Statistics & Probability Letters 15 (1992) 281–284
North-Holland
13 November 1992

A model selection test for an AR(1) versus an MA(1) model

Philip Hans Franses

Econometric Institute, Erasmus University Rotterdam, Netherlands

Received January 1992
Revised March 1992

Abstract: This paper proposes a model selection test statistic for the choice between an AR(1) and an MA(1) model. It is a function of the first two sample autocorrelations of a time series. This establishes that it can be compared directly with a statistic given in Burke, Godfrey and Tremayne (1990). From Monte Carlo evidence it appears that the new test meets its purpose more.

Keywords: Time series; model selection.

1. Introduction

The selection between a first order autoregressive model, AR(1), and a first order moving average model, MA(1), can be important in forecasting, see Magnus and Pesaran (1989), and in modeling disturbances in a linear regression model, see King and McAleer (1987). The latter study develops a selection procedure for the case where a residual series shows only positive correlation. Recently, in Burke et al. (1990) a test procedure is proposed which does not need this assumption. From various simulations it emerges that its empirical behavior can however be unsatisfactory, see also Smith and Tremayne (1990).

In this paper an alternative test procedure is proposed for the null hypothesis of an AR(1) model, which is based on the R^2. The resulting model selection test statistic turns out to be a function of the first two sample autocorrelations of the time series only, and hence no model has to be estimated. This facilitates a direct comparison with the statistic in Burke et al. (1990). Section 2 formulates the theoretical results. Section 3 deals with a Monte Carlo investigation of the empirical performance of the test. Section 4 concludes.

Correspondence to: Dr. Philip Hans Franses, Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR, Rotterdam, Netherlands.

2. Testing AR(1) versus MA(1)

Consider the AR(1) and MA(1) models for a zero mean time series y_t,

$$y_t - \phi y_{t-1} = \varepsilon_t, \tag{1}$$

$$y_t = \varepsilon_t + \theta\varepsilon_{t-1}, \tag{2}$$

where ε_t is an uncorrelated zero mean process with variance σ_ε^2. In the sequel it is assumed that $0 < |\phi|,\ |\theta| < 1$, which establishes stationarity and invertibility and excludes the white noise model. This exclusion is reasonable since in case $y_t = \varepsilon_t$ both models can be valid. The autocorrelations of the AR(1) model are $\rho_j = \phi^j$ for $j = 0, 1, \ldots$, and the only nonzero autocorrelations for the MA(1) model are $\rho_0 = 1$ and $\rho_1 = \theta/(1 + \theta^2)$.

In a way, both models (1) and (2) contain one variable. Hence, the R^2, defined as $R^2 = 1 -$

$\sigma_\varepsilon^2/\sigma_y^2$, where σ_y^2 is the variance of y_t and σ_ε^2 is the error variance of (1) or (2), seems a logical model selection device. To investigate whether its use yields proper inference, one may consider the asymptotic distribution of the estimated R^2, $\hat{R}^2 = 1 - \hat{\sigma}_\varepsilon^2/\hat{\sigma}_y^2$. Hosking (1979) derives the asymptotic result

$$\hat{R}^2 \sim \mathrm{N}\left\{R^2, 4\left(\sigma_\varepsilon^2/\sigma_y^2\right) \cdot \sum_{\tau=1}^{\infty} \rho_\tau^2/n\right\}, \tag{3}$$

where n denotes the number of observations. For an AR(1) model it applies that $\sigma_y^2 = \sigma_\varepsilon^2/(1-\phi^2)$ and $R^2 = \phi^2$, and that

$$\sum_{\tau=1}^{\infty} \rho_\tau^2 = \rho_1^2 + \rho_2^2 + \rho_3^2 + \cdots$$
$$= \phi^2/(1-\phi^2). \tag{4}$$

Combining these results gives that for an AR(1) model

$$\hat{R}^2 \sim \mathrm{N}\{\phi^2, 4\phi^2/n\}, \tag{5}$$

or equivalently that

$$Q = n^{1/2}\left(\hat{R}^2 - \rho_2\right)/(2\rho_1) \sim \mathrm{N}(0, 1). \tag{6}$$

It is convenient to design a test procedure for the null hypothesis of an AR(1) model, since it can be observed from (6) that the distribution of the estimated R^2 includes only autocorrelations. This in contrast to the distribution of the $\hat{R}^2$ under the MA(1) hypothesis, where it includes ρ_1 as well as the parameter θ.

An empirical test statistic for an AR(1) versus an MA(1) model may now be given by the sample analogue of Q. Given that the $\hat{R}^2$ under the null hypothesis of an AR(1) model can be estimated by r_1^2, where r_i denotes the sample autocorrelation at lag i, the appropriate test statistic is

$$Q_{\mathrm{AR}} = n^{1/2}\left(r_1^2 - r_2\right)/(2r_1), \tag{7}$$

which is a two-sided test statistic. Large absolute values of Q_{AR} imply the rejection of the AR(1) null hypothesis.

The statistic Q_{AR} is a model selection test statistic, which implies that its size can not be controlled. This in contrast to the test statistic for the same null hypothesis given in Burke et al. (1990). That statistic is based on the second order partial autocorrelation coefficient ψ_{22} of an AR(p) process, or $\psi_{22} = (\rho_2 - \rho_1^2)/(1-\rho_2)$, see Box and Jenkins (1970, p. 65). For an AR(1) model it can be shown that approximately $\hat{\psi}_{22} \sim \mathrm{N}(0, n^{-1})$, and a test statistic is now given by

$$t_{\mathrm{AR}} = n^{1/2}\left(r_2 - r_1^2\right)/(1-r_2), \tag{8}$$

and it should asymptotically follow a standard normal distribution under the AR(1) null hypothesis. Note that this t_{AR} implies one-sided testing since under the MA(1) hypothesis the $\rho_2 - \rho_1^2$ does not exceed zero, while $1-\rho_2$ is always positive.

3. Empirical performance

Before empirically comparing these two test statistics, it is convenient to look more closely to the issue of interest. The R^2 of an AR(1) model is ϕ^2, and the R^2 of an MA(1) model is $R_{\mathrm{MA}}^2 = \theta^2/(1+\theta^2) = \theta\rho_1$. It is now interesting to carry out some misspecification analysis, i.e. to evaluate the effects of considering one model while the other is the data generating process (DGP). In case an AR(1) model is used while the MA(1) model is true, its R^2, $R_{\mathrm{AR}\,\|\,\mathrm{MA}}^2$, equals ρ_1^2. Since for the MA(1) model always holds that $|\theta| > |\rho_1|$, it is easy to recognize that $R_{\mathrm{MA}\,\|\,\mathrm{MA}}^2 > R_{\mathrm{AR}\,\|\,\mathrm{MA}}^2$, see also Nelson (1976). When the AR(1) is the DGP the $R_{\mathrm{AR}\,\|\,\mathrm{AR}}^2$ is equal to ρ_1^2. (In case an MA(1) is erroneously used then

$$R_{\mathrm{MA}\,\|\,\mathrm{AR}}^2 = \theta\rho_1 = \tfrac{1}{2} \pm \tfrac{1}{2}\left(1 - 4\rho_1^2\right)^{1/2}, \tag{9}$$

since the parameter θ can be derived from the ρ_1. If $|\rho_1| \leqslant 0.5$ it applies that $R_{\mathrm{MA}\,\|\,\mathrm{AR}}^2 \geqslant R_{\mathrm{AR}\,\|\,\mathrm{AR}}^2$, or that the MA(1) model has a better fit whether it is the DGP or not. Moreover, if $|\rho_1|$ exceeds 0.5, an MA(1) model can not be appropriate anyhow.

These results should have an impact on the design and performance of test statistics for an AR(1) hypothesis. An ideal situation would be that when the parameter $|\phi|$ of the AR(1) approaches unity, the rejection rate of a test goes to zero. From (8) it can be seen that this is not the case for the t_{AR} statistic, but (7) indicates that it may be the case for the Q_{AR} statistic. In fact, when $|\phi|$ is close to 1, the ρ_2 will also be close to 1, providing that the denominator $1-\rho_2$ becomes small, that the value of t_{AR} becomes large, and hence that the null hypothesis will be rejected too often in case it is obviously correct. Empirical evidence of this problem for the t_{AR} test can be found in Burke et al. (1990, p. 141) and in Smith and Tremayne (1990, Tables 1 through 3, columns 4). It is for this occurence that Burke et al. (1990) propose to first check whether ρ_1 exceeds 0.5, and then to apply the t_{AR} statistic. Furthermore, it can be deduced from the expression in (7) that the empirical rejection rate of the Q_{AR} test is likely to increase for an AR(1) model with a small parameter value. Finally, the denominator in (8) is equal to 1 under the alternative MA(1) model, while that of Q_{AR} is smaller than or equal to 1. This implies that, at

Table 1
Empirical rejection rate when H_0: $y_t = \phi y_{t-1} + \varepsilon_t$ is tested against H_1: $y_t = \varepsilon_t + \theta\varepsilon_{t-1}$, $n = 25$ and 100, 1000 replications

Parameter values (ϕ)	Nominal size	Rejection rate [a]			
		t_{AR}		Q_{AR}	
		$n=25$	$n=100$	$n=25$	$n=100$
0.9	0.05	0.004	0.027	0.000	0.000
	0.10	0.041	0.078	0.000	0.000
0.8	0.05	0.003	0.027	0.005	0.000
	0.10	0.043	0.077	0.007	0.000
0.7	0.05	0.005	0.030	0.030	0.000
	0.10	0.066	0.084	0.039	0.001
0.6	0.05	0.002	0.021	0.050	0.006
	0.10	0.039	0.077	0.069	0.013
0.5	0.05	0.005	0.026	0.084	0.032
	0.10	0.033	0.082	0.125	0.054
0.4	0.05	0.003	0.026	0.174	0.097
	0.10	0.042	0.075	0.222	0.160
0.3	0.05	0.007	0.022	0.328	0.225
	0.10	0.059	0.070	0.391	0.289
0.2	0.05	0.001	0.038	0.432	0.464
	0.10	0.035	0.085	0.484	0.529
0.1	0.05	0.002	0.025	0.500	0.668
	0.10	0.047	0.080	0.561	0.709

[a] Expressions for the test statistics t_{AR} and Q_{AR} are displayed in (8) and (7), respectively.

Table 2
Empirical rejection rate when H_0: $y_t = \phi y_{t-1} + \varepsilon_t$ is tested against H_1: $y_t = \varepsilon_t + \theta\varepsilon_{t-1}$, $n = 25$ and 100, 1000 replications

Parameter values (θ)	Nominal size	Rejection rate [a]			
		t_{AR}		Q_{AR}	
		$n=25$	$n=100$	$n=25$	$n=100$
0.9	0.05	0.128	0.947	0.294	0.729
	0.10	0.455	0.980	0.373	0.851
0.8	0.05	0.100	0.909	0.305	0.683
	0.10	0.429	0.976	0.391	0.801
0.7	0.05	0.101	0.817	0.295	0.635
	0.10	0.334	0.945	0.359	0.773
0.6	0.05	0.066	0.712	0.286	0.627
	0.10	0.279	0.871	0.354	0.745
0.5	0.05	0.022	0.471	0.265	0.522
	0.10	0.195	0.671	0.334	0.637
0.4	0.05	0.019	0.268	0.316	0.460
	0.10	0.147	0.469	0.385	0.558
0.3	0.05	0.014	0.118	0.396	0.435
	0.10	0.098	0.268	0.458	0.514
0.2	0.05	0.004	0.064	0.455	0.533
	0.10	0.059	0.174	0.484	0.605
0.1	0.05	0.003	0.033	0.517	0.672
	0.10	0.050	0.091	0.571	0.713

[a] Expressions for the test statistics t_{AR} and Q_{AR} are displayed in (8) and (7), respectively.

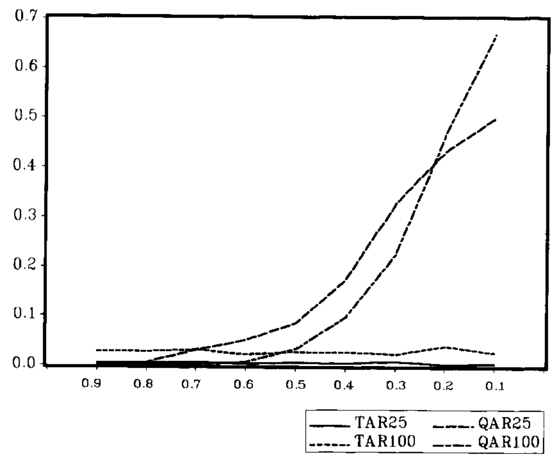

Fig. 1. Empirical rejection rate of the t_{AR} and Q_{AR} test statistics, for sample sizes of 25 and 100, when $y_t = \phi y_{t-1} + \varepsilon_t$, $\phi = 0.9, \ldots, 0.1$ is the data generating process and the nominal significance level is 0.05.

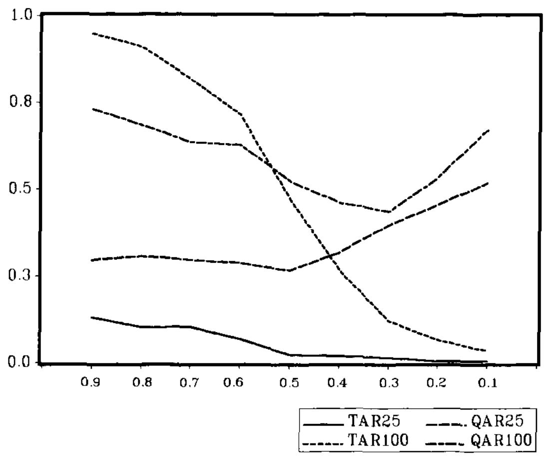

Fig. 2. Empirical rejection rate of the t_{AR} and Q_{AR} test statistics, for sample sizes of 25 and 100, when $y_t = \varepsilon_t + \theta\varepsilon_{t-1}$, $\theta = 0.9,\ldots,0.1$ is the data generating process and the nominal significance level is 0.05.

least theoretically, the empirical rejection rate of Q_{AR} under the MA(1) model is higher.

To investigate the performance of the Q_{AR} test relative to the t_{AR} test, several simulations have been carried out. For sample sizes of 25 and 100, 1000 replications of AR(1) and MA(1) processes have been generated. From the simulation results in Smith and Tremayne (1990) it appears that empirical size and power results are symmetrical across positive or negative values for the parameters, so only a selection of possible DGPs is considered. The results are displayed in Tables 1 and 2, and also in Figures 1 and 2 for a nominal significance level of 5%. Indeed, they show patterns which are close to what was expected. The Q_{AR} test, as opposed to the t_{AR} test, does not allow the selection of an MA(1) model when it cannot be the DGP, and its power remains reasonably constant while that of t_{AR} decays rapidly. For small parameter values of an AR(1) model, the Q_{AR} test suffers from an expected rejection rate deterioration, while that of t_{AR} behaves well.

4. Conclusion

The model selection test statistic for testing an AR(1) versus an MA(1) model developed in this paper seems to meet its purpose more than a rival selection device. This argument is emphasized with Monte Carlo evidence. Hence, the test statistic may be added to the usual diagnostic measures, such as tests for the presence of residual autocorrelation.

It may now also be worthwhile to consider testing an MA(1) versus an AR(1) model, see also Silvapulle and King (1991). A test statistic based on model selection should however recognize a result emerging from the misspecification analysis given in the present study, which is that the multiple correlation coefficient of an MA(1) model may in some cases be higher than that of an AR(1) model even when the AR(1) model is the data generating process.

Acknowledgements

I wish to thank Maxwell King, Teun Kloek and Jan Magnus for some helpful discussions and Joris Meijaard for computational assistance. Any errors are of course my own.

References

Box, G.E.P. and G.M. Jenkins (1970), *Time Series Analysis. Forecasting and Control* (Holden-Day, San Fransisco, CA).

Burke, S.P., L.G. Godfrey and A.R. Tremayne (1990), Testing AR(1) against MA(1) disturbances in the linear regression model: An alternative procedure, *Rev. Econom. Stud.* **57**, 135–145.

Hosking, J.R.M. (1979), The asymptotic distribution of R^2 for autoregressive-moving average time series models when parameters are estimated, *Biometrika* **66**, 156–157.

King, M.L. and M. McAleer (1987), Further results on testing AR(1) against MA(1) disturbances in the linear regression model, *Rev. Econom. Stud.* **54**, 649–663.

Magnus, J.R. and B. Pesaran (1989), Forecasting, misspecification and unit roots: the case of AR(1) versus ARMA(1,1), unpublished manuscript.

Nelson, C.R. (1976), The interpretation of R^2 in autoregressive-moving average time series models, *Amer. Statist.* **30**, 175–180.

Silvapulle, P. and M.L. King (1991), Testing moving average against autoregressive disturbances in the linear-regression model, *J. Bus. Econom. Statist.* **9**, 329–335.

Smith, M.D. and A.R. Tremayne (1990), Finite sample theory for testing AR(1) against MA(1) time series models, paper presented at *the 1991 Australasian Meeting of the Econometric Society, Sydney, Australia.*