Journal of Econometrics 47 (1991) 115–143. North-Holland

Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary or nonstationary AR(1) errors

Jean-Marie Dufour
Université de Montréal, Montréal H3C 3J7, Canada

Maxwell L. King
Monash University, Clayton, Vict. 3168, Australia

Inference on the autocorrelation coefficient ρ of a linear regression model with first-order autoregressive normal disturbances is studied. Both stationary and nonstationary processes are considered. Locally best and point-optimal invariant tests for any given value of ρ are derived. Special cases of these tests include tests for independence and tests for unit-root hypotheses. The powers of alternative tests are compared numerically for a number of selected testing problems and for a range of design matrices. The results suggest that point-optimal tests are usually preferable to locally best tests, especially for testing values of ρ greater than or equal to one.

1. Introduction

The first-order autoregressive [AR(1)] process is one of the most widely used models in econometrics. An important extension is the linear regression model with AR(1) disturbances. In this context, one usually meets the problem of making inferences about the autocorrelation coefficient. This problem can be of interest in itself (e.g., tests of random walk and stationarity

*This work was supported by the Centre de Recherche et Développement en Economique (C.R.D.E., Université de Montréal), CORE (Université Catholique de Louvain), CEPREMAP (Paris), the Australian Research Council, the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds FCAR (Government of Quebec). The authors thank Eric Ghysels, Marc Hallin, Grant Hillier, James Stock, Alain Trognon, William Veloce, Victoria Zinde-Walsh, and two anonymous referees for several useful comments.

0304-4076/91/$03.50 © 1991—Elsevier Science Publishers B.V. (North-Holland)

hypotheses) or can play a role in making inferences about the regression coefficients. In this paper, we develop finite-sample methods for testing whether the autocorrelation coefficient has any given value. We consider the general linear model

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

where y is an $n \times 1$ vector, X is an $n \times k$ matrix of fixed regressors with $\text{rank}(X) = k < n$, β is an unknown parameter vector, and u is an $n \times 1$ vector of disturbances which follow an AR(1) process with normal innovations:

$$u_t = \rho u_{t-1} + \varepsilon_t, \qquad t = 2, \ldots, n, \tag{2}$$

with $\varepsilon_2, \ldots, \varepsilon_n \sim \text{IN}(0, \sigma^2)$, $\sigma^2 > 0$. Further, it is necessary to make assumptions on the value of ρ and the distribution of the initial disturbance u_1. In this paper, we consider two main assumptions:

Assumption A (stationary process). $|\rho| < 1$, $u_1 \sim \text{N}(0, \sigma^2/(1-\rho^2))$, and u_1 is independent of $\varepsilon_2, \ldots, \varepsilon_n$.

Assumption B (unrestricted ρ). $-\infty < \rho < +\infty$ and $u_1 = d_1\varepsilon_1$, where d_1 is unknown, $\varepsilon_1 \sim \text{N}(0, \sigma^2)$, and ε_1 is independent of $\varepsilon_2, \ldots, \varepsilon_n$.

Clearly, Assumption A is a special case of Assumption B. Further, Assumption B includes the case of a random walk ($\rho = 1$) and explosive processes ($|\rho| > 1$) in the disturbances. Though most of our derivations will be based on these two assumptions, we will consider occasionally the more general assumption that u_1 follows an arbitrary distribution.

Assumption C. $-\infty < \rho < +\infty$, u_1 follows an arbitrary distribution with mean zero and u_1 is independent of $\varepsilon_2, \ldots, \varepsilon_n$.

If x_t' denotes the tth row of X and we take $x_t = 1$, (1) and (2) also include as a special case the simple stationary AR(1) process

$$y_t = \mu + \rho y_{t-1} + \varepsilon_t, \qquad t = 2, \ldots, n,$$

where $\mu = (1-\rho)\beta_1$ and $y_1 \sim \text{N}(\beta_1, \sigma^2/(1-\rho^2))$. By taking $x_t = (1, t)'$, we can get (under Assumption B)

$$y_t = \mu_1 + \mu_2 t + \rho y_{t-1} + \varepsilon_t, \qquad t = 2, \ldots, n,$$

where $\mu_1 = (1-\rho)\beta_1 + \rho\beta_2$, $\mu_2 = (1-\rho)\beta_2$, and $y_1 \sim \mathrm{N}(\beta_1 + \beta_2, d_1^2\sigma^2)$. With $\rho = 1$, this yields the random walk model

$$y_t = \beta_2 + y_{t-1} + \varepsilon_t, \qquad t = 2, \ldots, n,$$

where $y_1 \sim \mathrm{N}(\beta_1 + \beta_2, d_1^2\sigma^2)$; the mean and variance of y_1 and the drift coefficient β_2 can take arbitrary values. Finally, if we take $x_t = (1, \rho^t)'$ and $\rho \neq 0$ and 1, we get (under Assumption B)

$$y_t = \mu_1 + \rho y_{t-1} + \varepsilon_t, \qquad t = 2, \ldots, n,$$

where $\mu_1 = (1-\rho)\beta_1$ and $y_1 \sim \mathrm{N}(\beta_1 + \beta_2\rho, d_1^2\sigma^2)$; in contrast with the stationary case, y_1 has here arbitrary mean and variance.

Following Anderson (1948) and Durbin and Watson (1950, 1951), many authors have studied tests of $\rho = 0$; for a survey, see King (1987a). Attention has also been devoted to testing the random walk hypothesis $\rho = 1$, either in linear regressions [Sargan and Bhargava (1983)] or in simpler models [Evans and Savin (1981, 1984), Bhargava (1986), Nankervis and Savin (1985)].[1] On the other hand, very little has been done on testing whether ρ has a pre-specified value, possibly different from zero or one. For example, one may wish to test whether ρ has a value close to but not equal to one. A more basic reason why this problem is important is that tests of $\rho = \rho_0$ can be turned into confidence sets for ρ. Exact tests yield exact confidence sets which can be employed to obtain exact inference procedures (tests and confidence sets) for the regression coefficients [see Dufour (1990)]. In this context, we need efficient tests because the power of the test determines how short the intervals tend to be [see Lehmann (1986, ch. 3)]. Finally, in view of the unreliability of asymptotic critical values [see Park and Mitchell (1980) and Miyazaki and Griffiths (1984)], there is a potentially large benefit from developing finite-sample procedures.

In this paper, we construct optimal invariant tests of H_0: $\rho = \rho_0$ against alternatives of the form H_a^+: $\rho > \rho_0$ and H_a^-: $\rho < \rho_0$, where ρ_0 is any admissible value of ρ. We derive both locally best invariant (LBI) and point-optimal invariant (POI) tests, the latter being constructed as most

[1] Several authors have also studied asymptotic distributions in models with roots equal to or greater than one; see, for example, Anderson (1959), Rao (1978), Dickey and Fuller (1979, 1981), Satchell (1984), Fuller (1985), Phillips (1987a, 1987b), Phillips and Durlauf (1986), and the survey of Diebold and Nerlove (1988). In this paper, we concentrate on finite-sample procedures.

powerful invariant (MPI) tests against nonlocal point alternatives. We also discuss how to obtain two-sided tests against H_a: $\rho \neq \rho_0$. In section 2, we consider the stationary case (Assumption A), while the case of a nonstationary process (Assumption B) is studied in section 3. For all the statistics considered, we explain how exact critical values can be calculated. For the LBI test of $\rho = 1$, we also show that critical values may be obtained from tables of the central F distribution. We give expressions for the LBI test statistics as exponentially weighted averages of residual autocorrelations. In the nonstationary case, we stress the importance of dealing carefully with the distribution of the first disturbance, u_1. By considering tests invariant under appropriate transformation groups, we find LBI and POI tests whose null distributions are not influenced by the distribution of u_1. In sections 4 and 5, we report the results of power comparisons between LBI tests, POI tests, and Durbin–Watson (DW) tests based on appropriately transformed data. Among other things, the results suggest that POI tests can lead to substantial power improvements over alternative tests. The advantage of POI tests over LBI tests is especially striking when values of ρ equal to or greater than one are tested. Section 6, finally, summarizes our results and contains a few concluding remarks.

2. Stationary disturbances

In this section, we assume a stationary error process (Assumption A) and consider testing H_0: $\rho = \rho_0$ against the alternatives H_a^+: $\rho > \rho_0$, H_a^-: $\rho < \rho_0$, and H_a: $\rho \neq \rho_0$, ρ_0 is arbitrary and such that $|\rho_0| < 1$. We denote the problem of testing H_0 by PA(ρ_0). Under Assumption A, $u \sim N(0, \sigma^2\Sigma(\rho))$, where

$$\Sigma(\rho) = 1/(1-\rho^2)\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & & & \vdots \\ \rho^2 & \rho & 1 & & & \vdots \\ \vdots & & & \ddots & & \vdots \\ & & & & 1 & \rho \\ \rho^{n-1} & \cdots & \cdots & \cdots & \rho & 1 \end{bmatrix}.$$

The above problems are invariant to transformations of the form

$$y^{\ddagger} = \gamma_0 y + X\gamma, \tag{G1}$$

where $\gamma_0 > 0$ and γ is $k \times 1$. G1 is the transformation group used by Durbin and Watson (1971) to establish optimal properties of the DW test. This suggests studying tests of $\rho = \rho_0$ that are invariant under G1. For a general discussion of invariant tests, see Lehmann (1986).

Let

$$B(\rho) = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & \cdots & 0 \\ -\rho & 1 & 0 & & & 0 \\ 0 & -\rho & 1 & & & \vdots \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & 1 & 0 \\ 0 & 0 & \cdots & \cdots & -\rho & 1 \end{bmatrix},$$

$B_0 = B(\rho_0)$, and $\Sigma_0 = \Sigma(\rho_0)$. We can transform (1) by premultiplying by the nonsingular transformation B_0, so that

$$y^* = X^*\beta + u^*, \tag{3}$$

where $y^* = B_0 y$, $X^* = B_0 X$, $u^* = B_0 u$, and under H_0, $u^* \sim N(0, \sigma^2 I_n)$.

When $\rho \neq \rho_0$, $u^* \sim N(0, \sigma^2 B_0 \Sigma(\rho) B_0')$ and the transformed disturbances follow the ARMA(1,1) scheme, $u_t^* - \rho u_{t-1}^* = \varepsilon_t - \rho_0 \varepsilon_{t-1}$, $t = 2, \dots, n$. Also note that if G1 is transformed by premultiplying by B_0 we get

$$(y^*)^{\ddagger} = \gamma_0 y^* + X^*\gamma, \tag{4}$$

so that G1 and (4) are equivalent groups of transformations. The disturbances of (3) remain autocorrelated when $\rho \neq \rho_0$ which suggests that $\rho = \rho_0$ may be tested by checking whether the residuals of (3) are independent; see Dufour (1990). Any test for first-order autocorrelation in (3) may in principle be employed. Though they may have computational advantages, these procedures have no known optimal properties. For this reason, we study here tests with clear optimal properties against local and nonlocal alternatives. Note that King and Evans (1988) have shown that the DW test is approximately uniformly LBI against ARMA(1,1) disturbances. This is for H_0: $\rho = \phi = 0$ in $u_t - \rho u_{t-1} = \varepsilon_t - \phi \varepsilon_{t-1}$. Here we are interested in a different problem which involves testing H_0: $\rho = \phi$ when ϕ is *known*.

Theorem 1, which follows from King and Hillier (1985) [see Shively, Ansley, and Kohn (1989) for an alternative statement of the King–Hillier result], gives LBI tests of $\rho = \rho_0$ against one-sided alternatives.[2]

[2] Detailed proofs of the theorems that follow are available in an earlier version of this paper [Dufour and King (1989)].

Theorem 1. *Under (1), (2), and Assumption A, a LBI test of $\rho = \rho_0$ against $\rho > \rho_0$ ($\rho < \rho_0$) is to reject H_0 for small (large) values of*

$$D_1(\rho_0) = \hat{e}' A_0 \hat{e} / \hat{e}' \Sigma_0^{-1} \hat{e}, \tag{5}$$

where $\hat{e}$ is the generalized least-squares (GLS) residual vector from (1) corresponding to covariance matrix Σ_0 and

$$A_0 = \partial \Sigma(\rho)^{-1} / \partial \rho \big|_{\rho = \rho_0} = -2(1 - \rho_0) I_n + A_1 - 2\rho_0 C_1,$$

in which A_1 and C_1 are the $n \times n$ matrices

$$A_1 = \begin{bmatrix} 2 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & & & 0 \\ 0 & -1 & 2 & & & \vdots \\ \vdots & & & \ddots & 2 & -1 \\ 0 & 0 & \cdots & \cdots & -1 & 2 \end{bmatrix} \tag{6}$$

and

$$C_1 = \operatorname{diag}(1, 0, 0, \ldots, 0, 1). \tag{7}$$

Note that because $B_0 \hat{e} = z^*$, where z^* is the OLS residual vector from the transformed model (3), (5) can also be written as

$$\begin{aligned} D_1(\rho_0) &= z^{*\prime} (B_0^{-1})' A_0 B_0^{-1} z^* / z^{*\prime} z^* \\ &= u^{*\prime} M_0 (B_0^{-1})' A_0 B_0^{-1} M_0 u^* / u^{*\prime} M_0 u^*, \end{aligned}$$

where $M_0 = I_n - X^*(X^{*\prime} X^*)^{-1} X^{*\prime}$. Under H_0, $u^* \sim N(0, \sigma^2 I_n)$ so that $D_1(\rho_0)$ is a ratio of quadratic forms in normal variables and its distribution function can be computed using numerical methods developed for the DW test such as described by King (1987a, pp. 27–28) and Shively, Ansley, and Kohn (1989).

For the case $\rho_0 = 0$, the statistic $D_1(\rho_0)$ takes the form

$$D_1(0) = z' A_1 z / z' z - 2 = -2 r_1,$$

where z is the OLS residual vector from (1) and r_1 is the first-order autocorrelation coefficient of these residuals. More generally, it is possible to express $D_1(\rho_0)$ in a more intuitive form. After some tedious algebra (see appendix A), we can find two alternative expressions. The first relates $D_1(\rho_0)$

to the first-order autocorrelation of the GLS residuals $\hat{e}$:

$$D_1(\rho_0) = -2q[R_1(\rho_0) - \rho_0], \tag{8}$$

where

$$R_1(\rho_0) = \sum_{t=1}^{n-1} \hat{e}_t \hat{e}_{t+1} \Big/ \sum_{t=2}^{n-1} \hat{e}_t^2 \quad \text{and} \quad q = \sum_{t=2}^{n-1} \hat{e}_t^2 / z^{*\prime} z^*.$$

$R_1(\rho_0)$ is approximately the first-order autocorrelation of $\hat{e}$, while q can be viewed as an estimator of $\operatorname{var}(u_t)/\sigma^2 = 1/(1-\rho_0^2)$. When $\rho = \rho_0$, q converges to $1/(1-\rho_0^2)$ under fairly general regularity conditions. Roughly, H_0 is rejected against $\rho > \rho_0$ ($\rho < \rho_0$) when $R_1(\rho_0) - \rho_0$ is large (small). The second expression relates $D_1(\rho_0)$ to the autocorrelations of the OLS residuals, z^*, from the transformed model (3):

$$D_1(\rho_0) = -2 \sum_{k=1}^{n-1} \rho_0^{k-1} r_k^* + \eta, \tag{9}$$

where

$$r_k^* = \sum_{t=1}^{n-k} z_t^* z_{t+k}^* \Big/ z^{*\prime} z^*, \qquad 0^0 \equiv 1,$$

and

$$\eta = -2\left[\rho_0(1-\rho_0^2)^{-1}(z_1^*)^2 + \left\{(1-\rho_0^2)^{-1/2} - 1\right\} z_1^* \sum_{t=2}^{n} \rho_0^{t-2} z_t^*\right] \Big/ z^{*\prime} z^*.$$

In large samples, η is negligible (under standard regularity conditions on X), so that $D_1(\rho_0)$ is proportional to an exponentially weighted average of all the autorcorrelations, r_k^*. When $\rho_0 = 0$, (9) reduces to $-2r_1$, but not otherwise. This shows that looking only at low-order autocorrelations is not generally efficient.

While LBI tests have optimal power in the neighbourhood of H_0, they may have poor power away from H_0, i.e., when accepting H_0 is most damaging from the point of view of making reliable inferences. Power can even fall below the level of the test; see, for example, Krämer (1985). An attractive alternative is to use a point-optimal test, i.e., a test that optimizes power at a pre-determined point under the alternative hypothesis. The next theorem, which follows directly from King (1980), gives the MPI test of $\rho = \rho_0$ against a given alternative $\rho = \rho_1$, where $|\rho_0| < 1$, $|\rho_1| < 1$, and $\Sigma_1 = \Sigma(\rho_1)$.

Theorem 2. *Under (1), (2), and Assumption A, a MPI test of $\rho = \rho_0$ against $\rho = \rho_1$ is to reject H_0 for small values of*

$$S_1(\rho_0, \rho_1) = \tilde{e}'\Sigma_1^{-1}\tilde{e}/\hat{e}'\Sigma_0^{-1}\hat{e}, \tag{10}$$

where $\hat{e}$ and $\tilde{e}$ are the GLS residual vectors from (1) corresponding to covariance matrices Σ_0 and Σ_1, respectively.

To obtain a test of $\rho = \rho_0$ against $\rho > \rho_0$, we select a value of ρ_1 such that $\rho_0 < \rho_1 < 1$ and apply the test based on $S_1(\rho_0, \rho_1)$. For example, we may choose ρ_1 close to one or an intermediate value like $\rho_1 = (\rho_0 + 1)/2$. Similarly, against $\rho < \rho_0$, we select ρ_1 such that $-1 < \rho_1 < \rho_0$. Tests obtained in this way optimize power at $\rho = \rho_1$ and are known as POI tests. A survey by King (1987b) reveals that such tests often have substantially better power than LBI tests. Let $B_1 = B(\rho_1)$ and let

$$y^{\dagger} = X^{\dagger}\beta + u^{\dagger} \tag{11}$$

denote (1) transformed by premultiplying by B_1. The statistic $S_1(\rho_0, \rho_1)$ is easy to compute because it can also be written as $S_1(\rho_0, \rho_1) = z^{\dagger\prime}z^{\dagger}/z^{*\prime}z^{*}$, where $z^{\dagger}$ and z^{*} are the OLS residual vectors from (11) and (3), respectively. Further, we can write

$$S_1(\rho_0, \rho_1) = u^{*\prime}\left[\{B_1 B_0^{-1}\}' M_1 B_1 B_0^{-1}\right]u^{*}/u^{*\prime}M_0 u^{*},$$

where $u^{*} \sim \mathrm{N}(0, \sigma^2 I_n)$ under H_0 and $M_1 = I_n - X^{\dagger}(X^{\dagger\prime}X^{\dagger})^{-1}X^{\dagger\prime}$. Thus $S_1(\rho_0, \rho_1)$ is a ratio of quadratic forms in normal variables and its distribution function can be computed in a similar way to the DW test.

Theorems 1 and 2 describe tests against one-sided alternatives. To obtain confidence sets, one typically needs two-sided tests of $\rho = \rho_0$ against $\rho \neq \rho_0$. It is certainly possible to obtain LBI unbiased tests for this problem [see King and Hillier (1985)], although the test criterion does not reduce to a ratio of quadratic forms in normal variables. Instead, one gets forms of order 4 whose finite-sample distribution is unknown. For this reason, we suggest the combining of optimal one-sided tests.

Let $0 < \alpha_1 < 1$ and $0 < \alpha_2 < 1$ such that $\alpha_1 + \alpha_2 = \alpha$, e.g., $\alpha_1 = \alpha_2 = \alpha/2$. Using LBI tests, it is natural to reject H_0: $\rho = \rho_0$ against H_a: $\rho \neq \rho_0$ when

$$D_1(\rho_0) < c_1 \quad \text{or} \quad D_1(\rho_0) > c_1',$$

where c_1 and c_1' are chosen so that $\mathrm{P}[D_1(\rho_0) < c_1] = \alpha_1$ and $\mathrm{P}[D_1(\rho_0) > c_1'] = \alpha_2$ under H_0. Clearly this test has level α. Similarly, it is also possible to construct two-sided tests from POI tests. Choose ρ_1 and ρ_2 so that $-1 < \rho_1 < \rho_0 < \rho_2 < 1$. We reject H_0 against H_a when

$$S_1(\rho_0, \rho_1) < c_2 \quad \text{or} \quad S_1(\rho_0, \rho_2) < c_2',$$

where c_2 and c_2' are chosen so that $P[S_1(\rho_0, \rho_1) < c_2] = \alpha_1$ and $P[S_1(\rho_0, \rho_2) < c_2'] = \alpha_2$ under H_0. By the Bonferroni inequality, this test has level less than or equal to α.

3. Nonstationary disturbances

Assumption A is restrictive because it excludes $\rho = 1$ or $|\rho| > 1$ and requires the variance of u_1 to be $\sigma^2/(1-\rho^2)$. Even if $|\rho| < 1$, we may wish to allow more flexibility for the distribution of u_1. For example, the process may not have run long enough to become stationary. To obtain exact tests, it is important to take these difficulties into account. In this section, we do not impose any restriction on the value of ρ and simply assume that u_1 follows a normal distribution with arbitrary unknown variance (Assumption B). The normality assumption of u_1 is used mainly to derive tests with clear optimal properties. However, the procedures obtained in this way have correct sizes under weaker conditions (Assumption C).

Let $C(\rho)$ denote $B(\rho)$, with the top left element taking the value 1 instead of $\sqrt{1-\rho^2}$ and let J, E_1, and E_n be $n \times n$ matrices defined as $J = \text{diag}(d_1, 1, 1, \ldots, 1)$, $E_1 = \text{diag}(1, 0, 0, \ldots, 0)$, and $E_n = \text{diag}(0, 0, \ldots, 0, 1)$. Under Assumption B, $C(\rho)u = J\varepsilon$, where $\varepsilon \sim N(0, \sigma^2 I_n)$ so that $u \sim N(0, \sigma^2 \Omega(\rho, d_1))$ in which $\Omega(\rho, d_1) = C(\rho)^{-1} J^2 [C(\rho)^{-1}]'$. Observe that

$$\begin{aligned}\Omega(\rho, d_1)^{-1} &= C(\rho)' J^{-2} C(\rho) \\ &= (1-\rho)^2 I_n + \rho A_1 - \rho^2 E_n + \left(d_1^{-2} - 1\right) E_1.\end{aligned}$$

If $d_1^2 = (1-\rho^2)$, $\Omega(\rho, d_1) = \Sigma(\rho)$, while if $d_1 = 1$, the covariance matrix of u is identical to that used by Berenblut and Webb (1973). The assumption $d_1 = 1$ is, however, very stringent and usually implausible. We thus prefer to use $\Omega(\rho, d_1)$ with d_1 taken as unknown.

Under (1), (2), and Assumption B, $y \sim N(X\beta, \sigma^2 \Omega(\rho, d_1))$. We denote the problem of testing H_0: $\rho = \rho_0$ against H_a^+: $\rho > \rho_0$ or H_a^-: $\rho < \rho_0$ in this context by $PB(\rho_0)$. It is invariant to transformations in the group G1. Theorem 3, which again follows directly from King and Hillier (1985) and King (1980), gives LBI tests and POI tests of $\rho = \rho_0$ assuming d_1 is known. Even though such tests are rarely applicable, they provide useful benchmarks in the power comparisons that follow.

Theorem 3. *Let (1), (2), and Assumption B hold with the exception that d_1 is assumed known and $d_1 \neq 0$. A LBI test of $\rho = \rho_0$ against $\rho > \rho_0$ ($\rho < \rho_0$) is to reject H_0 for small (large) values of*

$$\bar{D}_1(\rho_0) = \hat{u}' A_2 \hat{u} / \hat{u}' \Omega(\rho_0, d_1)^{-1} \hat{u},$$

where $\hat{u}$ is the GLS residual vector from (1) corresponding to covariance matrix $\Omega(\rho_0, d_1)$ and

$$A_2 = -2(1-\rho_0) I_n + A_1 - 2\rho_0 E_n. \tag{12}$$

A POI test of $\rho = \rho_0$ against $\rho > \rho_0$ ($\rho < \rho_0$) that optimizes power at $\rho = \rho_1$ is to reject H_0 for small values of $\bar{S}_1(\rho_0, \rho_1) = \tilde{u}' \Omega(\rho_1, d_1)^{-1} \tilde{u} / \hat{u}' \Omega(\rho_0, d_1)^{-1} \hat{u}$, where $\tilde{u}$ is the GLS residual vector from (1) corresponding to covariance matrix $\Omega(\rho_1, d_1)$.

The $\bar{D}_1(\rho_0)$ and $\bar{S}_1(\rho_0, \rho_1)$ tests can be implemented like the tests of Theorems 1 and 2. In particular, both statistics can be written in forms involving OLS residuals from the transformed regression

$$J^{-1} C(\rho) y = J^{-1} C(\rho) X + J^{-1} C(\rho) u, \tag{13}$$

in which ρ takes the value ρ_0 or ρ_1. In practice, d_1 is usually unknown. To deal with this problem, we will try to find test statistics that satisfy two conditions:

(a) The value of d_1 is not required to compute the test statistic.
(b) The null distribution of the test statistic does not depend on d_1.

Let $C_0 = C(\rho_0)$ and consider the transformed regression model

$$C_0 y = C_0 X \beta + C_0 u. \tag{14}$$

Observe that under H_0 and Assumption B, $C_0 y \sim N(C_0 X\beta, \sigma^2 J^2)$. Thus, if we consider testing H_0: $\rho = \rho_0$ in the context of (14), conditions (a) and (b) above would be satisfied if the test statistic does not depend on d_1 and is invariant to the value of the first element of $C_0 u$. The testing problem expressed in terms of (14) is invariant to transformations of the form

$$(C_0 y)^{\ddagger} = \gamma_0 C_0 y + C_0 X \gamma, \tag{15}$$

where γ_0 is a positive scalar and γ is a $k \times 1$ vector. To make the test statistic invariant to the first element of $C_0 u$, it is sufficient to consider a statistic that is invariant to transformations of the form

$$(C_0 y)^{\ddagger} = \gamma_0 C_0 y + C_0 X\gamma + \gamma_{k+1} l_1, \tag{G2}$$

where γ_{k+1} is an arbitrary scalar and l_1 is the $n \times 1$ vector of zeros with one as the first element. Note that we are not claiming that the testing problem is invariant under G2.[3] However, we wish to consider tests invariant under G2 because such tests are invariant under (15) *and* satisfy (b).[4] In terms of y, transformations in the group G2 are equivalent to transformations of the form $y^{\ddagger} = \gamma_0 y + X\gamma + \gamma_{k+1} C_0^{-1} l_1$. Further, $C_0^{-1} l_1 = (1, \rho_0, \ldots, \rho_0^{n-1})'$. Thus tests invariant under G2 are invariant to transformations of the form $y_t^{\ddagger} = y_t + \gamma_{k+1}\rho_0^{t-1}$, $t = 1, \ldots, n$. The results of the test should not change when a solution of the homogeneous equation $Z_t - \rho_0 Z_{t-1} = 0$ is added to y.

A maximal invariant for the group G2 is $v/(v'v)^{1/2}$, where v is the OLS residual vector from the regression of $C_0 y$ on $[C_0 X, l_1]$. One can get v by adding a dummy variable for the first observation into the transformed model (14), or by introducing the regressor $(1, \rho_0, \ldots, \rho_0^{n-1})'$ into (1). Note that the extra regressor is not used because we believe it is a regressor in the model but simply to find a test that satisfies (b).

The matrix $[X, C_0^{-1} l_1]$ may not always have full column rank; an example being when $\rho_0 = 1$ and X contains a constant regressor. Let $\bar{X}$ denote $[X, C_0^{-1} l_1]$ where, if necessary, columns have been deleted until it has full column rank and let $\bar{X}_0 = C_0 \bar{X}$ denote $[C_0 X, l_1]$ with the analogous columns deleted. Then

$$v = \bar{M} C_0 y = \bar{M}\bar{u}, \tag{16}$$

where $\bar{M} = I_n - \bar{X}_0(\bar{X}_0'\bar{X}_0)^{-1}\bar{X}_0'$ and $\bar{u} = C_0 u$. Theorem 4, which can be proved along similar lines to King and Hillier's (1985) result, gives LBI tests of $\rho = \rho_0$ against one-sided alternatives.

Theorem 4. Under (1), (2), Assumption B, and assuming $d_1 \neq 0$, a LBI test of $\rho = \rho_0$ against $\rho > \rho_0$ ($\rho < \rho_0$) under the transformation group G2 is to reject H_0 for small (large) values of

$$D_2(\rho_0) = v'(C_0^{-1})' A_2 C_0^{-1} v / v'v = \tilde{v}' A_2 \tilde{v} / \tilde{v}'\Omega(\rho_0, 1)^{-1}\tilde{v}, \tag{17}$$

[3] A referee has pointed out to us that, because the normal distribution is complete, this is a necessary as well as sufficient condition provided the mean of the first element of $C_0 y$ is unknown.

[4] A similar technique was used by Dufour and Dagenais (1985) to obtain optimal autocorrelation tests in regression models with missing observations.

where A_2 is given by (12) and $\tilde{v} = C_0^{-1}v$ is the GLS residual vector from

$$y = \bar{X}\delta + u, \tag{18}$$

assuming covariance matrix $\Omega(\rho_0, 1)$.

Critical values and critical levels for tests based on $D_2(\rho_0)$ can be obtained by noting that

$$D_2(\rho_0) = \bar{u}'\bar{M}(C_0^{-1})'A_2C_0^{-1}\bar{M}\bar{u}/\bar{u}'\bar{M}\bar{u},$$

where $\bar{u} \sim \mathrm{N}(0, \sigma^2 J^2)$ when $\rho = \rho_0$. As a consequence of l_1 being a regressor, the first row and column of $\bar{M}$ are zero so that $v = \bar{M}\bar{u}$ is not a function of $\bar{u}_1 = u_1$, and hence under H_0 the distribution of $D_2(\rho_0)$ does not depend on d_1 or indeed on the distribution of u_1 (Assumption C). The test is thus applicable, in the sense that its level is correct, under the weaker Assumption C. Furthermore, when computing critical values and critical levels, one can assume $d_1 = 1$ or $\bar{u} \sim \mathrm{N}(0, \sigma^2 I_n)$. Also observe that $v_1 = 0$.

As in the stationary case, we can rewrite $D_2(\rho_0)$ in the form of (8) where now

$$R_1(\rho_0) = \sum_{t=2}^{n-1} \tilde{v}_t\tilde{v}_{t+1} \Big/ \sum_{t=2}^{n-1} \tilde{v}_t^2 \quad \text{and} \quad q = \sum_{t=2}^{n-1} \tilde{v}_t^2/v'v.$$

Also $D_2(\rho_0) = -2\sum_{k=1}^{n-1}\rho_0^{k-1}\tilde{r}_k$ where $\tilde{r}_k = \sum_{t=1}^{n-k} v_t v_{t+k}/v'v$. Thus $D_2(\rho_0)$ is proportional to an exponentially weighted average of all the autocorrelations $\tilde{r}_k$.

For $\rho_0 = 0$, $D_2(0) = -2\tilde{r}_1$ where $\tilde{r}_1$ is the first-order autocorrelation of the OLS residuals from the regression $y = X\beta + \beta_{k+1}l_1 + u$. For $\rho_0 = 1$ (random walk), the test statistic takes the form

$$D_2(1) = -2\sum_{k=1}^{n-1} \tilde{r}_k. \tag{19}$$

We reject $\rho = 1$ against $\rho > 1$ ($\rho < 1$) when the sum of all the autocorrelations of the OLS residuals is large (small). Critical values of $D_2(1)$ may be obtained from tables of the central F distribution. Provided $\bar{M}l_n \neq 0$,

$$1 - D_2(1) \sim \left[l_n'\bar{M}l_n\right]/\left[1 + (n - k_1 - 1)F(n - k_1 - 1, 1)\right],$$

where $l_n = (1, 1, \ldots, 1)'$, $k_1 = \mathrm{rank}[\bar{X}]$, and $F(n - k_1 - 1, 1)$ is a random variable that follows a central F distribution with $(n - k_1 - 1, 1)$ degrees of freedom; see appendix B for the proof. When $\bar{M}l_n = 0$, we have $D_2(1) = 1$ and the test is not applicable. This corresponds to a linear trend regressor in (1).

It is worth noting that Sargan and Bhargava (1983) considered the related problem of testing $\rho_0 = 1$ under Assumption B against the alternative hypothesis of Assumption A. As well as suggesting the use of the DW and Berenblut and Webb (1973) tests applied to (1) in a first-differenced form, they derived an approximately LBI test that is also approximately uniformly most powerful when the column space of X is spanned by k eigenvectors of $(A_1 - C_1)$ where A_1 and C_1 are given by (6) and (7), respectively.

As in the stationary case, LBI tests are not necessarily optimal against nonlocal alternatives. Theorem 5 allows the construction of POI tests that are invariant under G2 and thus have correct size for any d_1 value.

Theorem 5. Under (1), (2), Assumption B, and assuming $d_1 \neq 0$, a MPI test at $\rho = \rho_0$ against $\rho = \rho_1$, under the transformation group G2, is to reject H_0 for small values of

$$S_2(\rho_0, \rho_1, d_1) = \hat{v}'\Omega(\rho_1, d_1)^{-1}\hat{v}/\tilde{v}'\Omega(\rho_0, 1)^{-1}\tilde{v} = \hat{v}^{*\prime}\hat{v}^*/v'v,$$

where $\hat{v}$ and $\tilde{v}$ are the GLS residual vectors from (18) assuming covariance matrices $\Omega(\rho_1, d_1)$ and $\Omega(\rho_0, 1)$, respectively, and v and $\hat{v}^*$ are the OLS residual vectors defined by (16) and the regression of $J_n^{-1}C(\rho_1)y$ on $J_n^{-1}C(\rho_1)\bar{X}$, respectively. Further, if $\rho = \rho_0$ and d_1^* is any real number such that $d_1^* \neq 0$, the null distribution of $S_2(\rho_0, \rho_1, d_1^*)$ does not depend on d_1.

The proof of the first part of Theorem 5 follows directly from King (1980, 1987b) and the fact that $v/(v'v)^{1/2}$ is a maximal invariant for the group G2. Let $Q = J^{*-1}C(\rho_1)C_0^{-1}$, in which J^* denotes J with $d_1 = d_1^*$, and define $M_0 = I_n - Q\bar{X}_0(\bar{X}_0'Q'Q\bar{X}_0)^{-1}\bar{X}_0'Q'$. Then we can write

$$S_2(\rho_0, \rho_1, d_1^*) = \bar{u}'\bar{M}Q'M_0Q\bar{M}\bar{u}/\bar{u}'\bar{M}\bar{u} = v'Q'M_0Qv/v'v, \qquad (20)$$

so that $S_2(\rho_0, \rho_1, d_1^*)$ depends on u only through $v = \bar{M}\bar{u}$. When $\rho = \rho_0$, $S_2(\rho_0, \rho_1, d_1^*)$ is not a function of $\bar{u}_1 = u_1$ given that the first row and column of $\bar{M}$ are zero.

When applying Theorem 5, one needs to specify both ρ_1 and a value of d_1, say d_1^*. The resultant test is POI, optimizing power at $(\rho, d_1)' = (\rho_1, d_1^*)'$. It is not necessary that d_1^* be the true value of d_1 to get a valid test. Critical values can be computed from (20) in the usual way because we can assume $\bar{u} \sim N(0, \sigma^2 I_n)$ when $\rho = \rho_0$. As for the LBI test, this test remains applicable under Assumption C because it depends on u only through $v = \bar{M}\bar{u}$. For the special case in which $\rho_1 = 1$ and the model contains an intercept, $C(\rho_1)\bar{X}$

contains l_1 as a regressor so that $Q\bar{X}_0$, $Q'M_0Q$, and thus $S_2(\rho_0, 1, d_1^*)$ are invariant to the value of d_1^* selected. In this case $S_2(\rho_0, 1, d_1^*) = S_2(\rho_0, 1, d_1)$ and the choice of d_1^* is irrelevant.

The determination of d_1^* in $S_2(\rho_0, \rho_1, d_1^*)$ is usually arbitrary so it would be convenient to have a test that does not require specifying a value of d_1^*. A way to do this is to consider a larger invariance group such that the MPI test of $\rho = \rho_0$ against $\rho = \rho_1$ does not depend on d_1. Such a group is

$$y^{\ddagger} = \gamma_0 y + X\gamma + \gamma_{k+1} C(\rho_0)^{-1} l_1 + \gamma_{k+2} C(\rho_1)^{-1} l_1, \tag{G3}$$

where $\gamma_0, \gamma_{k+1}, \gamma_{k+2}$ are arbitrary scalars, such that $\gamma_0 > 0$ and γ is a $k \times 1$ vector. It can be shown that the MPI test of $\rho = \rho_0$ against $\rho = \rho_1$ under the transformation group G3 does not require specifying an arbitrary value of d_1 and has a null distribution that does not depend on the true value of d_1. However, power comparisons suggested that the power of this procedure is very inferior to that of the test based on $S_2(\rho_0, \rho_1, d_1^*)$. For this reason, we do not elaborate here on POI tests under G3.

Theorems 3, 4, and 5 all provide optimal one-sided tests for testing $\rho = \rho_0$. Corresponding two-sided tests may be obtained in a way analogous to the one used in the stationary case.

4. Empirical power comparisons

In order to study the small-sample properties of the above tests, their powers were calculated for testing problems *PA*(0.5), *PA*(0.9), *PB*(0.5), *PB*(0.9), *PB*(1.0), *PB*(1.1), and design matrices:

X1: ($n \times 1$; $n = 20, 60$). The constant dummy as the only regressor.

X2: ($n \times 3$; $n = 20, 60$). The first n observations of Durbin and Watson's (1951, p. 159) consumption of spirits example.

X3: ($n \times 3$; $n = 20, 60$). A constant, the quarterly Australian Consumer Price Index commencing 1959(1) and the same index lagged one quarter.

X4: ($n \times 4$; $n = 20, 60$). A constant, quarterly Australian private capital movements, the same series lagged one quarter and quarterly Australian Government capital movements commencing 1968(1).

X5: ($n \times 3$; $n = 20, 60$). Watson's X matrix with an intercept, i.e., a_1, $(a_2 + a_n)/\sqrt{2}$, $(a_3 + a_{n-1})/\sqrt{2}$ as regressors where $a_1, \ldots, a_n$ are the eigenvectors corresponding to the eigenvalues of the DW matrix $(A_1 - C_1)$ arranged in ascending order where A_1 and C_1 are defined by (6) and (7).

These design matrices cover a range of applications. *X1* is the special case of the Gaussian time-series model with unknown mean. *X2* is based on annual data, while *X3* is quarterly with a slight seasonal pattern. The two

Table 1

Summary of the empirical power comparison.

Problem	Alternative hypothesis	Tests	Values of ρ at which power computed
$PA(0.5)$	H_a^+	$DW_1(0.5)$, $D_1(0.5)$, $S_1(0.5, 0.75)$, $S_1(0.5, 0.999)$.	0.6, 0.7, 0.8, 0.9, 0.999.
$PA(0.5)$	H_a^-	$DW_1(0.5)$, $D_1(0.5)$, $S_1(0.5, 0.25)$, $S_1(0.5, 0)$.	0.4, 0.3, 0.2, 0.1, 0.
$PA(0.9)$	H_a^+	$DW_1(0.9)$, $D_1(0.9)$, $S_1(0.9, 0.95)$, $S_1(0.9, 0.999)$.	0.92, 0.94, 0.96, 0.98, 0.999.
$PA(0.9)$	H_a^-	$DW_1(0.9)$, $D_1(0.9)$, $S_1(0.9, 0.45)$, $S_1(0.9, 0)$.	0.8, 0.7, 0.6, 0.3, 0.
$PB(0.5)$	H_a^+	$\overline{DW}_1(0.5)$, $\bar{D}_1(0.5)$, $\bar{S}_1(0.5, 0.75)$, $\bar{S}_1(0.5, 1.0)$, $DW_2(0.5)$, $D_2(0.5)$, $S_2(0.5, 0.75, d_1^*)$, $S_2(0.5, 1.0, d_1^*)$.	0.6, 0.7, 0.8, 0.9, 1.0.
$PB(0.5)$	H_a^-	$\overline{DW}_1(0.5)$, $\bar{D}_1(0.5)$, $\bar{S}_1(0.5, 0.25)$, $\bar{S}_1(0.5, 0)$, $DW_2(0.5)$, $D_2(0.5)$, $S_2(0.5, 0.25, d_1^*)$, $S_2(0.5, 0, d_1^*)$.	0.4, 0.3, 0.2, 0.1, 0.
$PB(0.9)$	H_a^+	$\overline{DW}_1(0.9)$, $\bar{D}_1(0.9)$, $\bar{S}_1(0.9, 0.95)$, $\bar{S}_1(0.9, 1.0)$, $DW_2(0.9)$, $D_2(0.9)$, $S_2(0.9, 0.95, d_1^*)$, $S_2(0.9, 1.0, d_1^*)$.	0.92, 0.94, 0.96 0.98, 1.0.
$PB(0.9)$	H_a^-	$\overline{DW}_1(0.9)$, $\bar{D}_1(0.9)$, $\bar{S}_1(0.9, 0.45)$, $\bar{S}_1(0.9, 0)$, $DW_2(0.9)$, $D_2(0.9)$, $S_2(0.9, 0.45, d_1^*)$, $S_2(0.9, 0, d_1^*)$.	0.8, 0.7, 0.6, 0.3, 0.
$PB(1.0)$	H_a^+	$DW_2(1.0)$, $D_2(1.0)$, $S_2(1.0, 1.1, d_1^*)$, $S_2(1.0, 1.2, d_1^*)$.	1.025, 0.05, 1.1 1.115, 1.2.
$PB(1.0)$	H_a^-	SB, $DW_2(1.0)$, $D_2(1.0)$, $S_2(1.0, 0.5, d_1^*)$, $S_2(1.0, 0, d_1^*)$.	0.9, 0.75, 0.5, 0.25, 0.
$PB(1.1)$	H_a^+	$\overline{DW}_1(1.1)$, $\bar{D}_1(1.1)$, $\bar{S}_1(1.1, 1.15)$, $\bar{S}_1(1.1, 1.2)$, $DW_2(1.1)$, $D_2(1.1)$, $S_2(1.1, 1.15, d_1^*)$, $S_2(1.1, 1.2, d_1^*)$.	1.12, 1.14, 1.16, 1.18, 1.2.
$PB(1.1)$	H_a^-	$\overline{DW}_1(1.1)$, $\bar{D}_1(1.1)$, $\bar{S}_1(1.1, 1.0)$, $\bar{S}_1(1.1, 0.5)$, $DW_2(1.1)$, $D_2(1.1)$, $S_2(1.1, 1.0, d_1^*)$, $S_2(1.1, 0.5, d_1^*)$.	1.05, 1.0, 0.9, 0.5, 0.

capital movement series which make up *X4* are strongly seasonal with two seasonal peaks per year plus some large fluctuations. Watson (1955) found that within the class of orthogonal X matrices, OLS has minimum efficiency relative to the BLUE for *X5*. We therefore expect *X5* to show an extreme in the behaviour of the tests.

A summary of the tests and the ρ values at which their powers were calculated is given in table 1. For $PA(\rho_0)$ and against H_a^+ (H_a^-), $DW_1(\rho_0)$ denotes the one-sided DW test against positive (negative) autocorrelation applied to (3). In the case of $PB(\rho_0)$, two sets of power comparisons were made. The first involved applying each test assuming d_1 is known. This is unlikely to happen in practice, so these powers were only calculated to provide benchmarks. The tests in this case are based on $\overline{D}_1(\rho_0)$, $\overline{S}_1(\rho_0, \rho_1)$, and $\overline{DW}_1(\rho_0)$ which is the DW statistic applied to (13) with $\rho = \rho_0$. These tests are invariant under G1, but not necessarily under G2. The second set involved applying the tests $D_2(\rho_0)$, $S_2(\rho_0, \rho_1, d_1^*)$, and $DW_2(\rho_0)$ constructed to be invariant under G2.[5] This required $(1, \rho_0, \ldots, \rho_0^{n-1})'$ to be added as an additional regressor to (1) before the test was applied. For $PB(1.0)$, the two sets of tests are identical (provided the model contains an intercept). Also in this case, Sargan and Bhargava's (1983) approximately LBI test denoted SB was also included in the comparison. For the remaining $PB(\rho_0)$ testing problems except when $\rho_1 = 0$, the $S_2(\rho_0, \rho_1, d_1^*)$ tests require a choice of d_1^* value. The values used were $d_1^* = 0.1, 1.0, 10.0$. Although powers vary with d_1, in order to keep the computations manageable, all calculations were performed with $d_1 = 1.0$.

All test statistics can be expressed as ratios of quadratic forms of the disturbance vector, being of the form $u'Au/u'Bu$, where A and B are known $n \times n$ matrices. In order to calculate exact critical values and powers of tests based on such statistics, we need to be able to compute

$$\mathrm{P}[u'Au/u'Bu < c^*|u \sim \mathrm{N}(0, \Sigma)] = \mathrm{P}\left[\sum_{i=1}^{n} \lambda_i \zeta_i^2 < 0\right], \tag{21}$$

where c^* is the critical value, $\Sigma = \mathrm{E}(uu')$, $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $(\Sigma^{1/2})'(A - c^*B)\Sigma^{1/2}$, and $\zeta = (\zeta_1, \ldots, \zeta_n)' \sim \mathrm{N}(0, I_n)$. (21) may be evaluated using Imhof's (1961) algorithm, coded versions of which are given by Koerts and Abrahamse (1969) and Davies (1980). All power calculations were made using exact critical values at the five percent level.

We began by computing (21) using a modified version of Koerts and Abrahamse's FQUAD subroutine with maximum integration and truncation errors of 10^{-6}. This worked well for $PA(\rho_0)$ and $PB(\rho_0)$ with $\rho_0 = 0.5, 0.9$ but for $PB(1.0)$ and $PB(1.1)$, especially under H_a^1, it frequently failed to converge despite increasing the number of iterations. The lack of convergence seemed to be caused by there being one very large positive (negative) eigenvalue balanced by a number of smaller negative (positive) eigenvalues. A Fortran version of Davies' (1980) algorithm worked slightly better but also had similar problems. For the cases of nonconvergence, it seems that the numerical integration problem we were trying to solve is ill-conditioned and so accurate answers may not be possible. We successively lowered the maximum integration and truncation errors until convergence was achieved. This sometimes

[5] $DW_2(\rho_0)$ is the DW test based on the OLS residuals e given by (16).

Table 2

Calculated powers for *PA*(0.5) and *X3* against H_a^- and H_a^+.

Against H_a^-						
$\rho =$	0.5	0.4	0.3	0.2	0.1	0.0
		$n = 20$				
DW	0.050	0.080	0.125	0.190	0.275	0.381
$D_1(0.5)$	0.050	0.083	0.134	0.208	0.305	0.422
$S_1(0.5, 0.25)$	0.050	0.083	0.134	0.210	0.311	0.433
$S_1(0.5, 0)$	0.050	0.083	0.134	0.209	0.311	0.437
		$n = 60$				
DW	0.050	0.149	0.334	0.574	0.788	0.921
$D_1(0.5)$	0.050	0.169	0.400	0.677	0.878	0.968
$S_1(0.5, 0.25)$	0.050	0.169	0.401	0.681	0.883	0.971
$S_1(0.5, 0)$	0.050	0.168	0.399	0.680	0.884	0.972
Against H_a^+						
$\rho =$	0.5	0.6	0.7	0.8	0.9	0.999
		$n = 20$				
DW	0.050	0.094	0.163	0.254	0.357	0.452
$D_1(0.5)$	0.050	0.104	0.188	0.296	0.411	0.508
$S_1(0.5, 0.75)$	0.050	0.103	0.189	0.300	0.418	0.516
$S_1(0.5, 0.999)$	0.050	0.102	0.187	0.298	0.418	0.520
		$n = 60$				
DW	0.050	0.187	0.455	0.744	0.914	0.969
$D_1(0.5)$	0.050	0.219	0.532	0.814	0.946	0.981
$S_1(0.5, 0.75)$	0.050	0.219	0.534	0.818	0.949	0.983
$S_1(0.5, 0.999)$	0.050	0.215	0.528	0.815	0.949	0.984

meant integration and truncation errors as large as 10^{-3}. As a final check, the Monte Carlo method with ten thousand replications was used to recalculate all powers computed for the *X2* design matrix. We could find no significant difference between the two sets of results.

5. Results on power

In this section we discuss the results of the power comparison. A more detailed discussion can be found in an earlier version of this paper [Dufour and King (1989)]. We begin by discussing the results for $PA(\rho_0)$. Calculated powers for design matrix *X3* and *PA*(0.5) and *PA*(0.9) are given in tables 2 and 3, respectively.[6]

[6] To save space, only results for one design matrix are tabulated although this section discusses the results for all design matrices.

Table 3

Calculated powers for $PA(0.9)$ and $X3$ against H_a and H_a^+.

Against H_a

$\rho =$	0.9	0.8	0.7	0.6	0.3	0
		$n = 20$				
DW	0.050	0.067	0.092	0.126	0.300	0.579
$D_1(0.9)$	0.050	0.073	0.110	0.161	0.405	0.676
$S_1(0.9, 0.45)$	0.050	0.073	0.110	0.166	0.458	0.785
$S_1(0.9, 0)$	0.050	0.072	0.108	0.161	0.454	0.806
		$n = 60$				
DW	0.050	0.111	0.210	0.352	0.834	0.993
$D_1(0.9)$	0.050	0.184	0.458	0.747	0.990	0.999
$S_1(0.9, 0.45)$	0.050	0.181	0.465	0.781	0.999	1.000
$S_1(0.9, 0)$	0.050	0.175	0.447	0.763	0.999	1.000

Against H_a^+

$\rho =$	0.9	0.92	0.94	0.96	0.98	0.999
		$n = 20$				
DW	0.050	0.056	0.061	0.068	0.074	0.080
$D_1(0.9)$	0.050	0.060	0.071	0.084	0.097	0.111
$S_1(0.9, 0.95)$	0.050	0.060	0.071	0.084	0.098	0.113
$S_1(0.9, 0.999)$	0.050	0.060	0.071	0.084	0.098	0.114
		$n = 60$				
DW	0.050	0.064	0.081	0.102	0.126	0.148
$D_1(0.9)$	0.050	0.086	0.136	0.196	0.259	0.310
$S_1(0.9, 0.95)$	0.050	0.086	0.136	0.198	0.264	0.318
$S_1(0.9, 0.999)$	0.050	0.085	0.134	0.196	0.264	0.322

The powers of all tests increase as ρ moves away from H_0: $\rho = \rho_0$. Also, *ceteris paribus*, there is a noticeable drop in power going from *X1* to any of the other design matrices, reflecting the addition of extra regressors. With one minor exception, the $DW_1(\rho_0)$ test always has inferior power to the other tests. Generally the POI tests have almost identical power which is typically slightly higher than that of the $D_1(\rho_0)$ test. The spread of powers is greatest for $PA(0.9)$ against H_a^-, particularly for *X2–X5*. As expected, the $D_1(\rho_0)$ test is most powerful near ρ_0, while the $S_1(\rho_0, \rho_1)$ test is most powerful at ρ values near ρ_1. The best overall test is typically a POI test with ρ_1 taking a middle value.

Turning to the nonstationary case, we begin by discussing the results for $PB(0.5)$ and $PB(0.9)$. Calculated powers for *X3* and $PB(0.9)$ are given in table 4. With occasional exceptions for the $S_2(0.9, 0.45, 10.0)$ test, the powers of all tests nearly always increase as ρ moves away from ρ_0, *ceteris paribus*.

Table 4

Calculated powers for $PB(0.9)$ and $X3$ against H_a^- and H_a^+ with $d_1 = 1.0$

	Against H_a^-					
$\rho =$	0.9	0.8	0.7	0.6	0.3	0.0
			$n = 20$			
$\overline{DW}_1(0.9)$	0.050	0.067	0.093	0.127	0.299	0.577
$\bar{D}_1(0.9)$	0.050	0.073	0.110	0.161	0.391	0.650
$\bar{S}_1(0.9, 0.45)$	0.050	0.073	0.111	0.166	0.451	0.773
$\bar{S}_1(0.9, 0.0)$	0.050	0.072	0.107	0.160	0.446	0.800
$DW_2(0.9)$	0.050	0.065	0.086	0.114	0.254	0.503
$D_2(0.9)$	0.050	0.070	0.100	0.140	0.324	0.557
$S_2(0.9, 0.45, 0.1)$	0.050	0.069	0.100	0.141	0.342	0.613
$S_2(0.9, 0.45, 1.0)$	0.050	0.069	0.099	0.142	0.352	0.647
$S_2(0.9, 0.45, 10.0)$	0.050	0.042	0.036	0.033	0.038	0.056
$S_2(0.9, 0.0, 0.1)$	0.050	0.069	0.098	0.138	0.337	0.626
$S_2(0.9, 0.0, 1.0)$	0.050	0.068	0.097	0.136	0.348	0.678
$S_2(0.9, 0.0, 10.0)$	0.050	0.062	0.080	0.105	0.258	0.562
			$n = 60$			
$\overline{DW}_1(0.9)$	0.050	0.112	0.212	0.352	0.832	0.992
$\bar{D}_1(0.9)$	0.050	0.183	0.434	0.690	0.954	0.990
$\bar{S}_1(0.9, 0.45)$	0.050	0.177	0.448	0.760	0.999	1.000
$\bar{S}_1(0.9, 0.0)$	0.050	0.169	0.427	0.739	0.999	1.000
$DW_2(0.9)$	0.050	0.103	0.193	0.324	0.807	0.990
$D_2(0.9)$	0.050	0.155	0.381	0.660	0.981	0.999
$S_2(0.9, 0.45, 0.1)$	0.050	0.154	0.385	0.682	0.995	1.000
$S_2(0.9, 0.45, 1.0)$	0.050	0.153	0.384	0.686	0.997	1.000
$S_2(0.9, 0.45, 10.0)$	0.050	0.136	0.330	0.614	0.995	1.000
$S_2(0.9, 0.0, 0.1)$	0.050	0.150	0.373	0.668	0.995	1.000
$S_2(0.9, 0.0, 1.0)$	0.050	0.147	0.366	0.663	0.997	1.000
$S_2(0.9, 0.0, 10.0)$	0.050	0.143	0.351	0.642	0.997	1.000

	Against H_a^+					
$\rho =$	0.9	0.92	0.94	0.96	0.98	1.0
			$n = 20$			
$\overline{DW}_1(0.9)$	0.050	0.055	0.061	0.067	0.074	0.085
$\bar{D}_1(0.9)$	0.050	0.059	0.068	0.078	0.092	0.112
$\bar{S}_1(0.9, 0.95)$	0.050	0.058	0.068	0.080	0.096	0.122
$\bar{S}_1(0.9, 1.0)$	0.050	0.058	0.067	0.079	0.097	0.126
$DW_2(0.9)$	0.050	0.054	0.058	0.062	0.066	0.070
$D_2(0.9)$	0.050	0.058	0.065	0.072	0.079	0.089
$S_2(0.9, 0.95, 0.1)$	0.050	0.058	0.065	0.072	0.079	0.089
$S_2(0.9, 0.95, 1.0)$	0.050	0.058	0.065	0.072	0.079	0.089
$S_2(0.9, 0.95, 10.0)$	0.050	0.057	0.065	0.072	0.079	0.089
$S_2(0.9, 1.0, d_1^*)$	0.050	0.058	0.065	0.072	0.079	0.089
			$n = 60$			
$\overline{DW}_1(0.9)$	0.050	0.065	0.085	0.113	0.152	0.204
$\bar{D}_1(0.9)$	0.050	0.089	0.148	0.226	0.316	0.403
$\bar{S}_1(0.9, 0.95)$	0.050	0.089	0.148	0.229	0.324	0.418
$\bar{S}_1(0.9, 1.0)$	0.050	0.087	0.144	0.224	0.324	0.429
$DW_2(0.9)$	0.050	0.061	0.074	0.088	0.101	0.117
$D_2(0.9)$	0.050	0.077	0.110	0.146	0.181	0.221
$S_2(0.9, 0.95, 0.1)$	0.050	0.077	0.110	0.146	0.183	0.225
$S_2(0.9, 0.95, 1.0)$	0.050	0.077	0.110	0.146	0.183	0.225
$S_2(0.9, 0.95, 10.0)$	0.050	0.075	0.107	0.141	0.179	0.227
$S_2(0.9, 1.0, d_1^*)$	0.050	0.076	0.108	0.144	0.182	0.230

Each tests shows a loss of power going from *X1* to any other design matrix, *ceteris paribus*, reflecting the cost of including regressors. With the single exception[7] of the $\bar{D}_1(0.9)$ test against H_a^-, the G1-invariant tests, $\overline{DW}_1(\rho_0)$, $\bar{D}_1(\rho_0)$, and $\bar{S}_1(\rho_0,\rho_1)$, which require knowledge of d_1, are almost always more powerful than their respective G2-invariant tests. It also appears that knowledge of d_1 has greater potential to improve the power of a test, the further ρ_0 is away from zero. Because G1-invariant tests are typically nonoperational, our interest is in the G2-invariant tests.

In some circumstances, the choice of d_1^* in $S_2(\rho_0,\rho_1,d_1^*)$ has almost no effect on power, while in other situations such as testing against H_a^-, the choice of d_1^*, particularly that of $d_1^*=10.0$, can cause a severe loss of power. The choice seems to be less critical in large samples. In fact for $n=60$, setting $d_1^*=0.1$ can result in a slight power improvement for ρ values close to ρ_0. An explanation is that setting d_1^* at a lower than true value seems to have a similar effect on power as moving ρ_1 closer to ρ_0. In all cases, the results suggest that a good strategy if d_1 is unknown is to attempt to set d_1^* to a value that is likely to be below, but hopefully near, the true d_1 value. Of the tests that remain after exclusion of the POI tests with $d_1^*=10.0$, the $DW_2(\rho_0)$ test is nearly always the least powerful. For *PB*(0.5), the powers of the $D_2(0.5)$ and $S_2(0.5,\rho_1,d_1^*)$ tests with $d_1^*=0.1$ or 1.0 are nearly identical, particularly when $n=60$ against H_a^-. Against H_a^+, the results indicate that the $S_2(0.5,0.75,0.1)$ and $S_2(0.5,0.75,1.0)$ tests have the best overall power properties. For *PB*(0.9) and against H_a^+, the powers of the $S_2(0.9,\rho_1,d_1^*)$ tests with $d_1^*=0.1$ or 1.0 are almost identical and are typically slightly higher than those of the $D_2(0.9)$ test. Against H_a^-, the power differences are more distinctive with the $S_2(0.9,0.45,d_1^*)$ tests with $d_1^*=0.1$ and 1.0 possibly having the best overall power.

We now discuss the results of most interest, those for *PB*(1.0). Table 5 gives calculated powers for design matrix *X4* which are more representative than those for *X3*. This is because some tests exhibit uncharacteristic behaviour only for *X3* and occasionally also *X2*. For example, the $S_2(1.0,1.1,d_1^*)$ tests are found to be biased, but only for *X3* when $n=60$.

Ceteris paribus, the powers nearly always increase as ρ moves away from ρ_0. Each test typically shows a loss of power going from *X1* to any other design matrix, *ceteris paribus*. A feature of the results against H_a^- is the poor performance of the $D_2(1.0)$ test. Across all design matrices and ρ values, its maximum power is 0.217 when $n=20$ and 0.369 when $n=60$. In contrast, all other tests have maximum powers above 0.72 and 0.98, respectively. Once

[7]Although puzzling, situations in which knowledge of a parameter value can reduce the power of a test are not unknown. A related example is given by Krämer (1985) who shows that for certain regressions fitted through the origin, the power of the DW test can be improved by adding a superfluous intercept to the regression. $\bar{D}_1(0.9)$ is most powerful (among G1-invariant tests) only in the neighbourhood of $\rho=0.9$, not against $\rho=0.0$.

Table 5

Calculated powers for $PB(1.0)$ and $X4$ against H_a^- and H_a^+ with $d_1 = 1.0$.[a]

Against H_a^-						
$\rho =$	1.0	0.9	0.75	0.5	0.25	0.0
$n = 20$						
$DW_2(1.0)$	0.050	0.077	0.124	0.252	0.451	0.673
$D_2(1.0)$	0.050	0.091	0.122	0.148	0.166	0.184
SB	0.050	0.097	0.193	0.450	0.708	0.862
$S_2(1.0, 0.5, 0.1)$	0.050	0.104	0.213	0.433	0.614	0.738
$S_2(1.0, 0.5, 1.0)$	0.050	0.102	0.211	0.476	0.729	0.879
$S_2(1.0, 0.5, 10.0)$	0.050	0.056	0.080	0.205	0.431	0.667
$S_2(1.0, 0.0, 0.1)$	0.050	0.100	0.200	0.437	0.677	0.836
$S_2(1.0, 0.0, 1.0)$	0.050	0.096	0.188	0.442	0.738	0.917
$S_2(1.0, 0.0, 10.0)$	0.050	0.086	0.155	0.379	0.681	0.891
$n = 60$						
$DW_2(1.0)$	0.050	0.102	0.235	0.613	0.921	0.996
$D_2(1.0)$	0.050	0.150	0.196	0.223	0.239	0.256
SB	0.050	0.261	0.799	0.997	1.000	1.000
$S_2(1.0, 0.5, 0.1)$	0.050	0.301	0.750	0.921	0.962	0.983
$S_2(1.0, 0.5, 1.0)$	0.050	0.273	0.816	0.999	1.000	1.000
$S_2(1.0, 0.5, 10.0)$	0.050	0.217	0.734	0.999	1.000	1.000
$S_2(1.0, 0.0, 0.1)$	0.050	0.279	0.739	0.936	0.974	0.990
$S_2(1.0, 0.0, 1.0)$	0.050	0.236	0.749	0.999	1.000	1.000
$S_2(1.0, 0.0, 10.0)$	0.050	0.221	0.724	0.999	1.000	1.000

Against H_a^+						
$\rho =$	1.0	1.025	1.05	1.1	1.15	1.2
$n = 20$						
$DW_2(1.0)$	0.050	0.072	0.124	0.381	0.684	0.857
$D_2(1.0)$	0.050	0.120	0.238	0.533	0.751	0.871
SB	0.050	0.106	0.202	0.464	0.685	0.816
$S_2(1.0, 1.1, 0.1)$	0.050	0.111	0.232	0.577	0.803	0.908
$S_2(1.0, 1.1, 1.0)$	0.050	0.110	0.231	0.577	0.804	0.909
$S_2(1.0, 1.1, 10.0)$	0.050	0.107	0.224	0.574	0.807	0.913
$S_2(1.0, 1.2, 0.1)$	0.050	0.078	0.135	0.465	0.802	0.921
$S_2(1.0, 1.2, 1.0)$	0.050	0.079	0.140	0.481	0.804	0.921
$S_2(1.0, 1.2, 10.0)$	0.050	0.081	0.148	0.501	0.805	0.920
$n = 60$						
$DW_2(1.0)$	0.050	0.175	0.662	0.977	1.000	1.000
$D_2(1.0)$	0.050	0.417	0.789	0.981	1.000	1.000
SB	0.050	0.361	0.743	0.971	1.000	1.000
$S_2(1.0, 1.1, 0.1)$	0.050	0.085	0.555	0.991	0.999	1.000
$S_2(1.0, 1.1, 1.0)$	0.050	0.093	0.648	0.991	1.000	1.000
$S_2(1.0, 1.1, 10.0)$	0.050	0.125	0.757	0.991	0.999	1.000
$S_2(1.0, 1.2, 0.1)$	0.050	0.033	0.025	0.961	0.999	1.000
$S_2(1.0, 1.2, 1.0)$	0.050	0.036	0.030	0.979	1.000	1.000
$S_2(1.0, 1.2, 10.0)$	0.050	0.039	0.038	0.982	0.999	1.000

[a]For $PB(1.0)$, the tests $DW_2(1.0)$, $D_2(1.0)$, and $S_2(1.0, \rho_1, 1.0)$ are identical to $\overline{DW}_1(1.0)$, $\overline{D}_1(1.0)$, and $\overline{S}_1(1.0, \rho_1)$, respectively.

again, in order to apply a POI test when d_1 is unknown, it seems setting d_1^* to a value near or below the true d_1 value would be a sensible strategy, particularly if n is small. Against H_a^-, the *SB* test is clearly superior to the $DW_2(1.0)$ test. For most design matrices, the power of the $S_2(1.0, 0.5, 1.0)$ test dominates that of the *SB* test. The powers of the *SB* and $S_2(1.0, 0.0, 1.0)$ tests are identical for *X1*. We conclude that against H_a^-, the *SB* test has good power which can be improved by the use of an $S_2(1.0, 0.5, d_1^*)$ test with d_1^* chosen to be near or below the true d_1 value.

A feature of the *PB*(1.0) results against H_a^+ is the relatively good performance of the $D_2(1.0)$ test which almost totally dominates that of the *SB* and *DW* tests. Another feature is the insensitivity of the $S_2(1.0, \rho_1, d_1^*)$ tests to the choice of d_1^* when $n = 20$. At first sight, a surprising result is the general dominance, when $n = 60$, of the $S_2(1.0, \rho_1, 10.0)$ tests over POI tests with d_1^* set to 1.0 or 0.1. An explanation is that against H_a^+, an $S_2(1.0, \rho_1, 10.0)$ test has power properties similar to an $S_2(1.0, \rho_1^*, 1.0)$ test where $\rho_1^* < \rho_1$. The $S_2(1.0, 1.1, d_1^*)$ test has better overall power than the $S_2(1.0, 1.2, d_1^*)$ test for $d_1^* = 1.0, 10.0$. When $n = 20$, both the $S_2(1.0, 1.1, d_1^*)$ tests can be described as superior to the $D_2(1.0)$ test, although the picture is reversed when $n = 60$. In summary, against H_a^+, the $D_2(1.0)$ test has good power properties which can be improved upon by the use of an $S_2(1.0, \rho_1, d_1^*)$ test with $\rho_1 = 1.1$ for $n = 20$ and a lower value, say 1.025, when $n = 60$. The d_1^* value should be set near or above the true d_1 value.

Finally we discuss the results for *PB*(1.1). Calculated powers for *X3* are given in table 6. A feature is that not all tests increase in power as ρ moves away from ρ_0. For example, the power of nearly all the G2-invariant tests against H_a^- either first increases and then decreases, or first decreases and then increases, with the point of inflection being around $\rho = 1.0$. Also, when $n = 60$, the powers of the $\bar{D}_1(1.1)$ test against H_a^- and the $D_2(1.1)$ test against H_a^+ always decrease rapidly to zero. Further calculations at ρ values closer to 1.1 showed that the powers of both tests first increase and then decrease. Increasing n does not always improve power. In general, if a test's power when $n = 20$ is below (above) 0.05, then its power when $n = 60$ is further below (above) 0.05. Typically, all tests show a loss of power going from *X1* to any other design matrix, except when the power is below 0.05 in which case it invariably increases. The G1-invariant tests are generally more powerful than their respective G2-invariant tests. Against H_a^-, the $S_2(1.1, 1.0, d_1^*)$ test is invariant to the choice of d_1^* value, while for the $S_2(1.1, 0.5, d_1^*)$ test, the choice can have a substantial impact on power when $n = 20$. In view of these results, we recommend a choice of d_1^* near or below the true d_1 value when n is small, and near or above when n is large. Against H_a^+, the power functions of the $S_2(1.1, \rho_1, d_1^*)$ tests are insensitive to the choice of d_1^* value and this insensitivity increases with sample size.

Table 6

Calculated powers for $PB(1.1)$ and $X3$ against H_a^- and H_a^+ with $d_1 = 1.0$.

			Against H_a^-			
$\rho =$	1.1	1.05	1.0	0.9	0.5	0.0
			$n = 20$			
$\overline{DW}_1(1.1)$	0.050	0.066	0.072	0.082	0.203	0.599
$\bar{D}_1(1.1)$	0.050	0.097	0.066	0.034	0.006	0.000
$\bar{S}_1(1.1, 1.0)$	0.050	0.120	0.182	0.188	0.107	0.026
$\bar{S}_1(1.1, 0.5)$	0.050	0.098	0.143	0.195	0.600	0.951
$DW_2(1.1)$	0.050	0.046	0.046	0.052	0.134	0.483
$D_2(1.1)$	0.050	0.069	0.063	0.042	0.009	0.000
$S_2(1.1, 1.0, d_1^*)$	0.050	0.069	0.064	0.043	0.005	0.000
$S_2(1.1, 0.5, 0.1)$	0.050	0.045	0.045	0.052	0.158	0.528
$S_2(1.1, 0.5, 1.0)$	0.050	0.045	0.045	0.052	0.161	0.576
$S_2(1.1, 0.5, 10.0)$	0.050	0.053	0.053	0.046	0.030	0.030
			$n = 60$			
$\overline{DW}_1(1.1)$	0.050	0.017	0.035	0.070	0.546	0.993
$\bar{D}_1(1.1)$	0.050	0.000	0.000	0.000	0.000	0.000
$\bar{S}_1(1.1, 1.0)$	0.050	1.000	1.000	1.000	1.000	1.000
$\bar{S}_1(1.1, 0.5)$	0.050	0.998	0.982	1.000	1.000	1.000
$DW_2(1.1)$	0.050	0.014	0.028	0.056	0.493	0.989
$D_2(1.1)$	0.050	0.551	0.136	0.032	0.009	0.001
$S_2(1.1, 1.0, d_1^*)$	0.050	0.620	0.222	0.031	0.000	0.000
$S_2(1.1, 0.5, 0.1)$	0.050	0.013	0.024	0.061	0.855	0.996
$S_2(1.1, 0.5, 1.0)$	0.050	0.012	0.023	0.060	0.911	1.000
$S_2(1.1, 0.5, 10.0)$	0.050	0.012	0.026	0.054	0.856	1.000
			Against H_a^+			
$\rho =$	1.1	1.12	1.14	1.16	1.18	1.2
			$n = 20$			
$\overline{DW}_1(1.1)$	0.050	0.097	0.208	0.374	0.541	0.678
$\bar{D}_1(1.1)$	0.050	0.174	0.352	0.521	0.652	0.748
$\bar{S}_1(1.1, 1.15)$	0.050	0.169	0.369	0.559	0.698	0.791
$\bar{S}_1(1.1, 1.2)$	0.050	0.135	0.316	0.538	0.702	0.805
$DW_2(1.1)$	0.050	0.047	0.061	0.140	0.330	0.539
$D_2(1.1)$	0.050	0.053	0.047	0.032	0.020	0.012
$S_2(1.1, 1.15, 0.1)$	0.050	0.049	0.087	0.233	0.459	0.651
$S_2(1.1, 1.15, 1.0)$	0.050	0.048	0.086	0.240	0.475	0.665
$S_2(1.0, 1.15, 10.0)$	0.050	0.045	0.078	0.234	0.474	0.666
$S_2(1.1, 1.2, 0.1)$	0.050	0.046	0.083	0.241	0.479	0.669
$S_2(1.1, 1.2, 1.0)$	0.050	0.046	0.083	0.240	0.479	0.669
$S_2(1.1, 1.2, 10.0)$	0.050	0.045	0.081	0.238	0.478	0.669
			$n = 60$			
$\overline{DW}_1(1.1)$	0.050	0.935	0.987	0.997	0.999	1.000
$\bar{D}_1(1.1)$	0.050	0.885	0.927	0.000	0.000	0.000
$\bar{S}_1(1.1, 1.15)$	0.050	0.170	1.000	1.000	1.000	1.000
$\bar{S}_1(1.1, 1.2)$	0.050	0.026	0.988	1.000	1.000	1.000
$DW_2(1.1)$	0.050	0.938	0.988	0.997	0.999	1.000
$D_2(1.1)$	0.050	0.011	0.002	0.000	0.000	0.000
$S_2(1.1, 1.15, 0.1)$	0.050	0.969	0.994	0.999	1.000	1.000
$S_2(1.1, 1.15, 1.0)$	0.050	0.969	0.994	0.999	1.000	1.000
$S_2(1.1, 1.15, 10.0)$	0.050	0.969	0.994	0.999	1.000	1.000
$S_2(1.1, 1.2, 0.1)$	0.050	0.967	0.994	0.999	1.000	1.000
$S_2(1.1, 1.2, 1.0)$	0.050	0.968	0.994	0.999	1.000	1.000
$S_2(1.1, 1.2, 10.0)$	0.050	0.968	0.994	0.999	1.000	1.000

Against H_a^-, no one G2-invariant test performs well over both the subregions *R1*: $1.0 \leq \rho < 1.1$ and *R2*: $0 \leq \rho < 1.0$. When $n = 20$, the power curves of the $D_2(1.1)$ and $S_2(1.1, 1.0, d_1^*)$ tests are almost identical, while, for $n = 60$, the $S_2(1.1, 1.0, d_1^*)$ tests have a definite power advantage. These tests dominate the $DW_2(1.1)$ and $S_2(1.1, 0.5, d_1^*)$ tests over *R1*, while the reverse is the case over *R2*. The $S_2(1.1, 0.5, 1.0)$ test can be regarded as the best test over *R2*. While we have found that no test performs well over both *R1* and *R2*, it may be that a better choice of ρ_1 in the $S_2(1.1, \rho_1, d_1^*)$ test will produce such a test. For example, the most stringent POI test which involves choosing ρ_1 to minimize the maximum power difference with the power envelope could be such a test.

A feature of the G2-invariant results against H_a^+ is the poor performance of the $D_2(1.1)$ test. It seems that while the powers of the other tests tend to one as n increases, the power of the $D_2(1.1)$ test tends to zero. The powers of the remaining G2-invariant tests are very similar when $n = 60$ with the POI tests being slightly superior. Overall, we recommend the use of the $S_2(1.1, 1.15, d_1^*)$ test, with d_1^* chosen to be close to or above the true d_1 value, for testing $\rho = 1.1$ against $\rho > 1.1$.

6. Concluding remarks

In this paper, we considered the linear regression model with AR(1) disturbances and derived optimal invariant tests for the hypothesis that the autoregressive coefficient ρ has any given value. In the nonstationary case, we stressed the importance of getting test statistics that do not depend on the distribution of the first disturbance (which is typically unknown). We dealt with this problem by considering tests invariant under a larger transformation group than the one used by Durbin and Watson (1971). In practice, this can be done in a simple way by adding an artificial 'regressor', which depends on ρ_0, to the X matrix.

We also presented power comparisons between alternative tests. For the stationary models, our results suggest that both LBI and POI tests are usually superior to DW tests (based on transformed data under H_0), sometimes by wide margins, while the power differences between LBI and POI tests are relatively small, with the biggest difference favouring POI tests. For the nonstationary models, the same situation seems to hold when testing values of ρ less than (and not too close to) one. On the other hand, when testing values of ρ equal to or greater than one, LBI tests (under G2) have poor power relative to other tests. An exception is the LBI test of $\rho = 1$ against $\rho > 1$. The advantage of using POI tests (under G2) is especially strong for testing values of ρ equal to or greater than one. However, the results indicate that choosing a test which optimizes power at a particular point gives no guarantee about power at nonneighbouring points. In fact, when testing

$\rho = 1.1$ against $\rho < 1.1$, our results suggest it is extremely difficult, if not impossible, to find a test which has good power over both $1.0 < \rho < 1.1$ and $0 < \rho < 1.0$. Concerning the choice of d_1^*, our results suggest that it can have a sizable effect on the performance of POI tests. In general, selecting d_1^* below the true d_1, rather than above it, appears to be a wise choice. Further research on whether this extends to $d_1^* = 0$ is currently planned. It is also useful to note that G2-invariant point-optimal tests of $\rho = \rho_0$ against $\rho = 1$, based on the statistic $S_2(\rho_0, 1, d_1^*)$, are invariant to the value of d_1^*, provided the model contains an intercept.

Given a general procedure for testing hypotheses of the form $\rho = \rho_0$, we can obtain confidence regions for ρ by finding the set of admissible values ρ_0 that are acceptable at a given significance level α [see Dufour (1990)]. It is easy to see that the probability that the true value ρ be contained in this set is $1 - \alpha$. Numerical methods for constructing these confidence sets and comparisons between alternative testing procedures are the topic of on-going research.

Appendix A: Alternative forms of LBI tests

We show here how the LBI test statistic, $D_1(\rho_0)$, can be put into the forms (8) and (9). From (5),

$$\begin{aligned} D_1(\rho_0) &= \hat{e}' A_0 \hat{e} / z^{*\prime} z^* \\ &= \left\{ 2\rho_0 \left(\hat{e}'\hat{e} - \hat{e}_1^2 - \hat{e}_n^2 \right) - 2 \sum_{t=1}^{n-1} \hat{e}_t \hat{e}_{t+1} \right\} \Big/ z^{*\prime} z^* \\ &= 2q \left[\rho_0 - \sum_{t=1}^{n-1} \hat{e}_t \hat{e}_{t+1} \Big/ \sum_{t=2}^{n-1} \hat{e}_t^2 \right] \\ &= -2q \left[R_1(\rho_0) - \rho_0 \right], \end{aligned} \tag{A.1}$$

which establishes (8). Because $\hat{e}_t = \rho_0 \hat{e}_{t-1} + z_t^*$, $t = 2, \ldots, n$, we have

$$R_1(\rho_0) = \left\{ \rho_0 \sum_{t=1}^{n-1} \hat{e}_t^2 + \sum_{t=1}^{n-1} \hat{e}_t z_{t+1}^* \right\} \Big/ \sum_{t=2}^{n-1} \hat{e}_t^2. \tag{A.2}$$

Let $w_1 = \hat{e}_1$ and $w_t = z_t^*$, $t = 2, \ldots, n$. Then (setting $0^0 \equiv 1$) we have

$$\hat{e}_t = \sum_{k=0}^{t-1} \rho_0^k w_{t-k}, \qquad t = 1, \ldots, n,$$

and

$$\sum_{t=1}^{n-1} \hat{e}_t z_{t+1}^* = \sum_{t=1}^{n-1} \sum_{k=0}^{t-1} \rho_0^k w_{t-k} w_{t+1} = \sum_{t=2}^{n} \sum_{k=0}^{t-2} \rho_0^k w_{t-k-1} w_t$$

$$= \sum_{t=2}^{n} \sum_{k=1}^{t-1} \rho_0^{k-1} w_{t-k} w_t = \sum_{k=1}^{n-1} \rho_0^{k-1} \left[\sum_{t=1}^{n-k} w_t w_{t+1} \right]$$

$$= \sum_{k=1}^{n-1} \rho_0^{k-1} C_k(w), \tag{A.3}$$

where

$$C_k(w) = \sum_{t=1}^{n-k} w_t w_{t+k} = [w_1 - z_1^*] z_{k+1}^* + \sum_{t=1}^{n-k} z_t^* z_{t+k}^*$$

$$= \left[\left(1 - \rho_0^2\right)^{-1/2} - 1 \right] z_1^* z_{k+1}^* + C_k(z^*), \qquad 1 \le k \le n-1. \tag{A.4}$$

From (A.2), (A.3), and (A.4), we get

$$R_1(\rho_0) - \rho_0 = \rho_0 \left(\hat{e}_1^2 \Big/ \sum_{t=2}^{n-1} \hat{e}_t^2 \right) + \left\{ z^{*\prime} z^* \Big/ \sum_{t=2}^{n-1} \hat{e}_t^2 \right\} \left\{ \sum_{t=1}^{n-1} \hat{e}_t z_{t+1}^* / z^{*\prime} z^* \right\}$$

$$= \rho_0 \left(\hat{e}_1^2 \Big/ \sum_{t=2}^{n-1} \hat{e}_t^2 \right)$$

$$+ \left\{ q^{-1} \left[\left(1 - \rho_0^2\right)^{-1/2} - 1 \right] z_1^* \left[\sum_{k=1}^{n-1} \rho_0^{k-1} z_{k+1}^* / z^{*\prime} z^* \right] \right\}$$

$$+ q^{-1} \sum_{k=1}^{n-1} \rho_0^{k-1} r_k^*.$$

Then, using (A.1),

$$D_1(\rho_0) = -2\left\{\sum_{k=1}^{n-1} \rho_0^{k-1} r_k^* + \rho_0\left[\hat{e}_1^2/z^{*\prime}z^*\right]\right.$$

$$\left. + \left[\left(1-\rho_0^2\right)^{-1/2} - 1\right] z_1^* \left[\sum_{t=2}^{n} \rho_0^{t-2} z_t^*/z^{*\prime}z^*\right]\right\}$$

$$= -2\sum_{k=1}^{n-1} \rho_0^{k-1} r_k^* + \eta,$$

so that expression (9) is thus proved.

In the nonstationary case, analogous expressions for $D_2(\rho_0)$ can be derived in a similar way and by noting that $v_1 = 0$.

Appendix B: Distribution of the LBI statistic in the nonstationary case with $\rho_0 = 1$

From (19), we have $D_2(1) = 1 - \psi$, where

$$\psi = (l_n'v)^2/v'v = \bar{u}'\overline{M}l_n l_n'\overline{M}\bar{u}/\bar{u}'\overline{M}\bar{u},$$

in which $\rho_0 = 1$ so that $\bar{u} = C(1)u \sim \mathrm{N}(0, \sigma^2 I_n)$ when $\rho = 1$. (Since the distribution of v does not depend on d_1, we can take $d_1 = 1$.) If $\overline{M}l_n = 0$, we have $\psi = 0$ and the test is not applicable. This corresponds to the case where the untransformed model contains a linear trend. If we exclude this case and suppose that $\overline{M}l_n \neq 0$, we have

$$\psi/\left[l_n'\overline{M}l_n\right] = q_1/(q_1+q_2) = \left[1 + (q_2/q_1)\right]^{-1},$$

where $q_i = \bar{u}'M_i\bar{u}$, $i = 1, 2$, $M_1 = \overline{M}l_n l_n' M/[l_n'\overline{M}l_n]$, $M_2 = \overline{M} - M_1$. M_1 and M_2 are idempotent matrices with $M_1 M_2 = 0$. Further, because $l_n l_n'$ has rank 1 and $\overline{M}l_n \neq 0$, M_1 has rank 1 and

$$\mathrm{rank}[M_2] = \mathrm{tr}[M_2] = n - \mathrm{rank}[\overline{X}] - 1.$$

Thus, q_1 and q_2 are independent chi-square random variables with 1 and $n - k_1 - 1$ degrees of freedom, respectively, where $k_1 = \mathrm{rank}[\overline{X}]$. Hence $q_2/q_1 \sim (n - k_1 - 1)F(n - k_1 - 1, 1)$, where $F(n - k_1 - 1, 1)$ is a random variable that follows a central F distribution with $(n - k_1 - 1, 1)$ degrees of freedom.

References

Anderson, T.W., 1948, On the theory of testing serial correlation, Skandinavisk Aktuarietiedskrift 31, 88–116.

Anderson, T.W., 1959, On asymptotic distributions of parameters of stochastic difference equations, Annals of Mathematical Statistics 30, 676–687.

Berenblut, I.I. and G.I. Webb, 1973, A new test for autocorrelated errors in the linear regression model, Journal of the Royal Statistical Society B 35, 33–50.

Bhargava, A., 1986, On the theory of testing for unit roots in observed time series, Review of Economic Studies 53, 369–384.

Davies, R.B., 1980, Algorithm AS155: The distribution of a linear combination of χ^2 random variables, Applied Statistics 29, 323–333.

Dickey, D.A. and W.A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, Journal of the American Statistical Association 74, 427–431.

Dickey, D.A. and W.A. Fuller, 1981, Likelihood ratio statistics for autoregressive time series with a unit root, Econometrica 49, 1057–1072.

Diebold, F.X. and M. Nerlove, 1988, Unit roots in economic time series: A selective survey, in: T.B. Fomby and G.F. Rhodes, eds., Advances in econometrics: Co-integration, spurious regressions and unit roots (JAI Press, Greenwich, CT) forthcoming.

Dufour, J.-M., 1990, Exact tests and confidence sets in linear regressions with autocorrelated errors, Econometrica 58, 475–494.

Dufour, J.-M. and M. Dagenais, 1985, Durbin–Watson tests for serial correlation in regressions with missing observations, Journal of Econometrics 27, 371–381.

Dufour, J.-M. and M.L. King, 1989, Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary and nonstationary AR(1) errors, Cahier 2889 (Centre de Recherche et Développement en Économique, Université de Montréal, Montréal).

Durbin, J. and G.S. Watson, 1950, Testing for serial correlation in least squares regression I, Biometrika 37, 409–428.

Durbin, J. and G.S. Watson, 1951, Testing for serial correlation in least squares regression II, Biometrika 38, 159–178.

Durbin, J. and G.S. Watson, 1971, Testing for serial correlation in least squares regression III, Biometrika 58, 1–19.

Evans, G.B.A. and N.E. Savin, 1981, Testing for unit roots: I, Econometrica 49, 753–779.

Evans, G.B.A. and N.E. Savin, 1984, Testing for unit roots: II, Econometrica 52, 1241–1269.

Fuller, W.A., 1985, Nonstationary autoregressive time series, in: E.J. Hannan, P.R. Krishnaiah, and M.M. Rao, eds., Handbook of statistics, Vol. 5 (Elsevier Science, Amsterdam) 1–23.

Imhof, J.P., 1961, Computing the distribution of quadratic forms in normal variables, Biometrika 48, 419–426.

King, M.L., 1980, Robust tests for spherical symmetry and their application to least squares regression, Annals of Statistics 8, 1265–1271.

King, M.L., 1987a, Testing for autocorrelation in linear regression models: A survey, in: M.L. King and D.E.A. Giles, eds., Specification analysis in the linear model (Routledge and Kegan Paul, London) 19–73.

King, M.L., 1987b, Towards a theory of point optimal testing, Econometric Reviews 6, 169–218.

King, M.L. and M.A. Evans, 1988, Locally optimal properties of the Durbin–Watson test, Econometric Theory 4, 509–516.

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

Koerts, J. and A.P.J. Abrahamse, 1969, On the theory and application of the general linear model (Rotterdam University Press, Rotterdam).

Krämer, W., 1985, The power of the Durbin–Watson test for regressions without an intercept, Journal of Econometrics 28, 363–370.

Lehmann, E.L., 1986, Testing statistical hypotheses, 2nd ed. (Wiley, New York, NY).

Miyazaki, S. and W.E. Griffiths, 1984, The properties of some covariance matrix estimators in linear models with AR(1) errors, Economics Letters 14, 351–356.

Nankervis, J.C. and N.E. Savin, 1985, Testing the autoregressive parameter with the t statistic, Journal of Econometrics 27, 143–161.

Park, R.E. and B.M. Mitchell, 1980, Estimating the autocorrelated error model with trended data, Journal of Econometrics 13, 185–201.

Phillips, P.C.B., 1987a, Time series regression with a unit root, Econometrica 55, 277–301.

Phillips, P.C.B., 1987b, Towards a unified asymptotic theory for autoregression, Biometrika 74, 535–547.

Phillips, P.C.B. and S.N. Durlauf, 1986, Multiple time series regression with integrated processes, Review of Economic Studies 53, 473–495.

Rao, M.M., 1978, Asymptotic distribution of an estimator of the boundary parameter of an unstable process, Annals of Statistics 6, 185–190 and 8, 1403.

Sargan, J.D. and A. Bhargava, 1983, Testing residuals from least squares regression for being generated by the Gaussian random walk, Econometrica 51, 153–174.

Satchell, S.F., 1984, Approximation to the finite sample distribution for nonstable first order stochastic difference equations, Econometrica 52, 1271–1289.

Shively, T.S., C.F. Ansley, and R. Kohn, 1989, Fast evaluation of the distribution of the Durbin–Watson and other invariant test statistics in regression, Paper presented at the Australasian meeting of the Econometric Society, Armidale.

Watson, G.S., 1955, Serial correlation in regression analysis I, Biometrika 42, 327–341.