Pertanika J. Sci. & Technol. **16 (2)**: 171 - 176 (2008)
ISSN: 0128-7680


# Performance of Autoregressive Order Selection Criteria: A Simulation Study

**Venus Khim-Sen Liew[1], Mahendran Shitan[2*], Choong Chee Keong[3] and Hooy Chee Wooi[3]**

[1]*Labuan School of International Business and Finance, Universiti Malaysia Sabah, Jalan Sungai Pagar, 87000 Labuan, Malaysia*

[2]*Department of Mathematics/ Mathematics Research Institute, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*

[3]*Faculty of Accountancy and Management, Universiti Tunku Abdul Rahman, Jalan Genting Kelang, 53000 Kuala Lumpur, Malaysia*

## ABSTRACT

Proper selection of autoregressive order plays a crucial role in econometrics modeling cycles and testing procedures. This paper compares the performance of various autoregressive order selection criteria in selecting the true order. This simulation study shows that Schwarz information criterion (SIC), final prediction error (FPE), Hannan-Qiunn criterion (HQC) and Bayesian information criterion (BIC) have considerable high performance in selecting the true autoregressive order, even if the sample size is small, whereas Akaike's information criterion (AIC) over-estimated the true order with a probability of more than two-thirds. Further, this simulation study also shows that the probability of these criteria (except AIC) in correctly estimating the true order approaches one as sample size grows. Generally, these findings show that the most commonly used AIC might yield misleading policy conclusions due to its unsatisfactory performance. We note here that out of a class of commonly used criteria, BIC performs the best for a small sample size of 25 observations.

**Keywords: Autoregressive, order selection criteria, simulation**

## INTRODUCTION

Most econometric models are formulated based on the Autoregressive (AR) process. In particular, AR process forms the main building block of the celebrated Integrated Autoregressive Moving Average (ARIMA) model, Vector Autoregressive (VAR) model, Vector Error Correction (VEC) model, Autoregressive Distributive Lag (ARDL) model, Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models, Threshold Autoregressive (TAR) model and the Smooth Transition Autoregressive (STAR) model, to name a few. A major problem with these econometric models is that the true order ($p$) of the AR ($p$) process is always unknown and the optimal order has to be determined by certain order selection criteria. Moreover, several econometric test procedures in empirical research such as the unit root tests, causality tests and cointegration tests are sensitive to the choice of autoregressive order. Thus, proper selection of autoregressive order is an important task in econometric modeling cycles and testing procedures.

The final prediction error (FPE) criterion, Schwarz information criterion (SIC), Hannan-Qiunn criterion (HQC) and Akaike's information criterion (AIC), among

Received : 14 July 2007
Accepted : 3 December 2007
* Corresponding Author

others, are commonly employed by empirical models researchers to determine the optimal autoregressive order. One interesting question that concerns many researchers is which criterion has the capability to provide the most reliable optimal order. However, there is no universally accepted solution to this question. In this study, firstly we attempt to compare via a simulation study, the performance of these criteria and second, to examine the effects of various sample sizes on the performance of these criteria in the selection of order.

We find that all selection criteria under study (except AIC) performed quite well in estimating the true order even if the sample size is small. We note that this probability is less than 0.25 for AIC even when the sample size grows to the very large value of 100,000.

## AUTOREGRESSIVE ORDER SELECTION CRITERIA

There has been considerable literature published on AR order selection criteria. A brief discussion of these criteria is available in Beveridge and Oickle (1994), whereas Brockwell and Davis (1996) presented greater theoretical and practical detail and additional references for many of these criteria.

Suppose the data set $\{X_1, \ldots, X_T\}$ are in fact observations of an AR process of order $p$, denoted by AR($p$) and defined as

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t \quad (1)$$

where $\phi_i$, $i = 1, \ldots, p$ are autoregressive parameters to be estimated and $\varepsilon_t$, $t = 1, \ldots, T$ is an independent and identically distributed (iid) disturbance or error term of a zero mean and a finite variance $\sigma^2$, which can be easily calculated by $\hat{\sigma}^2 = (T - p - 1)^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t^2$, where symbol ^ stands for estimated value.

The final prediction error (FPE) criterion, which was originally proposed by Akaike (1969) for AR($p$) order determination, is given by FPE = $\hat{\sigma}^2 (T - p)^{-1}(T + p)$. Having found that FPE is asymptotically inconsistent, Akaike (1973) came up with the so-called Akaike information criterion (AIC), defined as AIC=$\ln(\hat{\sigma}^2)+2p/T$. Disregarding the fact that AIC is only asymptotically unbiased (Jones, 1975) and has the tendency to pick models which are over-parameterized (Shibata, 1976), AIC is a very commonly used criterion for order selection (Basci and Zaman, 1998). Another statistics due to Akaike (1979) is the Bayesian information criterion, which is defined as BIC = $(T-p)\text{In}[(T-p)^{-1}T\hat{\sigma}^2]+T[1+\text{In}\sqrt{2\pi})]$ $+ p\,\text{In}[p^{-1}(\sum_{t=1}^{T}X_t^2 - T\hat{\sigma}^2)$. There is evidence to suggest that the BIC is better than the AIC as order selection criterion (Hannan, 1980).

Schwarz (1978) derived an information criterion, known as Schwarz information criterion (SIC) defined as SIC = $\ln(\hat{\sigma}^2)+ T^{-1} p \ln(T)$. SIC is also called Schwarz-Rissanen criterion as Rissanen (1978) also arrived at the same criterion independently, using different methodology. However, SIC has the advantage of being consistent over AIC (Basci and Zaman, 1998).

Hannan and Quinn (1979) and Hannan (1980) constructed the Hannan-Quinn criterion (HQC) from the law of the iterated logarithm. It provides a penalty function, which decreases as fast as possible for a strongly consistent estimator, as sample size increases. Hannan-Quinn criterion is given by HQC = $\ln \hat{\sigma}^2 + 2T^{-1}p \ln(\ln T)$. Hannan and Rissanen (1982) replaced the term $\ln(\ln T)$ by $\ln T$ to speed up the convergence of HQC. This revised version of HQ, however, was found to overestimate the model orders (Kavalieris, 1991).

## SIMULATION RESULTS

To achieve our objective, we simulated the AR process with normally distributed random errors of zero mean and a finite variance $\sigma^2$. We arbitrarily set the true lag length $p = 4$ and generate $\phi_i$, $i = 1, 2, 3, 4$ from a uniform distribution in the region (–0.25, 0.25). The choice of this region allows us to avoid undesired nonstationary processes. Meanwhile $c$ may take any value and is generated from an arbitrarily chosen uniform distribution in the region (–100,100). We simulate data sets for various sample sizes: 25, 50, 100, 250, 500, 1000, 10000, 100000. For each sample size, we simulated 1000 independent series for the purpose of order estimation. The estimated order $\hat{p}$ can be any integer from 1 to 20. The probability of estimating the true order, denoted by $P(\hat{p} = p)$ is then determined. In addition, this study also investigated the probabilities of under- and over-estimating the true order [denoted by $P(\hat{p} > p)$ and $P(\hat{p} < p)$ respectively], in which the estimated order based on the selection criteria is less than and more than the true order, respectively. The probabilities (in percentages) of various criteria incorrectly, under- and over- estimating the true order are presented in *Figs. 1* to *3*.

The most important finding is that all selection criteria under study (except AIC) perform quite well in estimating the true order even if the sample size is small[1]. It is noteworthy that the performance of all criteria in correctly estimating the true order increases as the sample size increases. In particular, $P(\hat{p} = p)$ almost converges to 1 with a sample size of 100,000 observations for all criteria except in the case of AIC, where $P(\hat{p} = p)$ is less than 0.25 even for this extraordinary large sample size (*Fig. 1*). Taken together, these findings warrant us that the most commonly used AIC might yield misleading policy conclusions due to its unsatisfactory performance. We note here that out of a class of commonly used criteria, BIC performs the best, followed by HQC, FPE and SIC, whereas AIC shows the worst performance.

[1]Note that AR-type models (i.e. dynamic models with lagged dependent variables) are quite problematic in small samples (of less than 50 observations for instance) since both Ordinary Least Squares and Maximum Likelihood have only large small and the usual inference machinery like *t*-statistics are also only asymptotically justified in this case. However, in circumstance (for example, many economic time series hardly have more than 50 annual observations) that leaves us no choice, knowing the small sample performance of these criteria in identifying the order $p$ become useful.

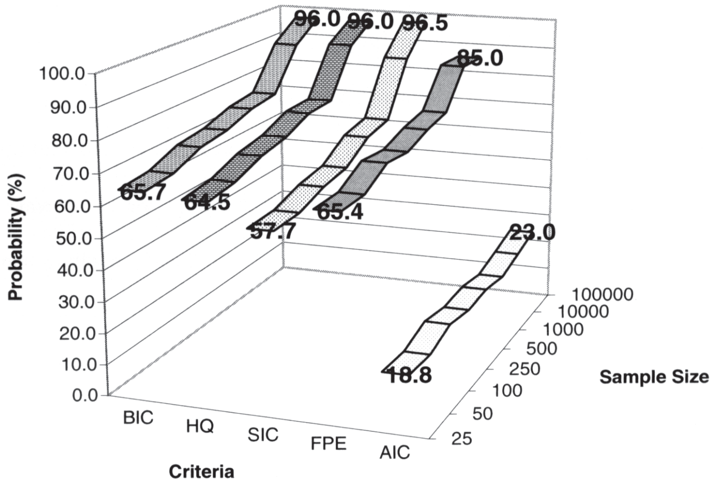


*Fig. 1: Probability of correctly estimating the true order*

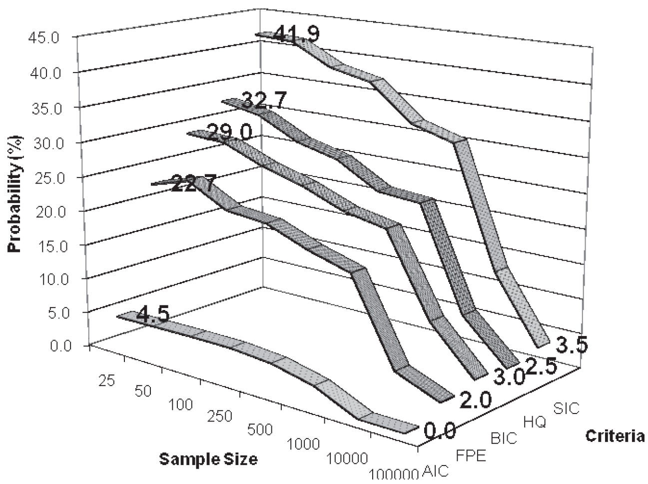


*Fig. 2: Probability of under-estimating the true order*

Another important finding of this study is that whenever FPE, BIC, HQC and SIC fail to estimate the true order, it is due to under-estimation rather than over-estimation. The reverse is true for AIC. For illustration, $P(\hat{p} < p)$ for these criteria are in the order, 22.7, 29.0, 32.7, 41.9 and 4.5 (*Fig. 2*), whereas $P(\hat{p} > p)$, in the same order is 11.9, 5.3, 2.0, 0.4 and 76.7 (*Fig. 3*) for a sample size of 25 observations. It is noted that in general (excluding the case of AIC), the probabilities of both under- and over-estimation reduces as sample size increases, thereby providing room for improvement in the performance of these criteria.

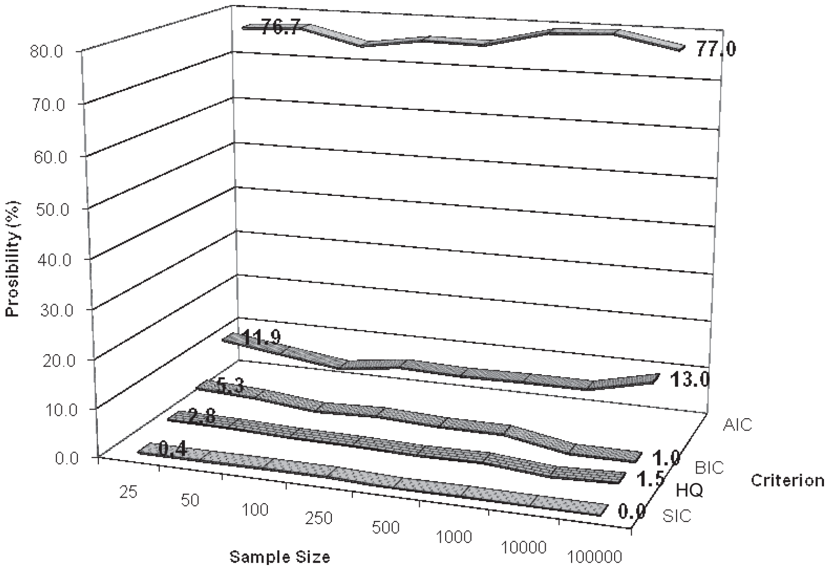


*Fig. 3: Probability of over-estimating the true order*

## CONCLUSIONS

Econometric modeling cycles and testing procedures mostly involves the predetermination of the autoregressive order ($p$). In reality, the true order ($p$) of the autoregressive (AR) process is unknown and the best a researcher can do is to estimate it using certain commonly accepted order selection criteria. The resulting estimate is known as the optimal order. Among others, the final prediction error (FPE) criterion, Schwarz information criterion (SIC), Hannan-Qiunn criterion (HQC), Akaike's information criterion (AIC) are commonly employed in empirical study. Nonetheless, there is no consensus on which of these criteria has the capability to provide the most reliable optimal order. As these criteria have important implications on the reliability of the estimated autoregressive order and thereby the subsequent findings based on this estimation, it is important to understand their performances. In this regard, this study attempts to compare via a simulation study, the performance of these criteria. Further, the effects of various sample sizes on the performance of these criteria in the selection of order were also analyzed.

This simulation study showed that SIC, FPE, HQC and BIC display considerably good performance in picking the true autoregressive order, even if the sample size is small, whereas AIC over-estimated the true order with a probability of more than two-third. Further, this simulation study also showed that the probability of these criteria (with the exception of AIC) correctly estimating the true order approaches one as sample size grows. In general, these findings show that the most commonly used AIC might yield misleading policy conclusions due to its unsatisfactory low performance. We note here that out of a class of commonly used criteria, BIC performs the best for a small sample size of 25 observations.

**ACKNOWLEDGEMENTS**

The authors would like to thank the anonymous referees who have provided useful comments and suggestions on the previous version of this paper.

**REFERENCES**

AKAIKE, H. (1969). Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics, 21,* 243–247.

AKAIKE, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Perron and F. Csaki (Eds.), *2nd International Symposium in Information Theory* (pp. 267–281). Budapest: Akademial Kiado.

AKAIKE, H. (1979). A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. *Biometrika, 66,* 237–242.

BASCI, S. and ZAMAN, A. (1998). Effects of skewness and kurtosis on model selection criteria. *Economic Letters, 59,* 17–22.

BEVERIDGE, S. and OICKLE, C. (1994). A comparison of Box-Jenkins and objective methods for determining the order of a non-seasonal ARMA model. *Journal of Forecasting, 13,* 419 – 434.

BROCKWELL, P. J. and DAVIS, R. A. (1996). *Introduction to Time Series and Forecasting*. New York: Springer.

HANNAN, E. J. (1980). The estimation of the order of an ARMA process. *Annals of Statistics, 8,* 1071 – 1081.

HANNAN, E. J. and QUINN, B. G. (1979). The determination of the order of an autoregression. *Journal of Royal Statistical Society, 41*, 190–195.

HANNAN, E. J. and RISSANEN, J. (1982). Recursive estimation of mixed autoregressive-moving average order. *Biometrika, 69,* 81–94.

HURVICH, C. M. and TSAI, C. L. (1989). Regression and time series model selection in small samples. *Biometrika, 76,* 297–307.

JONES, R. H. (1975). Fitting autoregressions. *Journal of American Statistics Association, 70,* 590– 592.

KAVALIERIS, L. (1991). A note on estimating autoregressive-moving average order. *Biometrika, 78,* 920–922.

SCHWARZ, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6,* 461–464.

SHIBATA, R. (1976). Selection of the order of an autoregressive model by Akaike's information criterion. *Biometrika, 63,* 117–126.

RISSANEN, J. (1978). Modelling by shortest data description. *Automatica, 14,* 467–471.