Testing for spatial autocorrelation: the regressors that make the power disappear

Article

Accepted Version

Testing for Spatial Autocorrelation: the Regressors that Make the Power Disappear

Federico Martellosio
University of Reading
June 25, 2010

Abstract

We show that for any sample size, any size of the test, and any weights matrix outside a small class of exceptions, there exists a positive measure set of regression spaces such that the power of the Cliff-Ord test vanishes as the autocorrelation increases in a spatial error model. This result extends to the tests that define the Gaussian power envelope of all invariant tests for residual spatial autocorrelation. In most cases, the regression spaces such that the problem occurs depend on the size of the test, but there also exist regression spaces such that the power vanishes regardless of the size. A characterization of such particularly hostile regression spaces is provided.

Keywords: Cliff-Ord test; point optimal tests; power; spatial error model; spatial lag model; spatial unit root.
JEL Classification: C12, C21.

Running Head: The Regressors that Make the Power Disappear

Mailing address: School of Economics, University of Reading, URS Building, Whiteknights, Reading RG6 6AW, UK. Tel: +44 (0) 118 378 6033.

E-mail: f.martellosio@reading.ac.uk

1 Introduction

In recent years, applied economists have become increasingly aware of the consequences of incorrectly ignoring spatial autocorrelation in cross-sectional regression studies. One of these consequences, for example, is that the OLS estimator of the slope parameters may be inefficient or inconsistent, depending on how the spatial autocorrelation enters the regression model; see, e.g., Anselin (1988). Moreover, even in cases when the OLS estimator of the slope parameters is consistent and does not involve a serious loss of efficiency compared to (a feasible version of) the best linearly unbiased estimator, neglecting spatial autocorrelation may lead to poor assessment of the estimator precision; see, e.g., Cordy and Griffith (1993). To avoid faulty inferences, testing for spatial autocorrelation is now common practice in many economic applications; e.g., Case (1991), De Long and Summers (1991), Besley and Case (1995).

The power of tests for spatial autocorrelation depends, among other things, on the regressors included in the model. In this paper, we are concerned with the impact of regressors on the limiting power achieved by tests of residual spatial autocorrelation as the spatial autocorrelation increases. More specifically, by limiting power we mean the power as the autocorrelation parameter of a first-order simultaneous autoregressive (SAR(1)) model approaches the right boundary of the parameter space. The study of such a limiting power is important for several reasons. Firstly, there are many empirical applications in economics where unobservable factors lead to large spatial autocorrelation in a regression model; see, e.g., Militino et al. (2004), Parent and LeSage (2007).[1] Secondly, the properties of inferential procedures that neglect spatial autocorrelation can be very poor when the autocorrelation is large. Thirdly, a SAR(1) model with autocorrelation parameter approaching the right boundary of the parameter space has an intrinsic theoretical interest for econometricians, because of the analogy to a near unit root process in time series; see, e.g., Fingleton (1999) and Lee and Yu (2008).

The key contribution on the limiting power of tests for spatial autocorrelation is Krämer (2005). Krämer focuses on a Gaussian SAR(1) error process with symmetric weights matrix, and on test statistics that can be expressed as a ratio of quadratic forms in regression errors. The main message of Krämer (2005) is that, for some combinations of the regressor matrix and of the spatial weights matrix, the power of such tests may vanish as the autocorrelation increases. That is, there are circumstances in which it may be very difficult to detect spatial autocorrelation when the autocorrelation is in fact large. Martellosio (2010a) shows that Krämer's results can be extended to any invariant test of spatial autocorrelation, and also to other models, including a spatial lag model.

Both Krämer (2005) and Martellosio (2010a) contain simulation results suggesting

[1] Lee and Yu (2008) give several other references to the empirical economic literature. Regression models with strongly autocorrelated errors are also relevant in many applications outside economics; see, e.g., Basu and Reinsel (1994) and Jones et al. (2008).

that, at least in some cases, a vanishing limiting power of tests for spatial autocorrelation may have some empirical relevance. In the present paper, we complement those simulation results with a theoretical investigation of whether, for any fixed weights matrix, there always are regression spaces (i.e., column spaces of the regressor matrix) such that the limiting power vanishes. For simplicity, we restrict attention to a spatial error model, and to the following tests: the Cliff-Ord test, which is the most popular test for residual spatial autocorrelation, and point optimal invariant tests, which define the Gaussian power envelope of invariant tests. Our main result is that for any fixed sample size, any fixed size of the tests, and any fixed weights matrix outside a small class of exceptions, the vanishing of the power is an event with positive probability (according to a suitable measure), in the sense that there exists a positive measure set of regression spaces such that the limiting power disappears.

What is more, and somewhat surprisingly, there are also regression spaces such that the limiting power vanishes for all values of the size of a test. We provide a characterization of such regression spaces, which are particularly "hostile" from the point of view of testing for large spatial autocorrelation. The characterization is similar in nature to characterizations of the regressor matrix that minimizes the efficiency of the OLS estimator (see Watson, 1955), and proves to be relevant for interpretative purposes.

The rest of the paper is organized as follows. Section 2 introduces the set-up. Section 3 contains our main results. Section 4 presents the characterization of the particularly hostile regression spaces. Section 5 concludes and indicates possible extensions of our analysis. Proofs and auxiliary lemmata are collected in the appendices.

2 The Testing Problem

We consider a linear regression model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}, \tag{1}$$

where $\boldsymbol{X}$ is a fixed $n \times k$ matrix of rank $k < n$, $\boldsymbol{\beta}$ is a $k \times 1$ vector of unknown parameters, and the error vector $\boldsymbol{u}$ follows a SAR(1) process

$$\boldsymbol{u} = \rho \boldsymbol{W} \boldsymbol{u} + \boldsymbol{\varepsilon} \tag{2}$$

(see, e.g., Cliff and Ord, 1973). Here, ρ is a scalar unknown parameter; $\boldsymbol{W}$ is a fixed $n \times n$ matrix of weights chosen to reflect a priori information on the spatial relations among the n observational units; $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of "innovations" with

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \boldsymbol{0}, \;\; \mathrm{var}(\boldsymbol{\varepsilon}) = \sigma^2 \boldsymbol{V},$$

where σ^2 is an unknown positive scalar parameter and $\boldsymbol{V}$ is a fixed $n \times n$ symmetric and positive definite matrix.

The results in this paper require only minimal additional restrictions on the above model. As far as $\boldsymbol{u}$ is concerned, we assume that its density is positive everywhere

on $\mathbb{R}^n$, is larger at $\mathbf{0}$ than anywhere else, and is continuous in both $\boldsymbol{u}$ and ρ. As for $\boldsymbol{W}$, we assume, for simplicity, that it has at least one (real) positive eigenvalue, and that the largest of the positive eigenvalues of $\boldsymbol{W}$, to be denoted by $\lambda_{\max}$, has geometric multiplicity one. Such an assumption is virtually always satisfied in applications of spatial autoregressions to connected irregular lattices.[2]

In the context of model (1)-(2), the null hypothesis of no residual spatial autocorrelation is

$$\mathrm{H}_0 : \rho = 0.$$

The empirically most relevant alternative hypothesis is

$$\mathrm{H}_1 : 0 < \rho < \lambda_{\max}^{-1},$$

which represents positive spatial autocorrelation if, as it is usually the case, all the entries of $\boldsymbol{W}$ are nonnegative. From now on, we set $\boldsymbol{V} = \boldsymbol{I}_n$, the identity matrix of order n, because this does not involve any loss of generality when testing H_0 against H_1 (if $\boldsymbol{V} \neq \boldsymbol{I}_n$, just premultiply $\boldsymbol{y}$ by $\boldsymbol{V}^{-1/2}$).

In this paper we are concerned with the behavior of the power function of tests for spatial autocorrelation when ρ is far from 0 (and n is fixed). More specifically, we will consider the limiting power of tests for H_0 against H_1 as $\rho \to \lambda_{\max}^{-1}$. Observe that, when $\rho = \lambda_{\max}^{-1}$, $\boldsymbol{u}$ cannot be written as a function of $\boldsymbol{\varepsilon}$, and $\mathrm{var}(\boldsymbol{u})$ does not exist.[3]

One nice property of the above testing problem is that it is unchanged under the transformations $\boldsymbol{y} \to \gamma\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\delta}$, with $\gamma \in \mathbb{R}\backslash\{0\}$ and $\boldsymbol{\delta} \in \mathbb{R}^k$. Accordingly, it is natural to require that a test for that problem is *invariant*, that is, is based on a statistic that is invariant under the same transformations; see, e.g., Lehmann and Romano (2005). It is simple to show that any invariant test for our testing problem is free of nuisance parameters both under H_0 (that is, the tests are similar) and under H_1; see, e.g., King (1980).

Model (1)-(2) is often referred to as a *spatial error model*. An alternative model, which is popular in economics, is the so-called *spatial lag model* $\boldsymbol{y} = \rho\boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ (see, e.g., Anselin, 2002, for a comparison of the two models). In the latter model, contrary to what happens in the former, ρ affects also $\mathrm{E}(\boldsymbol{y})$. Because this changes the problem of testing $\rho = 0$ quite significantly, all formal results in this paper are confined to the spatial error model. Extensions to the spatial lag model are discussed in Section 5.

A few, mainly notational, remarks are in order. We denote the size of a test by α, and, to avoid trivial cases, we assume $0 < \alpha < 1$. Note that, because of the invariance

[2] In particular, by the Perron-Frobenius theorem (e.g., Horn and Johnson, 1985, Ch. 8), our assumption is certainly satisfied if $\boldsymbol{W}$ is entrywise nonnegative and irreducible (see Section 3.3). Extensions of our set-up to the cases when $\lambda_{\max}$ is not defined (i.e., $\boldsymbol{W}$ is nilpotent) or has geometric multiplicity larger than one (e.g., $\boldsymbol{W}$ is block diagonal) are straightfoward.

[3] More generally, in order for model (2) to be invertible (so that $\boldsymbol{u} = (\boldsymbol{I}_n - \rho\boldsymbol{W})^{-1}\boldsymbol{\varepsilon}$), ρ must be different from the reciprocal of the nonzero real eigenvalues of $\boldsymbol{W}$. All such non-admissible values of ρ are outside H_1.

with respect to the transformations $\boldsymbol{y} \rightarrow \boldsymbol{y} + \boldsymbol{X}\boldsymbol{\delta}$, the power of any invariant test depends on $\boldsymbol{X}$ only through its column space $\text{col}(\boldsymbol{X})$, often referred to as the regression space. All matrices considered in this paper are real. For a $q \times q$ symmetric matrix $\boldsymbol{Q}$, we denote by $\lambda_1(\boldsymbol{Q}) \leq \lambda_2(\boldsymbol{Q}) \leq ... \leq \lambda_n(\boldsymbol{Q})$ its eigenvalues; by $\boldsymbol{f}_1(\boldsymbol{Q}), ..., \boldsymbol{f}_q(\boldsymbol{Q})$ a set of corresponding orthonormal eigenvectors; by $E_i(\boldsymbol{Q})$ the eigenspace associated to $\lambda_i(\boldsymbol{Q})$; by $m_i(\boldsymbol{Q})$ the (algebraic and geometric) multiplicity of $\lambda_i(\boldsymbol{Q})$.

3 Main Results

This section investigates the existence of pairs $(\boldsymbol{W}, \boldsymbol{X})$ such that the limiting power of tests for residual spatial autocorrelation vanishes. We shall see that such pairs always exist, as long as $\boldsymbol{W}$ is outside a small class of exceptions. Most importantly, the zero limiting power is a positive probability event, in a sense to be made clear below. Section 3.1 is devoted to the Cliff-Ord test. Section 3.2 shows that the results concerning the Cliff-Ord test extend to point optimal tests with only a minor modification. Section 3.3 discusses the exceptions to our main results.

Before we proceed, it is important to point out that the analysis to follow is not directly relevant if $\boldsymbol{W}$ is row-standardized, or, more generally, has constant row-sums. Indeed, if $\boldsymbol{W}$ has constant row-sums, the limiting power of any invariant test cannot vanish as long as an intercept is included in the regression; see Section 3.2.2 of Martellosio (2010a). A discussion of the possible consequences of our results for the important case of a row-standardized $\boldsymbol{W}$ is deferred to Section 5. For now, we point out that in some applications it may be preferable not to row-standardized weights matrices; see, e.g., Bell and Bockstael (2000), p. 74, and Kelejian and Prucha (2010).

3.1 The Cliff-Ord Test

The most popular test for residual spatial autocorrelation is the Cliff-Ord test. It consists of rejecting H_0 for large values of

$$I := \frac{\boldsymbol{y}'\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{W}\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}}{\boldsymbol{y}'\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}}, \tag{3}$$

where $\boldsymbol{M}_{\boldsymbol{X}} := \boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$; see Cliff and Ord (1973) and Kelejian and Prucha (2001).[4] When the distribution of $\boldsymbol{u}$ is elliptically symmetric, the Cliff-Ord test is locally best invariant for our testing problem (see King, 1980 and 1981). Critical values for the test can be obtained from the exact null distribution of I, or from the asymptotic null distribution of a suitably normalized version of I.

The issue of the existence of pairs $(\boldsymbol{W}, \boldsymbol{X})$ such that the limiting power of the Cliff-Ord test vanishes is considered in Krämer (2005). Theorem 1 in Krämer (2005) states that, in a spatial error model, "given any matrix $\boldsymbol{W}$ of weights, and independently

[4]Note that $\boldsymbol{y}'\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y} = 0$ if and only if $\boldsymbol{y}$ belongs to the set $\{\boldsymbol{0}\} \cup \text{col}(\boldsymbol{X})$, which, since $k < n$, has zero measure. Hence, I is defined almost surely.

of sample size, there is always some regressor $\boldsymbol{X}$ such that for the Cliff–Ord test the limiting power disappears". The statement is formulated under the assumptions that the model is Gaussian, and that $\boldsymbol{W}$ is symmetric. Unfortunately, Krämer's proof contains an incorrect argument, which has the consequence that the pairs $(\boldsymbol{W}, \boldsymbol{X})$ constructed in that proof do not need to cause the limiting power to vanish.[5] We now aim to settle the issue and place it in a more general context.

Let $\boldsymbol{f}_{\max}$ be one of the two normalized (so that $\boldsymbol{f}'_{\max}\boldsymbol{f}_{\max} = 1$) eigenvectors of $\boldsymbol{W}$ associated to $\lambda_{\max}$.[6] We need the following definition.

Definition 3.1 *$\mathcal{C}$ is the class of weights matrices $\boldsymbol{W}$ such that $m_1(\boldsymbol{W} + \boldsymbol{W}') = n - 1$ and $\boldsymbol{f}_{\max}$ is an eigenvector of $\boldsymbol{W}'$.*

The class $\mathcal{C}$ contains the exceptions to Lemma 3.2 below. More details about the class $\mathcal{C}$ will be given in Section 3.3, where it will become clear that that the weights matrices used in applications are generally not in $\mathcal{C}$ (see in particular Proposition 3.8). At this stage we limit ourselves to point out that the only members of $\mathcal{C}$ that have some empirical relevance are possibly those with $(\boldsymbol{W})_{ij}$ equal to some constant positive scalar if $i \neq j$, to 0 if $i = j$. We refer to such matrices as *equal weights* matrices. Equal weights matrices have recently attracted considerable attention in the spatial econometric literature; see Kelejian and Prucha (2002), Baltagi (2006), Kelejian et al. (2006), Baltagi and Liu (2009 and 2010), Smith (2009), and Martellosio (2010b).

Lemma 3.2 *Consider testing $\rho = 0$ in the context of a spatial error model. For any weights matrix $\boldsymbol{W} \notin \mathcal{C}$, any number of regressors $k > 0$, and any size α, there exists at least one k-dimensional regression space* $\mathrm{col}(\boldsymbol{X})$ *such that the limiting power of the Cliff-Ord test vanishes.*

Lemma 3.2 establishes that the statement from Krämer (2005) reported earlier is correct if $\boldsymbol{W} \notin \mathcal{C}$, for any n, k and α, and generalizes it to nonsymmetric $\boldsymbol{W}$ and to non-Gaussian models. Both generalizations are achieved by application of Lemma A.1 in Appendix A, which does not require symmetry of $\boldsymbol{W}$ or Gaussianity. Note that the extension to nonsymmetric weights matrices is important, despite the fact the Cliff-Ord test statistic can be rewritten in terms of the symmetric matrix $(\boldsymbol{W} + \boldsymbol{W}')/2$, because assuming a symmetric $\boldsymbol{W}$ in a SAR(1) model does cause loss of generality.

Although it holds for any $\boldsymbol{W} \notin \mathcal{C}$, Lemma 3.2 has little practical relevance when $\boldsymbol{W}$ has constant row-sums. As we have mentioned above, in that case—and only in that case—the restriction that $\mathrm{col}(\boldsymbol{X})$ contains an intercept is sufficient to circumvent

[5] The problem lies in inequality (12) of Krämer (2005). In most cases, the critical value d_1 in that inequality can be positive or negative depending on α, and hence Krämer's proof holds only for sufficiently small α. In addition, there are weights matrices such that $d_1 < 0$ for any α; e.g., a $\boldsymbol{W}$ with zero diagonal entries and constant off-diagonal entries. For such matrices, inequality (12) is incorrect for all values of α.

[6] The two normalized eigenvectors are the same up to sign. Throughout the paper, it is irrelevant which of the two eigenvectors is chosen. Also, note that the normalization of $\boldsymbol{f}_{\max}$ is made here only for convenience, and will not be relevant until Section 4.

the zero limiting power problem. In other words, when $\boldsymbol{W}$ has constant row-sums, the regression spaces identified by Lemma 3.2 must not contain an intercept, and hence typically do not occur in applications.

Given any $\boldsymbol{W}$ with non-constant row-sums, Lemma 3.2 says that, for any n, k and α, there is at least one possibility that the Cliff-Ord test is unable to reject the null hypothesis when ρ is very far from the null hypothesis. This is a negative and unusual feature of a statistical test. It is therefore natural to wonder whether the set of regression spaces causing the limiting power to vanish has *zero measure*. In that case, Lemma 3.2, which only says that such a set is *nonempty*, would be immaterial for applications. We denote by $G_{k,n}$ the set—usually called a Grassmann manifold—of all k-dimensional subspaces of $\mathbb{R}^n$. We refer to the unique rotationally invariant measure on $G_{k,n}$; see Section 4.6 of James (1954) for details.[7]

Theorem 3.3 *Consider testing $\rho = 0$ in the context of a spatial error model. For any weights matrix $\boldsymbol{W} \notin \mathcal{C}$, any number of regressors $k > 0$, and any size α, the set of k-dimensional regression spaces such that the limiting power of the Cliff-Ord test vanishes has positive measure.*

Theorem 3.3 asserts that, as $\boldsymbol{X}$ is free to vary without restrictions (in the sense that $\mathrm{col}(\boldsymbol{X})$ has positive density almost everywhere on $G_{k,n}$), the zero limiting power has a positive probability of occurring. The main practical consequence of this result is that the zero limiting power is always a threat in applications, regardless of how large $n - k$ or α are (provided that $\boldsymbol{W} \notin \mathcal{C}$).

How likely it is in a given application to run into a regression space causing the limiting power to vanish will depend to a very large extent on $\boldsymbol{W}$, $n - k$, and α. Some simulation experiments analyzing this issue are reported in Krämer (2005) and Martellosio (2010a). Here, we stress that Theorem 3.3 implies that in any simulation study of the power properties of the Cliff-Ord test for a fixed $\boldsymbol{W} \notin \mathcal{C}$ and when $\boldsymbol{X}$ is drawn from a distribution supported on the whole $\mathbb{R}^{n\times k}$, there must be repetitions such that the limiting power vanishes, provided only that the number of repetitions is large enough.

3.2 The Point Optimal Invariant Tests

Martellosio (2010a) shows that the zero limiting power problem is not due to the form of a particular invariant test statistic. More specifically, he shows that if an invariant critical region does not contain $\boldsymbol{f}_{\max}$ in its closure, then its limiting power is bound to vanish (see Lemma A.1 in Appendix A). This geometric interpretation suggests that the results in the previous section can be extended to any other invariant test of

[7] Of course, $\boldsymbol{X}$ is assumed to be nonstochastic when constructing the Cliff-Ord test. We are now equipping $G_{k,n}$ with a probability measure only as a device to assess the practical relevance of the zero limiting power problem. One may think of an experiment where $\boldsymbol{W}$ is fixed, $\boldsymbol{X}$ is random, and the Cliff-Ord test is constructed for each realization of $\boldsymbol{X}$.

spatial autocorrelation. We shall here focus on the tests that, under the assumption of elliptical symmetry, define the *power envelope* of all invariant tests.

Consider testing $\rho = 0$ against the specific alternative that $\rho = \bar{\rho}$, for some fixed $0 < \bar{\rho} < \lambda_{\max}^{-1}$. When the distribution of $\boldsymbol{u}$ is elliptically symmetric, the Neyman-Pearson lemma implies that the most powerful invariant test rejects $\rho = 0$ for small values of

$$P_{\bar{\rho}} := \frac{\boldsymbol{y}'\boldsymbol{C}'\left[\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}'\right]^{-1}\boldsymbol{C}\boldsymbol{y}}{\boldsymbol{y}'\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}}, \tag{4}$$

where $\boldsymbol{\Sigma}(\bar{\rho}) := \text{var}(\boldsymbol{y}) = [(\boldsymbol{I}_n - \bar{\rho}\boldsymbol{W}')(\boldsymbol{I}_n - \bar{\rho}\boldsymbol{W})]^{-1}$, and $\boldsymbol{C}$ is an $(n-k) \times n$ matrix such that $\boldsymbol{C}\boldsymbol{C}' = \boldsymbol{I}_{n-k}$ and $\boldsymbol{C}'\boldsymbol{C} = \boldsymbol{M}_{\boldsymbol{X}}$ (see King, 1980 and 1988). In econometrics, the most powerful tests against a specific alternative are usually called point optimal invariant (POI) tests. With an abuse of language, we shall refer to a test based on (4) as a POI test, irrespective of whether the distribution of $\boldsymbol{u}$ is elliptically symmetric. Under elliptical symmetry, the POI tests define the power envelope of invariant tests. Of course, it could be argued that, if the distribution of $\boldsymbol{u}$ is not far from being elliptically symmetric, then the power function of a test based on (4) should be close to the power envelope.

In order to state the analog of Theorem 3.3 for POI tests, we need to define a slightly modified class of exceptions.[8]

Definition 3.4 *$\mathcal{C}^*$ is the class of weights matrices $\boldsymbol{W} \in \mathcal{C}$ such that $m_1(\boldsymbol{W}'\boldsymbol{W}) \geq n - 1$.*

Theorem 3.5 *Consider testing $\rho = 0$ in the context of a spatial error model. For any weights matrix $\boldsymbol{W} \notin \mathcal{C}^*$, any number of regressors $k > 0$, and any size α, there is a positive measure set of k-dimensional regression spaces such that the limiting power of a POI test vanishes.*

Theorem 3.5 is even more surprising than the corresponding result for the Cliff-Ord test. To see why this is the case, consider, under the assumption of elliptical symmetry, the extreme case of a POI test when $\bar{\rho}$ is close to $\lambda_{\max}^{-1}$ and the size α is very large. Since a very large α means that the critical region covers almost the whole sample space, one might expect the limiting power to be large. In fact, Theorem 3.5 asserts that, even in the extreme case when both $\bar{\rho}$ and α are large, regressors can be found such that the probability content of the critical region vanishes as $\rho \to \lambda_{\max}^{-1}$.

At this point, some remarks about how a zero limiting power is achieved (when $\boldsymbol{W} \notin \mathcal{C}$) are in order. By simulation one can see that often the power function is increasing at first, and then is collapsing to zero quite close to $\lambda_{\max}^{-1}$. Indeed, under elliptical symmetry, the power function of the Cliff-Ord test must have positive derivative at $\rho = 0$ (because it is locally best invariant), and the power of the POI test must be

[8] Note that $\mathcal{C}^*$ is strictly smaller than $\mathcal{C}$, because the condition $m_1(\boldsymbol{W}'\boldsymbol{W}) \geq n-1$ is not necessarily satisfied by a matrix $\boldsymbol{W}$ in $\mathcal{C}$.

larger than α at $\rho = \bar{\rho}$ (by the Neyman-Pearson Lemma). Note, in particular, that if both $\bar{\rho}$ and α are large, then the power function of the POI test must drop to zero very quickly after $\rho = \bar{\rho}$. Although this goes beyond the scope of the present paper, it would be particularly interesting to study what happens when the assumption of elliptical symmetry is not satisfied; in this case it should even be possible to find pairs $(\boldsymbol{W}, \boldsymbol{X})$ such that the power function of the Cliff-Ord test or of a test based on (4) is monotonically decreasing from α to zero.

3.3 Exceptions and Equal Weights Matrices

The reason why the weights matrices in $\mathcal{C}$ (resp. $\mathcal{C}^*$) constitute exceptions to our theorems above is that, in their presence, the limiting power of the Cliff-Ord (resp. a POI) test can never be zero. This is established in the following proposition.

Proposition 3.6 *Consider testing $\rho = 0$ in the context of a spatial error model. For any $\boldsymbol{W} \in \mathcal{C}$, any $\boldsymbol{X}$, and any α, the Cliff-Ord test statistic is degenerate (i.e., it does not depend on $\boldsymbol{y}$) if $\boldsymbol{f}_{\max} \in \text{col}(\boldsymbol{X})$, whereas it has limiting power equal to 1 if $\boldsymbol{f}_{\max} \notin \text{col}(\boldsymbol{X})$. The same result holds for a POI test, with $\mathcal{C}$ replaced by $\mathcal{C}^*$.*

The most important part of Proposition 3.6 is the one concerning the case $\boldsymbol{f}_{\max} \in \text{col}(\boldsymbol{X})$.[9] Indeed, we are now going to show that, under some conditions that are often satisfied in practice, the eigenvector $\boldsymbol{f}_{\max}$ of a matrix $\boldsymbol{W} \in \mathcal{C}$ is a vector of identical entries, and hence is in $\text{col}(\boldsymbol{X})$ as long as the regression contains an intercept. Consider the following two conditions.

Condition 1 $(\boldsymbol{W})_{ij} \geq 0$ *with* $(\boldsymbol{W})_{ii} = 0$, *for* $i, j = 1, ..., n$.

Condition 2 $\boldsymbol{W}$ *is irreducible.*

Condition 1 is virtually always satisfied in applications. For the definition of an irreducible matrix, see e.g. Horn and Johnson (1985). Irreducibility requires the graph with adjacency matrix $\boldsymbol{W}$ (that is, the graph with n vertices and an edge from vertex i to vertex j if and only $(\boldsymbol{W})_{ij} \neq 0$) to have a path from any vertex i to any vertex j. This condition is often met in applications. In particular, in the case of an undirected graph, irreducibility simply rules out the presence of isolated vertices. We have the following result.

Proposition 3.7 *Assume that Conditions 1 and 2 hold. Then, if $\boldsymbol{W} \in \mathcal{C}$, $\boldsymbol{f}_{\max}$ is a vector of identical entries.*

The following result is also of interest, because it gives an idea of how small the class $\mathcal{C}$ is.

[9] This part of the Proposition 3.6 can be seen as a generalization of Proposition 5 in Smith (2009), which asserts that when $\boldsymbol{W}$ is an equal weights matrix and the regression contains an intercept, the Cliff-Ord statistic is degenerate. For a discussion of the associated identification problem, see Arnold (1979), Kariya (1980), and Martellosio (2010b).

Proposition 3.8 *Assume that Conditions 1 and 2 hold. Then, if* $\boldsymbol{W} \in \mathcal{C}$, $\boldsymbol{W} + \boldsymbol{W}'$ *is an equal weights matrix.*

We are now in a position to also explain why, as mentioned in Section 3.1, the equal weights matrices are particularly important members of $\mathcal{C}$. Consider the following condition.

Condition 3 All eigenvalues of $\boldsymbol{W}$ are real.

Condition 3 is certainly satisfied when $\boldsymbol{W}$ is symmetric or a row-standardized version of a symmetric matrix, whereas it may not be satisfied in applications to directed networks.[10]

Proposition 3.9 *Assume that Conditions 1, 2 and 3 hold. Then,* $\boldsymbol{W} \in \mathcal{C}$ *if and only if* $\boldsymbol{W}$ *is an equal weights matrix.*

4 The Particularly Hostile Regression Spaces

Suppose that for a certain pair $(\boldsymbol{W}, \boldsymbol{X})$ the limiting power of an invariant critical region $\Phi \subset \mathbb{R}^n$ vanishes. By Lemma A.1, a zero limiting power occurs if $\boldsymbol{f}_{\max}$ is outside Φ. One obvious way to try and increase the power is to increase the size of Φ, because the limiting power becomes positive if Φ includes $\boldsymbol{f}_{\max}$. However, the minimum size α such that the limiting power does not vanish may be very large. What is more, there are cases when $\boldsymbol{f}_{\max}$ remains outside Φ for all values of α (less than 1), so that the limiting power vanishes irrespective of α. In this section, we characterize the regression spaces such that, for a fixed $\boldsymbol{W}$, the limiting power of the Cliff-Ord test or of a POI test vanishes regardless of α. Such regression spaces are referred to as particularly hostile.

We begin by formalizing the notion of a particularly hostile $\text{col}(\boldsymbol{X})$. For a fixed $\boldsymbol{W}$ and α, and for the Cliff-Ord test or a POI test, let $H_k(\alpha)$ denote the set of k-dimensional regression spaces that cause the limiting power to disappear. The set of particularly hostile k-dimensional regression spaces is the intersection of all sets $H_k(\alpha)$, for $\alpha \in (0, 1)$. Recall from Section 3 that any set $H_k(\alpha)$ is nonempty, provided that $\boldsymbol{W}$ is outside a class of exceptions ($\mathcal{C}$ for for the Cliff-Ord test, $\mathcal{C}^*$ for a POI test). Since $H_k(\alpha_2) \subseteq H_k(\alpha_1)$ for any $0 < \alpha_1 \leq \alpha_2 < 1$, it follows that the set of particularly hostile regression spaces is nonempty, for any k and any $\boldsymbol{W}$ not in $\mathcal{C}$ or $\mathcal{C}^*$.

The following theorem provides some information on the particularly hostile $\text{col}(\boldsymbol{X})$'s. This is achieved by making the simplifying assumption that $\boldsymbol{W}$ is symmetric, and by confining attention to the case when k is not greater than the multiplicity $m_1(\boldsymbol{W})$ of the smallest eigenvalue of $\boldsymbol{W}$. The most common value of $m_1(\boldsymbol{W})$ in applications to irregular spatial configurations is 1.[11] For the case $k = m_1(\boldsymbol{W}) = 1$, the theo-

[10] An example of a $\boldsymbol{W} \in \mathcal{C}^* \subset \mathcal{C}$ that does not satisfy Condition 3 is the 3×3 matrix with rows $(0, 1, 2)$, $(2, 0, 1)$, and $(1, 2, 0)$.

[11] In general, $m_1(\boldsymbol{W}) > 1$ requires $\boldsymbol{W}$ to satisfy some symmetries; see Biggs (1993), Ch. 15. An emblematic example is a block diagonal matrix $\boldsymbol{W}$ whose r blocks are all equal to an $n \times n$ equal weights matrix. For such a matrix, $m_1(\boldsymbol{W}) = n - r$.

rem provides a complete characterization of the particularly hostile col($\boldsymbol{X}$)'s. Some consequences of nonsymmetry of $\boldsymbol{W}$ will be discussed later, by means of an example.

Theorem 4.1 *Consider a spatial error model with symmetric weights matrix $\boldsymbol{W}$ such that $m_1(\boldsymbol{W}) = 1$, and with a single regressor ($k = 1$) that is a scalar multiple of the vector $\boldsymbol{f}_1(\boldsymbol{W}) + \omega \boldsymbol{f}_{\max}$, for some $\omega \in \mathbb{R}$. Let*

$$\omega_1 := \left[\frac{\lambda_{\max} - \lambda_2(\boldsymbol{W})}{\lambda_2(\boldsymbol{W}) - \lambda_1(\boldsymbol{W})}\right]^{\frac{1}{2}}; \ \omega_2 := \frac{1 - \bar{\rho}\lambda_1(\boldsymbol{W})}{1 - \bar{\rho}\lambda_{\max}} \left[\frac{2 - \bar{\rho}\left(\lambda_{\max} + \lambda_2(\boldsymbol{W})\right)}{2 - \bar{\rho}\left(\lambda_2(\boldsymbol{W}) + \lambda_1(\boldsymbol{W})\right)}\right]^{\frac{1}{2}}.$$

Then, the limiting power of the Cliff-Ord test vanishes for all values of α if $|\omega| \geq \omega_1$, and the limiting power of a POI test vanishes for all values of α if $|\omega| \geq \omega_1\omega_2$.

More generally, consider a spatial error model with a symmetric $\boldsymbol{W} \notin \mathcal{C}$, and with $k \leq m_1(\boldsymbol{W})$. For any k linearly independent eigenvectors $\boldsymbol{g}_1, ..., \boldsymbol{g}_k \in E_1(\boldsymbol{W})$, let Θ be the set of hyperplanes in $\text{span}(\boldsymbol{g}_1, ..., \boldsymbol{g}_k, \boldsymbol{f}_{\max})$ that do not contain $\boldsymbol{f}_{\max}$ and are not in $E_1(\boldsymbol{W})$. Then, for the Cliff-Ord test or a POI test, any set Θ contains regression spaces such that the limiting power vanishes for all values of α.

The result in Theorem 4.1 suggests the following interpretation. For a fixed symmetric $\boldsymbol{W}$, let

$$I_0 := \frac{\boldsymbol{v}'\boldsymbol{W}\boldsymbol{v}}{\boldsymbol{v}'\boldsymbol{v}},$$

where $\boldsymbol{v}$ is a realization of an n-dimensional random vector $\boldsymbol{z}$.[12] The statistic I_0 is the particular case of the Cliff-Ord statistic I when $\text{E}(\boldsymbol{z}) = \boldsymbol{0}$. Suppose that, as it is typically the case in applications, $\boldsymbol{W}$ has zero diagonal entries. Then, $\lambda_1(\boldsymbol{W}) < 0$, because $\lambda_{\max} > 0$ by definition and $\text{tr}(\boldsymbol{W}) = \sum_{i=1}^{n} \lambda_i(\boldsymbol{W}) = 0$. If we regard I_0 as a measure of autocorrelation, then, by Lemma A.2, any $\boldsymbol{v} \in E_1(\boldsymbol{W})$ represents a strongly negatively autocorrelated vector, whereas $\boldsymbol{v} = \boldsymbol{f}_{\max}$ represents a strongly positively autocorrelated vector.[13] For simplicity, let us now focus on the case $k = m_1(\boldsymbol{W}) = 1$, so that there is a single regressor, to be denoted by $\boldsymbol{x}$. Theorem 4.1 asserts that it is particularly difficult to detect large positive spatial autocorrelation when $\boldsymbol{x}$ can be written as the sum of a strongly positively autocorrelated component (according to I_0) and a strongly negatively autocorrelated component (according to I_0). One could say that the tests get confused in the presence of such an $\boldsymbol{x}$.

There are similarities between Theorem 4.1 and contributions in the time-series literature concerning the so-called Watson's $\boldsymbol{X}$ matrix (see, in particular, Watson, 1955). This is the regressor matrix that minimizes the efficiency of the OLS estimator of $\boldsymbol{\beta}$ relative to the best linear unbiased estimator. The similarities are not surprising,

[12] Here, for clarity and contrary to what is done in the rest of the paper, we use different notation for a random variable and its realizations.

[13] Of course, if $\text{E}(\boldsymbol{z}) \neq \boldsymbol{0}$, one should demean the data before constructing an autocorrelation index. That is, in practice, one would not use I_0, but $I_{\boldsymbol{\iota}} := \boldsymbol{v}'\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{W}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{v}/\boldsymbol{v}'\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{v}$, where $\boldsymbol{\iota}$ denotes the $n \times 1$ vector of all ones. In general, $\boldsymbol{v} \in E_1(\boldsymbol{W})$ (resp. $\boldsymbol{v} = \boldsymbol{f}_{\max}$) will be associated to a small (resp. large) value of $I_{\boldsymbol{\iota}}$.

in view of results in Tillman (1975) indicating that, in the presence of Watson's $\boldsymbol{X}$ matrix, the Durbin-Watson test has low power as the autocorrelation coefficient of an AR(1) process approaches 1.

Next, we provide a graphical representation of the hostile regression spaces. Such a representation is helpful to better understand the characterization in Theorem 4.1, and to appreciate what happens when $\boldsymbol{W}$ is nonsymmetric. We take $n = 3$ and $k = 1$, so that the regression spaces are lines in $\mathbb{R}^3$ through the origin. Without loss of generality, we normalize the regressors to have fixed length, so that they are points on a sphere in $\mathbb{R}^3$ (of arbitrary radius). We consider the weights matrix

$$\boldsymbol{W} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \tag{5}$$

Figure 1 displays 5000 random points from each of the three regions $H_1(0.9) \subset H_1(0.3) \subset H_1(0.1)$, for the Cliff-Ord test. We only plot the positive octant in the coordinate system of the eigenvectors of $\boldsymbol{W}$, because the symmetry of $\boldsymbol{W}$ implies that all regions $H_1(\alpha)$ are symmetric with respect to the coordinate planes in that coordinate system. It can be seen from Figure 1 that, as stated in Theorem 4.1, the particularly hostile regressors belong to the plane spanned by $\boldsymbol{f}_1(\boldsymbol{W})$ and $\boldsymbol{f}_{\max}$, and lie between the vector $\boldsymbol{h} := \boldsymbol{f}_1(\boldsymbol{W}) + \omega_1 \boldsymbol{f}_{\max}$ and $\boldsymbol{f}_{\max}$.

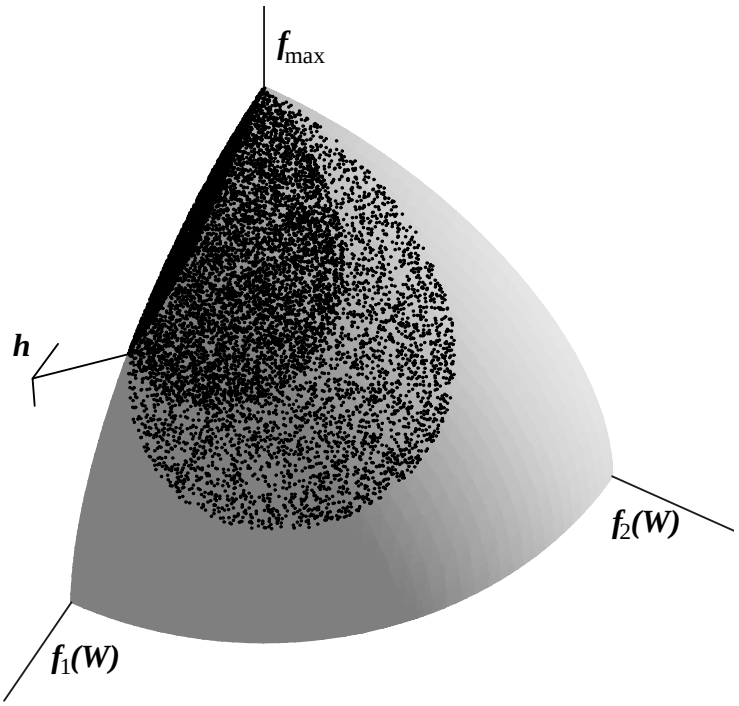

Figure 1: 5000 random points from each of the regions $H_1(0.9)$, $H_1(0.3)$ and $H_1(0.1)$, for the weights matrix in (5).

Let us now turn our attention to nonsymmetric weights matrices. Figure 2 is the

analog of Figure 1 for the weights matrix

$$\boldsymbol{Q} = \begin{bmatrix} 0 & 1 & 0 \\ 4 & 0 & 1 \\ 0 & 4 & 0 \end{bmatrix}. \tag{6}$$

The coordinate system for Figure 2 is the same as for Figure 1.[14] Since $\boldsymbol{Q}$ is nonsymmetric, the regions $H_1(\alpha)$ are no longer symmetric with respect to the coordinate planes (but are still symmetric with respect to the origin), so we now look at a whole hemisphere rather than at an octant as in Figure 1. We do not plot $H_k(0.1)$, as this region would cover almost the whole hemisphere (in this artificial illustration, it makes sense to consider large values of α, as we have only 2 degrees of freedom).

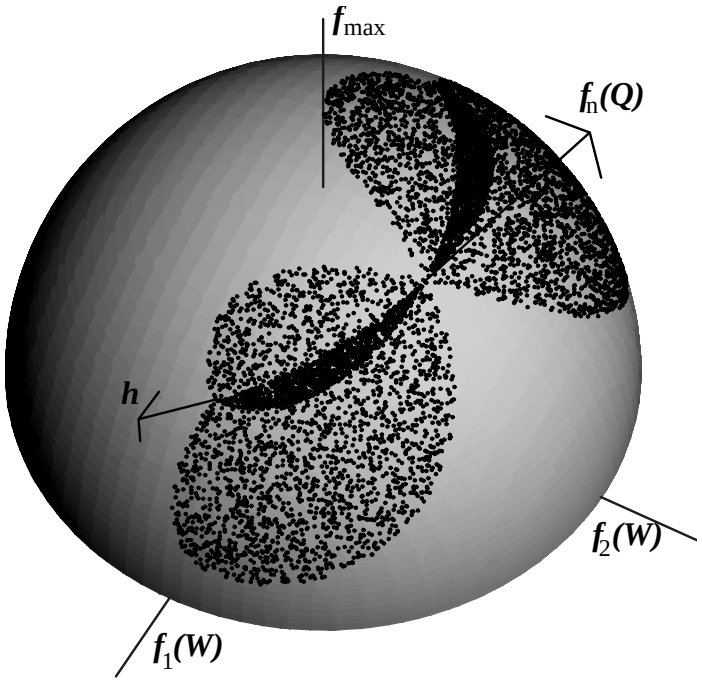

Figure 2: 10000 random points from each of the regions $H_1(0.9)$ and $H_1(0.3)$, for the weights matrix in (6).

The characterization in Theorem 4.1 requires symmetry of $\boldsymbol{W}$. Figure 2 suggests that a similar characterization should hold when $\boldsymbol{W}$ is nonsymmetric. Indeed, the particularly hostile regressors are still between $\boldsymbol{h}$ and the eigenvector $\boldsymbol{f}_n(\boldsymbol{Q})$ associated to the largest eigenvalue of $\boldsymbol{Q}$. However, contrary to the case of matrix (5), they do not lie on the plane spanned by $\boldsymbol{h}$ and $\boldsymbol{f}_n(\boldsymbol{Q})$. Because of this reason, a characterization similar to that in Theorem 4.1 is likely to be significantly more complicated in the case of nonsymmetric $\boldsymbol{W}$. To understand what happens when we move from a symmetric to a nonsymmetric weights matrix, it is useful to replace the two 4's in matrix (6) with

[14] In terms of $\boldsymbol{Q}$, the coordinate system is that of the eigenvectors of the matrix $\boldsymbol{Q} + \boldsymbol{Q}'$, which is a scalar multiple of (5).

a general scalar s. For any s, the particularly hostile regressors are between $\boldsymbol{h}$ and $\boldsymbol{f}_n(\boldsymbol{Q})$. When $s = 1$ (the case of Figure 1) they belong to the plane spanned by $\boldsymbol{h}$ and $\boldsymbol{f}_n(\boldsymbol{Q})$. As s moves away from 1, $\boldsymbol{f}_n(\boldsymbol{Q})$ moves away from $\boldsymbol{f}_{\max} = \boldsymbol{f}_n(\boldsymbol{Q} + \boldsymbol{Q}')$, and the curve described on the sphere by the particularly hostile regression spaces moves away from the plane spanned by $\boldsymbol{h}$ and $\boldsymbol{f}_n(\boldsymbol{Q})$.

5 Discussion

In Krämer (2005) and Martellosio (2010a) conditions are given for tests of spatial autocorrelation to have zero limiting power, where the limit is taken as the autocorrelation increases. The present paper has continued that work by addressing the question of whether it is always possible to run into regressors such that the limiting power vanishes. Provided that the weights matrix is not in a small class of exceptions, the answer is positive. As a consequence, in applications there is always a chance that detecting large autocorrelation by means of a certain test is extremely difficult.

Fortunately, since the regression spaces that cause a zero limiting power depend on the test, the problem can be tackled. Indeed, there always exists a test of spatial autocorrelation with nonvanishing limiting power (by Lemma A.1, any test such that the closure of its critical region does not contain $\boldsymbol{f}_{\max}$ has nonvanishing limiting power). Hence, a practical recommendation is to check whether the limiting power disappears using the condition in Lemma A.1. If the power disappears, one can try a different critical region and check the same condition again.

For simplicity, in this paper we have confined attention to the spatial error model. In applications, one is also often interested in testing for a spatially lagged dependent variable in a spatial lag model, or in testing for residual autocorrelation in a spatial autoregressive model with autoregressive disturbances (see, e.g., Kelejian and Prucha, 2001). Such testing problems have a more complicated structure than the one considered in this paper: first, they are not invariant under the group of transformations $\boldsymbol{y} \to \gamma\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\delta}$; second, the distributions of the test statistics for those problems generally depend on nuisance parameters. While these complications would certainly make an extension of the results concerning the spatial error model more involved analytically, there is no reason to believe that they would impede it.

At the beginning of Section 3, we have mentioned that the limiting power of an invariant test must be positive in the context of a spatial error model with a row-standardized $\boldsymbol{W}$, provided that the regression contains an intercept. It is worth remarking that, even if it is positive, the limiting power may be very small. An obvious extension of our analysis would be to establish whether, when $\boldsymbol{W}$ is row-standardized and in the presence of an intercept, there exist regression spaces such that the limiting power is smaller than some given positive number. It should also be noted that, in the context of a spatial lag model or a spatial autoregressive model with autoregressive disturbances, the power of a test for autocorrelation can vanish even when $\boldsymbol{W}$ is row-standardized and an intercept is included among the regressors (see Martellosio,

2010a).

Appendix A Auxiliary Lemmata

The first lemma, reported here for convenience, is Corollary 1 of Martellosio (2010a). In that paper, the result was derived under the assumption that $\boldsymbol{W}$ is nonnegative and irreducible, but it is clear that it also holds under the weaker assumption maintained in the present paper that $\lambda_{\max}$ has geometric multiplicity one. We denote by $\text{int}(S)$, $\text{bd}(S)$ and $\text{cl}(S)$, the interior, the boundary, and the closure of a set S, respectively.

Lemma A.1 (Martellosio, 2010a) *In a spatial error model, the limiting power of an invariant critical region Φ for testing $\rho = 0$ against $\rho > 0$ is: 1 if $\boldsymbol{f}_{\max} \in \text{int}(\Phi)$; in $(0, 1)$ if $\boldsymbol{f}_{\max} \in \text{bd}(\Phi)$; 0 if $\boldsymbol{f}_{\max} \notin \text{cl}(\Phi)$.*

The next lemma is proved, for instance, in Horn and Johnson (1985).

Lemma A.2 (Rayleigh-Ritz Theorem) *For a $q \times q$ symmetric matrix $\boldsymbol{Q}$,*

$$\lambda_1(\boldsymbol{Q})\boldsymbol{x}'\boldsymbol{x} \leq \boldsymbol{x}'\boldsymbol{Q}\boldsymbol{x} \leq \lambda_q(\boldsymbol{Q})\boldsymbol{x}'\boldsymbol{x},$$

for all $\boldsymbol{x} \in \mathbb{R}^q$. The equalities on the left and on the right are attained if and only if $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{Q}$ associated to, respectively, $\lambda_1(\boldsymbol{Q})$ and $\lambda_q(\boldsymbol{Q})$.

Lemma A.3 *For any $n \times n$ symmetric matrix $\boldsymbol{Q}$, and for any $k \geq 1$, $\lambda_{n-k}(\boldsymbol{C}\boldsymbol{Q}\boldsymbol{C}') = \lambda_1(\boldsymbol{Q})$ if and only if $\text{col}(\boldsymbol{X})$ contains all eigenvectors of $\boldsymbol{Q}$ associated to the eigenvalues different from $\lambda_1(\boldsymbol{Q})$.*

Proof. Consider the spectral decomposition $\boldsymbol{Q} = \sum_{i=1}^{s} \eta_i(\boldsymbol{Q})\boldsymbol{G}_i$, where $\eta_1(\boldsymbol{Q}) < \eta_2(\boldsymbol{Q}) < ... < \eta_s(\boldsymbol{Q})$ are the $s \leq n$ distinct eigenvalues of $\boldsymbol{Q}$, and $\boldsymbol{G}_1, ..., \boldsymbol{G}_s$ are the corresponding eigenprojectors. Since $\sum_{i=1}^{s} \boldsymbol{G}_i = \boldsymbol{I}_n$,

$$\boldsymbol{Q} = \eta_1(\boldsymbol{Q})\left(\boldsymbol{I}_n - \sum_{i=2}^{s} \boldsymbol{G}_i\right) + \sum_{i=2}^{s} \eta_i(\boldsymbol{Q})\boldsymbol{G}_i = \eta_1(\boldsymbol{Q})\boldsymbol{I}_n + \sum_{i=2}^{s} (\eta_i(\boldsymbol{Q}) - \eta_1(\boldsymbol{Q}))\boldsymbol{G}_i,$$

and hence

$$\boldsymbol{C}\boldsymbol{Q}\boldsymbol{C}' = \eta_1(\boldsymbol{Q})\boldsymbol{I}_{n-k} + \sum_{i=2}^{s} (\eta_i(\boldsymbol{Q}) - \eta_1(\boldsymbol{Q}))\boldsymbol{C}\boldsymbol{G}_i\boldsymbol{C}'. \tag{7}$$

Observe that if $\lambda_{n-k}(\boldsymbol{C}\boldsymbol{Q}\boldsymbol{C}') = \lambda_1(\boldsymbol{Q})$ then $\boldsymbol{C}\boldsymbol{Q}\boldsymbol{C}' = \lambda_1(\boldsymbol{Q})\boldsymbol{I}_{n-k}$, which in turn implies, by (7), that $\boldsymbol{C}\boldsymbol{G}_i\boldsymbol{C}' = \boldsymbol{O}_n$, for $i = 2, ..., s$. The necessity of the condition in the proposition is thus established. To prove the sufficiency, suppose that $\text{col}(\boldsymbol{X})$ contains all eigenvectors of $\boldsymbol{Q}$ associated to the eigenvalues different from $\lambda_1(\boldsymbol{Q})$. Then $\boldsymbol{M}_{\boldsymbol{X}}$ has an eigenspace spanned by k orthogonal eigenvectors of $\boldsymbol{Q}$ that are in $\text{col}(\boldsymbol{X})$, and an eigenspace spanned by $n - k$ orthogonal eigenvectors of $\boldsymbol{Q}$ that are not in $\text{col}(\boldsymbol{X})$. The

former eigenspace is associated to the eigenvalue 0 and the latter to the eigenvalue 1. Hence, $\boldsymbol{M_X Q}$ has the eigenvalue 0 with multiplicity k and the eigenvalue $\lambda_1(\boldsymbol{Q})$ with multiplicity $n-k$. But, since the nonzero eigenvalues of the product of two matrices are independent of the order of the factors (e.g., Theorem 1.3.20 in Horn and Johnson, 1985), the eigenvalues of $\boldsymbol{M_X Q}$ are the same as those of $\boldsymbol{CQC'}$, except for k zeros. Thus, we must have $\lambda_{n-k}(\boldsymbol{CQC'}) = \lambda_1(\boldsymbol{Q})$, and the proof is completed. ∎

Lemma A.4 *Any $\boldsymbol{W} \in \mathcal{C}$ is normal.*

Proof. Since $\boldsymbol{W}$ is real, we need to show that $\boldsymbol{WW'} = \boldsymbol{W'W}$, for any $\boldsymbol{W} \in \mathcal{C}$. Write $\boldsymbol{W} = \boldsymbol{A} + \boldsymbol{B}$, where $\boldsymbol{A} := (\boldsymbol{W} + \boldsymbol{W'})/2$ is symmetric and $\boldsymbol{B} := (\boldsymbol{W} - \boldsymbol{W'})/2$ is antisymmetric. For any $\boldsymbol{W} \in \mathcal{C}$, $\boldsymbol{A}$ has only two eigenvalues: one of them is $\lambda_{\max}$, associated to the eigenvector $\boldsymbol{f}_{\max}$. Letting θ be the other eigenvalue, and $\boldsymbol{G}$ the corresponding eigenprojector, we have the spectral decomposition

$$\boldsymbol{A} = \lambda_{\max}\boldsymbol{f}_{\max}\boldsymbol{f}'_{\max} + \theta\boldsymbol{G}.$$

Since $\boldsymbol{G} = \boldsymbol{I}_n - \boldsymbol{f}_{\max}\boldsymbol{f}'_{\max}$, $\boldsymbol{A} = \theta\boldsymbol{I}_n + (\lambda_{\max} - \theta)\boldsymbol{f}_{\max}\boldsymbol{f}'_{\max}$. Observe that, for any $\boldsymbol{W} \in \mathcal{C}$, $\boldsymbol{B}\boldsymbol{f}_{\max} = (\boldsymbol{W} - \boldsymbol{A})\boldsymbol{f}_{\max} = 0$. Then,

$$\begin{aligned}\boldsymbol{WW'} &= (\theta\boldsymbol{I}_n + (\lambda_{\max} - \theta)\boldsymbol{f}_{\max}\boldsymbol{f}'_{\max} + \boldsymbol{B})(\theta\boldsymbol{I}_n + (\lambda_{\max} - \theta)\boldsymbol{f}_{\max}\boldsymbol{f}'_{\max} - \boldsymbol{B}) \\ &= \theta^2\boldsymbol{I}_n + (2(\lambda_{\max} - \theta)\theta + (\lambda_{\max} - \theta)^2)\boldsymbol{f}_{\max}\boldsymbol{f}'_{\max} - \boldsymbol{B}^2 = \boldsymbol{W'W}.\end{aligned}$$

∎

Lemma A.5 *For any $\boldsymbol{W} \in \mathcal{C}$ and any $\boldsymbol{X}$ such that $\boldsymbol{f}_{\max} \notin \mathrm{col}(\boldsymbol{X})$, the Cliff-Ord test statistic evaluated at $\boldsymbol{y} = \boldsymbol{f}_{\max}$ is*

$$I(\boldsymbol{f}_{\max}) = \frac{1}{2}\lambda_{n-k}(\boldsymbol{C}(\boldsymbol{W} + \boldsymbol{W'})\boldsymbol{C'}).$$

Proof. For any $\boldsymbol{W} \in \mathcal{C}$, the matrix $\boldsymbol{A} := (\boldsymbol{W} + \boldsymbol{W'})/2$ admits the spectral decomposition $\lambda_1(\boldsymbol{A})\boldsymbol{G}_1 + \lambda_n(\boldsymbol{A})\boldsymbol{G}_n$ where $\boldsymbol{G}_1$ and $\boldsymbol{G}_n$ are the spectral projectors associated with $\lambda_1(\boldsymbol{A})$ and $\lambda_n(\boldsymbol{A})$, respectively. Since $\boldsymbol{f}_n(\boldsymbol{A}) = \boldsymbol{f}_{\max}$ for any $\boldsymbol{W} \in \mathcal{C}$, and the spectral projectors must sum to $\boldsymbol{I}_n$, we can write

$$\boldsymbol{A} = \lambda_1(\boldsymbol{A})\boldsymbol{I}_n + (\lambda_n(\boldsymbol{A}) - \lambda_1(\boldsymbol{A}))\boldsymbol{f}_{\max}\boldsymbol{f}'_{\max}. \tag{8}$$

Assume that $\boldsymbol{f}_{\max} \notin \mathrm{col}(\boldsymbol{X})$, and consider an arbitrary vector $\boldsymbol{v} \in \mathbb{R}^{n-k}$ that is orthogonal to $\boldsymbol{C}\boldsymbol{f}_{\max}$. From (8) we have

$$\boldsymbol{CAC'v} = \lambda_1(\boldsymbol{A})\boldsymbol{CC'v} + (\lambda_n(\boldsymbol{A}) - \lambda_1(\boldsymbol{A}))\boldsymbol{C}\boldsymbol{f}_{\max}\boldsymbol{f}'_{\max}\boldsymbol{C'v},$$

and hence, since $\boldsymbol{CC'} = \boldsymbol{I}_n$ and $\boldsymbol{f}'_{\max}\boldsymbol{C'v} = 0$,

$$\boldsymbol{CAC'v} = \lambda_1(\boldsymbol{A})\boldsymbol{v}.$$

Thus, $\boldsymbol{CAC}'$ has an $(n-1)$-dimensional eigenspace (the orthogonal complement of $\boldsymbol{Cf}_{\max}$). But, since $\boldsymbol{CAC}'$ is symmetric because $\boldsymbol{A}$ is, the other eigenspace of $\boldsymbol{CAC}'$ must be spanned by $\boldsymbol{Cf}_{\max}$. The eigenvalue of $\boldsymbol{CAC}'$ pertaining to such an eigenspace cannot be smaller than $\lambda_1(\boldsymbol{A})$ by the Poincaré separation Theorem (e.g., Horn and Johnson, 1985). It follows that

$$\boldsymbol{CAC}'\boldsymbol{Cf}_{\max} = \lambda_{n-k}(\boldsymbol{CAC}')\boldsymbol{Cf}_{\max}, \tag{9}$$

which in turn implies that $I(\boldsymbol{f}_{\max}) = \lambda_{n-k}(\boldsymbol{CAC}')$. ∎

Lemma A.6 *For any $\boldsymbol{W} \in \mathcal{C}^*$ and any $\boldsymbol{X}$ such that $\boldsymbol{f}_{\max} \notin \text{col}(\boldsymbol{X})$, the POI test statistic evaluated at $\boldsymbol{y} = \boldsymbol{f}_{\max}$ is*

$$P_{\bar{\rho}}(\boldsymbol{f}_{\max}) = \lambda_{n-k}^{-1}(\boldsymbol{C\Sigma}(\bar{\rho})\boldsymbol{C}').$$

Proof. For any $\boldsymbol{W} \in \mathcal{C}^*$, the matrices $\boldsymbol{W} + \boldsymbol{W}'$ and $\boldsymbol{W}'\boldsymbol{W}$ are simultaneously diagonalizable, because they are diagonalizable and, by Lemma A.4, they commute (e.g., Horn and Johnson, 1985, Theorem 1.3.12). Recall that $\boldsymbol{\Sigma}(\bar{\rho}) = (\boldsymbol{I}_n - \bar{\rho}(\boldsymbol{W}+\boldsymbol{W}') + \bar{\rho}^2\boldsymbol{W}'\boldsymbol{W})^{-1}$ and that, for any $\boldsymbol{W} \in \mathcal{C}^*$, $m_1(\boldsymbol{W}+\boldsymbol{W}') = n-1$ and $m_1(\boldsymbol{W}'\boldsymbol{W}) \geq n-1$. Also, observe that $m_1(\boldsymbol{W}'\boldsymbol{W}) \geq n-1$ if and only if $\boldsymbol{W}'\boldsymbol{W}$ is a scalar multiple of $\boldsymbol{I}_n$. Thus, for any $\boldsymbol{W} \in \mathcal{C}^*$, $m_1(\boldsymbol{\Sigma}(\bar{\rho})) = n-1$. We can now proceed similarly to the proof of Lemma A.5. More specifically, on replacing $\boldsymbol{A}$ in that proof with $\boldsymbol{\Sigma}^{-1}(\bar{\rho})$, expression (9) becomes

$$\boldsymbol{C\Sigma}(\bar{\rho})\boldsymbol{C}'\boldsymbol{Cf}_{\max} = \lambda_{n-k}(\boldsymbol{C\Sigma}(\bar{\rho})\boldsymbol{C}')\boldsymbol{Cf}_{\max},$$

or, equivalently,

$$(\boldsymbol{C\Sigma}(\bar{\rho})\boldsymbol{C}')^{-1}\boldsymbol{Cf}_{\max} = \lambda_1((\boldsymbol{C\Sigma}(\bar{\rho})\boldsymbol{C}')^{-1})\boldsymbol{Cf}_{\max}.$$

Using the last expression, we obtain

$$P_{\bar{\rho}}(\boldsymbol{f}_{\max}) = \frac{\boldsymbol{f}'_{\max}\boldsymbol{C}'\left[\boldsymbol{C\Sigma}(\bar{\rho})\boldsymbol{C}'\right]^{-1}\boldsymbol{Cf}_{\max}}{\boldsymbol{f}'_{\max}\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{f}_{\max}} = \lambda_1((\boldsymbol{C\Sigma}(\bar{\rho})\boldsymbol{C}')^{-1}) = \lambda_{n-k}^{-1}(\boldsymbol{C\Sigma}(\bar{\rho})\boldsymbol{C}'),$$

which is the desired conclusion. ∎

Appendix B Proofs

Proof of Lemma 3.2. Consider some arbitrary n, k, α, and $\boldsymbol{W} \notin \mathcal{C}$. Let $\Phi = \{\boldsymbol{y} \in \mathbb{R}^n : I > c\}$ be the critical region associated to the Cliff-Ord test. It is readily established that the closure of Φ is $\text{col}(\boldsymbol{X}) \cup \{\boldsymbol{y} \in \mathbb{R}^n : I \geq c\}$. Hence, by Lemma A.1, the limiting power of the Cliff-Ord test vanishes if $\boldsymbol{f}_{\max} \notin \text{col}(\boldsymbol{X})$ and $I(\boldsymbol{f}_{\max}) < c$, where $I(\boldsymbol{f}_{\max})$ denotes the Cliff-Ord test statistic evaluated at $\boldsymbol{y} = \boldsymbol{f}_{\max}$. Letting

$\boldsymbol{A} := (\boldsymbol{W} + \boldsymbol{W}')/2$, we obtain from Lemma A.2 that $I \geq \lambda_1(\boldsymbol{A})$, for all $\boldsymbol{y} \in \mathbb{R}^n$ and all $\boldsymbol{X} \in \mathbb{R}^{n\times k}$. Let $\boldsymbol{C}$ be an $(n-k)\times n$ matrix such that $\boldsymbol{C}\boldsymbol{C}' = \boldsymbol{I}_{n-k}$ and $\boldsymbol{C}'\boldsymbol{C} = \boldsymbol{M}_{\boldsymbol{X}}$. On writing $I = \boldsymbol{t}'\boldsymbol{C}\boldsymbol{W}\boldsymbol{C}'\boldsymbol{t}/\boldsymbol{t}'\boldsymbol{t}$, where $\boldsymbol{t} := \boldsymbol{C}\boldsymbol{y}$, Lemma A.2 also implies that $I \leq \lambda_{n-k}(\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}')$ for all $\boldsymbol{y} \in \mathbb{R}^n$. Thus, in order to prove the lemma, it suffices to show that there exists at least one $\text{col}(\boldsymbol{X})$ such that the following three properties are satisfied: (i) $\boldsymbol{f}_{\max} \notin \text{col}(\boldsymbol{X})$; (ii) $I(\boldsymbol{f}_{\max})$ is arbitrarily close to $\lambda_1(\boldsymbol{A})$; (iii) $\lambda_{n-k}(\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}')$ is not arbitrarily close to $\lambda_1(\boldsymbol{A})$. Note that, since we are assuming $\alpha < 1$, we do not need to prove the existence of a $\text{col}(\boldsymbol{X})$ such that $I(\boldsymbol{f}_{\max}) = \lambda_1(\boldsymbol{A})$. Also, observe that condition (iii) is necessary, because without it the limiting power could be 1 even if there exists a $\text{col}(\boldsymbol{X})$ that satisfies (ii).

Consider now, for some $\boldsymbol{g} \in E_1(\boldsymbol{A})$, a sequence $\{\boldsymbol{X}_l\}_{l=1}^{\infty}$ such that the vector $(\boldsymbol{M}_{\boldsymbol{X}_l}\boldsymbol{f}_{\max})^*$ converges to $\boldsymbol{g}^*$, in the sense that

$$\lim_{l\to\infty} \|(\boldsymbol{M}_{\boldsymbol{X}_l}\boldsymbol{f}_{\max})^* - \boldsymbol{g}^*\| = 0,$$

where $\|\cdot\|$ is an arbitrary norm on $\mathbb{R}^n$, and a '$*$' indicates that a vector $\boldsymbol{v} \in \mathbb{R}^n$ has been normalized with respect to $\|\cdot\|$, i.e., $\boldsymbol{v}^* := \boldsymbol{v}/\|\boldsymbol{v}\|$ (the arbitrariness of the norm follows from the fact that the convergence of a sequence of vectors in $\mathbb{R}^n$ is independent of the choice of the norm; see Corollary 5.4.6 of Horn and Johnson, 1985). By Lemma A.2

$$\lim_{l\to\infty} I(\boldsymbol{f}_{\max}) = \lim_{l\to\infty} \frac{\boldsymbol{f}'_{\max}\boldsymbol{M}_{\boldsymbol{X}_l}\boldsymbol{A}\boldsymbol{M}_{\boldsymbol{X}_l}\boldsymbol{f}_{\max}}{\boldsymbol{f}'_{\max}\boldsymbol{M}_{\boldsymbol{X}_l}\boldsymbol{f}_{\max}} = \lambda_1(\boldsymbol{A}). \tag{10}$$

Expression (10) implies the existence of at least one $\text{col}(\boldsymbol{X})$ that satisfies (i) and (ii). In order to establish that there exists at least one $\text{col}(\boldsymbol{X})$ that satisfies (i), (ii) and (iii), we need to show that it is possible to choose a $\boldsymbol{g} \in E_1(\boldsymbol{A})$ in such a way that

$$\lim_{l\to\infty} \lambda_{n-k}(\boldsymbol{C}_l\boldsymbol{A}\boldsymbol{C}_l') \neq \lambda_1(\boldsymbol{A}), \tag{11}$$

where $\boldsymbol{C}_l$ is an $(n-k)\times n$ matrix such that $\boldsymbol{C}_l\boldsymbol{C}_l' = \boldsymbol{I}_{n-k}$ and $\boldsymbol{C}_l'\boldsymbol{C}_l = \boldsymbol{M}_{\boldsymbol{X}_l}$. This is trivial if $m_1(\boldsymbol{A}) < n-k$ (because in that case (11) is satisfied for any $\boldsymbol{g} \in E_1(\boldsymbol{A})$, by the Poincaré separation Theorem; e.g., Horn and Johnson, 1985), but not more generally. Observe, however, that, as long as $\boldsymbol{W} \notin \mathcal{C}$, it is always possible to find a sequence $\{\boldsymbol{X}_l\}_{l=1}^{\infty}$ such that the following two properties are satisfied: (a) $\lim_{l\to\infty} \text{col}(\boldsymbol{X}_l)$ does not contain all eigenvectors of $\boldsymbol{A}$ associated to the eigenvalues other than $\lambda_1(\boldsymbol{A})$; (b) $\lim_{l\to\infty} \|(\boldsymbol{M}_{\boldsymbol{X}_l}\boldsymbol{f}_{\max})^* - \boldsymbol{g}^*\| = 0$ for some $\boldsymbol{g} \in E_1(\boldsymbol{A})$. The existence of a $\boldsymbol{g} \in E_1(\boldsymbol{A})$ such that (11) holds then follows from Lemma A.3, and this completes the proof. ∎

Proof of Theorem 3.3. Consider some arbitrary n, k, α, and $\boldsymbol{W} \notin \mathcal{C}$. In the proof of Lemma 3.2 it is shown that a $\text{col}(\boldsymbol{X})$ that minimizes $I(\boldsymbol{f}_{\max})$, regarded as a function from $G_{k,n}$ to $\mathbb{R}$, always exists and causes a zero limiting power of the Cliff-Ord test vanishes. Since $I(\boldsymbol{f}_{\max})$ is continuous at its points of minimum, it follows that (for any $\alpha < 1$), it is possible to find a neighborhood, defined according to some arbitrary distance on $G_{k,n}$, of the points of minimum such that any $\text{col}(\boldsymbol{X})$ in this

neighborhood causes the limiting power of the size-α Cliff-Ord test to disappear. Any such neighborhood has nonzero invariant measure on $G_{k,n}$ (see James, 1954), and hence the proof is complete. ∎

Proof of Theorem 3.5. The proof is similar to the proofs of Lemma 3.2 and Theorem 3.3. First, we need to establish some bounds on $P_{\bar{\rho}}$. By Lemma A.2, for all $\text{col}(\boldsymbol{X}) \in G_{k,n}$ and all $\boldsymbol{y} \in \mathbb{R}^n$,

$$P_{\bar{\rho}} \leq \lambda_{n-k}(\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}')^{-1}.$$

Noting that $\lambda_{n-k}(\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}')^{-1} = \lambda_1^{-1}(\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}')$ and that, by Poincaré separation Theorem (e.g., Horn and Johnson, 1985), $\lambda_1(\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}') \geq \lambda_1(\boldsymbol{\Sigma})$, we have

$$P_{\bar{\rho}} \leq \lambda_1^{-1}(\boldsymbol{\Sigma}(\bar{\rho})).$$

On writing $I = \boldsymbol{t}'(\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}')^{-1}\boldsymbol{t}/\boldsymbol{t}'\boldsymbol{t}$, where $\boldsymbol{t} := \boldsymbol{C}\boldsymbol{y}$, Lemma A.2 also implies that $P_{\bar{\rho}} \geq \lambda_{n-k}^{-1}(\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}')$, for all $\boldsymbol{y} \in \mathbb{R}^n$. Given the above bounds on $P_{\bar{\rho}}$, and following the same reasoning as in the proof of Lemma 3.2, we need to show that there exists at least one $\text{col}(\boldsymbol{X})$ such that: (i) $\boldsymbol{f}_{\max} \notin \text{col}(\boldsymbol{X})$; (ii) the POI statistic evaluated at $\boldsymbol{y} = \boldsymbol{f}_{\max}$, to be denoted by $P_{\bar{\rho}}(\boldsymbol{f}_{\max})$, is arbitrarily close to $\lambda_1^{-1}(\boldsymbol{\Sigma}(\bar{\rho}))$; (iii) $\lambda_{n-k}^{-1}(\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}')$ is not arbitrarily close to $\lambda_1^{-1}(\boldsymbol{\Sigma}(\bar{\rho}))$.

Consider a sequence $\{\boldsymbol{X}_l\}_{l=1}^{\infty}$ defined as in the proof of Lemma 3.2, but with $\boldsymbol{A}$ replaced by $\boldsymbol{\Sigma}(\bar{\rho})$. We have

$$\lim_{l\to\infty} \boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{M}_{\boldsymbol{X}_l}\boldsymbol{f}_{\max} = \lambda_1(\boldsymbol{\Sigma}(\bar{\rho}))\boldsymbol{M}_{\boldsymbol{X}_l}\boldsymbol{f}_{\max}. \tag{12}$$

Premultiplying both sides of (12) by $\boldsymbol{C}_l$, we obtain

$$\lim_{l\to\infty} (\boldsymbol{C}_l\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}_l')\boldsymbol{C}_l\boldsymbol{f}_{\max} = \lambda_1(\boldsymbol{\Sigma}(\bar{\rho}))\boldsymbol{C}_l\boldsymbol{f}_{\max},$$

and hence

$$\lim_{l\to\infty} (\boldsymbol{C}_l\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}_l')^{-1}\boldsymbol{C}_l\boldsymbol{f}_{\max} = \lambda_1^{-1}(\boldsymbol{\Sigma}(\bar{\rho}))\boldsymbol{C}_l\boldsymbol{f}_{\max}.$$

It follows that, by Lemma A.2,

$$\lim_{l\to\infty} P_{\bar{\rho}}(\boldsymbol{f}_{\max}) = \lim_{l\to\infty} \frac{\boldsymbol{f}'_{\max}\boldsymbol{C}_l'(\boldsymbol{C}_l\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}_l')^{-1}\boldsymbol{C}_l\boldsymbol{f}_{\max}}{\boldsymbol{f}'_{\max}\boldsymbol{M}_{\boldsymbol{X}_l}\boldsymbol{f}_{\max}} = \lambda_1^{-1}(\boldsymbol{\Sigma}(\bar{\rho})). \tag{13}$$

Expression (13) implies the existence of at least one $\text{col}(\boldsymbol{X})$ that satisfies (i) and (ii). In order to establish that there exists at least one $\text{col}(\boldsymbol{X})$ that satisfies (i), (ii) and (iii), we need to show that it is possible to choose a $\boldsymbol{g} \in E_1(\boldsymbol{\Sigma}(\bar{\rho}))$ in such a way that

$$\lim_{l\to\infty} \lambda_{n-k}(\boldsymbol{C}_l\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}_l') \neq \lambda_1(\boldsymbol{\Sigma}(\bar{\rho})). \tag{14}$$

But, as long as $\boldsymbol{W} \notin \mathcal{C}^*$, it is possible to find a sequence $\{\boldsymbol{X}_l\}_{l=1}^{\infty}$ such that $\lim_{l\to\infty} \text{col}(\boldsymbol{X}_l)$ does not contain all eigenvectors of $\boldsymbol{\Sigma}(\bar{\rho})$ associated to the eigenvalues of $\boldsymbol{\Sigma}(\bar{\rho})$ different

from $\lambda_1(\boldsymbol{\Sigma}(\bar{\rho}))$, and $\lim_{l\to\infty} \|(\boldsymbol{M}_{\boldsymbol{X}_l}\boldsymbol{f}_{\max})^* - \boldsymbol{g}^*\| = 0$ for some $\boldsymbol{g} \in E_1(\boldsymbol{\Sigma}(\bar{\rho}))$. The existence of a $\boldsymbol{g} \in E_1(\boldsymbol{\Sigma}(\bar{\rho}))$ such that (14) holds then follows from Lemma A.3. We have thus established that there exists a nonempty set of k-dimensional regression spaces such that the limiting power of a POI test vanishes. That such a set has positive invariant measure on $G_{k,n}$ follows by the same argument used in the proof of Theorem 3.3 for the Cliff-Ord test. ∎

Proof of Proposition 3.6. We start from the Cliff-Ord test. Write $I = \boldsymbol{t}'\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}'\boldsymbol{t}/\boldsymbol{t}'\boldsymbol{t}$, with $\boldsymbol{t} := \boldsymbol{C}\boldsymbol{y}$ and $\boldsymbol{A} := (\boldsymbol{W} + \boldsymbol{W}')/2$. Then, by Lemma A.2,

$$I \leq \lambda_{n-k}(\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}'), \tag{15}$$

for all $\boldsymbol{y} \in \mathbb{R}^n$. Consider some arbitrary α, $\boldsymbol{X}$ and $\boldsymbol{W} \in \mathcal{C}$. Suppose first that $\boldsymbol{f}_{\max} \notin \text{col}(\boldsymbol{X})$. By Lemma A.5, $I(\boldsymbol{f}_{\max}) = \lambda_{n-k}(\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}')$, which implies that $\boldsymbol{f}_{\max}$ is in the interior of the Cliff-Ord critical region. The limiting power of the Cliff-Ord test is thus 1, by Lemma A.1. Let us now suppose that $\boldsymbol{f}_{\max} \in \text{col}(\boldsymbol{X})$. For any $\boldsymbol{W} \in \mathcal{C}$, $m_1(\boldsymbol{A}) = n - 1$, and hence the application of Lemma A.3 with $\boldsymbol{Q} = \boldsymbol{A}$ yields $\lambda_{n-k}(\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}') = \lambda_1(\boldsymbol{A})$. Thus, $I \leq \lambda_1(\boldsymbol{A})$ by (15). But, by Lemma A.2, $I \geq \lambda_1(\boldsymbol{A})$, for any $\boldsymbol{y}$ and any $\boldsymbol{X}$. Thus I is degenerate because it equals $\lambda_1(\boldsymbol{A})$ for any $\boldsymbol{y}$ and any $\boldsymbol{X}$.

The proof for a POI test is similar. By Lemma A.2,

$$P_{\bar{\rho}} \geq \lambda^{-1}_{n-k}(\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}'). \tag{16}$$

If $\boldsymbol{f}_{\max} \notin \text{col}(\boldsymbol{X})$, then, by Lemma A.6, $P_{\bar{\rho}}(\boldsymbol{f}_{\max}) = \lambda^{-1}_{n-k}(\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}')$, and hence the limiting power of a POI test is 1, by Lemma A.1. Recall that $\boldsymbol{\Sigma}(\bar{\rho}) = (\boldsymbol{I}_n - \bar{\rho}(\boldsymbol{W} + \boldsymbol{W}') + \bar{\rho}^2\boldsymbol{W}'\boldsymbol{W})^{-1}$ and that, for any $\boldsymbol{W} \in \mathcal{C}^*$, $m_1(\boldsymbol{W} + \boldsymbol{W}') = n - 1$ and $m_1(\boldsymbol{W}'\boldsymbol{W}) \geq n - 1$. When $\boldsymbol{W} \in \mathcal{C}^*$, $\boldsymbol{W} + \boldsymbol{W}'$ and $\boldsymbol{W}'\boldsymbol{W}$ commute, by Lemma A.4, and hence they are simultaneously diagonalizable. It follows that $m_1(\boldsymbol{\Sigma}(\bar{\rho})) = n-1$ for any $\boldsymbol{W} \in \mathcal{C}^*$. Thus, when $\boldsymbol{f}_{\max} \notin \text{col}(\boldsymbol{X})$ we can apply Lemma A.3 with $\boldsymbol{Q} = \boldsymbol{\Sigma}(\bar{\rho})$, to obtain $\lambda_{n-k}(\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}') = \lambda_1(\boldsymbol{\Sigma}(\bar{\rho}))$. But then, using again Lemma A.2 as for the Cliff-Ord test, we reach the conclusion that $P_{\bar{\rho}}$ does not depend on $\boldsymbol{y}$ if $\boldsymbol{f}_{\max} \in \text{col}(\boldsymbol{X})$, which completes the proof. ∎

Proof of Proposition 3.7. For any $\boldsymbol{W} \in \mathcal{C}$, the symmetric matrix $\boldsymbol{W} + \boldsymbol{W}'$ admits the spectral decomposition

$$\boldsymbol{W} + \boldsymbol{W}' = 2\lambda_{\max}\boldsymbol{f}_{\max}\boldsymbol{f}'_{\max} + \lambda_1(\boldsymbol{W} + \boldsymbol{W}')(\boldsymbol{I}_n - \boldsymbol{f}_{\max}\boldsymbol{f}'_{\max}). \tag{17}$$

Since $(\boldsymbol{W})_{ii} = 0$ for $i = 1, ..., n$, $\boldsymbol{W} + \boldsymbol{W}'$ has zero trace. Hence the sum of the eigenvalues of $\boldsymbol{W}+\boldsymbol{W}'$ must be zero, which implies that $2\lambda_{\max} = -(n-1)\lambda_1(\boldsymbol{W}+\boldsymbol{W}')$. From (17) we obtain

$$\begin{aligned}(\boldsymbol{W} + \boldsymbol{W}')_{i,i} &= \lambda_1(\boldsymbol{W} + \boldsymbol{W}')\left[(1-n)(\boldsymbol{f}_{\max})_i^2 + (1 - (\boldsymbol{f}_{\max})_i^2)\right] \\ &= \lambda_1(\boldsymbol{W} + \boldsymbol{W}')(1 - n(\boldsymbol{f}_{\max})_i^2),\end{aligned}$$

which is 0 for all i if and only if $(\boldsymbol{f}_{\max})_i^2 = 1/n$ for all i. But, since $\boldsymbol{W}$ is a nonnegative irreducible matrix, it follows by the Perron-Frobenius theorem that $\boldsymbol{f}_{\max}$ is entrywise positive or entrywise negative (e.g., Horn and Johnson, 1985, Ch. 8). Hence, for any $\boldsymbol{W} \in \mathcal{C}$, $(\boldsymbol{f}_{\max})_i$ is independent of i, which is the desired conclusion. ∎

Proof of Proposition 3.8. Follows immediately from expression (17) and Proposition 3.7. ∎

Proof of Proposition 3.9. For any $c > 0$, the matrix $c(\boldsymbol{J}_n - \boldsymbol{I}_n)$, where $\boldsymbol{J}_n$ denotes the $n \times n$ matrix of all ones, has the simple eigenvalue $(n-1)c$ and the eigenvalue $-c$ with multiplicity $n-1$. Hence, any such matrix is a weights matrix in $\mathcal{C}$, establishing the sufficiency of the condition in the lemma. To prove the necessity, we start by observing that if a real normal matrix has only real eigenvalues, then it is symmetric (Horn and Johnson, 1985, p. 109). Thus, by Lemma A.4, we only need to show that if $\boldsymbol{W} \in \mathcal{C}$ is symmetric and has zero diagonal entries, then it has identical and positive off-diagonal entries. If $\boldsymbol{W} \in \mathcal{C}$ is symmetric, we can write

$$\boldsymbol{W} = \lambda_1(\boldsymbol{W})\boldsymbol{I}_n + (\lambda_n(\boldsymbol{W}) - \lambda_1(\boldsymbol{W}))\boldsymbol{f}_{\max}\boldsymbol{f}'_{\max}. \tag{18}$$

Such an expression shows that if $(\boldsymbol{W})_{ii} = 0$ for $i = 1, ..., n$, then $(\boldsymbol{f}_{\max})_i^2$ is independent of i. Since $\boldsymbol{W}$ is nonnegative and irreducible, it follows that $\boldsymbol{f}_{\max}$ has identical entries, by the same argument used at the end of the proof of Proposition 3.7. Then, by (18), the off-diagonal entries of $\boldsymbol{W}$ are identical and positive, which completes the proof. ∎

Proof of Theorem 4.1. We start from the part of the theorem relative to the Cliff-Ord test, when $k \leq m_1(\boldsymbol{W})$. The first step in the proof corresponds to a simplification of the first part of the proof of Lemma 3.2, for the case when $\boldsymbol{W}$ is symmetric. By replacing the bound $I \geq \lambda_1(\boldsymbol{A})$ (that holds for all $\boldsymbol{y} \in \mathbb{R}^n$ and all $\boldsymbol{X} \in \mathbb{R}^{n\times k}$) with the bound $I \geq \lambda_1(\boldsymbol{CWC}')$ (that holds for all $\boldsymbol{y} \in \mathbb{R}^n$ and for a fixed $\boldsymbol{X} \in \mathbb{R}^{n\times k}$), we obtain that the limiting power of the Cliff-Ord test vanishes for all values of α if (i) $\boldsymbol{f}_{\max} \notin \text{col}(\boldsymbol{X})$, (ii) $I(\boldsymbol{f}_{\max}) = \lambda_1(\boldsymbol{CWC}')$, and (iii) $\lambda_1(\boldsymbol{CWC}') \neq \lambda_{n-k}(\boldsymbol{CWC}')$. Observe that $\boldsymbol{W} \notin \mathcal{C}$, because in order for the size α of an invariant critical region to be less than 1, $n > 2$. Any $\text{col}(\boldsymbol{X}) \in \Theta$ satisfies (i) by definition. We are now going to show that there are $\text{col}(\boldsymbol{X}) \in \Theta$ such that $\boldsymbol{C}\boldsymbol{f}_{\max} \in E_1(\boldsymbol{CWC}')$, which is equivalent to (ii). Since $\boldsymbol{W}$ is symmetric, its eigenvectors $\boldsymbol{f}_i(\boldsymbol{W})$, $i = 1, ..., n$, are pairwise orthogonal. Thus, if $\text{col}(\boldsymbol{X}) \in \Theta$, $\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{f}_i(\boldsymbol{W}) = \boldsymbol{f}_i(\boldsymbol{W})$, for $i = m_1(\boldsymbol{W}) + 1, ..., n-1$. It follows that, for any $\text{col}(\boldsymbol{X}) \in \Theta$ and for $i = m_1(\boldsymbol{W}) + 1, ..., n-1$,

$$\boldsymbol{CWC}'\boldsymbol{C}\boldsymbol{f}_i(\boldsymbol{W}) = \boldsymbol{CW}\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{f}_i(\boldsymbol{W}) = \boldsymbol{CW}\boldsymbol{f}_i(\boldsymbol{W}) = \lambda_i(\boldsymbol{W})\boldsymbol{C}\boldsymbol{f}_i(\boldsymbol{W}).$$

That is, the $(n-k) \times (n-k)$ matrix $\boldsymbol{CWC}'$ admits the $n - m_1(\boldsymbol{W}) - 1$ eigenpairs $(\lambda_i(\boldsymbol{W}), \boldsymbol{C}\boldsymbol{f}_i(\boldsymbol{W}))$, $i = m_1(\boldsymbol{W}) + 1, ..., n-1$. But then, by the symmetry of $\boldsymbol{CWC}'$ and the fact that the vectors $\boldsymbol{C}\boldsymbol{f}_i(\boldsymbol{W})$, $i = m_1(\boldsymbol{W}) + 1, ..., n-1$ are pairwise orthogonal (because the $\boldsymbol{f}_i(\boldsymbol{W})$'s are), the remaining eigenvectors of $\boldsymbol{CWC}'$ must be in the subspace spanned by the vectors $\boldsymbol{C}\boldsymbol{f}_1(\boldsymbol{W}), ..., \boldsymbol{C}\boldsymbol{f}_{m_1(\boldsymbol{W})}(\boldsymbol{W}), \boldsymbol{C}\boldsymbol{f}_{\max}$. Observe that,

for any $\text{col}(\boldsymbol{X}) \in \Theta$ and for any $\boldsymbol{g} \in E_1(\boldsymbol{W}) \cap \text{col}(\boldsymbol{X})$, $\boldsymbol{C}\boldsymbol{f}_{\max}$ and $\boldsymbol{C}\boldsymbol{g}$ are linearly dependent. Thus, $\boldsymbol{C}\boldsymbol{f}_{\max}$ must be an eigenvector of $\boldsymbol{C}\boldsymbol{W}\boldsymbol{C}'$, i.e.,

$$\boldsymbol{C}\boldsymbol{W}\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{f}_{\max} = \tilde{\lambda}\boldsymbol{C}\boldsymbol{f}_{\max}, \tag{19}$$

for some eigenvalue $\tilde{\lambda}$. The condition $\boldsymbol{C}\boldsymbol{f}_{\max} \in E_1(\boldsymbol{C}\boldsymbol{W}\boldsymbol{C}')$ is satisfied if and only if

$$\tilde{\lambda} \leq \lambda_{m_1(\boldsymbol{W})+1}(\boldsymbol{W}). \tag{20}$$

As $\text{col}(\boldsymbol{X}) \in \Theta$ approaches a subspace orthogonal to $E_1(\boldsymbol{W})$, $\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{f}_{\max}$ tends to a vector in $E_1(\boldsymbol{W})$, which implies, by (19), that $\tilde{\lambda} \to \lambda_1(\boldsymbol{W})$ (note that, by the definition of Θ, no $\text{col}(\boldsymbol{X}) \in \Theta$ can be orthogonal to $E_1(\boldsymbol{W})$). Thus, by the continuity of the eigenvalues of a matrix ($\boldsymbol{C}\boldsymbol{W}\boldsymbol{C}'$ here) in the entries of the matrix, plus the fact that $\lambda_1(\boldsymbol{W}) < \lambda_{m_1(\boldsymbol{W})+1}(\boldsymbol{W})$, there always are $\text{col}(\boldsymbol{X}) \in \Theta$ that satisfy (20) and hence condition (ii). Such regression spaces also satisfy condition (iii), because, in order for $\lambda_1(\boldsymbol{C}\boldsymbol{W}\boldsymbol{C}') = \lambda_{n-k}(\boldsymbol{C}\boldsymbol{W}\boldsymbol{C}')$, we should have that all the eigenvalues of $\boldsymbol{C}\boldsymbol{W}\boldsymbol{C}'$ are identical, but this is impossible when $\boldsymbol{W} \notin \mathcal{C}$.

So far, we have established the part of the theorem relative to the case $k \leq m_1(\boldsymbol{W})$, for the Cliff-Ord test. The extension to a POI test is straightforward, by relying precisely on the same modifications necessary to move from the proof of Lemma 3.2 to the proof of Theorem 3.5. We now turn to the part of the theorem concerned with the particular case $k = m_1(\boldsymbol{W}) = 1$. Let $\boldsymbol{X}$ be a scalar multiple of $\boldsymbol{f}_1(\boldsymbol{W}) + \omega\boldsymbol{f}_{\max}$, so that $\text{col}(\boldsymbol{X}) \in \Theta$ as long as $\omega \neq 0$. For the Cliff-Ord test, we need to establish which values of ω satisfy (20). When $m_1(\boldsymbol{W}) = 1$, (20) reads

$$\tilde{\lambda} \leq \lambda_2(\boldsymbol{W}), \tag{21}$$

where $\tilde{\lambda}$ is the eigenvalue of $\boldsymbol{C}\boldsymbol{f}_{\max}$ associated to the eigenvector of $\boldsymbol{C}\boldsymbol{W}\boldsymbol{C}'$. Observe that

$$\begin{aligned}\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{f}_{\max} &= \left[\boldsymbol{I}_n - \frac{1}{1+\omega^2}(\boldsymbol{f}_1(\boldsymbol{W}) + \omega\boldsymbol{f}_{\max})(\boldsymbol{f}_1'(\boldsymbol{W}) + \omega\boldsymbol{f}_{\max}')\right]\boldsymbol{f}_{\max} \\ &= \boldsymbol{f}_{\max} - \frac{\omega}{1+\omega^2}(\boldsymbol{f}_1(\boldsymbol{W}) + \omega\boldsymbol{f}_{\max}) = \frac{1}{1+\omega^2}(\boldsymbol{f}_{\max} - \omega\boldsymbol{f}_1(\boldsymbol{W})),\end{aligned}$$

where we have used the fact that $\boldsymbol{f}_1(\boldsymbol{W})$ and $\boldsymbol{f}_{\max}$ are normalized and orthogonal. Plugging the above expression for $\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{f}_{\max}$ in (19) gives

$$\frac{1}{1+\omega^2}\left[\lambda_{\max}\boldsymbol{C}\boldsymbol{f}_{\max} - \omega\lambda_1(\boldsymbol{W})\boldsymbol{C}\boldsymbol{f}_1(\boldsymbol{W})\right] = \tilde{\lambda}\boldsymbol{C}\boldsymbol{f}_{\max}. \tag{22}$$

Now, since $\boldsymbol{C}\boldsymbol{X} = \boldsymbol{O}_n$, $\boldsymbol{C}\boldsymbol{f}_1(\boldsymbol{W}) = -\omega\boldsymbol{C}\boldsymbol{f}_{\max}$. Hence, from (22) we obtain $\tilde{\lambda} = [\lambda_{\max} + \omega^2\lambda_1(\boldsymbol{W})]/(1+\omega^2)$. This expression can be used to solve (21) in terms of ω, which yields $|\omega| \geq [\lambda_{\max} - \lambda_2(\boldsymbol{W})]^{1/2} / [\lambda_2(\boldsymbol{W}) - \lambda_1(\boldsymbol{W})]^{1/2}$. The extension to a POI test can be performed by replacing $\boldsymbol{W}$ with $\boldsymbol{\Sigma}(\bar{\rho}) = (\boldsymbol{I}_n - \bar{\rho}\boldsymbol{W})^{-2}$. In this case we need to establish for which values of ω

$$\bar{\lambda} \leq \lambda_2(\boldsymbol{\Sigma}(\bar{\rho})), \tag{23}$$

where $\bar{\lambda}$ is the eigenvalue of $\boldsymbol{C}\boldsymbol{f}_{\max}$ associated to the eigenvector of $\boldsymbol{C}\boldsymbol{\Sigma}(\bar{\rho})\boldsymbol{C}'$. Proceeding exactly as for the Cliff-Ord test, (23) yields

$$|\omega| \geq [\lambda_n(\boldsymbol{\Sigma}(\bar{\rho})) - \lambda_2(\boldsymbol{\Sigma}(\bar{\rho}))]^{1/2}/[\lambda_2(\boldsymbol{\Sigma}(\bar{\rho})) - \lambda_1(\boldsymbol{\Sigma}(\bar{\rho}))]^{1/2}. \tag{24}$$

Using $\lambda_i(\boldsymbol{\Sigma}(\bar{\rho})) = (1 - \bar{\rho}\lambda_i(\boldsymbol{W}))^{-2}$, and after some straightforward algebra, from (24) we obtain $|\omega| \geq \omega_1\omega_2$, which completes the proof. ∎

Acknowledgements

I am grateful to two anonymous referees for their comments. I also wish to thank Ingmar Prucha for several valuable suggestions, Peter Burridge, Grant Hillier, Harry Kelejian, Patrick Marsh, Paolo Paruolo, Tony Smith, participants at the ESRC Econometrics Study Group, Bristol, 2008, and at the Spatial Econometrics Association World Conference, New York, 2008, for discussions and encouragement.

REFERENCES

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

Anselin, L. (2002). Under the hood: issues in the specification and interpretation of spatial regression models. *Agr. Econ.* 27:247–267.

Arnold, S. F. (1979). Linear models with exchangeably distributed errors. *J. Amer. Statistical Assoc.* 74:194-199.

Baltagi, B. H. (2006). Random effects and spatial autocorrelation with equal weights. *Econometric Th.* 22:973-984.

Baltagi, B. H., Liu, L. (2009). Spatial lag test with equal weights. *Economics Letters* 104:81-82.

Baltagi, B. H., Liu, L. (2010) Spurious spatial regression with equal weights. *Stat. Probabil. Lett.* 80, 1640-1642.

Basu, S., Reinsel, G. C. (1994). Regression models with spatially correlated errors. *J. Amer. Statistical Assoc.* 89:88-99.

Bell, P. K., Bockstael, N. E. (2000). Applying the generalized-moments estimation approach to spatial problems involving microlevel data. *Rev. Econ. Statist.* 82:72-82.

Besley, T., Case, A. (1995). Incumbent behavior: vote-seeking, tax-setting, and yardstick competition. *Amer. Econ. Rev.* 85:25-45.

Biggs, N. L. (1993). *Algebraic Graph Theory* (2nd ed.). Cambridge: Cambridge University Press.

Case, A. (1991). Spatial patterns in household demand. *Econometrica* 59:953-966.

Cliff, A. D., Ord, J. K. (1973). *Spatial Autocorrelation.* London: Pion.

Cordy, C. B., Griffith, D. A. (1993). Efficiency of least squares estimators in the presence of spatial autocorrelation. *Commun. Stat. Simulat.* 22:1161-1179.

De Long, J. B., Summers, L. H. (1991). Equipment investment and economic growth. *Quarterly J. Econ.* 106:445-502.

Fingleton, B. (1999). Spurious spatial regression: Some Monte Carlo results with a spatial unit root and spatial cointegration. *J. Reg. Sci.* 39:1-19.

Horn, R., Johnson, C.R. (1985). *Matrix Analysis.* Cambridge: Cambridge University Press.

James, A. T. (1954). Normal multivariate analysis and the orthogonal group. *Ann. Math. Stat.* 25:40-75.

Jones, C. S., Finn, J. M., Hengartner, N. (2008). Regression with strongly correlated data. *J. Multiv. Anal.* 99:2136-2153.

Kariya, T. (1980). Note on a condition for equality of sample variances in a linear model. *J. Amer. Statistical Assoc.* 75:701-703.

Kelejian, H. H., Prucha, I. R. (2001). On the asymptotic distribution of the Moran I test statistic with applications. *J. Econometrics* 104:219–257.

Kelejian, H. H., Prucha, I. R. (2002). 2SLS and OLS in a spatial autoregressive model with equal spatial weights. *Reg. Sci. Urban. Econ.* 32:691–707.

Kelejian, H. H., Prucha, I. R., Yuzefovich, Y. (2006). Estimation problems in models with spatial weighting matrices which have blocks of equal elements, *J. Reg. Sci.* 46:507-515.

Kelejian, H. H., Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *J. Econometrics* 157:53-67.

King, M. L. (1980). Robust tests for spherical symmetry and their application to least squares regression. *Ann. Stat.* 8:1265-1271.

King, M. L. (1981). A small sample property of the Cliff-Ord test for spatial autocorrelation. *J. Roy Statist. Soc.* B43:263-4.

King, M. L. (1988). Towards a theory of point optimal testing. *Econometric Rev.* 6:169–255.

Krämer, W. (2005). Finite sample power of Cliff–Ord-type tests for spatial disturbance correlation in linear regression. *J. Stat. Plan. Infer.* 128:489-496.

Lee, L. F., Yu, J. (2008). Near unit root in the spatial autoregressive model. Manuscript, The Ohio State University.

Lehmann, E. L., Romano, J. (2005). *Testing Statistical Hypotheses* (3rd ed.). New York: Springer.

Martellosio, F. (2010a). Power properties of invariant tests for spatial autocorrelation in linear regression. *Econometric Th.* 26:152-186.

Martellosio, F. (2010b). Non-testability of equal weights spatial dependence. *Econometric Th.,* forthcoming.

Militino, A. F., Ugarte, M. D., García-Reinaldos, L. (2004). Alternative models for describing spatial dependence among dwelling selling prices. *J. Real Estate Finance Econ.* 29:193-209.

Parent, O., LeSage, J. P.(2008). Using the variance structure of the conditional autoregressive spatial specification to model knowledge spillovers. *J. Appl. Econometrics* 23:235-256.

Smith, T. E. (2009). Estimation bias in spatial models with strongly connected weight matrices. *Geogr. Anal.* 41:307-332.

Tillman, J. A. (1975). The power of the Durbin-Watson test. *Econometrica* 43:959-74.

Watson, G. S. (1955). Serial correlation in regression analysis I. *Biometrika* 42:327-341.