Cairo University

Faculty of Economics and Political Sciences

Department of Statistics

On Bayesian Identification for Moving Average Models

By

Amal Mohamed Abdel Fattah

Supervised by

Prof. Samir M. Shaarawy

Department of Statistics

Faculty of Economics and Political Sciences

Cairo University

Assoc. Prof. Emad E. A. Soliman

Department of Statistics

Faculty of Economics and Political Sciences

Cairo University

A Thesis Submitted to the Faculty of Economics and Political Science

Department of Statistics

In Partial Fulfilment of the Requirements for the M.Sc. Degree in Statistics

(2010)

Name Amal Mohamed Abdel Fattah **Nationality** Egyptian

Date and place of birth 21/11/1984, Libya.

Degree M.Sc. in Statistics

Specialization Statistics

Supervisors Prof. Samir M. Shaarawy Assoc. Prof. Emad E. A. Soliman

Department of Statistics

Faculty of Economics and Political Sciences

Title of Thesis On Bayesian Identification for Moving Average Models

Summary The study assesses the first stage in time series analysis which is the model identification from the Bayesian point of view. Two analytical Bayesian identification techniques are considered, the direct and the indirect techniques, using two approximations for the error term, the Newbold and the Broemeling-Shaarawy approximations. The two proposed techniques are developed using each approximation and evaluated for some moving average (MA) models. The behaviour of the Bayesian techniques is checked and compared via a comprehensive simulation study. The simulation study shows that the two techniques are efficient in identifying the moving average (MA) models. The direct technique dominates the indirect one. The Newbold approximation helps each technique to perform slightly better than the Broemeling and Shaarawy approximation.

Supervisors:

Prof. Samir M. Shaarawy Assoc. Prof. Emad E. A. Soliman

Department of Statistics

Faculty of Economics and Political Sciences

Cairo University

Abstract

The main objective of this study is to assess the first stage in time series analysis which is the model identification from the Bayesian point of view. Two analytical Bayesian identification techniques are considered, the direct and the indirect techniques, using two approximations for the error term, the Newbold and the Broemeling-Shaarawy approximations. The two proposed techniques are developed using each approximation and evaluated for some moving average (MA) models. The behaviour of the Bayesian techniques is checked and compared via a comprehensive simulation study. The simulation study shows that the two techniques are efficient in identifying the moving average (MA) models. The direct technique dominates the indirect one. The Newbold approximation helps each technique to perform slightly better than the Broemeling and Shaarawy approximation.

Keywords:

Moving average (MA) models- Prior distribution -Direct Bayesian identification -Indirect Bayesian identification- Newbold approximation- Broemeling and Shaarawy approximation- Posterior density function – Posterior mass function.

Supervisors:

Prof. Samir M. Shaarawy

Assoc. Prof. Emad E. A. Soliman

Department of Statistics

Faculty of Economics and Political Sciences

Cairo University

Acknowledgment

I would like to express my deepest gratitude for Prof. Samir Shaarawy for his gracious guidance, encouragement, and valuable advices. I would like also record my great gratitude to Dr. Emad Abdel-Salam for his great help, effort throughout the preparation of this research and to him I shall be always indebted. Their help and encouragement were supportive, and provided a constant source of inspiration.

I would also express my great appreciation for the generous assistance of Dr. Rasha El Souda, and Ms. Neveen El Zayat, whose helpful suggestions and discussions were very helpful.

I would like also to acknowledge the assistance of all the members of the Time Series Group in the Statistics Department, Faculty of Economics and Political science, Cairo University.

Finally, I would like to express my gratefulness to my lovely husband and for my parents for their constant support and prayers

Abstract

Identification plays an important role in the time series analysis since it's the first step in the time series analysis and the accuracy of all the preceeding steps depends on it. Identifying a moving average model means determining the order of the model. This step is a very difficult one since there is no optimal method that is widely accepted in the literature.

The moving average models (MA) are nonlinear in their coefficients. Thus, the errors sum of squares is not quadratic in the coefficients. A problem causing the likelihood function to be analytically intractable and lead to nonstandard posterior distributions. Several assertions tried to treat this problem numerically and analytically.

The current thesis aims to handle the model identification step for the moving average models from the Bayesian point of view. Two Bayesian analytical identification techniques are considered, namely the direct and the indirect techniques. Diaz and Farah (1981) introduced the direct technique for autoregressive (AR) models. In this technique the order of the model is considered a random variable with known maximum and its posterior mass function is derived. After that, the posterior probabilities are computed to choose the order with the maximum probability as a point estimate for the order of the model. The indirect technique, proposed by Broemeling and Shaarawy (1987), considers the orders of the autoregressive moving average (ARMA) models as unknown constants with known maximums. Therefore, this technique derives the posterior distribution for the coefficients instead of the orders using some analytical approximation for the error, which lead to standard posterior density. After that, the significance of each coefficient is tested. This technique determines the order of the model by keeping only the significant coefficients and removing the insignificant ones.

Among various well known approximations in the literature, this study highlights two approximations. These approximations were proposed by Newbold (1973), and Broemeling and Shaarawy (1988). Using such approximations, the problems concerning the posterior densities of time series models are considered. The first approximation by Newbold (N) expands the errors as linear functions in the coefficients around their nonlinear least squares estimators

NLSE's using Taylor's expansion, whereas, the second approximation by Broemeling and Shaarawy (B-S), approximates the errors as linear functions in the coefficients using their nonlinear least squares estimates NLSE's.

The identification of moving average models (MA) is studied via the direct and the indirect Bayesian identification techniques. The two proposed approximations are used to simplify the likelihood function of the model. The relationships between the approximate posterior densities based on the two approximations are investigated.

Comprehensive simulation studies are established to check the goodness of the two identification techniques based on the considered approximations. The simulation results show that the direct technique using Newbold approximation was the best to identify the order of the MA models. Moreover, the efficiency of the identification technique is affected by the approximation used in its derivations.

It is worth mentioning that this study is the first trial to study the effect of the approximation on the goodness of the identification technique for the MA models.

Table of Contents

page

Chapter 1
Introduction

Chapter 2
Review of the Literature

Chapter 3
Time series models and Bayesian concepts

Chapter 4
Bayesian Identification for MA models

Chapter 5
The Numerical Analysis and Simulation

Chapter 6
Discussions and Conclusion