FEA Working Paper No. 2000-3

A NEW APPROACH TO TESTING
A COMPOSITE NULL
AGAINST A COMPOSITE ALTERNATIVE

Nelufa Begum and Maxwell L. King*

Department of Econometrics and Business Statistics
Monash University
Clayton
Victoria 3800
AUSTRALIA

*E-Mail: Max.King@BusEco.monash.edu.au

January 2000

A New Approach to Testing a Composite Null against a Composite Alternative[1]

by

Nelufa Begum and Maxwell L. King

Department of Econometrics and Business Statistics
Monash University
Clayton, Vic. 3800
Australia

Abstract

This paper presents a new approach to testing a composite null hypothesis against a composite alternative based on the generalised Neyman-Pearson lemma. The test maximizes the average power function subject to the average size being controlled. We illustrate the new test procedure by applying it to the problem of testing AR(1) against MA(1) errors in the linear regression model. The practical difficulty of finding the critical values of the test is discussed.

Key Words: Generalised Neyman-Pearson Lemma, AR(1), MA(1), average power test.

[1] This paper is based on research for the first author's Ph.D. thesis. It has been presented at the Faculty of Business and Economics Ph.D. Conference at Monash University and at a seminar at the University of Malaya. The authors are grateful for comments from both audiences. Any remaining errors are our own.

1. Introduction

The theory of hypothesis testing is well developed in the case of testing a simple null hypothesis against a simple alternative hypothesis and a simple null hypothesis against a composite alternative hypothesis. Much less developed is the problem of testing a composite null hypothesis against a composite alternative. There are a number of procedures that involve techniques for reducing a composite null hypothesis to a simple null hypothesis. These include invariant techniques to simplify the testing problem and eliminate most or all of the nuisance parameters in the model under the null hypothesis. In any hypothesis testing problem involving nuisance parameters for which sufficient statistics exist under the null hypothesis, similar critical regions will exist and they may be constructed by finding regions of Neyman structure (Cox and Hinkley, 1974, p.135).

We know, when testing a simple null hypothesis against a simple alternative hypothesis, the fundamental Neyman-Pearson (N-P) lemma provides the most powerful test. Similarly, when testing a simple null hypothesis against a composite (one-sided) alternative hypothesis, the N-P lemma can be used to obtain a test (point optimal test) which is most powerful in the neighbourhood of a point in the alternative parameter space. King (1987) reviewed the theory of point optimal invariant (POI) tests, which have been found to work well for some composite hypothesis testing problems in finite samples. There are some situations where it is almost impossible to obtain point optimal tests, particularly when testing a composite null hypothesis. An example of such a situation is that of testing for moving average against autoregressive disturbances in the linear regression model. Silvapulle and King (1991) could not find a POI test for this composite null hypothesis testing problem, so they constructed an approximately POI test. In this regard we can say that when testing a composite null hypothesis against a composite alternative, the N-P lemma does not provide a test, which is point optimal.

The aim of this paper is to explore a general solution to testing a composite null hypothesis against a composite alternative. This approach involves the use of the generalised N-P lemma to construct a new test, called the 'optimal average power test'. The test maximizes the average power function subject to the average size being controlled. A composite null hypothesis brings with it the possibility that the size will vary with different parameter values under the null hypothesis. Typically we hope that any test we construct will have constant size across the null hypothesis parameter space. If this is the case then it is relatively easy to control the size of the test. Unfortunately, there are many circumstances in which size will vary across the parameter space. Controlling size is extremely difficult and time consuming in these cases. The standard approach is to control the maximum size to be less than or equal to some desired constant level (see, for example, Lehmann and Stein (1948)). The other aim of this paper is to consider controlling average size to ease this problem.

We illustrate the new test procedure by applying it to the problem of testing AR(1) against MA(1) errors in the linear regression model. King (1983) constructed and investigated the properties of a test for this problem, called a pseudo POI (PPOI) test by applying the N-P lemma to a maximal invariant statistic. King and McAleer (1987) further compared the small sample properties of the Cox test, some linearized Cox tests, and an approximate point optimal test, as well as the Lagrange multiplier test of AR(1) against ARMA(1,1) disturbances in the linear regression model. From the existing literature on hypothesis testing, the POI test seems to be the more powerful test than the PPOI tests. As Silvapulle and King (1991) could not find a POI test for a composite null hypothesis testing problem in the linear regression model, this motivates us to find a general solution for testing a composite null hypothesis against a composite alternative. In this paper, the finite sample behaviour of the average power test is analysed using the Monte Carlo simulation method.

The plan of the paper is as follows. The theory and the general testing procedures are introduced in section 2. The application of the approach to testing AR(1) against MA(1) errors in the linear regression model is described in section 3. A Monte Carlo experiment is outlined in section 4. The results of the simulation study are presented in section 5. Section 6 contains some concluding remarks.

2. Theory and testing procedure

Let x be an observable $n \times 1$ vector and suppose we wish to test,

H_0: x has density $f(x, \theta)$,

where θ is a $j \times 1$ vector of parameters restricted to the set Θ, against

H_a: x has density $g(x, \phi)$,

where ϕ is a $k \times 1$ vector of parameters restricted to the set Φ. This is a very general form of testing problem and it is assumed that any knowledge about the possible range of parameter values has been used to keep the parameter sets, Θ and Φ, as small as possible.

Our aim is to find the critical region ω for which when $x \in \omega$ we reject H_0 and which maximizes average power subject to controlling average size of the test. For our test, controlling average size may cause problems of large sizes in some parts of the parameter space and small sizes elsewhere. If we break up the parameter space into smaller subspaces, and control average size over these smaller subspaces, we may end up with a better test. In this regard, we break up the parameter space into subspaces with smaller sizes and larger sizes for the test in which average size is controlled over the entire space. Under H_0, we therefore select m disjoint regions of Θ, namely, $\Theta_1, \Theta_2, \ldots, \Theta_m$, so that

$$\Theta = \Theta_1 \cup \Theta_2 \cup \ldots \cup \Theta_m ,$$

with the aim of controlling size over each of the Θ_i regions under H_0.

Note that $\int_\omega f(x,\theta)dx$ is the size of the test for a given θ value under H_0. The average size of the test for a particular region of the parameter space Θ_i is

$$\int_{\Theta_i}\int_\omega f(x,\theta)\, dx P_{1i}(\theta)\, d\theta$$

$$= \int_\omega \int_{\Theta_i} f(x,\theta)\, P_{1i}(\theta)\, d\theta dx \qquad \text{(by Fubini's Theorem)}$$

$$= \int_\omega f_i(x)dx, \qquad i = 1, 2, \ldots, m,$$

where $P_{1i}(\theta)$ is the prior distribution of θ for $\theta \in \Theta_i$, and

$$f_i(x) = \int_{\Theta_i} f(x,\theta)\, P_{1i}(\theta)\, d\theta, \quad i = 1, 2, \ldots, m,$$

is the integrated likelihood function under the null hypothesis.

The power function of the test for a given ϕ is $\int_\omega g(x,\phi)dx$ and the average power function of the test is

$$\int_\Phi \int_\omega g(x,\phi)\, dx P_2(\phi)\, d\phi$$

$$= \int_\omega \int_\Phi g(x,\phi)\, P_2(\phi)\, d\phi dx \qquad \text{(by Fubini's Theorem)}$$

$$= \int_\omega f_{m+1}(x)dx,$$

where $P_2(\phi)$ is the prior distribution of ϕ for $\phi \in \Phi$, and

$$f_{m+1}(x) = \int_{\Phi} g(x,\phi)\, P_2(\phi)\, d\phi,$$

is the integrated likelihood function under the alternative hypothesis.

Now, we explain the basic idea behind the generalised N-P lemma and how this can be used to construct a maximum average power test. The lemma is given in the Appendix. If $f_1(x),\ f_2(x),\ \ldots\ ,\ f_{m+1}(x)$ are integrable functions over the sample space and if we wish to maximize

$$\int_{\omega} f_{m+1}(x)dx, \tag{1}$$

subject to

$$\int_{\omega} f_i(x)dx \le c_i, \qquad i = 1,2,\ldots,m, \tag{2}$$

then the generalised N-P lemma implies that this can be done by using the critical region

$$\omega = \left\{ f_{m+1}(x) \ge \sum_{i=1}^{m} k_i f_i(x) \right\} \tag{3}$$

provided that the constants $k_1, k_2, \ldots, k_m$ exist such that (2) holds. Therefore, under H_0, we have m size conditions in order to solve for the m constants $k_1, k_2, \ldots, k_m$. Then ω is the region, which maximizes $\int_{\omega} f_{m+1}(x)dx,$ our average power.

The average power test proposed in this paper for testing a composite null hypothesis against a composite alternative is based on (3). Thus the null hypothesis is rejected if

$$f_{m+1}(x) \ge \sum_{i=1}^{m} k_i f_i(x), \tag{4}$$

where $k_1, k_2, \ldots, k_m$ are chosen such that the size conditions are satisfied.

3. Application to testing AR(1) against MA(1) errors in the linear regression model

Consider the linear regression model

$$y = X\beta + u , \tag{5}$$

where y is an $n \times 1$ vector, X is an $n \times k$ non-stochastic matrix of rank $k < n$, and β is a $k \times 1$ vector of parameters. If the components of the $n \times 1$ disturbance vector u are generated by the stationary AR(1) process,

$$u_t = \rho u_{t-1} + e_t , \quad |\rho| < 1, \quad t = 1, 2, \ldots, n, \tag{6}$$

where $u_0 \sim N(0, \sigma^2 / (1 - \rho^2))$ and $e = (e_1, e_2, \ldots, e_n)' \sim N(0, \sigma^2 I_n)$, then $u \sim N(0, \sigma^2 \Omega(\rho))$ in which $\Omega(\rho)$ is an $n \times n$ matrix whose (i, j) th element is $\rho^{|i-j|} / (1 - \rho^2)$.

If the components of u are generated by the MA(1) process

$$u_t = \gamma e_{t-1} + e_t , \qquad t = 1, 2, \ldots, n, \tag{7}$$

where $e^* = (e_0, e_1, \ldots, e_n)' \sim N(0, \sigma^2 I_{n+1})$, then $u \sim N(0, \sigma^2 \Sigma(\gamma))$, where $\Sigma(\gamma)$ is the tri-diagonal matrix with $1 + \gamma^2$ as the main diagonal elements and γ as the non-zero off-diagonal elements. For this problem, β and σ^2 are nuisance parameters.

Our main concern is with testing

$$H_0 : u \sim N(0, \sigma^2 \Omega(\rho)), \quad 0 \le \rho < 1, \tag{8}$$

against

$$H_a : u \sim N(0, \sigma^2 \Sigma(\gamma)), \quad 0 < \gamma \le 1. \tag{9}$$

The nuisance parameters in the above problem, β and σ^2, can be eliminated by using invariant transformations. The testing problem (5) is invariant to transformations of the form,

$$y^* = \eta_0 y + X\eta, \tag{10}$$

where η_0 is a positive scalar and η is a $k \times 1$ vector. Therefore, the test statistic we consider should also be invariant to transformations of the form (10). With respect to (10), the vector

$$v = \frac{Pz}{(z'z)^{1/2}} \tag{11}$$

is a maximal invariant, where $z = My$ is the OLS residual vector from (5), $M = I_n - X(X'X)^{-1}X'$, and P is any $(n-k)\times n$ matrix such that $P'P = M$ and $PP' = I_{n-k}$. Let $m = n-k$. The joint density function of v when $u \sim N(0,\sigma^2\Psi)$ has been found by King (1979) to be

$$f(v) = \frac{1}{2}\Gamma(m/2)\pi^{-\frac{m}{2}}\left|P\Psi P'\right|^{-\frac{1}{2}}(v'(P\Psi P')^{-1}v)^{-\frac{m}{2}} \tag{12}$$

with respect to the uniform measure on $\{x: x \in R^m,\ x'x = 1\}$.

King (1983) observed that because v is a maximal invariant, any statistic invariant to transformations of the form (10) can be expressed as a function of v. Now we can consider v as the observed sample and the nuisance parameters, β and σ^2, are avoided.

Our testing problem given by (8) and (9) is equivalent to testing

$$H_0: f_0(v) = \frac{1}{2}\Gamma(m/2)\pi^{-\frac{m}{2}}\left|P\Omega(\rho)P'\right|^{-\frac{1}{2}}(v'(P\Omega(\rho)P')^{-1}v)^{-\frac{m}{2}} \tag{13}$$

against

$$H_a: f_a(v) = \frac{1}{2}\Gamma(m/2)\pi^{-\frac{m}{2}}\left|P\Sigma(\gamma)P'\right|^{-\frac{1}{2}}(v'(P\Sigma(\gamma)P')^{-1}v)^{-\frac{m}{2}} \quad . \tag{14}$$

From lemma 2 in King (1980)

$$v'(P\Omega(\rho)P')^{-1}v = \frac{\tilde{u}'\Omega(\rho)^{-1}\tilde{u}}{z'z} \ ,$$

where z is the OLS residual vector from (5) and $\tilde{u}$ is the GLS residual vector assuming $u \sim N(0,\Omega(\rho))$. Also $\tilde{u}'\Omega(\rho)^{-1}\tilde{u} = \tilde{z}'\tilde{z}$, where $\tilde{z}$ is the OLS residual vector from the regression

$$\Omega(\rho)^{-\frac{1}{2}}y = \Omega(\rho)^{-\frac{1}{2}}X\beta + \Omega(\rho)^{-\frac{1}{2}}u \, ,$$

where $\Omega(\rho)^{-\frac{1}{2}}$ is the inverse of $\Omega(\rho)^{\frac{1}{2}}$ such that $\Omega(\rho)^{\frac{1}{2}}\Omega(\rho)^{\frac{1}{2}\prime} = \Omega(\rho)$.

Therefore

$$v'(P\Omega(\rho)P')^{-1}v = \frac{\tilde{z}'\tilde{z}}{z'z} \, .$$

It is possible to show, as pointed out by Verbyla (1990), that

$$\left|P\Omega(\rho)P'\right| = \left|X'X\right|^{-1}\left|\Omega(\rho)\right|\left|X'\Omega(\rho)^{-1}X\right| .$$

Using these above two results, (13) and (14) can be written as

$$H_0: f_0(v) = \frac{1}{2}\Gamma(m/2)\pi^{-\frac{m}{2}}\left|X'X\right|^{\frac{1}{2}}\left|\Omega(\rho)\right|^{-\frac{1}{2}}\left|X'\Omega(\rho)^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}} \tag{15}$$

against

$$H_a: f_a(v) = \frac{1}{2}\Gamma(m/2)\pi^{-\frac{m}{2}}\left|X'X\right|^{\frac{1}{2}}\left|\Sigma(\gamma)\right|^{-\frac{1}{2}}\left|X'\Sigma(\gamma)^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}} . \tag{16}$$

For (16), $\tilde{z}$ is the OLS residual vector from the regression

$$\Sigma(\gamma)^{-\frac{1}{2}} y = \Sigma(\gamma)^{-\frac{1}{2}} X\beta + \Sigma(\gamma)^{-\frac{1}{2}} u ,$$

where $\Sigma(\gamma)^{-\frac{1}{2}}$ is the inverse of $\Sigma(\gamma)^{\frac{1}{2}}$ such that $\Sigma(\gamma)^{\frac{1}{2}} \Sigma(\gamma)^{\frac{1}{2}\prime} = \Sigma(\gamma)$.

Now the fundamental N-P lemma implies that the test that maximizes the average power subject to controlling average size over the entire null hypothesis parameter space, can be based on the critical region of the form

$$\int_0^1 \frac{1}{2}\Gamma(m/2)\pi^{-\frac{m}{2}} |X'X|^{\frac{1}{2}} |\Sigma(\gamma)|^{-\frac{1}{2}} |X'\Sigma(\gamma)^{-1} X|^{-\frac{1}{2}} \left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}} P(\gamma)d\gamma$$
$$\geq c_\alpha \int_0^1 \frac{1}{2}\Gamma(m/2)\pi^{-\frac{m}{2}} |X'X|^{\frac{1}{2}} |\Omega(\rho)|^{-\frac{1}{2}} |X'\Omega(\rho)^{-1} X|^{-\frac{1}{2}} \left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}} P(\rho)d\rho$$

or,

$$\frac{\int_0^1 |\Sigma(\gamma)|^{-\frac{1}{2}} |X'\Sigma(\gamma)^{-1} X|^{-\frac{1}{2}} \left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}} d\gamma}{\int_0^1 |\Omega(\rho)|^{-\frac{1}{2}} |X'\Omega(\rho)^{-1} X|^{-\frac{1}{2}} \left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}} d\rho} \geq c_\alpha \qquad (17)$$

where $P(\rho)$ and $P(\gamma)$ are prior distributions of ρ and γ respectively, and c_α is the critical value at the α level of significance. The uniform distribution would appear to be a good choice of prior for ρ and γ because both ρ and γ are restricted to lie in an interval. In equation (17), $P(\rho)$ and $P(\gamma)$ have been replaced by uniform density functions which cancel.

For simplicity of calculations, we convert the integrals to numerical integrals by replacing ρ and γ by the function $\left(\frac{l}{q} - \frac{1}{2q}\right)$, where l is positive integer, namely $l = 1, 2, \ldots, q$, q is a large number, say 100 and $\left(\frac{l}{q} - \frac{1}{2q}\right)$ is a point that exits in [0,1].

These sums over l give numerical approximations to the integrals. Thus the above ratio of integrals (17) is approximately equivalent to the ratio of numerical integrals (Conte 1965, Theorem 4.3, p.120)

$$\frac{\sum_{l=1}^{q}\left|\Sigma(\frac{l}{q}-\frac{1}{2q})\right|^{-\frac{1}{2}}\left|X'\Sigma(\frac{l}{q}-\frac{1}{2q})^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}}}{\sum_{l=1}^{q}\left|\Omega(\frac{l}{q}-\frac{1}{2q})\right|^{-\frac{1}{2}}\left|X'\Omega(\frac{l}{q}-\frac{1}{2q})^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}}} \geq c_{\alpha}. \tag{18}$$

Using equation (18), we calculated the true level (sizes) of the average power test for different values of ρ, say, ρ = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9. On the basis of these true levels of the test, we first divided the composite null hypothesis $(0 \leq \rho < 1)$ into two disjoint intervals, one of ρ values with larger test sizes and one of ρ values with smaller test sizes. We then controlled average size over these two smaller subspaces, by finding the value of the constants $k_i \; (i=1,\, 2)$ for which the size conditions are satisfied. Using the values of the constants we calculated sizes and powers of this test. The results obtained were not totally satisfactory. We then divided the null hypothesis parameter space further into three disjoint intervals. Thus, equation (4) gives the resultant test that maximizes average power subject to $m=3$ size conditions and involves a critical region of the form

$$\begin{aligned}
&\sum_{l=1}^{q}\left|\Sigma(\frac{l}{q}-\frac{1}{2q})\right|^{-\frac{1}{2}}\left|X'\Sigma(\frac{l}{q}-\frac{1}{2q})^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}} \\
&\geq \Bigg\{ k_1\sum_{l=1}^{l_1}\left|\Omega(\frac{l}{q}-\frac{1}{2q})\right|^{-\frac{1}{2}}\left|X'\Omega(\frac{l}{q}-\frac{1}{2q})^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}} \\
&+k_2\sum_{l=l_2}^{l_3}\left|\Omega(\frac{l}{q}-\frac{1}{2q})\right|^{-\frac{1}{2}}\left|X'\Omega(\frac{l}{q}-\frac{1}{2q})^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}} \\
&+k_3\sum_{l=l_4}^{q}\left|\Omega(\frac{l}{q}-\frac{1}{2q})\right|^{-\frac{1}{2}}\left|X'\Omega(\frac{l}{q}-\frac{1}{2q})^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}}\Bigg\}
\end{aligned}$$

or,

$$\Bigg\{\Bigg[\sum_{l=1}^{q}\left|\Sigma\left(\frac{l}{q}-\frac{1}{2q}\right)\right|^{-\frac{1}{2}}\left|X'\Sigma\left(\frac{l}{q}-\frac{1}{2q}\right)^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}}$$

$$-k_2\sum_{l=l_2}^{l_3}\left|\Omega\left(\frac{l}{q}-\frac{1}{2q}\right)\right|^{-\frac{1}{2}}\left|X'\Omega\left(\frac{l}{q}-\frac{1}{2q}\right)^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}}$$

$$-k_3\sum_{l=l_4}^{q}\left|\Omega\left(\frac{l}{q}-\frac{1}{2q}\right)\right|^{-\frac{1}{2}}\left|X'\Omega\left(\frac{l}{q}-\frac{1}{2q}\right)^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}}\Bigg]$$

$$\Bigg/\Bigg[\sum_{l=1}^{l_1}\left|\Omega\left(\frac{l}{q}-\frac{1}{2q}\right)\right|^{-\frac{1}{2}}\left|X'\Omega\left(\frac{l}{q}-\frac{1}{2q}\right)^{-1}X\right|^{-\frac{1}{2}}\left(\frac{\tilde{z}'\tilde{z}}{z'z}\right)^{-\frac{m}{2}}\Bigg]\Bigg\}>k_1. \qquad (19)$$

Thus, on the basis of (19), we calculated the sizes and powers of the test for appropriate choices of the constants k_i $(i=1, 2, 3)$. A description of how to choose the values of the constants, is presented in the following section.

4. Monte Carlo experiment

In order to investigate the small sample properties of our test, we conducted a Monte Carlo experiment for the following design matrices.

$X1$: $(n\times 2;\ n=20, 30, 50, 75)$. Intercept and a linear time trend.

$X2$: $(n\times 3;\ n=30, 50, 69)$. The first n observations of Durbin and Watson's (1951, p.159) consumption of spirits example: that is, a constant, plus annual data on the price of spirits and household income in the UK.

$X3$: $(n\times 5;\ n=30, 50, 70)$. A constant dummy, the quarterly Australian Consumer Price Index commencing 1959(1), the same index lagged one quarter, two quarters and three quarters as additional regressors.

$X4$: $(n \times 3;\ n = 30,\ 50)$. A constant dummy, plus quarterly Australian private capital and Government capital movements commencing 1986(1).

Using three thousand replications, we calculated the value of the test statistic and applied the test. For this, first we have to find the value of the constants $k_i\ (i = 1, 2, 3)$, so that the average size conditions are simultaneously satisfied. Based on the above design matrices, ranges and the corresponding values of l are given in Table 1 below.

Table 1
Ranges of the three disjoint intervals and the corresponding values of l

n	Θ_1	Θ_2	Θ_3
For design matrix $X1$:			
20	$0 \le \rho < 0.2$ $(l = 1,2,\ldots,20)$	$0.2 \le \rho < 0.8$ $(l = 21,22,\ldots,80)$	$0.8 \le \rho < 1$ $(l = 81,82,\ldots,q = 100)$
30	$0 \le \rho < 0.2$ $(l = 1,2,\ldots,20)$	$0.2 \le \rho < 0.725$ $(l = 21,22,\ldots,72)$	$0.725 \le \rho < 1$ $(l = 73,74,\ldots,q = 100)$
50	$0 \le \rho < 0.165$ $(l = 1,2,\ldots,16)$	$0.165 \le \rho < 0.65$ $(l = 17,18,\ldots,65)$	$0.65 \le \rho < 1$ $(l = 66,67\ldots,q = 100)$
75	$0 \le \rho < 0.125$ $(l = 1,2,\ldots,13)$	$0.125 \le \rho < 0.6$ $(l = 14,15,\ldots,60)$	$0.6 \le \rho < 1$ $(l = 61,62,\ldots,q = 100)$
For design matrix $X2$:			
30	$0 \le \rho < 0.22$ $(l = 1,2,\ldots,22)$	$0.22 \le \rho < 0.76$ $(l = 23,24,\ldots,76)$	$0.76 \le \rho < 1$ $(l = 77,78,\ldots,q = 100)$
50	$0 \le \rho < 0.24$ $(l = 1,2,\ldots,24)$	$0.24 \le \rho < 0.63$ $(l = 25,26,\ldots,63)$	$0.63 \le \rho < 1$ $(l = 64,65,\ldots,q = 100)$
69	$0 \le \rho < 0.16$ $(l = 1,2,\ldots,16)$	$0.16 \le \rho < 0.6$ $(l = 17,18,\ldots,60)$	$0.6 \le \rho < 1$ $(l = 61,62,\ldots,q = 100)$
For design matrix $X3$:			
30	$0 \le \rho < 0.22$ $(l = 1,2,\ldots,22)$	$0.22 \le \rho < 0.71$ $(l = 23,24,\ldots,71)$	$0.71 \le \rho < 1$ $(l = 72,73,\ldots,q = 100)$
50	$0 \le \rho < 0.18$ $(l = 1,2,\ldots,18)$	$0.18 \le \rho < 0.65$ $(l = 19,20,\ldots,65)$	$0.65 \le \rho < 1$ $(l = 66,67,\ldots,q = 100)$
70	$0 \le \rho < 0.12$ $(l = 1,2,\ldots,12)$	$0.12 \le \rho < 0.62$ $(l = 13,14,\ldots,62)$	$0.62 \le \rho < 1$ $(l = 63,64,\ldots,q = 100)$
For design matrix $X4$:			
30	$0 \le \rho < 0.18$ $(l = 1,2,\ldots,18)$	$0.18 \le \rho < 0.7$ $(l = 19,20,\ldots,70)$	$0.7 \le \rho < 1$ $(l = 71,72,\ldots,q = 100)$
50	$0 \le \rho < 0.19$ $(l = 1,2,\ldots,19)$	$0.19 \le \rho < 0.63$ $(l = 20,21,\ldots,63)$	$0.63 \le \rho < 1$ $(l = 64,65,\ldots,q = 100)$

For the problem under consideration, the proposed average power test under H_0 and H_a is invariant with respect to β and σ^2. We choose, $\beta = 0$ and $\sigma^2 = 1$ for the whole experiment and the computer program was written in Gauss 3.2.18 version.

The following steps were used to calculate the value of the constants k_1, k_2 and k_3.

Step 1: Fix the values of k_2 and k_3 to some reasonable initial values. We draw the parameters from Θ_1 and generate the data set for this drawing, and calculate the left-hand-side of (19). Repeat this three thousand times. Sorting these left-hand-side values in ascending order we find the ninety-fifth percentile value, which is k_1.

Step 2: Now, taking the values of k_1, k_2 and k_3, we draw the parameters from Θ_2 and generate the data set for this drawing. By doing this three thousand times, we calculate the probability of Type I error, that is, the average size over Θ_2 (we call it s_2) of the test. If s_2 is equal to the nominal level ($\alpha = 0.05$), go to step 3. If s_2 is greater than 0.05, we make k_2 bigger, otherwise, we make k_2 smaller and k_3 remains fixed, and we repeat steps 1 and 2.

Step 3: When s_2 = 0.05, again using the values of k_1, k_2 and k_3, we draw the parameters from Θ_3 and generate the data set for this drawing. By repeating this three thousand times, we calculate the average size over Θ_3 (we call it s_3) of the test. If s_3 is equal to the nominal level ($\alpha = 0.05$), we have the final values of the constants k_1, k_2 and k_3. If s_3 is greater than 0.05, we make k_3 bigger, otherwise, we make k_3 smaller, and repeat steps 1, 2 and 3.

Once we obtained the values of the constants, we calculated the sizes and powers of the test, which are discussed in the following section. The values of the constants k_1, k_2 and k_3 for different design matrices are given in Table 2.

Table 2
Final values of the constants k_1, k_2 and k_3.

Sample size	Values of the constants		
n	k_1	k_2	k_3
For design matrix $X1$:			
20	-0.1022	3.20	-0.16
30	0.0879	2.42	-0.0083
50	0.111004	1.85	-0.000005
75	0.08634965434	1.81	-0.00000000005
For design matrix $X2$:			
30	0.129311	2.17	-0.0113
50	0.228497	1.64	-0.0000011
69	0.130247299	1.61	-0.000000007
For design matrix $X3$:			
30	0.074187	2.5	-0.005
50	0.126533	1.9	-0.000004
70	0.0784357971	1.58	-0.0000000001
For design matrix $X4$:			
30	0.03175	2.75	-0.00025
50	0.137332786	2.084	-0.000000004

From Table 2, we observe that the k_2 values are always far larger than the k_1 and k_3 values for all design matrices. From the values of the constants, we also observe that there is a decreasing trend for the k_2 and k_3 values when the sample size is increasing. But there is no trend for k_1. Also k_1 always takes positive values with one exception being for $X1$ with n = 20. Moreover, for some cases, the k_1 and k_3 values are very close to zero. In all cases, k_3 takes a negative value. When the sample size is increasing k_3 takes very small negative values. Note that Lehmann (1986) required the constants $k_1, k_2, \ldots, k_m$ to all be positive numbers. Arthanari and Dodge (1980, Result 4.5.1, p.196) have been able to generalize the lemma to only require the existence of real numbers $k_1, k_2, \ldots, k_m$.

Finally, we calculated the sizes and power of our proposed average power test for different values of ρ and γ respectively via a Monte Carlo simulation experiment. The results are discussed in the following section.

4. Results

The calculated sizes and powers of the optimal average power test are presented in Table 3. We first investigate whether the rejection probabilities of the test under the true H_0 are significantly different from the nominal level. For 3000 replications, the rejection probabilities outside the range [0.0397, 0.0603] are significantly different from five percent at the 0.01 level. The sizes of our test fall inside this range for both $n = 20$ and $n = 30$, for all cases. But there are some points that fall outside this range when the sample size increases. The size of any reasonable test with $\rho > 0.5$ should tend to zero as the sample size increases (see King (1983) for detailed arguments of why this is the case). Unfortunately, almost all asymptotic tests and our test do not have this property. We have forced the average size in third part of the parameter space to be 0.05 (see Table 1), we could have allowed for a lower size for Θ_3.

In our cases, the size function under H_0, show a clear tendency to decrease to zero as n increases, except for the endpoint. This may have happened due to the negative value of k_3.

With respect to the power function under H_a, it always increases as n increases. So, under H_a, the power function of the test is well behaved.

Table 3

Calculated sizes and powers at 5 percent significance level (3000 replications).

n	$H_0: u_t = \rho u_{t-1} + e_t$, $0 \le \rho < 1$					$H_a: u_t = \gamma e_{t-1} + e_t$, $0 < \gamma \le 1$				
	$\rho = .1$	0.3	0.5	0.7	0.9	$\gamma = .1$	0.3	0.5	0.7	0.9
For design matrix $X1$:										
20	0.057	0.046	0.047	0.043	0.044	0.056	0.077	0.184	0.404	0.672
30	0.046	0.050	0.059	0.047	0.049	0.048	0.099	0.295	0.654	0.910
50	0.037	0.050	0.051	0.027	0.040	0.041	0.125	0.452	0.858	0.992
75	0.029	0.044	0.048	0.016	0.041	0.034	0.150	0.586	0.957	0.999
For design matrix $X2$:										
30	0.050	0.059	0.066	0.053	0.045	0.053	0.113	0.312	0.660	0.909
50	0.042	0.049	0.049	0.028	0.033	0.048	0.141	0.451	0.861	0.991
69	0.034	0.043	0.046	0.022	0.036	0.038	0.147	0.572	0.945	0.998
For design matrix $X3$:										
30	0.042	0.046	0.059	0.048	0.041	0.045	0.083	0.250	0.563	0.853
50	0.042	0.044	0.056	0.037	0.037	0.045	0.116	0.416	0.823	0.982
70	0.041	0.057	0.061	0.025	0.023	0.043	0.167	0.596	0.948	0.998
For design matrix $X4$:										
30	0.042	0.052	0.060	0.038	0.035	0.044	0.097	0.263	0.551	0.764
50	0.038	0.042	0.046	0.025	0.019	0.041	0.104	0.400	0.816	0.976

5. Concluding remarks

This paper discusses a new test, called the optimal average power test, which is based on the generalised N-P lemma for testing a composite null hypothesis against a composite alternative. In particular, the new approach is applied to the problem of testing for AR(1) disturbances against MA(1) disturbances in the context of a linear regression model, with encouraging results. The small sample sizes and powers of the new test seem to be most satisfactory. The main practical difficulty is with finding the values of the constants for which the average size conditions of the test are simultaneously satisfied. In this approach, we only used the uniform prior distribution. It can easily be applied using a non-uniform prior distribution.

References

Arthanari, T.S., and Dodge, Y. (1980), *Mathematical Programming in Statistics*, John Wiley & Sons, New York.

Conte, C.D. (1965), *Elementary Numerical Analysis*, McGraw-Hill Book Company, Kōgakusha Company Ltd., Tokyo.

Cox, D.R., and Hinkley, D.V. (1974), *Theoretical Statistics*, Imperial College, London.

Ferguson, T.S. (1967), *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, New York and London.

Judge, G.G., Griffiths, W.E., Hill, R.C., and Lee, T.C. (1980), *The Theory and Practice of Econometrics*, John Wiley and Sons, New York.

King, M.L. (1979), Some aspects of statistical inference in the linear regression model. Unpublished Ph.D. thesis, University of Canterbury, Christchurch.

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. (1983), Testing for autoregressive against moving average errors in the linear regression model, *Journal of Econometrics*, 21, 35-51.

King, M.L. (1987), Towards a theory of point optimal testing, *Econometric Reviews*, 6, 169-218.

King, M.L. (1989), Testing for fourth-order autocorrelation in regression disturbances when first-order autocorrelation is present, *Journal of Econometrics*, 41, 285-301.

King, M.L., and McAleer, M., (1987), Further results on testing AR(1) against MA(1) disturbances in the linear regression model, *Review of Economic Studies*, 54, 649-663.

Lehmann, E.L. (1986), *Testing Statistical Hypotheses*, 2nd ed, John Wiley and Sons, New York.

Lehmann, E.L., and Stein, C. (1948), Most powerful tests of composite hypotheses, I. Normal distributions, *Annals of Mathematical Statistics*, 19, 495-516.

McAleer, M. (1995), The significance of testing empirical non-nested models, *Journal of Econometrics*, 67, 149-171.

Pesaran, M.H. (1974), On the general problem of model selection, *Review of Economic Studies,* 41, 153-171.

Silvapulle, P., and King, M.L. (1991), Testing moving average against autoregressive disturbances in the linear regression model, *Journal of Business & Economic Statistics*, 9, 329-335.

Verbyla, A.P. (1990), A conditional derivation of residual maximum likelihood, *Australian Journal of Statistics*, 32, 227-230.

APPENDIX: The Generalised Neyman-Pearson Lemma

Theorem (Lehmann, 1986, p.96): Let $f_1, f_2, \ldots, f_{m+1}$ be real valued functions on a Euclidean space **A** and integrable μ, and suppose that for given constants $c_1, c_2, \ldots, c_m$ there exists a critical region ω of the form H_0 is rejected if $x \in \omega$ satisfying

$$\int_\omega f_i d\mu = c_i, \quad i = 1, 2, \ldots, m. \tag{A1}$$

Let ζ be the class of critical regions, ω, for which (A1) holds.

1. Among all members of ζ there exists one that maximize

$$\int_\omega f_{m+1} d\mu. \tag{A2}$$

2. A sufficient condition for a member ζ to maximize (A2) is the existence of constants $k_1, k_2, \ldots, k_m$ such that

$$\begin{aligned} x \in \omega \ \ when \ \ f_{m+1}(x) \geq \sum_{i=1}^{m} k_i f_i(x) \\ x \notin \omega \ \ when \ \ f_{m+1}(x) < \sum_{i=1}^{m} k_i f_i(x). \end{aligned} \tag{A3}$$

3. If a member of ζ satisfies (A3) with $k_1, k_2, \ldots, k_m \geq 0$, then it maximize (A2) among all critical functions satisfying

$$\int_\omega f_i d\mu \leq c_i, \quad i = 1, 2, \ldots, m. \tag{A4}$$

4. The set M of points in m-dimensional space whose coordinates are $(\int_\omega f_1 d\mu, \ldots, \int_\omega f_m d\mu)$ for some critical region ω is convex and closed. If $(c_1, c_2, \ldots, c_m)$ is an inner point of M, then there exist constants $k_1, k_2, \ldots, k_m$ and a test ω satisfying (A1) and (A3), and a necessary condition of ζ to maximize (A2) is that (A3) holds.

FEA Working Paper Series*

2000-1 Jones, G.W. "Education, Equity and Exuberant Expectations: Reflections on South-East Asia", September 1999.

2000-2 Goh, K.L. "Problems of the Wald Test in Non-Linear Settings and Some Solutions", October 1999.

2000-3 Begum, N. and M.L. King. "A New Approach to Testing a Composite Null against a Composite Alternative", January 2000.

2000-4 Noor Azina, I. and A.N. Pettitt. "Monitoring Hospital Outcomes Using Markov Chain Theory", July 2000.

* Papers are available at: http://www.cc.um.edu.my/FEP/

FEA Working Paper Series

Objective and Scope:

The Faculty of Economics and Administration (FEA) Working Paper Series is published to encourage the dissemination and facilitate discussion of research findings related to economics, development, public policies, administration and statistics. Both empirical and theoretical studies will be considered. The FEA Working Paper Series serves mainly as an outlet for research on Malaysia and other ASEAN countries. However, works on other regions that bear important implications or policy lessons for countries in this region are also acceptable.

Information to Paper Contributors:

1) Two copies of the manuscript should be submitted to:
   Chairperson
   Publications Committee
   Faculty of Economics and Administration
   University of Malaya
   50603 Kuala Lumpur
   MALAYSIA

2) The manuscript must be typed in double spacing throughout on one side of the paper only, and should preferably not exceed 30 pages of A4 size paper, including tables, diagrams, footnotes and references.

3) The first page of the manuscript should contain
   (i) the title,
   (ii) the name(s) and institutional affiliation(s) of the author(s), and
   (iii) the postal and email address of the corresponding author.

   This cover page will be part of the working paper document.

4) The electronic file of the manuscript must be submitted. The file can be a Word, Word Perfect, pdf or post-script document. This will be posted at the Faculty's website (http://www.cc.um.edu.my/FEP/) for public access.

5) Contents of the manuscript shall be the sole responsibility of the authors and publication does not imply the concurrence of the FEA or any of its agents. Manuscripts must be carefully edited for language by the authors. Manuscripts are vetted and edited, if necessary, but not refereed. The author is, in fact, encouraged to submit a concise version for publication in academic journals.

6) When published, the copyright of the manuscript remains with the authors. Submission of the manuscript will be taken to imply permission accorded by the authors for FEA to publicise and distribute the manuscript as a FEA Working Paper, in its hardcopy as well as electronic form.