Estimation of Autoregressive Model in the Presence of Measurement Errors

Article *in* Economics Bulletin · June 2006

Source: RePEc

3 authors:

Venus Khim-Sen Liew
University Malaysia Sarawak

Adam Chi Leung Wong
Lingnan University

Terence Tai Leung Chong
The Chinese University of Hong Kong

E C O N O M I C S   B U L L E T I N

Estimation of Autoregressive Model in the Presence of Measurement Errors

Terence Chong
The Chinese University of Hong Kong

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 comments.
Citation: Chong, Terence, (2006) "Estimation of Autoregressive Model in the Presence of Measurement Errors." *Economics Bulletin*, Vol. 3, No. 4 pp. 1−10
Submitted: June 1, 2005. **Accepted:** February 2, 2006.
URL: http://www.economicsbulletin.com/2006/volume3/EB−05C20019A.pdf

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

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