«ΣΠΟΥΔΑΙ», Τόμος 55, Τεύχος 4ο, (2005) / «SPOUDAI», Vol. 55, No 4, (2005), University of Piraeus, pp. 38-46

DETERMINING THE ORDER OF DIFFERENCING AND THE TRUE GENERATING MODEL USING INFORMATION CRITERIA FOR ARIMA(0,1,1) AND AR(1) PROCESSES

By

*Christos Agiakloglou**
Department of Economics, University of Piraeus, Greece

Abstract

When a unit root test is applied to an ARIMA(0, 1, 1) process with large and moderate values of the moving average parameter serious size distortions will appear. Contrary, the test has almost no power if it is implemented to an AR(1) process with large and moderate values of the autoregressive parameter. This study investigates the possibility of determining the order of differencing and the true generating model for these two simple processes using the *AIC* and the *SBC* information criteria and it finds that the use of these criteria will assist the analyst to successfully determine the order of differencing and the true model. JEL Classification: C12; C22.

Keywords: Unit root test, stationary and non-stationary processes, information criteria

1. Introduction

The determination as to whether or not a series should be differenced is known as the unit root test and in the last few years a large volume of theoretical and empirical work has been appeared in the economic literature dealing with this issue. Actually, the presence of such a root does not only have consequences for the asymptotic distributions of the estimators and test statistics, but also it affects the possibility of modeling correctly economic relationships among time series (see, for example, Granger and Newbold (1974), Engle and Granger (1987), Park and Phillips (1988, 1989), Sims, Stock and Watson (1990) and Johansen (1991)).

This paper examines the performance of the *AIC* and *SBC* information

* I would like to thank Paul Newbold for his useful comments.

criteria in terms of selecting the order of differencing and the true generating model for series generated by the ARIMA(0, 1, 1) process, where the null hypothesis of a unit autoregressive root is true, and by the AR(1) process, where the null is false. These two processes have been repeatedly used in the literature to investigate the performance of unit root tests.

2. The choice of differencing

In practice, it is quite difficult to determine an appropriate value for the difference parameter. Part of this problem is based on the fact that the autoregressive polynomial may have roots that are close to one. In this case, the "augmented" Dickey-Fuller test, known as the *ADF* test, will fail to successfully determine the appropriate value of the difference parameter. Moreover, as reported, simulation evidence indicates that the *ADF* test has unsatisfactory performance for moderately large samples even for the simplest possible model with one moving average parameter, the ARIMA(0, 1, 1) process:

$$X_t - X_{t-1} = \varepsilon_t - \theta\varepsilon_{t-1} \tag{1}$$

for which the unit root hypothesis is true for all values of θ strictly less than one (see, for example, Schwert (1989), Agiakloglou and Newbold (1992), Hall (1994) and Ng and Perron (1995)). In fact, the performance of the *ADF* test is strongly affected not only by the order of the approximating autoregression, but also by the value of the moving average parameter. For small and moderate values of θ the test performs well, whereas for large values of θ the empirical significance levels are grossly inflated. The test improves its performance for all values of θ as long as a large order of the approximating autoregression is selected for the implementation of the *ADF* test, but the power of the test will be very low against simple alternatives. The trade-off between size distortions and power loss of the *ADF* test is discussed in Agiakloglou and Newbold (1996), whereas a recent work by Ng and Perron (2001) is made to alleviate the size distortions of unit root tests.

In fact, the issue of identifying the presence of a unit autoregressive root of a single time series may not always be the main concern in time series analysis. A good example is the case of short-term forecasting following the methodology of Box and Jenkins (1976) and this is true simply because competing models will generate similar, if not identical, forecasts. To illustrate, consider two simple competing models, i.e., the random walk process without drift:

$$(1-B)X_t = \varepsilon_t \tag{2}$$

and the first-order autoregressive process with mean:

$$(1-\varphi B)(X_t - \mu) = \varepsilon_t \tag{3}$$

If the value of the autoregressive parameter is close to one the stationary AR(1) process (3) will look very similar to a non-stationary random walk process (2) and similar forecast will be generated by both processes. However, this particular case, which tends to appear very often in practice, has totally different behavior towards testing. The null hypothesis of a unit autoregressive root will not be rejected very frequently when the value of φ is close to one, whereas at the same time the power of a test will of course be very low. The test will have high power only when φ is zero, i.e., the generating process is white noise.

Broadly speaking, any stationary model with $d = 0$ is arbitrary close to some other non-stationary model with $d = 1$. Thus, selecting a specific model from a general class of ARIMA models can be easily implemented given the available computing power and the fact that several different models can be easily estimated. The choice of the best-fitted model can be obtained through an order selection criterion such as the Akaike Information Criterion, known as *AIC*:

$$AIC = \ln\hat{\sigma}^2 + 2k/n \tag{4}$$

and the Schwarz Bayesian Criterion, known as *SBC*:

$$SBC = \ln\hat{\sigma}^2 + k\ln(n)/n \tag{5}$$

where k is the number of parameters, n is the number of observations of residuals and $\hat{\sigma}^2$ is the estimated error variance, uncorrected for degrees of freedom.

The two information criteria measure how well the model fits the series taking into account that a more elaborate model is expected to fit the series better. The model with the lowest *AIC* or *SBC* is the best. When the interest is in short-term forecasting, the principle of parsimony, i.e., the model used should require the smallest possible number of parameters that will adequate represent the data, is a useful element in model selection.

Therefore, given the fact that the unit root test has unsatisfactory performance even for simple models, it is interesting to investigate whether or not the two information criteria will substantially assist an analyst to determine correctly the value of the difference parameter and the true generating process. It should be emphasized, however, that the objective of this paper is not to examine the finite sample performance of these information criteria, but to investigate their efficiency in terms of selecting the true order in lieu of the existing unit root testing procedures.

3. Monte Carlo Study

Consider the ARIMA (0, 1, 1) model where the null hypothesis of a unit autoregressive root is true. Series of 101 observations are generated of model (1) for $X_0 = 0$ and for values of the moving average parameter equal to 0.9, 0.8, 0.6, 0.4 and 0.2. Next, ARIMA (*p*, 1, *q*) with no constant and ARIMA (*p*, 0, *q*) with constant models are fitted for all possible combinations of $p \leq 2$ and $q \leq 2$, where the first observation of the undifferenced series is deleted to ensure comparability.[1]

In both cases, estimation was through maximum likelihood using SPSS and the best-fitted model was selected according to the minimum value of the *AIC* and the *SBC* criteria. The results of this simulation process are reported in Table 1.

The decision to estimate only 16 different models is based primarily on the principle of parsimony, i.e., a small model is always preferable to a large one. This will also eliminate the possibility of having a large model been selected by *AIC*, since it is known that *AIC* has the tendency to select a more elaborate model. Moreover, dealing with only 16 models it will be relatively easy to deeply analyze the results and draw general conclusions.

TABLE 1

Number of order and model selection in 1,000 trials
for series of 100 observations generated by ARIMA(0, 1, 1) models

θ	Order Selection $d = 1$		True Model Selection				Mean Estimate of θ	Common Models
			Known order $d = 1$		Unknown order $d = 1$ or $d = 0$			
	AIC	*SBC*	*AIC*	*SBC*	*AIC*	*SBC*		
0.9	389	649	715	956	274	629	0.9141	500 (274)
0.8	639	873	716	937	474	832	0.8116	588 (474)
0.6	811	969	658	911	540	888	0.6030	619 (540)
0.4	859	965	535	735	470	716	0.4017	678 (470)
0.2	861	944	382	575	345	546	0.2075	861 (345)

Note: Numbers in parenthesis are the number of times that the true generating model is selected by both criteria.

The issue of determining the right order using information criteria does not seem to exist even for large values of the moving average parameter, although this cannot be seen directly from Table 1. For example, for $\theta = 0.9$, which is the most interesting case, since the unit root test has large size distortions, *SBC* will select a non-stationary model 649 times out of which 629 times will be the true model. It will also select 180 times the AR(1) model and 135 times the MA(1) model as the second and third best-fitted models respectively. Among those 180 AR(1) models, only 30 had statistically significant estimates of a small value of the autoregressive parameter and among 135 MA(1) models, only 20 had statistically significant estimates of the moving average estimates. Thus, *SBC* chooses almost with certainty not only the right order of differencing, but also the true generating process. Contrary, general remarks for $\theta = 0.9$ cannot be made for *AIC*, since the diversification of model selection is relatively high. The truth is that *AIC* selects a non-stationary model 389 times, almost 60% less than *SBC*, and 274 times the true model.

Similar conclusions for the selection of the order of differencing and the true model can be made for values of $\theta = 0.8$ and 0.6. Both information criteria improve their performance in terms of selecting directly more often the true model as the best-fitted model. For example, for $\theta = 0.6$ the true model is selected 888 and 540 times by *SBC* and *AIC* respectively. However, for small values of the moving average parameter, although both criteria select the right order very frequently, it is difficult to believe that they will also select that often the true model. For example, for $\theta = 0.2$ *SBC* selects 546 times the true ARIMA(0, 1, 1) model as the best-fitted model, 357 times the ARIMA(1, 1, 0) model as the second best-fitted model and 53 times the ARIMA(1, 0, 0) model as the third-best fitted model. Among 357 ARIMA(1, 1, 0) models, 201 had statistically significant estimates with a small negative value of the autoregressive parameter and all 53 ARIMA(1, 0, 0) models had statistically significant estimates of the autoregressive parameter with very large values close enough to one.

In general, for all values of the moving average parameter *SBC* selects more frequently than *AIC* a model with $d = 1$. This statement is also true in the case of determining the true generating model when the order of differencing is known (or even unknown), i.e., for $d = 1$ *SBC* selects the true model as high as 956 and 937 times for $\theta = 0.9$ and 0.8 respectively. However, this number decreases as the value of the moving average parameter decreases, indicating that for small values of θ it will be less likely to select the true model, even when the order of differencing is known.

Consider next the AR(1) process where the null hypothesis of a unit root is false for all values of the autoregressive parameter strictly less than one. Series of 101 observations are generated of model (3) for $X_0 = 0, \mu = 0$ and for values of the autoregressive parameter equal to 0.95, 0.9, 0.8, 0.5 and 0.2. Following the same methodology, Table 2 reports the simulation results for this series. Unlike the previous case, *AIC* selects more often than *SBC* a stationary process for all values of the autoregressive parameter. However, when the order is known, i.e., $d = 0$, *SBC* selects more often than *AIC* the true AR(1) model. It is interesting to note that for very large values of the autoregressive parameter, *SBC*, in contrast to *AIC*, selects the true model almost with certainty. i.e., 946 and 940 times for $\varphi = 0.95$ and 0.9 respectively.

TABLE 2

Number of order and model selection in 1,000 trials
for series of 100 observations generated by ARIMA(1, 0, 0) models

Φ	Order Selection $d = 0$		True Model Selection				Mean Estimate of φ	Common Models
			Known order $d = 0$		Unknown order $d = 0$ or $d = 1$			
	AIC	*SBC*	*AIC*	*SBC*	*AIC*	*SBC*		
0.95	363	161	654	946	235	154	0.9102	492 (105)
0.9	671	455	644	940	433	433	0.8625	481 (300)
0.8	912	693	642	928	602	647	0.7657	543 (463)
0.5	981	970	569	795	557	773	0.4739	717 (556)
0.2	804	748	314	472	259	354	0.1812	692 (251)

Note: Numbers in parenthesis are the number of times that the true generating model is selected by both criteria.

The performance of *SBC* remains satisfactory even in this case of selecting the right model when the order is not known, although this cannot be seen directly from Table 2. For example, for $\varphi = 0.95$, a case in which the unit root test will have no power, *SBC* selects a stationary model 161 times, at which 154 times is the true model. Even though this number is indeed very small, careful examination of the simulation results will lead us to the right model. First, it should be pointed out that only for this case the AR(1) model is not selected by *SBC* as the best-fitted model, but as the third one. Actually, *SBC* picks 458 times the ARIMA (0, 1, 1) model as the best-fitted model and 318 times the ARIMA (1, 1, 0) model as the second best-fitted model. Among those 458 ARIMA (0, 1, 1) models only 48 had statistically significant estimates of a moving average parameter with a small value and among those 318 ARIMA (1, 1, 0) models only 22 had statistically significant estimates of a very small value of the autoregressive parameter. Thus, even in this extreme case *SBC* will select the right model. However, as in the case of the ARIMA (0, 1, 1) process with large value of the moving average parameter, general remarks cannot be easily made using the *AIC* criterion. The truth is that *AIC* picks the AR(1) model as the best-fitted model 235 times, the ARIMA (0, 1, 1) model as the second best-fitted model 189 times and the ARIMA (1, 1, 0) model as the third best-fitted model 145 times.

For large and moderate values of the autoregressive parameter both information criteria and especially *SBC* will lead us to the right model fairly easier than the previous extreme case. For example, for $\varphi = 0.9$ *SBC* selects 433 times the AR(1) model as the best –fitted model, 311 times the ARIMA (0, 1, 1) model, where 55 of them had statistically significant estimates and 180 times the ARIMA (1, 1, 0) model, where only 17 of them had significant estimates. However, for small values of the autoregressive parameter this statement cannot be supported. For $\varphi = 0.2$ for example both criteria select the AR(1), the MA(1) and the ARIMA (0, 1, 1) model as first, second and third best-fitted model. Excluding the case of the ARIMA (0, 1, 1) model in which the estimates of the moving average parameter were very close to one indicating over differencing, several of the selected MA(1) models had statistically significant estimates of small, in absolute terms, values of the moving average parameter. Thus, among those 334 MA(1) models selected by *SBC* 187 had significant estimates, meaning that overall an AR(1) process with a small value of the autoregressive parameter will have approximately 80% chances to be selected.

4. Conclusion

This study illustrates that the use of the information criteria will assist the analyst to successfully determine the order of differencing and the true generating process of any given time series. Moreover, this methodology should be applied in practice using a larger combination of autoregressive and moving average parameters, say, for example, for all possible combinations of $p + q \leq 5$, so that to allow the possibility of having a large model been selected by *AIC*.

References

Agiakloglou, C. and Newbold, P. (1992) Empirical evidence on Dickey-Fuller –type tests, *Journal of Time Series Analysis*, **13**, 471-483.

Agiakloglou, C. and Newbold, P. (1996) The balance between size and power in Dickey-Fuller tests with data-dependent rules for the choice of truncation lag, *Economic Letters*, **52**, 229-234.

Akaike, H. (1974) A new look at the statistical model identification, *IEEE Transaction on Automatic Control*, **AC-19**, 716-723.

Box, G. E. P. and Jenkins, G. M. (1976) *Time Series Analysis, Forecasting and Control*, Holden Day, San Francisco.

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

Engle, R. F. and Granger, C. W. J., (1987) Co-Integration and error correction: representation, estimation and testing, *Econometrica*, **55**, 251-276.

Fuller, W. A. (1996) *Introduction to Statistical Time Series*, second edition, John Wiley and Sons, Inc., New York.

Granger, C. W. J. and Newbold, P. (1974) Spurious regressions in econometrics, *Journal of Econometrics*, **2**, 111-120.

Hall, A. (1994) Testing for a unit root in time series with pretest data based model selection, *Journal of Business and Economic Statistics*, **12**, 461-470.

Johansen, S. (1991) Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models, *Econometrica*, **59**, 1551-1580.

Ng, S. and Perron, P. (1995) Unit root tests in ARMA models with data-dependent method for the selection of the truncation lag, *Journal of the American Statistical Association*, **90**, 268-281.

Ng, S. and Perron, P. (2001) Lag length selection and the construction of unit root tests with good size and power, *Econometrica*, **69**, 1519-1554.

Park, J. Y. and Phillips, P. C. B., (1988) Statistical inference in regressions with integrated processes: part 1, *Econometric Theory*, **4**, 468-497.

Park, J. Y. and Phillips, P. C. B., (1989) Statistical inference in regressions with integrated processes: part 2, *Econometric Theory*, **5**, 95-131.

Said, S. E. and Dickey, D. A. (1984) Testing for unit roots in autoregressive-moving average models of unknown order, *Biometrika*, **71**, 599-607.

Schwarz, G. (1978) Estimating the dimension of a model, *Annals of Statistics*, **6**, 461-464.

Schwert, G. W. (1989) Tests for unit roots: A monte carlo investigation, *Journal of Business and Economic Statistics*, **7**, 147-160.

Sims, C. A., Stock, J. H. and Watson, M. W. (1990) Inference in linear time series models with some unit roots, *Econometrica*, **58**, 113-144.