Improved Lagrange Multiplier Tests in Spatial Autoregressions

Peter M. Robinson[†] and Francesca Rossi[‡]
[†]*Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, UK*
E-mail: p.m.robinson@lse.ac.uk
[‡]*Department of Economics, University of Southampton, Southampton SO17 1BJ, UK*
E-mail: F.Rossi@soton.ac.uk

October 16, 2013

Summary For testing lack of correlation against spatial autoregressive alternatives, Lagrange multiplier tests enjoy their usual computational advantages, but the (χ^2) first-order asymptotic approximation to critical values can be poor in small samples. We develop refined tests for lack of spatial error correlation in regressions, based on Edgeworth expansion. In Monte Carlo simulations these tests, and bootstrap ones, generally significantly outperform χ^2- based tests.

JEL classification: *C29*

Keywords*:* Spatial autocorrelation; Lagrange multiplier test; Edgeworth expansion; bootstrap; finite-sample corrections.

1 INTRODUCTION

The spatial autoregressive (SAR) model is a parsimonious tool for describing spatial correlation, conveniently depending only on economic distances rather than geographical locations, which may be unknown or irrelevant. It thus provides a convenient, widely-usable class of alternatives in testing the null hypothesis of spatial uncorrelatedness which, if true, considerably simplifies statistical inference. A linear regression with SAR disturbances is given by

$$y = X\beta + u, \quad u = \lambda W u + \epsilon, \tag{1.1}$$

where y =(y_i) is a $n \times 1$ vector of observations, X =(x_{ij}) is an $n \times k$ matrix of non-stochastic regressors, β is a $k \times 1$ vector of unknown parameters, $\epsilon = (\epsilon_i)$ is a $n \times 1$ vector of unobservable, mutually independent, random variables, with zero mean and unknown variance σ^2, λ is an unknown scalar, and $W = (w_{ij})$ is a given $n \times n$ "weight" matrix, such that w_{ii}= 0, $1 \leq i \leq n$, and typically satisfying normalization restrictions (which aid identification of λ). A special case of (1.1) is "pure" SAR, or SAR for y, when $\beta = 0$ a priori,

$$y = \lambda W y + \epsilon, \tag{1.2}$$

and SAR for y with constant mean, when $k = 1$ and $X = l$, the $n \times 1$ vector of 1's, i.e.

$$y - \beta l = \lambda W (y - \beta l) + \epsilon, \tag{1.3}$$

When W is row normalized such that $Wl = l$, (1.3) becomes the intercept model

$$y = \alpha l + \lambda W y + \epsilon, \tag{1.4}$$

where $\alpha = (1 - \lambda)\beta$.

When $\lambda \neq 0$, (1.1) implies the y_i are spatially correlated, but under the null hypothesis

$$H_0 : \lambda = 0 \tag{1.5}$$

they are mutually independent. Various tests of (1.5) have been discussed in the literature (see e.g. Moran (1950), Cliff and Ord (1972), Burridge (1980), Kelejian and Robinson (1992), Pinkse (2004)). For example Wald and likelihood-ratio (LR) tests have been developed assuming that the ϵ_i are normally distributed (e.g. Ord (1975)). However these involve the maximum likelihood (ML) estimate of λ, β and σ^2, which isnot defined in closed form, and the likelihood need not necessarily be unimodal. Though Lagrange multiplier (LM) tests, following Moran (1950), are not guaranteed consistent against all violations of (1.5), and can have low power near inconsistent alternatives, they share the optimal local efficiency properties of Wald and likelihood ratio tests while being computationally simpler, involving closed form estimates of β and σ^2. Anselin (2001) surveyed LM testing in SAR models.

Under (1.5) and regularity conditions, LM, Wald and LR statistics against the two sided alternative

$$H_1 : \lambda \neq 0, \tag{1.6}$$

each have a null limiting χ_1^2 distribution as $n \to \infty$, and provide consistent tests. Frequently, however, spatial economic data sets are not very large, and the χ^2 approximation may be inaccurate. This is of particular concern in the SAR setting where convergence to the limit distribution can be slower than the classical parametric rate (as found for the ML estimate in SAR models by Lee (2004)). Table 1 reports simulated sizes of Wald, LR and LM tests of (1.5) for SAR y, (1.2) with $\epsilon_i \sim \mathcal{N}(0,1)$ and 1000 replications, and W follows the Case (1991) specification

$$W = I_r \otimes B_m, \quad B_m = \frac{1}{(m-1)}(l_m l_m' - I_m), \tag{1.7}$$

where I_s denotes the $s \times s$ identity matrix and l_m the $m \times 1$ vector of ones, so $n = mr$; in (1.7), r might represent the number of districts and m the number of households per district, so households are neighbours if and only if they belong to the same district, and neighbours are equally weighted. The four (m, r) combinations in Table 1, corresponding to $n = 40, 96, 198, 392,$ are designed to reflect an asymptotic regime where convergence is slower than the parametric rate, as discussed subsequently. The empirical sizes are to be compared with the nominal 5%, so the χ^2 approximation is not very good, with Wald and LR being over-sized and LM under-sized, and Wald and LM exhibiting little improvement with increasing n, and LR none. Thus, the issue of constructing tests that enjoy good size properties in modest samples seems worth pursuing.

(Table 1 about here)

In this paper we start from the LM statistic because of its computational advantages and local efficiency, noting also that its signed square root is locally best invariant (King and Hillier (1985)). *Ad hoc* finite sample corrections for LM tests have already been derived in the spatial econometrics literature. Robinson (2008) considers a wide class of residual-based, asymptotically χ^2 statistics which include LM ones for testing (1.5) in SAR models as special cases, and suggests transformed statistics which are still asymptotically χ^2, but have exactly the mean and variance of a χ^2 variate and are therefore expected to have improved finite sample properties. Baltagi and Yang (2012), in line with Koenker (1981), derive a standardised version of the square root of the LM statistic for testing (1.5) in a broad class of SAR-type models, which brings the mean exactly to zero and the variance closer to that of the normal limiting variate. Our main contribution is to develop tests based on the Edgeworth expansion of the distribution function of the LM statistic. We focus on tests against (1.6), but results for one-sided alternatives are simple corollaries. Our Edgeworth-corrected tests are also compared in Monte Carlo simulations with bootstrap-based tests, which are expected to achieve a similar refinement (see e.g. Singh (1981), Hall (1992)). Despite the advantages of bootstrap-based tests, we believe that our analytical approach is worthwhile as it sheds light on the magnitude of correction terms and offers insight on the adequacy of the standard χ^2 approximation for different choices of W, while our refined test statistics are still relatively simple and require no further nuisance parameter estimates, and perform comparably to bootstrap ones in small and moderately-sized Monte Carlo samples.

The derivation of the Edgeworth expansion for the distribution of LM under (1.5) and corrected tests are the focus of the following section, Theorem proofs being left to an Appendix. In Robinson and Rossi (2012) (hereafter RR), Edgeworth-corrected tests of (1.5) in (1.2) and (1.4) are developed, based on the least squares estimate of λ. While this estimate converges in probability to zero under (1.5), it is inconsistent, not converging in probability to λ when $\lambda \neq 0$. In Section 3 we derive the finite sample corrections of Robinson (2008) in the SAR case, so as to compare performance with Edgeworth-corrected tests. Some results on local power are presented in Section 4. A Monte Carlo comparison of the various tests is reported in Section 5. Section 6 contains final comments.

2 EDGEWORTH EXPANSION AND CORRECTED TESTS

The LM statistic for testing (1.5) in (1.1) against (1.6) is

$$LM = T^2, \quad T = \frac{n}{\sqrt{tr(W^2 + WW')}} \frac{y'PWPy}{y'Py}, \tag{2.1}$$

where $P = I - X(X'X)^{-1}X'$, $I = I_n$; in (1.2) $P = I$ and in (1.3) $P = I - l(l'l)^{-1}l'$. The statistic LM was derived by Burridge (1980) who noted that it is equivalent to that of Cliff and Ord (1972), which in turn is related to a statistic of Moran (1950); for extensions to more general models, see also Anselin (1988, 2001), Baltagi and Li (2004), Pinkse (2004). As noted by Burridge (1980), (2.1) is also the LM statistic for testing (1.5) against the spatial moving average model $u = \epsilon + \lambda W \epsilon$ (a corresponding equivalence to that found with time series models).

The derivation of (2.1) is based on a Gaussian likelihood but as usual its first order limit distribution obtains more generally. Under suitable conditions we have as $n \to \infty$

$$P(LM \leq \eta) = \Psi(\eta) + o(1) \tag{2.2}$$

for any $\eta > 0$, where Ψ denotes the distribution function (df) of a χ_1^2 random variable. Thus (1.5) is rejected in favour of (1.6) if LM exceeds the appropriate percentile of the χ_1^2 distribution. We can likewise test (1.5) against a one-sided alternative, $\lambda > 0 \;\; (< 0)$, by comparing T $(-T)$ with the appropriate $\mathcal{N}(0, 1)$ upper (lower) percentile. The present paper mainly focusses on two-sided tests.

We omit mild sufficient conditions for (2.2), because we wish to consider statistics with better finite-sample properties and we only justify these under the restrictive

Assumption 1 *The ϵ_i are independent $\mathcal{N}\left(0, \sigma^2\right)$ random variables.*

The normality assumption is common in higher-order asymptotic theory since Edgeworth expansions and resulting test statistics are otherwise complicated by the presence of cumulants of ϵ_i.

For a real matrix $A = (a_{ij})$, 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. $||A||_\infty = \max_i \sum_j |a_{ij}|$, where i and j vary respectively across all rows and columns of A). We introduce:

Assumption 2

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

(ii) As $n \to \infty$, $||W||_\infty + ||W'||_\infty = O(1)$.

(iii) As $n \to \infty$, $w_{ij} = O(1/h)$, *uniformly in i, j, where $h = h_n$ is bounded away from zero for all n and $h/n \to 0$ as $n \to \infty$.*

If W is row normalized such that $Wl = l$, with $w_{ij} = w_{ji} \geq 0$, all i, j, (as in (1.7)), part (ii) is automatically satisfied. The sequence h defined in (iii) can be bounded or divergent, and this distinction affects the rate of convergence to the null distribution, the order of the leading Edgeworth correction term being

h/n. For W given by (1.7), $h \sim m$, explaining our remark that the (m, r) used in Table 1, where m increases, slowly, with n, correspond to slow convergence.

In addition, we impose a standard boundedness and lack-of-multicollinearity condition on X. Throughout, K denotes a finite generic constant. We introduce

Assumption 3 *Uniformly in* i, j, n, $|x_{ij}| \leq K$, *and as* $n \to \infty$, $||\left(X'X/n\right)^{-1}||^{-1} = O\left(1\right)$.

For notational convenience define

$$a = \frac{h}{n}tr(W'W + W^2), \quad b = \frac{h}{n}tr((W + W')^3), \quad c = \frac{h}{n}tr((W + W')^4), \tag{2.3}$$

$$d = tr(X'(W + W')^2X(X'X)^{-1}), \quad e = tr((X'X)^{-1}X'WX), \tag{2.4}$$

$$f = tr(X'(W + W')X(X'X)^{-1}X'(W' + W)X(X'X)^{-1})/2. \tag{2.5}$$

To ensure that leading terms appearing in the theorem below are well defined, we introduce:

Assumption 4

$$\varliminf_{n\to\infty} a > 0. \tag{2.6}$$

Under Assumption 2, a, b and c in (2.3) are $O(1)$, since $tr(WA) = O(n/h)$ for any real A such that $||A||_\infty = O(1)$. Assumption 4 ensures that (the non-negative) a is positive in the limit. Also, under Assumptions 2 and 3, d, e and f are $O(1)$. Now define

$$\psi(x) = \frac{1}{\sqrt{2\pi}}x^{-1/2}e^{-x/2}, \quad x > 0, \tag{2.7}$$

$$v_1 = \left(\frac{3}{a^2}\left(\frac{c}{4} - \frac{eb}{3}\right) - \frac{e^2 + f - d}{a}\right), \quad v_2 = \frac{1}{a^2}\left(\frac{c}{4} - \frac{eb}{3}\right), \tag{2.8}$$

$$\omega_1(\eta) = v_1\eta - v_2\eta^2, \tag{2.9}$$

$$\omega_2(\eta) = h\omega_1(\eta) - 2(k + 2)\eta + 2\eta^2. \tag{2.10}$$

Both $\omega_1(.)$ and $\omega_2(.)$ are generally non-homogeneous quadratic functions of η with known coefficients. The proof of the following theorem 1 is in the Appendix.

Theorem 1 *Let (1.1) and Assumptions 1-4 hold. Under* H_0 *(1.5), for any real*

$\eta > 0$, the df of LM in (2.1) admits the formal Edgeworth expansion

$$P(LM \leq \eta) = \Psi(\eta) + \frac{h}{n}\omega_1(\eta)\psi(\eta) + o\left(\frac{h}{n}\right), \tag{2.11}$$

in case $h \to \infty$ as $n \to \infty$, and

$$P(LM \leq \eta) = \Psi(\eta) + \frac{1}{n}\omega_2(\eta)\psi(\eta) + o\left(\frac{1}{n}\right), \tag{2.12}$$

in case $h = O(1)$ as $n \to \infty$, and

$$\omega_1(\eta) = O(1), \quad \omega_2(\eta) = O(1), \tag{2.13}$$

as $n \to \infty$.

Since (2.11) and (2.12) entail better approximations than (2.2) and depend on known quantities, they can be used directly in approximating the df of LM. The two outcomes in Theorem 1 create a dilemma for the practitioner because it cannot be determined for finite n whether to treat h as "divergent" or "bounded". However, (2.12) is justified also when h is divergent because the extra term in the expansion, $-2((k+2)\eta - \eta^2)/n$, is $o(h/n)$. We retain both (2.11) and (2.12) to stress the possible dependence of our expansion on both n and h, which is peculiar in SAR models, and the slow convergence of LM in case h is divergent.

Theorem 1 holds for the pure SAR model (1.2) on setting $d = e = f = k = 0$ in (2.11) and (2.12). In (1.3) d, e and f can be likewise simplified, in particular, when $Wl = l$, $d = 2(1 + l'WW'l/n)$, $e = 1$ and $f = 2$.

To derive corrected tests, define w_α such that $P(LM \leq w_\alpha) = 1 - \alpha$, so a test that rejects (1.5) when $LM > w_\alpha$ has exact size α. Let $\Phi(z_\alpha) = 1 - \alpha$, where Φ denotes the standard normal df. From (2.2), a test based on (2.1) that rejects H_0 in (1.5) against (1.6) when

$$LM > z^2_{\alpha/2} \tag{2.14}$$

has approximate size α. Theorem 1 can be used to derive approximations of w_α that are more accurate than $z^2_{\alpha/2}$ (cf Cordeiro and Ferrari (1991), for example). For h divergent and bounded define, respectively,

$$s_\alpha = z^2_{\alpha/2} - \frac{h}{n}\omega_1(z^2_{\alpha/2}), \tag{2.15}$$

and

$$s_\alpha = z^2_{\alpha/2} - \frac{1}{n}\omega_2(z^2_{\alpha/2}). \tag{2.16}$$

From Theorem 1, we obtain:

Corollary 1 *Let (1.1) and Assumptions 1-4 hold. Under H_0 (1.5),*

$$w_\alpha = z^2_{\alpha/2} + O\left(\frac{h}{n}\right), \tag{2.17}$$

$$w_\alpha = s_\alpha + o\left(\frac{h}{n}\right), \tag{2.18}$$

as $n \to \infty$, with s_α defined in (2.15)/(2.16) in case h is divergent/bounded.

The proof of Corollary 1 is in the Appendix. When h is bounded, the remainders in (2.17) and (2.18) are $O(1/n) = O(h/n)$ and $o(1/n) = o(h/n)$, respectively. The use of (2.18) in (2.18) is justified also when h diverges, since the extra terms in ω_2 are $o(h/n)$. From Corollary 1, we conclude that a test that rejects H_0 in (1.5) against (1.6) when

$$LM > s_\alpha \tag{2.19}$$

has size which is closer to α than (2.14).

As an alternative to correcting critical values, we can apply Theorem 1 to construct a monotonic transformation of LM whose distribution better approximates χ^2_1 than that of LM itself (see e.g. Kakizawa (1996)).

Corollary 2 *Let (1.1) and Assumptions 1-4 hold. Under H_0 (1.5),*

$$P(v(LM) > z^2_{\alpha/2}) = \alpha + o\left(\frac{h}{n}\right), \tag{2.20}$$

where

$$v(x) = x + \frac{h}{n}\omega_1(x) + \left(\frac{h}{n}\right)^2\left(\frac{1}{4}v_1^2 x + \frac{1}{3}v_2^2 x^3 - \frac{1}{2}v_1 v_2 x^2\right), \tag{2.21}$$

when $h \to \infty$ as $n \to \infty$, and

$$\begin{aligned} v(x) =& x + \frac{1}{n}\omega_2(x) + \frac{1}{n^2}(\frac{1}{4}(hv_1 - 2(k+2))^2 x + \frac{1}{3}(2 - v_2 h)^2 x^3 \\ &+ \frac{1}{2}(hv_1 - 2(k+2))(2 - v_2 h)x^2), \end{aligned} \tag{2.22}$$

when $h = O(1)$ *as* $n \to \infty$.

The remainder in (2.20) is $o(1/n) = o(h/n)$ when h is bounded. From (2.20) we deduce that a test that rejects H_0 when

$$v(LM) > z^2_{\alpha/2} \tag{2.23}$$

is more accurate than (2.14).

3 MOMENTS-BASED CORRECTION

Robinson (2008) proposed both mean-adjusted and mean-and-variance-adjusted variants of (2.1), which might be expected to have better finite sample properties than (2.1), while still being asymptotically χ^2_1. Since mean adjusting alone might, for smallish n, increase variance, offsetting the gain in accuracy from centering, we focus on the mean-and-variance correction. Such corrected statistics are theoretically convenient since under (1.5), (2.1) depends on the ratio $\epsilon' PWP\epsilon/\epsilon' P\epsilon$, which is independent of its denominator, so its moments can be explicitly calculated (Pitman (1937)).

The mean-and-variance-adjusted statistic in Robinson (2008) starts from

$$\left(\frac{2}{Var(LM)}\right)^{1/2} (LM - E(LM)) + 1, \tag{3.1}$$

then replacing $E(LM)$ and $Var(LM)$ by approximations. Under Assumptions 1-4 and (1.5),

$$E(LM) = 1 + \frac{h}{na}(e^2 + f - d) + o\left(\frac{h}{n}\right), \tag{3.2}$$

when $h \to \infty$ as $n \to \infty$, and

$$E(LM) = 1 + \frac{h}{na}(e^2 + f - d) - \frac{2(1-k)}{n} + o\left(\frac{1}{n}\right), \tag{3.3}$$

when $h = O(1)$ as $n \to \infty$. By formulae for moments of normal quadratic forms (see e.g. Ghazal (1996)),

$$Var(LM) = 2 + \frac{h}{na}\left(4(e^2 + f - d) + \frac{3c - be}{a}\right) + o\left(\frac{h}{n}\right), \tag{3.4}$$

when h is divergent, and

$$Var(LM) = 2 + \frac{h}{na}\left(4(e^2 + f - d) + \frac{3c - be}{a}\right) - \frac{8(4-k)}{n} + o\left(\frac{1}{n}\right), \quad (3.5)$$

when h is bounded. Thus

$$(3.1) = LM_1 + o_p\left(\frac{h}{n}\right) \quad (3.6)$$

when $h \to \infty$, and

$$(3.1) = LM_2 + o_p\left(\frac{1}{n}\right) \quad (3.7)$$

when $h = O(1)$, where

$$LM_1 = LM - \frac{h}{na}\left((e^2 + f - d)LM + \frac{3c - be}{4a}(LM - 1)\right), \quad (3.8)$$

$$LM_2 = LM - \frac{h}{na}\left((e^2 + f - d)LM + \frac{3c - be}{4a}(LM - 1)\right) + \frac{1}{n}(-6 + 2(4-k)LM). \quad (3.9)$$

By construction, LM_1 and LM_2 have mean and variance which are closer to those of a χ_1^2 random variable than LM, so we expect the test that rejects H_0 when

$$LM_i > z^2_{\alpha/2}, \quad (3.10)$$

where $i = 1$ for h divergent and $i = 2$ for h bounded, will have size closer to α than (2.14). Though LM_1 is computationally simpler, LM_2 is valid also when h is divergent, since $(-6 + 2(4-k)LM)/n$ is $o_p(h/n)$. The finite sample performance of (3.10) is compared to (2.19) and (2.23) in Section 6.

4 ANALYSIS OF LOCAL POWER

We now focus on testing (1.5) in (1.1) against the local alternatives

$$H_1: \quad \lambda_n = \left(\frac{h}{n}\right)^{1/2}\delta, \quad \delta \neq 0. \quad (4.1)$$

It follows from (1.1) that

$$y = X\beta + S^{-1}(\lambda_n)\epsilon, \quad (4.2)$$

where $S(x) = I - xW$, because for n large enough $|\lambda_n| < 1$ and existence of $S^{-1}(\lambda_n)$ is guaranteed by Assumption 2. For $Z \sim \mathcal{N}(0,1)$, denote by $\Psi(x;\nu)$ the df of $(Z+\nu)^2$, the non-central χ_1^2 random variable with noncentrality parameter ν, its probability density function (pdf) being

$$\psi(x;\nu) = \frac{1}{\sqrt{2\pi}} x^{-1/2} \cosh\left(\nu x^{1/2}\right) \exp\left(-\left(x+\nu^2\right)/2\right), \ x > 0. \tag{4.3}$$

Define also

$$\tau(x;\nu) = \sqrt{\frac{2}{\pi}} \sinh\left(\nu x^{1/2}\right) \exp\left(-\frac{1}{2}\left(x+\nu^2\right)\right), \ x > 0, \tag{4.4}$$

$$p = \frac{h}{n} tr(W^2 W'), \tag{4.5}$$

Theorem 2 *Let (1.1) and Assumptions 1-4 hold. Under H_1 (4.1), for any real $\eta > 0$, the df of LM in (2.1) admits the formal Edgeworth expansion*

$$\begin{aligned} P(LM \le \eta) =& \Psi\left(\eta; a^{1/2}\delta\right) + \left(\frac{h}{n}\right)^{1/2} \left(a^{-1/2}\left(e+\delta^2 p\right) - \frac{b\left(a\delta^2+1\right)}{6a^{3/2}}\right) \tau\left(\eta; a^{1/2}\delta\right) \\ & - \left(\frac{h}{n}\right)^{1/2} \frac{b\delta}{2a} \eta \psi\left(\eta; a^{1/2}\delta\right) - \left(\frac{h}{n}\right)^{1/2} \frac{b}{6a^{3/2}} \eta \tau\left(\eta; a^{1/2}\delta\right) + o\left(\left(\frac{h}{n}\right)^{1/2}\right), \end{aligned} \tag{4.6}$$

where

$$a^{-1/2}\left(e+\delta^2 p\right) - \frac{b\left(a\delta^2+1\right)}{6a^{3/2}} = O(1), \quad \frac{b\delta}{2a} = O(1), \quad \frac{b}{6a^{3/2}} = O(1) \tag{4.7}$$

as $n \to \infty$.

The first-order asymptotic approximation to the df of LM under H_1 (4.1) has error $O((h/n)^{1/2})$. Terms of higher order could be derived at expense of considerable algebraical complication.

Theorem 2 can be used to derive a more accurate approximation for the local power of the LM test of H_0 against (4.1). Define the power function $\Pi(x) = P(LM > x) = 1 - P(LM \le x)$. From Theorem 2 the test in (2.14) has local power

$$\Pi(z^2_{\alpha/2}) = 1 - \Psi\left(z^2_{\alpha/2}; a^{1/2}\delta\right) - \left(\frac{h}{n}\right)^{1/2}\left(a^{-1/2}\left(e + \delta^2 p\right) - \frac{b\left(a\delta^2+1\right)}{6a^{3/2}}\right)\tau\left(z^2_{\alpha/2}; a^{1/2}\delta\right)$$
$$+\left(\frac{h}{n}\right)^{1/2}\frac{b\delta}{2a}z^2_{\alpha/2}\psi\left(z^2_{\alpha/2}; a^{1/2}\delta\right) + \left(\frac{h}{n}\right)^{1/2}\frac{b}{6a^{3/2}}z^2_{\alpha/2}\tau\left(z^2_{\alpha/2}; a^{1/2}\delta\right) + o\left(\left(\frac{h}{n}\right)^{1/2}\right). \tag{4.8}$$

Even the signs of the correction terms can vary with W, but the terms can be numerically evaluated for any given W. It is therefore possible to establish whether the actual local power of (2.14) is likely to be higher or lower than that of (2.14). It is worth stressing that (4.8) holds also in case of tests (2.19), (2.23) and (3.10) since the extra terms implied by the size-corrections would be of order $o((h/n)^{1/2})$. Hence, tests (2.19), (2.23) and (3.10) have sizes which are closer to α than (2.14), which has local power as in (4.8). The paper is concerned with refinements of the LM test and a comparison between its higher-order power with other existing tests of (1.5) is beyond our scope, but Theorem 2 can be useful for further studies on higher-order efficiency of tests of H_0 (1.5) in SAR models, along the lines of e.g. Peers (1971), Taniguchi (1991) or Rao and Mukerjee (1994).

5 BOOTSTRAP CORRECTION AND SIMULATIONS

Monte Carlo simulations to investigate finite sample performance of the tests developed above, and bootstrap tests, were carried out. The Monte Carlo design, and initial bootstrap specification, correspond to those in RR, except that they focussed only on the models (1.2) and (1.3), which have no varying regressors. Our bootstrap test against (1.6) was obtained by computing the independent bootstrap null statistics

$$LM^*_j = (nh/a)(u^{*'}_j PWPu^*_j / u^{*'}_j Pu^*_j)^2, \quad j = 1, ..., 199, \tag{5.1}$$

each u^*_j being a vector of independent $\mathcal{N}(0, y'Py/n)$ variables. For $\alpha = 0.05$, denote by w^*_α the largest value solving $\sum_{j=1}^{199} 1(LM^* \leq w^*_\alpha)/199 \leq 1-\alpha$, $1(.)$ denoting the indicator function. We reject H_0 (1.5) against (1.6) when

$$LM > w^*_\alpha. \tag{5.2}$$

We choose W as in (1.7), whence $h = m-1$, W is symmetric, satisfies $Wl = l$ and has non-negative elements. Since the tests derived in the previous sections can vary depending on whether h is divergent or bounded, we reflect both cases in our choices of (m, r). We choose $(m, r) = (8, 5)$, $(12, 8)$, $(18, 11)$ and $(28, 14)$ (as in Table 1, and corresponding to $n = 40, 96, 198, 392$) to represent "divergent" h, and $(m, r) = (5, 8)$, $(5, 20)$, $(5, 40)$ and $(5, 80)$ (which correspond to $n = 40, 100, 200, 400$) to represent "bounded" h.

As in Table 1 the ϵ_i were generated as $\mathcal{N}(0, 1)$, and results are based on 1000 replications. In the Tables we denote by "chi square", "Edgeworth", "transformation", "mean-variance correction" and "bootstrap" the empirical sizes of tests (2.14), (2.19), (2.23), (3.10) and (5.2), respectively; in the text we use the respective abbreviations C, E, T, MV and B. Tables 2-7 report empirical sizes of the tests for models (1.1), (1.2) and (1.3).

(Tables 2 and 3 about here)

Tables 2 and 3 concern the regression model with SAR disturbances (1.1), where $k = 3$, with X having first column l, and elements of the other two columns generated independently and uniformly $[0, 1]$, when h (and thus m in (1.7)) is "divergent" and "bounded", respectively. The standard test C is considerably under-sized in both cases, and the overall pattern of the results is consistent with the results in Theorem 1, where the df of LM converges at rate n when h is bounded and at the slower n/h when h is divergent. Indeed, from the first row of Table 2, as n increases from $n = 40$ to $n = 392$, the deviation between empirical and nominal sizes only decreases by 47%, while from the first row of Table 3 such deviation decreases by 85% when n increases from $n = 40$ to $n = 400$. The Edgeworth-corrected tests E and T seem to perform very well in both cases, offering an average (across sample sizes considered) respective improvement over C of 52% and 54% when h is "divergent", and of 52% and 50% when h is "bounded". The MV test is very under-sized, the discrepancy between actual and nominal values decreasing by only 2% and 18% for "divergent" and "bounded" h, respectively, compared to C. The average improvement offered by B is 71% when h is "divergent, and 50% when h is "bounded" and its performance is comparable (or even superior, in case h is "divergent") to E and T. Overall, E, T and B perform very well.

(Tables 4 and 5 about here)

Tables 4 and 5 concern pure SAR (1.2) for "divergent" and "bounded" h, respectively. Although less severely than in Tables 2 and 3, C is under-sized for all n. When h is "divergent "and as n increases from $n = 40$ and $n = 392$, the deviation between actual and nominal values decreases by 40%, while when h is "bounded" and n increases from $n = 40$ to $n = 400$ it decreases by 75%, consistently with Theorem 1 (with $d = e = f = k = 0$). Also, when h is "divergent" sizes for E, T, MV and B are, respectively, on average across the sample sizes considered, 57%, 42%, 14% and 69% closer to 0.05 than those for C. Such figures become 61%, 51%, 22% and 60% when h is "bounded". In both cases the performance of E, T and B is satisfactory, with B and E offering the greatest improvement when h is "divergent" and "bounded", respectively. The test MV, again, is less satisfactory than T, E and B, even though its performance is slightly better than that in Tables 2 and 3.

(Tables 6 and 7 about here)

Tables 6 and 7 concern the intercept model (1.3)/(1.4) for "divergent" and "bounded" h, respectively. The pattern remains similar. On average across the sample sizes considered, for E, T and B the discrepancies between actual and nominal values are reduced by 65%, 46% and 74% when h is "divergent", and by 57%, 88% and 52% when h is "bounded". Overall, E, T and B perform well, with B offering the highest improvement when h is "divergent" and T considerably outperforming both E and B when h is "bounded". Surprisingly, when h is "divergent" the MV test is outperformed by C: on average the empirical sizes for C are 28% closer to the nominal values than those for MV. However, when h is "bounded" MV offers an average improvement of 45% over C.

In Tables 8-13 we examine powers of (the non-size-corrected tests) C, E, T, MV and B against

$$H_1 : \lambda = \bar{\lambda} \neq 0, \tag{5.3}$$

for $\bar{\lambda} = 0.1, 0.5$ and 0.8.

(Tables 8 and 9 about here)

Tables 8 and 9 concern the same regression setting as in Tables 2 and 3. We observe that C, E, T and B perform well for all n, with C slightly the worst. The few exceptions occur for $\bar{\lambda} = 0.1$, where E and T are outperformed by C for $(m, r) = (18, 11)$ and $(m, r) = (5, 80)$, respectively. MV, instead, is outperformed by C for all sample sizes in almost all settings. Overall, B seems to offer the highest power.

(Tables 10 and 11 about here)

Tables 10 and 11 concern pure SAR (1.2). Again, MV has overall the lowest power. More interestingly, in case h is "divergent", for $\bar{\lambda} = 0.1$ and $\bar{\lambda} = 0.5$ E and T offer a slightly lower power than the standard test C for some sample sizes. C, in turn, is outperformed by B for all sample sizes and all choices of $\bar{\lambda}$. When h is "bounded", instead, E, T and B have comparable performances and are superior to C.

(Tables 12 and 13 about here)

Tables 12 and 13 concern the intercept model (1.3)/(1.4). Similarly to Tables 8-11, MV overall performs worst. When h is "divergent" C has lower power than E, T and B, with few exceptions in which E and T perform slightly worse than C (i.e for $\bar{\lambda} = 0.5$ when $(m, r) = (12, 8)$ and $(m, r) = (28, 14)$). Overall, when h is "divergent", B seems to have the highest power. The pattern of the results for "bounded" h is similar to Table 11, with E, T and B having similar performance and offering higher power than C.

Comparisons can be made with the Monte Carlo results reported in RR. The settings only overlap to a limited extent, because RR studied only (1.2) and (1.4), not more general regression models, they did not look at MV-type tests, and on the other hand they included tests of the one-sided alternative $\lambda > 0$. Subject to this we can compare the results in our Tables 4-7 with results of RR. Generally their tests corresponding to our C tests are very over-sized, especially for the intercept model. Their Edgeworth and transformation tests are much improved, though still quite poorly sized for the smallest n, and on the whole ours perform better here also. The bootstrap results are closer, with the LM tests doing better in 10 out of 16 cases.

(Tables 14 and 15 about here)

In Tables 14-17 we assess the performance of our tests against (1.6) for SAR for y, (1.1) when ϵ_i is non-normal. We generate ϵ_i as Laplace, with pdf

$$\text{pdf}(x) = 2^{-1/2} exp(-2^{1/2}|x|). \tag{5.4}$$

We compare the Edgeworth-corrected tests (2.19) and (2.23) with a bootstrap test. The 199 bootstrap statistics are obtained as in (5.1), but with each u_j^* generated by resampling with replacement from the (centred) empirical distribution of Py.

Tables 14 and 15 report empirical sizes when h is "divergent" and "bounded", respectively. The Edgeworth-corrected tests improve on C, indeed, when h is "divergent" the empirical sizes of E and T are 51% and 41% closer to 0.05, on average across sample sizes considered, but improve less when h is "bounded" (by 29% and 24%). As expected, B offers the greatest improvements since bootstrap critical values do not reflect distributional assumptions. On average across n, the sizes obtained by bootstrap critical values are 62% and 56% closer to 0.05 than those based on C. Our results suggest that in the present setting our normality-based Edgeworth-corrected tests E and T provide a "partial" correction when normality does not hold, and perform at least as well as C.

(Tables 16 and 17)

Finally, Tables 16 and 17 display empirical powers of the tests of H_0 in (1.5) for the regression setting of Tables 2 ad 3 when h is "divergent" and "bounded" respectively. For all n, the performance is similar to that in Tables 8 and 9. Except when $(m, r) = (5, 80)$ and $\bar{\lambda} = 0.1$, E and T are more powerful than C.

6 FINAL COMMENTS

We have derived refined LM tests of lack of correlation against spatial autoregressive error correlation in regression models, using Edgeworth expansion, examined their local power, and compared their finite sample performance with other tests. The tests are based on asymptotic thory, but they do seem to improve on standard, uncorrected, tests in modest sample sizes. They are relatively simple to compute, partly due to imposing normality. Edgeworth expansions without distributional assumptions can be derived, in terms of higher order cumulants (e.g. Knight (1985)), but estimates of the latter tend to be imprecise except in very large samples. As Ogasawara (2006a,b) found in other settings, our normal-based tests will remain valid under only slight relaxation of normality, with certain equality restrictions holding (e.g. zero fourth cumulants). Bootstrap-based tests will be valid much more generally, and rival our higher-order improvements, but bootstrap statistics do vary with implementation, and we believe that empirical researchers are still likely to report the standard LM statistic and compare it with χ^2 critical values, in which case it costs little more to carry out our tests, which do not require estimation of any further nuisance parameters. The paper makes other restrictive assumptions.

The requirement of deterministic regressors is quite standard in the SAR literature, but our results should hold after conditioning on stochastic regressors that are independent of errors. Relaxing exogeneity then becomes an issue, but Edgeworth expansions allowing endogeneity would be considerably more complicated. Allowing endogeneity of the weight matrix is also an important issue, but so far as we know serious progress on allowing this, in the context of first-order theory, has begun only recently, see Qu and Lee (2013). Other assumptions will be more straightforward to relax, such as linearity of the regression and homoscedasticity of the innovations ϵ_i.

ACKNOWLEDGMENTS

We thank two referees for comments which have led to significant improvements. The first author's research was supported by ESRC Grant ES/J007242/1.

REFERENCES

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

Anselin, L. (2001). Rao's score test in spatial econometrics, *Journal of Statistical Planning and Inference* **97**, 113-39.

Baltagi, B.H. and D. Li (2004). Testing for linear and log-linear models against Box-Cox alternatives with spatial lag dependence. In Lesage, J.P. and R.K. Pace (ed.) *Spatial and Spatiotemporal Econometrics (Advances in Econometrics)* **18**, 35-74.

Baltagi, B.H. and Z. Yang (2012). Standardized LM tests for spatial error dependence in linear or panel regressions. Forthcoming in *Econometrics Journal*.

Burridge, P. (1980). On the Cliff-Ord test for spatial correlation. *Journal of the Royal Statistical Society, Series B* **42**, 107-8.

Case, A.C. (1991). Spatial Patterns in Household Demand. *Econometrica* **59**, 953-65.

Cliff, A. and J.K. Ord (1972). Testing for spatial autocorrelation among regression residuals. *Geographical Analysis* **4**, 267-84.

Cordeiro, G.M. and S. L. de P. Ferrari (1991). A modified score test statistic having chi-squared distribution to order n^{-1}. *Biometrika* ***78**, 573-82.*

Davis, A.W. (1976). Statistical distributions in univariate and multivariate Edgeworth populations. *Biometrika* **63**, 661-70.

Ghazal, G.A. (1996) Recurrence formula for expectation of products of quadratic forms. *Statistics and Probability Letters* **27**, 101-9.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* Springer-Verlag.

Hillier, G.H. and M. L. King (1985). Locally best invariant tests of the error covariance matrix of the linear regression model. *Journal of the Royal Statistical Society, Series B* **47**, 98-102.

Kakizawa, Y. (1996). Higher order monotone Bartlett-type adjustment for some multivariate test statistics. *Biometrika* **83**, 923-7.

Kelejian, H.H. and D.P. Robinson (1992). Spatial autocorrelation - a new computationally simple test with an application to the per capita county police expenditures. *Regional Science and Urban Economics* **22**, 317-31.

Knight, J.L. (1985). The joint characteristic function of linear and quadratic forms of non-normal variables. *SankhyÄ: The Indian Journal of Statistics, Series A* **47**, 231-8.

Koenker, R. (1981). A note on studentizing a test for heteroscedasticity. *Journal of Econometrics* **17**, 107-12.

Lee, L.F. (2004). Asymptotic distribution of quasi-maximum likelihood estimates for spatial autoregressive models. *Econometrica* **72**, 1899-925.

Moran, P.A.P. (1950). A test for the serial dependence of residuals. *Biometrika* **37**, 178-81.

Ogasawara, H. (2006a). Asymptotic expanson of the the sample correlation coefficient under nonnormality. *Computational Statistics and Data Analysis* **50**, 891-910.

Ogasawara, H. (2006b). Asymptotic expanson and conditional robustness for the sample multipe correlation coefficient under nonnormality. *Communications in Statistics - Simulation and Computation* **35**, 177-199.

Ord, J.K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* **70**, 120-6.

Peers, H.W. (1971). Likelihood ratio and associated test criteria. *Biometrika* **58**, 577-87.

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

Pinkse, J. (2004). Moran-flavoured tests with nuisance parameters: examples. In Anselin, L., Florax, R.G.M., Rey, S. (eds.). *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Springer-Verlag, Berlin, 66-77.

Pitman, E.G.J. (1937). Significant tests which may be applied to samples from any population. Supplement to *Journal of the Royal Statistical Society* **4**, 119-30.

Qu, X., and L.-F. Lee (2013). Estimating a spatial autoregressive model with an endogenous spatial weight matrix. Preprint

Rao, C.R. and R. Mukerjee (1995). Comparison of Bartlett-type adjustments for the efficient score statistic. *Journal of Statistical Planning and Inference* **46**, 137-46.

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

Robinson, P.M. and F. Rossi (2012). Improved tests for spatial correlation. Submitted for publication.

Taniguchi, M. (1991). Third order asymptotic properties of a class of test statistics under a local alternative. *Journal of Multivariate Analysis* **37**, 223-38.

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

APPENDIX

Proof of Theorem 1

Since T in (2.1) is a continuous random variable,

$$P(LM \leq \eta) = P(T \leq \eta^{1/2}) - P(T \leq -\eta^{1/2}). \tag{A.1}$$

Thus we derive the formal Edgeworth expansion of the df of T under H_0,notation following that in the proof of Theorem 1 of RR. Similarly to Phillips (1977),

$$P(T \leq \zeta) = P\left(\frac{(nh)^{1/2}\epsilon' PWP\epsilon}{a^{1/2}\epsilon' P\epsilon} \leq \zeta\right) = P(\epsilon' C\epsilon \leq 0),$$

where

$$C = \frac{1}{2}P(W + W')P - \left(\frac{a}{nh}\right)^{1/2} P\zeta \tag{A.2}$$

and ζ is any real number.

Under Assumption 1, the characteristic function (cf) of $\epsilon' C\epsilon$ is

$$\begin{aligned} E(e^{it(\epsilon' C\epsilon)}) =& \frac{1}{(2\pi)^{n/2}\sigma^n}\int_{\Re^n} e^{it(\xi' C\xi)} e^{-\frac{\xi'\xi}{2\sigma^2}} d\xi = \frac{1}{(2\pi)^{n/2}\sigma^n}\int_{\Re^n} e^{-\frac{1}{2\sigma^2}\xi'(I-2it\sigma^2 C)\xi} d\xi \\ =& det(I - 2it\sigma^2 C)^{-1/2} = \prod_{j=1}^{n}(1 - 2it\sigma^2\gamma_j)^{-1/2}, \end{aligned} \tag{A.3}$$

where $det(A)$ denotes the determinant of a square matrix A, the γ_j are eigenvalues of C and $i = \sqrt{-1}$. From (A.3) the cumulant generating function (cgf) of $\epsilon' C\epsilon$ is

$$\begin{aligned} \psi(t) =& -\frac{1}{2}\sum_{j=1}^{n} ln(1 - 2it\sigma^2\gamma_j) = \frac{1}{2}\sum_{j=1}^{n}\sum_{s=1}^{\infty}\frac{(2it\sigma^2\gamma_j)^s}{s} \\ =& \frac{1}{2}\sum_{s=1}^{\infty}\frac{(2it\sigma^2)^s}{s}\sum_{j=1}^{n}\gamma_j^s = \frac{1}{2}\sum_{s=1}^{\infty}\frac{(2it\sigma^2)^s}{s} tr(C^s) \end{aligned}$$

and thus the s-th cumulant, κ_s, of $\epsilon' C\epsilon$ is

$$\kappa_1 = \sigma^2 tr(C), \tag{A.4}$$

$$\kappa_2 = 2\sigma^4 tr(C^2), \tag{A.5}$$

$$\kappa_s = \frac{\sigma^{2s} s! 2^{s-1} tr(C^s)}{s}, s > 2. \tag{A.6}$$

The cgf of $(\epsilon' C\epsilon - \kappa_1)/\kappa_2^{1/2}$ is

$$\psi^c(t) = -\frac{1}{2}t^2 + \sum_{s=3}^{\infty} \frac{\kappa_s^c (it)^s}{s!}, \tag{A.7}$$

where $\kappa_s^c = \kappa_s/\kappa_2^{s/2}$. Hence,

$$\begin{aligned}
E(e^{it(\epsilon' C\epsilon - \kappa_1)/\kappa_2^{1/2}}) =& e^{-\frac{1}{2}t^2} \exp\{\sum_{s=3}^{\infty} \frac{\kappa_s^c (it)^s}{s!}\} \\
=& e^{-\frac{1}{2}t^2}\{1 + \sum_{s=3}^{\infty} \frac{\kappa_s^c (it)^s}{s!} + \frac{1}{2!}(\sum_{s=3}^{\infty} \frac{\kappa_s^c (it)^s}{s!})^2 + \frac{1}{3!}(\sum_{s=3}^{\infty} \frac{\kappa_s^c (it)^s}{s!})^3 + ...\} \\
=& e^{-\frac{1}{2}t^2}\{1 + \frac{\kappa_3^c (it)^3}{3!} + \frac{\kappa_4^c (it)^4}{4!} + \frac{\kappa_5^c (it)^5}{5!} + \{\frac{\kappa_6^c}{6!} + \frac{(\kappa_3^c)^2}{(3!)^2}\}(it)^6 + ...\}.
\end{aligned}$$

Denote by $\phi(\zeta)$ the normal pdf. By Fourier inversion, formally,

$$P((\epsilon' C\epsilon - \kappa_1)/\kappa_2^{1/2} \le z) = \int_{-\infty}^{z} \phi(z)dz + \frac{\kappa_3^c}{3!}\int_{-\infty}^{z} H_3(z)\phi(z)dz + \frac{\kappa_4^c}{4!}\int_{-\infty}^{z} H_4(z)\phi(z)dz + ...,$$

where $H_i(z)$ is the $i-th$ Hermite polynomial. Collecting the above results,

$$\begin{aligned}
P(T \le \zeta) =& P(\epsilon' C\epsilon \le 0) = P((\epsilon' C\epsilon - \kappa_1)/\kappa_2^{1/2} \le -\kappa_1^c) \\
=& \Phi(-\kappa_1^c) - \frac{\kappa_3^c}{3!}\Phi^{(3)}(-\kappa_1^c) + \frac{\kappa_4^c}{4!}\Phi^{(4)}(-\kappa_1^c) + ...,
\end{aligned}$$

where $q^{(i)}$ denotes the ith derivative of the function q.

From (A.3)-(A.5),

$$\kappa_1 = \sigma^2\left(tr(PW) - \left(\frac{a}{nh}\right)^{1/2} tr(P)\zeta\right) = -\sigma^2\left(e + \left(\frac{na}{h}\right)^{1/2}\zeta - \frac{a^{1/2}k}{(nh)^{1/2}}\zeta\right)$$

and

$$\kappa_2 = \sigma^4(tr(W^2) + tr(W'W) + \frac{1}{2}tr(X'(W+W')X(X'X)^{-1}X'(W'+W)X(X'X)^{-1})$$
$$- tr(X'(W+W')^2X(X'X)^{-1}) + 2\frac{(n-k)a}{nh}\zeta^2 + \frac{4tr((X'X)^{-1}X'WX)a^{1/2}}{(nh)^{1/2}}\zeta)$$
$$= \sigma^4\left(\frac{n}{h}a + f - d + 2\left(\frac{a}{h} - \frac{ak}{nh}\right)\zeta^2 + \frac{4ea^{1/2}}{(nh)^{1/2}}\zeta)\right),$$

where a, d, e and f are defined in (2.3), (2.4) and (2.5). Thus

$$\kappa_1^c = \left(-\zeta - e(h/n)^{1/2}a^{-1/2} + \frac{k}{n}\zeta\right)\left(1 + \frac{h}{n}a^{-1}(f-d) + 2\frac{\zeta^2}{n} + o\left(\frac{1}{n}\right)\right)^{-1/2}.$$

By Taylor expansion, we deduce

$$\kappa_1^c = -\zeta - e\left(\frac{h}{n}\right)^{1/2}a^{-1/2} + \frac{1}{2}\frac{h}{n}a^{-1}(f-d)\zeta + o\left(\frac{h}{n}\right), \tag{A.8}$$

when h is divergent, and

$$\kappa_1^c = -\zeta - e\left(\frac{h}{n}\right)^{1/2}a^{-1/2} + \frac{1}{2}\frac{h}{n}a^{-1}(f-d)\zeta + \frac{1}{n}\left(\zeta^2 + k\right)\zeta + o\left(\frac{1}{n}\right), \tag{A.9}$$

when h is bounded.

Also, from (A.6) and (A.5),

$$\kappa_3^c = \frac{8\sigma^6 tr(C^3)}{\kappa_2^{3/2}} = \frac{tr((P(W+W')P)^3)}{(n/h)^{3/2}a^{3/2}} + o\left(\frac{h}{n}\right) = \left(\frac{h}{n}\right)^{1/2}\frac{b}{a^{3/2}} + o\left(\frac{h}{n}\right), \tag{A.10}$$

when h is divergent, and

$$\kappa_3^c = \frac{tr((P(W+W')P)^3)}{(n/h)^{3/2}a^{3/2}} - \frac{6htr(((W+W')P)^2)\zeta}{n^2a} + o\left(\frac{1}{n}\right)$$
$$= \left(\frac{h}{n}\right)^{1/2}\frac{b}{a^{3/2}} - \frac{12\zeta}{n} + o\left(\frac{1}{n}\right), \tag{A.11}$$

when h is bounded. Similarly, for h either divergent or bounded,

$$\kappa_4^c = \frac{48\sigma^8 tr(C^4)}{\kappa_2^2} = \frac{3tr(((W+W')P)^4)}{(n/h)^2a^2} + o\left(\frac{h}{n}\right) = \frac{h}{n}\frac{3c}{a^2} + o\left(\frac{h}{n}\right),$$

where c is defined in (2.3).

From (A.8) and (A.9) and by Taylor expansion, we obtain, respectively,

$$\Phi(-\kappa_1^c) = \Phi(\zeta) + \left(\frac{h}{n}\right)^{1/2} ea^{-1/2}\phi(\zeta) + \frac{h}{n}\frac{f-d+e^2}{2a}\Phi^{(2)}(\zeta) + o\left(\frac{h}{n}\right),$$

when h is divergent, and

$$\Phi(-\kappa_1^c) = \Phi(\zeta) + \frac{1}{n^{1/2}}h^{1/2}ea^{-1/2}\phi(\zeta) + \frac{1}{n}\left(\frac{h(e^2+f-d)}{2a} + \zeta^2 + k\right)\Phi^{(2)}(\zeta) + o\left(\frac{1}{n}\right),$$

when h is bounded. Similarly,

$$\Phi^{(3)}(-\kappa_1^c) = \Phi^{(3)}(\zeta) + \left(\frac{h}{n}\right)^{1/2} ea^{-1/2}\Phi^{(4)}(\zeta) + O\left(\frac{h}{n}\right)$$

whether h is divergent or bounded.

Noting that

$$\Phi^{(2)}(x) = -x\phi(x), \quad \Phi^{(3)}(x) = (x^2-1)\phi(x), \quad \Phi^{(4)}(x) = (3x - x^3)\phi(x),$$

and collecting the above results, under H_0 when h is divergent:

$$\begin{aligned}
P\left(T \le \zeta\right) = {} & \Phi(\zeta) + \left(\frac{h}{n}\right)^{1/2}\left(\frac{e}{a^{1/2}}\phi(\zeta) - \frac{b}{6a^{3/2}}\Phi^{(3)}(\zeta)\right) \\
& + \frac{h}{n}\left(\frac{e^2+f-d}{2a}\Phi^{(2)}(\zeta) + \frac{1}{2a^2}\left(\frac{c}{4} - \frac{eb}{3}\right)\Phi^{(4)}(\zeta)\right) + o\left(\frac{h}{n}\right)
\end{aligned}$$

$$\begin{aligned}
& = \Phi(\zeta) + \left(\frac{h}{n}\right)^{1/2}\left(\frac{e}{a^{1/2}} - \frac{b}{6a^{3/2}}(\zeta^2-1)\right)\phi(\zeta) \\
& + \frac{h}{n}\left(-\frac{e^2+f-d}{2a}\zeta - \frac{1}{2a^2}\left(\frac{c}{4} - \frac{eb}{3}\right)(\zeta^3 - 3\zeta)\right)\phi(\zeta) + o\left(\frac{h}{n}\right) \\
& = \Phi(\zeta) + \left(\frac{h}{n}\right)^{1/2}\left(\frac{e}{a^{1/2}} - \frac{b}{6a^{3/2}}(\zeta^2-1)\right)\phi(\zeta) \\
& + \frac{h}{n}\frac{\omega_1\left(\zeta^2\right)}{2\zeta}\phi(\zeta) + o\left(\frac{h}{n}\right),
\end{aligned}$$

since

$$-\frac{e^2+f-d}{2a}\zeta - \frac{1}{2a^2}\left(\frac{c}{4} - \frac{eb}{3}\right)(\zeta^3 - 3\zeta) = \frac{1}{2}v_1\zeta - \frac{1}{2}v_2\zeta^3 = \frac{\omega_1\left(\zeta^2\right)}{2\zeta};$$

when h is bounded:

$$
\begin{aligned}
&P\left(T \leq \zeta\right) = \Phi(\zeta) + \left(\frac{h}{n}\right)^{1/2} \left(\frac{e}{a^{1/2}}\phi(\zeta) - \frac{b}{6a^{3/2}}\Phi^{(3)}(\zeta)\right) \\
&+ \frac{1}{n}\left(\left(\frac{h(e^2+f-d)}{2a} + \zeta^2 + k\right)\Phi^{(2)}(\zeta) + 2\zeta\Phi^{(3)}(\zeta) + \frac{h}{2a^2}\left(\frac{c}{4} - \frac{eb}{3}\right)\Phi^{(4)}(\zeta)\right) + o\left(\frac{1}{n}\right) \\
=&\Phi(\zeta) + \left(\frac{h}{n}\right)^{1/2}\left(\frac{e}{a^{1/2}} - \frac{b}{6a^{3/2}}(\zeta^2-1)\right)\phi(\zeta) \\
&+ \frac{1}{n}\left(-\left(\frac{h\left(e^2+f-d\right)}{2a} + \zeta^2 + k\right)\zeta + 2\zeta(\zeta^2-1) - \frac{h}{2a^2}\left(\frac{c}{4} - \frac{eb}{3}\right)(\zeta^3 - 3\zeta)\right)\phi(\zeta) + o\left(\frac{h}{n}\right) \\
=&\Phi(\zeta) + \left(\frac{h}{n}\right)^{1/2}\left(\frac{e}{a^{1/2}} - \frac{b}{6a^{3/2}}(\zeta^2-1)\right)\phi(\zeta) + \frac{h}{n}\frac{\omega_2\left(\zeta^2\right)}{2\zeta}\phi(\zeta) + o\left(\frac{h}{n}\right),
\end{aligned}
$$

since

$$
\begin{aligned}
&-\left(\frac{h\left(e^2+f-d\right)}{2a} + \zeta^2 + k\right)\zeta + 2\zeta(\zeta^2-1) - \frac{h}{2a^2}\left(\frac{c}{4} - \frac{eb}{3}\right)(\zeta^3 - 3\zeta) \\
=&\frac{h}{2}v_1\zeta - \frac{h}{2}v_2\zeta^3 - \zeta(k+2) + \zeta^3 = \frac{\omega_2\left(\zeta^2\right)}{2\zeta}.
\end{aligned}
$$

Now for $Z \sim \mathcal{N}(0,1)$,

$$
\Phi(\eta^{1/2}) - \Phi(-\eta^{1/2}) = P\left(|Z| \leq \eta^{1/2}\right) = P\left(Z^2 \leq \eta\right) = \Psi\left(\eta\right),
$$

while, from (2.8),

$$
\eta^{-1/2}\left(\phi(\eta^{1/2}) + \phi(-\eta^{1/2})\right) = 2\psi\left(\eta\right).
$$

Thus, from (A.1),

$$
P(LM \leq \eta) = \Phi(\eta^{1/2}) - \Phi(-\eta^{1/2}) + \frac{h}{n}\frac{\omega_i(\eta)}{2\eta^{1/2}}(\phi(\eta^{1/2}) + \phi(-\eta^{1/2})) + o\left(\frac{h}{n}\right),
$$

for $i = 1$ when h is divergent and $i = 2$ when h is bounded, to give (2.12) and (2.13).

Proof of Corollary 1

Let h be divergent. By inverting (2.11), we can expand w_α as

$$w_\alpha = z^2_{\alpha/2} + p_1(z^2_{\alpha/2}) + o\left(\frac{h}{n}\right), \tag{A.12}$$

where $p_1(z^2_{\alpha/2})$ is a polynomial whose coefficients have exact order h/n, and can be determined from $1-\alpha = P(LM \le w_\alpha)$ and (2.11). Thus, using (2.11),

$$1-\alpha = P(LM \le w_\alpha) = \Psi(w_\alpha) + \frac{h}{n}\omega_1(w_\alpha)\psi(w_\alpha) + o\left(\frac{h}{n}\right).$$

Substituting (A.12), this is

$$\begin{aligned} P(LM \le w_\alpha) =& \Psi(z^2_{\alpha/2}) + p_1(z^2_{\alpha/2})\psi(z^2_{\alpha/2}) + \frac{h}{n}\omega_1(z^2_{\alpha/2})\psi(z^2_{\alpha/2}) + o\left(\frac{h}{n}\right) \\ =& 1-\alpha + p_1(z^2_{\alpha/2})\psi(z^2_{\alpha/2}) + \frac{h}{n}\omega_1(z^2_{\alpha/2})\psi(z^2_{\alpha/2}) + o\left(\frac{h}{n}\right). \end{aligned}$$

The latter is $1-\alpha+o(h/n)$ (rather than $1-\alpha+O(h/n)$) when we take $p_1(x) = -h\omega_1(z^2_{\alpha/2})/n$, which has exact order h/n. Hence, (2.17) and (2.18) follow from (A.12). The corresponding result for bounded h follows analogously from (2.12).

Proof of Theorem 2

The prof is similar to that of Theorem 1 so some details will be omitted. In view of (A.1), we derive the Edgeworth expansion of T under H_1. Write

$$P(T \le \zeta) = P(\epsilon' C \epsilon \le 0),$$

with

$$C = \frac{1}{2}S^{-1}(\lambda_n)'P(W+W')PS^{-1}(\lambda_n) - \frac{1}{(hn)^{1/2}}\zeta a^{1/2}S^{-1}(\lambda_n)'PS^{-1}(\lambda_n).$$

The cumulants κ_j of $\epsilon' C \epsilon$ are

$$\begin{aligned}
\kappa_1 =& \sigma^2 tr(C) \\
=& \sigma^2 \left(\frac{1}{2} tr \left(\sum_{t=0}^{\infty} (\lambda_n W')^t P (W + W') P \sum_{t=0}^{\infty} (\lambda_n W)^t \right) - \frac{\zeta a^{1/2}}{(nh)^{1/2}} tr \left(\sum_{t=0}^{\infty} (\lambda_n W')^t P \sum_{t=0}^{\infty} (\lambda_n W)^t \right) \right) \\
=& \sigma^2 \left(- \left(\frac{n}{h} \right)^{1/2} a^{1/2} (\zeta - \delta a^{1/2}) - e + \frac{h}{n} \delta^2 tr(W^3 + 2W^2 W') \right) + O\left(\left(\frac{h}{n} \right)^{1/2} \right),
\end{aligned}$$

and similarly

$$\kappa_2 = 2\sigma^4 tr(C^2) = \sigma^4 \left(\frac{n}{h} a + 2\delta \left(\frac{h}{n} \right)^{1/2} tr(W^3 + 3W^2 W') \right) + O(1),$$

where a and e are defined in (2.3) and (2.4), respectively. Thus, the first centred cumulant of $\epsilon' C \epsilon$ is

$$\kappa_1^c = -(\zeta - \delta a^{1/2}) + \left(\frac{h}{n} \right)^{1/2} a^{-1/2} \left(-e - \delta^2 p + \frac{1}{2} \delta a^{-1/2} b \zeta \right) + O\left(\frac{h}{n} \right),$$

and accordingly,

$$\begin{aligned}
\Phi(-\kappa_1^c) =& \Phi(\zeta - \delta a^{1/2}) \\
& - \left(\frac{h}{n} \right)^{1/2} a^{-1/2} \left(-e - \delta^2 p + \frac{1}{2} \delta a^{-1/2} b \zeta \right) \phi(\zeta - \delta a^{1/2}) + O\left(\frac{h}{n} \right),
\end{aligned}$$

where b and p are defined in (2.3) and (4.5), respectively. Under H_1, the leading term of the third centred cumulant of $\epsilon' C \epsilon$ is identical to that in (A.10)/(A.11), which is

$$\kappa_3^c = \frac{8\sigma^6 tr(C^3)}{\kappa_2^{3/2}} = \left(\frac{h}{n} \right)^{1/2} \frac{b}{a^{3/2}} + O\left(\frac{h}{n} \right).$$

Proceeding as in the proof of Theorem 1, under H_1,

$$\begin{aligned}
P(T \leq \zeta) =& \Phi(\zeta - a^{1/2}\delta) - \left(\frac{h}{n}\right)^{1/2} a^{-1/2}\left(-e - \delta^2 p + \frac{1}{2}\delta a^{-1/2} b\zeta\right)\phi(\zeta - a^{1/2}\delta) \\
& - \left(\frac{h}{n}\right)^{1/2} \frac{b}{6a^{3/2}} \Phi^{(3)}(\zeta - a^{1/2}\delta) + O\left(\frac{h}{n}\right) \\
=& \Phi(\zeta - a^{1/2}\delta) - \left(\frac{h}{n}\right)^{1/2} a^{-1/2}\left(-e - \delta^2 p + \frac{1}{2}\delta a^{-1/2} b\zeta\right)\phi(\zeta - a^{1/2}\delta) \\
& - \left(\frac{h}{n}\right)^{1/2} \frac{b}{6a^{3/2}}\left(\left(\zeta - a^{1/2}\delta\right)^2 - 1\right)\phi(\zeta - a^{1/2}\delta) + O\left(\frac{h}{n}\right) \\
=& \Phi(\zeta - a^{1/2}\delta) + \left(\frac{h}{n}\right)^{1/2} a^{-1/2}\left(e + \delta^2 p\right) - \left(\frac{h}{n}\right)^{1/2} \frac{\delta b}{2a}\zeta\phi(\zeta - a^{1/2}\delta) \\
& - \left(\frac{h}{n}\right)^{1/2} \frac{b}{6a^{3/2}}\left(\zeta^2 - 2a^{1/2}\delta\zeta + a\delta^2 + 1\right)\phi(\zeta - a^{1/2}\delta) + O\left(\frac{h}{n}\right) \\
=& \Phi(\zeta - a^{1/2}\delta) + \left(\frac{h}{n}\right)^{1/2}\left(a^{-1/2}\left(e + \delta^2 p\right) - \frac{b\left(a\delta^2 + 1\right)}{6a^{3/2}}\right)\phi(\zeta - a^{1/2}\delta) \\
& - \left(\frac{h}{n}\right)^{1/2} \frac{b\delta}{6a}\zeta\phi(\zeta - a^{1/2}\delta) - \left(\frac{h}{n}\right)^{1/2} \frac{b}{6a^{3/2}}\zeta^2\phi(\zeta - a^{1/2}\delta) + o\left(\left(\frac{h}{n}\right)^{1/2}\right).
\end{aligned}$$

Noting that

$$\begin{aligned}
\Phi(\zeta - \nu) - \Phi(-\zeta - \nu) =& P(Z \leq \zeta - \nu) - P(Z \leq -\zeta - \nu) \\
=& P((Z + \nu)^2 \leq \zeta^2) = \Psi\left(\zeta^2; \nu\right),
\end{aligned}$$

$$\begin{aligned}
\phi(\zeta - \nu) + \phi(-\zeta - \nu) =& \frac{1}{\sqrt{2\pi}}\left(\exp\left(-\frac{1}{2}(\zeta - \nu)^2\right) + \exp\left(-\frac{1}{2}(\zeta + \nu)^2\right)\right) \\
=& \sqrt{\frac{2}{\pi}}\cosh(\nu\zeta)\exp\left(-\frac{1}{2}\left(\zeta^2 + \nu^2\right)\right) \\
=& 2\zeta\psi\left(\zeta^2; \nu\right),
\end{aligned}$$

$$\begin{aligned}
\phi(\zeta - \nu) - \phi(-\zeta - \nu) =& \frac{1}{\sqrt{2\pi}}\left(\exp\left(-\frac{1}{2}(\zeta - \nu)^2\right) - \exp\left(-\frac{1}{2}(\zeta + \nu)^2\right)\right) \\
=& \sqrt{\frac{2}{\pi}}\sinh(\nu\zeta)\exp\left(-\frac{1}{2}\left(\zeta^2 + \nu^2\right)\right) = \tau\left(\zeta^2; \nu\right),
\end{aligned}$$

and therefore

$$\begin{aligned}
P(LM \leq \eta) =& \Phi(\eta^{1/2} - a^{1/2}\delta) - \Phi(-\eta^{1/2} - a^{1/2}\delta) \\
&+ \left(\frac{h}{n}\right)^{1/2} \left(a^{-1/2}\left(e + \delta^2 p\right) - \frac{b\left(a\delta^2+1\right)}{6a^{3/2}}\right)\left(\phi(\eta^{1/2} - a^{1/2}\delta) - \phi(-\eta^{1/2} - a^{1/2}\delta)\right) \\
&- \left(\frac{h}{n}\right)^{1/2} \frac{b\delta}{6a}\eta^{1/2}\left(\phi(\eta^{1/2} - a^{1/2}\delta) + \phi(-\eta^{1/2} - a^{1/2}\delta)\right) \\
&- \left(\frac{h}{n}\right)^{1/2} \frac{b}{6a^{3/2}}\eta\left(\phi(\eta^{1/2} - a^{1/2}\delta) - \phi(-\eta^{1/2} - a^{1/2}\delta)\right) + O\left(\frac{h}{n}\right) \\
=& \Psi\left(\eta; a^{1/2}\delta\right) + \left(\frac{h}{n}\right)^{1/2}\left(a^{-1/2}\left(e + \delta^2 p\right) - \frac{b\left(a\delta^2+1\right)}{6a^{3/2}}\right)\tau\left(\eta; a^{1/2}\delta\right) \\
&- \left(\frac{h}{n}\right)^{1/2} \frac{b\delta}{2a}\eta\psi\left(\eta; a^{1/2}\delta\right) - \left(\frac{h}{n}\right)^{1/2} \frac{b}{6a^{3/2}}\eta\tau\left(\eta; a^{1/2}\delta\right) + o\left(\left(\frac{h}{n}\right)^{1/2}\right),
\end{aligned}$$

to conclude the proof.

	$m = 8$ $r = 5$	$m = 12$ $r = 8$	$m = 18$ $r = 11$	$m = 28$ $r = 14$
Wald	0.112	0.119	0.088	0.090
LR	0.075	0.071	0.070	0.075
LM	0.030	0.031	0.035	0.038

Table 1: Empirical sizes of standard Wald, LR and LM tests of H_0 (1.5) against H_1 (1.6) for pure SAR (1.2) when h is "divergent". $\alpha = 5\%$.

	$m = 8$ $r = 5$	$m = 12$ $r = 8$	$m = 18$ $r = 11$	$m = 28$ $r = 14$
chi square	0.016	0.020	0.028	0.032
Edgeworth	0.035	0.036	0.039	0.041
transformation	0.033	0.038	0.043	0.039
mean-variance correction	0.015	0.022	0.029	0.032
bootstrap	0.040	0.055	0.056	0.058

Table 2: Empirical sizes of tests of H_0 (1.5) against H_1 (1.6) for regression with SAR disturbances (1.1) when h is "divergent". $\alpha = 5\%$.

	$m = 5$ $r = 8$	$m = 5$ $r = 20$	$m = 5$ $r = 40$	$m = 5$ $r = 80$
chi square	0.024	0.031	0.044	0.046
Edgeworth	0.045	0.046	0.053	0.054
transformation	0.044	0.045	0.047	0.046
mean-variance correction	0.032	0.039	0.041	0.052
bootstrap	0.039	0.045	0.048	0.046

Table 3: Empirical sizes of tests of H_0 (1.5) against H_1 (1.6) for regression with SAR disturbances (1.1) when h is "bounded". $\alpha = 5\%$.

	$m = 8$ $r = 5$	$m = 12$ $r = 8$	$m = 18$ $r = 11$	$m = 28$ $r = 14$
chi square	0.030	0.031	0.035	0.038
Edgeworth	0.042	0.042	0.043	0.045
transformation	0.039	0.037	0.040	0.045
mean-variance correction	0.022	0.032	0.033	0.037
bootstrap	0.057	0.045	0.047	0.055

Table 4: Empirical sizes of tests of H_0 (1.5) against H_1 (1.6) for the pure SAR model (1.2) when h is "divergent". $\alpha = 5\%$.

	$m = 5$ $r = 8$	$m = 5$ $r = 20$	$m = 5$ $r = 40$	$m = 5$ $r = 80$
chi square	0.030	0.038	0.039	0.045
Edgeworth	0.043	0.045	0.052	0.047
transformation	0.041	0.046	0.048	0.045
mean-variance correction	0.035	0.036	0.041	0.048
bootstrap	0.063	0.052	0.054	0.048

Table 5: Empirical sizes of tests of H_0 (1.5) against H_1 (1.6) for the pure SAR model (1.2) when h is “bounded”. $\alpha = 5\%$.

	$m = 8$ $r = 5$	$m = 12$ $r = 8$	$m = 18$ $r = 11$	$m = 28$ $r = 14$
chi square	0.031	0.032	0.034	0.039
Edgeworth	0.040	0.041	0.053	0.048
transformation	0.061	0.045	0.041	0.058
mean-variance correction	0.020	0.023	0.032	0.040
bootstrap	0.055	0.045	0.049	0.045

Table 6: Empirical sizes of tests of H_0 (1.5) against H_1 (1.6) for the intercept model (1.3) when h is “divergent”. $\alpha = 5\%$.

	$m = 5$ $r = 8$	$m = 5$ $r = 20$	$m = 5$ $r = 40$	$m = 5$ $r = 80$
chi square	0.023	0.035	0.036	0.042
Edgeworth	0.040	0.044	0.040	0.052
transformation	0.046	0.049	0.048	0.051
mean-variance correction	0.023	0.041	0.045	0.048
bootstrap	0.057	0.040	0.043	0.046

Table 7: Empirical sizes of tests of H_0 (1.5) against H_1 (1.6) for the intercept model (1.3) when h is “bounded”. $\alpha = 5\%$.

	$\bar{\lambda}$	$m = 8$ $r = 5$	$m = 12$ $r = 8$	$m = 18$ $r = 11$	$m = 28$ $r = 14$
chi square	0.1	0.0620	0.0790	0.0920	0.0920
	0.5	0.5850	0.8100	0.8890	0.9530
	0.8	0.9750	0.9990	1	1
Edgeworth	0.1	0.0730	0.0840	0.0880	0.1120
	0.5	0.6230	0.8150	0.8980	0.9540
	0.8	0.9750	1	1	1
transformation	0.1	0.0790	0.0910	0.0990	0.1040
	0.5	0.5980	0.8160	0.8920	0.9540
	0.8	0.9820	0.9980	1	1
mean-variance correction	0.1	0.0510	0.0650	0.0780	0.0940
	0.5	0.5860	0.7540	0.8760	0.9410
	0.8	0.9770	0.9990	1	1
bootstrap	0.1	0.1080	0.1100	0.1260	0.1270
	0.5	0.6580	0.8230	0.9050	1
	0.8	0.9860	1	1	1

Table 8: Empirical powers of tests of H_0 (1.5) against H_1 (5.3), with $\bar{\lambda} = 0.1, 0.5, 0.8$, for regression with SAR disturbances (1.1) when h is "divergent". $\alpha = 5\%$

	$\bar{\lambda}$	$m = 5$ $r = 8$	$m = 5$ $r = 20$	$m = 5$ $r = 40$	$m = 5$ $r = 80$
chi square	0.1	0.0850	0.1290	0.2130	0.3390
	0.5	0.7800	0.9880	1	1
	0.8	1	1	1	1
Edgeworth	0.1	0.0860	0.1290	0.2340	0.3390
	0.5	0.7820	0.9920	1	1
	0.8	1	1	1	1
transformation	0.1	0.0890	0.1370	0.2180	0.3210
	0.5	0.7920	0.9900	1	1
	0.8	0.9990	1	1	1
mean-variance correction	0.1	0.0730	0.1260	0.2140	0.3370
	0.5	0.7730	0.9920	1	1
	0.8	0.9990	1	1	1
bootstrap	0.1	0.0910	0.1300	0.2290	0.3520
	0.5	0.8090	0.9890	1	1
	0.8	0.9980	1	1	1

Table 9: Empirical powers of tests of H_0 (1.5) against H_1 (5.3), with $\bar{\lambda} = 0.1, 0.5, 0.8$, for regression with SAR disturbances (1.1) when h is "bounded". $\alpha = 5\%$.

	$\bar{\lambda}$	$m = 8$ $r = 5$	$m = 12$ $r = 8$	$m = 18$ $r = 11$	$m = 28$ $r = 14$
chi square	0.1	0.1050	0.1200	0.1290	0.1312
	0.5	0.7110	0.8680	0.9180	0.9680
	0.8	0.9930	1	1	1
Edgeworth	0.1	0.1040	0.1140	0.1210	0.1315
	0.5	0.7220	0.8470	0.9210	0.9830
	0.8	0.9930	1	1	1
transformation	0.1	0.0960	0.1160	0.1180	0.1400
	0.5	0.7170	0.8470	0.9300	0.9560
	0.8	0.9960	1	1	1
mean-variance correction	0.1	0.0560	0.0910	0.1070	0.1130
	0.5	0.6300	0.8360	0.9112	0.9480
	0.8	0.9890	1	1	1
bootstrap	0.1	0.1070	0.1260	0.1340	0.1370
	0.5	0.7660	0.8790	0.9330	1
	0.8	0.9960	0.9980	1	1

Table 10: Empirical powers of tests of H_0 (1.5) against H_1 (5.3), with $\bar{\lambda} = 0.1, 0.5, 0.8$, for the pure SAR model (1.2) when h is "divergent". $\alpha = 5\%$.

	$\bar{\lambda}$	$m = 5$ $r = 8$	$m = 5$ $r = 20$	$m = 5$ $r = 40$	$m = 5$ $r = 80$
chi square	0.1	0.1000	0.1520	0.2400	0.3560
	0.5	0.8700	0.9940	1	1
	0.8	1	1	1	1
Edgeworth	0.1	0.1130	0.1580	0.2440	0.3460
	0.5	0.8650	0.9960	1	1
	0.8	1	1	1	1
transformation	0.1	0.1160	0.1800	0.2410	0.3640
	0.5	0.8930	0.9920	1	1
	0.8	1	1	1	1
mean-variance correction	0.1	0.0960	0.1610	0.2090	0.3370
	0.5	0.8620	0.9940	1	1
	0.8	1	1	1	1
bootstrap	0.1	0.1240	0.1660	0.2110	0.3450
	0.5	0.9000	0.9940	1	1
	0.8	1	1	1	1

Table 11: Empirical powers of tests of H_0 (1.5) against H_1 (5.3), with $\bar{\lambda} = 0.1, 0.5, 0.8$, for the pure SAR model in (1.2) when h is "bounded". $\alpha = 5\%$.

	$\bar{\lambda}$	$m = 8$ $r = 5$	$m = 12$ $r = 8$	$m = 18$ $r = 11$	$m = 28$ $r = 14$
chi square	0.1	0.0550	0.0800	0.0930	0.0980
	0.5	0.6050	0.7900	0.8800	0.9500
	0.8	0.9810	0.9990	1	1
Edgeworth	0.1	0.0720	0.0920	0.1030	0.1040
	0.5	0.6150	0.7740	0.9030	0.9360
	0.8	0.9810	0.9990	1	1
transformation	0.1	0.0700	0.0810	0.0930	0.0980
	0.5	0.6140	0.7850	0.8920	0.9400
	0.8	0.9810	0.9990	1	1
mean-variance correction	0.1	0.0510	0.0580	0.0770	0.0910
	0.5	0.5510	0.7750	0.8870	0.9250
	0.8	0.9690	0.9960	1	1
bootstrap	0.1	0.0800	0.1110	0.1190	0.1220
	0.5	0.6270	0.8330	0.8970	0.9530
	0.8	0.9820	1	1	1

Table 12: Empirical powers of tests of H_0 (1.5) against H_1 (5.3), with $\bar{\lambda} = 0.1, 0.5, 0.8$, for the intercept model (1.3) when the sequence h is "divergent". $\alpha = 5\%$.

	$\bar{\lambda}$	$m = 5$ $r = 8$	$m = 5$ $r = 20$	$m = 5$ $r = 40$	$m = 5$ $r = 80$
chi square	0.1	0.0690	0.1310	0.1990	0.3470
	0.5	0.7800	0.9930	1	1
	0.8	0.9990	1	1	1
Edgeworth	0.1	0.1090	0.1390	0.2080	0.3480
	0.5	0.8070	0.9920	1	1
	0.8	0.9990	1	1	1
transformation	0.1	0.0960	0.1370	0.2060	0.3560
	0.5	0.8080	0.9880	0.9990	1
	0.8	0.9980	1	1	1
mean-variance correction	0.1	0.6880	0.1380	0.2190	0.3480
	0.5	0.7970	0.9950	1	1
	0.8	1	1	1	1
bootstrap	0.1	0.0950	0.1440	0.2040	0.3470
	0.5	0.8450	0.9920	1	1
	0.8	1	1	1	1

Table 13: Empirical powers of tests of H_0 (1.5) against H_1 (5.3), with $\bar{\lambda} = 0.1, 0.5, 0.8$, for the intercept model (1.3) when h is "bounded". $\alpha = 5\%$.

	$m = 8$ $r = 5$	$m = 12$ $r = 8$	$m = 18$ $r = 11$	$m = 28$ $r = 14$
chi square	0.021	0.024	0.025	0.032
Edgeworth	0.030	0.031	0.042	0.046
transformation	0.027	0.031	0.035	0.046
bootstrap	0.041	0.068	0.056	0.055

Table 14: Empirical sizes of tests of H_0 (1.5) against H_1 (1.6) for regression with SAR disturbances (1.1) when h is "divergent" and the disturbances are generated as in (5.4). $\alpha = 5\%$.

	$m = 5$ $r = 8$	$m = 5$ $r = 20$	$m = 5$ $r = 40$	$m = 5$ $r = 80$
chi square	0.023	0.032	0.034	0.047
Edgeworth	0.038	0.039	0.048	0.055
transformation	0.035	0.037	0.038	0.047
bootstrap	0.042	0.056	0.048	0.047

Table 15: Empirical sizes of tests of H_0 (1.5) against H_1 (1.6) for regression with SAR disturbances (1.1) when h is "bounded" and the disturbances are generated as in (5.4). $\alpha = 5\%$.

	$\bar{\lambda}$	$m = 8$ $r = 5$	$m = 12$ $r = 8$	$m = 18$ $r = 11$	$m = 28$ $r = 14$
chi square	0.1	0.044	0.067	0.072	0.096
	0.5	0.615	0.787	0.888	0.947
	0.8	0.974	0.999	1	1
Edgeworth	0.1	0.070	0.083	0.099	0.104
	0.5	0.598	0.797	0.897	0.941
	0.8	0.982	0.999	1	1
transformation	0.1	0.073	0.085	0.093	0.108
	0.5	0.621	0.800	0.894	0.949
	0.8	0.989	0.998	1	1
bootstrap	0.1	0.083	0.093	0.110	0.944
	0.5	0.658	0.805	0.909	1
	0.8	0.982	0.999	1	1

Table 16: Empirical powers of tests of H_0 (1.5) against H_1 (5.3), with $\bar{\lambda} = 0.1, 0.5, 0.8$, for regression with SAR disturbances (1.1) when h is "divergent" and the disturbances are generated as in (5.4). $\alpha = 5\%$

	$\bar{\lambda}$	$m = 5$ $r = 8$	$m = 5$ $r = 20$	$m = 5$ $r = 40$	$m = 5$ $r = 80$
chi square	0.1	0.088	0.136	0.177	0.360
	0.5	0.796	0.986	1	1
	0.8	1	1	1	1
Edgeworth	0.1	0.101	0.156	0.227	0.361
	0.5	0.804	0.987	1	1
	0.8	0.998	1	1	1
transformation	0.1	0.085	0.145	0.198	0.358
	0.5	0.809	0.997	1	1
	0.8	0.999	1	1	1
bootstrap	0.1	0.107	0.141	0.198	0.350
	0.5	0.833	0.991	1	1
	0.8	0.999	1	1	1

Table 17: Empirical powers of tests of H_0 (1.5) against H_1 (5.3), with $\bar{\lambda} = 0.1, 0.5, 0.8$, for regression with SAR disturbances (1.1) when h is "bounded" and the disturbances are generated as in (5.4). $\alpha = 5\%$.