathbb{R}^n$ , with equality holding only if all the  $\lambda_i[\Sigma(\rho)]$  or all the  $\lambda_i^{-1}[\Sigma(\bar{\rho})]$ ,  $i \in H$ , are the same. Rearranging the terms of the above inequality, one finds that the statistic appearing in expression (17) is stochastically larger (e.g., Lehmann and Romano, 2005, p. 70) than that appearing in expression (18), and hence that  $\pi_{\bar{\rho}}(\rho) \geq \alpha$ , for any  $\bar{\rho} > 0$ , any  $\rho > 0$  and any size  $\alpha$ . If there are at least two indexes  $i, j \in H$  such that  $\lambda_i[\Sigma(\rho)] \neq \lambda_j[\Sigma(\rho)]$ , i.e., if  $\text{col}^{\perp}(X)$  is not a subset of an eigenspace of  $\Sigma(\rho)$ , then the last inequality is strict (as we are assuming  $\alpha \neq 0, 1$ ). The proof of the proposition is completed.

**Proof of Proposition 4.3** For a CAR or a symmetric SAR model,  $\lambda_i[\Sigma(\rho)] = (1 - \rho \lambda_i)^{-r}$ , for  $i = 1, \dots, n$ , and with  $r = 1$  for a CAR model,  $r = 2$  for a symmetric SAR model. Inserting such expressions in equation (17) from the proof of Proposition 4.2, we obtain that the power function of a POI c.r. is non-decreasing in  $\rho$  if the statistic

$$t_{\bar{\rho}}(\rho) = \left( \sum_{i \in H} \left( \frac{1}{1 - \rho \lambda_i} \right)^r z_i^2 \right)^{-1} \sum_{i \in H} \left( \frac{1 - \bar{\rho} \lambda_i}{1 - \rho \lambda_i} \right)^r z_i^2$$

is non-increasing in  $\rho$  for any vector  $z \in \mathbb{R}^n$ . Direct differentiation of  $t_{\bar{\rho}}(\rho)$  with respect to  $\rho$  and some simple manipulation show that such a condition is satisfied if

$$\sum_{i, j \in H} a_{i, j} z_i^2 z_j^2 \leq 0, \quad (19)$$

with the coefficients  $a_{i, j}$  defined by

$$a_{i, j} = \lambda_j \frac{(1 - \bar{\rho} \lambda_j)^r - (1 - \bar{\rho} \lambda_i)^r}{(1 - \rho \lambda_i)^r (1 - \rho \lambda_j)^{r+1}}.$$

It is immediately verified that, for each  $i, j \in H$  such that  $i \neq j$ ,  $a_{i, j} + a_{j, i} \leq 0$ , with strict inequality if  $\lambda_i \neq \lambda_j$ . Thus, given that  $a_{i, i} = 0$  for any  $i \in H$ , (19) holds, the inequality being strict if there exist at least one pair of distinct eigenvalues  $\lambda_i, \lambda_j$  with  $i, j \in H$ , i.e., if  $\text{col}^{\perp}(X)$  is not a subset of an eigenspace of  $W$ . The statement in the proposition relative to the POI c.r.s is therefore proved, and the one relative to LBI follows immediately.

**Proof of Proposition 4.4** Under Condition B, if  $f_i \notin \text{col}(X)$ , for  $i = 1, \dots, n$ , then  $f_i \in \text{col}^\perp(X)$ . It follows that, if  $f_i \notin \text{col}(X)$ , for  $i = 1, \dots, n$  and when  $\Sigma(\rho)$  is that of a CAR or symmetric SAR model,  $\Omega_\rho C f_i = C \Sigma(\rho) M f_i = C \Sigma(\rho) f_i = \lambda_i(\Sigma(\rho)) C f_i$ , i.e.,  $\{C f_i, f_i \notin \text{col}(X), j = 1, \dots, n\}$  is a set of  $n - k$  orthogonal eigenvectors of  $\Omega_\rho$ . Thus, in particular,  $E_{n-k}(\Omega_\rho)$  does not depend on  $\rho$ . The proposition now follows by Theorem 3.3, and Propositions 3.5 and 4.3.

## References

- Anderson, T.W. (1948) On the theory of testing serial correlation. *Skandinavisk Aktuarietidskrift* 31, 88-116.
- Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers.
- Bartels, R. (1992) On the power function of the Durbin-Watson test. *Journal of Econometrics* 51, 101-112.
- Besag, J.E. (1974) Spatial interaction and the statistical analysis of lattice data. *Journal of the Royal Statistical Society* B36, 192-236.
- Biggs, N.L. (1993) *Algebraic Graph Theory*, second edition. Cambridge University Press.
- Cvetković, D.M., M. Doob & H. Sachs (1980) *Spectra of Graphs*. Academic Press.
- Cliff, A.D. & J.K Ord (1981) *Spatial Autocorrelation*. London: Pion.
- Cox, D.R. & D.V. Hinkley (1974) *Theoretical Statistics*. Chapman & Hall.
- Cressie, N. (1993) *Statistics for Spatial Data*, revised edition. Wiley.
- Diggle, P.J. (1988) An approach to the analysis of repeated measurements. *Biometrics* 44, 959-971.
- Durbin, J. (1970) An alternative to the bounds test for testing for serial correlation in least-squares regression. *Econometrica* 38, 422-29.
- Florax, R.J.G.M. & T. de Graaff (2004) The performance of diagnostic tests for spatial dependence in linear regression models: a meta-analysis of simulation studies. In L. Anselin, R.J.G.M. Florax & S. J. Rey (eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications*, pp. 29-65. Springer.
- Gantmacher, F.R. (1974) *The Theory of Matrices*, vol. II. Chelsea.
- Golub, G.H. & C.F. Van Loan (1996) *Matrix Computations*, third edition. The Johns Hopkins University Press.

- Hardy, G., J.E. Littlewood & G. Pólya (1952) *Inequalities*, second edition. Cambridge University Press.
- Horn, R. & C.R. Johnson (1985) *Matrix Analysis*. Cambridge University Press.
- Hillier, G.H. (1987) Classes of similar regions and their power properties for some econometric testing problems. *Econometric Theory* 3, 1– 44.
- Imhof, J.P. (1961) Computing the distribution of quadratic forms in normal variables. *Biometrika* 48, 419-26.
- James, A.T. (1954) Normal multivariate analysis and the orthogonal group. *Annals of Mathematical Statistics* 25, 40-75.
- Kadiyala, K.R. (1970) Testing for the independence of regression disturbances. *Econometrica* 38, 97-117.
- Kariya, T. (1980) Locally robust tests for serial correlation in least squares regression. *Annals of Statistics* 8, 1065-1070.
- King, M.L. (1980) Robust tests for spherical symmetry and their application to least squares regression. *Annals of Statistics* 8, 1265-1271.
- King, M.L. (1981) A small sample property of the Cliff-Ord test for spatial autocorrelation. *Journal of the Royal Statistical Society B* 43, 263 - 4.
- King, M.L. (1988) Towards a theory of point optimal testing. *Econometric Reviews* 6, 169–255.
- King, M.L. & G.H. Hillier (1985) Locally best invariant tests of the error covariance matrix of the linear regression model. *Journal of the Royal Statistical Society B* 47, 98-102.
- Kleiber, C. & W. Krämer (2005) Finite-sample power of the Durbin-Watson test against fractionally integrated disturbances. *Econometrics Journal* 8, 406-417.
- Krämer, W. (1985) The power of the Durbin-Watson test for regressions without an intercept. *Journal of Econometrics* 28, 363-370.
- Krämer, W. (2005) Finite sample power of Cliff–Ord-type tests for spatial disturbance correlation in linear regression. *Journal of Statistical Planning and Inference* 128, 489-496.
- Lee, L.F. (2002) Consistency and efficiency of least squares estimation for mixed regressive, spatial autoregressive models. *Econometric Theory* 18, 252-277.

- Lehmann, E.L., and Romano, J. (2005) *Testing Statistical Hypotheses*, third edition. Springer.
- Moran, P.A.P. (1950) Notes on continuous stochastic phenomena. *Biometrika* 37, 17-23.
- Ord, J.K. (1975) Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* 70, 120-6.
- Pace, R.K. & J.P. LeSage (2002) Semiparametric maximum likelihood estimates of spatial dependence. *Geographical Analysis* 34, 76-90.
- Rahman, S. & M.L. King (1997) Marginal-likelihood score-based tests of regression disturbances in the presence of nuisance parameters. *Journal of Econometrics* 82, 81-106.
- Rao, C.R. (1973) *Linear Statistical Inference and Its Applications*, second edition. Wiley.
- Tiefelsdorf, M. & B. Boots (1995) The exact distribution of Moran's I. *Environment and Planning A* 27, 985-999.
- Tillman, J.A. (1975) The power of the Durbin-Watson test. *Econometrica* 43, 959-74.
- Tunncliffe Wilson, G. (1989) On the use of marginal likelihood in time series model estimation. *Journal of the Royal Statistical Society B* 51, 15-27.
- Whittle, P. (1954) On stationary processes in the plane. *Biometrika* 41, 434-449.
- Zeisel, H. (1989) On the power of the Durbin-Watson test under high autocorrelation. *Communications in Statistics: Theory and Methods* 18, 3907-3916.