Estimation of the Autoregressive Order in the Presence of Measurement Errors

Terence Tai-Leung Chong
The Chinese University of Hong Kong

Venus Liew
Universiti Putra Malaysi

Yuanxiu Zhang
University of British Columbia

Chi-Leung Wong
University of British Columbia

Abstract

Most of the existing autoregressive models presume that the observations are perfectly measured. In empirical studies, the variable of interest is unavoidably measured with various kinds of errors. Thus, misleading conclusions may be yielded due to the inconsistency of the parameter estimates caused by the measurement errors. Thus far, no theoretical result on the direction of bias of the lag order estimate is available in the literature. In this note, we will discuss the estimation an AR model in the presence of measurement errors. It is shown that the inclusion of measurement errors will drastically increase the complexity of the problem. We show that the lag lengths selected by the AIC and BIC are increasing with the sample size at a logarithmic rate.

We would like to thank Howell Tong, W.K. Li and N.H. Chan for helpful discussions and suggestions. All errors are ours.
Correspondence to: Terence T.L. Chong, Department of Economics, The Chinese University of Hong Kong, Shatin, Hong Kong. Email: chong2064@cuhk.edu.hk.
Citation: Chong, Terence Tai-Leung, Venus Liew, Yuanxiu Zhang, and Chi-Leung Wong, (2006) "Estimation of the Autoregressive Order in the Presence of Measurement Errors." *Economics Bulletin,* Vol. 3, No. 12 pp. 1-10
Submitted: February 8, 2006. **Accepted:** May 23, 2006.
URL: http://economicsbulletin.vanderbilt.edu/2006/volume3/EB-06C20003A.pdf

Estimation of the Autoregressive Order in the Presence of Measurement Errors

by

Terence Tai-Leung Chong,*
Department of Economics, The Chinese University of Hong Kong,
Hong Kong

Yuanxiu Zhang and Chi-Leung Wong,
Department of Economics, University of British Columbia,
Canada

Venus Khim-Sen Liew
Department of Economics, Universiti Putra Malaysia,
Malaysia

Abstract

Most of the existing autoregressive models presume that the observations are perfectly measured. In empirical studies, the variable of interest is unavoidably measured with various kinds of errors. Thus, misleading conclusions may be yielded due to the inconsistency of the parameter estimates caused by the measurement errors. Thus far, no theoretical result on the direction of bias of the lag order estimate is available in the literature. In this note, we will discuss the estimation an AR model in the presence of measurement errors. It is shown that the inclusion of measurement errors will drastically increase the complexity of the problem. We show that the lag lengths selected by the AIC and BIC are increasing with the sample size at a logarithmic rate.

Keywords: Autoregressive Process; Measurement Error; Akaike Information Criterion; Bayesian Information Criterion

*We would like to thank Howell Tong, W.K. Li and N.H. Chan for helpful discussions and suggestions. All errors are ours. Correspondence to: Terence T.L. Chong, Department of Economics, The Chinese University of Hong Kong, Shatin, Hong Kong. Email: chong2064@cuhk.edu.hk.

1. Introduction and the Model

Measurement errors are common in real-life data. For instance, variables that are related to expectations and unobservable characteristics like human capital, productivity and ability are often measured with errors. Many aggregate economic data also suffer from measurement errors. The errors can be caused by the aggregation procedures of the data collection agencies, or subtle differences in the definition of the economic variable across different countries. Applying standard estimation procedures to these variables with measurement errors will lead to a wrong conclusion, which has significant policy implications. This note considers time-series models contaminated by measurement errors. The existence of measurement errors not only affects the estimation of model parameters, but also the choice of the lag length. Measurement errors have two opposite effects on the lag order selection. On one hand, the model is misspecified and we tend to select a higher order. On the other hand, the selected order will tend to zero as measurement errors increase. Therefore, the direction of bias is unknown. We study how the model parameters and the variance of measurement error distort the selection of the lag length of an AR model. We will focus on the $AR(1)$ model for its tractability (Chong, 2001). Suppose our variable of interest, y_t^*, follows the process

$$(1-\beta L)y_t^* = \varepsilon_t \qquad (t = 1, 2, ..., T)\,, \tag{1}$$

where L is a lag operator such that $Ly_t^* = y_{t-1}^*$, $\varepsilon_t \sim$ i.i.d.$(0, \sigma_\varepsilon^2)$, $\sigma_\varepsilon^2 < \infty$. We assume $\beta \in (-1, 1)$ such that the process y_t^* is stationary. The true values of $\{y_t^*\}_{t=1}^T$ are not observable. Instead, we observe

$$y_t = y_t^* + u_t \qquad (t = 1, 2, ..., T)\,, \tag{2}$$

where $\{u_t\}_{t=1}^T$ is the measurement error process. For simplicity, we study the case where $u_t \sim$ i.i.d.$(0, \sigma_u^2)$, $\sigma_u^2 < \infty$, and u_t and ε_t are independent. It is readily verified that:

$$\begin{cases} \gamma_0 = \gamma_0^* + \sigma_u^2 = \dfrac{\sigma_\varepsilon^2}{1-\beta^2} + \sigma_u^2 \\ \gamma_1 = \beta\dfrac{\sigma_\varepsilon^2}{1-\beta^2} \\ \gamma_i = \beta\gamma_{i-1} (i > 1) \end{cases}, \tag{3}$$

where γ_j denotes $Cov(y_t, y_{t-j})$ and γ_j^* denotes $Cov(y_t^*, y_{t-j}^*)$. Let the true lag order and the estimated lag order be p_0 and $\hat{p}$ respectively. We examine the performance of the Akaike Information Criterion (Akaike, 1973) and Bayesian Information Criterion (Schwarz, 1978). We follow closely the notations of AIC and BIC in Hannan (1980). For an AR(p) model, the corresponding AIC and BIC are

$$AIC(p) = \ln \widehat{\sigma}_p^2 + 2p/T \tag{4}$$

and

$$BIC(p) = \ln \widehat{\sigma}_p^2 + p \ln T/T \tag{5}$$

respectively, where T is the sample size, σ_p^2 is defined as

$$\widehat{\sigma}_p^2 = \frac{RSS(p)}{T}, \tag{6}$$

where $RSS\,(p)$ is the residual sum of squares for an autoregression of order p. Note that the denominator in (6) should be $T - p - 1$ for the variance estimator to be unbiased, however, as we are interested in the asymptotic property of the lag-order estimator, we use T for simplicity. Define

$$\widehat{p}_{AIC} = \underset{p \in \{0,1,2,3,....\}}{Arg \min}\ AIC(p), \tag{7}$$

$$\widehat{p}_{BIC} = \underset{p \in \{0,1,2,3,....\}}{Arg \min}\ BIC(p), \tag{8}$$

$$B_{(p_1,p_2)} = \ln \widehat{\sigma}_{p_1}^2 - \ln \widehat{\sigma}_{p_2}^2, \tag{9}$$

$$C_{(p_1,p_2)(AIC)} = \frac{2(p_2 - p_1)}{T} \tag{10}$$

and

$$C_{(p_1,p_2)(BIC)} = (p_2 - p_1)\frac{\ln T}{T}. \tag{11}$$

In the selection of $AR(p_1)$ against $AR(p_2)$ via the BIC, we compare $B_{(p_1,p_2)}$ with $C_{(p_1,p_2)(BIC)}$. If $B_{(p_1,p_2)} > C_{(p_1,p_2)(BIC)}$, we select the $AR(p_2)$ model; otherwise, we select the $AR(p_1)$ model. Similar arguments apply to the AIC criterion. To study the behavior of $\hat{p}$ when $T \to \infty$, we first inspect the shape of the asymptotic $B_{(p,p+1)}$. From the Appendix, $B_{(p,p+1)}$ can be approximated by

$$\text{plim} B_{(p,p+1)} = \ln \frac{\gamma_0 - \Gamma_p' \boldsymbol{\Omega}_p^{-1} \Gamma_p}{\gamma_0 - \Gamma_{p+1}' \boldsymbol{\Omega}_{p+1}^{-1} \Gamma_{p+1}}, \tag{12}$$

where

$$\Gamma_p = \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_p \end{pmatrix}' \tag{13}$$

and

$$\mathbf{\Omega}_p = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & & \gamma_{p-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & & \gamma_{p-3} \\ \vdots & & & \ddots & \vdots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_0 \end{pmatrix}. \tag{14}$$

To examine the properties of $\hat{p}$, we introduce the following lemma:

Lemma 1 *If the observed process is* $AR(1)$ *with measurement errors, then for any* $p > 0$, $plim B_{(p,p+1)} > 0$.

The proof of Lemma 1 is given in the Appendix. Using Lemma 1, we have:

Proposition 1 *If the observed process is* $AR(1)$ *with measurement errors, then both* $\hat{p}_{AIC}$ *and* $\hat{p}_{BIC}$ *diverge to infinity as* $T \to \infty$.

The proof of Proposition 1 is provided in the Appendix. The significance of Proposition 1 merits emphasis. It implies that the selected order $\hat{p}$ is asymptotically unbounded. Thus, the BIC losses its appealing feature of consistency in the presence of measurement errors.

2. Approximating the Large Sample Effects

The measurement error variance σ_u^2, the variance of original error term σ_ε^2, as well as the autoregressive parameter β will all affect the estimation of the true order. Without loss of generality, we assume $\sigma_\varepsilon^2 = 1$, and study the effects of σ_u^2 and β. To begin with, we examine the effect of σ_u^2 on B. In large samples, $B_{(p,p+1)}$ can be approximated by $\text{plim}\, B_{(p,p+1)}$, which is a function of σ_u^2 and β. To illustrate how the order selection is affected by σ_u^2, we consider an example where $T = 2000$ and $\beta = 0.5$. Figure 1 plots $\text{plim}\, B_{(p,p+1)}$ for $p = 0$, 1 and 2. By visual inspection, $C_{(0,1)AIC} = C_{(1,2)AIC} = C_{(2,3)AIC} = 2/T$, $C_{(0,1)BIC} = C_{(1,2)BIC} = C_{(2,3)BIC} = \ln T/T$. As C_{AIC} is below C_{BIC}, for the same value of σ_u^2, we have $\hat{p}_{BIC} \leq \hat{p}_{AIC}$. For a simple illustration of Figure 1, suppose the variance of the measurement error is equal to 25, then the benefit (B) of adding one lag is always less then the cost (C) no matter which criterion is used. Thus, the estimated lag should be zero. We identify the following properties: (i) If $\sigma_u^2 > 0$ and $\beta \neq 0$, $\text{plim}\, B_{(p,p+1)}$ diminishes as p increases; (ii) When $\sigma_u^2 = 0$, $\text{plim}\, B_{(0,1)} > 0$ and $\text{plim}\, B_{(p,p+1)} = 0$ for $p \geq 1$; (iii) For all $\beta \in (-1, 1)$ and $p \geq 0$, $\text{plim}\, B_{(p,p+1)} \to 0$ as $\sigma_u^2 \to \infty$.

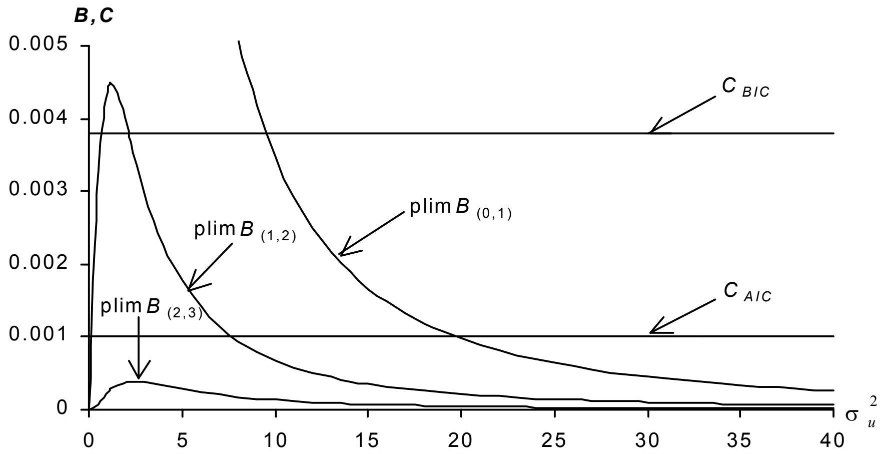

Figure 1: The Effect of σ_u^2 on $\hat{p}$ ($\beta = 0.5$, $T = 2000$)

The estimated lag orders are reported in Table 1. For example, the simulation results suggest that the lag length selected by AIC and BIC is respectively 2 and 1 for $\sigma_u^2 \in (0.17, 0.63)$, and for a variance larger than 20, the estimated lag orders from AIC and BIC are both zero.

Table 1: The effect of σ_u^2 on $\widehat{p}_{AIC}$ and $\widehat{p}_{BIC}$ ($T = 2000,\ \beta = 0.5$)

$\hat{p} \setminus \sigma_u^2$	$(0, 0.17)$	$(0.17, 0.63)$	$(0.63, 2.1)$	$(2.1, 7.7)$	$(7.7, 9.5)$	$(9.5, 20)$	$(20, \infty)$
$\widehat{p}_{AIC}$	1	2	2	2	1	1	0
$\widehat{p}_{BIC}$	1	1	2	1	1	0	0

To study the effect of effect of β on $\hat{p}$, we fix $\sigma_u^2 = 2$ and investigate the the plim$B_{(p,p+1)}$ for $p = 0$, 1 and 2 for $T = 2000$. The simulation results are plotted in Figure 2. It is suggestive from Figure 2 that for any given $\sigma_u^2 > 0$: (i) If $\beta \neq 0$, plim$B_{(p,p+1)}$ diminishes as p increases; (ii) For all $|\beta_1| < |\beta_2|$, plim$B_{(p,p+1)}(\sigma_u^2, \beta_1)$ <plim$B_{(p,p+1)}(\sigma_u^2, \beta_2)$; (iii) $\widehat{p}_{AIC}$ and $\widehat{p}_{BIC}$ are increasing step-functions of $|\beta|$; (iv) For any given $\beta \in (-1, 1)$, $\widehat{p}_{AIC} \geq \widehat{p}_{BIC}$. In short, $\widehat{p}_{AIC}$ and $\widehat{p}_{BIC}$ increase with the magnitude of the autoregressive parameter when the sample size is large.

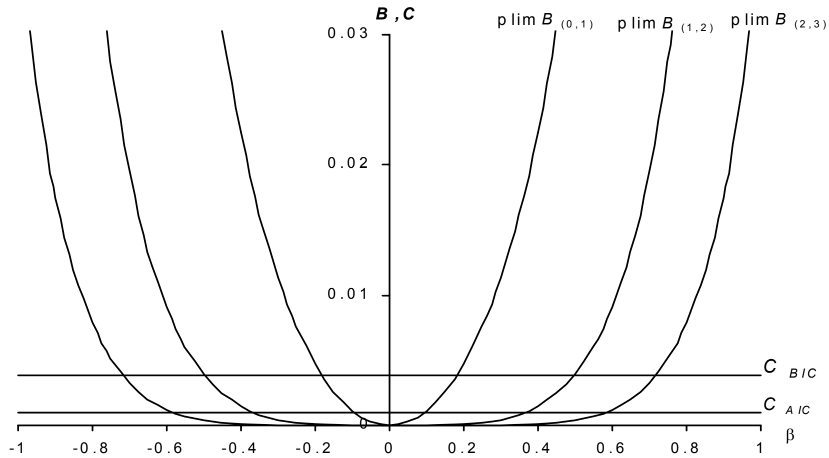

Figure 2: The Effect of β on $\hat{p}$ ($\sigma_u^2 = 2$, $T = 2000$)

3. Simulations

In this section, we conduct simulations to confirm the large sample results in the preceding section. We first simulate the mean of B, denoted by $\overline{B}$, for various values of σ_u^2 and β with sample size $T = 2000$ for 2000 replications. The difference between plim(B) and $\overline{B}$ is found to be small. We then simulate the average of $\hat{p}$, denoted by $\overline{\hat{p}}$, for various values of σ_u^2, β and T in Figure 3 to 5 respectively. In Figure 5, the sample sizes are from 2000 to 50000.

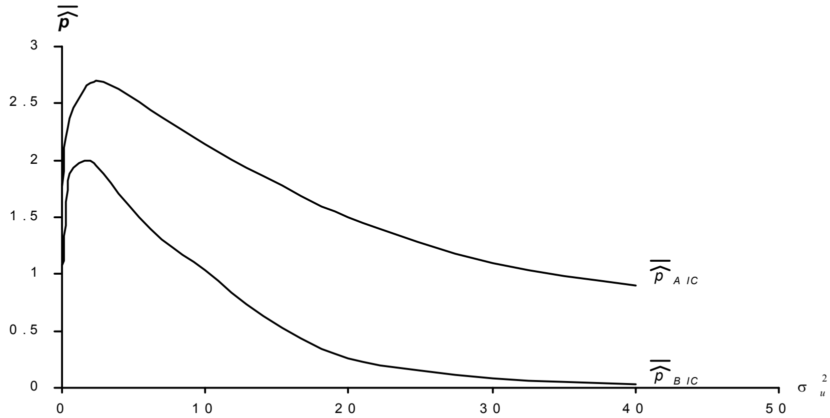

Figure 3: The Effect of σ_u^2 on $\overline{\hat{p}}$ ($\beta = 0.5$, $T = 2000$)

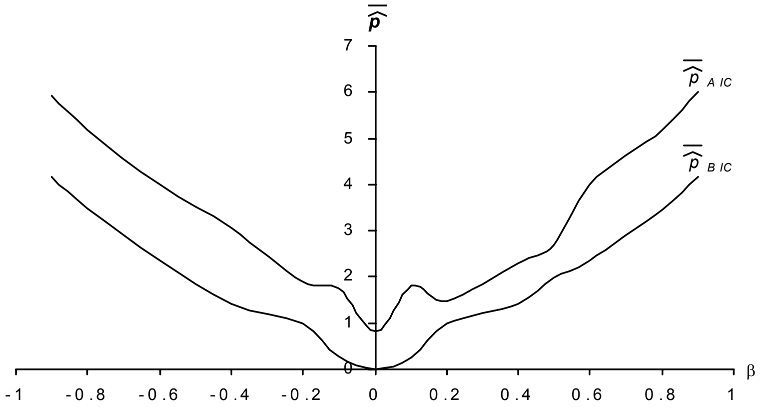

Figure 4: The Effect of β on $\overline{\hat{p}}$ ($\sigma_u^2 = 2$, $T = 2000$)

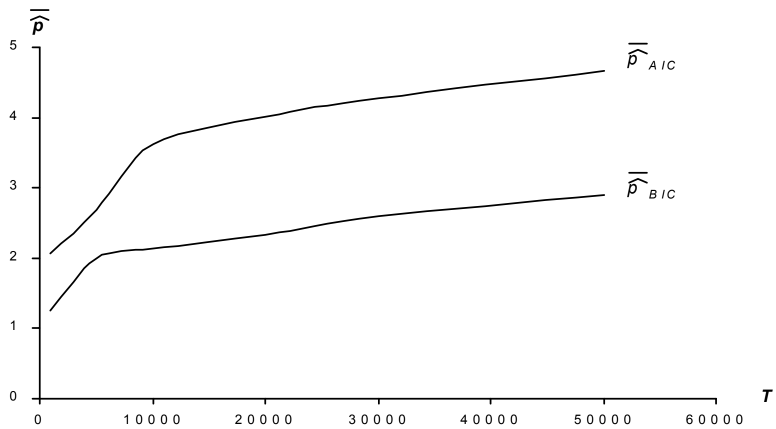

Figure 5: The Effect of T on $\overline{\hat{p}}$ ($\sigma_u^2 = 2$, $\beta = 0.5$, $T = 2000$ to 50000)

The results in Figures 3 to 5 can be summarized as follows: First, for any given $\beta \in (-1, 1)$, (i) As $\sigma_u^2 \to \infty$, we have $\widehat{p}_{AIC} \to 0$ and $\widehat{p}_{BIC} \to 0$; (ii) For any given $\sigma_u^2 > 0$, we have $\widehat{p}_{AIC} \geq \widehat{p}_{BIC}$. Second, for any given $\sigma_u^2 > 0$, the effect of β on $\hat{p}$ can be characterized as follows: (i) $\widehat{p}_{AIC}$ and $\widehat{p}_{BIC}$ are both weakly increasing with $|\beta|$; (ii) For any given $\beta \in (-1, 1)$, we have $\widehat{p}_{AIC} \geq \widehat{p}_{BIC}$. Lastly, if the observed process is $AR(1)$ with measurement errors, holding other factors constant, both $\widehat{p}_{AIC}$ and $\widehat{p}_{BIC}$ are increasing with the sample size at a logarithmic rate. Similar results are also obtained in a sample of size 100. Thus, in a finite sample, the pattern of $\hat{p}$ against σ_u^2 and β are reminiscent of their asymptotic counterparts.

4. Conclusion

Despite its prominent importance, the consequence of measurement errors on lag order selection is still a puzzle yet to be addressed. In this note, we make some steps towards the understanding of the effects of measurement errors on the order selection of autoregressive processes. In sharp contrast to the conventional finding on measurement-error models, which suggests that the parameter of interest has an attenuation bias towards zero, we show that the lag lengths selected by the AIC and BIC are increasing with the sample size at a logarithmic rate. It is concluded that the impact of the measurement error on the choice of lag length are similar regardless of the sample size. For any given sample size, the estimated lag length tends to be positively associated with the variance of measurement error if the variance is small, whereas they become negatively related when the variance is large. In addition, the selected order eventually approaches zero for any fixed sample size when the variance of the measurement error tends to infinity. Besides, in the presence of measurement errors, the magnitude of autoregressive parameter will also affect the choice of lag length. In particular, we tend to select higher order for larger magnitude of autoregressive parameter. For simplicity, we only examine the performance of the AIC and BIC. Other selection criteria, such as Akaike Information Corrected Criterion (AAIC) and Hannan-Quinn Criterion (HQC) would also be of interested.[1] The results in this note open the door to further investigations of the impact of measurement errors upon various generalizations of our model, e.g., the ARFIMA model of Chong (2000) and the structural-change model of Chong et al. (2005). Such extensions will be left for future research.

Appendix

Derivation of plim $B_{(p,p+1)}(\frac{\sigma_u^2}{\sigma_\varepsilon^2},\beta)$**:** For an autoregression of order p, let $\hat{\boldsymbol{\beta}}_p$ be the vector $\hat{\boldsymbol{\beta}}_p = \left(\hat{\beta}_{1,p}, \hat{\beta}_{2,p}, ..., \hat{\beta}_{p,p}\right)'$, where $\hat{\beta}_{i,p}$, $(i = 1, 2, .., p)$ denotes the OLS estimated coefficient of y_{t-i} in an $AR(p)$ regression without an intercept. It can be shown that for $p \geq 1$, $\text{plim}\widehat{\sigma}_p^2 = \gamma_0 - \Gamma_p' \text{plim}\hat{\boldsymbol{\beta}}_p$, where $\text{plim}\hat{\boldsymbol{\beta}}_p = \boldsymbol{\Omega}_p^{-1}\Gamma_p$. Thus, we have $\text{plim}\widehat{\sigma}_p^2 = \gamma_0 - \Gamma_p'\boldsymbol{\Omega}_p^{-1}\Gamma_p$ and $\text{plim}\, B_{(p,p+1)} = \ln \dfrac{\gamma_0 - \Gamma_p'\boldsymbol{\Omega}_p^{-1}\Gamma_p}{\gamma_0 - \Gamma_{p+1}'\boldsymbol{\Omega}_{p+1}^{-1}\Gamma_{p+1}}$.

Proof of Lemma 1: It is obvious from the OLS regression that $\text{plim}\,\hat{\sigma}_p^2 \geq \text{plim}\,\hat{\sigma}_{p+1}^2$ for all p, since $AR(p)$ model is a special case of AR$(p+1)$ model and $\hat{\sigma}^2$ is minimized over the estimated parameters. Thus, the remaining task is to show that $\text{plim}\,\hat{\sigma}_p^2 \neq \text{plim}\,\hat{\sigma}_{p+1}^2$ for all

[1] For more studies on the order selection problem, one is referred to Liew and Chong (2005), Rao and Wu (2001) and Pham (1988).

p. Since the first order conditions are linear in parameters, the solution must be unique. Thus, it suffices to show that plim $\hat{\beta}_{p+1,p+1} \neq 0$ for any $p > 0$, where $\hat{\beta}_{i,p+1}$ denotes the OLS estimated coefficient of y_{t-i} in an AR$(p+1)$ regression without an intercept. For any $p = 1$, it is readily verified that $p\lim \hat{\beta}_{1,1} = \dfrac{\gamma_1}{\gamma_0} = \dfrac{\gamma_1}{\gamma_0^* + \sigma_u^2} \neq 0$. For any $p > 1$, the first order condition of OLS estimation gives $\mathbf{\Omega}_{p+1}\text{plim}\hat{\boldsymbol{\beta}}_{p+1} = \Gamma_{p+1}$. We show plim $\hat{\beta}_{p+1,p+1} \neq 0$ by contradiction. Suppose plim $\hat{\beta}_{p+1,p+1} = 0$, then

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-2} & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-3} & \gamma_{p-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_0 & \gamma_1 \\ \gamma_{p-1} & \gamma_{p-2} & \cdots & \gamma_1 & \gamma_0 \\ \gamma_p & \gamma_{p-1} & \cdots & \gamma_2 & \gamma_1 \end{bmatrix} \begin{bmatrix} \text{plim}\hat{\beta}_{1,p+1} \\ \text{plim}\hat{\beta}_{2,p+1} \\ \vdots \\ \text{plim}\hat{\beta}_{p-1,p+1} \\ \text{plim}\hat{\beta}_{p,p+1} \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{p-1} \\ \gamma_p \\ \gamma_{p+1} \end{bmatrix}. \qquad (*)$$

Consider the p^{th} and $(p+1)^{th}$ rows in $(*)$:

$$\gamma_{p-1}\text{plim}\hat{\beta}_{1,p+1} + \gamma_{p-2}\text{plim}\hat{\beta}_{2,p+1} + ... + \gamma_1\text{plim}\hat{\beta}_{p-1,p+1} + \gamma_0\text{plim}\hat{\beta}_{p,p+1} = \gamma_p,$$

$$\gamma_p\text{plim}\hat{\beta}_{1,p+1} + \gamma_{p-1}\text{plim}\hat{\beta}_{2,p+1} + ... + \gamma_2\text{plim}\hat{\beta}_{p-1,p+1} + \gamma_1\text{plim}\hat{\beta}_{p,p+1} = \gamma_{p+1}.$$

Multiplying the first equation by β and subtracting it from the second, we obtain $(\gamma_1 - \beta\gamma_0)\text{plim}\,\hat{\beta}_{p,p+1} = 0 \Longrightarrow (\gamma_1 - \beta\gamma_0^* - \beta\sigma_u^2)\text{plim}\,\hat{\beta}_{p,p+1} = 0 \Longrightarrow -\beta\sigma_u^2\text{plim}\,\hat{\beta}_{p,p+1} = 0$. Since β and σ_u^2 are assumed to be non-zero, we get plim $\hat{\beta}_{p,p+1} = 0$. Plugging plim $\hat{\beta}_{p,p+1}$ into $(*)$, and repeating the same procedure on the $(p-2)^{th}$ and $(p-1)^{th}$ rows, we obtain plim $\hat{\beta}_{p-1,p+1} = 0$. By deduction, if plim $\hat{\beta}_{p+1,p+1} = 0$, we get plim $\hat{\beta}_{p,p+1}$ =plim $\hat{\beta}_{p-1,p+1} = ...$ =plim $\hat{\beta}_{1,p+1} = 0$, which contradicts $(*)$. Thus, plim $\hat{\sigma}_p^2$ >plim $\hat{\sigma}_{p+1}^2$ for all $p > 0$. Since the Logarithmic function is continuous and strictly monotonic, plim$B_{(p,p+1)}$ is unambiguously positive, and Lemma 1 is proved.

Proof of Proposition 1: We provide the proof for BIC. The proof of $\hat{p}_{AIC} \to \infty$ is essentially the same and is therefore skipped. For all $p > 0$, consider any two $AR(p)$ and $AR(p+1)$, we have $BIC(p) - BIC(p+1) \equiv B_{(p,p+1)} - C_{(p,p+1)(BIC)}$. By taking probability limits on both sides, we have

$$\begin{aligned} p\lim\left[BIC(p) - BIC(p+1)\right] &= p\lim\left[B_{(p,p+1)} - C_{(p,p+1)(BIC)}\right] = p\lim B_{(p,p+1)} \\ &= \ln\frac{\gamma_0 - \Gamma_p'\mathbf{\Omega}_p^{-1}\Gamma_p}{\gamma_0 - \Gamma_{p+1}'\mathbf{\Omega}_{p+1}^{-1}\Gamma_{p+1}} > 0. \qquad \text{(by Lemma 1)} \end{aligned}$$

Thus $lim_{T\to\infty}Pr\left(BIC(p) > BIC\left(p+1\right)\right) = 1$. By the definition of $\widehat{p}_{BIC}$, we select $p+1$ instead of p *if and only if* $BIC(p+1) < BIC(p)$. Thus, as the sample size goes to infinity, the probability of selecting $p+1$ instead of p equals one. Since p is arbitrary, Proposition 1 is proved.

References

1. Akaike, H. (1973) "Information Theory and an Extension of the Maximum Likelihood Principle" in *The Second International Symposium on Information Theory* by B. N. Petrov and F. Csake, Eds., Akademini Kiado, Hungary, 267-281.

2. Brockwell, P. J. and R.A. Davis (1991) *Time Series: Theory and Method*, Springer.

3. Chong, T. T. L. (2000) "Estimating the Differencing Parameter via the Partial Autocorrelation Function" *Journal of Econometrics* **97**, 365-381.

4. Chong, T. T. L. (2001) "Structural Change in AR(1) Models" *Econometric Theory* **17**, 87-155.

5. Chong, T. T. L., J. Bai, X. Wang and H. Chen (2005) "Generic Consistency of the Break-Point Estimators under Specification Errors in a Multiple-Break Model" *Working paper.*

6. Hannan, E. J. (1980) "The Estimation of the Order of an ARMA Process" *The Annals of Statistics* **8**, 1071-1081.

7. Hannan, E. J. and B. G. Quinn (1979) "The Determination of the Order of an Autoregression" *Journal of the Royal Statistical Society*, Series B, **41**, 190-195.

8. Liew, V. K. S. and T. T. L. Chong (2005) "Autoregressive Lag Length Selection Criteria in the Presence of ARCH Errors" *Economics Bulletin* **3**, no. 19, 1-5.

9. Pham, D. T. (1988) "Estimation of Autoregressive Parameters and Order Selection for ARMA Models" *Journal of Time Series Analysis* **9**, 265–279.

10. Rao, C. R. and Y. Wu (2001) "On Model Selection" in *IMS Lecture Notes Monograph. Series 38, Model Selection* by P. Lahiri, Eds., Institue of Mathematical Statistics, Beachwood, Ohio, 1–64.

11. Schwarz, G. (1978) "Estimating the Dimension of a Model" *The Annals of Statistics* **6**, 461-464.