

# On Autoregressive Order Selection Criteria

Venus Khim-Sen Liew

Faculty of Economics and Management, Universiti Putra Malaysia, 43400 UPM,  
Serdang, Malaysia

This version: 1 March 2004.

---

## **Abstract**

This study investigates the performance of various commonly applied order selection criteria in selecting order of Autoregressive (AR) process. The most important finding of this study is that Akaike's information criterion, Schwarz information criterion, Hannan-Quinn criterion, final prediction error and Bayesian information criterion perform considerably well in estimating the true autoregressive order, even in small sample. Besides, there is no significant gain in differentiating these criteria unless one has a considerable large sample size. This study contributes to the empirical literature by providing helpfully guidelines regarding the use of order selection criteria in determining the autoregressive order.

*JEL classification:* C22; C51

*Keywords:* Autoregressive process, order selection criteria, simulation study.

---

# On Autoregressive Order Selection Criteria

## 1. Introduction

Many order selection criteria have been employed in economic study to determine the Autoregressive (AR) order of time series variables. Briefly, an AR process of order  $p$  refers to a time series in which its current value is dependent on its first  $p$  lagged values and is normally denoted by AR ( $p$ ). Note that the AR order  $p$  is always unknown and therefore has to be estimated via various order selection criteria such as the Aikake's information criterion (AIC) (Akaike, 1973), Schwarz information criterion (SIC) (Schwarz, 1978) Hannan-Quinn criterion (HQC) (Hannan and Quinn, 1979), final prediction error (FPE) (Akaike, 1969), and Bayesian information criterion (BIC) (Akaike, 1979); see Liew (2000) for an overview of these criteria. These criteria have been popularly adopted in economic studies, see for examples the works of Sarantis (1999, 2001) and Baum *et al.* (2001) who employed the AIC, Ahmed (2000) who used the AIC and BIC, Tan and Baharumshah (1999) who deployed the FPE, Yamada (2000) who used AIC and HQC and Xu (2003) who utilized the SIC in their empirical research. However, no special study has been allocated to contrast the performances of these order selection criteria, although few empirical studies (Taylor and Peel, 2000; Baum *et al.*, 2001; Guerra, 2001) do notify the inconsistency of these criterion and their tendency to under estimate the autoregressive order<sup>1</sup>. In this respect, this study has taken the initiative to conduct a simulation study on the empirical performances of these order selection criteria. In particular, this study aims at comparing the performances of various selected criteria, in their ability to identify the correct autoregressive order.

---

<sup>1</sup> A related work by Liew (2000) studies the performance of an individual criteria, namely the Aikake's biased corrected information criterion. The current study is more comprehensive than Liew (2000) in the sense that more criteria are involved for the purpose of comparative study.

To preview our results, the major finding in the current study is that these criteria perform fairly well even in a relatively small sample with only 25 observations, in the sense that they managed to identify the correct order about 60% of the time and this performance increases as sample size grows. At other times, these criteria tend to under estimate the true order than to over estimate it. Consequently, in the spirit of parsimony, which is against the selection of too long the autoregressive order, these criteria can thus be safely applied. With the discovery of these empirical findings, this study contributes to the literature in providing useful insights on the use of order selection criteria in future empirical researches.

The rest of this paper is organized as follows. Section 2 briefly describes the AR process, the order selection criteria and simulation procedure. Section 3 presents and discusses the results of this simulation study. Section 4 offers a summary of this study.

## **2. Methodology of Study**

### *2.1 Autoregressive process*

Mathematically, an  $AR(p)$  process of a series  $y_t$  may be represented by

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p} + \varepsilon_t \quad (1)$$

where  $a_0$  is the intercept term and  $a_1, a_2, \dots, a_p$  are autoregressive parameters and  $\varepsilon_t$  are normally distributed random error terms with a zero mean and a finite variance  $\sigma^2$ .

The estimation of AR ( $p$ ) process involves 2 stages: First, identify the AR order  $p$  based on certain rules such as order selection criteria. Second, estimate the numerical values for intercept and parameters using regression analysis. This study is confined to the study of the performances of various commonly used order selection criteria in identifying the true order  $p$ . In particular, this study generates AR processes with  $p$  arbitrary fixed at a value of 3 and uses these criteria to determine the order of each generated series as if the order is unknown. The autoregressive parameters are independently generated from uniform distribution with values ranging from -1 to 1 exclusively. Measures are taken to ensure that the sum of these simulated autoregressive parameters is less than unity in magnitude ( $|a_1 + a_2 + a_3| < 1$ ) so as to avoid non-stationary AR process. The error term is generated from standard normal distribution, whereas the intercept term is omitted without loss of generality. We simulate data sets for various usable sample sizes,  $S$ : 25, 50, 100, 200, 400, 800 and 1600. For each combination of processes and sample sizes, we simulated 1000 independent series for the purpose of order estimation. In every case, the initial value,  $y_0$  is arbitrary set to zero. In our effort to minimize the initial effect, we simulate  $100 + S$  observations and discard the first 100 observations, leaving the last  $S$  observations for order estimation. The estimated order  $\hat{p}$  is allowed to be determined from any integer ranging from 1 to 20 inclusively. In this respect, we compute the values for all 20 orders for each specific criterion and  $\hat{p}$  is taken from the one that minimizes that criterion. Note that each criterion independently selects one  $\hat{p}$  for the same simulated series.

## 2.2 Order selection criteria

The order selection criteria to be evaluated include:

(a) Akaike information criterion,

$$AIC_p = -2T [\ln(\hat{\sigma}_p^2)] + 2p ; \quad (2)$$

(b) Schwarz information criterion,

$$SIC_p = \ln(\hat{\sigma}_p^2) + [p \ln(T)]/T ; \quad (3)$$

(c) Hannan-Quinn criterion,

$$HQC_p = \ln(\hat{\sigma}_p^2) + 2 T^{-1} p \ln[\ln(T)] ; \quad (4)$$

(d) the final prediction error,

$$FPE_p = \hat{\sigma}_p^2 (T - p)^{-1} (T + p) \text{ and} \quad (5)$$

(e) Bayesian information criterion,

$$\text{BIC}_p = (T-p) \ln[(T-p)^{-1} T \hat{\sigma}_p^2] + T[1 + \ln(\sqrt{2\pi})] + p \ln[p^{-1} (\sum_{t=1}^T y_t^2 - T \hat{\sigma}_p^2)], \quad (6)$$

where  $\hat{\sigma}_p^2 = (T-p-1)^{-1} \sum_{t=p}^T \hat{\varepsilon}_t^2$ ,  $\varepsilon_t$  is the model's residuals and  $T$  is the sample size.

Note that the cap sign (^) indicates an estimated value. [Liew \(2000\)](#) provides an overview on these criteria, whereas details are given in, for instance, [Brockwell and Davis \(1996\)](#) and the references therein.

The main task of this study is to compute the probability of each of these criteria in correctly estimated the true autoregressive order. Note that this probability takes a value between zero and one inclusively, with a probability of zero, means that the criterion fails to pick up any true order and thereby is a poor criterion. On the other hand, a probability of one implies that the criterion manages to correctly select the true order in all cases and hence is an excellent criterion.

Besides, we also inspect the selected orders of the estimated order for 1000 simulated series of known order (that is,  $p = 3$ ), so as to gain deeper understanding on the performance of various criteria. We will refer to the situation whereby a criterion selected lower orders than the true ones as under estimate, whereas over estimate would mean the selection of higher orders than the true ones.

### 2.3 Simulation procedure

Briefly, the simulation procedure involves three sub-routines: with the first sub-routine generates a series of from the AR process, whereas the second sub-routine selects the autoregressive order of the simulated series and the third sub-routine evaluates the performance of the order selection criteria. The algorithm for the simulation procedure for each combination of sample size  $S$  and AR order  $p$  is outlined as follows:

1. Independently generate  $a_1$ ,  $a_2$  and  $a_3$  from a uniform distribution in the range  $(-1, 1)$ , conditioned on  $|a_1 + a_2 + a_3| < 1$ .
2. Generate a series of size  $100 + S$  from the AR process as represented in Equation (1) of order  $p = 3$  with  $a_0 = 0.0$  and  $a_1$ ,  $a_2$  and  $a_3$  obtained from Step 1. Initialize the starting value,  $y_0 = 0.0$ . Discard the first 100 observations to minimize the effect of initial value.
3. Use each selection criterion to determine the autoregressive order ( $\hat{p}$ ) for the last  $S$  observations of the series simulated in Step 2. Six selection criteria are involved.
4. Repeat Step 1 to Step 3 for  $B$  times, where  $B$  is fixed at 1000 in this study.
5. Compute the probabilities of (i) correct estimate, which is computed as  $\#(\hat{p} = P) / B$ ; (ii) under estimate, which is computed as  $\#(\hat{p} < P) / B$ ; and (iii) over estimate, which is computed as  $\#(\hat{p} > P) / B$ , where  $\#(\bullet)$  denotes numbers of time event  $(\bullet)$  happens.

### 3. Results and discussions

The probability of various criteria in correctly estimated the true order of the AR, process is tabulated in Table 1. Generally, Table 1 shows that AIC, SIC, FPE, HQC and BIC perform considerably well in estimating the true autoregressive order, in all cases. For example, for the case of sample size equals 25, the probability for each of the above criterion, in that order, is 0.591, 0.584, 0.591, 0.603 and 0.596. This means that out of 1000 simulated series of known order, AIC, SIC, FPE, HQC and BIC respectively have correctly identified the true order 591, 584, 591, 603 and 596 times. Thus, 60% of the times, these criteria have successfully selected the true order. Thus, we may conclude that these criteria perform fairly well in picking up the true order.

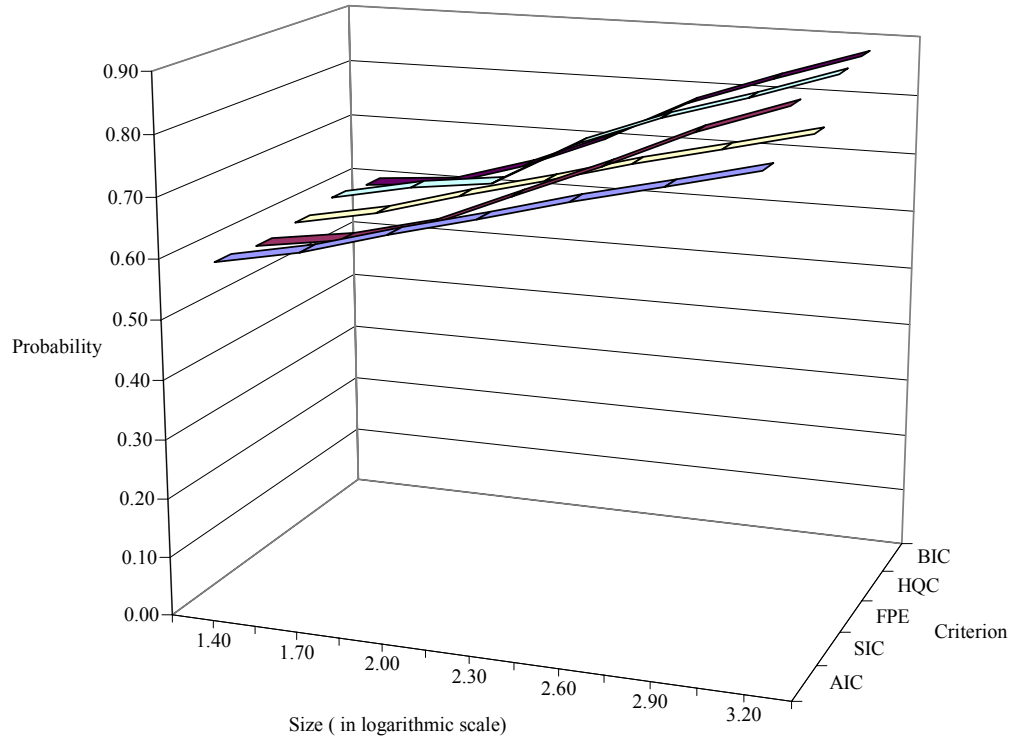
Table 1  
Probability of correctly estimated the true order of AR process.

Sample Size (Logarithmic Scale)	Order Selection Criteria				
	AIC	SIC	FPE	HQC	BIC
25 (1.40)	0.591	0.584	0.591	0.603	0.596
50 (1.70)	0.619	0.605	0.619	0.632	0.608
100 (2.00)	0.659	0.640	0.659	0.650	0.651
200 (2.30)	0.695	0.700	0.695	0.728	0.709
400 (2.60)	0.734	0.759	0.734	0.782	0.781
800 (2.90)	0.768	0.823	0.768	0.825	0.832
1600 (3.20)	0.803	0.873	0.803	0.873	0.879

Table 1 also shows that these criteria perform better and better as the sample size grows. With a sample size of 1600, the probability concerned for each of the same the five criteria has reached a value of 0.803, 0.873, 0.803, 0.873 and 0.879 respectively. This conclusion of improvement in performance for each of these five criteria as the sample size grows is clearly depicted in Figure 1.



Figure 1  
Performances of various criteria in correctly selected the true order.



The third finding revealed by Table 1 is that AIC and FPE (both constructed by Akaike) seems to have identical performance in terms of their ability to correctly locating the true order. In fact, a closer inspection on the selected order for each simulated series (results not shown) discovered that they consistently choose the same order at all times<sup>2</sup>. One would expect AIC to improve over FPE as it was proposed by Akaike to overcome the inconsistency of the latter (Akaike, 1973). However, such improvement is not observed in this study.

<sup>2</sup> Hence, these two criteria also have the same level of under estimation and over estimation as will be shown in Tables 2 and 3 later.

An interesting question in mind is whether we can identify the best criterion in selecting the AR order. However, it is difficult to just from Table 1 regarding this matter, as no criterion is found to consistently perform better than the rest in all cases. Nonetheless, both HQC and BIC do perform substantially better than others, in when the sample size is equal to or larger than 200. This suggests that differentiating between order selection criterion may be paid off only when one has large enough sample size (200 and above).

Further analysis of the distribution of the selected orders is conducted and the results are summarized in Tables 2 and 3. Table 2 reveals that for a relatively small sample size, AIC, SIC, FPE, HQC and BIC have under-estimated the true order with a probability of around one-third. Nonetheless, the probability of under estimation reduces as sample size grows, to the extent of negligible effect for a sample size as large as 1600. This finding is may be clearly seen from Figure 1.

Regarding over estimation, Table 3 shows that AIC, SIC, FPE, HQC and BIC is negligible in all cases regardless of small sample size. This empirical finding is in line with the built-in property of these criteria, which are designed in such a way that larger order is less preferable, in the spirit of parsimony (that is the simpler the better).

Table 2

Probability of under estimated the true order of AR process.

Sample Size (Logarithmic Scale)	Order Selection Criteria				
	AIC	SIC	FPE	HQC	BIC
25 (1.40)	0.312	0.401	0.312	0.355	0.364
50 (1.70)	0.284	0.381	0.284	0.333	0.368
100 (2.00)	0.256	0.355	0.256	0.316	0.328
200 (2.30)	0.208	0.296	0.208	0.244	0.275
400 (2.60)	0.159	0.232	0.159	0.189	0.204
800 (2.90)	0.118	0.170	0.118	0.146	0.158
1600 (3.20)	0.076	0.119	0.076	0.093	0.107

Figure 2

Performances of various criteria in under estimated the true order.

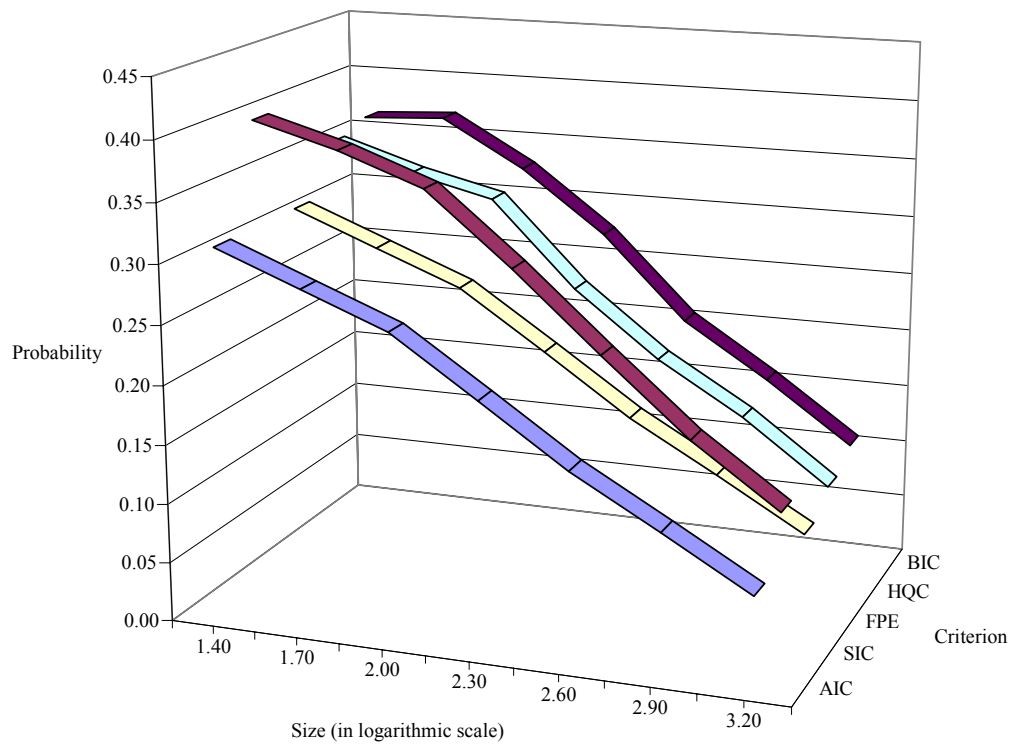


Table 3

Probability of over estimated the true order of AR process.

Sample Size (Logarithmic Scale)	Order Selection Criteria				
	AIC	SIC	FPE	HQC	BIC
25 (1.40)	0.097	0.015	0.097	0.042	0.040
50 (1.70)	0.097	0.014	0.097	0.035	0.024
100 (2.00)	0.085	0.005	0.085	0.034	0.021
200 (2.30)	0.097	0.004	0.097	0.028	0.016
400 (2.60)	0.107	0.009	0.107	0.029	0.015
800 (2.90)	0.114	0.007	0.114	0.029	0.010
1600 (3.20)	0.121	0.008	0.121	0.034	0.014

#### 4. Summary

The determination of autoregressive order for a time series is especially important in economics studies. Various order selection criteria such as the Aikake's information criterion (AIC), Schwarz information criterion (SIC), Hannan-Quinn criterion (HQC), final prediction error (FPE) and Bayesian information criterion (BIC) have been employed for this while by researchers in this respect. As the outcomes of these criteria may influence the ultimate findings of a study, a throughout understanding on the empirical performance of these criteria is warranted. This simulation study is specially conducted to shed light on this matter.

The current study independently simulate 1000 series from autoregressive process of known order ( $p = 3$ ) each of the various sample sizes ranging from 25 to 1600 observations in each series. Each order selection criterion is then allowed to independently estimate the autoregressive order for each simulated series, yielding some 1000 selected orders for each criterion. Based on these selected orders, we compute the

probabilities in which the true order is correctly identified, under estimate and over estimate. The results, which provide useful insights for empirical researchers are summarized as follows.

First, these criteria perform fairly well even in a relatively small sample ( $S$ ) with only 25 observations, in the sense that they managed to pick up the correct order about 60% of the time. Second, this performance increases as sample size grows. Third, with relatively small sample, there is no gain in differentiating between these criteria but HQC and BIC are found to outdo the rest for sample sizes reach 200 and beyond. Fourth, the probability of under estimation can be as high as one-third in the case of  $S=25$ , but this probability decreases as sample size grows. The problem of over estimation, however, is negligible in all cases.

To conclude, this study finds that depending on the sample size, at least 60% of the time, the order selection under study is able to correctly identify the true order of a given autoregressive process. Over estimation is negligible in all cases, whereas under estimation does occur but this problem tends to minimize as sample size grows. The findings in this simulation study may be taken as useful guidelines for future research in the use of order selection criteria in determining the autoregressive order.

## References

Ahmed, M. 2000. Money –income and money-price-causality in selected SAARC countries: some econometric exercises. *Indian Economic Journal*, 48, 55 – 62.

- 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. 2nd International Symposium on Information Theory, B. N. Petrov and F. Csaki (eds.), Akademiai Kiado, Budapest, 267 – 281.
- Akaike, H. 1979. A Bayesian extension of the minimum AIC procedure of Autoregressive model fitting. *Biometrika*, 66, 237 – 242.
- Brockwell, P. J., Davis, R. A., 1996. *Introduction to Time Series and Forecasting*. Springer, New York.
- Guerra, R. 2001. Nonlinear adjustment towards Purchasing Power Parity: the Swiss France - German Mark case. Working Paper, Department of Economics, University of Geneva.
- Hannan, E. J. and Quinn, B. G. 1978. The determination of the order of an autoregression. *Journal of Royal Statistical Society*, 41, 190 – 195.
- Liew, K. S. 2000. The performance of AICC as order determination criterion in the selection of ARMA time series models. Unpublished thesis, Department of Mathematics, Universiti Putra Malaysia.
- Sarantis, N. 1999. Modelling non-linearities in real effective exchange rates, *Journal of International Money and Finance*, 18, 27 – 45.
- Sarantis, N. 2001. Nonlinearities, cyclical behaviour and predictability in stock markets: international evidence. *International Journal of Forecasting*, 17, 439 – 482.
- Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics*, 6, 461 – 464.
- Tan, H. B. and Baharumshah, A. Z. 1999. Dynamic causal chain of money, output, interest rate and prices in Malaysia: evidence based on vector error-correction modeling analysis, *International Economic Journal*, 13, 103 – 120.
- Xu, Z. 2003. Purchasing power parity, price indices, and exchange rate forecasts. *Journal of International Money and Finance*, 22, 105 – 130.
- Yamada, H. 2000. M2 demand relation and effective exchange rate in Japan: a cointegration analysis. *Applied Economics Letters*, 7, 229 – 232.