AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

MONASH

WP 11/98

ISSN 1440-771X
ISBN 0 7326 1048 6

MONASH UNIVERSITY

ANCORA IMPARO

AUSTRALIA

LIBRARY
NOV 24 1998
WITHDRAWN
GIANNINI FOUNDATION OF
AGRICULTURAL ECONOMICS

Comparisons of Estimators and Tests Based on Modified Likelihood and Message Length Functions

Mizan R. Laskar and Maxwell L. King

Working Paper 11/98
August 1998

DEPARTMENT OF ECONOMETRICS
AND BUSINESS STATISTICS

COMPARISONS OF ESTIMATORS AND TESTS BASED ON MODIFIED LIKELIHOOD AND MESSAGE LENGTH FUNCTIONS

Mizan R. Laskar and Maxwell L. King
Department of Econometrics and Business Statistics
Monash University
Clayton, Victoria 3168
Australia

Summary

The presence of nuisance parameters causes unwanted complications in statistical and econometric inference procedures. A number of modified likelihood and message length functions have been developed for better handling of nuisance parameters but they are not equally efficient. In this paper, we empirically compare different modified likelihood and message length functions in the context of estimation and testing of parameters from linear regression disturbances that follow either first-order moving average or first-order autoregressive error processes. The results show that estimators based on the conditional profile likelihood and tests based on the marginal likelihood are best. If there is a minor identification problem, the sizes of the likelihood ratio and Wald tests based on simple message length functions are best. The true sizes of the Lagrange multiplier tests based on message length functions are rather poor because the score functions of message length functions are biased.

Key words: linear regression model; marginal likelihood; conditional profile likelihood; first-order moving average errors; first-order autoregressive errors.

1. Introduction

Satisfactory statistical analysis of non-experimental data is an important problem in statistics and econometrics. Econometric models usually involve a large number of influences, most of which are not of immediate interest and give rise to nuisance parameters. Their presence causes unwanted complications in statistical inference. The standard solution to this problem in likelihood based inference is to use the profile or concentrated likelihood function. It is obtained by replacing nuisance parameters by their respective maximum likelihood (ML) estimators conditional on the parameters of interest in the classical likelihood function. Unfortunately in such situations, estimators and tests can perform poorly in small samples (see for example Bewley 1986, Cox and Reid 1987, King 1987, King and McAleer 1987, Moulton and Randolph 1989, and Chesher and Austin 1991). Earlier, Neyman and Scott (1948) warned that nuisance parameters can seriously compromise likelihood based inference.

The question which then arises is what method should be used to deal with nuisance parameters in order to improve estimators and tests? The marginal likelihood is one such method. Estimators and tests based on this likelihood have better small sample properties compared to those based on the classical likelihood function (Ara 1995, Cordus 1986, and Rahman and King 1998). A related approach known as residual maximum likelihood (REML) for estimating variance components in the linear regression model was introduced by Patterson and Thompson (1971). The marginal likelihood and REML have some problems, the former cannot be constructed in all situations and latter applies only to the disturbance parameters in the linear model. Barndorff-Nielsen (1983) and Cox and Reid (1987) proposed the modified profile likelihood (MPL) and the conditional profile likelihood (CPL) respectively as alternatives to the above methods. Also, using the combination of REML and CPL, Laskar and King (1998) derived the conditional profile restricted log-likelihood function (CPRL) for parameters in the covariance matrix of linear regression disturbances. In the context of first-order moving average (MA(1)) and first-order autoregressive (AR(1)) regression disturbances, they investigated the small sample properties of estimators and tests based on this new likelihood function and three other modified likelihood functions and compared them with those based on the profile likelihood function.

An alternative approach, known as minimum message length (MML) (see Wallace and Freeman 1987), is a information theoretic criteria for parameter estimation and model selection. The MML principle needs a prior distribution of the parameters, the square root of the determinant of the information matrix for the parameters and a likelihood function. Recently, Laskar and King (1996) derived six different message length functions for the general linear regression model using different prior distributions of the parameters and combinations of CPL. They investigated the small sample properties of estimators based on these message length functions in the context of MA(1) regression disturbances. Additionally, Laskar and King (1997b) investigated the small sample properties of different tests based on these message length functions in a similar context. There are many different modified likelihood and message length functions for handling nuisance parameters but for econometric problems where estimation and diagnostic testing are of interest, they are not equally efficient. Thus, it is important to investigate and discover the best approaches for handling nuisance parameters.

The aim of this paper is to empirically compare all the likelihood and message length functions in the context of estimation and testing of parameters involved in the covariance matrix of linear regression disturbances by extending and comparing the various Monte Carlo results of Laskar and King (1996, 1997a, 1997b, 1998). This will enable us to recommend the best methods for estimation and testing problems. In section 2, different likelihood and message length functions are presented for the general linear model. A Monte Carlo experiment, conducted to compare the estimators and tests based on all the likelihood and message length functions is reported in section 3. Concluding remarks are made in section 4.

2. Theory

Consider the linear regression model with non-spherical disturbances

$$y = X\beta + u;\ u \sim N(0, \sigma^2\Omega(\theta)) \tag{1}$$

where y is $n\times1$, X is $n\times k$, nonstochastic and of rank $k < n$, β is a $k\times1$ vector, $\Omega(\theta)$ is a symmetric matrix and θ is a $p\times1$ vector. This model generalizes a wide range of disturbance processes of the linear regression model of interest to statisticians and econometricians. These include all parametric forms of autocorrelated disturbances, all parametric forms of heteroscedasticity (in which case $\Omega(\theta)$ is a diagonal matrix), and error components models including those that result from random regression coefficients. The likelihood and log-likelihood for this model (excluding constants) are respectively

$$L(y;\theta,\sigma^2,\beta) \propto \sigma^{-n}|\Omega(\theta)|^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma^2}(y - X\beta)'\Omega(\theta)^{-1}(y - X\beta)\}, \tag{2}$$

$$l(y;\theta,\sigma^2,\beta) \propto -\frac{n}{2}\log\sigma^2 - \frac{1}{2}\log|\Omega(\theta)| - \frac{1}{2\sigma^2}(y - X\beta)'\Omega(\theta)^{-1}(y - X\beta) \tag{3}$$

and the log-profile (or concentrated) likelihood is

$$l_p(y;\theta) \propto -\frac{n}{2}\log\hat{\sigma}_\theta^2 - \frac{1}{2}\log|\Omega(\theta)| \tag{4}$$

where $\hat{\sigma}_\theta^2 = (y - X\hat{\beta}_\theta)'\Omega(\theta)^{-1}(y - X\hat{\beta}_\theta)/n$ and $\hat{\beta}_\theta = (X'\Omega(\theta)^{-1}X)^{-1}X'\Omega(\theta)^{-1}y$.

2.1. Modified Likelihood Functions

Tunnicliffe Wilson (1989) derived the marginal likelihood for θ in (1) as

$$\ell_m(y;\theta) = -\frac{1}{2}\log|\Omega(\theta)| - \frac{1}{2}\log|X'\Omega(\theta)^{-1}X| - \frac{m}{2}\log(\hat{u}'\Omega(\theta)^{-1}\hat{u}) \tag{5}$$

where $\hat{u} = y - X\hat{\beta}_\theta$ and $m = n - k$. Using the combination of REML and CPL, Laskar and King (1998) derived the CPRL function of θ for model (1) as

$$\bar{\ell}_{cpr}^*(y;\theta) = -\frac{(m-2)}{2m}\left[\log|\Omega(\theta)| - \log|X'\Omega(\theta)^{-1}X| - m\log(\hat{u}'\Omega(\theta)^{-1}\hat{u})\right]. \tag{6}$$

Using the work of Cox and Reid (1987), Laskar and King (1998) derived the CPL for θ in (1) as

$$l_{cp}(y;\theta) = -\frac{1}{2}\log\left|X'\Omega(\theta)^{-1}X\right| - \frac{(n-2)}{2n}\log|\Omega(\theta)| - \frac{(m-2)}{2}\log(\hat{u}'\Omega(\theta)^{-1}\hat{u}). \tag{7}$$

Based on the work of Cox and Reid (1993), Laskar and King (1998) also derived an approximate conditional profile likelihood (ACPL) for θ in (1) as

$$l_{acp}(y;\theta) = -\frac{m-2}{2}\log(\hat{u}'\Omega(\theta)^{-1}\hat{u}) - \frac{1}{2}\log|\Omega(\theta)| - \frac{1}{2}\log\left|X'\Omega(\theta)^{-1}X\right|$$
$$+\frac{1}{n}tr\left[\Omega(\theta)^{-1}\frac{\partial\Omega(\theta)}{\partial\theta}\right]_{\theta=\hat{\theta}}(\theta - \hat{\theta}) \tag{8}$$

where $\hat{\theta}$ is the ML estimator of θ from (4). From (5) and (6)

$$\bar{l}_{cpr}^*(y;\theta) = \frac{(m-2)}{m}l_m(y;\theta)$$

so that for the purposes of estimating θ, the marginal likelihood and the CPRL are equivalent. This is not necessarily true for likelihood based tests of θ, because scores, Hessians and maximized likelihoods will be different, although any differences obviously disappear as n increases.

2.2. Message Length Functions

MML is a Bayesian method which chooses estimators to minimize the length of an encoded form of the data made up of a model and the deviations from that model (residuals). Wallace and Dowe (1993) state that the MML principle is that the best possible conclusion to draw from the data is the theory which maximizes the product of the probability of the data occurring in the light of the theory with the prior probability of that theory.

For model (1), an approximate message length function given by Wallace and Freeman (1987) is

$$-\log\left[\frac{\pi(\theta,\sigma^2,\beta)L(\theta,\sigma^2,\beta)}{\sqrt{F(\theta,\sigma^2,\beta)}}\right]+\frac{s}{2}(1+\log K_s) \qquad (9)$$

where $\pi(\theta,\sigma^2,\beta)$ is a prior density for $\gamma = (\theta',\sigma^2,\beta',)'$, $F(\theta,\sigma^2,\beta)$ is the determinant of the information matrix, $s = k+p+1$, K_s is the s dimensional lattice constant which is independent of parameters, as given by Conway and Sloan (1988, p. 59-61). Wallace and Dowe (1994) mentioned, maximizing (9) is equivalent to maximizing the average of the log-likelihood function over a region of size proportional to $1/\sqrt{F(\theta,\sigma^2,\beta)}$ while the ML estimator maximizes the likelihood function at a single point. The value of θ which minimizes (9) is the MML estimate of θ with accuracy

$\delta = 1/\sqrt{(K_s)^s F(\theta,\sigma^2,\beta)})$. Inclusion of $\pi(\theta,\sigma^2,\beta)$ and $\sqrt{F(\theta,\sigma^2,\beta)}$ help reduce the measure of uncertainty, their ratio is dimension free and invariant to reparameterization (Wallace and Dowe 1993). Since MML is a Bayesian method and depends on the choice of prior density of the parameters, there is scope in selecting the prior. Using different prior densities and combinations of CPL and message length functions, Laskar and King (1996) derived six different message length functions for the general linear regression model which are

$$ML_1 = \frac{m-1}{2}\log\sigma^2 + \frac{1}{2}\log|\Omega(\theta)| + \frac{1}{2\sigma^2}u'\Omega(\theta)^{-1}u + \frac{1}{2}\log\left|X'\Omega(\theta)^{-1}X\right|$$

$$+\frac{1}{2}\log\left(n\times tr\left[-\frac{\partial\Omega(\theta)^{-1}}{\partial\theta}\frac{\partial\Omega(\theta)}{\partial\theta}\right] - \left\{tr\left[\Omega(\theta)^{-1}\frac{\partial\Omega(\theta)}{\partial\theta}\right]\right\}^2\right)$$

$$+\frac{s}{2}(1+\log K_s) - \log 2, \tag{10}$$

$$ML_2 = \frac{m}{2}\log\sigma^2 + \frac{1}{2}\log|\Omega(\theta)| + \frac{1}{2\sigma^2}u'\Omega(\theta)^{-1}u + \frac{1}{2}\log\left|X'\Omega(\theta)^{-1}X\right|$$

$$+\frac{1}{2}\log\left(n\times tr\left[-\frac{\partial\Omega(\theta)^{-1}}{\partial\theta}\frac{\partial\Omega(\theta)}{\partial\theta}\right] - \left\{tr\left[\Omega(\theta)^{-1}\frac{\partial\Omega(\theta)}{\partial\theta}\right]\right\}^2\right)$$

$$+\frac{s}{2}(1+\log K_s) - \log 2, \tag{11}$$

$$CPML_1 = \frac{m-k-3}{2}\log\hat{\delta}_1 + \frac{k+1}{n+k+1}\log|\Omega(\theta)| + \log\left|X_\theta^{\dagger\prime}X_\theta^\dagger\right|$$

$$+\frac{1}{2}\log\left(n\times tr\left[-\frac{\partial\Omega(\theta)^{-1}}{\partial\theta}\frac{\partial\Omega(\theta)}{\partial\theta}\right] - \left\{tr\left[\Omega(\theta)^{-1}\frac{\partial\Omega(\theta)}{\partial\theta}\right]\right\}^2\right) \tag{12}$$

where $u_\theta^\dagger = y_\theta^\dagger - X_\theta^\dagger\beta$, $X_\theta^\dagger = D(\theta)^{-\frac{1}{2}}X$, $D(\theta) = \Omega(\theta)/|\Omega(\theta)|^{\frac{1}{(n+k+1)}}$, $y_\theta^\dagger = D(\theta)^{-\frac{1}{2}}y$, $\hat{\delta}_1 = \hat{u}_\theta^{\dagger\prime}\hat{u}_\theta^\dagger/(n-k-1)$, $\hat{u}_\theta^\dagger = y_\theta^\dagger - X_\theta^\dagger\hat{\beta}_\theta^\dagger$ and $\hat{\beta}_\theta^\dagger = (X_\theta^{\dagger\prime}X_\theta^\dagger)^{-1}X_\theta^{\dagger\prime}y_\theta^\dagger$.

$$CPML_2 = \frac{m-k-2}{2}\log\hat{\delta}_2 + \frac{k}{n+k}\log|\Omega(\theta)| + \log\left|X_\theta^{*\prime} X_\theta^*\right|$$

$$+\frac{1}{2}\log\left(tr\left[-\frac{\partial\Omega(\theta)^{-1}}{\partial\theta}\frac{\partial\Omega(\theta)}{\partial\theta}\right] - \left\{tr\left[\Omega(\theta)^{-1}\frac{\partial\Omega(\theta)}{\partial\theta}\right]\right\}^2\right) \quad (13)$$

where $\hat{\delta}_2 = \hat{u}_\theta^{*\prime}\hat{u}_\theta^* / m$, $\hat{u}_\theta^* = y_\theta^* - X_\theta^*\hat{\beta}_\theta^*$, $\hat{\beta}_\theta^* = (X_\theta^{*\prime} X_\theta^*)^{-1} X_\theta^{*\prime} y_\theta^*$, $X_\theta^* = G_1(\theta)^{-\frac{1}{2}} X$,

$y_\theta^* = G_1(\theta)^{-\frac{1}{2}} y$ and $G_1(\theta) = \Omega(\theta) / |\Omega(\theta)|^{\frac{1}{(n+k)}}$.

$$AML_1 = \frac{m-1}{2}\log\delta + \frac{1}{2\delta}u_\theta' u_\theta + \frac{1}{2}\log|X_\theta' X_\theta| + \frac{1}{2}\log|C(\theta)|, \quad (14)$$

$$AML_2 = \frac{m}{2}\log\delta + \frac{1}{2\delta}u_\theta' u_\theta + \frac{1}{2}\log|X_\theta' X_\theta| + \frac{1}{2}\log|C(\theta)| \quad (15)$$

where $u_\theta = y_\theta - X_\theta\beta$, $X_\theta = G(\theta)^{-\frac{1}{2}} X$, $y_\theta = G(\theta)^{-\frac{1}{2}} y$, $\sigma^2 = \delta / |\Omega(\theta)|^{\frac{1}{n}}$, $G(\theta)$ is an $n\times n$ matrix comprised of $\Omega(\theta)$ with each element divided by $|\Omega(\theta)|^{\frac{1}{n}}$ and the $(i,j)^{\text{th}}$ element of the $p\times p$ matrix $C(\theta)$ is $\frac{1}{2}tr\left[\frac{\partial^2 G(\theta)^{-1}}{\partial\theta_i\partial\theta_j}G(\theta)\right]$.

In order to empirically compare estimators and tests based on different likelihood and message length functions, we need to construct different test statistics. Details of the likelihood ratio (LR), Lagrange multiplier (LM), Wald, alternative Wald (AW) and null Wald (NW) tests based on all the likelihood and message length functions in the context of testing $H_0: \theta = \theta_0$ against $H_a: \theta \neq \theta_0$ in (1) are given in Laskar and King (1997a, 1997b, 1998). Laskar and King (1998) estimated the MA(1) disturbance parameter constrained between -1 to 1, because of the identification problem for MA(1) disturbances. It is well known that there is a non-zero probability of getting ML estimators of -1 or 1 for the MA(1) disturbance parameter (Shephard 1993). The score with respect to the MA(1) parameter is discontinuous and the information matrix is not well defined at those two points. As a result, Laskar and King (1998) faced the problem of nonmonotonicity of the power curve of the Wald test. They initially tackled this problem by rejecting the null hypothesis whenever the estimate of the MA(1) disturbance parameter is ± 1 and called this the AW test. Unfortunately the AW test cannot totally solve this problem because it takes into account boundary values of the parameter estimates only. Laskar and King (1997a) overcame this problem by replacing the unknown parameter values in the variance component of the Wald test with their null hypothesis values rather than their estimated values and called the resultant test the NW test.

3. Monte Carlo Experiment

Laskar and King (1998) investigated the small sample properties of estimators and the LR, LM, Wald and AW tests based on different likelihood functions in the context of MA(1) and AR(1) regression disturbances. Also, Laskar and King (1997a) investigated the properties of NW tests based on different modified likelihood functions in the context of MA(1) regression disturbances. With respect to message length function based estimation and testing, Laskar and King (1996, 1997b) investigated the properties of estimators and tests in the context of MA(1) regression disturbances.

In order to compare the properties of estimators and tests of MA(1) regression disturbances or AR(1) regression disturbances based on different likelihood and message length functions, we consolidated the results from above papers with a further Monte Carlo experiment conducted to investigate estimators and tests based on message length functions with the disturbances of (1) generated by the AR(1) process

$$u_t = \rho u_{t-1} + \varepsilon_t \tag{16}$$

in which $\varepsilon_t \sim IN(0,\sigma^2)$, $t = 1,\ldots,n$. Unfortunately all the message length functions are not defined at $\rho = \pm 1$. The best way of tackling this problem is to restrict ρ to the interval

$$-0.9999 \leq \rho \leq 0.9999\,. \tag{17}$$

For details of the implication of restriction (17) see Laskar and King (1998).

3.1. Experimental Design

The first part of the study involved a comparison of the different MML estimators for the AR(1) parameter. Estimates based on (i) ML_1, (ii) ML_2, (iii) $CPML_1$, (iv) $CPML_2$, (v) AML_1 and (vi) AML_2 when ρ = -0.8, -0.4, 0, 0.4, 0.8 were used for the first comparison. The second part involved a comparison of sizes of message length based LR, LM and Wald tests using asymptotic critical values at the five percent level. The third part of the experiment was divided into two parts. In the first part, the Monte Carlo method was used to estimate appropriate critical values of each of the tests in order to compare the powers of all tests at approximately the same level. These critical values were calculated using 2000 replications. In the second part, powers of all the tests were calculated using these (simulated) critical values.

All the calculations were repeated 2000 times using the GAUSS (1996) constrained optimization routine but with particular care taken in choosing starting values (see Laskar 1998). The following X matrices used by Laskar and King (1996, 1997a, 1997b, 1998) were used with $n = 30$ and $n = 60$:

$X1$: (k = 5). A constant, quarterly Australian private capital movements, Government capital movements commencing 1968(1) and these two variables lagged one quarter as two additional regressors.

$X2$: (k = 3). A constant, quarterly seasonally adjusted Australian household disposable income and private final consumption expenditure commencing 1959(4).

$X3$: (k = 3). The regressors are the eigenvectors corresponding to the three smallest eigenvalues of the $n \times n$ tridiagonal matrix whose main diagonal elements are 2, except for the top left and bottom right elements which are both 1 and whose elements in the leading off-diagonals are all -1.

$X4$: ($k = 2$). A constant and a linear trend.

These matrices reflect a variety of behaviour. The capital movements regressors in $X1$ are rapidly changing with a high degree of seasonality. This is in contrast to the relatively smooth regressors of $X2$ (seasonally adjusted quarterly data). The regressors in $X3$ are smoothly evolving and include an intercept. They cause the Durbin-Watson statistic, which is an approximately locally best one-sided test against both MA(1) and AR(1) disturbances (King and Evans 1988), to attain its upper bound.

3.2. Empirical Comparisons of Estimators Based on Likelihood and Message Length Functions

Estimated bias, standard deviation, skewness and kurtosis of all the estimators were computed and summarized using the loss function, $|\text{bias}| + \frac{1}{\lambda}(\text{standard deviation}) + \frac{1}{\lambda^2}|\text{skewness}| + \frac{1}{\lambda^3}|\text{kurtosis} - 3|$ where $\lambda = 3$ (Laskar and King 1998).

Then we compared all the estimators by ranking in turn of their absolute value of bias, standard deviation, the absolute value of skewness, the absolute value of (kurtosis - 3) and loss. These statistics of the estimators based on profile likelihood, marginal likelihood, CPL, ACPL, ML_1, $CPML_1$, ML_2, $CPML_2$, AML_1 and AML_2 were combined from this Monte Carlo experiment and those of Laskar and King (1996, 1998). For each X matrix, value of n and γ or ρ, they were ranked from 1 to 10 in ascending order, from the smallest to the largest values. A selection of averages of these ranks with their standard error in parentheses and the rank of this average rank are presented in Table 1 and Table 2. The following discussion is based on full set of results which are available from the authors on request.

The results in Table 1 for combined MA(1) and AR(1) processes with n = 30 and 60 indicate that the CPL based estimators are best, but not always. For MA(1) processes with n = 60 and n = 30, they are second and third best respectively, but for combined n = 30 and 60, they are best. For AR(1) processes with n = 60 and n = 30 they are best and second best respectively and for combined n = 30 and 60, they are second best. They are best for combined MA(1) and AR(1) processes with all sample sizes. The average ranks for individual statistics sheds more light on the performance of CPL based estimators. CPL based estimators reduce bias, but not very much. Their average ranks for bias vary from two to five, while for standard deviations it is typically eight or nine. With respect to skewness and kurtosis, CPL based estimators typically rank first.

The results in Table 1 suggest that marginal likelihood based estimators are second best, and are more uniform in their ranking performance than other estimators with the smallest standard error of 0.164, compared to 0.201 for the CPL based estimators. They perform relatively better for AR(1) processes than for MA(1) processes. For AR(1) processes with $n = 30$, $n = 60$ and combined $n = 30$ and 60, they are best. In contrast, for MA(1) processes with $n = 30$ and for combined $n = 30$ and 60, they are fourth best. For combined MA(1) and AR(1) processes with $n = 30$ and 60, they are second best, while for $n = 30$, they are third best. The marginal likelihood reduces estimators' bias compared to all other likelihood and message length functions. This is why Ara and King (1993) and Rahman and King (1998) found better small sample properties of marginal likelihood based tests compared to classical likelihood based tests.

The results in Table 1 suggest that ML_1 based estimators are the third best. They do well for MA(1) processes with $n = 30$ and combined $n = 30$ and 60 as well as for combined MA(1) and AR(1) processes with $n = 30$. For MA(1) processes with $n = 30$, skewness of ML_1 based estimators are smallest and for $n = 60$ and for combined $n = 30$ and 60, they are second best. There is a tendency for ML_1 based estimators to do slightly better when $n = 30$ while marginal likelihood based estimators do better for $n = 60$.

ML_2 based estimators are fourth best. They perform well for MA(1) processes with $n = 30$. Biases of ML_2 based estimators are smallest for MA(1) processes with $n = 30$ and for combined $n = 30$ and 60 while for $n = 60$ they are the second smallest.

ACPL based estimators are fifth best. Kurtosis of these estimators are relatively low for AR(1) processes but for all other cases their average ranks are typically large.

AML_2 based estimators are sixth best. These estimators have relatively low standard deviations for all processes and all sample sizes. Their average ranks based on skewness for AR(1) processes with combined $n = 30$ and 60 are smallest. AML_1 based estimators are seventh best overall.

Overall, this analysis suggests that CPL based estimators are best with marginal likelihood estimators being second best. Although the latter are second best, they perform more uniformly in terms of ranking on a range of criteria with the smallest standard error of 0.164 compared to 0.201 for the CPL based estimator. The biases of marginal likelihood estimators are typically smaller than those of CPL based estimators. On the other hand, skewness and kurtosis of the CPL based estimators are smaller. When the values of γ and ρ are closer to zero, both marginal likelihood and CPL based estimators perform relatively poorly compared to AML_1 and AML_2 based estimators, but the latter are rather nonuniform in terms of their ranking. It is not difficult to conclude that marginal likelihood and CPL based estimators are close competitors. ML_1, ML_2 and ACPL based estimators are third, fourth and fifth best respectively. Typically, $CPML_1$, $CPML_2$ and profile likelihood based estimators are the worst of all estimators considered.

3.3. Empirical Comparisons Among Likelihood and Message Length Based Tests

The sizes and powers of all tests based on profile likelihood, marginal likelihood, CPL, CPRL, ACPL, ML_1, $CPML_1$, ML_2, $CPML_2$, AML_1 and AML_2 were combined from this Monte Carlo experiment, and those of Laskar and King (1997a, 1997b, 1998). The absolute value of (size - 0.05) for all eleven tests are ranked from the smallest to the largest values, and powers are ranked from the largest to the smallest values for each X matrix and values of n, γ and ρ. The average rank of size of different tests with its standard error in parenthesis and the ranking of the average rank for combined MA(1) and AR(1) processes, MA(1) processes and AR(1) processes for $n = 30$ and $n = 60$ are presented in Tables 3 and 4 respectively. The average rank of power with its standard error in parenthesis and the ranking of the average rank for all the tests based on marginal likelihood, CPL, ACPL, ML_1, ML_2, AML_1 and AML_2 for combined MA(1) and AR(1) processes, MA(1) processes and AR(1) processes for $n = 30$ and $n = 60$ are presented in Tables 5 and 6 respectively. The marginal likelihood and CPRL are equivalent for the purpose of estimating γ and ρ, but they have scores and information matrices that differ by a multiplicative constant. As a result, the small sample sizes of asymptotic tests based on these two likelihood functions will differ, although the tests are identical if simulated (or exact) critical values are used. For these reasons, CPRL was not considered for power comparisons. Also $CPML_1$, $CPML_2$ and profile likelihood were dropped from the power comparison because they produce highly biased power curves. (Results are available from the authors on request.)

3.3.1. Comparisons of Sizes

The ranking of the average ranks for sizes of the LR tests reflects that for MA(1) processes with $n = 30$, sizes of the ML_1 based tests are best overall and those of CPRL and AML_1 are jointly second best. The fourth best are the sizes of the AML_2 based LR tests. However, for $n = 60$, the CPRL based sizes are closest overall to the nominal size and those of ML_1 and ACPL are second and third best respectively. The sizes of the profile likelihood based LR tests have the worst performance. It appears that for MA(1) processes, sizes of the marginal likelihood based LR tests are not as good as many of the tests already mentioned. This may be because of the identification problem associated with MA(1) processes. The sizes of the marginal likelihood based LR tests for AR(1) processes rank closest to the nominal size while those for MA(1) processes are ninth best, so clearly they depend on data generating processes. For AR(1) processes, sizes of the ML_1 based LR tests are second best and very similar both for MA(1) and AR(1) processes.

The sizes of the marginal likelihood based LM tests are the most promising, of the eleven tests, with sizes ranked closest to the nominal size for both MA(1) and AR(1) processes; those based on ACPL, CPL and CPRL are typically second, third and fourth best, respectively. In contrast, sizes of the message length based LM tests rank relatively poorly. These results are not surprising as Mahmood and King (1997) observed that the LM test based on an unbiased score function is likely to have good small sample properties and reported that the score functions based on marginal likelihood and CPRL are unbiased while those of message length functions are biased.

The Wald test was not constructed for MA(1) processes for all the likelihood and message length functions because of the identification problem mentioned in section 3.1. Sizes of the Wald test based on ML_2 are best for AR(1) processes with both n = 30 and 60 and those based on AML_1 for n = 30 and the marginal likelihood for n = 60 are second best. NW tests were constructed only for MA(1) processes and in terms of rankings, their sizes seem to be more accurate when modified likelihood functions are replaced by message length functions. For n = 30 and n = 60, sizes of the NW tests based on ML_2 are best and those based on ML_1 are second best. These results may be related to the identification problem outlined in section 3.1. When there is a lack of identification, the information matrix term in the message length functions appears to react and overcome this problem. In this regard, Martin (1997) used the Bayesian method for inference in autoregressive fractionally integrated moving average processes and reported that the use of Jeffreys prior helped to offset an identification problem in the likelihood function. Jeffreys prior is proportional to the determinant of the information matrix, one of the components of the message length function. This may explain why message length based Wald tests typically have better small sample sizes compared to those based on modified likelihood functions.

Overall, the results show that sizes of the marginal likelihood based LM tests for both MA(1) and AR(1) processes and those of LR tests for AR(1) processes are most accurate and closest to the nominal size. The LM test based on message length functions perform poorly and show strong size distortion. On the other hand, the LR test based on ML_1 is quite impressive, having best and second best sizes for MA(1) processes with n = 30 and n = 60, respectively and second best sizes for AR(1) processes. Wald and NW tests perform relatively better for message length functions and, particularly those tests based on ML_1, ML_2 and AML_1, have desirable sizes. Clearly, sizes of profile likelihood, $CPML_1$ and $CPML_2$ based tests are away from the nominal size, indicating their poor performance in testing problems.

Finally, if we take the average of all ranks for each test, and then use this to give an overall ranking based on size then we get the following order (from best to worst): marginal likelihood, ML_1, CPL and AML_1 (ranked equal third), CPRL, ACPL, AML_2, ML_2, $CPML_2$, $CPML_1$ and profile likelihood.

3.3.2. Comparisons of Powers

The ranking of the average ranks from Tables 5 and 6 reflects that for MA(1) processes, power curves of marginal likelihood and AML_1 based LR tests are best centred and second and fourth best, respectively. In contrast, power curves of the CPL based LR test show great bias, being best (worst) powers for positive (negative) values of γ and ρ. This pattern is completely the opposite for AR(1) processes, with highly biased power curves for marginal likelihood and AML_1 and relatively better centred power curves for CPL, ACPL and ML_1. Finally, for MA(1) processes, powers of the marginal likelihood based LR test rank highest overall and those based on ACPL and CPL are second and third highest respectively. For AR(1) processes, powers of the CPL based test rank best and those based on the marginal likelihood are second best overall. In contrast, for combined MA(1) and AR(1) processes, powers of the marginal likelihood based LR tests rank best and those based on CPL and ACPL are the second best.

There is no clear pattern for the powers of LM tests, but it appears that power curves of the CPL based tests are poorly centred compared to those of the marginal likelihood. The power curves of AML_1 and AML_2 based LM tests are best centred for AR(1) processes and those of CPL and ACPL are best and second best respectively for combined MA(1) and AR(1) processes, while those based on CPL and ACPL have the same rank and are best for AR(1) processes. It seems that power curves of the ACPL based LM test are bettered centred, the second best for combined MA(1) and AR(1) processes and best for MA(1) processes. The power curves based on ML_2 are also quite impressive, being the third best for combined MA(1) and AR(1) processes.

Those based on ML_1 are close competitors, having the fourth best power and very consistent rankings.

With respect to powers of Wald tests, when based on marginal likelihood and ML_1, they are highly biased and those based on AML_2 are best centred. Power curves of the NW tests based on AML_1 and AML_2 are best centred with relatively low power and they are sixth and seventh best respectively. There is no other clear pattern, but it seems that powers of the ML_2 based tests are best and those based on the ML_1 are second best.

When we consider the average ranks based on all tests, a slightly different pattern in powers is observed. Powers of the marginal likelihood based tests are consistently best, followed by the powers of the ACPL based tests which are second best for both MA(1) and AR(1) processes. Those based on CPL, AML_2, AML_1, ML_2 and ML_1 are third, fourth, fifth, sixth and seventh best respectively.

Using average ranks of all ranks for each tests based on power, we come to the following conclusions. Overall, the power results clearly favour marginal likelihood based tests. ACPL based tests appear to be second best while CPL based tests are third best. The powers of AML_2, AML_1, ML_2 and ML_1 based tests are fourth, fifth, sixth and seventh best respectively. Our ranking is very revealing in that modified likelihood functions rank ahead of message length functions in terms of power. There are some cases where message length based tests produce better centred power curves with relatively lower powers. Above all, it is easy to conclude in favour of marginal likelihood based tests when power is the criteria.

4. Conclusions

In this paper, we empirically compare estimators and tests of the parameters involved in the variance-covariance matrix of the linear regression disturbances based on eleven different likelihood and message length functions, most of which are designed for proper handling of nuisance parameters. The estimation results show that estimators based on the CPL are best overall out of the eleven estimation methods. The CPL reduces skewness and brings kurtosis close to three. The second best are marginal likelihood based estimators which show the best uniformity in terms of their ranking over the various criteria and circumstances considered. The average sizes of all likelihood and message length based tests show that overall, marginal likelihood based tests have the most accurate sizes; particularly for LM test, these are most impressive. All message length functions perform poorly based on power compared to likelihood functions. Powers of marginal likelihood based LR tests and CPL based LM tests are best for both MA(1) and AR(1) processes. Overall, powers of marginal likelihood based, ACPL based and CPL based tests are best, second best and third best, respectively.

In conclusion, it can be said that, at least in the context of MA(1) and AR(1) linear regression disturbances, the marginal likelihood may be the best likelihood for handling nuisance parameters. But, for more general inference problems, there are situations where marginal likelihood cannot be applied. Then the CPL and ACPL may be preferred alternatives. The former has the problem of non-uniqueness of reparameterization to achieve orthogonality and cannot always be found. The latter is applicable for scalar parameters and depends on the profile likelihood based estimators. In addition to these, MML is a Bayesian method and depends on the prior distribution of the parameters. Our results show that, in some situations, ML_1 and ML_2 perform very well and have the ability to deal with a minor identification problem. However, there is the need to choose an appropriate prior. Also, message length based LM tests have poor small sample properties due to their biased score functions. All this confirms that none of the methods is superior in all situations, and their performance may depend on the nature of the problem at hand. However, for the problems considered in this paper, we conclude in favour of the marginal likelihood.

Acknowledgment

This research was supported in part by an ARC grant and in part by a writing up scholarship jointly organized by the Faculty of Business and Economics and the Department of Econometrics and Business Statistics at Monash University. We are grateful for comments from and discussions with David L. Dowe, Gael M. Martin, C.S. Wallace and T. DiCiccio.

References

Ara, I. (1995). Marginal likelihood based tests of regression disturbances, unpublished Ph.D. thesis, Monash University.

Ara, I. and M.L. King (1993). Marginal likelihood based tests of regression disturbances, mimeo (Monash University).

Barndorff-Nielsen, O.E. (1983). On a formula for the distribution of the maximum likelihood estimator, *Biometrika* 70, 343-365.

Bewley, R. (1986). *Allocation Models: Specification, Estimation and Applications*, Ballinger, Boston.

Chesher, A. and G. Austin (1991). The finite sample distributions of heteroskedasticity robust Wald statistics, *J. Econometrics* 47, 153-173.

Conway, J.H. and N.J.A. Sloan (1988). *Sphere Packings, Lattices and Groups*, Springer-Verlag, London.

Cordus, M. (1986). The use of the marginal likelihood in testing for serial correlation in time series regression, unpublished M. Phil. thesis, University of Lancaster.

Cox, D.R. and N. Reid (1987). Parameter orthogonality and approximate conditional inference (with discussion), *J. R. Statist. Soc. Ser.* B 49, 1-39.

Cox, D.R. and N. Reid (1993). A note on the calculation of adjusted profile likelihood, *J. R. Statist. Soc. Ser.* B 55, 467-471.

GAUSS (1996). *GAUSS Constrained Optimization*: *Application Module*, Aptech System, Inc., Washington.

King, M.L. (1987). Testing for autocorrelation in linear regression models: A survey, *Specification Analysis in the Linear Model*, M.L. King and D.E.A. Giles (eds.), Routledge and Kegan Paul, London, 19-73.

King, M.L. (1996). Hypothesis testing in the presence of nuisance parameters, *J. Statist. Plann. Inference* 50, 103-120.

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 M. McAleer (1987). Further results on testing AR(1) against MA(1) disturbances in the linear regression model, *Rev. Econ. Studies* 54, 649-663.

Laskar, M.R. (1998). Estimation and testing of linear regression disturbances based on modified likelihood and message length functions, Unpublished Ph.D. thesis, Monash University.

Laskar, M.R. and M.L. King (1996). Estimation of regression disturbances based on minimum message length, *Proc. Con. ISIS'96: Information, Statistics and Induction in Science,* D.L. Dowe, K.B. Korb and J.J. Oliver (eds.), World Scientific, Singapore, 92-101.

Laskar, M.R. and M.L. King (1997a). Modified Wald test for regression disturbances, *Econ. Letters* 56, 5-11.

Laskar, M.R. and M.L. King (1997b). Testing of regression disturbances based on minimum message length, *Proc. Econom. S. Australasian Meeting*, P. Bardsley and V.L. Martin (eds.), University of Melbourne 2, 179-202.

Laskar, M.R. and M.L. King (1998). Estimation and testing of regression disturbances based on modified likelihood functions, *J. S. Plann. Inference.*, forthcoming.

Mahmood, M. and M.L. King (1997). Dealing with nuisance parameters in likelihood based estimating equations, mimeo (Monash University).

Martin, G.M. (1997), Fractional cointegration: Bayesian inference using Jeffreys prior, mimeo (Monash University).

Moulton, B.R. and W.C. Randolph, (1989). Alternative tests of the error components model, *Econometrica* 57, 685-693.

Neyman, J. and E.L. Scott (1948). Consistent estimates based on partially consistent observations, *Econometrica* 16, 1-32.

Patterson, H.D. and R. Thompson (1971). Recovery of interblock information when block sizes are unequal, *Biometrika* 58, 545-554.

Rahman, S. and M.L. King (1998). Marginal likelihood score based tests of regression disturbances in the presence of nuisance parameters, *J. Econometrics.* 82, 81-106.

Shephard, N. (1993). Distribution of the ML estimator of an MA(1) and a local level model, *Econometric Theory* 9, 377-401.

Tunnicliffe Wilson, G. (1989). On the use of marginal likelihood in time series model estimation, *J. R. Statist. Soc. Ser.* B 51, 15-27.

Wallace, C.S. and D.L. Dowe (1993). MML estimation of the von Mises concentration parameter, Technical Report No.93/193, Monash University.

Wallace, C.S. and D.L. Dowe (1994). Intrinsic classification by MML - the Snob program, *Proc. Seventh Australian Joint Conference on Artificial Intelligence*, 37-44.

Wallace, C.S. and P.R. Freeman (1987). Estimation and inference by compact coding, *J. R. Statist. Soc. Ser.* B 49, 240-265.

Table 1. Average ranks (Av.R), standard errors (SE) of average ranks and rank of average rank (R) of loss of estimators based on different likelihood and message length functions.

Function	n	MA(1) and AR(1)			MA(1) only			AR(1) only		
		Av.R	SE	R	Av.R	SE	R	Av.R	SE	R
Profile	30 and 60	7.050	0.369	8	8.325	0.409	10	5.775	0.548	6
	30	7.500	0.533	9	9.300	0.385	10	5.804	0.757	4
	60	6.600	0.537	8	7.350	0.662	8	5.850	0.828	6
Marginal	30 and 60	3.513	0.164	2	4.125	0.203	4	2.900	0.171	1
	30	3.700	0.217	3	4.450	0.336	4	5.019	0.254	1
	60	3.325	0.173	1	3.800	0.213	1	2.850	0.233	1
CPL	30 and 60	3.413	0.201	1	3.700	0.271	1	3.125	0.293	2
	30	3.350	0.290	2	3.300	0.391	3	5.313	0.421	2
	60	3.475	0.284	2	4.100	0.362	2	2.850	0.399	1
ACPL	30 and 60	4.563	0.221	5	5.600	0.248	6	3.525	0.284	3
	30	4.875	0.291	5	6.150	0.319	5	5.636	0.409	3
	60	4.250	0.288	3	5.050	0.344	7	3.450	0.394	3
ML_1	30 and 60	3.900	0.199	3	3.775	0.361	2	4.025	0.174	4
	30	3.250	0.240	1	2.800	0.345	1	6.925	0.273	6
	60	4.475	0.300	4	4.750	0.561	6	4.200	0.213	4
$CPML_1$	30 and 60	8.613	0.223	10	7.850	0.382	9	9.375	0.163	10
	30	8.425	0.338	10	7.100	0.657	9	8.861	0.165	10
	60	8.800	0.218	10	8.600	0.328	10	9.000	0.290	10
ML_2	30 and 60	4.438	0.210	4	3.850	0.331	3	5.025	0.225	5
	30	4.050	0.286	4	3.000	0.348	2	6.313	0.258	8
	60	4.825	0.310	5	4.700	0.503	4	4.950	0.373	5
$CPML_2$	30 and 60	7.825	0.259	9	6.975	0.350	8	8.675	0.335	9
	30	7.375	0.372	8	6.200	0.569	7	8.250	0.507	9
	60	8.275	0.286	9	7.750	0.339	9	8.800	0.439	9
AML_1	30 and 60	5.950	0.325	7	5.600	0.544	6	6.300	0.355	8
	30	6.300	0.425	7	6.500	0.587	8	6.250	0.514	7
	60	5.600	0.524	7	4.700	0.886	4	6.500	0.505	8
AML_2	30 and 60	5.600	0.328	6	5.175	0.533	5	6.025	0.378	7
	30	6.025	0.442	6	6.150	0.662	5	5.900	0.561	5
	60	5.175	0.492	6	4.200	0.793	3	6.150	0.514	7

Table 2. Average ranks (Av.R), standard errors (SE) of average ranks and rank of average rank (R) of estimated bias (B), standard deviation (SD), skewness (S) and kurtosis (K) of estimators based on different likelihood and message length functions for $n = 30$ and 60.

		MA(1) and AR(1)			MA(1) only			AR(1) only		
Function	Statistics	Av.R	SE	R	Av.R	SE	R	Av.R	SE	R
Profile	B	6.375	0.392	8	7.250	0.475	8	5.500	0.598	6
	SD	5.275	0.416	3	7.800	0.536	10	2.950	0.371	3
	S	6.713	0.349	9	7.975	0.444	10	5.450	0.464	7
	K	6.187	0.350	9	6.875	0.500	10	5.500	0.472	7
Marginal	B	3.088	0.182	1	3.525	0.322	3	2.650	0.146	1
	SD	6.213	0.196	8	6.750	0.335	7	5.675	0.169	6
	S	4.688	0.211	4	4.875	0.261	4	4.500	0.332	5
	K	4.625	0.236	2	4.800	0.365	2	4.450	0.304	3
CPL	B	3.725	0.234	3	4.252	0.328	5	2.925	0.285	2
	SD	6.313	0.215	9	5.725	0.302	5	6.900	0.279	9
	S	3.738	0.227	1	3.500	0.256	1	3.975	0.375	3
	K	3.962	0.243	1	4.125	0.266	1	3.800	0.410	1
ACPL	B	4.050	0.221	5	5.000	0.300	6	3.100	0.250	3
	SD	6.613	0.200	10	6.825	0.336	8	6.400	0.217	7
	S	4.938	0.229	5	5.700	0.310	5	4.175	0.295	4
	K	4.725	0.284	3	5.650	0.412	7	3.800	0.338	2
ML_1	B	3.588	0.153	2	3.425	0.258	2	3.750	0.163	4
	SD	6.038	0.171	6	6.700	0.266	6	5.375	0.159	5
	S	4.575	0.280	3	3.675	0.389	2	5.475	0.353	8
	K	5.112	0.294	5	5.125	0.461	4	5.100	0.371	5
$CPML_1$	B	8.825	0.186	10	8.600	0.318	10	9.050	0.189	10
	SD	6.188	0.412	7	3.850	0.404	3	8.525	0.495	10
	S	7.388	0.350	10	6.225	0.462	9	8.550	0.463	10
	K	7.112	0.395	10	6.650	0.549	9	7.575	0.566	10
ML_2	B	3.950	0.228	4	3.050	0.305	1	4.850	0.274	5
	SD	5.558	0.255	4	7.500	0.240	9	3.675	0.236	4
	S	4.388	0.262	2	4.200	0.364	3	4.575	0.379	6
	K	5.425	0.247	7	5.250	0.359	5	5.600	0.343	8
$CPML_2$	B	8.200	0.225	9	7.775	0.303	9	8.625	0.322	9
	SD	5.863	0.332	5	4.125	0.346	4	7.600	0.456	9
	S	5.675	0.352	8	5.775	0.436	6	5.575	0.558	9
	K	6.037	0.370	8	5.875	0.453	8	6.200	0.592	9
AML_1	B	6.088	0.301	7	5.550	0.492	7	6.625	0.333	8
	SD	1.789	0.109	1	1.800	0.157	1	1.775	0.154	2
	S	5.013	0.402	7	6.150	0.577	8	3.875	0.507	1
	K	5.100	0.347	4	5.150	0.522	3	5.050	0.465	4
AML_2	B	5.325	0.318	6	4.475	0.465	4	6.175	0.394	7
	SD	1.988	0.113	2	2.300	0.183	2	1.675	0.115	1
	S	5.000	0.379	6	6.075	0.556	7	3.925	0.462	2
	K	5.250	0.305	6	5.375	0.465	6	5.125	0.402	6

Table 3. Average ranks (Av.R), standard errors (SE) of average ranks and rank of average rank (R) of sizes of the LR, LM, Wald and NW tests based on different likelihood and message length functions for $n = 30$.

		MA(1) and AR(1)			MA(1) only			AR(1) only		
Function	Test	Av.R	SE	R	Av.R	SE	R	Av.R	SE	R
Profile	LR	10.00	0.627	11	10.75	0.250	11	9.250	1.180	9
	LM	7.750	1.048	7	8.500	0.500	8	7.000	2.120	6
	Wald*	9.875	0.875	11	11.00	000	11	8.750	1.650	5
Marginal	LR	4.750	1.319	4	8.000	0.913	9	1.500	0.500	1
	LM	1.125	0.125	1	1.250	0.250	1	1.000	000	1
	Wald*	6.250	1.114	8	8.750	0.629	9	3.750	1.110	3
CPL	LR	6.250	0.881	7	7.250	1.320	7	5.250	1.110	6
	LM	3.375	0.263	3	3.500	0.289	4	3.250	0.479	3
	Wald*	5.375	0.263	5	5.500	0.500	5	5.250	0.250	5
CPRL	LR	5.000	1.134	5	2.250	0.750	2	7.750	0.629	8
	LM	3.625	0.375	4	3.250	0.250	3	4.000	0.707	4
	Wald*	5.875	0.441	6	6.000	0.816	7	5.750	0.479	6
ACPL	LR	7.375	0.532	8	8.250	0.629	10	6.500	0.645	7
	LM	2.375	0.375	2	1.750	0.250	2	3.000	0.707	2
	Wald*	6.750	1.176	9	9.750	0.250	10	3.750	0.629	3
ML_1	LR	2.375	0.324	1	2.000	0.408	1	2.750	0.479	2
	LM	8.000	0.500	8	8.000	0.707	7	8.000	0.816	9
	Wald*	5.125	0.972	4	2.750	0.750	2	7.500	0.289	7
$CPML_1$	LR	8.375	0.865	9	7.000	1.290	6	9.750	0.750	10
	LM	9.000	1.180	10	9.750	1.250	11	8.250	2.130	10
	Wald*	7.625	1.068	10	5.250	1.110	4	10.00	0.577	11
ML_2	LR	5.000	0.598	5	6.250	0.479	5	3.750	0.629	4
	LM	9.125	0.350	11	9.00	0.707	9	9.250	0.250	11
	Wald*	4.75	1.424	3	1.000	000	1	8.500	0.289	9
$CPML_2$	LR	8.625	0.963	10	7.250	1.750	7	10.00	000	11
	LM	8.250	0.940	9	9.250	0.750	10	7.250	1.750	7
	Wald*	6.000	1.282	7	3.750	1.110	3	8.250	1.750	8
AML_1	LR	3.000	0.845	2	2.250	0.750	2	3.750	1.550	4
	LM	5.625	0.325	5	5.000	000	5	6.250	0.479	5
	Wald*	4.250	0.977	2	6.500	0.957	8	2.000	0.408	2
AML_2	LR	3.500	0.627	3	3.750	0.479	4	3.250	1.250	3
	LM	6.750	0.366	6	6.250	0.629	6	7.250	0.250	7
	Wald*	3.625	0.962	1	5.500	1.190	5	1.750	0.750	1

* Denotes Wald test for AR(1) and NW test for MA(1).

Table 4. Average ranks (Av.R), standard errors (SE) of average ranks and rank of average rank (R) of sizes of the LR, LM, Wald and NW tests based on different likelihood and message length functions for $n = 60$.

Function	Test	MA(1) and AR(1)			MA(1) only			AR(1) only		
		Av.R	SE	R	Av.R	SE	R	Av.R	SE	R
Profile	LR	8.250	1.264	9	10.75	0.250	11	5.500	1.550	6
	LM	6.875	1.008	6	8.000	1.290	9	5.750	1.490	5
	Wald*	7.750	1.373	11	11.00	000	11	4.500	1.320	4
Marginal	LR	4.875	1.187	5	7.500	0.866	9	2.000	0.577	1
	LM	1.375	0.263	1	1.250	0.250	1	1.500	0.500	1
	Wald*	4.750	1.082	4	7.500	0.500	9	2.000	0.408	2
CPL	LR	4.875	0.479	5	4.500	0.866	5	5.000	0.577	4
	LM	3.375	0.460	3	3.500	0.645	3	3.250	0.750	3
	Wald*	4.625	0.680	3	3.500	1.040	2	5.750	0.479	6
CPRL	LR	4.750	1.292	4	1.500	0.500	1	8.000	0.707	9
	LM	4.000	0.779	4	4.000	1.410	4	4.000	0.913	4
	Wald*	5.125	0.515	5	4.500	0.645	5	5.750	0.750	6
ACPL	LR	3.375	0.778	2	4.000	2.380	3	5.250	0.479	5
	LM	2.875	0.398	2	3.000	0.707	2	2.750	0.479	2
	Wald*	4.500	0.707	2	3.750	1.250	4	5.250	0.629	5
ML_1	LR	2.500	0.756	1	2.500	1.500	2	2.500	0.645	2
	LM	6.500	0.627	5	7.250	0.854	7	5.750	0.854	5
	Wald*	6.750	1.306	8	3.500	0.645	2	10.00	0.707	11
$CPML_1$	LR	8.750	1.146	10	7.000	2.040	8	10.25	0.479	11
	LM	7.250	1.411	7	7.000	2.040	6	7.750	2.350	9
	Wald*	7.000	0.926	9	6.250	1.250	7	7.750	1.430	8
ML_2	LR	4.000	0.732	3	5.000	1.080	6	2.750	0.629	3
	LM	7.875	0.515	10	8.250	0.854	10	7.500	0.645	8
	Wald*	5.250	1.623	6	1.000	000	1	9.500	0.500	9
$CPML_2$	LR	8.875	0.875	11	7.500	1.500	9	10.00	0.408	10
	LM	7.750	1.398	9	8.500	1.890	11	6.750	2.130	7
	Wald*	7.250	0.881	10	6.000	1.470	6	8.500	0.645	9
AML_1	LR	5.000	1.086	7	4.250	1.370	4	5.750	1.790	7
	LM	7.250	0.959	7	6.250	1.700	5	8.250	0.854	10
	Wald*	6.000	1.309	7	9.250	0.750	10	2.750	0.629	3
AML_2	LR	6.250	0.977	8	6.000	1.080	4	6.250	1.790	8
	LM	8.500	0.707	11	7.500	0.866	8	9.500	0.957	11
	Wald*	4.375	1.362	1	7.250	1.750	8	1.500	0.289	1

* Denotes Wald test for AR(1) and NW test for MA(1).

Table 5. Average ranks (Av.R), standard errors (SE) of average ranks and rank of average rank (R) of powers of the LR, LM, Wald and NW tests based on different likelihood and message length functions for $n = 30$.

			MA(1) and AR(1)			MA(1) only			AR(1) only		
Function	Test	Value of γ&ρ	Av.R	SE	R	Av.R	SE	R	Av.R	SE	R
Marginal	LR	+ve	4.563	0.376	5	3.625	0.324	3	5.500	0.500	6
		−ve	2.000	0.291	1	2.750	0.453	2	1.250	0.164	1
	LM	+ve	4.750	0.393	6	5.375	0.460	7	4.125	0.581	4
		−ve	3.000	0.274	3	3.125	0.350	2	2.875	0.441	2
	Wald*	+ve	4.625	0.272	4	4.125	0.125	4	5.125	0.479	6
		−ve	1.813	0.292	1	2.500	0.463	2	1.125	0.125	1
CPL	LR	+ve	3.000	0.329	3	2.375	0.498	1	3.625	0.324	4
		−ve	4.375	0.483	6	5.625	0.498	6	3.125	0.398	3
	LM	+ve	2.250	0.347	1	2.750	0.411	2	1.750	0.526	1
		−ve	5.562	0.364	7	6.250	0.366	7	4.875	0.549	6
	Wald*	+ve	3.938	0.551	3	2.250	0.313	2	5.625	0.625	7
		−ve	3.750	0.470	3	5.000	0.534	6	2.500	0.463	2
ACPL	LR	+ve	4.625	0.499	6	5.250	0.901	7	4.000	0.378	5
		−ve	2.688	0.388	2	1.750	0.366	1	3.625	0.498	5
	LM	+ve	3.000	0.474	2	3.875	0.549	4	2.125	0.666	2
		−ve	2.937	0.495	2	1.625	0.324	1	4.250	0.674	3
	Wald*	+ve	5.250	0.296	6	5.625	0.324	6	4.875	0.479	5
		−ve	3.375	0.437	2	4.000	0.707	4	2.750	0.453	3
ML_1	LR	+ve	3.500	0.408	4	3.500	0.463	2	3.500	0.707	3
		−ve	4.250	0.423	5	5.875	0.295	7	2.625	0.324	2
	LM	+ve	3.500	0.303	3	4.125	0.398	5	2.875	0.350	3
		−ve	4.688	0.254	6	5.125	0.350	6	4.250	0.313	3
	Wald*	+ve	1.125	0.085	1	1.250	0.164	1	1.000	000	1
		−ve	4.938	0.347	7	3.750	0.313	3	6.125	0.125	6
ML_2	LR	+ve	5.500	0.408	7	5.000	0.423	6	6.000	0.681	7
		−ve	3.875	0.423	3	4.250	0.313	4	3.500	0.732	4
	LM	+ve	4.750	0.452	6	5.000	0.423	6	4.500	0.824	5
		−ve	2.063	0.322	1	3.125	0.350	2	1.000	000	1
	Wald*	+ve	2.313	0.176	2	2.250	0.250	2	2.375	0.263	2
		−ve	4.375	0.632	6	2.000	0.270	1	6.750	0.164	7
AML_1	LR	+ve	2.563	0.491	1	3.750	0.773	5	1.375	0.183	1
		−ve	5.438	0.536	7	4.250	0.921	4	6.625	0.375	7
	LM	+ve	4.250	0.536	5	2.875	0.693	3	5.625	0.460	6
		−ve	4.625	0.482	5	4.500	0.779	5	4.750	0.620	5
	Wald*	+ve	4.625	0.437	4	5.375	0.460	5	3.875	0.666	3
		−ve	4.000	0.456	4	4.250	0.861	5	3.750	0.366	5
AML_2	LR	+ve	2.813	0.502	2	3.625	0.925	3	2.000	0.189	2
		−ve	3.938	0.558	4	2.750	0.750	2	5.125	0.639	6
	LM	+ve	4.125	0.657	4	2.500	0.886	1	5.750	0.559	7
		−ve	4.188	0.668	4	3.250	0.921	4	5.125	0.833	7
	Wald*	+ve	5.313	0.425	7	6.500	0.189	7	4.125	0.581	4
		−ve	4.313	0.568	5	5.500	0.886	7	3.125	0.441	4

* Denotes Wald test for AR(1) and NW test for MA(1).

Table 6. Average ranks (Av.R), standard errors (SE) of average ranks and rank of average rank (R) of powers of the LR, LM, Wald and NW tests based on different likelihood and message length functions for $n = 60$.

Function	Test	Value of γ&ρ	MA(1) and AR(1)			MA(1) only			AR(1) only		
			Av.R	SE	R	Av.R	SE	R	Av.R	SE	R
Marginal	LR	+ve	2.938	0.588	4	2.125	0.441	3	3.750	0.299	6
		−ve	1.625	0.239	1	2.000	0.423	1	1.250	0.164	1
	LM	+ve	3.312	0.489	4	4.375	0.497	7	2.250	0.675	3
		−ve	2.750	0.335	2	3.375	0.263	3	2.125	0.548	2
	Wald*	+ve	3.500	0.626	5	2.875	0.718	5	4.125	1.025	7
		−ve	1.500	0.274	1	1.875	0.515	2	1.125	0.125	1
CPL	LR	+ve	1.813	0.379	1	1.000	000	1	2.625	0.187	4
		−ve	2.250	0.310	3	2.625	0.498	3	1.875	0.350	3
	LM	+ve	2.000	0.329	1	2.500	0.378	1	1.500	0.500	1
		−ve	4.187	0.607	6	5.625	0.625	7	2.750	0.773	3
	Wald*	+ve	2.313	0.435	3	1.375	0.263	1	3.250	0.701	5
		−ve	2.563	0.491	4	3.125	0.789	5	2.000	0.566	2
ACPL	LR	+ve	1.938	0.359	2	1.375	0.183	2	2.500	0.187	3
		−ve	2.188	0.356	2	2.375	0.625	2	2.000	0.378	4
	LM	+ve	2.250	0.403	2	2.875	0.441	2	1.625	0.625	4
		−ve	3.187	0.449	3	3.500	0.500	5	2.875	0.766	4
	Wald*	+ve	2.813	0.510	4	2.000	0.423	3	3.625	0.865	6
		−ve	2.188	0.400	2	2.375	0.653	4	2.000	0.500	2
ML_1	LR	+ve	4.813	0.526	7	5.125	0.639	6	4.500	0.247	7
		−ve	3.938	0.602	7	5.125	0.718	7	2.750	0.796	5
	LM	+ve	3.125	0.364	3	3.500	0.500	5	2.750	0.526	4
		−ve	4.313	0.575	7	5.125	0.515	6	3.500	0.982	5
	Wald*	+ve	1.375	0.155	1	1.375	0.263	1	1.375	0.183	2
		−ve	3.063	0.655	5	2.125	0.549	3	4.000	1.134	7
ML_2	LR	+ve	4.313	0.637	6	5.625	0.754	7	3.000	0.235	5
		−ve	3.250	0.588	4	5.000	0.681	6	1.500	0.378	2
	LM	+ve	4.250	0.536	6	4.000	0.655	6	4.500	0.886	5
		−ve	2.375	0.340	1	3.375	0.263	3	1.375	0.375	1
	Wald*	+ve	1.563	0.302	2	2.000	0.570	3	1.125	0.125	1
		−ve	2.313	0.506	3	1.380	0.260	1	3.250	0.811	6
AML_1	LR	+ve	2.563	0.418	3	4.000	0.378	4	1.125	0.036	1
		−ve	3.688	0.681	6	3.625	0.925	5	3.750	1.065	7
	LM	+ve	4.188	0.549	5	3.125	0.854	3	5.250	0.491	6
		−ve	3.625	0.531	5	2.375	0.680	2	4.875	0.549	6
	Wald*	+ve	4.438	0.555	6	6.250	0.164	6	2.625	0.595	3
		−ve	3.938	0.588	6	5.500	0.681	6	2.375	0.565	5
AML_2	LR	+ve	2.938	0.433	4	4.500	0.267	5	1.375	0.052	2
		−ve	3.375	0.598	5	3.375	0.844	4	3.375	0.905	6
	LM	+ve	4.313	0.675	7	3.125	0.953	3	5.500	0.802	7
		−ve	3.250	0.668	4	1.625	0.625	1	4.875	0.875	6
	Wald*	+ve	4.813	0.593	7	6.750	0.164	7	2.875	0.639	4
		−ve	4.250	0.655	7	6.500	0.267	7	2.000	0.567	2

* Denotes Wald test for AR(1) and NW test for MA(1).