Peter M Robinson and **Francesca Rossi**

Refined tests for spatial correlation

Article (Accepted version) (Refereed)

Original citation:
Robinson, Peter M. and Rossi, Francesca (2015) *Refined tests for spatial correlation*. Econometric Theory, 31 (6). pp. 1249-1280. ISSN 0266-4666
DOI: 10.1017/S0266466614000498

Refined Tests for Spatial Correlation

Peter M. Robinson and Francesca Rossi
London School of Economics and University of Southampton

January 7, 2016

Abstract

We consider testing the null hypothesis of no spatial correlation against the alternative of pure first order spatial autoregression. A test statistic based on the least squares estimate has good first-order asymptotic properties, but these may not be relevant in small or moderate-sized samples, especially as (depending on properties of the spatial weight matrix) the usual parametric rate of convergence may not be attained. We thus develop tests with more accurate size properties, by means of Edgeworth expansions and the bootstrap. Though the least squares estimate is inconsistent for the correlation parameter, we show that under quite general conditions its probability limit has the correct sign, and that least squares testing is consistent; we also establish asymptotic local power properties. The finite-sample performance of our tests is compared with others in Monte Carlo simulations.

JEL classifications: C12; C21
Keywords: Spatial Autocorrelation; Ordinary Least Squares; Hypothesis Testing; Edgeworth Expansion; Bootstrap.

1 Introduction

The modelling and analysis of spatially correlated data can pose significant complications and difficulties. Correlation across spatial data is typically a possibility, due to competition, spillovers, aggregation and other circumstances, and might be anticipated in observable variables or in the unobserved disturbances in an econometric model, or both. In, for example, a linear regression model with exogenous regressors, if only the regressors are spatially correlated the usual rules for large sample inference (based on least squares) are unaffected. However, if also the disturbances are spatially correlated then though the least squares estimate (LSE) of the regression coefficients is likely to retain its consistency, its asymptotic variance matrix reflects the correlation. This matrix needs to be consistently estimated in order to carry out statistical inference, and its estimation (whether parametric or nonparametric) offers greater challenges than when time series data are involved, due to the lack of ordering in spatial data, as well as possible irregular spacing or lack of reliable information on locations. In addition the LSE is rendered asymptotically inefficient by spatial correlation, and developing generalized least squares estimates is similarly beset by ambiguities.

A sensible first step in data analysis is therefore to investigate whether or not there is evidence of spatial correlation, by carrying out a statistical test of the null hypothesis of no spatial correlation. A number of such tests have been developed, see e.g. Burridge (1980), Cliff and Ord (1981), Lee and Yu (2012), Kelejian and Prucha (2001), Li et al (2007), Martellosio (2012), Moran (1950), Pinkse (2004). A number of them have been directed against the (first-order) spatial autoregression (SAR). For simplicity we stress the case of zero mean observable data, and will also allow for an unknown intercept, but our work can be extended to test for lack of spatial correlation in unobservable disturbances in more general models, such as regressions. Given the $n \times 1$ vector of observations $y = (y_1, ..., y_n)'$, the prime denoting transposition, the SAR model is

$$y = \lambda W y + \epsilon, \tag{1.1}$$

where $\epsilon = (\epsilon_1, ..., \epsilon_n)'$ consists of unobservable, uncorrelated random variables with zero mean and unknown variance σ^2, λ is an unknown scalar, and W is an $n \times n$ user-specified "weight" matrix, having (i, j)-th element w_{ij}, where $w_{ii} = 0$ for all i and (in order to identify λ) normalization restrictions may be applied.

Such restrictions imply that in general each element w_{ij} changes as n increases, implying that W, and thus y, form triangular arrays (i.e. $W = W_n = (w_{ijn})$, $y = y_n = (y_{in})$) but we suppress reference to the n subscript. The element w_{ij} can be regarded as a (scaled) inverse economic distance between locations i and j, where symmetry of W is not necessarily imposed. Thus knowledge of actual locations is not required, extending the applicability of the model beyond situations when they are known, and entailing simpler modelling and theory than is typically possible when one attempts to incorporate locations of irregularly spaced geographical observations.

The null hypothesis of interest is

$$H_0 : \lambda = 0, \tag{1.2}$$

whence the y_i are uncorrelated (and homoscedastic). An obvious statistic for testing (1.2) is the statistic based on the LSE $\hat{\lambda}$ of λ, which is given by

$$\hat{\lambda} = \frac{y'Wy}{y'W'Wy}. \tag{1.3}$$

Due to the dependence between right-hand side observables and disturbances in (1.1), $\hat{\lambda}$ is inconsistent for λ, as discussed by Lee (2002). However, $\hat{\lambda}$ does converge in probability to zero when $\lambda = 0$, so a test for (1.2) based on $\hat{\lambda}$ might be expected to be asymptotically valid. In particular, under (1.1), (1.2) and regularity conditions a central limit theorem for independent non-identically distributed random variables gives

$$\left[tr\left(WW'\right) / \left\{ tr\left(W^2 + WW'\right) \right\}^{1/2} \right] \hat{\lambda} \rightarrow_d \mathcal{N}(0,1), \tag{1.4}$$

as $n \to \infty$. Since the square-bracketed norming factor can be directly computed, asymptotically valid tests against one-sided ($\lambda > 0$ or $\lambda < 0$) or two-sided ($\lambda \neq 0$) hypotheses are readily carried out.

The accuracy of such tests is dependent on the magnitude of n, and the normal approximation might not be expected to be good for smallish n. Moreover, under conditions described later and as shown by Lee (2004) for the Gaussian maximum likelihood estimate (MLE) of λ, the rate of convergence in (1.4) can be less than the usual parametric rate $n^{1/2}$, depending on the assumptions imposed on W as n increases. In particular if $w_{ij} = O\left(1/h\right)$ is imposed, where the positive sequence $h = h_n$ can increase no faster than n, the rate is $(n/h)^{1/2}$, which increases more slowly than $n^{1/2}$ unless h remains bounded. This outcome renders the usefulness of tests based on first-order asymptotics more dubious than in standard parametric situations.

Cliff and Ord (1971) noted that the limit distributions of tests for spatial independence can be innacurate, and proposed an *ad hoc* correction. Higher-order asymptotic expansions can offer theoretically justifiable improvements in finite samples. Bao and Ullah (2007) derived the second-order bias and mean square error of the Gaussian MLE of λ using a Nagar-type expansion, and Bao (2013) gave extensions to models with exogenous regressors and non-normal disturbances, but neither reference studied test statistics. Various refinements of the Moran I/LM statistics have been presented by Cliff and Ord (1981), Terui and Kikuchi (1994), Robinson (2008), Baltagi and Yang (2013) and Robinson and Rossi (2013). Validity of Edgeworth expansions for the distribution of statistics in models involving SAR(1) processes has been established by Jin and Lee (2012), Yang (2013). Earlier, in a quite general setting of irregularly-spaced spatial observations, García-Soidán (1996) studied the validity of Edgeworth expansions for studentized and unstudentized estimates of a scalar parameter, extending work Götze and Hipp (1983) for mixing time series to coverage processes defined on an expanding subset of $\Re^k$, under the assumption of exponentially decaying correlations.

Here we develop tests derived from Edgeworth expansion of the cumulative distribution function (cdf) of $\hat{\lambda}$, which unlike the MLE is advantageously explicitly defined. The cdf of $\hat{\lambda}$ can be computed by simulation or other numerical techniques, while bootstrap tests can be employed to match our higher-order improvements, but our analytical approach sheds light on theoretical features and the order of magnitude of corrections. This seems especially relevant for the SAR model (1.1) because asymptotic theory depends on h as well as n, and explains for example why tests against two-sided alternative can be more accurate than one-sided tests.

Formal Edgeworth expansions are established in the following section for both $\hat{\lambda}$ and for the LSE of an extended model that includes an unknown intercept. In Section 3 we deduce corrected critical values and corrected (asymptotically normal) test statistics and their properties against both one- and two-sided alternatives. Section 4 compares finite sample performance of our tests with bootstrap and Lagrange multiplier (LM) tests, and also compares Edgeworth approximations to the cdf with numerical calculations. The simple test based on (1.4) generally performs worse than the others, but Section 5 shows that under quite general conditions it is consistent against fixed alternatives, indeed the asymptotically biased $\hat{\lambda}$ actually exaggerates spatial dependence, while against local alternatives the left side of (1.4) has the same limit disribution as both the LM statistic and the Wald statistic based on the MLE, so it is efficient. Some final comments are offered in Section 6. Proofs are left to an appendix.

2 Edgeworth expansions for the least squares estimate

The present section develops a (third-order) formal Edgeworth expansion for $\hat{\lambda}$ (1.3) under the null hypothesis of no spatial correlation (1.2). We introduce some assumptions.

Assumption 1 *The ϵ_i are independent normal random variables with mean zero and unknown variance σ^2.*

Normality is an unnecessarily strong condition for the first-order result (1.4), but it provides some motivation for stressing a quadratic form objective function and is familiar in higher-order asymptotic theory. Edgeworth expansions and resulting test statistics are otherwise complicated by the presence of cumulants of ϵ_i. Assumption 1 implies that under (1.2) the y_i are spatially independent.

For a real matrix A, let $||A||$ be the spectral norm of A (i.e. the square root of the largest eigenvalue of $A'A$) and let $||A||_\infty$ be the maximum absolute row sums norm of A (i.e. $||A||_\infty = \max_i \sum_j |a_{ij}|$, in which a_{ij} is the (i,j)th element of A and i and j vary respectively across all rows and columns of A). Let K be a finite generic constant.

Assumption 2

(i) *For all n, $w_{ii} = 0$, $i = 1, ..., n$.*

(ii) *For all sufficiently large n, W is uniformly bounded in row and column sums in absolute value, i.e. $||W||_\infty + ||W'||_\infty \leq K$*

(iii) *For all sufficiently large n, uniformly in $i, j = 1, ..., n$, $w_{ij} = O(1/h)$, where $h = h_n$ is a positive sequence bounded away from zero for all n such that $h/n \to 0$ as $n \to \infty$.*

Parts (i) and (ii) of Assumption 2 are standard conditions on W imposed in the literature. In particular, (ii) was introduced by Kelejian and Prucha (1998) to keep spatial correlation manageable. If W is symmetric with non-negative elements and row normalized, such that $\Sigma_{j=1}^{n} w_{ij} = 1$ for all i, then Assumption 2(ii) is automatically satisfied. Part (iii) covers two cases which have rather different implications for our results: either h is bounded (when in (1.4) $\hat{\lambda}$ enjoys a parametric $n^{1/2}$ rate of convergence), or h is divergent (when $\hat{\lambda}$ has a slower than parametric, $(n/h)^{1/2}$, rate).

By way of illustration consider (see Case (1991)),

$$W_n = I_r \otimes B_m, \qquad B_m = \frac{1}{m-1}(l_m l_m' - I_m), \tag{2.1}$$

where I_m is the $m \times m$ identity matrix, l_m is the $m \times 1$ vector of 1's, and $\otimes$ denotes Kronecker product. Here W is symmetric with non-negative elements and row normalized, $n = mr$. Parts (i) and (ii) of Assumption 2 are satisfied, and $h \sim m$, where "$\sim$" throughout indicates that the ratio of left and right sides converges to a finite, nonzero constant. Thus in the bounded h case only $r \to \infty$ as $n \to \infty$, whereas in the divergent h case $m \to \infty$ and $r \to \infty$.

Now define

$$t_{ij} = \frac{h}{n} tr(W^i W'^j), \quad i \geq 0, \quad j \geq 0, \quad i + j \geq 1, \tag{2.2}$$

$$t = \frac{h}{n} tr((WW')^2). \tag{2.3}$$

Under Assumption 2 all t_{ij} in (2.2) and t are $O(1)$ (because, for any real A such that $||A||_\infty \leq K$, we have $tr(AW) = O(n/h)$). To ensure the leading terms of the expansion in the theorem below are well defined, we introduce

Assumption 3

$$\underline{\lim}_{n \to \infty} (t_{20} + t_{11}) > 0. \tag{2.4}$$

By the Cauchy inequality, Assumption 3 implies $\underline{\lim}_{n\to\infty} t_{11} > 0$, and the two conditions are equivalent when W is symmetric or when its elements are all non-negative. Assumption 3 is automatically satisfied under (2.1). It follows from Assumptions 2 and 3 that in (1.4) the norming factor

$$\frac{tr(WW')}{(tr(W^2 + WW'))^{1/2}} = \frac{t_{11}}{(t_{20} + t_{11})^{1/2}} \left(\frac{n}{h}\right)^{1/2} \sim \left(\frac{n}{h}\right)^{1/2}. \tag{2.5}$$

Now define

$$a = \frac{t_{11}}{(t_{20}+t_{11})^{1/2}}, \quad b = \frac{t_{21}}{(t_{20}+t_{11})^{1/2}t_{11}}, \quad c = \frac{2t_{30}+6t_{21}}{(t_{20}+t_{11})^{3/2}}, \tag{2.6}$$

$$d = \frac{t}{t_{11}^2}, \quad e = \frac{12(t_{31}+t_{22})}{(t_{20}+t_{11})t_{11}}, \quad f = \frac{6t_{40}+24t_{31}+6t_{22}+12t}{(t_{20}+t_{11})^2}, \quad g = \frac{1}{t_{20}+t_{11}} \tag{2.7}$$

and

$$U(\zeta) = 2b\zeta^2 - \frac{c}{6}H_2(\zeta), \tag{2.8}$$

$$V(\zeta) = \frac{1}{6}(e-6bc)\zeta H_2(\zeta) - (d-6b^2)\zeta^3 - \frac{1}{24}fH_3(\zeta) + \frac{1}{3}bc\zeta^2 H_3(\zeta) - 2b^2\zeta^5, \tag{2.9}$$

where $H_j(\zeta)$ is the jth Hermite polynomial, such that

$$H_2(\zeta) = \zeta^2 - 1 \quad H_3(\zeta) = \zeta^3 - 3\zeta. \tag{2.10}$$

Thus $U(\zeta)$ is an even, generally non-homogeneous, quadratic function of ζ, while $V(\zeta)$ is an odd, generally non-homogeneous, polynomial in ζ of degree 5.

Write $\Phi(\zeta) = Pr(Z \le \zeta)$ for a standard normal random variable Z, and $\phi(\zeta)$ for the probability density function (pdf) of Z. Let $F(\zeta) = P\left((n/h)^{1/2}a\hat{\lambda} \le \zeta\right)$.

Theorem 1 *Let (1.1) and Assumptions 1-3 hold. Under H_0 in (1.2), for any real ζ, $F(\zeta)$ admits the third order formal Edgeworth expansion*

$$F(\zeta) = \Phi(\zeta) + U(\zeta)\phi(\zeta)\left(\frac{h}{n}\right)^{1/2} + V(\zeta)\phi(\zeta)\frac{h}{n} + O\left(\left(\frac{h}{n}\right)^{3/2}\right), \tag{2.11}$$

where

$$U(\zeta) = O(1), \quad V(\zeta) = O(1), \tag{2.12}$$

as $n \to \infty$.

Generally, $U(\zeta)$ and $V(\zeta)$ are non-zero, whence there are leading correction terms of exact orders $(h/n)^{1/2}$ and h/n, and both terms are known functions of ζ.

A corresponding result to Theorem 1 is available for the pure SAR model with unknown intercept, i.e.

$$y = \mu l + \lambda W y + \epsilon, \tag{2.13}$$

where μ is an unknown scalar and $l = l_n$. The LSE of λ in (2.13) is

$$\tilde{\lambda} = \frac{y'W'Py}{y'W'PWy}, \tag{2.14}$$

where $P = I_n - ll'/n$. Under (1.2), the same kind of regularity conditions and the additional

Assumption 4 *For all n, $\Sigma_{j=1}^{n} w_{ij} = 1$, $i = 1, ..., n$,*

$\tilde{\lambda}$ has the same first-order limit distribution as $\hat{\lambda}$, so (1.4) holds with $\hat{\lambda}$ replaced by $\tilde{\lambda}$. However the second- and higher-order limit distributions differ. In case Assumption 4 is not satisfied also the first-order limit distribution of $\tilde{\lambda}$ under (1.2) differs from that of $\hat{\lambda}$ and, in particular, $\tilde{\lambda}$ converges to the true value at the standard $n^{1/2}$ rate whether h is bounded or divergent as $n \to \infty$. Since the main goal of this paper is to provide refined tests when the rate of convergence might be slower than the parametric rate $n^{1/2}$, the case of model (2.13) when W is not row-normalized is not considered here.

Define

$$\tilde{U}(\zeta) = U(\zeta) + g^{1/2} \tag{2.15}$$

and

$$\tilde{V}(\zeta) = V(\zeta) + \left\{ \frac{g}{2}(1+p) + 2bg^{1/2} - \frac{g^4}{2} \right\} \zeta - 2bg^2\zeta^3 + \frac{cg^{1/2}}{6} H_3(\zeta), \tag{2.16}$$

where

$$p = l'WW'l/n. \tag{2.17}$$

(When W is symmetric Assumption 4 implies $p = 1$). Let $\tilde{F}(\zeta) = P((n/h)^{1/2} a\tilde{\lambda} \leq \zeta)$.

Theorem 2 *Let (2.13) and Assumptions 1-4 hold. Under H_0 in (1.2), for any real ζ, $\tilde{F}(\zeta)$ admits the third order formal Edgeworth expansion*

$$\tilde{F}(\zeta) = \Phi(\zeta) + \tilde{U}(\zeta)\phi(\zeta)\left(\frac{h}{n}\right)^{1/2} + \tilde{V}(\zeta)\phi(\zeta)\frac{h}{n} + O\left(\left(\frac{h}{n}\right)^{3/2}\right), \tag{2.18}$$

where

$$\tilde{U}(\zeta) = O(1), \quad \tilde{V}(\zeta) = O(1), \tag{2.19}$$

as $n \to \infty$.

The second- and third-order correction terms are again generally non-zero, and of orders $(h/n)^{1/2}$ and h/n respectively. Notice that $\tilde{U}(\zeta) > U(\zeta)$, so the second-order approximate distribution function (df) of $\tilde{\lambda}$ is greater than that of $\hat{\lambda}$. The Edgeworth approximation in (2.18) is unaffected by μ (and the approximations in both (2.11) and (2.18) are unaffected by σ^2). Consequently results can be similarly obtained when there is a more general linear regression component than in (2.13), at least when regressors are non-stochastic or strictly exogenous. Indeed, similar techniques will yield approximations with respect to the model $y - \mu l = \lambda W(y - \mu l) + \epsilon$, or more general linear regression models with SAR disturbances.

Theorems 1 and 2 continue to hold after replacing ϵ in (1.1) by $w = f(\epsilon)\epsilon$, for almost surely nonzero, scalar functions f (so in general the elements of w form a triangular array). For example, if $f(\epsilon) = \left(\sigma^2 n/\epsilon'\epsilon\right)^{1/2}$ the elements of w have zero mean and variance σ^2 and are uncorrelated, but they are not independent, indeed having a singular distribution for each n (as therefore do the observations y_i), being uniformly distributed on the $n-$sphere with radius $\left(\sigma^2 n\right)^{1/2}$.

3 Improved tests for no spatial correlation

We consider first tests of the null hypothesis (1.2) against the alternative

$$H_1 : \quad \lambda > 0 \tag{3.1}$$

in the no-intercept model (1.1).

For $\alpha \in (0, 1)$ (for example $\alpha = 0.05$ or $\alpha = 0.01$) define the normal critical value z_α such that $1 - \alpha = \Phi(z_\alpha)$. Write $q = (n/h)^{1/2} a\hat{\lambda}$. On the basis of (1.4) a test that rejects (1.2) against (3.1) when

$$q > z_\alpha \tag{3.2}$$

has approximate size α. Theorem 1 readily yields more accurate tests that are simple to calculate because the coefficients of $U(\zeta)$ and $V(\zeta)$ are known, W being chosen by the practitioner.

Define the exact critical value w_α such that $1 - \alpha = F(w_\alpha)$, so a test that rejects when $q > w_\alpha$ has exact size α. Also introduce the Edgeworth corrected critical value

$$u_\alpha = z_\alpha - \left(\frac{h}{n}\right)^{1/2} U(z_\alpha). \tag{3.3}$$

Corollary 1 *Let (1.1) and Assumptions 1-3 hold. Under H_0 in (1.2), as $n \to \infty$*

$$\begin{aligned} w_\alpha =& z_\alpha + O\left(\left(\frac{h}{n}\right)^{1/2}\right) & (3.4) \\ =& u_\alpha + O\left(\frac{h}{n}\right). & (3.5) \end{aligned}$$

Corollary 1 follows follows immediately from Theorem 1. From Corollary 1, the test that rejects (1.2) against (3.1) when

$$q > u_\alpha \tag{3.6}$$

is more accurate than (3.2). Of course when the alternative of interest is $\lambda < 0$, the same conclusion can be drawn for the tests which reject when $q < -z_\alpha$, $q < -u_\alpha$, respectively.

Instead of correcting critical values we can derive from Theorem 1 a corrected test statistic that can be compared with z_α. Introduce the polynomial

$$G(\zeta) = \zeta + \left(\frac{h}{n}\right)^{1/2} U(\zeta) + \frac{h}{n}\frac{1}{3}\left(2b - \frac{c}{6}\right)^2 \zeta^3 \tag{3.7}$$

which has known coefficients (see Yanagihara et al. (2005)). Since $G(\zeta)$ has derivative $(1 + \zeta(2b - c/6)(h/n)^{1/2})^2 > 0$, it is monotonically increasing. Thus $F(\zeta) = P(G(q) \le G(\zeta))$ and we invert the expansion in Theorem 1 to obtain

Corollary 2 *Let (1.1) and Assumptions 1-3 hold. Under H_0, as $n \to \infty$*

$$P(G(q) > z_\alpha) = \alpha + O\left(\frac{h}{n}\right). \tag{3.8}$$

Thus the test that rejects when

$$G(q) > z_\alpha \tag{3.9}$$

has size that differs from α by smaller order than the size of (3.2).

Still more accurate tests can be deduced from Theorem 1 by employing also the third-order correction factor $V(\zeta)$, but the above tests have the advantage of simplicity. The V term, however, is especially relevant in deriving improved tests against the two-sided alternative hypothesis

$$H_0: \quad \lambda \neq 0. \tag{3.10}$$

Because $U(\zeta)$ is an even function it follows from Theorem 1 that

$$P(|q| \leq \zeta) = 2\Phi(\zeta) - 1 + 2\frac{h}{n}V(\zeta) + O\left(\left(\frac{h}{n}\right)^{3/2}\right). \tag{3.11}$$

Thence define the Edgeworth-corrected critical value for a two-sided test,

$$v_{\alpha/2} = z_{\alpha/2} - \frac{n}{h}V(z_{\alpha/2}), \tag{3.12}$$

noting that the approximate size-α two-sided test based on (1.4) rejects H_0 against (3.10) when

$$|q| > z_{\alpha/2}. \tag{3.13}$$

Also, define $s_{\alpha/2}$ such that $P(|q| \leq s_{\alpha/2}) = 1 - \alpha$.

Corollary 3 *Let (1.1) and Assumptions 1-3 hold. Under H_0, as $n \to \infty$*

$$s_{\alpha/2} = z_{\alpha/2} + O\left(\frac{h}{n}\right) \tag{3.14}$$

$$= v_{\alpha/2} + O\left(\left(\frac{h}{n}\right)^{3/2}\right). \tag{3.15}$$

Thus rejecting (1.2) against (3.10) when

$$|q| > v_{\alpha/2} \tag{3.16}$$

rather than (3.13) reduces the error to $O((h/n)^{3/2})$. In fact, Theorem 1 can be established to fourth-order, with fourth-order term that is even in ζ, and error $O((h/n)^2)$, so the error in (3.15) can be improved to $O((h/n)^2)$.

As with the one-sided alternative (3.1), a corrected test statistic that can be compared with $z_{\alpha/2}$ can be derived from Theorem 1. Define (Yanagihara et al.

(2005))

$$L(\zeta) = \zeta + \frac{h}{n}V(\zeta) + \left(\frac{h}{n}\right)^2 \frac{1}{4}\left(L_1^2\zeta + \frac{L_2^2\zeta^5}{5} + \frac{L_3^2\zeta^9}{9} + \frac{2}{3}L_1L_2\zeta^3 + \frac{2}{5}L_1L_3\zeta^5 + \frac{2}{7}L_2L_3\zeta^7\right), \tag{3.17}$$

where $L_1 = -\frac{1}{6}(e-6bc)+\frac{1}{8}f,\ L_2 = \frac{1}{2}(e-6bc)-3(d-6b^2)-\frac{1}{8}f-3bc$ and $L_3 = \frac{5}{3}bc - 10b^2$, so $L(\zeta)$ is a degree-7 polynomial in ζ with known coefficients. It is readily checked that $V(\zeta)$ has derivative $L_1 + L_2\zeta^2 + L_3\zeta^4$, where $L(\zeta)$ has derivative $(1+(h/n)(L_1+L_2\zeta^2+L_3\zeta^4)/2)^2 > 0$ and is thus monotonically increasing. Therefore, from (3.11), we obtain

Corollary 4 *Let (1.1) and Assumptions 1-3 hold. Under H_0, as $n \to \infty$*

$$P(L(|q|) > z_{\alpha/2}) = \alpha + O\left(\left(\frac{h}{n}\right)^{3/2}\right). \tag{3.18}$$

The transformation in (3.17) and Corollary 4 follow from (3.11) using a minor modification of Theorem 2 of Yanagihara et al. (2005). From the latter result, we conclude that the test that rejects H_0 against (3.10) when

$$L(|q|) > z_{\alpha/2} \tag{3.19}$$

has size which is closer to α than (3.13).

Improved tests can be similarly derived from Theorem 2 for the intercept model in (2.13). We first consider tests of H_0 in (1.2) against (3.1). Let $\tilde{q} = (n/h)^{1/2}a\tilde{\lambda}$. A standard test based on first order asymptotic theory rejects (1.2) against (3.1) at approximate level α when

$$\tilde{q} > z_\alpha. \tag{3.20}$$

Define the exact and Edgeworth-corrected critical values $\tilde{w}_\alpha$, such that $1-\alpha = \tilde{F}(\tilde{w}_\alpha)$, and $\tilde{u}_\alpha = z_\alpha - \tilde{U}(z_\alpha)(h/n)^{1/2} = u_\alpha - g^{1/2}(h/n)^{1/2}$, respectively.

Similarly to Corollaries 1 and 2, from Theorem 2 we deduce

Corollary 5 *Let (2.13) and Assumptions 1-4 hold. Under H_0 in (1.2), as*

$n \to \infty$

$$\tilde{w}_\alpha = z_\alpha + O\left(\left(\frac{h}{n}\right)^{1/2}\right) \tag{3.21}$$

$$= \tilde{u}_\alpha + O\left(\frac{h}{n}\right). \tag{3.22}$$

Notice that $\tilde{u}_\alpha < u_\alpha$ for any α, so that the second-order corrected critical value is lower for the intercept model.

Let

$$\tilde{G}(\zeta) = \zeta + \left(\frac{h}{n}\right)^{1/2} \tilde{U}(\zeta) + \frac{h}{n}\frac{1}{3}\left(2b - \frac{c}{6}\right)^2 \zeta^3 = G(\zeta) + \left(\frac{h}{n}\right)^{1/2} g^{1/2}. \tag{3.23}$$

Corollary 6 *Let (2.13) and Assumptions 1-4 hold. Under H_0 in (1.2), as $n \to \infty$*

$$P(\tilde{G}(\tilde{q}) > z_\alpha) = \alpha + O\left(\frac{h}{n}\right). \tag{3.24}$$

Thus, tests that reject (1.2) against (3.1) when either

$$\tilde{q} > \tilde{u}_\alpha \tag{3.25}$$

or

$$\tilde{G}(\tilde{q}) > z_\alpha, \tag{3.26}$$

are more accurate than (3.20).

Also, from Theorem 2 improved tests of (1.2) against (3.10) can be deduced. From (2.18), since $\tilde{U}(\zeta)$ is an even function we obtain,

$$P(|\tilde{q}| \leq \zeta) = 2\Phi(\zeta) - 1 + 2\frac{h}{n}\tilde{V}(\zeta) + O\left(\left(\frac{h}{n}\right)^{3/2}\right). \tag{3.27}$$

Define $\tilde{s}_{\alpha/2}$ such that $P(|\tilde{q}| \leq \tilde{s}_{\alpha/2}) = 1 - \alpha$ and $\tilde{v}_{\alpha/2} = z_{\alpha/2} - (n/h)\tilde{V}(z_{\alpha/2})$. A standard, approximate size α, two-sided test rejects (1.2) against (3.10) when

$$|\tilde{q}| > z_{\alpha/2}. \tag{3.28}$$

From (3.27) we deduce

Corollary 7 *Let (2.13) and Assumptions 1-4 hold. Under H_0, as $n \to \infty$*

$$\tilde{s}_{\alpha/2} = z_{\alpha/2} + O\left(\frac{h}{n}\right) \tag{3.29}$$

$$= \tilde{v}_{\alpha/2} + O\left(\left(\frac{h}{n}\right)^{3/2}\right). \tag{3.30}$$

Finally, define

$$\tilde{L}(\zeta) = \zeta + \frac{h}{n}\tilde{V}(\zeta) + \left(\frac{h}{n}\right)^2 \frac{1}{4}\left(\tilde{L}_1^2\zeta + \frac{\tilde{L}_2^2\zeta^5}{5} + \frac{L_3^2\zeta^9}{9} + \frac{2}{3}\tilde{L}_1\tilde{L}_2\zeta^3 + \frac{2}{5}\tilde{L}_1 L_3\zeta^5 + \frac{2}{7}\tilde{L}_2 L_3\zeta^7\right), \tag{3.31}$$

where $\tilde{L}_1 = L_1 + \frac{g}{2}(1+p) + 2bg^{1/2} - \frac{g^4}{2} - \frac{cg^{1/2}}{2}$, $\tilde{L}_2 = L_2 - 6bg^{1/2} + \frac{cg^{1/2}}{2}$.

Corollary 8 *Let (2.13) and Assumptions 1-4 hold. Under H_0, as $n \to \infty$*

$$P(\tilde{L}(|\tilde{q}|) > z_{\alpha/2}) = \alpha + O\left(\left(\frac{h}{n}\right)^{3/2}\right). \tag{3.32}$$

From Corollaries 7 and 8, we conclude that the tests that reject H_0 against (3.10) when either

$$|\tilde{q}| > \tilde{v}_{\alpha/2} \tag{3.33}$$

or

$$\tilde{L}(|\tilde{q}|) > z_{\alpha/2} \tag{3.34}$$

have sizes closer to α than that obtained from (3.28).

4 Monte Carlo comparison of finite sample performance

In this section we report and discuss a Monte Carlo investigation of the finite sample performance of the tests derived in Section 3 and of bootstrap tests, given that in many circumstances the bootstrap is known to achieve a first-order Edgeworth correction (see e.g. Singh (1981)). For the no-intercept model

(1.1) the bootstrap test is as follows (e.g Paparoditis and Politis (2005)). We construct 199 $n \times 1$ vectors ϵ_j^* for $j = 1, ..., 199$, where each ϵ_j^* is obtained by resampling with replacement from $y_i - \sum_1^n y_i/n$, $i = 1, ..., n$. The bootstrap test statistic is $q_j^* = (n/h)^{1/2} a \epsilon_j^{*'} W_j'^{*} / \epsilon_j^{*'} W_j'^{*}$, $j = 1, ..., 199$, its $(1-\alpha)$th percentile being u_α^* which solves $\sum_{j=1}^{199} 1(q_j^* \leq u_\alpha^*)/199 \leq 1 - \alpha$, where $1(.)$ indicates the indicator function. We reject (1.2) against the one-sided alternative (3.1) when

$$q > u_\alpha^*. \tag{4.1}$$

Defining the $(1-\alpha)$th percentile of $|q_j^*|$ as the value v_α^* solving $\sum_{j=1}^{199} 1(|q_j^*| \leq v_\alpha^*)/199 \leq 1 - \alpha$, we reject (1.2) against the two-sided alternative (3.10) if

$$|q| > v_\alpha^*. \tag{4.2}$$

For the intercept model (2.13) we define $\tilde{q}_j^* = (n/h)^{1/2} a \epsilon_j^{*'} W_j'^{*} / \epsilon_j^{*'} W_j'^{*}$, and the $(1-\alpha)$th quantiles of $\tilde{q}_j^*$ and $|\tilde{q}_j^*|$, $\tilde{u}_\alpha^*$ and $\tilde{v}_\alpha^*$, solve $\sum_{j=1}^{199} 1(\tilde{q}_j^* \leq \tilde{u}_\alpha^*)/199 \leq 1 - \alpha$, and $\sum_{j=1}^{199} 1(|\tilde{q}_j^*| \leq \tilde{v}_\alpha^*)/199 \leq 1 - \alpha$, respectively. We reject (1.2) against (3.1) or (3.10) when

$$\tilde{q} > \tilde{u}_\alpha^* \tag{4.3}$$

or

$$|\tilde{q}| > \tilde{v}_\alpha^*, \tag{4.4}$$

respectively.

We we also compare our tests with ones based on the (signed) square root of the LM statistic and its mean-variance corrected version (Moran (1950), Cliff and Ord (1981)), defined respectively as

$$L = \frac{n}{(tr(WW'^2))^{1/2}} \frac{y'MWMy}{y'My} \tag{4.5}$$

and

$$CL = \gamma^{-1/2} \left(\frac{y'MWMy}{y'My} - \beta \right), \tag{4.6}$$

where

$$\beta = -\frac{n}{n-k} tr((I-M)W), \quad \gamma = \frac{n^2 tr(MWMW + MWMW')}{(n-k)(n-k+2)} - \frac{2n^2 (tr((I-M)W))^2}{(n-k)^2(n-k+2)}, \tag{4.7}$$

in which we take $k = 0$, $M = I$ for the no-intercept model (1.1) and $k = 1$,

$M = P$ for the intercept model (2.13). For the respective statistics, we reject H_0 (1.2) against (3.1) when

$$L > z_{\alpha}, \tag{4.8}$$

and when

$$CL > z_{\alpha}, \tag{4.9}$$

while we reject H_0 (1.2) against (3.10) when

$$|L| > z_{\alpha/2}, \tag{4.10}$$

and when

$$|CL| > z_{\alpha/2}. \tag{4.11}$$

In the simulations we set $\sigma^2 = 1$ in Assumption 1, $\mu = 2$ in (2.13) and choose W as in (2.1), for various m and r. Recalling that orders of magnitudes in Theorems 1 and 2 are affected by whether h diverges or remains bounded as $n \to \infty$, we represent both cases by different choices of $m \sim h$. We choose $(m, r) = (8, 5)$, $(12, 8)$, $(18, 11)$, $(28, 14)$, i.e. $n = 40$, 96, 198, 392, to represent "divergent" h, and $(m, r) = (5, 8)$, $(5, 20)$, $(5, 40)$, $(5, 80)$, i.e. $n = 40$, 100, 200, 400 to represent "bounded" h. For each of these combinations we compute $\hat{\lambda}$ and $\tilde{\lambda}$ from the same realization of ϵ across 1000 replications. In all tests $\alpha = 0.05$.

Empirical sizes are displayed in Tables 1-8, in which "normal", "Edgeworth", "transformation", "bootstrap", "L" and "corrected L" refer respectively to tests using the standard normal approximation, Edgeworth-corrected critical values, Edgeworth-corrected test statistic, bootstrap critical values, LM statistic and LM corrected statistic, and the respective abbreviations N, E, T, B, L, CL will be extensively used in the text.

(Tables 1 and 2 about here)

Tables 1 and 2 cover one-sided tests (3.2), (3.6), (3.9), (4.1), (4.8), (4.9) in the no-intercept model (1.1), when h is respectively "divergent" and "bounded". Test N is drastically under-sized for each n in both tables. The sizes for E are somewhat better, and improve as n increases, in particular for "divergent" h the discrepancy between empirical and nominal sizes is 18% lower relative to N, on average across sample size. Both T and B perform well for all n. Indeed, on average, when h is "divergent" empirical sizes for T and B are 80% and 86%, respectively, closer to 0.05 than those for N, with a similar pattern in Table 2. From Table 2, the average improvements offered by E, T and B over N are about 41%, 89% and 88%, respectively. Overall, T and B perform best among the tests based on LSE. Tables 1 and 2 are consistent with Theorem 1 in which F converges to Φ at rate $n^{1/2}$ when h is bounded, but only at rate $(n/h)^{1/2}$ when h is divergent. Indeed, for N, when h is "bounded", the difference between empirical and nominal size decreases by 20% as n increases from $n = 40$ to $n = 400$, while this difference only decreases by 2% in case h is "divergent" as n increases from $n = 40$ to $n = 392$. From Tables 1 and 2, L and CL drastically outperform both N and E, but on average sizes for T are, respectively, 53% and 52% closer to 5% than those of L and CL when h is "divergent". The latter figures are 54% and 51% when h is "bounded".

(Tables 3 and 4 about here)

Tables 3 and 4 cover two-sided tests for the no-intercept model (1.1), namely (3.13), (3.16), (3.19), (4.2), (4.10) and (4.11). Again, N is very poor, though contrary to the one-sided test case the problem is now over-sizing, and E, T and B all offer notable improvements. Instead, both L and CL appear to be under-sized. When h is "divergent" the difference between empirical and nominal sizes is reduced respectively on average across sample sizes by 87%, 59% and 91% for E, T and B relative to N, and by 86%, 59% and 69% when h is "bounded". In case h is "divergent", L and CL are closer to the nominal 0.05 than both T and N. However, on average E outperforms L by 37%, while CL and E are comparable. When h is "bounded", again L and CL perform better than N and T, but E offers a significant improvement over both L and CL of 67% and 55%, respectively.

(Tables 5 and 6 about here)

Tables 5 and 6 contain results for one-sided tests for the intercept model (2.13), the N, E, T, B, L and CL tests being given in (3.20), (3.25), (3.26), (4.3) (4.8) and (4.9). The pattern is similar to that displayed in Tables 1 and 2. For "divergent" h, on average across sample sizes, empirical sizes for E, T and B are 12%, 65% and 89% closer to 5% than ones for N, with figures of 22%, 79% and 79% for "bounded" h. Overall, B performs best among the LSE-based tests for "divergent" h, but it is difficult to choose between B and T when h is "bounded". Also, similarly to Table 1, when h is "divergent" L and CL outperform both N and E, but empirical sizes for T are 27% closer to 5% than those for CL. For "bounded" h again L and CL are drastically better than N and E, but T offers respective improvements over L and CL of 14% and 41%.

(Tables 7 and 8 about here)

Tables 7 and 8 correspondingly describe two-sided tests given in (3.28), (3.33), (3.34), (4.4), (4.10) and (4.11). The improvements on average across sample sizes offered by E, T and B over N are 58%, 27% and 94%, respectively, when h is "divergent", and 64%, 64% and 47%, respectively, when h is "bounded". For "divergent" h B again comes out top overall among tests based on the LSE, followed by E, but for "bounded" h B is outperformed by both E and T. For both "divergent" and "bounded" h, unlike the results in Tables 1-6, L and CL perform better than N, E and T.

Monte Carlo results for our tests of (1.2) against a two-sided alternative can be compared with the LM-based Edgeworth-corrected tests derived in Robinson and Rossi (2013). Sizes for E and T in Tables 3 and 4 are relatively satisfactory also in light of those in Robinson and Rossi (2013). In particular, for "bounded" h the deviation of empirical sizes for both E and T from the nominal 5% appears to be similar to that for their Edgeworth-corrected tests, while for "divergent" h only E has a similar performance to theirs, T only offering a modest improvement over N. For the intercept model (2.13), Tables 7 and 8 show that Edgeworth-corrected tests in Robinson and Rossi (2013) outperform ours in terms of size properties when h is "divergent". However, for "bounded" h the performance of our tests appears to be comparable to that of Robinson and Rossi (2013), exept for very small sample sizes. In general the Edgeworth-corrected tests in Robinson and Rossi (2013) are under-sized, while from Tables 3-4 and 7-8 it is clear that the problem here is over-sizing.

(Figures 1 and 2 about here)

To illustrate the effect of the transformations $G(.)$ and $\tilde{G}(.)$ used in Section 3, in Figures 1 and 2 we plot the histograms with 100 bins of q and $G(q)$ (Figure 1) and of $\tilde{q}$ and $\tilde{G}(\tilde{q})$ (Figure 2) obtained from 1000 replications when $m = 28$ and $r = 14$. Both figures suggest that the densities of q and $\tilde{q}$ are very skewed to the left and that most of the skewness is removed by the transformations, as in Hall (1992).

(Tables 9-12 about here)

In Tables 9-12 we assess power against a fixed alternative, i.e.

$$H_1 : \lambda = \bar{\lambda} > 0. \tag{4.12}$$

Tables 9 and 10 display the empirical power of one-sided tests in the no-intercept model (1.1) when h is "divergent" and "bounded" respectively, while Tables 11 and 12 correspondingly contain results for the intercept model (2.13). These are non-size-corrected tests. Exept for the smallest sample size when h is "divergent", even N performs well for the largest $\bar{\lambda} = 0.8$, as do all other tests in all settings. N also does comparably well to E, T and B when h is bounded and $\bar{\lambda} = 0.5$. But overall N is outperformed by the other tests, with T and B offering the greatest power among the LSE-based tests. Tables 9 and 10 suggest that T is slightly outperformed by L and CL for almost all sample sizes considered, but the opposite holds true for the intercept model in (2.13), as shown by Tables 11 and 12.

A direct comparison with empirical powers of corrected-tests in Robinson and Rossi (2013) is not possible because here we are focussing on tests of H_0 against a one-sided alternative, (4.12), while in their Monte Carlo study empirical powers of tests of H_0 against $H_1 : \quad \lambda = \bar{\lambda} \neq 0$ are reported. From Robinson and Rossi (2013), however, it is clear that Edgeworth-corrected tests display similar power to that of the standard χ^2 test. Thus, we might expect that one-sided Edgeworth-corrected tests based on L would have similar power to that of L reported in Tables 9-12.

Finally we calculate numerically $F(\zeta)$ and $\tilde{F}(\zeta)$ for various ζ by means of Imhof's (1961) procedure and find that Theorems 1 and 2 approximations work fairly well. Numerical algorithms do have limitations; Lu and King (2002) surveyed the numerical evaluation of the cdf of normal quadratic forms. Let

$$F_3(\zeta) = \Phi(\zeta) + U(\zeta)\phi(\zeta)\left(\frac{h}{n}\right)^{1/2} + V(\zeta)\phi(\zeta)\frac{h}{n} \tag{4.13}$$

and

$$\tilde{F}_3(\zeta) = \Phi(\zeta) + \tilde{U}(\zeta)\phi(\zeta)\left(\frac{h}{n}\right)^{1/2} + \tilde{V}(\zeta)\phi(\zeta)\frac{h}{n} \tag{4.14}$$

be the third-order Edgeworth corrected cdf of q and $\tilde{q}$, respectively. Tables 13 and 14 compare $F_3(\zeta)$ for ζ representing one- and two-sided 5% normal critical values with numerical calculations for "divergent" and "bounded" h, respectively, and Tables 15 and 16 contain results under $\tilde{F}_3(\zeta)$. In the Tables "Edgeworth" refers to either (4.13) or (4.14), while "exact" refers to the numerical procedure.

(Tables 13-16 about here)

In all Tables 13-16, the "exact" results confirm that both $F(\zeta)$ and $\tilde{F}(\zeta)$ are heavily skewed to the left, all values being above the normal cdf ones, for all sample sizes considered. Although skewness decreases with increasing n, $F(\zeta)$ and $\tilde{F}(\zeta)$ converge quite slowly, especially for "divergent" h, confirming the theory. For very small n, both (4.13) and (4.14) return some values that slightly exceed 1, but the problem disappears as n increases. Otherwise, the agreement between Edgeworth-corrected and exact values in the lower tail leaves something to be desired, especially for (2.13) with "divergent" h, but it improves as n increases, and in the upper tail it is very satisfactory, for both (1.1) and (2.13) and both "divergent" and "bounded" h.

5 Consistency and local power of LSE-based testing

As previously remarked, $\hat{\lambda}$ and $\tilde{\lambda}$ are inconsistent when λ is non-zero. Therefore, if it should be the case that $p\lim_{n\to\infty}\hat{\lambda} < \lambda \quad (> \lambda)$ when $\lambda > 0$ ($\lambda < 0$), it might feared that for some $\lambda \quad p\lim_{n\to\infty}\hat{\lambda} = 0$ as $n \to \infty$, with the same possibility for $\tilde{\lambda}$. Then the standard and corrected tests would be inconsistent. The following theorem shows that under fairly general conditions the direction of inconsistency of $\hat{\lambda}$ and $\tilde{\lambda}$ follows the sign of λ, so that the tests are actually consistent against fixed alternatives. We relax Assumption 1 to:

Assumption 1' *The ϵ_i are independently and identically distributed with zero mean, variance σ^2 and finite fourth moment.*

Theorem 3 *Let Assumptions 1' and 2 (i) hold, let*

$$\underline{\lim}_{n\to\infty} t_{11} > 0 \tag{5.1}$$

with h as defined in Assumption 2 (iii), and let

$$\overline{\lim}_{n\to\infty} \|W\| \leq 1 \tag{5.2}$$

and

$$\lim_{n\to\infty} \frac{tr\left(W^2\right)}{tr\left(W'W\right)} \to \tau > 0. \tag{5.3}$$

If also (a) W has non-negative elements for all sufficiently large n, then for all $\lambda \in [0,1)$,

$$p\lim_{n\to\infty} \widetilde{\lambda} = p\lim_{n\to\infty} \widehat{\lambda} = \lambda(1+\xi(\lambda)), \quad (1-\lambda)^2 \tau \le \xi(\lambda) \le \frac{(1+\lambda)^2}{1-\lambda}, \tag{5.4}$$

or if

(b) W is similar to a symmetric matrix for all sufficiently large n, then for all $\lambda \in (-1,1)$

$$p\lim_{n\to\infty} \widetilde{\lambda} = p\lim_{n\to\infty} \widehat{\lambda} = \lambda(1+\xi(\lambda)), \quad \frac{(1-|\lambda|)^2}{1+|\lambda|} \le \xi(\lambda) \le \frac{(1+|\lambda|)^2}{1-|\lambda|}. \tag{5.5}$$

The proof is in the Appendix. Both sets of conditions on W are clearly satisfied in the case (2.1). An example of a non-symmetric W for which (b) holds arises when an initially symmetric matrix is row-normalized. Condition (5.3) holds automatically if W is symmetric, when $\tau = 1$. Note that we do not necessarily require that $\lim_{n\to\infty} t_{11}$ exists, and that as $|\lambda| \uparrow 1$ the upper bounds for $\xi(\lambda)$ tend to infinity while the lower bounds tend to zero, whereas as $\lambda \to 0$ both bounds under condition (b) tend to 1 while under (a) the upper and lower bounds respectively tend to 1 and $\tau \le 1$. But most significantly, for each $|\lambda| \in (0,1)$ the bounds are finite and positive, so the asymptotic bias $\lambda\xi(\lambda)$ of both $\widehat{\lambda}$ and $\widetilde{\lambda}$ is finite and shares the sign of λ, indeed $\widehat{\lambda}$ and $\widetilde{\lambda}$ tend to exaggerate the spatial correlation. We stress that (a) rules out any negative λ, w_{ij}, in which situation we have been unable to obtain a suitable lower bound under simple conditions, but $\lambda > 0$ seems the case of main practical relevance and negative λ, w_{ij} are covered under (b). Recalling that $q = (n/h)^{1/2} a\hat{\lambda}$, $a = t_{11}^{1/2}(t_{20}/t_{11}+1)^{-1/2}$, it follows from (5.1) and (5.3) and Theorem 3 that when $\lambda > 0$, $q \to_p \infty$ as $n \to \infty$, so for (1.1) $P(q > z_\alpha) \to 1$, $P(q > u_\alpha) \to 1$ and $P(G(q) > z_\alpha) \to 1$ as $n \to \infty$, while for (2.13) $P(\tilde{q} > z_\alpha) \to 1$, $P(\tilde{q} > \tilde{u}_\alpha) \to 1$ and $P(\tilde{G}(\tilde{q}) > z_\alpha) \to 1$ as $n \to \infty$, with similar results when $\lambda < 0$ under condition (b), and obvious implications for two-sided tests.

We now consider behaviour in the presence of local alternatives, namely

$$\lambda = \lambda_n = \delta \left(\frac{h}{n}\right)^{1/2}, \quad \delta \neq 0. \tag{5.6}$$

Theorem 4 *Let Assumptions 1' , 2, (5.2) and (5.3) hold, and let*

$$\omega = \lim_{n\to\infty} t_{11} \tag{5.7}$$

exist and be positive. Then as $n \to \infty$,

$$\frac{tr(WW')}{(tr(W^2 + WW'))^{1/2}}\hat{\lambda} \to_d \mathcal{N}(\omega^{1/2}(1+\tau)^{1/2}\delta, 1), \tag{5.8}$$

$$\frac{tr(WW')}{(tr(W^2 + WW'))^{1/2}}\tilde{\lambda} \to_d \mathcal{N}(\omega^{1/2}(1+\tau)^{1/2}\delta, 1). \tag{5.9}$$

Since $\omega(1+\tau)$ is the asymptotic information, it follows that the limit distributions in (5.8) and (5.9) are the same as those for the Wald statistic based on the MLE of λ and for the LM statistic, so tests based on the normal approximation for $\hat{\lambda}$ and $\widetilde{\lambda}$ are efficient.

6 Final comments

We have developed tests for lack of spatial correlation based on the LSE of the correlation parameter λ in pure SAR and SAR with intercept, that have improved higher-order properties, and compared their finite-sample performance with other tests via Monte Carlo simulations. Though the LSE is inconsistent, we have shown under quite general conditions that LSE-based tests are consistent against fixed alternatives and locally efficient, to add to their computational appeal. Our methods can straightforwardly extended to derive improved LSE-based tests in models involving regressors.

Acknowledgments We thank the editor and three referees for helpful comments. The first author's research was supported by ESRC Grant ES/J007242/1.

Appendix

Proof of Theorem 1

Under H_0, $\hat{\lambda} = \epsilon' W' \epsilon / \epsilon' W' W \epsilon$ and thus $P(\hat{\lambda} \leq x) = P(\varsigma \leq 0)$, where $\varsigma = \epsilon'(C+C')\epsilon/2$, $C = W' - xW'W$ and x is any real number. We proceed much as in, e.g., Phillips (1977). Under Assumption 1, the characteristic function (cf) of ς is

$$E(e^{\frac{it}{2}\epsilon'(C+C')\epsilon}) = \frac{1}{(2\pi)^{n/2}\sigma^n}\int\limits_{\Re^n} e^{\frac{it}{2}\xi'(C+C')\xi}e^{-\frac{\xi'\xi}{2\sigma^2}}d\xi$$

$$= \frac{1}{(2\pi)^{n/2}\sigma^n}\int\limits_{\Re^n} e^{-\frac{1}{2\sigma^2}\xi'^2(C+C'))\xi}d\xi$$

$$= det(I - it\sigma^2(C+C'^{-1/2} = \prod_{j=1}^{n}(1 - it\sigma^2\eta_j)^{-1/2}, \qquad \text{(A.1)}$$

where the η_j are eigenvalues of $C + C'$ and $det(A)$ denotes the determinant of a generic square matrix A. From (A.1) the cumulant generating function (cgf) of ς is

$$\psi(t) = -\frac{1}{2}\sum_{j=1}^{n} ln(1 - it\sigma^2\eta_j) = \frac{1}{2}\sum_{j=1}^{n}\sum_{s=1}^{\infty}\frac{(it\sigma^2\eta_j)^s}{s}$$

$$= \frac{1}{2}\sum_{s=1}^{\infty}\frac{(it\sigma^2)^s}{s}\sum_{j=1}^{n}\eta_j^s = \frac{1}{2}\sum_{s=1}^{\infty}\frac{(it\sigma^2)^s}{s}tr((C+C'^s). \qquad \text{(A.2)}$$

Denoting by κ_s the $s-$th cumulant of ς, from (A.2)

$$\kappa_1 = \sigma^2 tr(C), \qquad \text{(A.3)}$$

$$\kappa_2 = \frac{\sigma^4}{2}tr((C+C'^2), \qquad \text{(A.4)}$$

$$\kappa_s = \frac{\sigma^{2s}s!}{2}\frac{tr((C+C'^s)}{s}, s > 2. \qquad \text{(A.5)}$$

Let $\varsigma^c = (\varsigma - \kappa_1)/\kappa_2^{1/2}$. The cgf of ς^c is

$$\psi^c(t) = -\frac{1}{2}t^2 + \sum_{s=3}^{\infty}\frac{\kappa_s^c(it)^s}{s!}, \qquad \text{(A.6)}$$

where

$$\kappa_s^c = \frac{\kappa_s}{\kappa_2^{s/2}}, \qquad \text{(A.7)}$$

so the cf of ς^c is

$$\begin{aligned}
E(e^{it\varsigma^c}) &= e^{-\frac{1}{2}t^2}\exp\{\sum_{s=3}^{\infty}\frac{\kappa_s^c(it)^s}{s!}\} \\
&= e^{-\frac{1}{2}t^2}\{1+\sum_{s=3}^{\infty}\frac{\kappa_s^c(it)^s}{s!}+\frac{1}{2!}(\sum_{s=3}^{\infty}\frac{\kappa_s^c(it)^s}{s!})^2+\frac{1}{3!}(\sum_{s=3}^{\infty}\frac{\kappa_s^c(it)^s}{s!})^3+.....\} \\
&= e^{-\frac{1}{2}t^2}\{1+\frac{\kappa_3^c(it)^3}{3!}+\frac{\kappa_4^c(it)^4}{4!}+\frac{\kappa_5^c(it)^5}{5!}+\{\frac{\kappa_6^c}{6!}+\frac{(\kappa_3^c)^2}{(3!)^2}\}(it)^6+.....\}.
\end{aligned} \tag{A.8}$$

Thus by Fourier inversion, formally

$$P(\varsigma^c \leq z) = \int_{-\infty}^{z}\phi(z)dz+\frac{\kappa_3^c}{3!}\int_{-\infty}^{z}H_3(z)\phi(z)dz+\frac{\kappa_4^c}{4!}\int_{-\infty}^{z}H_4(z)\phi(z)dz+.... \ . \tag{A.9}$$

Collecting the above results,

$$\begin{aligned}
P(\hat{\lambda} \leq x) = P(\varsigma \leq 0) = P(\varsigma^c\kappa_2^{1/2}+\kappa_1 \leq 0) = P(\varsigma^c \leq -\kappa_1^c) \\
= \Phi(-\kappa_1^c)-\frac{\kappa_3^c}{3!}\Phi^{(3)}(-\kappa_1^c)+\frac{\kappa_4^c}{4!}\Phi^{(4)}(-\kappa_1')+... \ .
\end{aligned} \tag{A.10}$$

From (A.3), (A.4) and (A.7),

$$\kappa_1^c = \frac{tr(C)}{(\frac{1}{2}tr((C+C'^2))^{1/2}}. \tag{A.11}$$

The numerator of κ_1^c is

$$tr(W)-xtr(WW') = -xtr(WW') = -\frac{n}{h}xt_{11}, \tag{A.12}$$

while its denominator is

$$\begin{aligned}
(\frac{1}{2}tr(C+C'^2)^{1/2} = (tr(W^2)+tr(WW'^2W'^2tr((WW'^2))^{1/2}. \\
= \left(\frac{n}{h}\right)^{1/2}\left(t_{20}+t_{11}-4xt_{21}+2x^2t\right)^{1/2}.
\end{aligned} \tag{A.13}$$

Thus

$$\kappa_1^c = \frac{-xt_{11}(n/h)^{1/2}}{(t_{20}+t_{11}-4xt_{21}+2x^2t)^{1/2}} = \frac{-xt_{11}(n/h)^{1/2}}{(t_{20}+t_{11})^{1/2}(1-\frac{4xt_{21}-2x^2t}{(t_{20}+t_{11})})^{1/2}}. \tag{A.14}$$

Choose

$$x = \left(\frac{h}{n}\right)^{1/2} \frac{(t_{20}+t_{11})^{1/2}}{t_{11}}\zeta = (\frac{h}{n})^{1/2} a^{-1}\zeta, \tag{A.15}$$

where a was defined in (2.6). By Taylor expansion

$$\begin{aligned}\kappa_1^c = -\zeta\left(1 - \frac{4xt_{21} - 2x^2t}{(t_{20}+t_{11})}\right)^{-1/2} &= -\zeta - 2\left(\frac{h}{n}\right)^{1/2} \frac{t_{21}}{t_{11}(t_{20}+t_{11})^{1/2}}\zeta^2 \\ + \frac{h}{n}\frac{t}{t_{11}^2}\zeta^3 - 6\frac{h}{n}\left(\frac{t_{21}}{(t_{20}+t_{11})^{1/2}t_{11}}\right)^2\zeta^3 + O\left(\left(\frac{h}{n}\right)^{3/2}\right) \\ = -\zeta - 2\left(\frac{h}{n}\right)^{1/2} b\zeta^2 + \frac{h}{n}d\zeta^3 - 6\frac{h}{n}b^2\zeta^3 + O\left(\left(\frac{h}{n}\right)^{3/2}\right),\end{aligned} \tag{A.16}$$

where b and d were defined in (2.6) and (2.7). Then by Taylor expansion and using

$$(-d/dx)^j\Phi(x) = -H_{j-1}(x)\phi(x), \tag{A.17}$$

we have

$$\begin{aligned}\Phi(-\kappa_1^c) &= \Phi\left(\zeta + 2\left(\frac{h}{n}\right)^{1/2} b\zeta^2 - \frac{h}{n}d\zeta^3 + 6\frac{h}{n}b^2\zeta^3 + O\left(\left(\frac{h}{n}\right)^{3/2}\right)\right) \\ &= \Phi(\zeta) + \left(2\left(\frac{h}{n}\right)^{1/2} b\zeta^2 - \frac{h}{n}d\zeta^3 + 6\frac{h}{n}b^2\zeta^3\right)\phi(\zeta) + 2\frac{h}{n}b^2\zeta^4\Phi^{(2)}(\zeta) + O\left(\left(\frac{h}{n}\right)^{3/2}\right) \\ &= \Phi(\zeta) + 2\left(\frac{h}{n}\right)^{1/2} b\zeta^2\phi(\zeta) + \frac{h}{n}\left(-d\zeta^3 + b^2(6\zeta^3 - 2\zeta^4 H_1(\zeta))\right)\phi(\zeta) + O\left(\left(\frac{h}{n}\right)^{3/2}\right) \\ &= \Phi(\zeta) + 2\left(\frac{h}{n}\right)^{1/2} b\zeta^2\phi(\zeta) + \frac{h}{n}\left(-d\zeta^3 + b^2(6\zeta^3 - 2\zeta^5)\right)\phi(\zeta) + O\left(\left(\frac{h}{n}\right)^{3/2}\right).\end{aligned} \tag{A.18}$$

Similarly,

$$\begin{aligned}\Phi^{(3)}(-\kappa_1^c) &= \Phi^{(3)}(\zeta) + 2\left(\frac{h}{h}\right)^{1/2} b\zeta^2\Phi^{(4)}(\zeta) + O\left(\frac{h}{n}\right) \\ &= \left(H_2(\zeta) - 2\left(\frac{h}{h}\right)^{1/2} b\zeta^2 H_3(\zeta)\right)\phi(\zeta) + O\left(\frac{h}{n}\right).\end{aligned} \tag{A.19}$$

From (A.5), (A.7),

$$\kappa_3^c = \frac{tr((C+C'^3)}{(\frac{1}{2}tr((C+C'^2))^{3/2}}.$$

By standard algebra, for x defined in (A.15),

$$\begin{aligned}\frac{1}{2}tr((C+C'^2) &= \frac{n}{h}\left(t_{20}+t_{11}-4\left(\frac{h}{n}\right)^{1/2}\frac{(t_{20}+t_{11})^{1/2}t_{21}}{t_{11}}\zeta+O\left(\frac{h}{n}\right)\right)\\ &= \frac{n}{h}(t_{20}+t_{11})-4\left(\frac{n}{h}\right)^{1/2}\frac{(t_{20}+t_{11})^{1/2}t_{21}}{t_{11}}\zeta+O(1), \end{aligned} \quad \text{(A.20)}$$

$$\begin{aligned}tr((C+C'^3) &= \frac{n}{h}\left(2t_{30}+6t_{21}-12\left(\frac{h}{n}\right)^{1/2}\frac{(t_{20}+t_{11})^{1/2}(t_{31}+t_{22})}{t_{11}}\zeta+O\left(\frac{h}{n}\right)\right)\\ &= \frac{n}{h}(2t_{30}+6t_{21})-12\left(\frac{n}{h}\right)^{1/2}\frac{(t_{20}+t_{11})^{1/2}(t_{31}+t_{22})}{t_{11}}\zeta+O(1)\end{aligned} \quad \text{(A.21)}$$

and thus

$$\begin{aligned}\kappa_3^c &= \frac{\frac{n}{h}(2t_{30}+6t_{21})-12\left(\frac{n}{h}\right)^{1/2}(t_{20}+t_{11})^{1/2}(t_{31}+t_{22})t_{11}^{-1}\zeta+O(1)}{\left(\frac{n}{h}\right)^{3/2}(t_{20}+t_{11})^{3/2}\left(1-4\left(\frac{h}{n}\right)^{1/2}t_{21}t_{11}^{-1}(t_{20}+t_{11})^{-1/2}\zeta+O\left(\frac{h}{n}\right)\right)^{3/2}}\\ &= \left(\left(\frac{h}{n}\right)^{1/2}\frac{2t_{30}+6t_{21}}{(t_{20}+t_{11})^{3/2}}-12\frac{h}{n}\frac{t_{31}+t_{22}}{t_{11}(t_{20}+t_{11})}\zeta+O\left(\left(\frac{h}{n}\right)^{3/2}\right)\right)\\ &\times\left(1+6\left(\frac{h}{n}\right)^{1/2}\frac{t_{21}}{t_{11}(t_{20}+t_{11})^{1/2}}\zeta+O\left(\frac{h}{n}\right)\right)\\ &= \left(\frac{h}{n}\right)^{1/2}\frac{2t_{30}+6t_{21}}{(t_{20}+t_{11})^{3/2}}-12\frac{h}{n}\frac{t_{31}+t_{22}}{t_{11}(t_{20}+t_{11})}\zeta+\frac{h}{n}\frac{6(2t_{30}+6t_{21})t_{21}}{(t_{20}+t_{11})^2t_{11}}\zeta+O\left(\left(\frac{h}{n}\right)^{3/2}\right)\\ &= \left(\frac{h}{n}\right)^{1/2}c-\frac{h}{n}(e-6bc)\zeta+O\left(\left(\frac{h}{n}\right)^{3/2}\right),\end{aligned} \quad \text{(A.22)}$$

where b, c and e were defined in (2.6) and (2.7).

Similarly,

$$3tr((C+C'^4) = \frac{n}{h}(6t_{40}+24t_{31}+12t+6t_{22})+O\left(\left(\frac{n}{h}\right)^{1/2}\right) \quad \text{(A.23)}$$

and thus

$$\kappa_4^c = \frac{h}{n}\frac{6t_{40}+24t_{31}+12t+6t_{22}}{(t_{20}+t_{11})^2} + O\left(\left(\frac{h}{n}\right)^{3/2}\right) = \frac{h}{n}f + O\left(\left(\frac{h}{n}\right)^{3/2}\right), \tag{A.24}$$

where f was defined in (2.7).

Substituting (A.15), (A.18), (A.19), (A.22) and (A.24) in (A.10) and rearranging using (2.8) and (2.9) completes the proof.

Proof of Theorem 2

Under H_0 and by Assumption 2(i), $\hat{\lambda} = \epsilon' W' P\epsilon / \epsilon' W' PW\epsilon$. Proceeding as before, $P(\tilde{\lambda} \leq x) = P(\varsigma \leq 0)$, which can be written as the right side of (A.10), with $\varsigma = \epsilon'(C + C')\epsilon/2$ and

$$C = W'P(I - xW). \tag{A.25}$$

Derivation of the cumulants κ_j of ς is very similar to that in the proof of Theorem 1, and so is not described in detail. From (A.25), (2.2) and (2.17),

$$\kappa_1 = \sigma^2 tr(C) = -\sigma^2\left(1 + xtr(W'W) - \frac{x}{n}(l'WW'l)\right) = -\sigma^2\left(1 + x\frac{n}{h}t_{11} - xp\right). \tag{A.26}$$

Similarly, since

$$l'^i W'^j l = O(n) \quad \text{for all} \quad i \geq 0, \quad j \geq 0, \tag{A.27}$$

$$\begin{aligned}
\kappa_2 &= \frac{\sigma^4}{2}tr((C + C'^2) \\
&= \sigma^4\left(tr(W^2) + tr(W'W) - 1 - \frac{1}{n}l'W'Wl - 4x(tr(WW'^2(tr((W'^2) + O(1))\right) \\
&= \sigma^4\left(\frac{n}{h}(t_{20} + t_{11}) - 1 - p - 4x\left(\frac{n}{h}t_{21} + O(1)\right) + 2x^2\left(\frac{n}{h}t + O(1)\right)\right).
\end{aligned} \tag{A.28}$$

Proceeding as in the proof of Theorem 1, the first centred cumulant of ς is

$$\kappa_1^c = \frac{-x\frac{n}{h}t_{11} - 1 + xp}{\left(\frac{n}{h}(t_{20}+t_{11})\right)^{1/2}} \left(1 - \frac{1 + p + 4x\left(\frac{n}{h}t_{21} + O(1)\right) - 2x^2\left(\frac{n}{h}t + O(1)\right)}{\frac{n}{h}(t_{20}+t_{11})}\right)^{-1/2}. \tag{A.29}$$

Setting x as in (A.15) and by Taylor expansion,

$$\begin{aligned}
\kappa_1^c &= -\left(\zeta + \frac{(h/n)^{1/2}}{(t_{20}+t_{11})^{1/2}} - \frac{h}{n}\frac{p}{t_{11}}\zeta\right) \\
&\quad \times \left(1 + \left(\frac{h}{n}\right)^{1/2} \frac{2t_{21}}{t_{11}(t_{20}+t_{11})^{1/2}}\zeta + \frac{h}{n}\left(\frac{1}{2(t_{20}+t_{11})} + \frac{1}{2}\frac{p}{t_{20}+t_{11}} - \frac{t}{t_{11}^2}\zeta^2 + \frac{6t_{21}^2}{t_{11}^2(t_{20}+t_{11})}\zeta^2\right)\right) \\
&\quad + O\left(\left(\frac{h}{n}\right)^{3/2}\right) \\
&= -\left(\zeta + \left(\frac{h}{n}\right)^{1/2} g^{1/2} - \frac{h}{n}\frac{p}{t_{11}}\zeta\right)\left(1 + \left(\frac{h}{n}\right)^{1/2} 2b\zeta + \frac{h}{n}\left(\frac{g}{2} + \frac{g}{2}p - d\zeta^2 + 6b^2\zeta^2\right)\right) \\
&\quad + O\left(\left(\frac{h}{n}\right)^{3/2}\right) \\
&= -\zeta - \left(\frac{h}{n}\right)^{1/2}(2b\zeta^2 + g^{1/2}) - \frac{h}{n}\left(\frac{g}{2}\zeta + \frac{g}{2}p\zeta - d\zeta^3 + 6b^2\zeta^3 + 2bg^{1/2}\zeta\right) + O\left(\left(\frac{h}{n}\right)^{3/2}\right),
\end{aligned} \tag{A.30}$$

with b, d, g and p defined in (2.6), (2.7) and (2.17). Similarly, by standard algebra and using (A.27),

$$tr((C + C'^3) = \frac{n}{h}(2t_{30} + 6t_{21}) - 12\left(\frac{n}{h}\right)^{1/2} \frac{(t_{20}+t_{11})^{1/2}(t_{31}+t_{22})}{t_{11}}\zeta + O(1), \tag{A.31}$$

agreeing with the corresponding formula in the proof of Theorem 1, so that the third centred cumulant of ς, κ_3^c, is (A.22), whereas the fourth centred cumulant of ς, κ_4^c, is again (A.24).

Next,

$$\Phi(-\kappa_1^c) = \Phi(\zeta) + \left(\frac{h}{n}\right)^{1/2}(2b\zeta^2 + g^{1/2})\phi(\zeta) + \frac{h}{n}\left(\frac{g}{2}\zeta + \frac{g}{2}p\zeta - d\zeta^3 + 6b^2\zeta^3 + 2bg^{1/2}\zeta\right)\phi(\zeta)$$
$$+ \frac{1}{2}(2b\zeta^2 + g^{1/2})^2\Phi^{(2)}(\zeta) + O\left(\left(\frac{h}{n}\right)^{3/2}\right)$$
$$= \Phi(\zeta) + \left(\frac{h}{n}\right)^{1/2}(2b\zeta + g^{1/2})\phi(\zeta)$$
$$+ \frac{h}{n}\left(\frac{g}{2}\zeta + \frac{g}{2}p\zeta - d\zeta^3 + 6b^2\zeta^3 + 2bg^{1/2}\zeta - \frac{1}{2}(2b\zeta^2 + g^{1/2})^2 H_1(\zeta)\right)\phi(\zeta)$$
$$+ O\left(\left(\frac{h}{n}\right)^{3/2}\right) \tag{A.32}$$

and

$$\Phi^{(3)}(-\kappa_1^c) = \Phi^{(3)}(\zeta) + \left(\frac{h}{h}\right)^{1/2}(2b\zeta^2 + g^{1/2})\Phi^{(4)}(\zeta) + O\left(\frac{h}{n}\right)$$
$$= \left(H_2(\zeta) - \left(\frac{h}{h}\right)^{1/2}(2b\zeta^2 + g^{1/2})H_3(\zeta)\right)\phi(\zeta) + O\left(\frac{h}{n}\right). \tag{A.33}$$

Substituting (A.15), (A.22), (A.24), (A.32) and (A.33) in the right side of (A.10) complete the proof.

Proof of Theorem 3

Define $S(x) = I - xW$. We have

$$\tilde{\lambda} - \lambda = \frac{\epsilon' PWS^{-1}(\lambda)\epsilon}{\epsilon' S^{-1\prime}(\lambda)W'PWS^{-1}(\lambda)\epsilon}$$
$$= \frac{\epsilon' WS^{-1}(\lambda)\epsilon - \epsilon' ll'WS^{-1}(\lambda)\epsilon/n}{\epsilon' S^{-1\prime}(\lambda)W'WS^{-1}(\lambda)\epsilon - \epsilon' S^{-1\prime}(\lambda)W'll'WS^{-1}(\lambda)\epsilon/n}. \tag{A.34}$$

Now

$$E\left|\epsilon' ll'WS^{-1}(\lambda)\epsilon\right| \le \left(E\left(\epsilon' l\right)^2 E\left(l'WS^{-1}(\lambda)\epsilon\right)^2\right)^{1/2} = \sigma^2 n^{1/2}\left(l'WS^{-1}(\lambda)S^{-1\prime}(\lambda)W'l\right)^{1/2}$$
$$\le \sigma^2 n \left\|S^{-1}(\lambda)\right\| \left\|W\right\|. \tag{A.35}$$

From (5.2),

$$\overline{\lim_{n\to\infty}} \left\| S^{-1}(\lambda) \right\| \leq \sum_{j=0}^{\infty} |\lambda|^j \overline{\lim_{n\to\infty}} \|W\|^j \leq (1-|\lambda|)^{-1}, \tag{A.36}$$

and thus (A.35)$= (1-|\lambda|)^{-1} O(n)$. Likewise

$$\begin{aligned} E(\epsilon' S^{-1\prime}(\lambda) W' l l' W S^{-1}(\lambda)\epsilon) =& \sigma^2 l' W S^{-1}(\lambda) S^{-1\prime}(\lambda) W' l \leq \sigma^2 n \left\| S^{-1}(\lambda)\right\|^2 \|W\|^2 \\ =& (1-|\lambda|)^{-2} O(n). \end{aligned} \tag{A.37}$$

Thus

$$\tilde{\lambda} - \lambda = \frac{(h/n)\epsilon' W S^{-1}(\lambda)\epsilon + (1-|\lambda|)^{-1} O_p(h/n)}{(h/n)\epsilon' S^{-1\prime}(\lambda) W' W S^{-1}(\lambda)\epsilon + (1-|\lambda|)^{-2} O_p(h/n)}. \tag{A.38}$$

The subsequent proof will show that the leading terms of both numerator and denominator are of larger order than the remainder terms, and so the latter may be ignored and it suffices to examine

$$\hat{\lambda} - \lambda = \frac{\epsilon' W S^{-1}(\lambda)\epsilon}{\epsilon' S^{-1\prime}(\lambda) W' W S^{-1}(\lambda)\epsilon}. \tag{A.39}$$

Now

$$Var(\epsilon' W S^{-1}(\lambda)\epsilon) = \sigma^4 tr\left(W S^{-1}(\lambda) W S^{-1}(\lambda) + W S^{-1}(\lambda) S^{-1\prime}(\lambda) W'\right) + \kappa \sum_{i=1}^{n} u_i^2,$$

where κ is the 4th cumulant of ϵ_i and u_i is the ith diagonal element of $U = W S^{-1}(\lambda)$. The first term on the right is bounded by $2\sigma^4 tr\left(W S^{-1}(\lambda) S^{-1\prime}(\lambda) W'\right) \leq K \left\|S^{-1}(\lambda)\right\|^2 tr(W'W)$, using the Cauchy inequality and the inequality $tr(A'B'BA) \leq \|B\|^2\ tr(A'A)$. Denoting by w_i' the ith row of W and e_i the $n \times 1$ vector whose ith element is 1 and remaining elements are 0, we have $u_i^2 = \left(w_i' S^{-1}(\lambda) e_i\right)^2 \leq \|w_i\|^2 \left\|S^{-1}(\lambda)\right\|^2$ and so $\sum_{i=1}^n u_i^2 \leq K \left\|S^{-1}(\lambda)\right\|^2 tr(W'W)$ also. Thus (A.40) is bounded by $K \left\|S^{-1}(\lambda)\right\|^2 tr(W'W) = (1-|\lambda|)^{-2} O(n/h)$. Likewise

$$Var\left(\epsilon' S^{-1\prime}(\lambda) W' W S^{-1}(\lambda)\epsilon\right) = 2\sigma^4 tr\left(S^{-1\prime}(\lambda) W' W S^{-1}(\lambda) S^{-1\prime}(\lambda) W' W S^{-1}(\lambda)\right) + \kappa \sum_{i=1}^{n} v_i^2,$$

where v_i is the ith diagonal element of $V = U'U$. Proceeding as before, the first term on the right is bounded by $K \left\|S^{-1}(\lambda)\right\|^4 \|W\|^2 tr(W'W)$. Since $v_i =$

$e_i'Ve_i \leq \|V\|$,

$$\begin{aligned}
\sum_{i=1}^{n} v_i^2 \leq \|V\| \sum_{i=1}^{n} v_i \\
&= \|V\| \sum_{i=1}^{n} e_i' S^{-1\prime}(\lambda) \sum_{j=1}^{n} w_j w_j' S^{-1}(\lambda) e_i \\
&= \|V\| \sum_{j=1}^{n} w_j' S^{-1}(\lambda) \sum_{i=1}^{n} e_i e_i' S^{-1\prime}(\lambda) w_j \\
&= \|V\| \sum_{j=1}^{n} w_j' S^{-1}(\lambda) S^{-1\prime}(\lambda) w_j \\
&\leq \|V\| \left\| S^{-1}(\lambda) \right\|^2 \sum_{j=1}^{n} \|w_j\|^2 \\
&\leq \|W\|^2 \left\| S^{-1}(\lambda) \right\|^4 tr\left(W'W\right). \qquad \text{(A.42)}
\end{aligned}$$

It follows that (A.41) is bounded by $K \left\| S^{-1}(\lambda) \right\|^4 tr\left(W'W\right) = (1 - |\lambda|)^{-4} O(n/h)$.

Thus

$$\hat{\lambda} - \lambda = \frac{tr\left(WS^{-1}(\lambda)\right) + (1 - |\lambda|)^{-1} O_p\left((n/h)^{1/2}\right)}{tr\left(S^{-1\prime}(\lambda)W'WS^{-1}(\lambda)\right) + (1 - |\lambda|)^{-2} O_p\left((n/h)^{1/2}\right)}. \qquad \text{(A.43)}$$

We obtain upper and lower bounds for the leading terms in denominator and numerator. First, as already found,

$$tr\left(S^{-1\prime}(\lambda)W'WS^{-1}(\lambda)\right) \leq (1 - |\lambda|)^{-2} tr\left(W'W\right), \qquad \text{(A.44)}$$

whereas

$$tr\left(S^{-1\prime}(\lambda)W'WS^{-1}(\lambda)\right) \geq \|S(\lambda)\|^{-2} tr\left(W'W\right) \geq (1 + |\lambda|)^{-2} tr\left(W'W\right), \qquad \text{(A.45)}$$

since

$$\|S(\lambda)\| \leq 1 + |\lambda|. \qquad \text{(A.46)}$$

Next note that

$$tr\left(WS^{-1}(\lambda)\right) = tr\left(W\sum_{j=0}^{\infty}\lambda^j W^j\right) = \lambda tr\left(W^2\sum_{j=0}^{\infty}\lambda^j W^j\right) = \lambda tr\left(W^2 S^{-1}(\lambda)\right). \tag{A.47}$$

Suppose Now under condition (a) all elements of $S^{-1}(\lambda)$ are non-negative so $tr\left(W^2S^{-1}(\lambda)\right) \geq 0$, and by the Cauchy inequality

$$\begin{aligned} tr\left(W^2S^{-1}(\lambda)\right) &\leq \left(tr\left(W'W\right) tr\left(WS^{-1}(\lambda)S^{-1\prime}(\lambda)W'\right)\right)^{1/2} \\ &\leq \left\|S^{-1}(\lambda)\right\| tr\left(W'W\right) \leq (1-\lambda)^{-1} tr\left(W'W\right). \end{aligned} \tag{A.48}$$

Under condition (b), W is similar to a symmetric matrix W^*, so $W = Q^{-1}W^*Q$, for some non-singular matrix Q. Thus with $S^*(\lambda) = I_n - \lambda W^*$, we have $S^{-1}(\lambda) = Q^{-1}S^{*-1}(\lambda)Q$ and $tr\left(W^2S^{-1}(\lambda)\right) = tr\left(W^{*2}S^{*-1}(\lambda)\right)$. The Cauchy inequality gives

$$\begin{aligned} tr\left(W^{*2}S^{*-1}(\lambda)\right) &\leq \left(tr\left(W^{*2}S^{*-2}(\lambda)\right) tr\left(W^{*2}\right)\right)^{1/2} \\ &\leq \left\|S^{*-1}(\lambda)\right\| tr\left(W^{*2}\right) = (1-|\lambda|)^{-1} tr\left(W^{2}\right) \end{aligned} \tag{A.49}$$

As for lower bounds, under condition (a) all elements of $S(\lambda) - I_n$ are non-negative, so

$$tr\left(W^2S^{-1}(\lambda)\right) \geq tr\left(W^2\right), \tag{A.50}$$

whereas under condition (b),

$$\begin{aligned} tr\left(W^{*2}S^{*-1}(\lambda)\right) &\geq \left\|S^*(\lambda)\right\|^{-2} tr\left(W^{*2}\right) \geq (1+|\lambda|)^{-1} tr\left(W^{*2}\right) \\ &= (1+|\lambda|)^{-1} tr\left(W^2\right). \end{aligned} \tag{A.51}$$

Thus under both sets of conditions

$$\hat{\lambda} - \lambda = \frac{tr\left(WS^{-1}(\lambda)\right)\left(1 + (1 - |\lambda|)^{-1} O_p\left((n/h)^{1/2}/tr\left(WS^{-1}(\lambda)\right)\right)\right)}{tr\left(S^{-1\prime}(\lambda)W'WS^{-1}(\lambda)\right)\left(1 + (1 - |\lambda|)^{-2} O_p\left((n/h)^{1/2}/tr\left(S^{-1\prime}(\lambda)W'WS^{-1}(\lambda)\right)\right)\right)}$$
$$= \frac{(h/n)tr\left(WS^{-1}(\lambda)\right)\left(1 + (1 - |\lambda|)^{-1} O_p\left((h/n)^{1/2}\right)\right)}{(h/n)tr\left(S^{-1\prime}(\lambda)W'WS^{-1}(\lambda)\right)\left(1 + (1 - |\lambda|)^{-2} O_p\left((h/n)^{1/2}\right)\right)}$$
$$= \frac{tr\left(WS^{-1}(\lambda)\right)}{tr\left(S^{-1\prime}(\lambda)W'WS^{-1}(\lambda)\right)}\left(1 + \frac{(1 - |\lambda|)^{-2} o_p(1)}{1 + (1 - |\lambda|)^{-2} o_p(1)}\right)$$
$$= \frac{tr\left(WS^{-1}(\lambda)\right)}{tr\left(S^{-1\prime}(\lambda)W'WS^{-1}(\lambda)\right)}\left(1 + (1 - |\lambda|)^{-2} o_p(1)\right). \tag{A.52}$$

Further, from the previous calculations, under condition (a)

$$\frac{tr\left(W^2\right)}{(1 - \lambda)^{-2} tr\left(W'W\right)} \leq \frac{tr\left(W^2 S^{-1}(\lambda)\right)}{tr\left(S^{-1\prime}(\lambda)W'WS^{-1}(\lambda)\right)} \leq \frac{(1 - \lambda)^{-1} tr\left(W'W\right)}{(1 + \lambda)^{-2} tr\left(W'W\right)}, \tag{A.53}$$

and under condition (b)

$$\frac{(1 + |\lambda|)^{-1} tr\left(W^2\right)}{(1 - |\lambda|)^{-2} tr\left(W'W\right)} \leq \frac{tr\left(W^2 S^{-1}(\lambda)\right)}{tr\left(S^{-1\prime}(\lambda)W'WS^{-1}(\lambda)\right)} \leq \frac{(1 - |\lambda|)^{-1} tr\left(W'W\right)}{(1 + |\lambda|)^{-2} tr\left(W'W\right)}. \tag{A.54}$$

For all sufficiently large n, the factors $tr\left(W'W\right)$ cancel in the upper bounds in both (A.53) and (A.54), since $\eta(n/h) \leq (n/h)\underline{\lim}_{n\to\infty} t_{11} \leq tr\left(W'W\right) \leq n\left\|W\right\|^2 \leq n$ for some $\eta > 0$, and application of ((5.3) completes the proof.

Proof of Theorem 4

In view of previous calculations we give only the proof of (5.8). We have

$$\frac{tr(WW')}{\left(tr(W^2 + WW')\right)^{1/2}}\hat{\lambda} = \left(\frac{n}{h}\right)^{1/2} a\frac{\epsilon' S^{-1\prime}(\lambda_n)WS^{-1}(\lambda_n)\epsilon}{\epsilon' S^{-1\prime}(\lambda_n)W'^{-1}(\lambda_n)\epsilon}. \tag{A.55}$$

This can be written

$$
\begin{aligned}
&\left(\frac{n}{h}\right)^{1/2} a\frac{\epsilon'(I+\lambda_n W+\lambda_n^2 W^2 S^{-1}(\lambda_n))'W(I+\lambda_n W+\lambda_n^2 W^2 S^{-1}(\lambda_n))\epsilon}{\epsilon'(I+\lambda_n W+\lambda_n^2 W^2 S^{-1}(\lambda_n))'W(I+\lambda_n W'W+\lambda_n^2 W^2 S^{-1}(\lambda_n))\epsilon}\\
=&\left(\frac{n}{h}\right)^{1/2} a\frac{\epsilon' W\epsilon+\lambda_n\epsilon'(WW'^2)\epsilon+O_p(1)}{\epsilon' W'W\epsilon+\lambda_n\epsilon'(W'W'W+W'WW)\epsilon+O_p(1)}\\
=&a\frac{(h/n)^{1/2}\epsilon' W\epsilon+\delta h\epsilon'(WW'^2)\epsilon/n+O_p((h/n)^{1/2})}{h\epsilon' W'W\epsilon/n+\lambda_n h\epsilon'(W'W'W+W'WW)\epsilon/n+O_p(h/n)}\\
=&(\omega^{1/2}(1+\tau)^{-1/2}+o(1))\frac{(h/n)^{1/2}\epsilon'^2\delta\omega(1+\tau)+o_p(1)}{\sigma^2\omega+o_p(1)}\\
=&\omega^{-1/2}(1+\tau)^{-1/2}\left(\frac{h}{n}\right)^{1/2}\frac{\epsilon' W\epsilon}{\sigma^2}+\delta\omega^{1/2}(1+\tau)^{1/2}+o_p(1), \qquad \text{(A.56)}
\end{aligned}
$$

from which (5.8) readily follows.

References

- Baltagi, B.H. & Z. Yang (2013) Standardized LM tests for spatial error dependence in linear or panel regressions. *Econometrics Journal* **16**, 103-34.

- Bao, Y. (2013) Finite-sample bias of the QMLE in spatial autoregressive models. *Econometric Theory* **29**, 68-88.

- Bao, Y. & A. Ullah (2007) Finite sample properties of maximum likelihood estimator in spatial models. *Journal of Econometrics* **137**, 396-413.

- Case, A.C. (1991) Spatial patterns in household demand. *Econometrica* **59**, 953-65.

- Cliff, A. & J.K. Ord (1971) Evaluating the percentage points of a spatial autocorrelation coefficient. *Geograpical Analysis **3,*** 51-62.

- Cliff, A. & J.K. Ord (1981) *Spatial Processes: Models & Applications.* Pion.

- García-Soidán, P. (1996) Asymptotic expansions for statistics computed from spatial data. *Test* **5**, 61-76.

- Götze, F. & C. Hipp (1983) Asymptotic expansions for sums of weakly dependent random vectors. *Z. Wahrsch. Verw. Gebeite* **64**, 211-39.

- Hall, P. (1992) On the removal of skewness by transformation. *Journal of the Royal Statistical Society Series B* **54**, 221-28.

- Imhof, J.P. (1961) Computing the distribution of quadratic forms in normal variables. *Biometrika* **48**, 266-83.

- Jin, F. & L.F. Lee (2012) On the bootstrap for spatial econometric models. Preprint.

- Kelejian, H.H. & I.R. Prucha (1998) A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics* **17**, 99-121.

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

- Lee, L.F. (2002) Consistency and efficiency of least squares estimation for mixed regressive, spatial autoregressive models. *Econometric Theory* **18**, 252-77.

- Lee, L.F. (2004) Asymptotic distribution of quasi-maximum likelihood estimates for spatial autoregressive models. *Econometrica* **72**, 1899-1925.

- Lee, L.F., & J. Yu (2012). The $C(\alpha)$ -type gradient test for spatial dependence in spatial autoregressive models. *Letters in Spatial and Resource Sciences* **5**, 119-135.

- Li, H., C.A. Calder & N. Cressie (2007). Beyond Moran's I: testing for spatial dependence based on the SAR model. *Geographical Analysis* **39**, 357-375.

- Lu, Z.H. & M.L.King (2002) Improving the numerical techniques for computing the accumulated distribution of a quadratic form in normal variables. *Econometric Reviews* **21**, 149-65.

- Martellosio, F., 2012, Testing for spatial autocorrelation: The regressors that make the power disappear, Econometric Reviews, 31, 215-240.

- Moran, P.A.P. (1950) Notes on continuous stochastic phenomena. *Biometrika* **9,** 17-23.

- Paparoditis, E.& D.N. Politis (2005) Bootstrap hypothesis testing in regression models. *Statistics & Probability Letters* **74**, 356-65.

- Phillips, P.C.B. (1977) Approximations to some finite sample distributions associated with a first-order stochastic difference equation. *Econometrica* **45**, 463-85.

- Robinson, P.M. (2008) Correlation testing in time series, spatial and cross-sectional data. *Journal of Econometrics* **147**, 5-16.

- Robinson, P.M. & F. Rossi (2013) Improved Lagrange multiplier tests for spatial autoregressions. Forthcoming in *Econometrics Journal*.

- Singh, K. (1981) On the asymptotic accuracy of Efron's bootstrap. *Annals of Statistics* **9**, 1187-95.

- Terui, N. & M. Kikuchi (1994) The size-adjusted critical region of Moran's I test statistics for spatial autocorrelation and its application to geographical areas. *Geographical Analysis* **26**, 213-27.

- Yanagihara, H. & K. Yuan (2005) Four improved statistics for contrasting means by correcting skewness and kurtosis. *British Journal of Mathematical and Statistical Psychology* **58**, 209-37.

- Yang, Z. (2013) LM tests of spatial dependence based on bootstrap critical values. SMU Economics & Statistics Working Papers Series.