

TESTING FOR A UNIT ROOT IN AN AR(p) SIGNAL
OBSERVED WITH MA(q) NOISE AND
MODEL MISSPECIFICATION

By

DONG-BIN JEONG

Bachelor of Science
Chung-Ang University
Seoul, South Korea
1989

Master of Science
Chung-Ang University
Seoul, South Korea
1991

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 1996

TESTING FOR A UNIT ROOT IN AN AR(p) SIGNAL
OBSERVED WITH MA(q) NOISE AND
MODEL MISSPECIFICATION

Thesis Approved:

Jakobel Awan

Thesis Advisor

Barry Kent Moser

W. W. Jandy

Kevin M. Currier

Thomas C. Collins

Dean of the Graduate College

PREFACE

A question that arises often in time series analysis is that whether the time series should be differenced. This is equivalent to asking if autoregressive (AR) part of the model being fitted has a unit root. Time series models with unit roots are known to provide good stochastic approximations for many nonstationary time series. The vector autoregressive moving average (ARMA) model has wide applications in various fields such as economics, engineering and survey sampling. Often the model parameters are restricted by a number of constraints. In this thesis a simple and easy-to-compute Newton-Raphson estimator (Shin and Sarkar, 1995) will be discussed that approximates the restricted maximum likelihood (RML) estimator and takes full advantage of the information contained in the restrictions. In order to obtain the initial parameter estimator in the Newton-Raphson iterative method, two stages of multivariate version of Hannan and Rissanen's (1982) procedure can be used to methods of both Kohn (1979) and Shin and Sarkar (1995). In the first part (Chapters 2, 3, 4) of this thesis we study the problem of testing for a unit root in an $AR(p)$ signal observed with $MA(q)$ noise by using three different estimation methods (Hannan-Rissanen, Kohn and Shin-Sarkar).

In the second part of this thesis (Chapters 5 and 6) we study various model misspecification problems. Model misspecification is a common problem in statistical data analysis. This is no exception in case of nonstationary time series data. A time series

model can be misspecified in various ways. One may regress a nonstationary time series on another unrelated nonstationary time series; or one may misspecify the order of a nonstationary autoregressive or a polynomial regression model; or one may misspecify a deterministic trend as a nonstationary AR process or vice versa. In the misspecified models, inference using the usual statistics such as t , Durbin-Watson (DW) and R^2 can be very often misleading. This topic of model misspecification is discussed in Chapter 5 and Chapter 6.

This thesis consists of six chapters. The motivation and purpose of this research will be introduced in Chapter 1. Previous works in testing for a unit root without parametric constraints in time series will be discussed in Chapter 2. The autoregressive moving average with nonlinear parametric restrictions will be introduced in Chapter 3 in case that parameters of both autoregressive part and moving average (MA) part are unknown. The general restricted model will be considered in Section 3.2 and three different estimation methods (Hannan-Rissanen, Kohn and Shin-Sarkar) will be introduced in Section 3.3. In Chapter 4 each method will be applied to test for a unit root in an AR(1) signal observed with MA(1) noise in a Monte Carlo experiment.

In Section 5.1 general regression model with integrated errors and one system of integrated regressors will be introduced, the limiting distributions of least squares (LS) estimators and the usual LS statistics such as $\hat{\sigma}^2$, t , DW and R^2 will be proposed in Section 5.2, and three different kinds of model misspecifications will be suggested. These are spurious regression problem, misspecification of nonstationary AR and polynomial regression models, and misspecification of orders in a nonstationary AR and polynomial regression models in Sections 5.3-5.5. In Chapter 6 we will analyze three different model

misspecification problems through a Monte Carlo study and examine each model misspecification problem by using the usual LS statistics as diagnostic tools.

I would like to express my sincere appreciation to my major adviser Professor Sahadeb Sarkar for his superb guidance, encouragement and support throughout this research. During my doctoral studies, Dr. Sarkar acted not only as my mentor but a loyal friend as well. I am certain that the accomplishment of this dissertation would not have been possible without his inspiration and assistance.

I also wish to express my sincere gratitude to my Ph. D. advisory committee members, Dr. William D. Warde, Dr. Barry K. Moser and Dr. Kevin M. Currier for their impact on my education and research. I would like to thank my parents Dr. Gab-Won Jeong and Won-Kyong Kim, and my brothers Dong-Hyuk, Byong-Gye Lee and Dong-Phil, and my sister Gee-Yeon for their inspiration, encouragement, understanding and support. I also thank my Seuk-Jong friends, Nobuyuki Nezu, Konrad Brandemuhl and Craig Wilson for their encouragement and heartfelt friendship. I wish to dedicate this dissertation to my family and hope my effort in graduate school has been worthy of my family's effort. Finally, I would like to express my appreciation to my wife Kyung-Hwa Lee (Oh Gam Do) for her endless love and concern.

TABLES OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
1.1. Unit Root Tests and Model Misspecification Problems	1
1.2. Research Purpose	4
1.3. Some Definitions	9
II. PREVIOUS WORKS IN TESTING FOR A UNIT ROOT WITHOUT PARAMETRIC CONSTRAINTS IN TIME SERIES	15
2.1. Previous Results on ARIMA(p,1,q) Models	15
2.1.1. Unit Roots in Univariate Time Series	15
2.1.2. Unit Roots in Multivariate Time Series	35
2.2. Previous Results on Measurement Error Models	46
III. TESTING FOR A UNIT ROOT IN AN AR(p) SIGNAL WITH MA(q) NOISE	51
3.1. Previous Works in Testing for a Unit Root with Parametric Constraints	51
3.2. General Restricted Model	59
3.3. Parameter Estimation	63
3.3.1. Hannan and Rissanen's Method	64
3.3.2. Kohn's Method	64
3.3.3. Shin and Sarkar's Method	70
3.4. The Limiting Distributions	71
IV. SIMULATION STUDY	74
4.1. Algorithms	75
4.1.1. Kohn's Method	75
4.1.2. Shin and Sarkar's Method	77

Chapter	Page
4.1.3. Hannan and Rissanen's Method	77
4.2. Results and Discussion	78
 V. MODEL MISSPECIFICATION PROBLEMS	 109
5.1. General Regression Model	109
5.2. Asymptotic Properties	112
5.3. Spurious Regression Problem	121
5.3.1. Previous Works	121
5.3.2. Spurious Regression	124
5.4. Misspecification of Nonstationary AR and Polynomial Regression Models	128
5.4.1. Previous Works	129
5.4.2. Misspecification of a Nonstationary AR Model as a Polynomial Regression Model	130
5.4.3. Misspecification of a Polynomial Regression Model as a Nonstationary AR Model	131
5.5. Underspecification of orders in Nonstationary AR and Polynomial Regression Models	135
5.5.1. Previous Works	135
5.5.2. Underspecification of the Order in a Nonstationary AR Model	135
5.5.3. Underspecification of the Order in a Polynomial Regression Model	138
 VI. AN EMPIRICAL STUDY	 140
6.1. Algorithms	140
6.1.1. Spurious Regression	141
A. Case 1: $k = 1, p = 2$	142
B. Case 2: $k = 2, p = 1$	142
C. Case 2: $k = 2, p = 2$	143
6.1.2. Misspecification of Nonstationary AR and Polynomial Regression Models	144
A. Misspecification of a Nonstationary AR(2) Model as a Polynomial Regression Model of Order 2	144
B. Misspecification of a Polynomial Regression Model of Order 2 as a Nonstationary AR(2) Model	145
6.1.3. Underspecification of the Orders in Nonstationary AR and Polynomial Regression Models	145
A. Underspecification of the Order in a Nonstationary AR Model	145

Chapter	Page
B. Underspecification of the Order in a Polynomial Regression Model	146
6.2. Results and Discussion	148
REFERENCES	156
APPENDICES	161
APPENDIX A - FORTRAN PROGRAM FOR HANNAN AND RISSANEN'S METHOD BASED ON KOHN'S APPROACH	162
APPENDIX B - FORTRAN PROGRAM FOR HANNAN AND RISSANEN'S METHOD BASED ON SHIN AND SARKAR'S APPROACH	172
APPENDIX C - FORTRAN PROGRAM FOR	176
KOHN'S METHOD	
APPENDIX D - FORTRAN PROGRAM FOR SHIN	191
AND SARKAR'S METHOD	
APPENDIX E - FORTRAN PROGRAM FOR EMPIRICAL	198
CUMULATIVE DISTRIBUTION OF UNIT ROOTS TEST STATISTICS BASED ON SHIN AND SARKAR'S METHOD	
APPENDIX F - FORTRAN PROGRAM FOR	214
SPURIOUS REGRESSION	
APPENDIX G - FORTRAN PROGRAM FOR	232
MISSPECIFICATION OF A NONSTATIONARY AR(2) AND A POLYNOMIAL REGRESSION MODEL OF ORDER 2	
APPENDIX H - FORTRAN PROGRAM FOR	241
UNDERSPECIFICATION OF THE ORDER IN A NON- STATIONARY AR(2) AND POLYNOMIAL REGRESSION MODEL OF ORDER 2	

LIST OF TABLES

Table		Page
4.1	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.01 (Hannan-Rissanen's method based on Kohn's approach)	81
4.2	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.05 (Hannan-Rissanen's method based on Kohn's approach)	82
4.3	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.10 (Hannan-Rissanen's method based on Kohn's approach)	83
4.4	Empirical power (%) of $\hat{\tau}$ for nominal level 0.01 (Hannan-Rissanen's method based on Kohn's approach)	84
4.5	Empirical power (%) of $\hat{\tau}$ for nominal level 0.05 (Hannan-Rissanen's method based on Kohn's approach)	85
4.6	Empirical power (%) of $\hat{\tau}$ for nominal level 0.10 (Hannan-Rissanen's method based on Kohn's approach)	86
4.7	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.01 (Hannan-Rissanen's method based on Shin and Sarkar's approach)	87
4.8	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.05 (Hannan-Rissanen's method based on Shin and Sarkar's approach)	88
4.9	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.10 (Hannan-Rissanen's method based on Shin and Sarkar's approach)	89
4.10	Empirical power (%) of $\hat{\tau}$ for nominal level 0.01 (Hannan-Rissanen's method based on Shin and Sarkar's approach)	90
4.11	Empirical power (%) of $\hat{\tau}$ for nominal level 0.05 (Hannan-Rissanen's method based on Shin and Sarkar's approach)	91

Table	Page
4.12	Empirical power (%) of $\hat{\tau}$ for nominal level 0.10 92 (Hannan-Rissanen's method based on Shin and Sarkar's approach)
4.13	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.01 93 (Kohn's method)
4.14	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.05 94 (Kohn's method)
4.15	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.10 95 (Kohn's method)
4.16	Empirical power (%) of $\hat{\tau}$ for nominal level 0.01 96 (Kohn's method)
4.17	Empirical power (%) of $\hat{\tau}$ for nominal level 0.05 97 (Kohn's method)
4.18	Empirical power (%) of $\hat{\tau}$ for nominal level 0.10 98 (Kohn's method)
4.19	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.01 99 (Shin and Sarkar's method)
4.20	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.05 100 (Shin and Sarkar's method)
4.21	Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.10 101 (Shin and Sarkar's method)
4.22	Empirical power (%) of $\hat{\tau}$ for nominal level 0.01 102 (Shin and Sarkar's method)
4.23	Empirical power (%) of $\hat{\tau}$ for nominal level 0.05 103 (Shin and Sarkar's method)
4.24	Empirical power (%) of $\hat{\tau}$ for nominal level 0.10 104 (Shin and Sarkar's method)
4.25	Empirical cumulative distribution of $n(\hat{A}_1 - 1)$ for $A_1 = 1$ ($N = 25$) 105
4.26	Empirical cumulative distribution of $n(\hat{A}_1 - 1)$ for $A_1 = 1$ ($N = 50$) 105

Table	Page
4.27	Empirical cumulative distribution of $n(\hat{A}_1 - 1)$ for $A_1 = 1$ ($N = 100$) ... 106
4.28	Empirical cumulative distribution of $n(\hat{A}_1 - 1)$ for $A_1 = 1$ ($N = 250$) ... 106
4.29	Empirical cumulative distribution of $\hat{\tau}$ for $A_1 = 1$ ($N = 25$) 107
4.30	Empirical cumulative distribution of $\hat{\tau}$ for $A_1 = 1$ ($N = 50$) 107
4.31	Empirical cumulative distribution of $\hat{\tau}$ for $A_1 = 1$ ($N = 100$) 108
4.32	Empirical cumulative distribution of $\hat{\tau}$ for $A_1 = 1$ ($N = 250$) 108
6.1	Spurious regression ($k = 1, p = 2$) 151
6.2	Spurious regression ($k = 2, p = 1$) 152
6.3	Spurious regression ($k = 2, p = 2$) 153
6.4	Misspecification of a nonstationary AR(2) as 154 a polynomial regression model of order 2
6.5	Misspecification of a polynomial regression model of order 2 as 154 a nonstationary AR(2) model
6.6	Underspecification of 155 a nonstationary AR(2) model
6.7	Underspecification of a polynomial regression model 155 of order 2

NOMENCLATURE

Guide: Subscripts -- i, j, k, t

Scalars -- $a, A_1, C_1, z, \alpha, \beta, \beta_1, \delta_0, \mu, \theta_0, \theta_1, \phi_1, \sigma, \theta, \rho, \tau$

Vectors -- $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$

Estimators -- $\hat{A}_1, \hat{\delta}, \hat{\lambda}, \hat{\mu}, \hat{\theta}, \tilde{\rho}, \hat{\rho}, \tilde{\xi}$

Matrices -- $\mathbf{A}, \mathbf{B}, \mathbf{\Omega}$

$\mathbf{0}$ a vector of matrix of zeros

ACF Autocorrelation Function

ADF Augmented Dickey-Fuller

AR(p) Autoregressive of order p

ARIMA(p, d, q) Autoregressive Integrated Moving Average of order p, d and q

ARMA(p, q) Autoregressive Moving Average of order p and q

ARMAX Autoregressive Moving Average with exogenous variables

CLT Central Limit Theorem

CPS Current Population Survey

DSP Difference Stationary Process

DW Durbin-Watson

GLS Generalized Least Squares

I(d) Integrated of order d

IID Independent Identically Distributed

IV	<u>I</u> nstrumental <u>V</u> ariables
MA(q)	<u>M</u> oving <u>A</u> verage of order q
MLE	<u>M</u> aximum <u>L</u> ikelihood <u>E</u> stimator
MRMLE	<u>M</u> odified <u>R</u> estricted <u>M</u> aximum <u>L</u> ikelihood <u>E</u> stimator
NID	<u>N</u> ormally <u>I</u> dentically <u>D</u> istributed
OLS	<u>O</u> rdinary <u>L</u> east <u>S</u> quares
PSID	<u>P</u> anel <u>S</u> tudy of <u>I</u> ncome <u>D</u> ynamics
RMLE	<u>R</u> estricted <u>M</u> aximum <u>L</u> ikelihood <u>E</u> stimator
R ²	Coefficient of Determination
SLLN	<u>S</u> trong <u>L</u> aw of <u>L</u> arge <u>N</u> umbers
SNR	<u>S</u> ignal-to- <u>N</u> oise <u>R</u> atio
VARMA	<u>V</u> ector <u>A</u> utoregressive <u>M</u> oving <u>A</u> verage

CHAPTER I

INTRODUCTION

1.1. Unit roots tests and model specification problems

Suppose a process has a mean that is different for each time period. How could we estimate these means ? As usual, we use sample information. However, typically we have only one observation per time period for time series data. Therefore, we have only one observation at time t to estimate the mean at time t , one observation at time $t+1$ to estimate the mean at time $t+1$, and so on. An estimate of a mean based on only one observation is automatically not useful. The situation gets much worse if the variance of the process also is not constant through time. In such cases an appropriate degree of differencing is to be considered to make the data stationary in mean. The reason for making data stationary is to obtain useful estimates of the parameters in the time series model.

There are several ways to check the stationarity of a time series, i.e., to make a decision on the proper degree of differencing:

(1) Examine the realization visually (graphically) to see if either the mean or the variance appears to be changing over time. If different segments of each series behave much like the rest of the series after we allow for changes in level and/or slope, then this nonstationary realization can be transformed into a stationary series by differencing. In addition visual inspection of the data is perhaps the most practical way of measuring stationarity of the variance.

(2) Inspect the estimated autocorrelation function (ACF) to see if the autocorrelation move rapidly toward zero. In practice, “rapidly” means that the absolute t statistic values of the estimated autocorrelations should fall below approximately 1.6 by about lag 5 or 6 (Pankratz, 1983). These values are only guidelines, not the absolute rules. In general, if any examination of the ACF for a time series shows relatively large autocorrelations at high lags, differencing of the observations is probably appropriate.

(3) Scrutinize the estimated autoregressive (AR) coefficients to see if they satisfy the stationarity conditions.

However, the above methods have some drawbacks in the sense that they fail to quantify the extent of mean nonstationarity and they depend on subjective judgment. Moreover, the problem with determining the degree of differencing from plots is that in small samples the picture is often not clear cut.

On the contrary, the unit root tests provide a more formal method for determining the degree of differencing and gives a quantitative and rigorous way to check whether or not the given AR model in time series is stationary. The presence of a unit root indicates that the time series is not stationary and that differencing will make it stationary. To make matters precise, the value of d in $ARIMA(p,d,q)$, autoregressive integrated moving

average of order p , d , q , indicates the number of differencing necessary to make the series stationary and the number d also equals the number of unit roots in the autoregressive characteristic equation for the time series. Therefore, unit root tests provide a formal and quantitative method when the degree of differencing needs to be determined. Methods for detecting the presence of a unit root in parametric time series models have lately attracted a great deal of interest in both statistical theory and applications. One major field of application where the hypothesis of a unit root has important implications is economics. This is because a unit root is often a theoretical implication of models which postulate the rational use of information that is available to economic agents. Examples include various financial market variables such as futures contracts (Samuelson, 1965), stock price (Samuelson, 1973), dividends (Kleidon, 1986), spot and forward exchange rates (Meese and Singleton, 1983), and even aggregate variables like real consumption (Hall, 1978). Formal statistical tests of the unit root hypothesis are of additional interest to economists because they can help evaluate the nature of nonstationarity that most macroeconomic data exhibit. In particular, economists determine whether the trend is stochastic, through the presence of a unit root, or deterministic, through the presence of a polynomial time trend.

Next we consider the topic of model misspecification. Model misspecification is a common problem in statistical data analysis. A time series model can be misspecified in a variety of ways. One may regress a nonstationary time series on another unrelated nonstationary time series; or one may misspecify the order of a nonstationary AR or a polynomial regression model; or one may misspecify a polynomial regression as a nonstationary AR model or vice versa. In the misspecified models inference using the

usual regression statistics such as t , Durbin-Watson (DW) and R^2 can be very often misleading, and these statistics can be useful tools to detect the model misspecification at the same time.

1.2. Research purpose

We consider the following general restricted model.

$$\mathbf{a}(L)\mathbf{y}_t = \mathbf{b}(L)\mathbf{e}_t, \quad t = 1, 2, \dots, n,$$

where $\{\mathbf{e}_t\}$ is a sequence of independent identically distributed (iid) random vectors with variance covariance matrix Ω and L is the lag operator such that $L^k \mathbf{y}_t = \mathbf{y}_{t-k}$,

$$\mathbf{a}(L) = \mathbf{I} + \mathbf{A}_1 L + \dots + \mathbf{A}_p L^p, \quad \mathbf{b}(L) = \mathbf{I} + \mathbf{B}_1 L + \dots + \mathbf{B}_q L^q.$$

It is assumed that (p, q) are nonnegative integers and known, and $\mathbf{A} = (\mathbf{A}_1 | \mathbf{A}_2 | \dots | \mathbf{A}_p)$, $\mathbf{B} = (\mathbf{B}_1 | \mathbf{B}_2 | \dots | \mathbf{B}_q)$ are respectively $m \times mp$ and $m \times mq$ matrices of unknown parameters. Let $\{\mathbf{y}_t\}$, $t = 1, \dots, n$ is the set of $m \times 1$ observation vectors and $\{\mathbf{e}_t\}$ is a sequence of iid m -dimensional random vectors with mean zero and a nonsingular variance covariance matrix Ω . For stationarity, invertibility and identifiability, all the roots of $\det[\mathbf{a}(L)]$ are assumed to lie outside the unit circle and \mathbf{A}_p is of full rank. The vector of restrictions on parameters \mathbf{A} , \mathbf{B} and Ω is defined as

$$\mathbf{f}(\mathbf{A}, \mathbf{B}, \Omega) = \mathbf{0}$$

where \mathbf{f} is a k -dimensional vector of differential functions.

There are two different situations which we may model with the above general restricted model as follows:

Case 1. Testing for a unit root with parametric restrictions when parameters \mathbf{B} in general restricted model are known.

Case 2. Testing for a unit root with parametric restrictions when parameters \mathbf{B} in general restricted model are unknown.

As far as Case 1 is concerned, Shin and Sarkar (1994) showed that an ARIMA(p , 1, 0) signal distributed by MA(q) noise is an ARIMA(p , 1, $p+q+1$) process restricted by nonlinear constraints on parameters and for this model with a unit root they presented a modified restricted maximum likelihood estimator (MRMLE)(Shin and Sarkar, 1994, p. 2649), obtained by maximizing a modified likelihood function, that has the same limiting properties as the restricted MLE (Shin and Sarkar 1994, p. 2648) and is computationally much simpler. They showed that unit root tests based on the MRMLE perform very well in small samples and compare favorably with the Said and Dickey (1985) tests with respect to both sizes and powers through simulation study for the case $p = 0$ and $q = 1$. Shin and Sarkar (1995) considered the multivariate ARMA model with nonlinear parametric restriction in the context of Case 2 and proposed a simple and easy-to-compute Newton-Raphson estimator that approximates the restricted ML estimator which uses fully the information contained in the restriction.

As an application of the general restricted model and restrictions on parameters, we consider the case with $m = 1$, $p = 1$, $q = 1$, i.e., we model z_t and u_t as AR(1) and MA(1) respectively

$$y_t = z_t + u_t, \quad a(L)z_t = \varepsilon_t, \quad u_t = c(L)w_t \quad (1.2.1)$$

where

$$c(L) = 1 + C_1L, \quad a(L) = 1 + A_1L \quad (1.2.2)$$

with backshift operator L .

The model (1.2.1) can be expressed as follows:

$$y_t = -A_1z_{t-1} + \varepsilon_t + w_t + C_1w_{t-1}, \quad (1.2.3)$$

subject to the restriction $\mathbf{f} = (f_1, f_2)' = (0, 0)'$.

We will study through simulation experiments on model (1.2.1)-(1.2.3) the power functions of unit root tests based on the following estimation methods:

- (1) Hannan and Rissanen (1982)(HR) method.
- (2) Kohn (1979) method.
- (3) Shin and Sarkar (1995)(SS) method.

In order to obtain the initial parameter estimates in the Newton-Raphson iterative procedure, the first and second stages of multivariate version of Hannan-Rissanen's procedure can be used to both Kohn's and Shin-Sarkar's methods respectively.

Our objective is to consider the limiting null distribution of the unit roots and to check which of the above three different unit root tests perform well with respect to both

sizes and powers based on model (1.2.1)-(1.2.3) and thus apply these results to several fields such as the engineering science and economics. In chapter 3 and 4 three different estimation methods are introduced and each method will be applied to test for a unit root in an AR(1) signal observed with MA(1) noise, respectively. Finally, we will check the above three different unit root tests with respect to both sizes and powers based on the given model. Ten thousand replications will be simulated for sample size 25, 50, 100 and 250, for the coefficient of autoregressive signal $A_1 = 1.00, 0.99, 0.95, 0.90, 0.80$ and 0.70 (see equation (3.2.10), p. 62) and for nominal levels 1%, 5% and 10%. The normal random variables will be generated by the subroutine RNNOA of the IMSL subroutine library. The computed values of the tests $n(\hat{A}_1 - 1)$ and $\hat{\tau}$ based on three different methods will be compared to the theoretical 1%, 5% and 10% left tail critical values tabulated by Dickey and Fuller (see Fuller, 1976, p. 371).

Next we discuss our research on model misspecification. Shin and Sarkar (1996) considered a general regression model with integrated errors and one system of integrated regressors and worked asymptotic properties of the conventional regression statistics under this model. Specifically, they discussed three types of model misspecifications as applications based on the analysis of the general regression model (5.1.1): (a) how to detect existence of cointegration; (b) how to decide whether a nonstationary AR(p) model or a polynomial regression model gives a good fit; (c) the t-statistic has more power than its alternative test statistic discussed by Dickey and Fuller (1979) for testing the single unit root null hypothesis against the double unit roots hypothesis. They concluded that: the behavior of the regression statistics depends on whether the error term $U_{t,k}$ is stationary or nonstationary.

Consider a general regression model

$$y_t = \beta_0 + \beta_1 X_{t,1} + \dots + \beta_p X_{t,p} + U_{t,k}, \quad t = 1, \dots, n, \quad (1.2.4)$$

where

$$X_{t,j} = X_{1,j-1} + \dots + X_{t,j-1}, \quad j = 1, 2, \dots, p, \quad X_{t,0} = x_t,$$

$$U_{t,j} = U_{1,j-1} + \dots + U_{t,j-1}, \quad j = 1, \dots, k, \quad U_{t,0} = u_t.$$

If (1.2.4) is misspecified then $U_{t,k}$ is nonstationary; $\hat{\sigma}^2 = \sum_{t=1}^n \hat{U}_{t,k}^2 / (n-p-1)$ diverges; $DW = \sum (\hat{U}_{t,k} - \hat{U}_{t-1,k})^2 / \sum \hat{U}_{t,k}^2 \xrightarrow{p} 0$; $\hat{\beta}_j$ diverges for j ; if the probabilistic order of the regressors is not greater than that of the regression error, then the limit of

$$R^2 = \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - n\bar{y}^2 / (\mathbf{y}'\mathbf{y} - n\bar{y}^2)$$

is less than one, otherwise $R^2 \xrightarrow{p} 1$; and the t-statistic diverges, implying misleading high significance of the t-statistics (Theorems 1-2 of Shin and Sarkar, 1996). On the other hand, if the stationarity of y_t is well explained by $X_{t,j}$'s then the error term $U_{t,k}$ will be stationary. In this case $\hat{\beta}_j$ is consistent; $\hat{\sigma}^2$ converges; DW converges to a positive number; $R^2 \xrightarrow{p} 1$; t statistics converges (Theorem 3 of Shin and Sarkar, 1996). The above facts developed by Shin and Sarkar (1996) will provide the theoretical background for our simulation results. We will analyze the above three different situations through a Monte Carlo study and examine each model misspecification problem by using the usual

OLS statistics such as $\hat{\sigma}^2$, t, DW and R^2 as diagnostic tools. Ten thousand samples will be simulated for sample sizes 25, 50, 100 and 250 using the normal random numbers will be generated by the subroutine RNNOA of the IMSL subroutine library.

1.3. Some definitions

(1) What is Stationarity ?

When useful estimates of the parameters in time series process are expected to be obtained, the stationarity among a number of properties in satisfying the modeling are to be considered. Consider a finite set of random variables $\{Y_{t_1}, \dots, Y_{t_n}\}$ from stochastic process $\{Y_t; t = 0, \pm 1, \pm 2, \dots\}$. A process is said to be *strictly n-th order stationary* if the n dimensional distribution function F is time invariant, i.e., $F\{Y_{t_1}, \dots, Y_{t_n}\} = F\{Y_{t_1+k}, \dots, Y_{t_n+k}\}$ for $\forall (t_1, \dots, t_n)$, k is true for any $n = 1, 2, \dots$. However, there are some difficulties in verifying a distribution function, particularly a joint distribution function from an observed time series. Therefore, we use a weaker sense of stationarity in terms of the moments of the model. That is, a process is said to be *nth order weakly stationary* if all its joint moments up to order n exist and are time invariant. Throughout this paper we use the terminology 'stationary' for *covariance stationary*, which has constant mean and variance with the covariance and the correlation being functions of the time difference alone, i.e.,

$$E(y_t) = \mu$$

$$E[(y_t - \mu)^2] = \sigma_y^2 = \gamma(0)$$

and

$$E[(y_t - \mu)(y_{t-m} - \mu)] = \gamma(m), \quad m=1, 2, \dots$$

(2) What is White Noise ?

A time series in which successive values are highly dependent can be usefully regarded as generated from a series of independent shock e_t . These shocks are uncorrelated random variables from a fixed distribution, usually assumed normal and having mean zero and constant variance σ_e^2 . Such a sequence of random variables e_t, e_{t-1}, \dots is called a *white noise* process by engineers. Although this process hardly ever occurs in applied time series, it plays an crucial role as a basic building block in the construction of the time series model.

(3) What is Random Walk ?

Consider the following ARIMA(p, d, q) model:

$$\phi(L)(1 - L)^d y_t = \theta_0 + \theta(L)e_t$$

where $\phi(L) = (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)$, $\theta(L) = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)$ with backshift operator L and θ_0 is a fixed value. If $p = 0$, $d = 1$, $q = 0$ in above ARIMA model, $(1-L)y_t = e_t$, $y_t = y_{t-1} + e_t$, then this model is called *random walk*. This behavior is similar to following a drunken man whose position at time t is his position at time $(t-1)$ plus a step in a random direction at time t . Note that the random walk model is the limiting process of ARIMA(1, 0, 0), AR(1), process $(1 - \phi_1 L)y_t = e_t$ with $\phi_1 \rightarrow 1$. Moreover, random walk

has the distinct characteristic of having stationary mean and nonstationary variance based on AR(1) model with $\phi_1 = 1$ and white noise e_t . Finally, the random walk phenomenon can be characterized by large nonvanishing spikes in the same ACF of the original series $\{y_t\}$ and insignificant zero ACF for the differenced series $\{(1-L)y_t\}$.

(4) What is the Measurement Error Model ?

It is not always possible to observe a time series x_t directly. Instead of observing x_t one observes the sum

$$y_t = x_t + u_t \quad (1.3.1)$$

where u_t is a $N(0, \sigma_u^2)$ random variable, called *measurement error*. The observed variable y_t is sometimes called the *manifest variable* or the *indicator variable*. The unobserved variable x_t is called a *latent variable* in certain areas of application. Models with fixed x_t are called *functional models*, while models with random x_t are called *structural models*. As an example of a situation where x_t can not be observed, consider the relationship between the yield of corn and available nitrogen in the soil. Assume that there is an adequate approximation to the linear relationship between yield and nitrogen. In order to estimate the available soil nitrogen, it is necessary to sample the soil of the experimental plot and to perform a laboratory analysis on the selected sample. As a result of sampling and of the laboratory analysis, we do not observe x_t but observe an estimate of x_t . Therefore, we represent the observed nitrogen by y_t , where y_t satisfies (1.3.1) and u_t is the measurement error introduced by sampling and laboratory analysis. Estimation of the true value x_t is very important in signal measurement problems encountered in the

engineering sciences where x_t is known as signal and u_t is known as noise. It is also important in survey sampling in the situation where a population is sampled at a sequence of time points according to a sample scheme. In a survey sampling, y_t , x_t and u_t denote observed direct survey estimates, unobservable population values and the sampling errors respectively.

(5) What is the Autocovariance Generating Function ?

For each of the covariance-stationary processes for y_t , we calculate the sequence of autocovariances. If this sequence is absolutely summable, then one way of summarizing the covariances is through a scalar-valued function called the *autocovariance generating function*. Frequently this function is easy to calculate, in which case the autocovariance at lag k may be determined by identifying the coefficient of either z^k or z^{-k} . If $\{y_t\}$ is a stationary process with autocovariances, then its autocovariance generating function is defined by

$$g(z) = \sum_{j=-\infty}^{\infty} \gamma_j z^j,$$

provided the series converges for all z in some annulus $r^{-1} < |z| < r$ with $r > 1$.

For a covariance-stationary vector processes \mathbf{y}_t with an absolutely summable sequence of covariance matrices, the analogous matrix-valued autocovariance generating function $\mathbf{G}(z)$ is defined as

$$\mathbf{G}(z) = \sum_{j=-\infty}^{\infty} \Gamma_j z^j,$$

where $\Gamma_j = E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-j} - \boldsymbol{\mu})']$ with mean vector of \mathbf{y}_t , $\boldsymbol{\mu}$.

(6) What are Integration and Cointegration ?

If a series with no deterministic component has a stationary, invertible ARMA representation after differencing d times, it is said to be *integrated of order d* . If two series, y_{1t} and y_{2t} , are both $I(d)$, it will normally be the case that any linear combination is also $I(d)$. However, it is possible that there is a linear combination of the two series for which the order of the integration is smaller than d . In this case the series are said to be cointegrated. More generally, the components of the vector time series \mathbf{y}_t are said to be *cointegrated of order d, b* if (a) all components of \mathbf{y}_t are $I(d)$, and (b) there exists a non-null vector, \mathbf{a} , such that $\mathbf{a}'\mathbf{y}_t$ is $I(d-b)$ with $b > 0$. Cointegration means that although many developments can cause permanent changes in the individual elements of \mathbf{y}_t , there is some long-run equilibrium relation tying the individual components together, represented by the linear combination $\mathbf{a}'\mathbf{y}_t$ and thus certain constraints are needed to impose on the multivariate time series model. An example of such a system is the model of consumption spending. Although both consumption and income exhibit a unit root, over the long run consumption tends to be a roughly constant proportion of income, so that the difference between the log of consumption and the log of income appears to be a stationary process.

(7) What is Brownian Motion (or Wiener) process?

A stochastic process $[X(t), t \geq 0]$ is said to be a *Brownian motion process* if:

- (a) $X(0) = 0$;
- (b) $\{X(t), t \geq 0\}$ has stationary independent increments;

(c) for every $t > 0$, $X(t)$ is normally distributed with mean 0 and variance c^2t , where c is some positive constant.

When $c = 1$, the process is often called *standard Brownian motion*. As any Brownian motion can always be converted to the standard process by looking at $X(t)/c$, we shall suppose throughout that $c = 1$. This process has been used beneficially in such areas as statistical testing of goodness of fit, analyzing the price levels on the stock market, and quantum mechanics.

CHAPTER II

PREVIOUS WORKS IN TESTING FOR A UNIT ROOT WITHOUT PARAMETRIC CONSTRAINTS IN TIME SERIES

There has been a huge amount of work done on unit roots tests so far, but only a few papers are reviewed in this chapter. Section 2.1 discusses statistical inference for ARIMA(p, 1, q) models and Section 2.2 for measurement error models.

2.1. Previous results on ARIMA(p, 1, q) models

In Section 2.1.1 we review unit roots tests in univariate time series and in multivariate time series in Section 2.1.2.

2.1.1. Unit roots in univariate time series

Box and Jenkins (1976) proposed an “interactive modeling process” that consists

of model formulation (“identification”), estimation and diagnostic checking, one key element of which is the use of integrated, or “difference-stationary,” ARMA models,

$$(1 - L)^d \Phi(L)y_t = \Theta(L)e_t.$$

Thus, although y_t is nonstationary, it is assumed that its d -th difference is a stationary ARMA process. The appropriateness of the Box-Jenkins tactic of differencing to achieve stationarity depends on the existence of one or more unit roots in the autoregressive lag operator polynomial. The desirability of specific tests for unit roots is therefore apparent.

Dickey (1976) and Fuller (1976) used the Monte Carlo method to calculate the percentage points of the finite sample as well as the asymptotic distribution of the least squares (LS) estimator of the AR parameter as well as its “t-statistic,” when the true AR parameter is positive or negative one. Specifically, they consider the null model

$$H_0: y_t = y_{t-1} + e_t, \quad e_t \sim \text{NID}(0, \sigma^2), \quad y_0 = 0$$

and the alternative

$$\text{Model 1: } H_A: y_t = \rho y_{t-1} + e_t, \quad e_t \sim \text{NID}(0, \sigma^2), \quad \rho \neq 1, \quad y_0 = 0$$

and consider the LS estimator (which is also ML, under the normality assumption)

$$\hat{\rho}_{LS} = \frac{\sum y_t y_{t-1}}{\sum y_{t-1}^2}, \quad (2.1.1)$$

obtained from a sample of size n . Under the null hypothesis $(\hat{\rho}_{LS} - 1) = O_p(n^{-1})$; the convergence is faster than for $|\rho| < 1$, in which case $(\hat{\rho}_{LS} - \rho) = O_p(n^{-1/2})$. It follows that the proper quantity for which percentage points should be calculated under the null is $n(\hat{\rho}_{LS} - 1)$. Dickey and Fuller (1979) tabulate the asymptotic as well as the finite sample percentage points of $n(\hat{\rho} - 1)$ for various sample sizes. Under the null hypothesis of a unit root, the distribution of $n(\hat{\rho} - 1)$ is skewed to the left, and $\lim_{n \rightarrow \infty} P(\hat{\rho} < 1) = 0.68$. They also consider the usual “student’s t ” for testing $\rho = 1$,

$$\hat{\tau} = \frac{\hat{\rho} - 1}{\sqrt{s^2 \left(\sum_{t=2}^n y_{t-1}^2 \right)^{-1}}}$$

where

$$s^2 = \sum_{t=2}^n (y_t - \hat{\rho} y_{t-1})^2 / (n - 2).$$

Under the null, $\hat{\tau} = O_p(1)$, but it does not have the t distribution; again, they calculate its percentage points for various sample sizes using Monte Carlo simulation. Rewriting the first-order autoregression as

$$y_t - y_{t-1} = (\rho - 1)y_{t-1} + e_t. \quad (2.1.2)$$

Thus, $\hat{\tau}$ is the usual t -statistic in a regression of the first difference of y_t on the first lag of y_t .

Dickey and Fuller (1979) showed that $\hat{\tau}$ is a monotone function of the likelihood ratio for the null hypothesis of $\rho = 1$ vs. the two-sided alternative and they obtained the limiting distribution of $n(\hat{\rho} - 1)$. Despite the non-standard properties of the LS estimator in a nonstationary AR(1) model, a test of the hypothesis that the process is a random walk against the alternative that it is stationary can be carried out using $n(\hat{\rho} - 1)$ or the regression t-statistic $\hat{\tau}$ and comparing its value with tabulated percentiles of the distribution by Monte Carlo simulation. In a similar manner, unit root testing in two more general AR(1) model can also be developed:

Model 2: $y_t - \mu = \rho(y_t - \mu) + e_t$, $t = 1, \dots$, where $y_0 = 0$, $e_t \sim \text{NID}(0, \sigma^2)$

Model 3: $y_t - a - bt = \rho\{y_{t-1} - a - b(t-1)\} + e_t$, $t = 1, \dots$, where $y_0 = 0$, $e_t \sim \text{NID}(0, \sigma^2)$

and the limiting distributions of the three different LS estimators, $\hat{\rho}$, $\hat{\rho}_\mu$, and $\hat{\rho}_\tau$ respectively can be derived. The LS estimator $\hat{\rho}$ of Model 1 is given by (2.1.1) for the model when the true value of ρ is positive or negative one. The limiting distribution of $n(\hat{\rho} - 1)$ is

$$n(\hat{\rho} - 1) = \frac{\frac{1}{n} \sum_{t=1}^n y_{t-1} e_t}{\frac{1}{n^2} \sum_{t=1}^n y_{t-1}^2} \xrightarrow{L} \frac{\frac{1}{2} [W(1)^2 - 1]}{\int_0^1 W(r)^2 dr}$$

where $W(\cdot)$ is the standard Brownian motion on $[0, 1]$. We could define the alternatives to $H_1: |\rho| < 1$ or $H_1: |\rho| > 1$. For either of these cases, one-sided test is appropriate. If the alternative is $|\rho| < 1$, then the model is asymptotically stationary, since the effect of the

fixed initial condition is transient. The alternative $|\rho| > 1$, on the other hand, is non-stationary for all sample sizes. Now consider allowance for a nonzero mean μ under the above alternative (Model 2), or $y_t = \alpha + \rho y_{t-1} + e_t$, where $\alpha = \mu(1-\rho)$. If μ were known, we could simply center the data and proceed as before. In practice, μ must be estimated along with the other parameters. Although α vanishes under the null hypothesis $\rho = 1$, it is nevertheless present under the alternative, and so an intercept is included in the regression. The corresponding statistics $\hat{\tau}_\mu$ and $n(\hat{\rho}_\mu - 1)$ have been tabulated under the null $(\alpha, \rho) = (0, 1)$ by Dickey (1976) and Fuller (1976). Similarly, the alternative might include a trend (Model 3), or $y_t = \alpha + \beta t + \rho y_{t-1} + e_t$ where $\alpha = a(1-\rho) + b\rho$ and $\beta = b(1-\rho)$. Under the null hypothesis that $\rho = 1$,

$$y_t = b + y_{t-1} + e_t,$$

but under the trend-stationary alternative both the intercept and trend enter and so they must be included in the regression. Dickey (1976) and Fuller (1976) tabulated the distributions of normalized-bias and studentized statistics, $\hat{\tau}_\tau$ and $n(\hat{\rho}_\tau - 1)$, respectively, under $H_0: (\alpha, \beta) = (0, 0)$. Furthermore, they showed that the null distributions of the two test statistics are unaffected by the value of b in Model 3 and the representations for the limit distributions of all the test statistics under the null are invariant to the choice of y_0 .

Dickey and Fuller (1981) extended unit roots tests in AR(1) to higher order AR process. The AR(p) model may be reformulated as

$$(y_t - y_{t-1}) = \rho^* y_{t-1} + \sum_{j=1}^{p-1} \rho_j^* (y_{t-j} - y_{t-j+1}) + e_t, \quad e_t \sim \text{NID}(0, \sigma^2) \quad (2.1.3)$$

where $\rho^* = \sum_{k=j+1}^p \rho_k$, $j = 1, 2, \dots, p-1$ and $\rho^* = \sum_{k=1}^p \rho_k - 1$. When ρ^* is zero, the model reduces to an AR(p-1) in first differences, i.e., ARIMA(p-1, 1, 0). Thus, a test of $H_0: \rho^* = 0$ is the unit root test, which is called “*Augmented Dickey-Fuller (ADF) test.*” The test could also be carried out by regressing y_t on y_{t-1} to y_{t-p} and testing the restriction that the coefficients sum to one. The asymptotic distribution of the t-statistic associated with y_{t-1} in (2.1.3) is the same as that of (2.1.2). The t-statistics associated with the differenced variables in (2.1.3) are asymptotically standard normal, so that inference on the lags can be conducted in the usual way. In addition for the case of a nonzero mean, we have

$$(y_t - \mu) = \sum_{j=1}^p \alpha_j (y_{t-j} - \mu) + e_t, \quad e_t \sim \text{NID}(0, \sigma^2), \quad y_0 = 0.$$

We can put this in the form

$$y_t = k + \rho^* y_{t-1} + \sum_{j=1}^{p-1} \rho_j^* (y_{t-j} - y_{t-j+1}) + e_t$$

where $k = \mu(1 - \sum_{j=1}^p \alpha_j)$. Under the null of a unit root, the intercept vanishes because in

that case $\sum_{j=1}^p \alpha_j = -1$. The distribution of the studentized statistic for testing $\rho^* = 1$ in this

regression is asymptotically identical to that of $\hat{\tau}_\mu$. Consider the alternative of a linear trend,

$$(y_t - a - bt) + \sum_{j=1}^p \alpha_j (y_{t-j} - a - b(t-j)) + e_t, \quad e_t \sim \text{NID}(0, \sigma^2), \quad y_0 = 0.$$

or

$$y_t = k_1 + k_2 t + \rho^* y_{t-1} + \sum_{j=1}^{p-1} \rho_j^* (y_{t-j} - y_{t-j+1}) + e_t,$$

where $k_1 = a(1 + \sum_{i=1}^p \alpha_i) - b \sum_{i=1}^p i \alpha_i$, $k_2 = bt(1 + \sum_{i=1}^p \alpha_i)$. Under the null, $k_1 = -b \sum_{i=1}^p i \alpha_i$ and $k_2 = 0$. The use of AR processes as approximations raises the question of the applicability of the ADF test. In theory the ADF test is only valid if the underlying model is indeed a finite AR process. When an MA part is present in model, a bias is in general introduced into the test statistics.

Said and Dickey (1984) made an initial attempt to extend the unit root tests to the general ARMA(p, q) case, where the p, q are unknown, by approximating the ARMA model as a finite autoregression and showed that as long as p goes to infinity sufficiently slow relative to n, then the OLS t test of $\rho = 1$ can continue to be compared with the Dickey-Fuller values (Table 8.5.2, Fuller, 1976). The Said-Dickey result permits the researcher to use a large value of p on which to base this comparison the larger is the sample size n. However, this has the problem that one loses the degree of freedom by having to both estimate a number of nuisance parameters and condition on increasing number of observations n. LS method can be used to estimate the coefficients, and this procedure produces test statistics whose limit distributions are the same as those tabulated

in Fuller (1976). It will prove convenient to consider first the simple case of H_0 : ARIMA(1, 1, 1) vs. H_1 : ARMA(2, 1). Suppose that

$$y_t = \rho y_{t-1} + x_t, \quad x_t = \alpha x_{t-1} + e_t + \beta e_{t-1}, \quad |\alpha|, |\beta| < 1, \quad e_t \sim \text{NID}(0, \sigma^2), \quad x_0 = y_0 = 0.$$

If $|\rho| < 1$, then $\{y_t\}$ is a stationary ARMA(2, 1) process except for transitory effects. On the other hand, if $\rho = 1$, then $\{y_t\}$ is ARIMA(1, 1, 1). Now note that

$$e_t = \sum_{j=0}^{\infty} (-\beta)^j (x_{t-j} - \alpha x_{t-j-1})$$

$$y_t = \rho y_{t-1} + (\alpha + \beta)(x_{t-1} - \beta x_{t-2} + \beta^2 x_{t-3} - \dots) + e_t.$$

We can use these results to write

$$y_t - y_{t-1} = (\rho - 1)y_{t-1} + (\alpha + \beta)(x_{t-1} - \beta x_{t-2} + \beta^2 x_{t-3} - \dots) + e_t.$$

Under the null hypothesis of $\rho = 1$, $x_t = y_t - y_{t-1} \equiv \Delta y_t$, and so we write

$$\Delta y_t = (\alpha + \beta)(\Delta y_{t-1} - \beta \Delta y_{t-2} + \beta^2 \Delta y_{t-3} - \dots) + e_t.$$

We can then develop a test of the null by regressing Δy_t on $y_{t-1}, \Delta y_{t-1}, \dots, \Delta y_{t-k}$, where $k = o_p(n^{1/3})$. The LS estimates of ρ, α and β obtained by this technique are consistent under the null. However, the distribution of $n(\hat{\rho} - 1)$ depends on the nuisance parameters α and

β and so it is not useful. On the other hand, the studentized statistic $\hat{\tau}$ associated with $(\hat{\rho}-1)$ does not involve any unknown parameters, and it has asymptotically the original Dickey-Fuller $\hat{\tau}$ distribution. Now consider the general ARMA($p+1, q$) model, which is written as

$$y_t = \rho y_{t-1} + x_t,$$

$$x_t + \sum_{i=1}^p \alpha_i x_{t-i} = e_t + \sum_{j=1}^q \beta_j e_{t-j}, \quad x_0 = y_0 = 0, \quad e_t \sim \text{NID}(0, \sigma^2).$$

where x_t is stationary and invertible. Then we can proceed exactly as before, estimating by using regression

$$\Delta y_t = (\rho - 1)y_{t-1} - d_1 \Delta y_{t-1} - \dots - d_k \Delta x_{t-k} + e_t.$$

We still get consistency under the null, and the distribution of the t-statistic on $(\hat{\rho}-1)$ is the same as before. Also, if the nonzero mean is allowed under the alternative, $\hat{\tau}_\mu$ may be used rather than $\hat{\tau}$.

Said and Dickey (1985) studied a one step Gauss-Newton procedure of estimating ρ starting with $\rho = 1$ based on ARIMA($p, 1, q$) model with known p and q . The model they suggested is as

$$y_t = \rho y_{t-1} + z_t \text{ and } z_t + \sum_{i=1}^p a_i z_{t-i} = e_t + \sum_{j=1}^q \beta_j e_{t-j} \text{ where } y_0 = 0, e_t \sim \text{NID}(0, \sigma^2).$$

(2.1.4)

Associated with the above z_t process are $(1 - a_1 L - \dots - a_p L^p)$, $(1 - \beta_1 L - \dots - \beta_q L^q)$, the roots of which are all assumed to be greater than one in absolute value. Viewing as a

difference equation subject to q initial conditions, $e_p = \delta_1, e_{p-1} = \delta_2, \dots, e_{p-q+1} = \delta_q$ holds, and thus model (2.1.4) can be rewritten as follows:

$$e_t = (y_t - \rho y_{t-1}) + a_1(y_{t-1} - \rho y_{t-2}) + \dots + a_p(y_{t-p} - \rho y_{t-p}) - \beta_1 e_{t-1} - \dots - \beta_q e_{t-q} \quad (2.1.5)$$

and let $\hat{\theta}' = (\hat{\rho}, a_1, \dots, a_p, \hat{\beta}_1, \dots, \hat{\beta}_q, \hat{\delta}_1, \dots, \hat{\delta}_q)$ and expand $e_t(\theta)$ about $\hat{\theta}$ to obtain

$$e_t(\theta) = e_t(\hat{\theta}) - v_t(\hat{\theta})(\rho - \hat{\rho}) - \sum_{i=1}^p x_{i,t}(\hat{\theta})(a_i - \hat{a}_i) - \sum_{j=1}^q w_{j,t}(\hat{\theta})(\beta_j - \hat{\beta}_j) - \Delta_{k,t}(\hat{\theta})(\delta_k - \hat{\delta}_k) + r_t \quad (2.1.6)$$

where r_t is the Taylor series remainder and $-v_t(\hat{\theta})$, $-x_{i,t}(\hat{\theta})$, $-w_{j,t}(\hat{\theta})$, and $-\Delta_{k,t}(\hat{\theta})$ are the derivatives of $e_t(\theta)$ with respect to ρ , a_i , β_j and δ_k , respectively, evaluated at $\hat{\theta}$. Ignoring the remainder term and rearranging (2.1.6) as

$$e_t(\hat{\theta}) = v_t(\hat{\theta})(\rho - \hat{\rho}) + \sum_{i=1}^p x_{i,t}(\hat{\theta})(a_i - \hat{a}_i) + \sum_{j=1}^q w_{j,t}(\hat{\theta})(\beta_j - \hat{\beta}_j) + \Delta_{k,t}(\hat{\theta})(\delta_k - \hat{\delta}_k) + e_t \quad (2.1.7)$$

suggests the regression of $e_t(\hat{\theta})$ on $v_t(\hat{\theta})$, $\{x_{i,t}(\hat{\theta})\}_{i=1}^p$, $\{w_{j,t}(\hat{\theta})\}_{j=1}^q$ and $\{\Delta_{k,t}(\hat{\theta})\}_{k=1}^q$ to estimate $\gamma' = (\rho - \hat{\rho}, a_1 - \hat{a}_1, \dots, a_p - \hat{a}_p, \beta_1 - \hat{\beta}_1, \dots, \beta_q - \hat{\beta}_q, \delta_1 - \hat{\delta}_1, \dots, \delta_q - \hat{\delta}_q)$. The derivatives in (2.1.7) can be computed recursively as follows:

$$\begin{aligned}
v_t(\hat{\theta}) &= y_{t-1} + \hat{\alpha}_1 y_{t-2} + \dots + \hat{\alpha}_p y_{t-p-1} - \hat{\beta}_1 v_{t-1} - \dots - \hat{\beta}_q v_{t-q}, \\
x_{i,t}(\hat{\theta}) &= -\hat{\beta}_1 x_{i,t-1} - \dots - \hat{\beta}_q x_{i,t-q} - z_{t-i}, \\
w_{j,t}(\hat{\theta}) &= e_{t-j} - \hat{\beta}_1 w_{j,t-1} - \dots - \hat{\beta}_q w_{j,t-q},
\end{aligned}$$

and

$$\Delta_{k,t} = \hat{\beta}_1 \Delta_{k,t-1} - \dots - \hat{\beta}_q \Delta_{k,t-q}$$

where $i = 1, \dots, p$; $j = 1, \dots, q$ and $k = 1, \dots, q$. Computation is performed by taking the initial value of ρ to be 1 and using the method of moment estimator of $(\beta_1, \dots, \beta_q)$. Furthermore, since the limiting distribution of $n(\hat{\rho}-1)$ is identical to that of Dickey and Fuller (1979) and Fuller (1976), the tables suggested by them can also be applied.

Phillips (1987) suggested a nonparametric approach with respect to nuisance parameters and thereby allows for a very wide class of time series models in which there is a unit root. This includes ARIMA models with MA part with heterogeneously as well as identically distributed innovations. The basic idea is to estimate a non-augmented Dickey-Fuller regression, $y_t = \rho y_{t-1} + e_t$ and then to correct the Dickey-Fuller normalized-bias $n(\hat{\rho}-1)$ and studentized statistic $\hat{\tau}$ for general forms of serial correlation and/or heteroskedasticity that might be present in e_t . Implicitly, the procedure amounts to semiparametric estimation of ρ , accounting for the infinite-dimensional nuisance parameter associated with e_t . Moreover, he modified the $\hat{\tau}$, $n(\hat{\rho}-1)$ tabulated by Dickey-Fuller when the data follow an ARMA(p, q) process, and developed new tests for unit roots that apply under general conditions. He defined the new statistics as

$$Z_\rho = n(\hat{\rho} - 1) - \frac{1}{2}(S_{nk}^2 - S_e^2)(n^{-2} \sum_{t=2}^n (y_{t-1} - \bar{y})^2)^{-1}$$

and

$$Z_\tau = \left(\sum_{t=1}^n y_{t-1}^2 \right) (\hat{\rho} - 1) / S_{nk} - \frac{1}{2} (S_{nk}^2 - S_e^2) \left(\sum_{t=1}^n y_{t-1}^2 \right)^{-1/2} S_{nk}^{-1} n^2 \quad (2.1.8)$$

where

$$S_e^2 = \sum_{t=1}^n (y_t - y_{t-1})^2 / n = \frac{1}{n} \sum_{t=1}^n e_t^2,$$

$$S_{nk}^2 = \sum_{t=1}^n e_t^2 / n + 2 \sum_{m=1}^k \omega_{mk} \sum_{t=m+1}^n e_t e_{t-m} / n,$$

and

$$\omega_{mk} = 1 - m/(1 + k).$$

The cutoff point k of the weighted autocovariances should grow with sample size, but at a slower rate, to ensure convergence to the asymptotic Dickey-Fuller distribution. In this case, Z_ρ is a transformation of the standardized estimator $n(\hat{\rho} - 1)$ and Z_τ is a transformation of regression t statistic t_ρ with

$$t_\rho = \frac{\sqrt{\sum_{t=1}^n y_{t-1}^2}}{\sqrt{\sum_{t=1}^n (y_t - \rho y_{t-1})^2 / n}} (\hat{\rho} - 1)$$

and

$$n(\hat{\rho} - 1) = \frac{\sum_{t=1}^n y_{t-1} (y_t - y_{t-1}) / n}{2 \sum_{t=1}^n y_{t-1}^2 / n} \quad \text{based on } y_t = \rho y_{t-1} + e_t, \rho = 1.$$

The limiting distribution of Z_τ is the same as that of the regression t statistics as long as $\lim_{n \rightarrow \infty} \left\{ \sum_t E(e_t^2 / n) \right\} = \lim_{n \rightarrow \infty} \left\{ \sum E(S_n^2 / n) \right\}$. Without depending on strong assumptions such as independence, homoskedasticity and normality, the given statistics Z_τ , Z_ρ allow for quite general weakly dependent and heterogeneously distributed innovations based on an asymptotic theory for the LS regression estimators and the associated regression t statistics in (2.1.8). This method seems to have significant advantages when there are MA components in the time series and, at least in this respect, offers a promising alternative to the Dickey and Fuller (1979) and Said and Dickey (1984) procedures.

Phillips and Perron (1988) extended the study of Phillips (1987) to the cases where (1) a random walk with a shift in mean included (2) a random walk with a shift in mean and a linear trend are included in the specified model and suggested a modified test statistic which employed a correction. The limiting distributions of the following new test statistics are expressed as functionals of standard Brownian motion and are identical to those given in Fuller (1976), and these tests may be used with existing tabulations even if they allow for much more general innovations specification (p. 341).

$$\begin{aligned}
 Z_{\hat{\rho}} &= n(\hat{\rho} - 1) - \hat{\lambda} / \bar{m}_{yy}, \\
 Z_{t_{\hat{\rho}}} &= \frac{\hat{s}}{\hat{\sigma}_{nl}} t_{\hat{\rho}} - \frac{\hat{\lambda} \hat{\sigma}_{nl}}{\sqrt{\bar{m}_{yy}}}, \\
 Z_{t_{\hat{\rho}}} &= \frac{\hat{s}}{\hat{\sigma}_{nl}} t_{\hat{\rho}} + \hat{\lambda} \hat{\sigma}_{nl} \frac{m_y}{\sqrt{\bar{m}_{yy} m_{yy}}} \\
 Z_{\tilde{\rho}} &= n(\tilde{\rho} - 1) - \frac{\tilde{\lambda}}{M}
 \end{aligned}$$

$$Z_{t_{\hat{\rho}}} = \frac{\tilde{s}}{\tilde{\sigma}_{nl}} t_{\hat{\rho}} - \tilde{\lambda} \tilde{\sigma}_{nl} / \sqrt{M}$$

$$Z_{t_{\hat{\mu}}} = \frac{\tilde{s}}{\tilde{\sigma}_{nl}} t_{\hat{\mu}} - \tilde{\lambda} \tilde{\sigma}_{nl} / \sqrt{M(M + m_y^2)}$$

$$Z_{t_{\hat{\beta}}} = \frac{\tilde{s}}{\tilde{\sigma}_{nl}} t_{\hat{\beta}} - \tilde{\lambda} \tilde{\sigma}_{nl} \left(\frac{1}{2} m_y - m_{ty} \right) / \sqrt{\frac{M}{12} \overline{m}_{yy}}$$

$$m_{yy} = \sum y_t^2 / n^2, \overline{m}^2 = \sum (y_t - \bar{y})^2 / n^2, m_y = \sum y_t / \sqrt[3]{n^2}, m_{ty} = \sum t y_t / \sqrt[5]{n^2}$$

where

$$M = (1 - n^{-2}) m_{yy} - 12 m_y^2 + 12(1 + n^{-1}) m_y m_y - (4 + 6n^{-1} + 2n^{-2}) m_y^2,$$

$$\hat{\lambda} = (\hat{\sigma}_{nl}^2 - \hat{s}^2) / 2, \tilde{\lambda} = \hat{\lambda} / \hat{\sigma}_{nl}^2, \tilde{\lambda} = (\tilde{\sigma}_{nl}^2 - \tilde{s}^2) / 2, \check{\lambda} = \tilde{\lambda} / \tilde{\sigma}_{nl}^2$$

based on the two LS regression equations

$$y_t = \hat{\mu} + \hat{\rho} y_{t-1} + \hat{\mu}_t$$

$$y_t = \tilde{\mu} + \tilde{\beta} \left(t - \frac{1}{2} n \right) + \tilde{\rho} y_{t-1} + \tilde{\mu}_t$$

where $(\hat{\mu}, \hat{\rho})$ and $(\tilde{\mu}, \tilde{\beta}, \tilde{\rho})$ are the least squares estimates.

Pantula (1991) characterized the asymptotic null distribution of a variety of unit root test statistics in the model

$$y_t = y_{t-1} + e_t - \theta_n e_{t-1}$$

as $\theta \rightarrow 1$ with sample size. Specifically, let $\theta_n = 1 - \gamma/n^\delta$, where $0 < \gamma < 2$ and $\delta \leq 0$. As θ approaches unity, of course, the process approaches stationarity, since the first differenced process would then have a unit MA root. The asymptotic null distributions of the various test statistics may then be shown to depend on the speed with which $\theta \rightarrow 1$, i.e.,

on the value of δ . Pantula's examination of asymptotic size in a sequence of null models getting arbitrarily close to the alternative nicely complements the "roots local to unity" literature, in which asymptotic power is investigated against a sequence of alternatives getting arbitrarily close to the null.

Hall (1989) investigated a strategy for testing for a unit root in the situations considered by Phillips (1987) and Phillips and Perron (1988) where the innovations constitute a MA process of order q , denoted by $MA(q)$. The rationale behind the test procedure is based on the observation that the generalization of the Dickey and Fuller (1981) tests to the random walk with the MA process is complicated by three factors: the unit root, the MA innovations, and the use of LS to estimate the model. He proposed instrumental variables (IV) estimator which uses y_{t-k} for $k > q$ as the instrument on y_{t-1} when the underlying model is an $ARIMA(0, 1, q)$. Let $\{y_t\}$ be a time series generated by the model

$$y_t = \rho y_{t-1} + \mu_t, \quad t = 1, \dots, n, \quad \rho = 1$$

where y_0 has a fixed distribution independent of t , sample size, and

$$\mu_t = e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q}.$$

It is assumed that $\{e_t\} \sim \text{iid}(0, \sigma^2)$ for all t and $\sup_t \{E(|\mu_t|^{\beta+\varepsilon})\} < \infty$ for some $\beta > 2$, $\varepsilon > 0$.

Then the corresponding parameters for the three different models are estimated using the IV approach respectively are as follows, summations being over $t = k, \dots, n$:

$$\begin{aligned}
(1) \quad y_t &= \hat{\rho}y_{t-1} + \hat{\mu}_t \Rightarrow \hat{\rho} = \frac{\sum y_{t-k}y_t}{\sum y_{t-k}y_{t-1}} \\
(2) \quad y_t &= \tilde{\mu} + \tilde{\rho}y_{t-1} + \tilde{\mu}_t \Rightarrow \begin{bmatrix} \tilde{\rho} \\ \tilde{\mu} \end{bmatrix} = \begin{pmatrix} \sum y_{t-k}y_{t-1} & \sum y_{t-k} \\ \sum y_{t-1} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \sum y_{t-k}y_t \\ \sum y_t \end{pmatrix} \\
(3) \quad y_t &= \bar{\mu} + \bar{\beta}(t - \frac{n}{2}) + \bar{\rho}y_{t-1} + \bar{\mu}_t \Rightarrow \begin{pmatrix} \bar{\rho} \\ \bar{\mu} \\ \bar{\beta} \end{pmatrix} = \begin{pmatrix} \sum y_{t-k}y_{t-1} & \sum y_{t-k} & \sum y_{t-k}(t - \frac{n}{2}) \\ \sum y_{t-1} & n & \sum(t - \frac{n}{2}) \\ \sum y_{t-1}(t - \frac{n}{2}) & \sum(t - \frac{n}{2}) & \sum(t - \frac{n}{2})^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum y_{t-k}y_t \\ \sum y_t \\ \sum y_t(t - \frac{n}{2}) \end{pmatrix}
\end{aligned}$$

It is apparent that the use of instrumental variables to reduce the bias in the LS estimates might deliver a test with improved finite-sample performance relative to the Said and Dickey (1984) and Phillips (1987) approaches. Furthermore, the limiting distribution of the estimator, properly normalized, of the coefficient of y_{t-1} converges to the distribution tabulated by Dickey (1976) and Fuller (1976) and when multiplied by a sample covariance ratio t statistics for the coefficient on y_{t-1} converges to the distribution tabulated by Dickey (1976) and Fuller (1976). This test may perform relatively favorably because it uses more information on model structure than Phillips's (1987) approach and avoids the loss of information inherent in Said and Dickey (1984) approach.

The Monte Carlo evidence in Schwert (1989) which examines the effects of model misspecification on the size of unit root tests for mixed ARIMA process indicated that the above modifications of test statistics may not be very successful. For instances, consider the ARIMA(0, 1, 1) model, $y_t = y_{t-1} + e_t - \theta e_{t-1}$, $t = -19, \dots, n$ where the errors $\{e_t\}$ are serially uncorrelated standard normal variables, the data are generated by setting e_{-20} and y_{-20} equal to zero and creating $n+20$ observations, discarding the first 20 observations to remove the effect of the initial conditions. A negative value of θ close to

minus one means that the series is not too far from being stationary white noise. In this case, all the above tests reject the null hypothesis of a unit root much more frequently than they should. Thus for $n = 100$ and $\theta = -0.8$, Pantula (1991) showed that for a nominal test size of 0.05, the ADF tests based on AR models with $p = 4$ and $p = 8$ have empirical sizes of 0.36 and 0.11 respectively, Hall's (1989) IV test has a size of 0.22, while Phillips-Perron (1988) tests have sizes of over 0.85. And simulations in Schwert (1989) showed that the tests for a unit root developed by Dickey and Fuller (1979, 1981) are sensitive to the assumption that the data are generated by a pure AR process. When the underlying process contains a MA part, the distribution of the unit root test statistics can be much different from the distribution defined by Dickey and Fuller (1979, 1981). Moreover the tests proposed by Said and Dickey (1984), Phillips (1987) and Phillips and Perron (1988) to correct the model misspecification problem do not appear to work well as long as the MA parameter is large. Both Pantula (1991) and Schwert (1989) recommended the ADF test, while conceding that the question of how to decide on appropriate value of p is difficult one to determine. The ADF test remains attractive.

Pantula and Hall (1991) suggested an approach based on an instrumental variable estimator for testing the null hypothesis that a process y_t is an ARIMA($p, 1, q$) against the alternative that it is stationary ARIMA($p+1, 0, q$) process, which is the extension of ARIMA($0, 1, q$) in Hall (1989). They also considered the limiting distribution of the instrumental variable estimator when the estimated model is either (1) the true model, (2) the true model with a shift in mean included, or (3) the true model with a shift in mean and a linear time trend included. Their simulation showed that the criteria based on the IV approach seem to perform as well as or better than the existing methods when the

model is specified correctly (or when p and q are known). However, if the model is overspecified, then empirical levels are higher than the nominal level in moderate-sized samples, whereas if the model is underspecified the IV estimators are inconsistent.

Shin and Pantula (1993) proposed problem of testing for a unit root in an AR model where the data are available only from each s -th period, called systematic sampling. In some applications, the observations may be available only at periodic intervals. For example, economic data may suggest a monthly model but only quartile (or yearly) data are available; or in a manufacturing process the data are collected only at regular time intervals because of cost considerations. They consider an AR(p) process x_t given by

$$x_t = \sum_{i=1}^p \phi_{t-i} x_{t-i} + e_t \text{ where } \{e_t\} \sim \text{iid}(0, \sigma^2).$$

Assume that y_i for only each s -th period $y_1 = x_s, y_2 = x_{2s}, \dots, y_n = x_{ns}$ are available, where s is a positive integer. They show that y_t may be modeled by a ARMA(p, q) where $q = \max\{q: q \leq p - (s/p)\}$. Using a Monte Carlo study they compared the IV approach by Pantula and Hall (1991) and nonlinear least squares estimators by Said and Dickey (1985) and Shin and Fuller (1992) and obtained results that methods based on approximated ARMA(1, $p-1$) performed better than those based on ARMA($p, p-1$).

We now discuss recursive estimation of ARMA(p, q) stationary process suggested by Hannan and Rissanen (1982) for the purpose of comparative study in Chapter 3. They considered a ARMA(p, q) stationary process generating a sequence y_t according to the

ARMA model $\sum_{j=0}^p \alpha_j y_{t-j} = \sum_{i=0}^q \beta_i \varepsilon_{t-i}$ for certain values, $p = p_0$, $q = q_0$, $\alpha_j = \alpha_{0,j}$, $\beta_j = \beta_{0,j}$ of the coefficients. It is assumed that either α_{0,p_0} or β_{0,q_0} is not zero with $E(\varepsilon_t) = 0$, and $E(\varepsilon_s \varepsilon_t) = \delta_{st} \sigma^2$. In order to estimate p_0 , q_0 , $\alpha_{0,j}$ and $\beta_{0,j}$ from observations y_1, \dots, y_T , we assume for the true system that the polynomials

$$a_{0,z} = \sum_{j=0}^p \alpha_{0,j} z^j, b_{0,z} = \sum_{j=0}^q \beta_{0,j} z^j, \alpha_{0,0} = \beta_{0,0} = 1,$$

have no common factors and that $a_{0,z} \neq 0$, $b_{0,z} \neq 0$ for $|z| \leq 1$. Then in the first stage estimates, $\hat{\varepsilon}_t$ are found by fitting a high order AR model for the known p_0, q_0 . Thus the $\hat{\varepsilon}_t$

are formed, for n large, as $\hat{\varepsilon}_t = \sum_{j=0}^n \hat{a}_{n,j} y_{t-j}$, $\hat{a}_{n,0} = 1$, where $\hat{a}_{n,j}$ are AR coefficients

estimated through the Yule-Walker equations. If $c_t = \frac{1}{T} \sum_{s=1}^{T-t} y_s y_{s+t}$, then $\hat{a}_{n,j}$ may be

recursively calculated through the equations

$$\hat{a}_{n,n} = -\sum_{j=0}^{n-1} \hat{a}_{n-1,j} c_{n-j} / \hat{\sigma}_{n-1}^2, \hat{a}_{n,j} = \hat{a}_{n-1,j} + \hat{a}_{n,n} \hat{a}_{n-1,n-j}, \hat{\sigma}_n^2 = (1 - \hat{a}_{n,n}^2) / \hat{\sigma}_{n-1}^2.$$

At this time, n is assumed to be chosen so as to increase slightly faster than $\log T$ but no faster than some power of $\log T$ by considering the criterion $\log \hat{\sigma}_m^2 + (m \log T) / T$, provided n was not allowed to be too large, where $m = \max(n+p+1, n+q+1)$. In the second stage we calculate, for each (p, q)

$$\tilde{\sigma}_{p,q}^2 = \inf \frac{1}{T} \sum_{t=m}^T \left(\sum_{j=0}^p \alpha_j y_{t-j} - \sum_{j=1}^q \beta_j \hat{\varepsilon}_{t-j} \right)^2 \text{ with } \alpha_0=1.$$

Then (\tilde{p}, \tilde{q}) are chosen to minimize

$$\log \tilde{\sigma}_{p,q}^2 + (p+q) \log T / T. \quad (2.1.9)$$

Once \tilde{p}, \tilde{q} are determined we may, at the third stage, use any of the available algorithms for the calculation of the maximum of the Gaussian likelihood, initiating the iterative calculation with the strongly consistent estimates, $\hat{\alpha}_j, \hat{\beta}_j$. The procedure should be iterated commencing from the produced regression coefficients $\tilde{\alpha}_j, \tilde{\beta}_j$ until the residual mean square stabilizes. They also showed that for \tilde{p}, \tilde{q} minimizing (2.1.9) and $\hat{\alpha}_j, \hat{\beta}_j, \tilde{\alpha}_j, \tilde{\beta}_j$, almost surely $\tilde{p} - p_0, \tilde{q} - q_0, \tilde{\alpha}_j - \alpha_{0,j}, \hat{\alpha}_j - \alpha_{0,j}, \hat{\beta}_j - \beta_{0,j} \rightarrow 0$ and also $T^{1/2}(\hat{\alpha}_j - \alpha_{0,j}), T^{1/2}(\hat{\beta}_j - \beta_{0,j})$ have the same asymptotic distribution as the maximum likelihood estimators, for p_0, q_0 known, on Gaussian assumptions.

2.1.2. Unit roots in multivariate time series

There are two approaches used to investigate unit roots in vector time series: one is the approach of Engle and Granger (1987) and the other is that of Fountis and Dickey (1989). The main difference between two approaches is that Engle and Granger (1987) look for evidence of a single cointegrating vector vs. the null hypothesis that all linear combinations of the vector entries are nonstationary while Fountis and Dickey's (1989) approach tests a null hypothesis of a single unit root canonical series. In addition Fountis and Dickey's (1989) distributional results for the test statistic rely on the assumption that there is exactly one unit root and the remaining parameters less than 1 in magnitude. The null hypothesis indicates a single unit root series mixed, by way of linear combinations, with several stationary series. We now introduce some ideas from Fountis and Dickey (1989) and we will discuss Engle and Granger (1987) in Chapter 5.

Let

$$\mathbf{Y}_t = \mathbf{A}\mathbf{Y}_{t-1} + \mathbf{e}_t \quad (2.1.10)$$

where \mathbf{e}_t is a dimension k multivariate normal column vector with mean ϕ and variance covariance matrix Σ . Assume that \mathbf{A} has one eigenvalue 1 and the rest less than 1 in magnitude. The right eigenvectors of \mathbf{A} and $\mathbf{A} - \mathbf{I}$ are identical, and there exists a matrix \mathbf{R} such that $\mathbf{R}^{-1}(\mathbf{A} - \mathbf{I})\mathbf{R}$ has 0 everywhere in the first row and column.

Let

$$\mathbf{V} = \mathbf{R}^{-1}\mathbf{A}\mathbf{R} = \begin{bmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{V}_{22} \end{bmatrix}$$

where all eigenvalues of \mathbf{V}_{22} are less than 1 in magnitude.

Regressing \mathbf{Y}_t on \mathbf{Y}_{t-1} produces the coefficient matrix

$$\hat{\mathbf{A}} = (\mathbf{Y}_{(0)} \mathbf{Y}'_{(-1)}) (\mathbf{Y}_{(-1)} \mathbf{Y}'_{(-1)})^{-1}$$

where $\mathbf{Y}_{(-1)} = (Y_0, Y_1, \dots, Y_{n-1})$, $\mathbf{Y}_{(0)} = (Y_1, Y_2, \dots, Y_n)$. Notice that $\hat{\mathbf{A}}$ is similar to $\mathbf{R}^{-1} \hat{\mathbf{A}} \mathbf{R}$, where

$$\mathbf{R}^{-1} \hat{\mathbf{A}} \mathbf{R} = [\mathbf{R}^{-1} \mathbf{Y}_{(0)} \mathbf{Y}'_{(-1)} (\mathbf{R}^{-1})'] [\mathbf{R}^{-1} \mathbf{Y}_{(-1)} \mathbf{Y}'_{(-1)} (\mathbf{R}^{-1})']^{-1},$$

and so the eigenvalues of $\hat{\mathbf{A}}$ are the same as those we would get if we could regress $\mathbf{R}^{-1} \mathbf{Y}_{(0)}$ on $\mathbf{R}^{-1} \mathbf{Y}_{(-1)}$. Thus if we can prove a distributional result for eigenvalues in this transformed regression it will also hold for the untransformed regression. Even though we do not know the true \mathbf{R} we can act as if we do for the purpose of studying distributional properties of the eigenvalues. Thus we assume

$$\begin{bmatrix} \mathbf{Z}_{1t} \\ \mathbf{Z}_{2t} \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{V}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{Z}_{1,t-1} \\ \mathbf{Z}_{2,t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{n}_{1t} \\ \mathbf{n}_{2t} \end{bmatrix} \quad (2.1.11)$$

is the true model on the transformed scale and estimate all estimates of the coefficient matrix (including the 0's). Let

$$\hat{\mathbf{V}} = \begin{bmatrix} \hat{\mathbf{V}}_{11} & \hat{\mathbf{V}}_{12} \\ \hat{\mathbf{V}}_{21} & \hat{\mathbf{V}}_{22} \end{bmatrix}$$

denote the estimate of the coefficient matrix as partitioned in (2.1.11). Each element of \hat{V}_{12} is $O_p(n^{-1})$ and each element of \hat{V}_{21} is $O_p(n^{-1/2})$. This implies that $n(\hat{V}_{11}-1)$ has the same limit distribution as that tabulated for $n(\hat{\rho}-1)$ in Fuller (1976).

The characteristic equation expanded on the first row is

$$|\hat{V} - m\mathbf{I}| = (\hat{V}_{11} - m) |\hat{V}_{22} - m\mathbf{I}| + O_p(n^{-3/2}) = 0,$$

and since the eigenvalues of \hat{V}_{22} are bounded away from 1, n large implies a solution $m = \hat{\lambda}_n$, which differs from \hat{V}_{11} by $O_p(n^{-3/2})$. Thus

$$n(\hat{\lambda}_n - 1) = n(\hat{V}_{11} - 1) + O_p(n^{-1/2}),$$

showing that the distribution of $\hat{\lambda}_n$ is also governed in the limit by the distribution for the Dickey-Fuller tests.

This procedure easily extends to higher-order cases. Let

$$\mathbf{Y}_t = \mathbf{A}_1 \mathbf{Y}_{t-1} + \mathbf{A}_2 \mathbf{Y}_{t-2} + \dots + \mathbf{A}_p \mathbf{Y}_{t-p} + \varepsilon_t, \quad t = 1, 2, \dots, n,$$

where \mathbf{Y}_t is a dimension k column vector. Reparameterize this as

$$\nabla \mathbf{Y}_t = -(\mathbf{I} - \mathbf{A}_1 - \dots - \mathbf{A}_p) \mathbf{Y}_{t-1} - (\mathbf{A}_2 + \mathbf{A}_3 + \dots + \mathbf{A}_p) \nabla \mathbf{Y}_{t-1} - \dots - \mathbf{A}_p \nabla \mathbf{Y}_{t-p+1} + \varepsilon_t,$$

where $\nabla \mathbf{Y}_{t-p} = (\mathbf{Y}_t - \mathbf{Y}_{t-p})$ and note that if the characteristic equation,

$$|L^p \mathbf{I} - L^{p-1} \hat{\mathbf{A}}_1 - \dots - \hat{\mathbf{A}}_p| = 0,$$

has a unique root $L = 1$ then the coefficient matrix on Y_{t-1} is of rank $k-1$. If estimate $\hat{\mathbf{A}}_i$ of \mathbf{A}_i are calculated by least squares method from n observations, then the solution $L = \hat{\lambda}_n$ of $|L^p \mathbf{I} - L^{p-1} \hat{\mathbf{A}}_1 - \dots - \hat{\mathbf{A}}_p| = 0$ closest in magnitude to 1 provides a test since, under the null hypothesis of a single unit root, $n(|\hat{\lambda}_n| - 1)$ has the same limit distribution as $n(\hat{\rho} - 1)$ given by Dickey and Fuller (1979). Computing the eigenvectors of $(\mathbf{I} - \hat{\mathbf{A}}_1 - \dots - \hat{\mathbf{A}}_p)$ associated with $\hat{\lambda}_n$, we can construct \mathbf{R} with the right eigenvector as its first column and the left eigenvector as the first row of \mathbf{R}^{-1} .

The following papers may provide the background to develop the unit roots in a vector process (Chapter 3).

Reinsel et al. (1992) expressed the explicit form of the gradient vector and Hessian matrix of the log likelihood function for the multivariate ARMA model and an explicit description of the Gauss-Newton iterative procedure to get ML estimates of the parameters. Their resulting computational procedure has the form of a generalized least squares (GLS) estimation involving lagged values of the observed vector series and of the residual series as independent variables. This explicit representation of the estimation procedure provides not only a convenient and appealing computational scheme relative to the general nonlinear maximization algorithms which are commonly used, but also proves useful in understanding and interpreting properties of the ML estimators. Their presentation also allows for easy comparison of the ML procedure with the GLS

procedure of Koreisha and Pukkila (1989) in the univariate case. Furthermore, their approach displayed the stage three estimation procedure by Hannan and Rissanen (1982) and Hannan and Kavalieris (1984a) in the univariate ARMA model context, and the stage four procedure by Hannan and Kavalieris (1984b) for the vector ARMA model as corresponding to the ML Newton-Raphson iteration procedure as represented in the GLS form. Consider estimation of parameters in the multivariate ARMA model

$$\mathbf{Y}_t - \sum_{i=1}^p \Phi_i \mathbf{Y}_{t-i} = \boldsymbol{\varepsilon}_t - \sum_{i=1}^q \Theta_i \boldsymbol{\varepsilon}_{t-i}$$

where $\mathbf{Y}_t = (y_{1t}, \dots, y_{kt})'$, $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})'$ are k -component vectors, and $\Phi_1, \dots, \Phi_p, \Theta_1, \dots, \Theta_q$ are $k \times k$ matrices of unknown parameters. The $\boldsymbol{\varepsilon}_t$ are assumed to be iid with mean vector zero and nonsingular covariance matrix Σ . All roots of $\det[\Phi(z)] = 0$ and $\det[\Theta(z)] = 0$ are assumed to lie outside the unit circle, and the necessary conditions for the identification of parameters for the above model are satisfied. On the assumption of normality of the $\boldsymbol{\varepsilon}_t$, since $\mathbf{e} = \text{vec}(\boldsymbol{\varepsilon}')$ is $N(\mathbf{0}, \mathbf{I}_n \otimes \Sigma)$, the approximate log likelihood function can be written as

$$\begin{aligned} l &= -\frac{n}{2} \log|\Sigma| - \frac{1}{2} \mathbf{e}' (\mathbf{I}_n \otimes \Sigma^{-1}) \mathbf{e} \\ &= -\frac{n}{2} \log|\Sigma| - \frac{1}{2} \mathbf{w}' \Theta^{-1} (\mathbf{I}_n \otimes \Sigma^{-1}) \Theta^{-1} \mathbf{w} \end{aligned}$$

where $\mathbf{w} = \mathbf{y} - \sum_{i=1}^p (\mathbf{L}^i \mathbf{Y} \otimes \mathbf{I}_k) \boldsymbol{\phi}_i$ with $\mathbf{L}^i \mathbf{Y} = (Y_{1-i}, \dots, Y_{n-i})'$ and $\mathbf{y} = \text{vec}(\mathbf{Y}') = (Y_1', \dots, Y_n')$.

Defining the vector $\boldsymbol{\delta} = (\boldsymbol{\phi}'_1, \dots, \boldsymbol{\phi}'_p, \boldsymbol{\theta}'_1, \dots, \boldsymbol{\theta}'_q)'$ with $\boldsymbol{\phi}_i = \text{vec}(\Phi_i)$ and $\boldsymbol{\theta}_i = \text{vec}(\Theta_i)$, $i =$

$1, \dots, q$ and matrix $\mathbf{Z} = [(\mathbf{L}\mathbf{Y} \otimes \mathbf{I}_k), \dots, (\mathbf{L}^p \mathbf{Y} \otimes \mathbf{I}_k), -(\mathbf{L}\boldsymbol{\varepsilon} \otimes \mathbf{I}_k), \dots, -(\mathbf{L}^q \boldsymbol{\varepsilon} \otimes \mathbf{I}_k)]$, we can express these derivatives collectively in a convenient form as

$$\partial l / \partial \boldsymbol{\delta} = \mathbf{Z}' \boldsymbol{\Theta}'^{-1} (\mathbf{I}_n \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{e} \quad (2.1.12)$$

or

$$\partial l / \partial \boldsymbol{\delta} = (-\partial \mathbf{e}' / \partial \boldsymbol{\delta}) (\mathbf{I}_n \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{e} \quad \text{with } \partial \mathbf{e}' / \partial \boldsymbol{\delta}' = -\boldsymbol{\Theta}^{-1} \mathbf{Z}.$$

These likelihood equations (2.1.12) are highly nonlinear in the parameters $\boldsymbol{\delta}$ (unless $q = 0$). Thus for $q > 0$ these equations need to be solved by iterative numerical procedures such as the Newton-Raphson equations for an approximate ML estimator $\hat{\boldsymbol{\delta}}$ are

$$-\left(\frac{\partial^2 l}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}'} \right) \Big|_{\boldsymbol{\delta} = \hat{\boldsymbol{\delta}}_0} (\hat{\boldsymbol{\delta}} - \hat{\boldsymbol{\delta}}_0) = \left(\frac{\partial l}{\partial \boldsymbol{\delta}} \right) \Big|_{\boldsymbol{\delta} = \hat{\boldsymbol{\delta}}_0} \quad (2.1.13)$$

where $\hat{\boldsymbol{\delta}}_0$ is an initial estimate of $\boldsymbol{\delta}$ and the estimate $\tilde{\boldsymbol{\Sigma}} = \hat{\boldsymbol{\varepsilon}}' \hat{\boldsymbol{\varepsilon}} / n$ from a previous iteration is used for $\boldsymbol{\Sigma}$. To carry out the iteration in (2.1.13) it is useful to have a convenient expression for the Hessian matrix of second partial derivatives. It can be shown that on neglecting terms which, when divided by n , converge to zero in probability as $n \rightarrow \infty$, we obtain the approximation

$$\begin{aligned} -\left(\frac{\partial^2 l}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}'} \right) &= \left(\frac{\partial \mathbf{e}'}{\partial \boldsymbol{\delta}} \right) (\mathbf{I}_n \otimes \boldsymbol{\Sigma}^{-1}) \left(\frac{\partial \mathbf{e}}{\partial \boldsymbol{\delta}'} \right) \\ &= \mathbf{Z}' \boldsymbol{\Theta}'^{-1} (\mathbf{I}_n \otimes \boldsymbol{\Sigma}^{-1}) \boldsymbol{\Theta}^{-1} \mathbf{Z} \end{aligned} \quad (2.1.14)$$

Using initial estimates $\delta_o = (\tilde{\phi}'_1, \dots, \tilde{\phi}'_p, \tilde{\phi}'_1, \dots, \tilde{\phi}'_q)' = \text{vec}(\tilde{\Phi}_1, \dots, \tilde{\Phi}_p, \tilde{\Theta}_1, \dots, \tilde{\Theta}_q)$, we let

$$\tilde{\Theta} = \mathbf{I} - \sum_{i=1}^q (\mathbf{L}^i \otimes \tilde{\Theta}_i), \quad \tilde{\mathbf{e}} = \tilde{\Theta}^{-1} [\mathbf{y} - \sum_{i=1}^p (\mathbf{L}^i \mathbf{Y} \otimes \mathbf{I}_k) \tilde{\phi}_i]$$
 with $\text{vec}(\tilde{\mathbf{e}}') = \tilde{\mathbf{e}}$, $\tilde{\Sigma} = \tilde{\mathbf{e}}' \tilde{\mathbf{e}} / n$ and let

$\tilde{\mathbf{Z}}$ denote the matrix \mathbf{Z} with $\mathbf{L}^j \tilde{\mathbf{e}}$ in place of $\mathbf{L}^j \mathbf{\varepsilon}$, $j = 1, \dots, q$. Then, using (2.1.12)-(2.1.14),

the modified Newton-Raphson equations for $\hat{\delta}$ have the solution of the form

$$\hat{\delta} = \delta_o + \{\bar{\mathbf{Z}}' (\mathbf{I}_n \otimes \tilde{\Sigma}^{-1}) \bar{\mathbf{Z}}\}^{-1} \bar{\mathbf{Z}}' (\mathbf{I}_n \otimes \tilde{\Sigma}^{-1}) \tilde{\mathbf{e}} \quad (2.1.15)$$

where $\bar{\mathbf{Z}} = \tilde{\Theta}^{-1} \tilde{\mathbf{Z}}$. Note that $\tilde{\mathbf{e}} = \text{vec}(\tilde{\mathbf{e}}')$ is easily computed recursively from

$$\tilde{\Theta} \tilde{\mathbf{e}} = \mathbf{y} - \sum_{i=1}^p (\mathbf{L}^i \mathbf{Y} \otimes \mathbf{I}_k) \tilde{\phi}_i = \mathbf{y} - \sum_{i=1}^p (\mathbf{I}_n \otimes \tilde{\Phi}_i) \mathbf{L}^i \mathbf{y}$$

as

$$\tilde{\mathbf{e}}_t = \mathbf{Y}_t - \sum_{i=1}^p \tilde{\Phi}_i \mathbf{Y}_{t-i} + \sum_{i=1}^q \tilde{\Theta}_i \tilde{\mathbf{e}}_{t-i}, \quad t = 1, \dots, n$$

with $\tilde{\mathbf{e}}_0 = \tilde{\mathbf{e}}_{-1} = \dots = \tilde{\mathbf{e}}_{1-q} = 0$. In addition the ML-GLS estimation procedure provides asymptotically efficient estimators and it involves the same degree of GLS computations for the univariate model as are involved in the univariate GLS estimation procedure of Koreisha and Pukkila (1989). The ML estimation through the GLS estimation in (2.1.15) corresponds essentially to the procedure stated by Hannan and Kavalieris (1984b) as their Stage 4 procedure for estimation of vector ARMA with exogenous variables (ARMAX) models. The ML-GLS procedure provided very good estimates after only a few iterations

using reasonable initial estimates, and its performance was not dependent on the exact nature of the initial estimates used.

Osborn (1977) worked exact and approximate ML estimators for vector MA processes, which is a generalization of that adopted by Box and Jenkins (1976) for the univariate case. Let a q th vector MA process be

$$\mathbf{W}_t = \boldsymbol{\varepsilon}_t + \theta_1 \boldsymbol{\varepsilon}_{t-1} + \dots + \theta_q \boldsymbol{\varepsilon}_{t-q} \quad (2.1.16)$$

where \mathbf{W}_t and $\boldsymbol{\varepsilon}_t$ are $g \times 1$ vectors of observations and disturbances respectively, while θ_i , $i = 1, \dots, q$, are $g \times g$ matrices of MA coefficients. The disturbances are assumed to be independent normally distributed about zero with variance covariance matrix Σ . With n sample observations assumed to be available on \mathbf{W}_t , we denote the vector of disturbances in (2.1.16) by $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}'_{1-q}, \dots, \boldsymbol{\varepsilon}'_0, \boldsymbol{\varepsilon}'_1, \dots, \boldsymbol{\varepsilon}'_n)'$. This vector is normally distributed with zero mean vector and variance-covariance matrix $\Omega = \mathbf{I} \otimes \Sigma$. Then the exact likelihood function is

$$f(\mathbf{W}) = (2\pi)^{-\frac{gn}{2}} |\mathbf{W}|^{-\frac{1}{2}} |\mathbf{K}'\mathbf{W}^{-1}\mathbf{K}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{M}\mathbf{W} + \mathbf{K}\hat{\boldsymbol{\varepsilon}}^*)'\Omega^{-1}(\mathbf{M}\mathbf{W} + \mathbf{K}\hat{\boldsymbol{\varepsilon}}^*)\right) \quad (2.1.17)$$

where $\hat{\boldsymbol{\varepsilon}}^* = (\boldsymbol{\varepsilon}'_{1-q}, \dots, \boldsymbol{\varepsilon}'_0)'$, $\mathbf{W} = (\mathbf{W}'_1, \dots, \mathbf{W}'_n)'$, $\boldsymbol{\varepsilon} = \mathbf{M}\mathbf{W} + \mathbf{K}\hat{\boldsymbol{\varepsilon}}^*$. By using the following two facts,

$$\left(-\frac{1}{2}(\mathbf{M}\mathbf{W} + \mathbf{K}\hat{\boldsymbol{\varepsilon}}^*)'\Omega^{-1}(\mathbf{M}\mathbf{W} + \mathbf{K}\hat{\boldsymbol{\varepsilon}}^*)\right) = \left(-\frac{1}{2}\hat{\boldsymbol{\varepsilon}}^*\Omega^{-1}\hat{\boldsymbol{\varepsilon}}^*\right)$$

and

$$|\Omega|^{-\frac{1}{2}} |\mathbf{K}'\Omega^{-1}\mathbf{K}|^{-\frac{1}{2}} = |\Sigma|^{-\frac{1}{2}(n+q)} |\mathbf{K}'(\mathbf{I} \otimes \Sigma^{-1})\mathbf{K}|^{-\frac{1}{2}},$$

equation (2.1.17) can be replaced by the following function L_1^*

$$L_1^* = -(n+q)\log|\Sigma| - \log|\mathbf{K}'(\mathbf{I} \otimes \Sigma^{-1})\mathbf{K}| - \sum_{t=1-q}^n \hat{\boldsymbol{\varepsilon}}_t' \Sigma^{-1} \hat{\boldsymbol{\varepsilon}}_t.$$

Therefore, exact ML estimates of Σ and θ_i , $i = 1, \dots, q$, may be achieved by maximizing the function L_1^* with respect to these parameters using an appropriate nonlinear optimization technique. Although this exact procedure may be preferred on the theoretical grounds, the computational burden it imposes may make it impractical in some circumstances.

Spliid (1983) suggested a very fast and simple algorithm for estimation of the parameters of large multivariate time series which includes distributed lag variables and showed that the distribution of the estimates are asymptotically normal and unbiased. Furthermore, the algorithm is applicable for estimation of large multivariate models where it is generally many times faster than maximization algorithms. In the model identification stage of the analysis of multivariate time series it is often desirable to estimate a number of alternative models and for this purpose this algorithm is very well suited. Consider a general linear model that includes lagged regression variables is denoted as

$$\phi(L)\mathbf{y}_t = \beta(L)\mathbf{x}_t + \theta(L)\mathbf{a}_t, \quad t = 1, 2, \dots \quad (2.1.18)$$

where $\beta(L) = \beta_0 + \beta_1 L + \dots + \beta_{r-1} L^{r-1}$ is a $k \times m$ matrix polynomial and \mathbf{x}_t is an m -variate series disturbances and (2.1.18) can be rewritten in the equivalence forms:

$$\mathbf{Y} = \mathbf{U}\delta + \mathbf{a} \quad (2.1.19)$$

where $\mathbf{U} = (-\mathbf{A}, \mathbf{Y}, \mathbf{X})$, $\mathbf{A} = (L\mathbf{a}, L^2\mathbf{a}, \dots, L^q\mathbf{a})$, $\mathbf{Y} = (L\mathbf{y}, L^2\mathbf{y}, \dots, L^p\mathbf{y})$, and $\mathbf{X} = (\mathbf{x}, L\mathbf{x}, \dots, L^{r-1}\mathbf{x})$ with the matrix of unknown residuals \mathbf{a} . Introducing the four stage procedures based on the above models (2.1.19), we can take the following steps:

Step 1. Construct $\mathbf{W} = (L\mathbf{y}, L^2\mathbf{y}, \dots, L^s\mathbf{y}, \mathbf{X})$ where s is a chosen order of the initial AR part of the model. Linear regression gives $\hat{\mathbf{a}}_0 = \mathbf{y} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{y}$ where $\hat{\mathbf{a}}_0$ denotes the estimate of \mathbf{a} obtained in iteration number 0.

Step 2. Compute recursively residuals for $t = 1, 2, \dots, n$:

$$\hat{\mathbf{a}}_{t,j} = \mathbf{y}_t - \sum_{i=1}^p \hat{\phi}_{i,j} \mathbf{y}_{t-i} + \sum_{i=1}^q \hat{\theta}_{i,j} \hat{\mathbf{a}}_{t-i,j} - \sum_{i=0}^{r-1} \hat{\beta}_{i,j} \mathbf{x}_{t-i}.$$

Step 3. Construct $\hat{\mathbf{A}}_j$, recall \mathbf{Y} and \mathbf{X} , create $\hat{\mathbf{U}}_j$, and compute new estimates by linear regression, by solving $\hat{\mathbf{U}}_j' \hat{\mathbf{U}}_j \hat{\delta}_{j+1} = \hat{\mathbf{U}}_j' \mathbf{y}$.

Step 4. If $\hat{\delta}_{j+1} \neq \hat{\delta}_j$, increase j by 1 and repeat steps 2 through 4. If $\hat{\delta}_{j+1} = \hat{\delta}_j$, let $\hat{\delta} = \hat{\delta}_{j+1}$ and stop. The algorithm stops if and only if the empirical autocorrelations and cross-correlations are zero ($\hat{\mathbf{U}}_j' \hat{\mathbf{a}}_j = 0$).

Koreisha and Pukkila (1989) considered explicit computation of the maximum likelihood estimation for the unrestricted multivariate model and presented three linear preliminary estimation method which, when used in conjunction with the exact ML

(Hillmer and Tiao, 1979), can reduce parameter estimation by more than an order of magnitude. Furthermore, because of the great speed and accuracy associated with parameter value calculations they can also be employed in process order identification, and these techniques are useful in the early model building-stage since they permit analysts to compare the characteristics of various competing models inexpensively.

2.2. Previous results on measurement error model

While the functional model, which has a fixed x_t in the time series, considered in the Section 2.1, the structural model with stochastic x_t in the time series can be applied to the following:

$$y_t = x_t + u_t \quad (2.2.1)$$

where u_t is a measurement error. Estimation of the true value x_t is very important in signal measurement problems encountered in the engineering areas where x_t is known as 'signal' and u_t as 'noise'. In addition, it is important in survey sampling in the situation where a population is sampled at a sequence of time points according to a sampling scheme and in macroeconomic survey data such as the Panel Study of Income Dynamics (PSID) and the U. S. Current Population Survey (CPS). We consider the measurement error model with the ARMA(p, q) process included. Some results on estimating the univariate version of model (2.2.1) available in the following literature.

The problem of estimating the AR process parameters based on the data corrupted by the unknown white noise has been studied by many authors. This problem was first treated by Walker (1960). Later, Parzen (1967) proposed some methods to estimate process parameters (PP) and signal-to-noise ratio (SNR) by means of spectral density and third order correlation. Kashyap (1970) worked the ML method for more general vector ARMA model with additional white noise.

Pagano (1974) suggested a two-stage nonlinear LS procedure to obtain consistent and efficient estimators with the situation where x_t is AR(p) and u_t is white noise:

$$y_t = x_t + u_t$$

$$\sum_{j=0}^p \phi_j x_{t-j} = e_t \quad (2.2.2)$$

where $\{x_t\}$ and $\{u_t\}$ are independent, $\{u_t\}$ and $\{e_t\}$ are iid with $N(0, \sigma_u^2)$, $N(0, \sigma_e^2)$, respectively, and $\phi_0 = 1$, $\phi_p \neq 0$. His basic idea is to introduce the first $(p+1)$ new parameters and obtain initial estimates of these new parameters and original $(p+2)$ parameters on the basis that the model (2.1.1) and (2.2.2) are then transformed into a mixed ARMA model. Having obtained these initial estimates, we can get better estimates by expressing these parameters in terms of the original parameters. Therefore, model (2.2.1) and (2.2.2) are rewritten (Pagano, 1974, p. 100), as a mixed ARMA(p, q),

$$\sum_{j=0}^p \phi_j x_{t-j} = \sum_{i=0}^q \gamma_i e_{t-i} \quad (2.2.3)$$

with $\gamma_0 = 1$ and $\gamma_q \neq 0$, $q \geq 1$. In this reparameterization, we have gone from the original $p+q+2$ parameters to $p+\max(p,q)+1$ parameters. Furthermore, Pagano presented the nonlinear regression method which is asymptotically consistent with the ML method and proved its properties (Theorem 4, p. 107).

Sakai and Arase (1979) considered a modified LS procedures for the problem of estimating parameters of an AR process based on the data corrupted by unknown white noise. Assume that the time series $\{x_t\}$ is a zero mean Gaussian p -th order AR process, (2.2.2) where $\{e_t\}$ is a sequence of white noise with $E(e_t) = 0$ and $E(e_t e_s) = \sigma_e^2 \delta_{t,s}$. Denote

the noisy observation sequence by $\{y_t\}$ with (2.2.1) where $\{u_t\}$ is a Gaussian white noise sequence uncorrelated with $\{x_t\}$ and σ_e^2 is known. There are several steps to develop the modified LS method so as to estimate $\phi_1, \dots, \phi_p, \sigma_e^2$ and σ_μ^2 recursively, when we have n successive samples y_1, \dots, y_n . Firstly, from (2.2.1) and (2.2.2), we get

$$\sum_{i=1}^p \phi_i y_{t-i} = e_t - \sum_{j=0}^p \phi_j \mu_{t-j}, \quad \phi_0 = 0, \quad (2.2.4)$$

or

$$y_t - \mathbf{Y}'_{t-1} \Phi = v_t \quad (2.2.5)$$

where $\mathbf{Y}_{t-1} = (y_{t-1}, \dots, y_{t-p})'$, $\Phi = (\phi_1, \dots, \phi_p)'$. $(\sum_{t=1}^n \mathbf{Y}_{t-1} \mathbf{Y}'_{t-1}) \hat{\mathbf{a}}_n = \sum_{t=1}^n \mathbf{Y}_{t-1} y_t$ is obtained by

applying the LS method to (2.2.5) where

$$\hat{\mathbf{a}}_n = (\hat{\alpha}_n^1, \dots, \hat{\alpha}_n^p)' = (\sum_{t=1}^n \mathbf{Y}_{t-1} \mathbf{Y}'_{t-1})^{-1} (\sum_{t=1}^n \mathbf{Y}_{t-1} y_t) \quad (2.2.6)$$

and $\hat{\mathbf{a}}_n$ is the LS estimate for Φ . Since the LS estimator is asymptotically biased, we may

apply their recursive modified LS method to this situation, the estimator $\hat{\Phi}_n$ is defined by

$$\hat{\Phi}_n = \hat{\mathbf{a}}_n + n \mathbf{P}_n \sigma_n^2 \hat{\Phi}_{n-1} \quad (2.2.7)$$

with $\mathbf{P}_n = (\sum \mathbf{Y}_{t-1} \mathbf{Y}'_{t-1})^{-1}$, $\hat{\Phi}_n = (\hat{\phi}_n^1, \dots, \hat{\phi}_n^p)'$. $\hat{\Phi}_n$ in (2.2.7) converges to Φ under the assumption that σ_n^2 is known. But in many practical cases it is necessary to estimate it since σ_n^2 is not known. Therefore, (2.2.7) can be modified to

$$\hat{\Phi}_n = \hat{\mathbf{a}}_n + n \mathbf{P}_n \hat{\sigma}_n^2 \hat{\Phi}_{n-1} \quad (2.2.8)$$

Moreover, the recursive formulae for \mathbf{P}_n , $\hat{\mathbf{a}}_n$ and $\sum_{t=1}^n \hat{\xi}_{t,n}^2$ where the residual at time n is $\hat{\xi}_{t,n} = y_t - \mathbf{Y}'_{t-1} \hat{\mathbf{a}}_n$ are given by (Sakai and Arase, 1979, p. 952). On the contrary, in the case that σ_e^2 is known, we estimate $\hat{\sigma}_n^2$, $\hat{\Phi}_n$, and $(\hat{\sigma}_e^2)_n$ as in (Sakai and Arase, 1979, p. 953). Finally, they showed that modified LS method dominates the Yuler-Walker and the bootstrap method for the problem of estimating AR process parameters based on the data corrupted by unknown white noise from the simulation results.

Dunsmuir (1979) investigated central limit theorem for parameter estimation in stationary vector time series and discussed its application to models for a stationary signal observed with noise. Furthermore, he derived the strong consistency and the limiting distribution of estimators by maximizing a frequency domain approximation of the likelihood function.

Binder and Hidioglou (1987) considered model $y_t = x_t + u_t$ where x_t is a ARMA(p, q) and u_t is a ARMA(p^*, q^*) and formulated it as a state space model and exploited an iterative numerical procedure to compute the MLE. The exact log likelihood function may be written as:

$$L = -\frac{1}{2} \sum_{t=1}^n \log |\mathbf{W}_{t|t-1}| - \frac{1}{2} \sum_{t=1}^n (\mathbf{Y}_t - \hat{\mathbf{Y}}_{t|t-1})' \mathbf{W}_{t|t-1}^{-1} (\mathbf{Y}_t - \hat{\mathbf{Y}}_{t|t-1}) \quad (2.2.9)$$

where

$$\mathbf{W}_{t|t-1} = \text{Var}(\mathbf{y}_t | \mathbf{Y}_i) = \mathbf{A}_t \mathbf{P}_{t|t-1} \mathbf{A}_t', \quad \hat{\mathbf{y}}_{t|t-1} = \mathbb{E}(\mathbf{y}_t | \mathbf{Y}_i) = \mathbf{x}_t \boldsymbol{\mu}_t + \mathbf{A}_t \hat{\mathbf{z}}_{t|t-1} \quad (2.2.10)$$

based on the given model $\mathbf{z}_0 \sim N(\mathbf{0}, \boldsymbol{\Sigma}_0)$, $\mathbf{z}_t = \mathbf{F}_t \mathbf{z}_{t-1} + \mathbf{G}_t \boldsymbol{\varepsilon}_t$ where $\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \mathbf{V}_t)$ and $\{\mathbf{z}_t\}$ is assumed to be a series of r -dimensional vectors. The model formulation continues with

$$\boldsymbol{\theta}_t = \boldsymbol{\mu}_t + \mathbf{h}' \mathbf{z}_t, \quad \mathbf{e}_t = \mathbf{k}_t \mathbf{z}_t, \quad \text{and} \quad \mathbf{Y}_t = \mathbf{x}_t \boldsymbol{\theta}_t + \mathbf{e}_t = \mathbf{x}_t \mathbf{u}_t + \mathbf{A}_t \mathbf{z}_t$$

where $\mathbf{A}_t = \mathbf{x}_t \mathbf{h}' + \mathbf{k}_t$. Maximizing this log likelihood function (2.2.9) with respect to $\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q, \sigma^2$ and $\gamma_1, \dots, \gamma_s$ (when $\mathbf{u}_t = \mathbf{c}_t' \boldsymbol{\gamma}$ for known \mathbf{c}_t) based on ARMA(p, q) process involves finding the first and second derivatives of (2.2.9) with respect to these parameters. Therefore, it is sufficient to get the first and second derivations of $\hat{\mathbf{y}}_{t|t-1}$ and $\mathbf{W}_{t|t-1}$ with respect to the known parameters $\{\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q, \sigma^2, \gamma_1, \dots, \gamma_s\}$. These can be obtained recursively by using the following along with (2.2.10):

$$\begin{aligned} \hat{\mathbf{z}}_{1|0} &= \mathbf{0}, \\ \mathbf{P}_{1|0} &= \mathbf{F}_1 \boldsymbol{\Sigma}_0 \mathbf{F}_1' + \mathbf{G}_1 \mathbf{V}_1 \mathbf{G}_1', \\ \hat{\mathbf{z}}_{t|t} &= \hat{\mathbf{z}}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{A}_t' \mathbf{W}_{t|t-1}^{-1} (\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}), \\ \hat{\mathbf{z}}_{t|t+1} &= \mathbf{F}_{t+1} \hat{\mathbf{z}}_{t|t}, \\ \mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{A}_t' \mathbf{W}_{t|t-1}^{-1} \mathbf{A}_t \mathbf{P}_{t|t-1}, \\ \mathbf{P}_{t+1|t} &= \mathbf{F}_{t+1} \mathbf{P}_{t|t} \mathbf{F}_{t+1}' + \mathbf{G}_{t+1} \mathbf{V}_{t+1} \mathbf{G}_{t+1}'. \end{aligned}$$

CHAPTER III

TESTING FOR A UNIT ROOT IN AN AR(p) SIGNAL WITH MA(q) NOISE

Section 3.1 describes previous works in testing for a unit root with parametric constraints and general restricted model is introduced in Section 3.2. In Section 3.3 we discuss the asymptotic distributions both of parameter estimators and of the unit root test statistics under the null hypothesis.

3.1. Previous works in testing for a unit root with parametric constraints

Shin (1993) considered ML estimation for stationary AR processes when the signal is subject to a MA noise, and also this ML estimator of the parameter vector was shown to be strongly consistent and to have a multivariate normal limiting distribution.

The model which he suggested is

$$y_t = x_t + u_t \quad (3.1.0)$$

$$x_t = \rho x_{t-1} + z_t \quad (3.1.1)$$

$$\sum_{i=0}^p \alpha_i z_{t-i} = \varepsilon_t, \quad \alpha_0 = 1$$

$$u_t = \sum_{j=0}^q \beta_j w_{t-j}, \quad \beta_0 = 1,$$

for $t = 1, \dots, n$, where p and q are known nonnegative integers and n is the sample size. In addition $\{y_t\}$ is a sequence of observations and $\{\varepsilon_t\}$ and $\{w_t\}$ are iid $N(0, \sigma_\varepsilon^2)$ and $N(0, \sigma_w^2)$ sequences respectively. The process $\{x_t\}$ is assumed to be stationary and $\alpha_p \neq 0$ and $\beta = (\beta_1, \dots, \beta_q)'$ and σ_w^2 are known. From (3.1.0) - (3.1.1) the following is obtained as

$$\sum_{j=0}^p \alpha_j (y_{t-j} - \rho y_{t-1-j}) = \sum_{j=0}^{p+1} (\alpha_j - \rho \alpha_{j-1}) u_{t-j} + \varepsilon_t \quad (3.1.2)$$

where $\alpha_0 = 1$ and $\alpha_{-1} = \alpha_{p+1} = 0$. Because the autocovariance function of the time series defined by the right hand side of (3.1.2) is zero for lags greater than $(p+q+1)$ it can be written as an MA($p+q+1$) process:

$$\sum_{j=0}^p \alpha_j (y_{t-j} - \rho y_{t-1-j}) = \sum_{j=0}^{p+q+1} \gamma_j e_{t-j}, \quad (3.1.3)$$

where $\{e_t\}$ is a sequence of iid $N(0, \sigma^2)$ random variables, $\sigma^2 > 0$ and $\gamma_0 = 1$. Let $\theta = (\alpha_1, \dots, \alpha_p, \gamma_1, \dots, \gamma_{p+q})'$, $\xi = (\theta', \sigma^2)'$, $\mathbf{Y} = (y_1, \dots, y_n)'$ and $\mathbf{X} = (x_1, \dots, x_n)'$. Then the negative log likelihood function is

$$L_n(\xi) = \mathbf{Y}'\text{Var}(\mathbf{Y})^{-1}\mathbf{Y} + \log[\det(\text{Var}(\mathbf{Y})/\sigma^2)] + n\log\sigma^2. \quad (3.1.4)$$

The ML estimator is defined by minimizing $L_n(\xi)$ with respect to ξ under the following restriction on ξ , for $h = 0, 1, \dots, p+q$,

$$\sigma^2 \sum_{j=0}^{p+q-h} \gamma_j \gamma_{j+h} - \sigma_e^2 \delta_{0h} - \sum_{i=0}^p \sum_{j=0}^p \alpha_i \alpha_j \sum_{s=0}^{q-|i-j-h|} \beta_s \beta_{s+|i-j-h|} = 0 \quad (3.1.5)$$

where $\gamma_0 = \alpha_0 = 1$ and δ_{0h} is the Kronecker delta. The modified ML estimator is defined by minimizing

$$L_n^M(\xi) = \sum \mathbf{e}_t^2(\mathbf{Y}, \theta)/\sigma^2 + n\log\sigma_e^2 \quad (3.1.6)$$

under the above restriction of (3.1.5) and it has the same limiting distribution as ML estimator based on (3.1.4) and is also strongly consistent. He showed that this modified ML estimator is much better than that of Said and Dickey (1984) in the sense that in general empirical powers are bigger and empirical sizes are much closer to the nominal levels through numerical results.

Miazaki and Dorea (1993) considered univariate ARMA model with nonlinear restrictions for analyzing a single time series data set from rotational sampling. If the sampling at regular time intervals is designed such that samples drawn on different occasions partially overlap, we can have a rotational sampling scheme. Under a rotational sampling scheme, they extended the ideas of Pagano (1974) in the white noise u_t case to

the MA(q) u_t case. Consider a signal detection model in which the signal x_t is a p th order AR process:

$$y_t = x_t + u_t, \quad t = 1, \dots, n,$$

$$\sum_{i=0}^p \alpha_i x_{t-i} = e_t, \quad \alpha_0 = 1 \quad (3.1.7)$$

$$u_t = \sum_{j=0}^q b_j v_{t-j}, \quad b_0 = 1,$$

where $(e_t, v_t)'$ is a sequence of independent $N(\mathbf{0}, \Sigma)$ random vectors with $\Sigma = \text{diag}(\sigma_e^2, \sigma_v^2)$. It is assumed that the roots of $\sum_{i=0}^p \alpha_i L^i = 0$ and of $\sum_{b=0}^q b_j L^j = 0$ are greater than one in absolute value. Then by Pagano's (1974) approach the model (3.1.7) can be written as an ARMA process,

$$\sum_{i=0}^p \alpha_i y_{t-i} = \sum_{j=0}^{p+q} \beta_j z_{t-j}, \quad \alpha_0 = \beta_0 = 1, \quad (3.1.8)$$

with restrictions

$$\sigma_e^2 + \sum_{j=0}^p \sum_{k=0}^p \alpha_j \alpha_k \gamma_u(k-j) - \sigma_z^2 \sum_{j=0}^{p+q} \beta_j^2 = 0,$$

$$\sum_{j=0}^p \sum_{k=0}^p \alpha_j \alpha_k \gamma_u(k-j+h) - \sigma_z^2 \sum_{j=0}^{p+q-h} \beta_j \beta_{j+h} = 0, \quad h = 1, \dots, p+q,$$

where $\gamma_\mu(h) = E(z_t z_{t+h})$. The new parameters $\beta = (\beta_1, \dots, \beta_{p+q})'$ and $\sigma_z^2 = V(z_t)$ are functions of $\alpha = (\alpha_1, \dots, \alpha_p)'$ and σ_e^2 from (3.1.7). Let $\theta = (\delta', \sigma_e^2)'$, $\delta = (\gamma', \sigma_z^2)'$, $\gamma = (\alpha'$,

$\beta')'$ and $\pi = (\alpha', \sigma_e^2)'$. Then we can develop a procedure to construct estimate of α and σ_e^2 as follows.

Step 1. Consider the restricted model (3.1.8) and the problem of estimating δ . Assume one has an initial estimator γ^0 of γ . Let

$$z_t(\mathbf{Y}, \gamma) = Y_t + \sum_{i=1}^p \gamma_i Y_{t-i} - \sum_{j=1}^{p+q} \gamma_{p+j} Z_{t-j}.$$

Expanding $z_t(\mathbf{Y}, \gamma)$ in a Taylor series about γ^0 and regressing on the vector of the partial derivatives of $z_t(\mathbf{Y}, \gamma)$ with respect to γ , $\mathbf{u}_t(\mathbf{Y}, \gamma)$, we obtain the one-step Gauss-Newton estimator of γ :

$$\tilde{\gamma} = \gamma^0 - \left(\sum_{t=1}^n \mathbf{u}_t'(\mathbf{Y}, \gamma^0) \mathbf{u}_t(\mathbf{Y}, \gamma^0) \right)^{-1} \sum_{t=1}^n \mathbf{u}_t'(\mathbf{Y}, \gamma^0) z_t(\mathbf{Y}, \gamma^0). \quad (3.1.9)$$

Step 2. Let δ^0 be an initial estimator of δ and \mathbf{G}^0 be its estimated covariance matrix. To estimate consider the model:

$$\theta: \delta^0 = \delta + \mathbf{a} \text{ subject to } \mathbf{f}(\theta) = \mathbf{0}, \quad (3.1.10)$$

where \mathbf{a} is a random column vector with zero mean and covariance matrix \mathbf{G}^0 . Find the value $\hat{\delta}$ that maximizes $(\delta^0 - \delta)'(\mathbf{G}^0)^{-1}(\delta^0 - \delta)$ subject to $\mathbf{f}(\theta) = \mathbf{0}$ and take $\hat{\sigma}_e^2$ such that

$$\mathbf{f}(\hat{\delta}, \hat{\sigma}_e^2) = \mathbf{0}, \quad \hat{\sigma}_e^2 = \hat{\sigma}_z^2 \sum_{j=0}^{p+q} \hat{\beta}_j^2 - \sum_{i=0}^p \sum_{k=0}^p \hat{\alpha}_i \hat{\alpha}_k \gamma_{\mu}(k-j) \text{ where } \hat{\alpha}_0 = \hat{\beta}_0 = 1.$$

In order to find the GLS estimator $\hat{\theta}$ for (3.1.10) the Gauss-Newton procedure for the model is proposed as

$$\begin{pmatrix} \mathbf{T}' & \mathbf{0} \\ \mathbf{0} & c\mathbf{I}_{p+q+1} \end{pmatrix} \begin{pmatrix} \delta^0 \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{T}' & \mathbf{0} \\ \mathbf{0} & c\mathbf{I}_{p+q+1} \end{pmatrix} \begin{pmatrix} \delta \\ f(\theta) \end{pmatrix} + \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix}$$

where $\mathbf{T}'\mathbf{T} = (\mathbf{G}^0)^{-1}$, c is an arbitrary large number and $(\mathbf{a}_1, \mathbf{a}_2)'$ is a random column vector with zero mean and identity estimated covariance matrix. Essentially, Steps 1-2 suggested by Miazaki and Dorea (1993) are Pagano's estimation method applied to model (3.1.1). It differs in that LS estimates are constructed in the first step and the set of statistics used at the second step are different. Pagano parameterized the unrestricted model in terms of α and covariance of $\sum_{j=0}^{p+q} \beta_j z_{t-j}$, $\beta_0 = 1$ whereas γ and σ_z^2 are taken as parameters in Miazaki and Dorea (1993). The parameterization suggested by Miazaki and Dorea (1993) has the advantage that the estimated covariance of $\tilde{\gamma}$ is obtained as a direct result of the computations when γ is estimated in Step 1.

Shin and Sarkar (1994) showed that an ARIMA(p, 1, 0) signal disturbed by MA(q) noise is a ARIMA(p, 1, p+q+1) process restricted by nonlinear constraints on parameters. They considered the following model:

$$x_t = \rho x_{t-1} + z_t \quad (3.1.11)$$

$$z_t + \alpha_1 z_{t-1} + \dots + \alpha_p z_{t-p} = a_t \quad (3.1.12)$$

and

$$u_t = w_t + \beta_1 w_{t-1} + \dots + \beta_p w_{t-p} \quad (3.1.13)$$

for $t = 1, 2, \dots, n$, where p and q are known nonnegative integers, $\alpha_p \neq 0$, the a_t and the w_t are independent unobservable processes, the a_t are iid $N(0, \sigma_a^2)$ and the w_t are iid $N(0, \sigma_w^2)$. The roots of $A(m) = 1 - \alpha_1 L - \dots - \alpha_p L^p = 0$ are assumed to lie outside the unit circle and $\beta = (\beta_1, \dots, \beta_q)'$ and σ_w^2 are assumed to be known. Without loss of generality, we assume $\sigma_w^2 = 1$. From (3.1.0) and (3.1.11)-(3.1.13) we have

$$\sum_{j=0}^p \alpha_j (y_{t-j} - \rho y_{t-1-j}) = \sum_{j=0}^{p+1} \alpha_j^* u_{t-j} + a_t \quad (3.1.14)$$

where $\alpha_j^* = \alpha_j - \rho \alpha_{j-1}$, $j = 0, 1, \dots, p+1$, $\alpha_0 = 1$ and $\alpha_{-1} = \alpha_{p+1} = 0$. Because the autocovariance function of the time series defined by the right hand side of (3.1.14) is zero for lags greater than $(p+q+1)$ it can be written as an MA($p+q+1$) process:

$$\sum_{j=0}^p \alpha_j (y_{t-j} - \rho y_{t-1-j}) = \sum_{j=0}^{p+q+1} \gamma_j e_{t-j}, \quad (3.1.15)$$

where $\{e_t\}$ is a sequence of iid $N(0, \sigma^2)$ random variables, σ^2 and $\gamma_0 = 1$. Equating the covariance function of the right hand side of (3.1.14) to that of the right hand side of (3.1.15), for $h = 0, 1, \dots, p+q+1$, we get

$$\sigma^2 \left(\sum_{j=0}^{p+q+1-h} \gamma_j \gamma_{j+h} \right) - \left\{ \sigma_a^2 \mathbf{I}_0(h) + \sum_{i=0}^{p+1} \sum_{j=0}^{p+1} \alpha_i^* \alpha_j^* \sum_{k=0}^{q-|i-j-h|} \beta_k \beta_{k+|i-j-h|} \right\} = 0, \quad (3.1.16)$$

where $\beta_0 = 1$, $I_0(h) = 1$ for $h = 0$ and $I_0(h) = 0$ for $h \neq 0$. Thus the parameters $(\gamma_1, \dots, \gamma_{p+q+1}, \sigma^2)$ are functionally related to $(\rho, \alpha_1, \dots, \alpha_p, \sigma_a^2)$ and the reparameterized (3.1.15) is a restricted ARIMA model of order $(p, 1, p+q+1)$. Therefore, the reparameterized model (3.1.15) of (3.1.0) and (3.1.11)-(3.1.13) is considered so that results for unrestricted ARIMA models can be used. Let $\theta = (\alpha_1, \dots, \alpha_p, \gamma_1, \dots, \gamma_{p+q+1})'$, $\psi = (\rho, \theta)'$. The negative log likelihood function of $\mathbf{y} = (y_1, \dots, y_n)'$, conditional on y_0 , up to a constant, is

$$L(\xi) = Q(\psi)/\sigma^2 + \log_e \{\det \Gamma(\theta)\} + n \log_e \sigma^2,$$

where

$$Q(\psi) = (\mathbf{y} - \rho \mathbf{y}_1)' \Gamma^{-1}(\theta) (\mathbf{y} - \rho \mathbf{y}_1), \quad \mathbf{y}_1 = (y_0, \dots, y_{n-1})'$$

and $\sigma^2 \Gamma_n(\theta)$ is the $n \times n$ covariance of $(y_1 - \rho y_0, \dots, y_n - \rho y_{n-1})'$. Let ξ^* denote as a ξ -value that minimizes $L(\xi)$ subject to the $(p+q+1)$ restrictions on ξ defined by

$$\mathbf{f} = (f_1, \dots, f_{p+q+1})' = \mathbf{0}. \quad (3.1.17)$$

Then, ξ^* is the restricted ML estimator (RMLE) of ξ . In addition, they considered an approximation $L_m(\xi)$ of $L(\xi)$ and the modified restricted ML estimator (MRMLE) which minimizes $L_m(\xi)$ under restrictions (3.1.17). The MRMLE is defined to be a value of ξ that minimizes

$$L_m(\xi) = S(\psi)/\sigma^2 + n \log_e \sigma^2, \quad (3.1.18)$$

subject to the restriction of (3.1.17), where $S(\psi) = \sum e_t^2(\mathbf{y}; \psi)$ with

$$e_t(\mathbf{y}; \psi) = - \sum_{j=0}^{p+q+1} \gamma_j e_{t-j}(\mathbf{y}; \psi) + \sum_{k=0}^p \alpha_k (y_{t-k} - \rho y_{t-1-k}).$$

Furthermore, they showed that the MRMLE and the RMLE have the same limiting distribution and that this RMLE of the unit root parameter is strongly consistent and it has the same distribution as the LS estimator of the unit root parameter in an AR(1) model tabulated by Dickey and Fuller (1979). Their simulation study showed that unit root tests based on the MRMLE performed very well for small samples and compare favorably with the Said and Dickey (1985) tests with respect to both sizes and powers under model (3.1.0) and (3.1.11)-(3.1.13) with $p = 0$ and $q = 1$.

3.2. General Restricted Model

To illustrate the general restricted model we first consider the vector version of the measurement error model

$$\mathbf{y}_t = \mathbf{z}_t + \mathbf{u}_t, \quad \mathbf{a}(L)\mathbf{z}_t = \boldsymbol{\varepsilon}_t, \quad \mathbf{u}_t = \mathbf{c}(L)\mathbf{w}_t \quad (3.2.1)$$

where

$$\mathbf{c}(L) = \mathbf{I} + \mathbf{C}_1 L + \dots + \mathbf{C}_q L^q, \quad \mathbf{a}(L) = \mathbf{I} + \mathbf{A}_1 L + \dots + \mathbf{A}_p L^p, \quad (3.2.2)$$

and \mathbf{z}_t and \mathbf{u}_t are AR(p) and MA(q) processes respectively. The $\{\varepsilon_t\}$ and $\{\mathbf{w}_t\}$ are two independent sequences of iid $m \times 1$ random vectors with variance covariance matrices $\Omega_{\varepsilon\varepsilon}$ and $\Omega_{\mathbf{w}\mathbf{w}}$ respectively. Assume that $\Omega_{\mathbf{w}\mathbf{w}}$ is known and without loss of generality, we can assume $\Omega_{\mathbf{w}\mathbf{w}} = \mathbf{I}$. For stationarity, invertibility and identifiability, all the roots of $\det[\mathbf{a}(L)]$ are assumed to lie outside the unit circle, $\det[\Omega_{\varepsilon\varepsilon}] \neq 0$, and \mathbf{A}_p is of full rank. We can express the stationary process $\{y_t\}$ of (3.2.1) as an ARMA process. Note that from (3.2.1), $y_t = \mathbf{a}^{-1}(L)\varepsilon_t + \mathbf{c}(L)\mathbf{w}_t$ and hence

$$\mathbf{a}(L)y_t = \varepsilon_t + \mathbf{a}(L)\mathbf{c}(L)\mathbf{w}_t \quad (3.2.3)$$

Since the autocovariance function of the right hand side of (3.2.3) is zero when the lag is greater than $q = p+r$, we can find a q th order moving average $\mathbf{e}_t + \sum_{j=1}^q \mathbf{B}_j \mathbf{e}_{t-j}$ whose autocovariance function is the same as that of the right-hand side of (3.2.3). Hence we can write model (3.2.3) as a special case of the *general restricted model*

$$\mathbf{a}(L)y_t = \mathbf{b}(L)\mathbf{e}_t, \quad t = 1, 2, \dots, n, \quad (3.2.4)$$

where $\{\mathbf{e}_t\}$ is a sequence of iid random vectors with variance covariance matrix Ω and L is the lag operator such that $L^k y_t = y_{t-k}$,

$$\mathbf{b}(L) = \mathbf{I} + \mathbf{B}_1 L + \dots + \mathbf{B}_q L^q. \quad (3.2.5)$$

It is assumed that (p, q) are nonnegative integers and known, and $\mathbf{A} = (\mathbf{A}_1 | \mathbf{A}_2 | \dots | \mathbf{A}_p)$, $\mathbf{B} = (\mathbf{B}_1 | \mathbf{B}_2 | \dots | \mathbf{B}_q)$ are respectively $m \times mp$ and $m \times mq$ matrices of unknown parameters. Let

$\{\mathbf{y}_t\}$, $t = 1, \dots, n$ is the set of $m \times 1$ observation vectors and $\{\mathbf{e}_t\}$ is a sequence of iid m -dimensional random vectors with mean zero and a nonsingular variance covariance matrix Ω . Furthermore, the following regularity conditions C1-C3 should be imposed to satisfy the stationarity, invertibility and identifiability of the model:

C1. All roots of $\det[\mathbf{a}(L)]$ and $\det[\mathbf{b}(L)]$ lie outside the unit circle;

C2. Suppose $\mathbf{a}(L) = \mathbf{c}(L)\mathbf{a}_1(L)$ and $\mathbf{b}(L) = \mathbf{c}(L)\mathbf{b}_1(L)$ for polynomials $\mathbf{a}_1(L)$, $\mathbf{b}_1(L)$ and $\mathbf{c}(L)$, then $\det[\mathbf{c}(L)]$ is independent of L ;

C3. The matrix $[\mathbf{A} \mid \mathbf{B}]$ is of full rank.

The condition for $\mathbf{a}(L)$ in C1 is a stationarity condition for the \mathbf{y}_t . The condition for $\mathbf{b}(L)$ in C1 is an invertibility condition for \mathbf{e}_t to be expressed in terms of $\{y_{t;j}\}$, $j = 0, \dots, \infty$, through $\mathbf{e}_t = \mathbf{b}^{-1}(L)\mathbf{a}(L)\mathbf{y}_t$, and this is used to approximate the error \mathbf{e}_t and the derivatives of the approximation. The role of conditions C2 and C3 is to make the parameters in \mathbf{A} and \mathbf{B} identifiable. Also, it is common to model the elements of Ω as functions of a few parameters with some functional constraints. See Judge et al.(1985, Section 14.5.2).

The vector of restrictions on parameters \mathbf{A} , \mathbf{B} and Ω is defined as

$$\mathbf{f}(\mathbf{A}, \mathbf{B}, \Omega) = \mathbf{0} \quad (3.2.6)$$

where \mathbf{f} is a k -dimensional vector of differential functions. With the reparameterized model (3.2.4), we can exploit the vector ARMA estimation procedure which will be discussed in Section 3.3. Since $\mathbf{b}(L)\mathbf{e}_t$ and $\varepsilon_t + \mathbf{a}(L)\mathbf{c}(L)\mathbf{w}_t$ should have the same autocovariance generating function, it follows that

$$\mathbf{b}(L)\Omega[\mathbf{b}(L^{-1})]' - \Omega_{\varepsilon\varepsilon} - \mathbf{a}(L)\mathbf{c}(L)[\mathbf{c}(L^{-1})]'[\mathbf{a}(L^{-1})]' = \mathbf{0}. \quad (3.2.7)$$

As an application of the above model (3.2.1) and restriction (3.2.3), we consider the case with $m = 1$, $p = 1$, $q = 1$, i.e., we model z_t and u_t as AR(1) and MA(1) respectively

$$y_t = z_t + u_t, \quad \mathbf{a}(L)z_t = \varepsilon_t, \quad u_t = \mathbf{c}(L)w_t \quad (3.2.8)$$

where

$$\mathbf{c}(L) = 1 + C_1L, \quad \mathbf{a}(L) = 1 + A_1L. \quad (3.2.9)$$

The model (3.2.8) can be expressed as follows:

$$y_t = -A_1z_{t-1} + \varepsilon_t + w_t + C_1w_{t-1}, \quad (3.2.10)$$

subject to the restrictions $\mathbf{f} = (f_1, f_2)' = (0, 0)'$.

3.3. Parameter estimation

In this section, three different methods are introduced to estimate unrestricted models and compared with one another so that we can take advantage of theoretical developments in a simulation study. Let $\alpha = (\alpha'_1, \dots, \alpha'_p)'$, $\alpha_i = \text{vec}(\mathbf{A}'_i)$, $\beta = (\beta'_1, \dots, \beta'_q)'$, $\beta_i = \text{vec}(\mathbf{B}'_i)$, $\theta = (\alpha', \beta)'$ and $\xi = (\alpha', \beta', \eta)'$. Letting $\mathbf{e}_t = \mathbf{0}$ for $t \leq 0$ and \mathbf{I}_m denote the m -dimensional identity matrix, we can solve (3.2.4) to get

$$\begin{aligned} \mathbf{e}_t(\theta) &= \mathbf{y}_t + \mathbf{A}_1 \mathbf{y}_{t-1} + \dots + \mathbf{A}_p \mathbf{y}_{t-p} - \mathbf{B}_1 \mathbf{e}_{t-1}(\theta) - \dots - \mathbf{B}_q \mathbf{e}_{t-p}(\theta) \\ &= \mathbf{y}_t + \sum_{i=1}^p (\mathbf{I}_m \otimes \mathbf{y}'_{t-p}) \alpha_i - \sum_{i=1}^q (\mathbf{I}_m \otimes \mathbf{e}'_{t-i}(\theta)) \beta_i \end{aligned} \quad (3.3.1)$$

using $\text{vec}(\mathbf{RST}) = (\mathbf{T}' \otimes \mathbf{R}) \text{vec}(\mathbf{S})$ (see Fuller, 1987, p. 387) with proper orders \mathbf{R} , \mathbf{S} and \mathbf{T} .

In order to develop the conditional Gaussian likelihood of $\{\mathbf{y}_t\}$, $t = 1, \dots, n$, the initial observations $\mathbf{y}_0, \mathbf{y}_{-1}, \dots, \mathbf{y}_{1-p}$ are assumed to be available and fixed and also the initial disturbances $\mathbf{e}_0, \mathbf{e}_{-1}, \dots, \mathbf{e}_{1-q}$ are assumed to be zero, respectively (Reinsel et al., 1992). Then the negative logarithm of the conditional Gaussian likelihood function of $\{\mathbf{y}_t\}$, $t = 1, \dots, n$ is approximated by (Shin, 1993)

$$L_n = L_n(\xi) = \frac{1}{2} \sum_{t=1}^n \mathbf{e}'_t(\theta) \Omega^{-1} \mathbf{e}_t(\theta) + \frac{1}{2} n \log \det(\Omega). \quad (3.3.2)$$

Based on the model (3.2.8)-(3.2.10) as a special case of (3.2.4)-(3.2.5) and (3.2.7) suggested by Shin and Sarkar (1994, p. 2646), (3.3.1) and (3.3.2) can be expressed as follows:

$$e_t^* = e_t^*(\theta) = y_t + A_1 y_{t-1} - B_1 e_{t-1}(\theta) - B_2 e_{t-2}(\theta) \quad (3.3.3)$$

and

$$L_n^* = L_n^*(\xi) = \sum e_t^*(\theta)^2 / 2\sigma_e^2 + n/2(\log_e \sigma_e^2). \quad (3.3.4)$$

3.3.1. Hannan and Rissanen (HR)'s method

As mentioned in Section 2.1.2, we may apply the first and second stages of the multivariate version of Hannan and Rissanen (1982) procedure in order to obtain an initial estimator $\tilde{\xi}$ of ξ . In the first stage, regressing y_t on y_{t-1}, \dots, y_{t-s} , where s is a suitably chosen lag size large enough for approximating y_t by previous observations, the residuals \tilde{e}_t are computed as estimates of the e_t . In the second stage, we estimate $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ by regressing y_t on $\{y_{t-1}, \dots, y_{t-p}\}$ and $\{\tilde{e}_{t-1}, \dots, \tilde{e}_{t-q}\}$. Then $\tilde{\theta} = \{(\text{vec}(\tilde{\mathbf{A}}_1'))', \dots, (\text{vec}(\tilde{\mathbf{A}}_p'))', (\text{vec}(\tilde{\mathbf{B}}_1'))', \dots, (\text{vec}(\tilde{\mathbf{B}}_q'))'\}$, $\tilde{\Omega} = n^{-1} \sum e_t(\tilde{\theta}) e_t'(\tilde{\theta})$ and $\tilde{\eta} = \text{vech}(\tilde{\Omega})$.

3.3.2. Kohn's method

For a general vector linear time series model Kohn (1979) proved the strong consistency and asymptotic normality of parameter estimates obtained by maximizing a

particular time domain approximation to a Gaussian likelihood, where the observations are not necessarily assumed to be normally distributed. To solve the normal equations he set up a constrained Gauss-Newton iteration and obtained the properties of the iterates when the sample size is large. A Newton-Raphson procedure which takes restriction (3.2.6) into account is used to estimate the parameters based on model (3.2.4)-(3.2.6). In particular, an efficient estimator is obtained if we initialize the iteration with a \sqrt{n} -consistent estimator.

Now we evaluate the first partial derivatives $\mathbf{L}_\xi = \partial \mathbf{L}_n / \partial \xi$ of \mathbf{L}_n . Letting $\mathbf{W}_{\theta,t} = \partial \mathbf{e}_t(\theta) / \partial \theta'$, $\mathbf{W}_{\theta,t} = (\mathbf{W}_{\alpha,t} \mid \mathbf{W}_{\beta,t})$ is a $m \times \{m^2(p+q)\}$ matrix, and $\mathbf{L}_\theta = \partial \mathbf{L}_n / \partial \theta = \Sigma \mathbf{W}'_{\theta,t} \Omega^{-1} \mathbf{e}_t(\theta)$. When the restriction (3.2.6) involves the elements of Ω , estimates of α and β are functionally related to the parameter estimates of Ω . In case of $p = q = 1$ and $\Omega \mathbf{B}_1' = \mathbf{A}_1'$ the distribution of the estimator of \mathbf{A}_1 clearly depends on the distribution of the estimator of Ω . This means that it may not be proper to follow the approach of finding a minimizer $\Omega(\theta)$ of $\mathbf{L}_n(\xi)$ with respect to Ω for a fixed θ , and then minimizing $\mathbf{L}_n(\theta, \Omega(\theta))$, which works in the restricted case (Reinsel et al., 1992). Hence we need to find expressions and the asymptotic behavior of the derivatives $\mathbf{L}_\eta = \partial \mathbf{L}_n / \partial \eta$ and $\mathbf{L}_{\eta\eta} = \partial^2 \mathbf{L}_n / \partial \eta \partial \eta'$ with respect to $\eta = \text{vech}(\Omega)$. In handling the derivatives \mathbf{L}_η and $\mathbf{L}_{\eta\eta}$ our analysis is based on Theorem 4.B.2 of Fuller (1987, p. 398), which gives the limiting distribution of the ML estimator of the variance covariance matrix of the multivariate normal distribution based on iid observations. In the sequel, the range for t in the summations will be 1 to n . By (4.B.20) of Fuller (1987, p. 400),

$$\mathbf{L}_\xi = -\Gamma^{-1} \text{vech}[\Sigma \mathbf{e}_t(\theta) \mathbf{e}_t'(\theta) - n\Omega]$$

where $\Gamma = 2\Psi(\Omega \otimes \Omega)\Psi'$ and Ψ is a transformation matrix defined by $\text{vech}(\Omega) = \Psi \text{vec}(\Omega)$. Let ω_{ij} denote the (i, j) th element of Ω , let δ_{ij} be the Kronecker's delta such that $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$, and let $\Psi(ij, ks)$ be the element of Ψ corresponding to ω_{ij} and ω_{ks} of $\text{vech}(\Omega)$ and $\text{vec}(\Omega)$ respectively. Then, from (4.A.3) of Fuller (1987) we have

$$\Psi(ij, ks) = (\delta_{ks}\delta_{si} + \delta_{ki}\delta_{sj})/2, \quad j \leq i.$$

Therefore,

$$\mathbf{L}_\xi = \text{diag}[\Sigma \mathbf{W}'_{\theta,t} | -\{\text{vech}(\Sigma \mathbf{e}_t(\theta) \mathbf{e}_t'(\theta) - n\Omega)\}' \Gamma^{-1}].$$

A Newton-Raphson procedure for estimating the model (3.2.4)-(3.2.5), without restriction (3.2.6), is

$$\xi^+ = \tilde{\xi} - (\tilde{\mathbf{H}}_{\xi\xi})^{-1} \tilde{\mathbf{H}}_{\xi} \quad (3.3.5)$$

where $\tilde{\xi}$ is an initial estimator of ξ . In order to get estimator $\tilde{\xi}$, the first and second stages of multivariate version of Hannan-Rissanen (1982) procedure may be applied to (3.3.5) as mentioned in Section 3.2.1. Now we modify the Newton-Raphson estimator ξ^+ in (3.3.5) to accommodate the restriction (3.2.6). The Lagrangian multiplier method for the vector linear time series models is to be used as a tool of developing the above estimators. The maximum likelihood estimator in this case minimize L_n subject to the restriction (3.2.6) and can be obtained by minimizing

$$\mathbf{L}_n + \mathbf{f}'\lambda, \lambda \in \mathbb{R}^k$$

with respect to ξ and λ . Setting the first partial derivatives of $(\mathbf{L}_n + \mathbf{f}'\lambda)$ with respect to ξ and λ to zero, we get

$$\mathbf{L}_\xi + \mathbf{G}'\lambda = \mathbf{0} \text{ and } \mathbf{f} = \mathbf{0}, \quad (3.3.6)$$

where $\mathbf{G} = \partial \mathbf{f} / \partial \xi'$. Combining the Taylor expansions

$$\mathbf{L}_\xi = \tilde{\mathbf{L}}_\xi + \tilde{\mathbf{L}}_{\xi\xi}(\xi - \tilde{\xi}), \quad \mathbf{f} = \tilde{\mathbf{f}} + \tilde{\mathbf{G}}(\xi - \tilde{\xi}), \quad (3.3.7)$$

we have

$$\tilde{\mathbf{L}}_{\xi\xi}(\xi - \tilde{\xi}) + \tilde{\mathbf{G}}'\lambda = -\tilde{\mathbf{L}}_\xi$$

and

$$\tilde{\mathbf{G}}(\xi - \tilde{\xi}) = -\tilde{\mathbf{f}}$$

where $\tilde{\xi}$ is an initial estimator, and $\tilde{\mathbf{L}}_\xi$, $\tilde{\mathbf{L}}_{\xi\xi}$, $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{f}}$ are the values of \mathbf{L}_ξ , $\mathbf{L}_{\xi\xi}$, \mathbf{G} and \mathbf{f} evaluated at $\xi = \tilde{\xi}$. Note that we have used $\mathbf{f} = \mathbf{0}$. Therefore, the Newton-Raphson estimator is obtained as

$$\begin{bmatrix} \hat{\xi} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \tilde{\xi} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \tilde{\mathbf{L}}_{\xi\xi} & \tilde{\mathbf{G}}' \\ \tilde{\mathbf{G}} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{L}}_\xi \\ \tilde{\mathbf{f}} \end{bmatrix}. \quad (3.3.8)$$

In order for the square matrix in (3.3.8) to be nonsingular, we need to make an assumption that \mathbf{G} is of full rank, which is equivalent to there being no redundancy in $\mathbf{f} = \mathbf{0}$.

When some parameters are restricted to zero, we need to remove only the rows and columns corresponding to the zero-restricted parameters in the vectors and matrices in (3.3.8) instead of incorporating the zero-restrictions into \mathbf{G} and \mathbf{f} .

Applying above to the model (3.2.8)-(3.2.10), the following are obtained:

$$\mathbf{L}_\xi = \partial \mathbf{L}_n / \partial \xi = \begin{bmatrix} \sum W_{A_1,t} e_t^* / \sigma_e^2 \\ \sum W_{B_1,t} e_t^* / \sigma_e^2 \\ \sum W_{B_2,t} e_t^* / \sigma_e^2 \\ -\frac{\sum e_t^{*2}}{2(\sigma_e^2)^2} + \frac{n}{2\sigma_e^2} \end{bmatrix}, \quad (3.3.9)$$

and by differentiating (3.3.3) the elements of $W_{A_1,t}$, $W_{B_1,t}$ and $W_{B_2,t}$ are computed recursively as

$$W_{A_1,t} = y_{t-1} - B_1 W_{A_1,t-1} - B_2 W_{A_1,t-2}$$

$$W_{B_1,t} = -e_{t-1}^* - B_1 W_{B_1,t-1} - B_2 W_{B_1,t-2}$$

and

$$W_{B_2,t} = -e_{t-2}^* - B_1 W_{B_2,t-1} - B_2 W_{B_2,t-2}.$$

Note that $\mathbf{L}_{\xi\xi} = \partial^2 \mathbf{L}_n / \partial \xi \partial \xi'$ can be expressed as follows:

$$\mathbf{L}_{\xi\xi} = \begin{bmatrix} \mathbf{P} & \mathbf{q}' \\ \mathbf{q} & s \end{bmatrix},$$

where

$$\mathbf{P} = \begin{bmatrix} \Sigma(W_{A_1,t}^2 + e_t^* WW_{A_1,t}) / \sigma_e^2 & \Sigma(W_{A_1,t} W_{B_1,t} + e_t^* WW_{A_1,B_1}) / \sigma_e^2 & \Sigma(W_{A_1,t} W_{B_2,t} + e_t^* WW_{A_1,B_2,t}) / \sigma_e^2 \\ \Sigma(W_{A_1,t} W_{B_1,t} + e_t^* WW_{A_1,B_1,t}) / \sigma_e^2 & \Sigma(W_{B_1,t}^2 + e_t^* WW_{B_1,t}) / \sigma_e^2 & \Sigma(W_{B_1,t} W_{B_2,t} + e_t^* WW_{B_1,B_2}) / \sigma_e^2 \\ \Sigma(W_{A_1,t} W_{B_2,t} + e_t^* WW_{A_1,B_2,t}) / \sigma_e^2 & \Sigma(W_{B_1,t} W_{B_2,t} + e_t^* WW_{B_1,B_2}) / \sigma_e^2 & \Sigma(W_{B_2,t}^2 + e_t^* WW_{B_2,t}) / \sigma_e^2 \end{bmatrix}$$

with

$$WW_{A_1,B_1} = \partial W_{A_1,t} / \partial B_1 = -W_{A_1,t-1} - B_1 WW_{A_1,B_1,t-1} - B_2 WW_{A_1,B_1,t-2},$$

$$WW_{A_1,B_2,t} = \partial W_{A_1,t} / \partial B_2 = -W_{A_1,t-2} - B_1 WW_{A_1,B_2,t-1} - B_2 WW_{A_1,B_2,t-2},$$

$$WW_{B_1,t} = \partial W_{B_1,t} / \partial B_1 = -2W_{B_1,t-1} - B_1 WW_{B_1,t-1} - B_2 WW_{B_1,t-2},$$

$$WW_{B_1,B_2,t} = \partial W_{B_1,t} / \partial B_2 = -W_{B_2,t-1} - B_1 WW_{B_1,B_2,t-1} - W_{B_1,t-2} - B_2 WW_{B_1,B_2,t-2},$$

$$WW_{B_2,t} = \partial W_{B_2,t} / \partial B_2 = -2W_{B_2,t-1} - B_1 WW_{B_2,t-1} - B_2 WW_{B_2,t-2},$$

$$\mathbf{q} = \left[-\sum e_t^* W_{A_1,t} / (\sigma_e^2)^2 \quad -\sum e_t^* W_{B_1,t} / (\sigma_e^2)^2 \quad -\sum e_t^* W_{B_2,t} / (\sigma_e^2)^2 \right]$$

and

$$s = \Sigma(e_t^*)^2 / (\sigma_e^2)^3 - n / (\sigma_e^2)^2.$$

Furthermore,

$$\mathbf{f} = (f_1, f_2)' = [\sigma_e^2(B_1 + B_1 B_2) - (A_1 + C_1 + A_1^2 C_1 + A_1 C_1^2), \sigma_e^2 B_2 - A_1 C_1]',$$

and

$$\mathbf{G} = \partial \mathbf{f} / \partial \xi' = \begin{bmatrix} -1 - 2A_1 C_1 & \sigma_e^2 + B_2 \sigma_e^2 & B_1 \sigma_e^2 & B_1 + B_1 B_2 \\ -C_1 & 0 & \sigma_e^2 & B_2 \end{bmatrix}.$$

Now we can compute an estimate for model (3.2.1), by applying (3.3.8) with \mathbf{G} , \mathbf{f} and $\mathbf{L}_{\xi\xi}$

described above to the reparameterized model (3.2.4).

3.3.3. Shin and Sarkar's method

Based on the Kohn's method, Shin and Sarkar (1995) considered an approximation $\mathbf{H}_{\xi\xi}$ of $\mathbf{L}_{\xi\xi}$ in (3.3.8), in order to compute an easier Newton-Raphson estimator that approximates the restricted ML estimator. They defined

$$\mathbf{H}_{\xi\xi} = \text{diag}[\Sigma \mathbf{W}'_{\theta,t} \Omega^{-1} \mathbf{W}_{\theta,t}, -n\Gamma^{-1}]. \quad (3.3.10)$$

Approximation $\Sigma \mathbf{W}'_{\theta,t} \Omega^{-1} \mathbf{W}_{\theta,t}$ of $\partial^2 L_n / \partial \theta \partial \theta'$ is well explained in Reinsel et al. (1992) and by (4.B.22) of Fuller (1987) $\mathbf{L}_{\eta\eta} = -n\Gamma^{-1} + o_p(n)$. Hence, $\mathbf{L}_{\xi\xi} = \mathbf{H}_{\xi\xi} + o_p(n)$. Applying (3.3.10) to model (3.2.8)-(3.2.10), the following is obtained:

$$\mathbf{H}_{\xi\xi} = \begin{bmatrix} \sum W_{A,t}^2 / \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sum W_{B_1,t}^2 / \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sum W_{B_2,t}^2 / \sigma_e^2 & 0 \\ 0 & 0 & 0 & \frac{n}{2(\sigma_e^2)^2} \end{bmatrix}.$$

Therefore, by replacing $\tilde{\mathbf{L}}_{\xi\xi}$ with its approximation $\tilde{\mathbf{H}}_{\xi\xi}$, the Newton-Raphson estimator is obtained as

$$\begin{bmatrix} \hat{\xi} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \tilde{\xi} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \tilde{\mathbf{H}}_{\xi\xi} & \tilde{\mathbf{G}}' \\ \tilde{\mathbf{G}} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{L}}_{\xi} \\ \tilde{\mathbf{f}} \end{bmatrix}. \quad (3.3.11)$$

3.4. The limiting distribution

Assume that $\sqrt{n}(\hat{\xi} - \xi) = O_p(1)$. Combining the Taylor expansion

$$\begin{bmatrix} \tilde{\mathbf{H}}_{\xi} \\ \tilde{\mathbf{f}} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{\xi} \\ \mathbf{f} \end{bmatrix} - \begin{bmatrix} \tilde{\mathbf{H}}_{\xi\xi} & \tilde{\mathbf{G}}' \\ \tilde{\mathbf{G}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\xi} - \xi \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} O_p(1) \\ o_p(1) \end{bmatrix}$$

with (3.3.8) replaced $\mathbf{L}_{\xi\xi}$ by $\mathbf{H}_{\xi\xi}$ and $\mathbf{f} = \mathbf{0}$, we get

$$\begin{bmatrix} \hat{\xi} - \xi \\ \hat{\lambda} \end{bmatrix} = - \begin{bmatrix} \tilde{\mathbf{H}}_{\xi\xi} & \tilde{\mathbf{G}}' \\ \tilde{\mathbf{G}} & \mathbf{0} \end{bmatrix}^{-1} \left\{ \begin{bmatrix} \mathbf{H}_{\xi} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} O_p(1) \\ o_p(1) \end{bmatrix} \right\}.$$

By Dunsmuir and Hannan (1976) together with theorem 4.B.2 of Fuller (1987),

$$\frac{1}{\sqrt{n}} \mathbf{H}_{\xi} \xrightarrow{d} N(\mathbf{0}, \mathbf{V}) \quad \text{and} \quad \frac{1}{n} \mathbf{H}_{\xi\xi} \xrightarrow{p} \mathbf{V},$$

where $\mathbf{V} = \text{diag}[\mathbf{I}(\theta), \Gamma^{-1}]$ and $\mathbf{I}(\theta)$ is the information matrix of θ in the unrestricted model (3.2.4)-(3.2.5). The normality of the \mathbf{e}_t is used for convergence in distribution of $n^{-1/2} \mathbf{H}_{\eta}$, the elements of $n^{-1/2} \mathbf{H}_{\xi}$ corresponding to the parameters η . Since \mathbf{G} is a continuous function of ξ , $\tilde{\mathbf{G}} \xrightarrow{p} \mathbf{G}$. Then from Shin and Sarkar (1995) the limiting distribution of these estimator $\hat{\xi}$ follows.

Theorem 3.1. Assume that the model (3.2.4)-(3.2.6) holds with regularity conditions C1-C3. And assume that the \mathbf{e}_t 's are normal, \mathbf{f} is continuously differentiable and G is nonsingular. Suppose that the initial estimator $\hat{\xi}$ of ξ satisfies $\sqrt{n}(\hat{\xi} - \xi) = O_p(1)$. Let $(\hat{\xi}, \hat{\lambda})$ be as defined in (3.3.11). Then,

$$\begin{bmatrix} \sqrt{n}(\hat{\xi} - \xi) \\ \frac{1}{\sqrt{n}}\hat{\lambda} \end{bmatrix} \xrightarrow{d} N\left(0, \begin{bmatrix} \mathbf{V} & \mathbf{G}' \\ \mathbf{G} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{V} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V} & \mathbf{G}' \\ \mathbf{G} & 0 \end{bmatrix}^{-1}\right). \quad (3.4.1)$$

Using the formula for the inverse of a partitioned matrix (see, for example, Judge et al., 1985, p. 947), it can easily be shown that

$$\begin{bmatrix} \mathbf{V} & \mathbf{G}' \\ \mathbf{G} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{V} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V} & \mathbf{G}' \\ \mathbf{G} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{V}^{-1} - \mathbf{V}_g' \mathbf{V}_{gg} \mathbf{V}_g & 0 \\ 0 & \mathbf{V}_{gg} \end{bmatrix} \quad (3.4.2)$$

where $\mathbf{V}_g = \mathbf{G}\mathbf{V}^{-1}$ and $\mathbf{V}_{gg} = (\mathbf{G}\mathbf{V}^{-1}\mathbf{G}')^{-1}$.

Note that the upper left block of the covariance matrix in the right hand side of (3.4.13) corresponding to ξ is the same as the upper-left block of the matrix $\begin{bmatrix} \mathbf{V} & \mathbf{G}' \\ \mathbf{G} & 0 \end{bmatrix}^{-1}$.

Since $I(\theta)$ can be consistently estimated by $\sum \hat{\mathbf{W}}'_{0,t} \hat{\Omega}^{-1} \hat{\mathbf{W}}_{0,t}$, the variance covariance matrix of $\hat{\xi}$ is estimated by the upper-left block of the matrix $\begin{bmatrix} \hat{\mathbf{H}}_{\xi\xi} & \hat{\mathbf{G}}' \\ \hat{\mathbf{G}} & 0 \end{bmatrix}^{-1}$. Then the

limit distributions of $n(\hat{A}_1 - 1)$ and $\hat{\tau} = \hat{A}_1 / \text{s.e.}(\hat{A}_1)$ under the unit root null hypothesis are given by the following.

Theorem 3.2. Given the model (3.2.10) with restrictions $f = (f_1, f_2)' = (0, 0)'$, where $\sigma_e^2 = V(e_t)$ and $f_1 = \sigma_e^2(B_1 + B_1B_2) - (A_1 + C_1 + A_1^2C_1 + A_1C_1^2)$, $f_2 = \sigma_e^2B_2 - A_1C_1$, the limiting distribution of $n(\hat{A}_1 - 1)$ and $\hat{\tau}$ under the null hypothesis $H_0: A_1 = -1$ are

given by $\frac{\frac{1}{2}\{W(1)^2 - 1\}}{\int_0^1 W(r)^2 dr}$ and $\frac{\frac{1}{2}\{W(1)^2 - 1\}}{\sqrt{\int_0^1 W(r)^2 dr}}$ respectively, where $W(\cdot)$ is the standard

Brownian motion on $[0, 1]$.

Observe that Theorem 3.1 and Theorem 3.2 give asymptotics of a unit root parameter estimator and unit root tests based on Shin and Sarkar's method. The limiting distributions of unit root tests are the same as in Theorem 3.2 when unit root parameter estimates are based on Kohn's and Hannan-Rissanen's methods.

CHAPTER IV

SIMULATION STUDY

We now consider a Monte Carlo study on the power functions of our unit root tests under model (3.2.8)-(3.2.10). Ten thousand replications were simulated for sample sizes 25, 50, 100 and 250, for $A_1 = 1.00, 0.99, 0.95, 0.90$ and 0.70 , and for nominal levels 1%, 5% and 10%. The normal random numbers $\{\varepsilon_t\}$ and $\{w_t\}$ were generated by the subroutine DRNNOA and the calculation of inverse matrix are performed by the subroutine DLINRG of the IMSL package. The value of σ_w^2 was set to one. The value of C_1 was set to $-0.5, 0$ and 0.5 , and the value of σ_e^2 was set to $0.2, 1.0$ and 5.0 . For different (n, C_1, σ_e^2) combinations independent samples were used. The computed values of the tests $n(\hat{A}_1 - 1)$ and $\hat{\tau}$ were compared to the theoretical 1%, 5% and 10% left tail critical values tabulated by Dickey and Fuller (see Fuller 1976, p. 371 and p. 373). Section 4.1 discusses algorithms based on three different estimation methods in Section 3.3, and Section 4.2 gives simulation results and conclusive remarks.

4.1. Algorithms

Three different testing algorithms are introduced to perform the comparative study on the basis of Section 3.3.

4.1.1. Kohn's method

Step 1. Generate random numbers, $\{e_t\}$ and $\{w_t\}$, as follows:

$$\{e_t\} \sim \text{NID}(0, \sigma_e^2), \{w_t\} \sim \text{NID}(0, 1)$$

and

$$\{e_t\} \text{ and } \{w_t\} \text{ are independent}$$

where $t = 1, \dots, n$ for $\sigma_e^2 = 0.2, 1.0, 5.0$ and for sample size $n = 25, 50, 100, 250$.

Step 2. Compute the following y_t 's based on random variables $\{e_t\}$ and $\{w_t\}$ obtained in Step 1:

$$y_t = z_t + u_t$$

where $z_t = -A_1 z_{t-1} + e_t$ and $u_t = w_t + C_1 w_{t-1}$ with $e_t = w_t = 0$ if $t \leq 0$, $A_1 = 1.0, 0.99, 0.95, 0.9, 0.8$ and 0.7 , $C_1 = -0.5, 0.0$ and 0.5 .

Step 3. Estimate initial values $\tilde{\xi} = (\tilde{A}_1, \tilde{B}_1, \tilde{B}_2, \tilde{\sigma}_e^2)'$ of $\xi = (A_1, B_1, B_2, \sigma_e^2)'$ using the following steps:

(3.1) Regress y_t on $(y_{t-1}, \dots, y_{t-s})$ to obtain the residuals, \tilde{e}_t , where $\tilde{e} = \mathbf{Y} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ with $\mathbf{Y} = (y_1, \dots, y_n)'$ and the \mathbf{X} matrix, for $s = 5$ for example, is given by

$$\mathbf{X} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ y_1 & 0 & 0 & 0 & 0 \\ y_2 & y_1 & 0 & 0 & 0 \\ y_3 & y_2 & y_1 & 0 & 0 \\ y_4 & y_3 & y_2 & y_1 & 0 \\ y_5 & y_4 & y_3 & y_2 & y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{n-1} & y_{n-2} & y_{n-3} & y_{n-4} & y_{n-5} \end{pmatrix}.$$

(3.2) Regress y_t on $(y_{t-1}, \tilde{e}_{t-1}, \tilde{e}_{t-2})$ to estimate the initial values $(\tilde{A}_1, \tilde{B}_1, \tilde{B}_2)$ of (A_1, B_1, B_2) , where $\{\tilde{e}_t\}$ is obtained from Step (3.1).

(3.3) Compute the initial estimate of σ_e^2 as $\tilde{\sigma}_e^2 = ((e_1^*)^2 + \dots + (e_t^*)^2)/n$, $t = 1, \dots, n$ where $e_t^* = y_t + \tilde{A}_1 y_{t-1} - \tilde{B}_1 e_{t-1}^* - \tilde{B}_2 e_{t-2}^*$ with $e_t^* = 0$, $t \leq 0$.

Step 4. Obtain the final estimates $\hat{\xi} = (\hat{A}_1, \hat{B}_1, \hat{B}_2, \hat{\sigma}_e^2)'$ of ξ , by using the Newton-Raphson method (3.3.8) based on the initial estimates, $\tilde{\xi} = (\tilde{A}_1, \tilde{B}_1, \tilde{B}_2, \tilde{\sigma}_e^2)'$ in Step 3. In the numerical computations, the following revised Newton-Raphson estimator instead of (3.3.8) will be examined:

$$\begin{bmatrix} \hat{\xi} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \tilde{\xi} \\ \mathbf{0} \end{bmatrix} - k \begin{bmatrix} \tilde{\mathbf{L}}_{\xi\xi} & \tilde{\mathbf{G}}' \\ \tilde{\mathbf{G}} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{L}}_{\xi} \\ \tilde{\mathbf{f}} \end{bmatrix} \quad (4.1.1)$$

where k is a scalar controlling the step length. We calculate the left-hand side and the associated value for the log likelihood for various values of k and choose as the estimate the left-hand side value that produces the biggest value for the log likelihood as defined in equation (3.3.2).

Step 5. Compute statistics $n(\hat{A}_1 - 1)$ and $\hat{\tau}$ and compare their values with the tabulated percentiles of the distribution given by Dickey and Fuller (1979) and check if the values of $n(\hat{A}_1 - 1)$ and $\hat{\tau}$ value are less than the corresponding tabulated percentiles.

Step 6. Repeat Step 1 through Step 5 ten thousand times and thus count the relative frequencies of $n(\hat{A}_1 - 1)$ and $\hat{\tau}$ smaller than the corresponding tabulated percentiles.

4.1.2. Shin and Sarkar (SS)'s method

All steps are the same as those of Kohn's method except that we replace $\tilde{\mathbf{L}}_{\xi\xi}$ with its approximation $\tilde{\mathbf{H}}_{\xi\xi}$ in (4.1.1).

4.1.3. Hannan and Rissanen (HR)'s method

We follow Steps 1, 2, 3, 5, and 6 of Kohn's (Section 4.1.1) method in obtaining the statistic $n(\hat{A}_1 - 1)$. In particular, in order to obtain the standard error of \hat{A}_1 when $\hat{\tau}$ is

calculated, the matrix $\begin{bmatrix} \tilde{\mathbf{L}}_{\xi\xi} & \tilde{\mathbf{G}}' \\ \tilde{\mathbf{G}} & \mathbf{0} \end{bmatrix}^{-1}$ of (4.1.1) is computed.

4.2. Results and Discussion

Simulation results on the empirical level and power of the test statistics $n(\hat{A}_1 - 1)$ and $\hat{\tau}$ are presented in Tables 4.1-4.28. Note that Tables 4.1-4.12 show the empirical powers of the two test statistics based on Hannan and Rissanen's method. In preparing for $\hat{\tau}$, we needed to compute the standard error of \hat{A}_1 . For Tables 4.4-4.6 these were done by using

$$\begin{bmatrix} \tilde{\mathbf{L}}_{\xi\xi} & \tilde{\mathbf{G}}' \\ \tilde{\mathbf{G}} & \mathbf{0} \end{bmatrix}^{-1}$$

of (4.1.1) and for Tables 4.10-4.12 by using

$$\begin{bmatrix} \tilde{\mathbf{H}}_{\xi\xi} & \tilde{\mathbf{G}}' \\ \tilde{\mathbf{G}} & \mathbf{0} \end{bmatrix}^{-1}.$$

While Tables 4.13-4.18 present simulation numbers based on Kohn's method, and Tables 4.19-4.24 based on Shin and Sarkar's method. We now discuss our findings based on the tables.

Common points among three methods

(1) When we fix n and increase σ_e^2 or fix σ_e^2 and increase n , the empirical powers are getting bigger and sizes are closer to the nominal level in terms of both $n(\hat{A}_1 - 1)$ and $\hat{\tau}$. In particular, In case of large samples ($n = 100, 250$), the empirical sizes of each of three methods are almost close to the desired nominal levels except that $\sigma_e^2 = 0.2$.

(2) For $\sigma_e^2 = 0.2$ (i.e., when the signal is weak), especially when $C_1 = 0.5$, the size tends to be larger than the nominal level. For fixed (n, C_1, σ_e^2) the powers are monotone functions of A_1 .

(3) The comparisons for nominal levels 1%, 5%, and 10% are similar.

What are the differences among three methods ?

(1) For the “normalized” unit root test statistic $n(\hat{A}_1 - 1)$, the empirical sizes using Kohn’s estimates are usually slightly smaller than those using Shin and Sarkar’s estimates; and sizes are close to one another. As expected, the empirical powers based on Shin and Sarkar’s method generally are larger than those based on Kohn’s due to the larger empirical sizes. There is no big differences between tests based on Kohn’s and Shin and Sarkar’s estimates.

(2) For the t-test statistic $\hat{\tau}$, the sizes using Kohn’s estimates are in general much smaller than those using Shin and Sarkar’s. The powers for Shin and Sarkar’s method become closer as the sample size gets larger and C_1 gets smaller.

Summary

For small sample sizes ($n = 25, 50$), all three different methods (H-R, Kohn, S-S) lead to test statistics for which the tabulated critical values given by Dickey and Fuller (1979) are not appropriate. Numerical results of Schwert (1989) show that the distribution of unit root test statistics in unrestricted ARMA models (containing a nonzero MA component) can be very different from those tabulated by Dickey and Fuller (1979). Therefore, we prepared Table 4.25 - Table 4.32 to obtain the empirical critical values for

both normalized unit root test statistic and t-test statistic in order to apply the unit root tests to real time series data.

The advantage of using Shin and Sarkar's method lies in the fact that it is much easier than Kohn's for computation in terms of inverting the Hessian matrix. If p , q , m are bigger than the case considered in our simulation study ($p=1$, $q=1$, $m=1$), gain in computational ease will be even more noticeable. In terms of statistic $n(\hat{A}_1 - 1)$, the Shin and Sarkar method can be a good alternative to the Kohn's method when we test for a unit root in an AR(1) signal observed with MA(1) noise. On the other hand, the Kohn's method is preferable while using the t-test statistic $\hat{\tau}$.

Table 4.1

Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.01

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	17.1	18.6	20.9	6.9	7.5	8.8	3.6	4.2	4.3
	0.99	18.8	19.6	24.4	7.5	7.9	9.6	3.8	4.7	5.3
	0.95	26.5	28.0	35.8	11.7	11.3	14.5	6.5	6.5	7.4
	0.90	34.8	37.6	45.7	17.9	18.0	20.2	10.6	10.6	11.8
	0.80	46.3	55.7	66.1	29.0	31.2	38.5	19.8	21.2	23.1
	0.70	53.7	64.9	77.9	40.9	46.0	54.4	32.6	33.9	36.5
50	1.00	11.4	10.4	12.8	4.2	4.1	4.5	2.4	2.6	2.7
	0.99	14.3	12.3	14.3	5.4	5.0	5.5	3.4	3.0	3.4
	0.95	26.8	23.1	28.6	10.9	10.0	11.5	6.7	6.5	6.9
	0.90	42.7	40.4	47.0	22.1	21.2	23.6	15.6	16.2	16.8
	0.80	63.6	64.8	72.1	47.4	48.5	53.2	40.8	41.2	42.5
	0.70	72.5	77.4	83.3	64.8	67.0	71.3	62.1	62.5	63.8
100	1.00	6.3	5.5	7.7	2.6	2.4	2.8	1.9	1.6	1.8
	0.99	9.6	8.4	11.5	3.7	3.5	3.8	2.4	2.6	2.5
	0.95	29.8	27.9	34.5	15.4	14.5	16.3	11.9	11.5	12.0
	0.90	58.3	57.1	65.0	40.9	40.8	44.5	36.7	35.9	36.6
	0.80	80.6	84.3	88.3	79.0	80.8	83.8	81.3	81.8	81.8
	0.70	87.3	90.5	93.4	88.7	90.9	92.6	92.7	93.3	93.2
250	1.00	2.8	2.3	3.2	1.4	1.5	1.3	1.3	1.2	1.0
	0.99	7.3	6.7	9.4	4.3	4.3	5.0	3.8	3.8	4.0
	0.95	54.2	53.5	60.4	50.5	50.8	52.0	49.1	49.3	49.2
	0.90	83.3	86.5	89.9	93.8	94.6	95.2	96.3	96.4	96.0
	0.80	91.8	94.5	95.7	99.3	99.7	99.7	100.0	100.0	100.0
	0.70	92.6	95.2	96.2	99.0	99.6	99.7	100.0	100.0	100.0

Table 4.2

Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.05

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	26.1	26.7	29.2	13.6	14.0	15.3	8.9	9.4	10.2
	0.99	29.2	28.5	33.1	14.7	14.6	16.4	10.0	10.3	11.1
	0.95	40.1	38.8	46.8	22.2	21.5	25.4	15.6	15.3	16.9
	0.90	50.4	50.9	58.3	32.8	30.9	34.7	24.4	23.0	24.5
	0.80	63.5	68.6	76.1	49.4	49.9	55.8	40.2	40.5	43.5
	0.70	69.5	75.6	84.1	61.7	64.8	70.4	55.3	56.8	58.8
50	1.00	18.2	18.8	20.3	9.6	9.5	10.6	7.5	7.3	7.6
	0.99	22.7	20.1	23.3	12.4	11.5	12.4	9.9	8.8	9.3
	0.95	39.3	36.0	42.5	23.1	23.1	24.3	19.2	18.4	18.7
	0.90	57.7	55.3	60.9	42.2	41.6	44.3	37.3	37.8	37.8
	0.80	74.4	75.2	79.1	69.2	69.9	73.4	69.2	68.7	70.2
	0.70	80.1	82.7	86.7	79.3	80.8	82.7	82.2	82.3	82.4
100	1.00	13.0	11.7	14.9	7.7	6.8	8.1	6.6	5.6	6.3
	0.99	18.8	17.5	22.6	10.7	10.9	12.0	9.0	9.5	9.4
	0.95	49.3	48.3	55.5	37.6	37.8	40.3	34.5	34.8	35.0
	0.90	75.2	75.3	81.5	71.8	72.2	75.3	72.2	72.0	72.3
	0.80	87.7	90.5	92.3	92.4	94.1	95.3	95.9	96.5	96.3
	0.70	90.7	92.6	94.9	94.1	96.1	96.5	98.0	98.3	98.1
250	1.00	7.9	7.1	8.5	5.9	5.7	6.0	5.5	5.4	5.3
	0.99	19.6	18.9	23.0	15.9	15.9	17.3	15.3	15.2	15.7
	0.95	76.9	79.6	84.2	86.1	86.5	86.9	87.7	87.3	86.8
	0.90	90.9	94.0	95.8	99.5	99.6	99.7	99.9	99.9	99.9
	0.80	93.7	95.9	96.9	99.8	99.9	99.9	100.0	100.0	100.0
	0.70	93.6	95.9	96.9	99.5	99.8	99.8	100.0	100.0	100.0

Table 4.3

Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.10

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	32.1	32.3	34.8	18.5	19.1	20.2	14.0	14.6	14.8
	0.99	35.7	34.5	39.1	20.3	20.5	22.7	15.5	15.9	16.7
	0.95	47.6	45.9	53.7	30.0	29.4	33.6	23.7	23.1	24.9
	0.90	58.4	58.8	64.5	43.1	40.7	44.4	35.3	33.9	35.4
	0.80	71.2	74.2	80.2	60.6	60.5	65.9	53.8	54.0	56.6
	0.70	76.1	79.9	86.4	71.1	72.9	76.8	67.7	68.6	69.9
50	1.00	23.4	21.6	25.9	14.7	14.6	15.8	12.3	12.4	12.6
	0.99	28.9	25.7	29.7	19.0	17.6	19.0	15.9	14.8	15.6
	0.95	48.6	45.0	51.0	33.8	33.8	36.3	30.4	30.2	30.1
	0.90	66.2	63.6	68.1	55.9	55.8	57.9	54.0	54.4	53.6
	0.80	78.4	79.1	82.0	78.7	79.5	81.8	82.0	81.7	82.2
	0.70	82.4	84.8	87.8	84.2	85.6	86.5	88.9	89.0	88.6
100	1.00	18.9	17.3	21.4	12.7	11.9	13.4	11.5	10.8	11.3
	0.99	26.3	25.2	31.2	18.5	19.0	20.2	16.6	17.0	17.1
	0.95	61.5	61.2	68.9	55.4	54.9	57.6	53.8	53.1	54.0
	0.90	81.9	83.1	87.7	85.1	86.3	88.1	87.5	87.8	87.8
	0.80	89.6	92.6	93.8	95.6	97.0	97.3	98.6	98.9	98.8
	0.70	91.9	93.5	95.4	95.8	97.3	97.6	99.0	99.2	99.0
250	1.00	12.7	12.1	14.4	10.8	10.6	11.2	10.1	10.2	10.3
	0.99	31.4	30.5	35.4	28.9	28.5	29.5	27.6	27.3	28.5
	0.95	85.8	88.7	91.8	95.7	96.0	96.2	96.9	96.8	96.5
	0.90	93.3	95.8	97.2	99.9	99.9	99.9	100.0	100.0	100.0
	0.80	94.5	96.6	97.3	99.8	100.0	100.0	100.0	100.0	100.0
	0.70	93.9	96.2	97.1	99.6	99.9	99.9	100.0	100.0	100.0

Table 4.4

Empirical power (%) of $\hat{\tau}$ for nominal level 0.01

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	12.3	6.7	11.6	5.5	3.1	3.3	3.4	3.2	2.5
	0.99	13.7	6.8	13.6	5.9	3.2	3.9	3.5	3.4	2.9
	0.95	18.5	10.0	19.6	9.1	4.6	5.6	6.0	4.6	4.3
	0.90	23.0	13.1	25.3	13.1	7.3	8.2	9.4	8.0	7.0
	0.80	27.8	20.0	40.0	20.2	13.5	15.9	17.7	16.3	15.3
	0.70	30.4	26.8	52.1	26.9	21.6	24.2	27.8	26.7	26.2
50	1.00	6.7	2.3	3.5	2.8	1.5	1.8	2.1	2.0	1.7
	0.99	8.1	3.0	4.2	3.4	2.3	2.1	2.9	2.4	2.3
	0.95	15.1	6.2	8.8	7.0	4.1	4.4	6.0	5.2	4.8
	0.90	24.3	11.1	16.4	14.4	9.8	9.9	13.9	13.1	12.3
	0.80	35.3	21.5	32.5	30.4	25.3	26.7	36.4	35.5	34.0
	0.70	39.3	29.5	47.0	42.0	40.4	40.7	55.8	55.5	55.17
100	1.00	2.7	1.2	2.6	1.8	1.2	1.6	1.7	1.4	1.4
	0.99	3.9	1.6	4.1	2.3	1.6	2.2	2.1	2.0	1.9
	0.95	12.4	6.5	11.6	9.5	8.1	9.8	10.5	9.6	9.5
	0.90	26.1	16.1	22.9	26.2	25.0	29.0	33.1	31.8	31.4
	0.80	44.4	32.7	37.8	57.2	61.0	66.3	76.7	77.1	75.6
	0.70	51.6	41.2	48.1	68.1	74.5	79.3	89.6	90.4	90.3
250	1.00	1.1	0.9	3.0	1.1	1.0	0.9	1.2	1.0	0.9
	0.99	3.2	3.2	8.8	3.2	3.00	3.9	3.8	3.4	3.5
	0.95	27.5	31.3	52.3	41.2	42.2	45.1	47.8	47.7	46.6
	0.90	50.3	63.0	80.3	87.9	90.1	91.3	95.5	95.5	95.1
	0.80	62.1	71.3	81.0	97.9	99.1	99.3	100.0	100.0	100.0
	0.70	64.3	69.6	73.4	97.2	99.1	99.3	100.0	100.0	100.0

Table 4.5

Empirical power (%) of $\hat{\tau}$ for nominal level 0.05

Sample Size	C_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	18.5	11.6	16.5	10.4	7.8	8.0	8.3	8.1	7.7
	0.99	20.1	11.9	18.8	11.0	7.9	8.7	9.3	8.8	8.6
	0.95	26.6	17.3	27.2	16.5	11.8	13.7	14.5	13.1	12.8
	0.90	32.9	22.5	34.2	24.2	17.9	18.8	22.6	20.2	19.5
	0.80	38.6	33.1	49.6	33.9	29.6	32.8	36.9	35.9	36.5
	0.70	41.3	39.3	60.8	41.4	41.1	44.8	49.7	51.0	50.6
50	1.00	11.1	7.0	9.7	7.3	5.9	6.9	6.9	6.6	6.4
	0.99	13.5	7.7	10.9	8.8	7.4	7.9	9.3	7.7	7.8
	0.95	24.3	14.8	20.4	17.0	15.3	16.2	17.8	16.5	16.1
	0.90	35.8	24.4	31.0	31.3	29.0	30.7	35.0	34.7	33.5
	0.80	48.3	38.8	47.9	52.8	53.3	56.9	65.5	65.1	64.8
	0.70	51.6	46.7	60.2	61.9	65.7	68.6	78.5	79.0	78.6
100	1.00	7.0	5.5	12.0	6.1	5.2	6.6	6.1	5.3	5.7
	0.99	10.1	8.8	18.5	8.5	8.6	9.8	8.5	8.7	8.7
	0.95	27.7	26.0	41.7	30.1	30.8	34.4	33.2	33.3	32.8
	0.90	45.8	44.6	58.8	61.0	62.8	68.0	70.0	69.8	69.2
	0.80	59.2	59.2	65.2	84.0	88.3	91.0	94.8	95.3	95.1
	0.70	63.5	62.0	69.3	85.7	91.4	93.5	97.3	98.0	97.5
250	1.00	4.9	5.4	9.4	5.1	5.1	5.5	5.3	5.2	4.9
	0.99	12.8	14.5	24.8	14.1	14.2	16.0	15.0	14.7	15.5
	0.95	61.6	69.4	81.6	82.5	82.9	84.6	86.9	86.4	85.8
	0.90	78.4	87.8	93.3	98.7	99.2	99.5	99.8	99.8	99.8
	0.80	78.6	88.4	91.2	99.5	99.8	99.9	100.0	100.0	100.0
	0.70	75.5	84.4	86.3	98.9	99.7	99.7	100.0	100.0	100.0

Table 4.6

Empirical power (%) of $\hat{\tau}$ for nominal level 0.10

Sample Size	C_1 A_1	$\sigma_\varepsilon^2 = 0.2$			$\sigma_\varepsilon^2 = 1.0$			$\sigma_\varepsilon^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	23.5	16.6	21.6	15.1	13.1	13.1	13.4	13.3	12.6
	0.99	25.5	17.3	24.4	16.2	13.5	14.9	14.9	14.4	14.5
	0.95	33.7	24.1	34.3	24.1	20.5	22.5	22.8	21.3	22.0
	0.90	40.6	31.1	41.6	34.0	28.6	30.5	33.9	31.8	31.6
	0.80	47.2	43.6	56.9	45.9	43.9	47.9	51.3	50.6	51.7
	0.70	49.8	49.0	66.9	53.2	55.1	59.5	63.4	64.6	65.2
50	1.00	15.6	12.8	17.5	11.9	11.1	12.5	11.7	11.6	11.4
	0.99	19.0	14.0	19.6	15.4	13.8	15.1	15.6	13.9	14.3
	0.95	32.3	25.2	34.0	27.7	27.2	29.8	29.5	28.7	28.0
	0.90	45.2	37.7	45.1	47.0	46.0	48.9	52.4	52.3	51.0
	0.80	56.5	52.2	59.1	67.3	69.6	73.2	79.9	79.4	79.3
	0.70	59.0	58.3	68.9	72.2	77.0	78.7	86.7	87.0	86.8
100	1.00	12.8	12.6	21.5	11.4	10.4	12.2	11.5	10.3	10.9
	0.99	18.3	18.6	31.2	16.4	16.8	18.8	16.5	16.7	16.4
	0.95	43.7	46.5	62.8	50.2	50.9	55.0	53.4	53.0	53.0
	0.90	61.4	64.8	77.0	80.3	81.9	85.3	86.5	86.9	86.4
	0.80	68.8	73.6	77.6	92.4	94.9	96.1	98.2	98.5	98.3
	0.70	70.8	73.2	78.7	91.7	95.3	96.1	98.6	98.9	98.7
250	1.00	9.7	10.6	15.8	10.0	9.9	10.9	10.0	10.0	10.3
	0.99	24.7	27.1	38.6	27.4	27.4	29.1	27.9	27.4	28.0
	0.95	76.8	83.0	90.1	94.2	94.7	95.3	96.5	96.0	96.0
	0.90	87.3	92.9	95.7	99.7	99.8	99.9	100.0	100.0	100.0
	0.80	84.8	92.5	93.7	99.8	99.9	99.9	100.0	100.0	100.0
	0.70	80.8	89.4	90.3	99.3	99.8	99.8	100.0	100.0	100.0

Table 4.7

Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.01

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	17.0	19.5	23.6	6.9	7.5	9.5	3.6	4.2	4.4
	0.99	18.9	20.0	26.6	7.4	7.7	10.1	3.8	4.7	5.3
	0.95	26.3	28.6	37.9	11.6	11.2	15.3	6.5	6.4	7.4
	0.90	34.7	38.8	48.6	17.7	18.0	21.2	10.6	10.5	11.8
	0.80	45.9	54.6	68.9	28.6	30.9	39.7	19.7	21.1	23.0
	0.70	51.7	65.0	79.3	39.9	45.5	55.7	32.6	33.7	36.4
50	1.00	11.5	10.3	12.0	4.2	3.8	4.7	2.5	2.6	3.0
	0.99	14.4	13.0	14.2	5.4	5.2	5.4	3.2	3.3	3.1
	0.95	25.4	23.7	27.4	10.7	9.5	11.7	6.5	6.4	7.1
	0.90	42.8	41.4	46.3	22.0	21.6	25.1	15.6	15.8	16.4
	0.80	62.3	65.0	72.9	47.6	48.3	53.9	41.0	40.9	43.2
	0.70	70.3	76.4	82.8	64.7	66.1	72.6	62.6	62.5	63.9
100	1.00	6.4	5.4	6.4	2.6	2.4	2.7	1.9	1.6	1.8
	0.99	9.8	8.3	9.7	3.7	3.6	3.7	2.4	2.6	2.5
	0.95	29.9	27.9	30.2	15.3	14.7	16.2	11.9	11.5	12.0
	0.90	58.1	57.3	60.5	41.3	40.8	45.0	36.7	35.9	36.7
	0.80	80.6	83.5	86.7	79.0	81.2	84.5	81.3	81.5	81.8
	0.70	86.0	89.6	93.2	88.9	91.0	92.6	92.7	93.3	93.2
250	1.00	2.9	2.2	3.1	1.4	1.5	1.3	1.3	1.2	1.1
	0.99	7.4	6.2	9.4	4.4	4.2	5.0	3.8	3.8	4.0
	0.95	54.4	51.6	59.8	50.5	50.6	52.3	49.1	49.3	49.2
	0.90	83.1	84.7	88.8	93.9	94.5	95.2	96.3	96.4	96.0
	0.80	91.3	93.7	94.6	99.3	99.7	99.7	100.0	100.0	100.0
	0.70	92.0	94.5	95.3	99.0	99.7	99.6	100.0	100.0	100.0

Table 4.8

Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.05

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	25.8	27.8	31.2	13.6	13.9	15.9	8.9	9.4	10.2
	0.99	29.0	29.2	34.4	14.7	14.5	16.9	10.0	10.3	11.1
	0.95	39.7	39.5	47.7	22.1	21.5	26.0	15.5	15.3	17.0
	0.90	49.5	51.8	59.1	32.4	31.0	35.5	24.4	23.0	24.5
	0.80	62.6	67.4	76.8	49.0	49.7	56.9	40.2	40.5	43.5
	0.70	67.8	75.4	84.2	60.5	64.3	70.9	55.3	56.6	58.6
50	1.00	18.5	16.3	18.5	9.5	9.4	10.1	7.4	7.5	7.6
	0.99	22.7	20.4	21.5	12.2	11.7	12.3	9.5	9.0	9.3
	0.95	38.5	36.0	39.6	23.4	22.4	24.1	18.9	18.6	18.9
	0.90	57.7	55.7	58.5	42.5	41.9	44.3	37.8	37.3	37.8
	0.80	73.8	74.3	78.9	68.8	69.9	73.4	69.1	69.1	69.3
	0.70	78.2	81.2	86.3	78.8	80.0	83.1	82.2	82.4	82.9
100	1.00	13.0	11.3	13.5	7.7	6.8	7.8	6.6	5.6	6.3
	0.99	18.9	17.0	20.9	10.7	10.9	11.9	9.0	9.5	9.4
	0.95	48.8	47.1	52.5	37.6	38.0	39.8	34.5	34.8	35.0
	0.90	74.4	74.3	78.7	72.0	72.2	75.3	72.2	72.0	72.3
	0.80	86.9	89.4	91.3	92.4	94.4	95.0	95.9	96.5	96.3
	0.70	89.1	92.2	94.7	94.2	95.9	96.5	98.0	98.3	98.1
250	1.00	8.0	6.7	8.6	5.9	5.7	5.9	5.5	5.4	5.3
	0.99	19.7	17.7	23.0	15.8	15.9	17.0	15.3	15.2	15.7
	0.95	79.8	78.4	84.1	86.1	86.5	86.9	87.7	87.3	86.8
	0.90	90.6	93.3	95.2	99.5	99.6	99.7	99.9	99.9	99.9
	0.80	93.3	95.4	96.0	99.8	99.9	99.9	100.0	100.0	100.0
	0.70	93.2	95.5	95.9	99.4	99.8	99.8	100.0	100.0	100.0

Table 4.9

Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.10

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	31.7	33.3	36.0	18.5	19.1	20.8	14.0	14.6	14.8
	0.99	35.4	34.9	40.1	20.3	20.5	22.9	15.5	15.8	16.7
	0.95	46.6	46.1	54.1	29.9	29.4	34.0	23.7	23.0	24.9
	0.90	57.9	58.9	64.8	42.7	40.9	45.1	35.4	33.8	35.4
	0.80	69.9	74.1	80.0	60.2	60.2	66.4	53.8	54.0	56.5
	0.70	73.8	79.4	86.3	70.0	72.5	76.9	67.5	68.5	69.7
50	1.00	23.8	21.1	23.5	15.1	13.9	15.1	12.6	12.4	12.6
	0.99	28.8	26.7	28.4	18.5	17.7	18.8	16.0	15.3	15.7
	0.95	47.3	44.6	48.0	34.3	33.3	35.2	30.4	29.8	29.8
	0.90	65.9	63.5	65.5	55.9	56.8	57.9	53.8	54.3	53.8
	0.80	77.8	77.9	81.6	78.4	79.5	81.5	81.9	81.8	82.4
	0.70	80.5	83.2	87.4	83.8	84.9	86.7	88.7	88.8	89.1
100	1.00	18.9	16.7	20.1	12.7	12.0	13.1	11.5	10.8	11.3
	0.99	26.5	24.5	29.4	18.5	19.0	19.9	16.6	17.0	17.1
	0.95	61.1	60.0	66.5	55.3	55.0	57.2	53.8	53.1	54.0
	0.90	80.8	82.2	85.9	85.3	86.3	88.0	87.5	87.8	87.8
	0.80	89.0	91.5	93.0	95.6	97.1	97.2	98.6	98.9	98.8
	0.70	90.3	93.1	95.3	95.6	97.1	97.5	99.0	99.2	99.0
250	1.00	12.7	11.6	14.5	10.8	10.6	11.0	10.1	10.2	10.3
	0.99	31.3	29.1	35.4	28.8	28.4	29.2	27.6	27.3	28.5
	0.95	85.5	88.1	91.7	95.7	96.0	96.2	96.9	96.8	96.5
	0.90	92.9	95.0	96.6	99.9	99.9	99.9	100.0	100.0	100.0
	0.80	94.0	96.0	96.6	99.9	99.9	99.9	100.0	100.0	100.0
	0.70	93.6	95.8	96.1	99.6	99.9	99.8	100.0	100.0	100.0

Table 4.10

Empirical power (%) of $\hat{\tau}$ for nominal level 0.01

Sample Size	C_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	21.1	26.3	31.0	8.1	9.5	14.1	4.0	4.6	5.5
	0.99	23.0	27.2	34.5	8.8	9.8	15.0	4.2	5.1	6.3
	0.95	32.7	37.5	47.2	13.7	14.5	23.0	7.3	7.3	9.4
	0.90	42.4	49.6	58.4	20.7	22.2	31.9	11.7	11.7	14.2
	0.80	57.6	66.2	74.8	33.7	37.9	52.3	21.5	23.1	27.1
	0.70	64.9	74.9	82.3	46.6	54.0	66.8	35.1	36.9	41.2
50	1.00	13.9	14.8	14.3	4.7	4.7	8.0	2.8	2.7	3.4
	0.99	17.6	18.6	17.2	6.1	6.4	9.3	3.4	3.5	3.8
	0.95	30.8	32.5	30.8	12.1	11.9	19.0	7.1	7.1	8.1
	0.90	50.8	52.6	47.6	25.1	26.5	36.4	17.1	17.4	18.8
	0.80	70.5	71.9	71.0	52.9	55.7	66.7	43.2	43.4	47.2
	0.70	77.2	78.5	80.7	70.2	73.0	78.6	64.8	64.9	68.0
100	1.00	8.1	9.0	6.8	3.0	2.9	5.4	2.0	1.8	2.1
	0.99	12.3	13.4	10.9	3.9	4.5	7.6	2.60	2.7	2.6
	0.95	37.4	38.5	28.3	16.9	17.9	26.9	12.4	12.3	13.2
	0.90	67.0	62.0	49.8	44.9	47.5	60.9	37.8	37.5	39.5
	0.80	83.2	80.5	78.5	83.4	86.7	85.8	82.4	82.8	84.6
	0.70	86.6	87.2	89.9	91.2	91.9	89.8	93.6	94.1	94.4
250	1.00	3.5	3.3	4.1	1.5	1.6	2.5	1.3	1.2	1.1
	0.99	9.1	8.5	11.8	4.7	4.8	8.2	4.0	3.8	4.2
	0.95	59.8	43.6	62.7	52.9	54.6	64.7	50.5	51.0	51.6
	0.90	80.4	72.7	84.8	94.7	96.1	91.9	96.4	96.7	96.6
	0.80	87.8	89.7	89.5	99.1	99.2	97.5	100.0	100.0	100.0
	0.70	89.5	93.0	93.2	98.4	99.2	98.3	100.0	100.0	100.0

Table 4.11

Empirical power (%) of $\hat{\tau}$ for nominal level 0.05

Sample Size	C_1 A_1	$\sigma_\epsilon^2 = 0.2$			$\sigma_\epsilon^2 = 1.0$			$\sigma_\epsilon^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	29.8	33.9	35.9	15.0	16.6	20.6	9.6	10.1	11.4
	0.99	33.0	35.4	40.0	16.1	17.3	22.6	10.5	11.1	12.5
	0.95	44.9	47.1	52.6	24.5	25.3	33.4	16.6	16.4	19.1
	0.90	56.2	59.9	63.3	36.1	36.2	44.8	26.0	25.0	27.6
	0.80	70.2	73.6	77.5	54.0	56.1	65.4	42.6	43.2	47.6
	0.70	75.0	79.2	84.1	66.0	69.6	76.2	57.8	59.4	62.7
50	1.00	21.0	20.6	20.6	10.4	10.6	13.8	7.7	7.7	8.3
	0.99	25.9	25.8	24.1	13.2	13.3	16.7	9.9	9.4	10.0
	0.95	43.8	43.2	40.8	25.5	25.5	32.0	19.8	19.4	20.6
	0.90	63.4	61.3	56.5	45.6	47.3	53.7	39.7	39.2	40.7
	0.80	77.3	75.1	76.0	72.9	75.0	76.9	70.7	70.9	72.5
	0.70	80.0	80.4	83.9	81.7	82.7	82.9	83.4	83.6	84.9
100	1.00	15.1	14.7	16.1	8.1	7.8	10.7	6.7	5.9	6.6
	0.99	22.1	21.1	24.2	11.5	12.4	16.1	9.2	9.9	10.0
	0.95	55.2	51.1	52.5	40.2	42.2	48.4	35.5	36.0	37.0
	0.90	77.1	71.1	72.8	74.8	77.2	79.5	73.2	73.2	74.3
	0.80	85.5	86.2	87.7	93.1	93.9	92.0	96.2	96.6	96.9
	0.70	87.9	90.5	93.2	93.9	94.8	94.8	98.1	98.4	98.2
250	1.00	8.2	7.3	10.8	6.1	6.0	7.3	5.5	5.4	5.4
	0.99	21.3	18.1	28.3	16.4	16.9	20.8	15.5	15.3	16.3
	0.95	75.0	68.0	84.7	86.6	87.8	87.3	87.8	87.3	87.3
	0.90	86.2	87.4	92.8	99.4	99.5	97.2	99.9	99.9	99.8
	0.80	91.4	93.4	93.0	99.6	99.7	99.0	100.0	100.0	100.0
	0.70	91.8	94.5	94.8	99.2	99.6	99.1	100.0	100.0	100.0

Table 4.12

Empirical power (%) of $\hat{\tau}$ for nominal level 0.10

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	35.4	38.6	39.8	20.2	21.8	25.3	14.7	15.4	16.6
	0.99	39.1	40.4	44.1	22.2	23.2	28.0	16.5	16.9	18.4
	0.95	51.4	52.5	57.1	32.8	33.7	40.0	25.2	24.8	27.8
	0.90	63.3	64.9	67.0	46.3	46.3	53.3	37.6	36.3	39.0
	0.80	75.1	76.4	79.6	64.6	65.6	72.1	56.8	56.8	60.7
	0.70	78.5	80.7	85.4	74.1	76.0	80.1	69.4	70.7	73.0
50	1.00	26.3	24.9	25.5	15.8	15.3	18.5	13.0	12.8	13.5
	0.99	32.1	30.9	30.3	19.6	19.6	22.8	16.2	15.9	16.9
	0.95	52.3	50.0	40.1	36.7	36.9	41.9	31.5	31.2	32.1
	0.90	69.4	65.6	62.4	59.2	60.9	63.8	55.5	56.0	56.5
	0.80	79.5	76.7	79.4	80.8	82.2	81.7	82.8	83.0	84.1
	0.70	81.0	81.8	85.8	84.8	85.3	85.2	89.2	89.4	89.9
100	1.00	21.2	19.4	23.7	13.4	13.2	15.5	12.0	10.8	11.7
	0.99	29.8	27.7	34.3	19.5	20.8	24.1	17.3	17.8	18.0
	0.95	64.9	60.3	67.4	57.6	59.1	62.9	55.2	55.1	56.0
	0.90	80.6	78.0	82.7	87.2	88.3	87.4	88.0	88.5	88.7
	0.80	87.2	89.4	90.6	95.4	96.6	95.2	98.6	98.9	98.9
	0.70	88.9	92.1	94.3	95.1	96.3	96.4	98.8	99.1	98.9
250	1.00	12.8	11.7	17.1	10.8	11.1	12.6	10.3	10.3	10.6
	0.99	32.0	28.3	41.0	29.5	29.8	32.9	28.4	27.9	29.0
	0.95	81.4	80.6	91.3	95.7	96.1	94.1	96.7	96.3	96.4
	0.90	89.5	91.5	94.7	99.8	99.8	98.6	100.0	100.0	100.0
	0.80	92.5	94.6	94.4	99.7	99.8	99.4	100.0	100.0	100.0
	0.70	92.6	95.1	95.3	99.4	99.7	99.4	100.0	100.0	100.0

Table 4.13

Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.01

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	16.5	18.3	20.8	6.6	7.5	8.7	3.6	4.1	4.3
	0.99	18.1	21.8	28.8	7.3	8.2	10.3	3.8	4.7	5.3
	0.95	25.6	29.5	39.2	11.3	11.5	14.9	6.6	6.5	7.4
	0.90	32.3	39.3	49.2	17.2	18.0	20.8	10.5	10.6	11.7
	0.80	44.5	55.7	69.9	27.8	31.3	38.5	20.0	21.4	23.0
	0.70	50.9	65.9	80.1	39.7	45.5	54.3	33.2	34.0	36.2
50	1.00	11.2	10.2	12.5	3.9	3.8	4.4	2.4	2.5	2.5
	0.99	14.2	14.2	18.4	5.5	5.1	5.6	3.4	3.2	3.5
	0.95	25.8	25.2	32.0	10.7	10.0	11.2	7.1	6.5	6.9
	0.90	41.7	42.7	51.0	22.0	21.3	23.2	16.1	16.0	16.1
	0.80	63.1	67.6	75.3	47.3	47.5	51.0	41.6	41.5	42.9
	0.70	71.3	78.3	85.1	64.7	66.7	70.1	62.8	62.9	64.5
100	1.00	6.0	5.2	7.4	2.5	2.3	2.4	1.8	1.6	1.7
	0.99	9.3	8.1	11.1	3.6	3.4	3.5	2.5	2.7	2.5
	0.95	28.3	27.1	32.7	15.2	14.2	15.2	11.9	11.5	11.8
	0.90	56.7	55.8	62.6	40.5	39.6	41.2	36.7	35.9	36.5
	0.80	80.3	84.9	87.9	79.2	80.4	82.2	81.4	81.6	81.8
	0.70	86.7	91.6	93.5	88.9	91.0	92.5	92.8	93.3	93.2
250	1.00	2.4	2.0	2.5	1.3	1.4	1.2	1.3	1.1	1.0
	0.99	6.5	5.9	7.5	4.4	4.3	4.7	3.8	3.8	4.0
	0.95	51.5	51.0	55.0	50.5	50.5	50.6	49.2	49.3	49.1
	0.90	83.5	87.2	90.0	93.9	94.6	95.2	96.3	96.4	96.1
	0.80	92.1	94.7	96.0	99.3	99.7	99.7	100.0	100.0	100.0
	0.70	93.4	95.8	96.4	99.0	99.2	99.7	100.0	100.0	100.0

Table 4.14

Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.05

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	25.4	26.6	29.0	13.2	13.9	15.2	8.8	9.4	10.0
	0.99	28.2	30.6	37.3	14.6	14.9	16.7	10.1	10.5	11.0
	0.95	38.7	40.3	50.1	22.0	21.7	25.2	15.8	15.4	16.8
	0.90	48.3	52.9	61.9	32.3	30.6	34.9	24.5	23.3	24.5
	0.80	62.8	68.4	79.2	48.7	49.9	55.4	40.8	41.2	43.8
	0.70	67.5	76.6	86.0	61.1	64.6	70.0	56.3	57.6	59.2
50	1.00	18.0	17.8	20.0	9.2	9.1	9.8	7.4	7.2	7.4
	0.99	22.1	22.1	27.4	12.2	11.3	12.2	9.6	9.1	9.2
	0.95	38.3	38.0	45.2	23.2	22.5	23.5	19.1	18.5	18.9
	0.90	57.2	58.1	64.5	42.0	41.7	43.1	37.5	37.7	37.7
	0.80	74.5	78.0	83.0	69.3	69.6	72.0	68.8	68.8	69.8
	0.70	79.1	84.2	89.0	79.6	81.8	83.6	82.2	82.7	83.1
100	1.00	12.3	10.8	13.7	7.5	6.6	7.3	6.5	5.5	6.2
	0.99	18.0	16.8	20.7	10.6	10.7	11.1	9.0	9.5	9.3
	0.95	47.4	46.6	51.5	37.3	37.3	38.2	34.6	34.8	34.9
	0.90	74.2	74.5	79.3	71.8	71.9	73.2	72.2	72.0	72.2
	0.80	87.8	91.4	92.8	92.5	94.1	95.1	96.0	96.5	96.3
	0.70	90.6	94.0	95.2	94.3	96.3	96.6	98.0	98.3	98.1
250	1.00	7.1	6.5	7.3	5.8	5.6	5.9	5.4	5.3	5.2
	0.99	17.7	17.7	19.7	16.0	15.8	16.6	15.3	15.2	15.8
	0.95	76.8	79.6	82.6	86.4	86.6	86.5	87.7	87.3	86.9
	0.90	91.7	94.6	95.9	99.6	99.6	99.8	99.9	99.9	99.9
	0.80	94.2	96.1	97.1	99.8	99.9	99.9	100.0	100.0	100.0
	0.70	94.5	96.5	97.1	99.6	99.9	99.9	100.0	100.0	100.0

Table 4.16

Empirical power (%) of $\hat{\tau}$ for nominal level 0.01

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	11.8	6.5	11.5	5.2	3.1	3.3	3.3	3.1	2.4
	0.99	12.9	6.9	14.3	5.5	3.2	4.0	3.4	3.4	2.9
	0.95	17.1	9.2	19.1	8.7	4.5	5.5	6.0	4.5	4.2
	0.90	20.7	12.1	24.5	12.2	6.8	7.9	9.2	7.8	6.8
	0.80	25.7	17.9	39.2	18.7	12.7	15.0	17.6	16.0	14.7
	0.70	26.9	23.2	48.6	25.0	19.5	22.1	27.4	26.1	25.0
50	1.00	6.6	2.2	3.4	2.5	1.3	1.5	2.1	2.0	1.7
	0.99	8.0	3.4	5.1	3.7	2.1	2.2	3.1	2.5	2.4
	0.95	14.1	6.0	8.9	6.8	4.5	4.8	6.3	5.2	4.5
	0.90	23.1	11.2	15.4	13.4	10.0	9.2	13.7	12.8	11.8
	0.80	33.3	20.2	30.9	29.8	24.8	25.0	36.9	35.4	33.5
	0.70	36.5	26.2	43.9	41.0	39.0	38.5	56.4	56.1	54.9
100	1.00	2.6	1.1	2.5	1.7	1.1	1.3	1.6	1.4	1.4
	0.99	3.8	1.6	4.1	2.2	1.6	2.0	2.2	2.0	1.9
	0.95	11.8	6.4	10.9	9.3	7.9	8.9	10.5	9.6	9.4
	0.90	24.9	15.5	21.5	25.7	24.0	25.9	33.2	31.9	31.1
	0.80	42.9	32.0	36.5	56.6	59.5	62.0	76.8	77.1	75.5
	0.70	50.1	40.2	46.3	67.5	73.6	77.6	89.7	90.6	90.4
250	1.00	1.0	0.8	2.0	1.0	1.0	0.9	1.1	1.0	1.0
	0.99	2.9	2.8	6.0	3.2	2.9	3.5	3.8	3.4	3.6
	0.95	24.6	27.8	41.9	40.8	41.8	43.0	47.8	47.7	46.4
	0.90	48.2	60.6	75.2	87.8	90.1	90.9	95.5	95.5	95.2
	0.80	60.6	71.0	81.7	97.9	99.1	99.3	100.0	100.0	100.0
	0.70	63.3	69.2	79.2	98.0	99.2	99.4	100.0	100.0	100.0

Table 4.17

Empirical power (%) of $\hat{\tau}$ for nominal level 0.05

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	17.6	11.4	16.2	10.0	7.8	7.8	8.2	8.0	7.5
	0.99	19.0	12.2	19.7	10.7	7.9	8.6	9.4	8.9	8.5
	0.95	24.9	16.5	26.9	15.9	11.7	13.3	14.6	13.0	12.7
	0.90	30.4	21.7	33.2	23.3	17.4	18.4	22.7	20.2	19.2
	0.80	36.2	30.1	48.7	32.5	28.6	31.2	37.2	36.1	36.0
	0.70	37.8	36.5	58.1	39.8	39.0	41.8	50.2	51.2	50.0
50	1.00	10.7	6.9	9.6	6.6	5.6	6.2	6.8	6.4	6.2
	0.99	13.3	8.4	12.0	9.2	7.4	7.7	8.6	7.9	7.7
	0.95	22.4	14.8	20.2	16.6	14.7	15.1	18.0	16.6	15.8
	0.90	34.4	24.3	30.4	30.2	28.5	29.3	35.4	34.8	33.1
	0.80	46.1	36.7	46.6	52.5	52.0	53.4	65.5	64.9	63.9
	0.70	48.5	43.6	56.9	60.4	63.2	66.3	78.4	78.9	78.3
100	1.00	6.6	5.0	9.4	5.8	4.9	5.7	6.0	5.2	5.6
	0.99	9.5	8.3	16.0	8.3	8.3	8.8	8.5	8.7	8.5
	0.95	25.9	24.6	36.7	30.0	29.9	31.6	33.3	33.3	32.6
	0.90	43.6	42.5	54.7	60.3	61.5	63.9	69.9	69.7	69.1
	0.80	57.4	58.4	64.0	83.9	88.0	89.7	94.8	95.3	95.1
	0.70	61.8	60.8	68.0	85.6	91.2	93.2	97.3	98.0	97.5
250	1.00	4.3	4.7	7.0	5.0	5.0	5.3	5.2	5.1	4.9
	0.99	11.1	13.0	18.5	14.1	14.1	15.2	15.1	14.7	15.4
	0.95	58.8	66.4	74.9	82.5	83.0	83.6	86.9	86.5	85.8
	0.90	77.7	87.5	91.7	98.8	99.2	99.5	99.8	99.8	99.8
	0.80	77.8	88.7	91.4	99.5	99.8	99.9	100.0	100.0	100.0
	0.70	78.0	88.5	91.6	99.6	99.8	99.9	100.0	100.0	100.0

Table 4.18

Empirical power (%) of $\hat{\tau}$ for nominal level 0.10

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	22.5	16.5	21.3	14.6	12.9	12.7	13.3	13.2	12.3
	0.99	24.1	17.6	25.4	15.9	13.6	14.5	15.0	14.5	14.4
	0.95	31.6	23.4	33.8	23.7	20.6	21.7	22.9	21.3	21.7
	0.90	37.9	30.7	41.0	33.1	27.9	29.7	34.0	31.9	31.4
	0.80	44.8	40.3	56.1	44.4	42.8	45.8	63.8	50.9	51.4
	0.70	46.3	46.3	64.3	51.7	53.1	56.2	70.8	64.9	64.8
50	1.00	15.3	12.0	17.0	11.7	10.6	11.6	11.5	11.5	11.2
	0.99	18.4	14.8	20.4	14.9	13.2	14.2	15.1	13.9	14.3
	0.95	30.5	24.9	32.4	27.5	26.5	27.4	29.4	28.6	28.2
	0.90	43.3	37.5	44.1	46.0	45.4	46.6	52.5	51.9	50.9
	0.80	54.5	49.1	57.7	66.8	68.7	70.1	79.3	79.4	79.5
	0.70	56.1	54.8	65.9	71.2	74.9	77.8	86.7	87.5	86.6
100	1.00	11.9	11.4	18.4	11.2	10.1	11.1	11.4	10.2	10.8
	0.99	17.1	17.2	26.4	16.2	16.4	17.3	16.5	16.7	16.4
	0.95	41.0	43.6	55.3	49.7	50.1	52.0	53.4	52.9	52.9
	0.90	59.0	62.7	73.1	80.0	81.4	82.8	86.5	86.9	86.4
	0.80	67.2	73.1	77.1	92.4	94.9	95.7	98.2	98.5	98.3
	0.70	69.1	73.3	77.6	92.6	95.3	96.0	98.6	98.9	98.7
250	1.00	8.9	9.5	12.4	10.0	9.9	10.5	10.0	10.0	10.2
	0.99	22.1	24.9	31.1	27.3	27.3	27.9	27.9	27.4	28.0
	0.95	75.5	81.9	86.2	94.2	94.7	95.0	96.5	96.0	96.0
	0.90	87.4	92.9	94.6	99.7	99.8	99.9	100.0	100.0	100.0
	0.80	87.5	92.8	93.8	99.8	99.9	99.9	100.0	100.0	100.0
	0.70	87.7	91.3	91.6	99.6	99.8	99.9	100.0	100.0	100.0

Table 4.19

Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.01

Sample Size	C_1 A_1	$\sigma_\epsilon^2 = 0.2$			$\sigma_\epsilon^2 = 1.0$			$\sigma_\epsilon^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	17.0	19.5	23.6	6.8	7.4	9.5	3.6	4.1	4.3
	0.99	18.9	20.1	26.7	7.5	7.7	10.2	3.8	4.7	5.3
	0.95	26.4	28.6	38.1	11.6	11.3	15.4	6.6	6.4	7.4
	0.90	34.9	38.9	48.8	17.7	18.0	21.3	10.7	10.5	11.8
	0.80	46.0	54.8	69.1	28.7	30.9	39.7	19.8	21.1	23.0
	0.70	52.0	65.3	79.4	39.9	45.6	55.7	32.6	33.8	36.5
50	1.00	11.5	10.3	12.0	4.1	3.8	4.7	2.5	2.5	2.9
	0.99	14.5	13.1	14.3	5.4	5.2	5.5	3.2	3.3	3.2
	0.95	25.4	23.7	28.0	10.7	9.5	11.7	6.5	6.4	7.2
	0.90	42.8	41.3	46.6	22.1	21.6	25.1	15.6	15.9	16.4
	0.80	62.3	65.8	72.9	47.7	48.3	53.9	41.0	40.9	43.2
	0.70	70.6	76.4	83.0	64.6	66.8	72.6	62.6	62.6	63.9
100	1.00	6.3	5.3	6.1	2.6	2.4	2.7	1.9	1.6	1.8
	0.99	9.8	8.4	9.5	3.7	3.6	3.7	2.5	2.7	2.6
	0.95	29.8	27.9	29.8	15.4	14.7	16.2	11.9	11.6	12.0
	0.90	58.1	57.2	59.7	41.3	40.8	44.9	36.8	35.9	36.7
	0.80	80.6	83.5	86.0	79.0	81.3	84.1	81.3	81.6	81.8
	0.70	86.0	89.6	93.1	88.9	91.3	93.0	92.7	93.3	93.2
250	1.00	2.8	2.1	2.9	1.4	1.4	1.2	1.3	1.1	1.1
	0.99	7.4	6.2	9.1	4.4	4.3	5.0	3.8	3.8	4.0
	0.95	54.1	51.6	58.6	50.5	50.6	52.3	49.2	49.3	49.2
	0.90	83.3	84.4	88.3	93.9	94.5	95.2	96.3	96.4	96.0
	0.80	91.5	93.9	94.8	99.3	99.7	99.8	100.0	100.0	100.0
	0.70	92.2	94.7	95.5	99.1	99.7	99.7	100.0	100.0	100.0

Table 4.20

Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.05

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	25.7	27.7	31.1	13.5	13.9	15.8	8.8	9.3	10.1
	0.99	28.9	29.2	34.4	14.7	14.5	16.9	10.1	10.3	11.1
	0.95	39.7	39.5	47.6	22.1	21.5	26.1	15.6	15.3	17.0
	0.90	49.6	52.0	59.1	32.4	31.0	35.5	24.5	23.0	24.5
	0.80	62.5	67.5	76.8	49.1	49.7	56.8	40.2	40.5	43.6
	0.70	67.8	75.7	84.1	60.5	64.5	70.9	55.3	56.6	58.7
50	1.00	18.4	16.3	18.2	9.5	9.3	10.1	7.3	7.4	7.5
	0.99	22.7	20.3	21.2	12.2	11.7	12.3	9.5	9.0	9.3
	0.95	38.4	36.2	39.7	23.4	22.4	24.2	18.9	18.6	18.9
	0.90	57.6	56.1	58.6	42.5	41.8	44.3	37.8	37.3	37.8
	0.80	73.2	74.5	79.0	68.9	69.8	73.2	69.1	69.1	69.4
	0.70	78.4	81.5	86.2	79.2	80.8	82.9	82.2	82.4	82.9
100	1.00	12.9	11.2	12.8	7.6	6.8	7.8	6.5	5.6	6.3
	0.99	18.8	17.0	19.9	10.7	11.0	11.9	9.0	9.5	9.4
	0.95	48.7	47.1	50.8	37.6	38.0	39.8	34.6	34.8	35.0
	0.90	74.2	74.3	77.5	72.0	72.2	75.3	72.2	72.1	72.3
	0.80	86.8	89.4	91.2	92.4	94.4	94.9	95.9	96.5	96.3
	0.70	89.2	92.2	94.9	94.2	95.9	96.7	98.0	98.3	98.1
250	1.00	7.9	6.6	8.1	5.8	5.7	5.8	5.5	5.4	5.2
	0.99	19.8	17.8	22.0	15.9	15.9	17.0	15.3	15.2	15.8
	0.95	79.7	79.2	82.6	86.1	86.5	86.9	87.7	87.3	86.8
	0.90	90.8	93.7	95.9	99.5	99.6	99.7	99.9	99.9	99.9
	0.80	93.4	95.5	96.7	99.8	99.9	99.9	100.0	100.0	100.0
	0.70	93.5	95.7	96.4	99.4	99.8	99.9	100.0	100.0	100.0

Table 4.21

Empirical power (%) of $n(\hat{A}_1 - 1)$ for nominal level 0.10

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	31.5	33.2	35.7	18.4	19.1	20.7	13.9	14.6	14.7
	0.99	35.2	34.8	39.9	20.3	20.5	22.9	15.6	15.9	16.7
	0.95	46.6	46.0	53.6	29.9	29.4	34.1	23.8	23.1	24.9
	0.90	57.8	59.0	64.3	42.6	40.9	45.1	35.4	33.9	35.4
	0.80	69.9	74.1	80.0	60.2	60.2	66.3	53.8	54.1	56.5
	0.70	73.7	79.6	86.2	70.0	72.5	76.8	67.6	68.5	69.7
50	1.00	23.7	21.0	22.8	15.1	13.9	15.1	12.6	12.3	12.6
	0.99	28.8	26.7	27.5	18.5	17.7	18.9	16.0	15.3	15.7
	0.95	47.1	44.6	47.3	34.2	33.2	35.2	30.4	29.9	29.8
	0.90	65.8	63.6	65.4	55.9	56.8	58.0	53.8	54.4	53.8
	0.80	77.8	78.0	81.5	78.4	79.5	81.3	81.9	81.8	82.4
	0.70	80.7	83.4	87.4	84.1	85.4	86.3	88.8	88.8	89.2
100	1.00	18.8	16.5	19.0	12.6	11.9	13.0	11.5	10.7	11.3
	0.99	26.3	24.4	29.4	18.5	19.0	19.9	16.6	17.0	17.1
	0.95	60.7	59.6	66.5	55.4	54.9	57.1	53.8	53.1	54.0
	0.90	80.4	81.9	85.9	85.3	86.3	88.0	87.5	87.9	87.8
	0.80	88.9	91.4	93.0	95.9	97.1	97.3	98.6	99.0	98.8
	0.70	90.5	93.2	95.5	95.7	97.2	97.7	99.0	99.2	99.0
250	1.00	12.6	11.5	13.8	10.7	10.6	10.9	10.0	10.1	10.2
	0.99	31.5	29.1	34.0	28.9	28.4	29.2	27.7	27.4	28.6
	0.95	85.5	88.4	91.0	95.7	96.1	96.2	96.9	96.8	96.5
	0.90	92.9	96.1	97.9	99.9	99.9	99.9	100.0	100.0	100.0
	0.80	94.1	96.3	97.5	99.9	100.0	100.0	100.0	100.0	100.0
	0.70	93.8	96.0	97.1	99.6	99.8	99.9	100.0	100.0	100.0

Table 4.22

Empirical power (%) of $\hat{\tau}$ for nominal level 0.01

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	20.8	26.1	30.2	8.0	9.4	14.0	3.9	4.6	5.5
	0.99	22.8	27.1	33.8	8.8	9.7	15.0	4.2	5.1	6.3
	0.95	32.6	37.2	46.3	13.7	14.5	22.9	7.3	7.3	9.4
	0.90	42.4	49.5	57.5	20.6	22.1	31.8	11.7	11.7	14.2
	0.80	57.3	66.0	74.5	33.7	37.9	52.3	21.5	23.1	27.1
	0.70	64.9	74.8	81.9	46.4	54.0	66.7	35.2	36.9	41.2
50	1.00	13.8	14.6	13.0	4.6	4.6	8.0	2.7	2.7	3.3
	0.99	17.6	18.4	15.9	6.1	6.4	9.3	3.4	3.5	3.8
	0.95	30.8	32.4	29.1	12.1	11.9	19.0	7.1	7.1	8.2
	0.90	50.8	52.2	46.7	25.1	26.5	36.4	17.1	17.4	18.8
	0.80	70.5	71.2	70.2	52.9	55.6	66.7	43.2	43.4	47.2
	0.70	77.2	78.5	80.6	70.1	73.5	78.6	64.8	64.9	68.0
100	1.00	8.0	8.7	5.7	2.9	2.9	5.3	1.9	1.7	2.0
	0.99	12.1	13.2	9.4	3.9	4.5	7.6	2.60	2.7	2.6
	0.95	37.1	38.1	25.5	16.9	17.9	26.9	12.4	12.3	13.2
	0.90	66.6	61.6	69.8	44.9	47.5	60.9	37.8	37.5	39.6
	0.80	82.9	80.5	76.6	82.9	86.9	85.7	82.4	82.8	84.6
	0.70	86.6	86.9	89.5	91.2	92.2	89.8	93.6	94.1	94.4
250	1.00	3.3	3.2	3.7	1.5	1.5	2.4	1.2	1.2	1.1
	0.99	9.7	8.5	11.1	4.7	4.8	8.3	4.1	3.9	4.3
	0.95	59.7	43.2	59.7	52.9	54.6	64.7	50.5	51.1	51.6
	0.90	80.6	71.2	82.5	94.7	96.1	91.9	96.4	96.7	96.6
	0.80	87.9	90.0	88.9	99.1	99.2	97.5	100.0	100.0	100.0
	0.70	89.7	93.2	93.4	98.6	99.2	98.3	100.0	100.0	100.0

Table 4.23

Empirical power (%) of $\hat{\tau}$ for nominal level 0.05

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	29.2	33.5	34.4	14.9	16.4	20.3	9.5	10.0	11.4
	0.99	32.5	35.1	38.6	16.0	17.2	22.5	10.5	11.1	12.5
	0.95	44.4	46.6	51.0	24.4	25.2	33.3	16.7	16.4	19.1
	0.90	55.8	59.5	61.7	35.8	36.0	44.6	26.0	25.0	27.6
	0.80	69.5	73.2	76.8	53.9	55.9	65.1	42.6	43.2	47.6
	0.70	74.4	78.7	83.4	65.9	69.4	75.9	57.8	59.4	62.7
50	1.00	20.7	20.2	18.3	10.3	10.5	13.7	7.6	7.6	8.2
	0.99	25.6	25.4	22.0	13.2	13.3	16.7	9.9	9.4	10.0
	0.95	43.2	43.0	38.0	25.5	25.4	32.0	19.9	19.5	20.6
	0.90	63.7	61.5	54.6	45.5	47.2	53.5	39.7	39.2	40.7
	0.80	76.8	74.8	74.6	72.9	74.9	76.7	70.7	70.9	72.5
	0.70	79.9	80.3	83.7	81.9	82.9	81.7	83.4	83.6	84.9
100	1.00	15.0	14.4	14.2	8.0	7.6	10.6	6.6	5.8	6.6
	0.99	22.0	20.6	21.4	11.6	12.5	16.1	9.2	9.9	10.0
	0.95	54.6	50.3	48.3	40.3	42.2	48.4	35.5	36.0	37.0
	0.90	76.5	70.4	69.8	74.7	77.2	79.4	73.2	73.2	74.3
	0.80	85.2	86.0	86.4	93.5	94.5	91.8	96.2	96.6	96.9
	0.70	88.0	90.3	92.8	93.6	95.0	94.5	98.1	98.4	98.2
250	1.00	8.2	7.1	9.9	6.0	5.9	7.3	5.4	5.4	5.3
	0.99	22.0	18.0	26.4	16.5	16.9	20.8	15.6	15.3	16.3
	0.95	75.2	67.3	81.4	86.5	87.8	87.4	87.8	87.3	87.3
	0.90	86.4	87.1	91.0	99.4	99.5	97.2	99.9	99.9	99.8
	0.80	91.6	93.5	93.0	99.6	99.7	99.0	100.0	100.0	100.0
	0.70	92.1	94.7	95.1	99.2	99.6	99.1	100.0	100.0	100.0

Table 4.24

Empirical power (%) of $\hat{\tau}$ for nominal level 0.10

Sample Size	C_1 A_1	$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
		-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
25	1.00	34.8	38.1	38.0	20.1	21.6	24.9	14.6	15.3	16.5
	0.99	38.5	39.9	42.2	22.1	23.1	27.8	16.5	16.9	18.4
	0.95	50.7	51.7	55.1	32.7	33.5	39.8	25.3	24.8	27.8
	0.90	62.8	64.4	65.0	46.0	46.1	52.9	37.6	36.3	39.0
	0.80	74.3	75.7	78.7	64.2	65.2	71.6	56.8	56.8	60.6
	0.70	77.6	80.1	84.7	73.7	75.7	79.5	69.4	70.6	72.9
50	1.00	26.0	24.4	23.0	15.8	15.3	18.4	12.9	12.8	13.4
	0.99	31.7	30.5	27.8	19.6	19.6	22.9	16.2	15.9	16.9
	0.95	51.6	49.9	45.2	36.6	36.8	41.8	31.5	31.3	32.1
	0.90	69.1	65.6	60.9	59.2	60.9	63.5	55.5	56.0	56.5
	0.80	78.6	76.8	78.1	80.8	82.1	81.0	82.8	83.0	84.1
	0.70	80.7	81.7	85.7	85.2	85.5	84.3	89.2	89.4	89.9
100	1.00	21.1	18.9	21.2	13.3	13.0	15.4	11.9	10.7	11.6
	0.99	29.6	27.1	31.0	19.5	20.8	24.1	17.3	17.8	18.0
	0.95	64.2	59.2	63.0	57.6	59.0	62.7	55.3	55.1	56.0
	0.90	80.0	77.1	79.7	87.1	88.2	87.3	88.1	88.5	88.6
	0.80	87.0	89.0	89.8	95.6	96.6	95.2	98.6	99.0	98.9
	0.70	89.0	91.9	94.1	94.9	96.3	96.2	98.9	99.1	98.9
250	1.00	12.5	11.5	15.8	10.7	11.1	12.5	10.2	10.3	10.6
	0.99	32.9	28.1	38.4	29.5	29.8	32.8	28.4	27.9	29.1
	0.95	81.2	80.2	88.2	95.6	96.1	94.2	96.7	96.3	96.4
	0.90	89.7	91.8	93.7	99.8	99.8	98.6	100.0	100.0	100.0
	0.80	92.5	94.9	94.6	99.7	99.8	99.4	100.0	100.0	100.0
	0.70	93.4	95.8	96.2	99.4	99.7	99.4	100.0	100.0	100.0

Table 4.25

Empirical cumulative distribution of $n(\hat{A}_1 - 1)$ for $A_1 = 1$ ($N = 25$)

(σ_e^2, C_1)	Probability of a Smaller Value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-46.78	-33.53	-25.30	-17.58	2.93	5.65	10.10	20.77
(0.2, 0.0)	-61.67	-42.89	-31.44	-21.93	2.84	5.63	10.47	20.42
(0.2, 0.5)	-57.06	-43.22	-35.27	-26.87	3.24	6.60	11.62	20.78
(1.0,-0.5)	-30.25	-20.47	-14.21	-9.32	1.63	2.51	3.57	6.13
(1.0, 0.0)	-32.31	-22.39	-15.72	-9.75	1.58	2.36	3.61	6.40
(1.0, 0.5)	-36.23	-25.48	-18.24	-11.42	1.58	2.42	3.65	6.26
(5.0,-0.5)	-21.78	-13.93	-10.15	-6.77	1.40	1.93	2.60	3.66
(5.0, 0.0)	-22.27	-15.42	-10.85	-7.02	1.39	1.95	2.57	3.57
(5.0, 0.5)	-22.89	-15.78	-11.13	-7.37	1.41	1.99	2.60	3.71

Table 4.26

Empirical cumulative distribution of $n(\hat{A}_1 - 1)$ for $A_1 = 1$ ($N = 50$)

(σ_e^2, C_1)	Probability of a Smaller Value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-58.75	-38.39	-25.78	-14.72	1.89	3.15	5.73	12.97
(0.2, 0.0)	-54.25	-35.53	-23.34	-13.30	1.84	3.05	6.03	11.93
(0.2, 0.5)	-61.03	-43.09	-29.56	-15.89	2.32	4.87	9.58	16.65
(1.0,-0.5)	-25.83	-16.61	-11.50	-7.45	1.19	1.67	2.21	2.90
(1.0, 0.0)	-24.29	-15.84	-11.03	-7.31	1.19	1.65	2.10	2.80
(1.0, 0.5)	-27.41	-17.84	-12.42	-7.75	1.22	1.70	2.25	3.08
(5.0,-0.5)	-18.78	-12.96	-9.44	-6.41	1.10	1.51	1.94	2.46
(5.0, 0.0)	-17.86	-13.02	-9.42	-6.32	1.12	1.54	1.91	2.46
(5.0, 0.5)	-18.63	-13.72	-9.76	-6.53	1.14	1.62	2.03	2.59

Table 4.27

Empirical cumulative distribution of $n(\hat{A}_1 - 1)$ for $A_1 = 1$ ($N = 100$)

(σ_e^2, C_1)	Probability of a Smaller Value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-38.05	-24.53	-15.73	-9.77	1.16	1.67	2.22	3.02
(0.2, 0.0)	-32.98	-20.88	-14.00	-8.76	1.12	1.62	2.13	2.79
(0.2, 0.5)	-33.37	-20.89	-15.05	-9.57	1.06	1.54	2.09	3.04
(1.0,-0.5)	-19.60	-13.61	-9.95	-6.67	1.00	1.42	1.80	2.32
(1.0, 0.0)	-18.52	-13.07	-9.28	-6.24	1.01	1.43	1.89	2.34
(1.0, 0.5)	-19.72	-13.76	-10.10	-6.76	1.00	1.39	1.74	2.24
(5.0,-0.5)	-16.90	-11.72	-9.02	-6.19	1.00	1.39	1.76	2.23
(5.0, 0.0)	-15.84	-11.44	-8.39	-5.88	1.01	1.40	1.84	2.35
(5.0, 0.5)	-15.89	-11.92	-8.87	-6.09	1.0	1.39	1.74	2.17

Table 4.28

Empirical cumulative distribution of $n(\hat{A}_1 - 1)$ for $A_1 = 1$ ($N = 250$)

(σ_e^2, C_1)	Probability of a Smaller Value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-20.53	-14.41	-10.22	-6.76	1.04	1.44	1.87	2.39
(0.2, 0.0)	-17.86	-12.92	-9.46	-6.33	1.03	1.43	1.83	2.35
(0.2, 0.5)	-20.21	-14.24	-10.46	-7.02	0.97	1.39	1.78	2.31
(1.0,-0.5)	-15.25	-11.41	-8.59	-5.97	0.98	1.35	1.67	2.10
(1.0, 0.0)	-14.78	-11.17	-8.54	-5.91	0.95	1.33	1.72	2.14
(1.0, 0.5)	-14.62	-11.02	-8.54	-5.95	0.94	1.32	1.68	2.12
(5.0,-0.5)	-14.44	-11.05	-8.28	-5.73	0.97	1.34	1.65	2.06
(5.0, 0.0)	-14.28	-10.80	-8.26	-5.78	0.96	1.33	1.71	2.16
(5.0, 0.5)	-13.86	-10.45	-8.23	-5.80	0.96	1.32	1.67	2.09

Table 4.29

Empirical cumulative distribution of $\hat{\tau}$ for $A_1 = 1$ ($N = 25$)

(σ_e^2, C_1)	Probability of a Smaller Value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-20.21	-9.47	-6.08	-4.17	4.34	20.11	74.92	4804
(0.2, 0.0)	-318.0	-15.72	-8.78	-5.63	4.06	17.90	70.03	4775
(0.2, 0.5)	-186.4	-28.78	-16.07	-10.22	4.80	23.64	109.2	12684
(1.0,-0.5)	-6.98	-4.41	-3.24	-2.40	1.92	3.59	7.90	25.87
(1.0, 0.0)	-7.79	-4.94	-3.67	-2.58	1.88	3.49	6.87	26.10
(1.0, 0.5)	-12.66	-7.29	-5.11	-3.39	1.81	3.23	6.28	22.32
(5.0,-0.5)	-4.33	-3.05	-2.45	-1.91	1.51	2.41	3.51	6.10
(5.0, 0.0)	-4.49	-3.35	-2.59	-1.96	1.57	2.57	3.64	6.30
(5.0, 0.5)	-5.00	-3.59	-2.73	-2.08	1.51	2.42	3.60	6.41

Table 4.30

Empirical cumulative distribution of $\hat{\tau}$ for $A_1 = 1$ ($N = 50$)

(σ_e^2, C_1)	Probability of a Smaller Value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-9.85	-6.37	-4.53	-3.16	2.21	4.90	20.79	217.5
(0.2, 0.0)	-10.04	-6.78	-4.97	-3.49	2.12	4.28	20.67	193.5
(0.2, 0.5)	-17.93	-11.97	-8.53	-4.18	2.80	15.86	96.14	9403
(1.0,-0.5)	-4.27	-3.20	-2.55	-1.97	1.29	1.99	2.78	3.95
(1.0, 0.0)	-4.24	-3.21	-2.57	-2.00	1.25	1.88	2.62	4.09
(1.0, 0.5)	-5.86	-4.17	-3.24	-2.31	1.30	1.96	2.70	3.98
(5.0,-0.5)	-3.33	-2.67	-2.21	-1.78	1.12	1.70	2.28	2.97
(5.0, 0.0)	-3.23	-2.68	-2.22	-1.76	1.15	1.63	2.18	2.94
(5.0, 0.5)	-3.46	-2.86	-2.34	-1.81	1.17	1.71	2.23	3.01

Table 4.31

Empirical cumulative distribution of $\hat{\tau}$ for $A_1 = 1$ ($N = 100$)

(σ_e^2, C_1)	Probability of a Smaller Value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-5.35	-4.00	-3.15	-2.40	1.14	1.76	2.47	3.94
(0.2, 0.0)	-6.09	-4.72	-3.47	-2.44	1.11	1.73	2.31	3.33
(0.2, 0.5)	-7.92	-3.32	-2.71	-2.21	1.06	1.70	2.37	3.51
(1.0,-0.5)	-3.32	-2.71	-2.26	-1.81	1.00	1.52	1.96	2.54
(1.0, 0.0)	-3.32	-2.70	-2.21	-1.77	1.02	1.50	1.98	2.55
(1.0, 0.5)	-4.18	-3.29	-2.68	-2.00	1.00	1.46	1.93	2.52
(5.0,-0.5)	-2.95	-2.45	-2.10	-1.71	0.98	1.44	1.85	2.33
(5.0, 0.0)	-2.90	-2.41	-2.03	-1.67	1.01	1.46	1.86	2.41
(5.0, 0.5)	-2.92	-2.48	-2.12	-1.71	1.01	1.41	1.82	2.34

Table 4.32

Empirical cumulative distribution of $\hat{\tau}$ for $A_1 = 1$ ($N = 250$)

(σ_e^2, C_1)	Probability of a Smaller Value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-3.55	-2.85	-2.33	-1.84	0.98	1.44	1.89	2.42
(0.2, 0.0)	-4.00	-2.90	-2.20	-1.73	1.01	1.43	1.86	2.38
(0.2, 0.5)	-3.32	-2.82	-2.41	-1.95	0.96	1.44	1.89	2.39
(1.0,-0.5)	-2.77	-2.38	-2.04	-1.67	0.94	1.37	1.79	2.21
(1.0, 0.0)	-2.77	-2.37	-2.05	-1.68	0.94	1.33	1.74	2.15
(1.0, 0.5)	-3.13	-2.58	-2.17	-1.75	0.94	1.37	1.74	2.16
(5.0,-0.5)	-2.68	-2.32	-1.99	-1.63	0.93	1.35	1.76	2.17
(5.0, 0.0)	-2.65	-2.29	-1.99	-1.63	0.93	1.32	1.72	2.09
(5.0, 0.5)	-2.64	-2.26	-1.98	-1.65	0.93	1.35	1.71	2.11

CHAPTER V

MODEL MISSPECIFICATION PROBLEMS

A general regression model with integrated errors and one system of integrated regressors is introduced in Section 5.1. Section 5.2 develops the basic asymptotic results of parameter estimates and the usual OLS statistics, which is applied to subsequent Sections 5.3-5.5. Spurious regression problem is discussed in Section 5.3 and misspecification of nonstationary autoregressive and polynomial regression models considered in Section 5.4. Section 5.5 discusses underspecification of orders in nonstationary AR and polynomial regression models.

5.1. General Regression Model

We consider a *general regression model* with integrated errors and one system of integrated regressors as follows:

$$y_t = \beta_0 + \beta_1 X_{t,1} + \dots + \beta_p X_{t,p} + U_{t,k}, \quad t = 1, \dots, n, \quad (5.1.1)$$

where y_t is the regressand, the regressors consist of an intercept and a system of integrated variables defined by

$$X_{t,j} = X_{1,j-1} + \dots + X_{t,j-1}, \quad j = 1, 2, \dots, p, \quad X_{t,0} = x_t, \quad (5.1.2)$$

for a given deterministic or random sequence $\{x_t\}$, the regression error $U_{t,k}$ is a k -integrated error defined by

$$U_{t,j} = U_{1,j-1} + \dots + U_{t,j-1}, \quad j = 1, \dots, k, \quad U_{t,0} = u_t$$

and p and k are the levels of integration. The two processes $\{x_t\}$ and $\{u_t\}$ are assumed to have certain limiting behavior based on the following Assumption 5.1 suggested by Shin and Sarkar (1996, pp. 4-5). For simplicity of presentation, we discuss one system of integrated regressors.

Assumption 5.1. *There exist sequences of real numbers a_n and d_n and continuous non-zero functions f and g on $[0,1]$, possibly random, such that as $n \rightarrow \infty$ $(\sum_{t=1}^{[nr]} x_t a_n^{-1}, \sum_{t=1}^{[ns]} u_t d_n^{-1})$ converges in distribution to $(g(r), f(s))$, $0 \leq r, s \leq 1$, where $[nr]$ denotes the largest integer not greater than nr .*

Usually, each of the functions g and f in Assumption 5.1 is either a deterministic function, or a standard Brownian motion. The latter holds with $w_t = x_t$ or $w_t = u_t$ under conditions such as

CC1. $E(w_t) = 0$ for all t , $\text{Sup}_t E|w_t|^{\xi+\nu} < \infty$ for some $\xi > 2$ and $\nu > 0$, $\sigma^2 = \lim_{n \rightarrow \infty} E(w_1 + \dots + w_n)^2/n$ exists and $\sigma^2 > 0$, $\{w_t\}$ is a strong-mixing with mixing coefficients ψ_m satisfying $\sum \psi_m^{1-2/\xi} < \infty$; or

CC2. $w_t = \sum_{j=0}^{\infty} \phi_j e_{t-j}$ where $\{e_t\}$ is an iid sequence with $E(e_t) = 0$, $E(e_t^2) = \sigma_e^2 < \infty$, $\sum_j^2 \phi_j^2 < \infty$ and $\sigma^2 = \sigma_e^2 (\sum_{j=0}^{\infty} \phi_j)^2 > 0$.

The conditions CC1 and CC2 were formulated by Herrndorf (1984) and Phillips and Solo (1992) respectively.

If $X_{t,1}, \dots, X_{t,p}$ linearly explain nothing of y_t , then $(\beta_0 + \beta_1 X_{t,1} + \dots + \beta_p X_{t,p})$ is said to be *totally misspecified* and $U_{t,k}$ represents the nonstationary process y_t . If $X_{t,1}, \dots, X_{t,p}$ explain the nonstationarity of y_t partially in the sense that $U_{t,k}$ is still nonstationary with $k < p$. Then the model (5.1.1) is said to be *partially misspecified*. The conventional t statistics diverge asymptotically and the Durbin-Watson (DW) statistic converges in probability to zero in cases of totally or partially misspecified models. On the other hand, model (5.1.1) is said to be *well specified* when y_t is well explained by $X_{t,1}, \dots, X_{t,p}$ in the sense that $U_{t,k}$ is stationary with $k = 0$. In this case, t statistic has nondegenerate limiting distributions and the limits of DW and R^2 are positive and one respectively. If $k = 0$ in model (5.1.1) and $x_t = u_{t-1}$ in (5.1.2), then model (5.1.1) becomes the AR model of order p

$$y_t = \phi_0 + \sum_{j=1}^p \phi_j y_{t-j} + u_t, \quad (5.1.3)$$

where ϕ_j 's are linear combination of β_j 's whose characteristic roots are all one. Similarly, if $k = 0$ in model (5.1.1) and $x_t = 1$ in (5.1.2), then (5.1.1) becomes the polynomial regression model

$$y_t = \delta_0 + \sum_{j=1}^p \delta_j t^j / j! + u_t \quad (5.1.4)$$

where δ_j 's are linear combinations of β_j 's.

5.2. Asymptotic properties

In this section we discuss the asymptotics of the LS estimator and other regression statistics in model (5.1.1) under proper normalization. Assumption 5.1 is imposed on the two processes $\{x_t\}$ and $\{u_t\}$ to derive the asymptotic results.

Let $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$ and let $\mathbf{y} = (y_1, \dots, y_n)'$, $\mathbf{X}_t = (1, X_{t,1}, \dots, X_{t,p})'$, $\mathbf{X} = (\mathbf{X}_1 | \mathbf{X}_2 | \dots | \mathbf{X}_n)'$, $\mathbf{U} = (U_{1,k}, \dots, U_{n,k})'$, $\bar{y} = \sum_{t=1}^n y_t / n$ and $\bar{U} = \sum_{t=1}^n U_{t,k} / n$. Then the LS estimator $\hat{\beta}$ of β satisfies

$$(\hat{\beta} - \beta) = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{U}.$$

To establish the limiting distribution of suitably normalized $(\hat{\beta} - \beta)$, the joint limiting distributions of the elements in $\mathbf{X}'\mathbf{X}$ and $\mathbf{X}'\mathbf{U}$ of Lemma 5.1 below (Shin and Sarkar, 1996, pp. 6-7) are exploited.

Lemma 5.1. Let model (5.1.1) hold with Assumption 5.1. Assume that $k \geq 1$ and $p \geq 1$.

Then, jointly

$$(1) \quad n^{-j+1} a_n^{-1} X_{[nr],j} \xrightarrow{L} g_j(r), \quad j = 1, \dots, p;$$

$$n^{-k+1} d_n^{-1} U_{[nr],k} \xrightarrow{L} f_k(r);$$

$$(2) \quad n^{-i-j+1} a_n^{-2} \sum_{t=1}^n X_{t,i} X_{t,j} \xrightarrow{L} \int_0^1 g_i(r) g_j(r) dr, \quad i, j = 1, \dots, p;$$

$$(3) \quad n^{-j-k+1} d_n^{-1} a_n^{-1} \sum_{t=1}^n X_{t,j} U_{t,k} \xrightarrow{L} \int_0^1 g_j(r) f_k(r) dr, \quad j = 1, 2, \dots, p,$$

where $g_j(r) = \int_0^r g_{j-1}(s) ds$, $j = 2, \dots, p$, $g_1(r) = g(r)$, $f_j(r) = \int_0^r f_{j-1}(s) ds$, $j = 2, \dots, k$, $f_1(r) =$

$f(r)$.

Proof. Noting $a_n^{-1} X_{[nr],1} = a_n^{-1} \sum_{t=1}^{[nr]} x_t \xrightarrow{L} g_1(r) = g(r)$, we have

$$X_{[nr],2} n^{-1} a_n^{-1} = \sum_{t=1}^{[nr]} X_t n^{-1} a_n^{-1} \xrightarrow{L} g_2(r) = \int_0^r g_1(s) ds$$

because of the continuity of the map $g_1(r) = \int_0^r g_1(s) ds$ and the continuous mapping theorem (Billingsley 1968, Theorem 5.1). The proof of part (1) is completed by repeatedly applying similar arguments. Let D be the set of all right continuous functions on $[0, 1]$ having the left hand limits, equipped with Skorokhod topology. Then the proof

of parts (2), (3) and the joint convergence is verified by using part (1) and continuity of the map $D^2 \rightarrow D^{p+k} \times \mathfrak{R}^{p(p+1)}$ defined by

$$\mathfrak{T}:(f, g) \rightarrow [f_1, \dots, f_k, g_1, \dots, g_p, (\int_0^1 g_i g_j dr, i = 1, \dots, p, j = 1, \dots, p), (\int_0^1 g_j f_k dr, j = 1, \dots, p)]$$

and continuous mapping theorem. since $(g(r), f(s))$ is continuous, the convergence in Skorokhod topology is equivalent to the uniform convergence (Billingsley 1968, p. 112).

Hence, it suffices to show the continuity of the map \mathfrak{T} in the uniform topology, which is

verified by observing that, for example, $|\int_0^1 g_j f_k dr| \leq \|g_j g_k\|$, where $\|\cdot\|$ is the uniform norm

$\|g\| = \sup_{0 \leq t \leq 1} |g(t)|$ that induces the uniform topology on $C[0,1]$, the set of all continuous functions on $[0, 1]$. Since boundedness of a map on $C[0, 1]$ to \mathfrak{R} is equivalent to

continuity, the map $(g_i, g_j) \rightarrow \int_0^1 g_i g_j dr$ is continuous. Q.E.D.

Now let

$$\mathbf{A}_n = \text{diag}[1, a_n(1, n, \dots, n^{p-1})], \quad \mathbf{G}(r) = [1, g_1(r), \dots, g_p(r)]',$$

$$\mathbf{H} = \int_0^1 \mathbf{G}(r)\mathbf{G}(r)' dr, \quad \mathbf{V} = \int_0^1 \mathbf{G}(r)f_k(r) dr, \quad z = \int_0^1 f_k^2 dr, \quad (5.2.1)$$

and let

$$\hat{\sigma}^2 = \sum_{t=1}^n \hat{U}_t^2 / (n-p-1)$$

where $\hat{U}_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 X_{t,1} - \dots - \hat{\beta}_p X_{t,p}$. Then,

$$(n-p-1)\hat{\sigma}^2 = \mathbf{U}'\mathbf{U} - \mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U}.$$

Now from the above Lemma 5.1 it follows that

$$n^{-1}\mathbf{A}^{n-1}(\mathbf{X}'\mathbf{X})\mathbf{A}_n^{-1} \xrightarrow{L} \mathbf{H}, \quad \mathbf{A}_n^{-1}(\mathbf{X}'\mathbf{U})/(n^k d_n) \xrightarrow{L} \mathbf{V}$$

and

$$\mathbf{U}'\mathbf{U}/(n^{2k-1} d_n^2) \xrightarrow{L} z. \quad (5.2.2)$$

We now define conventional regression statistics: the Durbin-Watson statistic

$$DW = \sum_{t=2}^n (\hat{U}_t - \hat{U}_{t-1})^2 / \sum_{t=1}^n \hat{U}_t^2;$$

the coefficient of determination

$$R^2 = [\mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - n\bar{y}^2] / (\mathbf{y}'\mathbf{y} - n\bar{y}^2);$$

and the t-statistic

$$t_{\beta_j} = (\hat{\beta}_j - \beta_j) / s_{\hat{\beta}_j},$$

where $s_{\hat{\beta}_j}$ denotes the standard error of $\hat{\beta}_j$.

Now we state three theorems containing the limiting properties of the parameter estimators and other conventional regression statistics, which are given by Shin and Sarkar (1996).

Theorem 5.1. Let model (5.1.1) hold with Assumption 5.1. Assume $k \geq 1$, $p \geq 1$ and that \mathbf{H} is nonsingular. Then

- (1) $\mathbf{A}_n(\hat{\beta} - \beta) / (n^{k-1} d_n) \xrightarrow{L} \mathbf{H}^{-1} \mathbf{V}$, i.e.,
- $$d_n^{-1} a_n n^{j-k} (\hat{\beta}_j - \beta_j) \xrightarrow{L} (\mathbf{H}^{-1} \mathbf{V})_{j+1,1} \quad \text{for } j = 1, \dots, p;$$
- (2) $(n-p-1)\hat{\sigma}^2 / (n^{2k-1} d_n^2) \xrightarrow{L} (z - \mathbf{V}'\mathbf{H}^{-1}\mathbf{V})$.

Theorem 5.2. Let model (5.1.1) hold with Assumption 5.1. Assume that $k \geq 1$, $p \geq 1$ and \mathbf{H} is nonsingular. also assume that f_k is not a linear combination of g_1, g_2, \dots, g_p . Define $\tau = 0$ if $\beta = \mathbf{0}$, otherwise define $\tau = \lim_{n \rightarrow \infty} (n^\kappa a_n) / (n^k d_n)$ where $\kappa = \max\{0 \leq j \leq p: \beta_j \neq 0\}$ and $h(r) = \tau \beta_\kappa g_\kappa(r) + f_k(r)$. then

- (1) $DW \xrightarrow{p} 0$;
- (2) if $\tau = \infty$, $R^2 \xrightarrow{p} 1$;
- if $0 \leq \tau < \infty$,

$$R^2 \xrightarrow{L} 1 - \left[\int_0^1 h^2(r) dr - \left(\int_0^1 h(r) dr \right)^2 \right]^{-1} [z - \mathbf{V}'\mathbf{H}^{-1}\mathbf{V}];$$

- (3) for $j = 1, \dots, p$,

$$t_{\beta_j} / \sqrt{n} \xrightarrow{L} (\mathbf{H}^{-1} \mathbf{V})_{j+1,1} [(\mathbf{H}^{-1})_{j+1,j+1} (z - \mathbf{V}'\mathbf{H}^{-1}\mathbf{V})]^{-1/2}.$$

Proof of (1). Note that

$$\hat{U}_t = \sum_{j=0}^p (\beta_j - \hat{\beta}_j) X_{t,j} + U_{t,k}, \quad \hat{U}_t - \hat{U}_{t-1} = \sum_{j=1}^p (\beta_j - \hat{\beta}_j) X_{t,j-1} + U_{t,k-1}$$

Now

$$\sum_{t=1}^n (\hat{U}_t - \hat{U}_{t-1})^2 = \mathbf{U}^* \mathbf{U}^* - 2(\hat{\beta} - \beta)' \mathbf{X}^* \mathbf{U}^* + (\hat{\beta} - \beta)' \mathbf{X}^* \mathbf{X}^* (\hat{\beta} - \beta)$$

where

$$\mathbf{X}_t^* = (0, x_t, X_{t,1}, \dots, X_{t,p-1})', \quad \mathbf{X}^* = (\mathbf{X}_1^* | \mathbf{X}_2^* | \dots | \mathbf{X}_n^*)', \quad \text{and } \mathbf{U}^* = (U_{1,k-1}, \dots, U_{n,k-1})'.$$

Then, by the continuous mapping theorem and (5.2.2) we have

$$\mathbf{A}_n^{-1} (\mathbf{X}^* \mathbf{X}^*) \mathbf{A}_n^{-1} n^{-1} \xrightarrow{p} \mathbf{0}, \quad \mathbf{A}_n^{-1} (\mathbf{X}^* \mathbf{U}^*) / (n^k d_n) \xrightarrow{p} \mathbf{0}, \quad \mathbf{U}^* \mathbf{U}^* / (n^{2k-1} d_n^2) \xrightarrow{p} \mathbf{0} \quad (5.2.3)$$

and the result follows as a direct consequence of (5.2.3) and Theorem 5.1. Q.E.D.

Proof of (2). Since $\hat{\mathbf{y}}' \hat{\mathbf{y}} = \mathbf{y}' \mathbf{y} - (n-p-1) \hat{\sigma}^2$, where $\hat{\mathbf{y}} = (\hat{y}_1, \dots, \hat{y}_n)'$, $R^2 = 1 - (\mathbf{y}' \mathbf{y} - n \bar{y}^2)^{-1} (n-p-1) \hat{\sigma}^2$. By (5.2.2) and Theorem 5.1-(2), $(n-p-1) \hat{\sigma}^2$ is of the same order in probability as $\mathbf{U}' \mathbf{U}$. Now if $\tau = \infty$, $y_t \cong \beta_\kappa X_{t,\kappa}$ and by Lemma 5.1-(1)

$$n^{1-\kappa} \mathbf{a}_n^{-1} \mathbf{y}_{[nr]} \xrightarrow{L} \beta_\kappa \mathbf{g}_\kappa(\mathbf{r}) \neq \mathbf{0}.$$

Hence by Theorem 5.1- (2) and the continuous mapping theorem

$$(n-p-1)\hat{\sigma}^2/(y'y - n\bar{y}^2) = O_p(n^{2k-1}d_n^2/n^{2k-1}a_n^2) = o_p(1)$$

and $R^2 \xrightarrow{p} 1$. If $0 \leq \tau < \infty$, $y_t \cong \beta_\kappa X_{t,\kappa} + U_{t,\kappa}$ and

$$\begin{aligned} n^{k-1}d_n^{-1}y_{[nr]} &\cong (n^\kappa a_n/n^k d_n)\beta_\kappa n^{1-\kappa}X_{[nr],\kappa} + n^{1-k}d_n^{-1}U_{[nr],\kappa} \\ &\xrightarrow{L} \tau \beta_\kappa g_\kappa(r) + f_k(r) = h(r). \end{aligned} \quad (5.2.4)$$

Hence, the limiting distribution of R^2 follows from Theorem 5.1-(2) and (5.2.4) and continuous mapping theorem. Q.E.D.

Proof of (3). The result follow from

$$t_{\beta_j} = \frac{n^{1-k}d_n^{-1}a_n n^{j-1}(\hat{\beta}_j - \beta_j)}{\sqrt{n(a_n n^{j-1})^2 \{(X'X)^{-1}\}_{j+1,j+1} n^{1-2k}d_n^{-2}\hat{\sigma}^2}},$$

(5.2.2) and Theorem 5.1. Q.E.D.

The condition that f_k is not a linear combination of g_1, g_2, \dots, g_p is equivalent to the fact that model (5.1.1) is misspecified in our sense as described in Section 5.1 after two conditions are defined. If model (5.1.1) is totally or partially misspecified, then $U_{t,\kappa}$ is an integrated error representing the nonstationarity of y_t unexplained by the regressors. Therefore, f_k can not be a linear combination of g_1, g_2, \dots, g_p , which makes $(z - \mathbf{V}'\mathbf{H}^{-1}\mathbf{V}) > 0$ and the limits in part (3) of Theorem 5.2 well defined. This is the case for all the misspecification examples considered in Sections 5.3, 5.4.1, and 5.5. In view of Lemma

5.1-(1), τ defined in the above Theorem 5.2 represents the probabilistic order of $(\beta_0 + \beta_1 X_{t,1} + \dots + \beta_p X_{t,p})$ relative to the error $U_{t,k}$. Hence $\tau = \infty$ corresponds to the situation when the regressors have greater probabilistic order than the error, and in such a situation $R^2 \xrightarrow{P} 1$. On the other hand, $0 \leq \tau < \infty$ corresponds to the case when the regressors do not have greater probabilistic order than the error and in such a case R^2 converges to a random variable which is less than one.

When (5.1.1) is a well specified model in the sense that the nonstationarity of y_t is captured by $X_{t,j}$'s then the regression error $U_{t,k}$ is a stationary process with $k = 0$. This situation can not be handled using Theorem 5.1 and 5.2 and to derive asymptotics in the case we need some additional conditions.

Theorem 5.3. Let model (5.1) hold with Assumption 5.1. Assume that \mathbf{H} is nonsingular and also $k = 0$, $p \geq 1$, $d_n = n^{1/2}$ and $\mathbf{A}_n^{-1} \mathbf{X}' \mathbf{U} / d_n$ converges in distribution to a nonzero finite random vector, say, ζ and $\mathbf{U}' \mathbf{U} / n$ converges in probability to a positive real number, say, ψ . Then

- (1) for $j = 1, \dots, p$, $n^{j-1/2} a_n (\hat{\beta}_j - \beta_j) \xrightarrow{L} (\mathbf{H}^{-1} \zeta)_{j+1,1}$
- (2) $\hat{\sigma}^2 \xrightarrow{P} \psi$;
- (3) if $\sum_{t=1}^n U_{t,0} U_{t-1,0} \xrightarrow{P} \phi$, then $DW \xrightarrow{P} 2(\psi - \phi) / \psi$;
- (4) $R^2 \xrightarrow{P} 1$;
- (5) for $j = 1, \dots, p$, $t_{\beta_j} \xrightarrow{L} (\mathbf{H}^{-1} \zeta)_{j+1,1} [(\mathbf{H}^{-1})_{j+1,j+1} \cdot \psi]^{-1/2}$

Proof. Proof is similar to that of Theorem 5.1 and 5.2 and is, therefore, omitted. Q.E.D.

As seen from the above theorems, the behavior of the regression statistics depends on whether the error term $U_{t,k}$ is stationary or nonstationary. If model (5.1.1) is misspecified then $U_{t,k}$ is nonstationary and $\hat{\beta}_j$ diverges for j such that $d_n n^{k-j}/a_n \uparrow \infty$; $\hat{\sigma}^2$ diverges; $DW \xrightarrow{P} 0$. If the probabilistic order of the regressors is not greater than that of the regression error, then the limit of R^2 is less than one, otherwise $R^2 \xrightarrow{P} 1$; and t_{β_j} diverges, implying misleading high significance of the t_{β_j} .

On the other hand, if the nonstationarity of y_t is well explained by $X_{t,j}$'s then the error term $U_{t,k}$ will be stationary. In this case $\hat{\beta}_j$ is consistent; $\hat{\sigma}^2$ converges; DW converges to the positive number because $\psi > \varphi$; $R^2 \xrightarrow{P} 1$; t_{β_j} converges. One implication is that a small value of DW along with a moderate R^2 value may be taken as an indication of possible misspecification in the sense that the error term $U_{t,k}$ in model (5.1.1) is nonstationary and needs to be explained further. Furthermore, we will state several corollaries based on the above three theorems in the next sections, which are given by Shin and Sarkar (1996).

5.3. Spurious regression problem

5.3.1. Previous works

A small value of DW together with a moderate R^2 value may be taken as an indication of possible misspecification in the sense that the error term $U_{t,k}$ in model (5.1.1) is nonstationary. Granger and Newbold (1974) showed, through simulation results, the danger of acceptance of spurious relationships if the traditional significant test statistics are used. If autocorrelated errors in time series regression equations are ignored, problems arise involving inefficient parameter estimates, and invalid significant tests and sub-optimal results when the fitted equations are used to derive forecasts. Provided that a regression equation relating variables is found to have strongly autocorrelated residuals, equivalent to a low Durbin-Watson value, the only conclusion that can be reached is that the equation is misspecified, whatever the value of R^2 observed. Furthermore, the form of the misspecification can be considered to be either (a) the omission of relevant variables or (b) the inclusion of irrelevant variables or (c) autocorrelated residuals. There are three ways suggested by them in which the problems associated with spurious regression can be avoided. The first approach is to include lagged values of both the dependent and independent variable in the regression. A second approach is to difference the data before estimating the relation. A third approach is to assume a simple first-order autoregressive form for the residual of the equation. Because the second approach avoids the spurious regression problem as well as the nonstandard distributions for certain hypotheses, many researchers recommend routinely differencing apparently nonstationary variables before

estimating regressions. However, there are two different situations in which it might be inappropriate. First, if the data are really stationary, then differencing the data can result in a misspecified regression. Second, even if both series y_{1t} and y_{2t} are truly *integrated* of order 1 ($I(1)$), i.e., both series with no deterministic component has a stationary, invertible ARMA representation after differencing once, there is a class of models, so called *cointegrated models*, for which the bivariate dynamic relation between y_{1t} and y_{2t} will be misspecified if the researcher simply differences both y_{1t} and y_{2t} .

Newbold and Davies (1978) suggested that a reasonable alternative to AR(1), in certain circumstances, is the ARIMA(0, 1, 1) process in terms of error structure and that inference in regressions involving time series can be greatly affected by the error structure assumed.

Phillips (1986) developed an asymptotic theory for regression with integrated processes that explains analytically the empirical findings of Granger and Newbold (1974) and furthermore worked the asymptotics of various statistics such as regression coefficients, t statistics and F statistics for the significance of the regression coefficients in a spurious regression, and showed that the usual t and F test statistics do not possess limiting distributions but actually diverge as the sample size n approaches ∞ . for illustration, consider the simplest case:

$$y_{1,t} = y_{1,t-1} + \varepsilon_{1t}$$

$$y_{2,t} = y_{2,t-1} + \varepsilon_{2t}$$

$$\mathbf{y}_t = (y_{1t}, y_{2t})'$$

where ε_{1t} 's are iid with mean zero and variance σ_1^2 , ε_{2t} 's are iid with mean zero and variance σ_2^2 , and ε_{1t} is independent of ε_{2t} for all t . First, an OLS regression of y_{1t} on y_{2t} and a constant,

$$y_{1t} = \alpha + \beta y_{2t} + e_t, \quad (5.3.1)$$

produces estimates $\hat{\alpha}_n$ and $\hat{\beta}_n$ characterized by

$$\begin{bmatrix} \hat{\alpha}_n / \sqrt{n} \\ \hat{\beta}_n \end{bmatrix} \xrightarrow{L} \begin{bmatrix} \sigma_1 h_1 \\ (\sigma_1 / \sigma_2) h_2 \end{bmatrix}$$

where

$$\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \equiv \begin{bmatrix} 1 & \int W^*(r) dr \\ \int W^*(r) dr & \int (W^*(r))^2 dr \end{bmatrix}^{-1} \begin{bmatrix} \int W^*(r) dr \\ \int (W^*(r))^2 dr \end{bmatrix}$$

and the integral sign indicates integration over r from 0 to 1, $W^*(r)$ and denotes scalar standard Brownian motion. Usually, the LS estimates are consistent with $\hat{\beta}_n \xrightarrow{P} \mathbf{0}$ and must be multiplied by some increasing function of n in order to obtain a nondegenerate asymptotic distribution. Here, however, neither estimate is consistent—different arbitrarily large samples will have randomly differing estimates $\hat{\beta}_n$. Indeed, the estimate of the constant term $\hat{\alpha}_n$ actually diverges, and must be divided by $n^{1/2}$ to obtain a random variable with a well-specified distribution. The estimate $\hat{\alpha}_n$ itself is likely to get farther and farther from the true value of zero as the sample size increases. Second, the usual LS estimate of the variance of e_t , $\hat{\sigma}_n^2 = (\text{Residual SS})/(n-2)$, again diverges as n goes to

infinity. To obtain an estimate that does not grow with the sample size, the residual sum of squares has to be divided by n^2 rather than n . In this respect, the residuals $\hat{\epsilon}_n$ from a spurious regression behave like a unit root process; if z_t is a $I(1)$ series, then $\sum z_t^2/n$ diverges and $\sum z_t^2/n^2$ converges. Third, OLS t or F tests based on the spurious regression (5.3.1) also diverge; the regression F -statistic must be divided by n to obtain a variable that does not grow with the sample size. Since an F test of a single restriction is the square of the corresponding t test, any t statistic would have to be divided by $n^{1/2}$ to obtain a convergent variable. Thus, as the sample size n becomes larger, it becomes likely that the absolute value of a t test statistic will exceed any arbitrary finite value. Note that above results can be obtained only when all of the elements of \mathbf{y}_t are $I(1)$ with zero drift and when the vector \mathbf{y}_t is not cointegrated.

5.3.2. Spurious regression

We now discuss a few of the implications of the unit root tests literature for regression with integrated variables. Suppose first that one is interested in making inference on a particular regression coefficient (which may be a vector). The following situations arise naturally:

(1) If the disturbance term is serially uncorrelated and uncorrelated with the regressors, and the coefficient(s) of interest can be expressed as coefficients on zero-mean stationary variables, then LS estimators are consistent and the standard asymptotic distribution theory applies. If, on the other hand, the coefficient(s) of interest cannot be expressed as

coefficients on zero-mean stationary variables, then consistency still obtains, but the asymptotic distribution is nonstandard.

(2) If an integrated regressand is cointegrated with at least one regressor, so that the disturbance term is stationary but not necessarily serially uncorrelated or independent of the regressors, and if the coefficient(s) of interest can be expressed as coefficients on zero-mean stationary variables, then LS estimates are inconsistent. If, however, the coefficient of interest is on an integrated regressor and cannot be written as a coefficient on a zero-mean stationary variable, then the estimator is consistent (but it has a nonstandard asymptotic distribution).

(3) If an integrated regressand is not cointegrated with any regressor, so that the disturbance term is integrated, then the LS estimates of coefficients on integrated regressors will be inconsistent; it is in this situation that the Granger-Newbold “*spurious regression*” phenomenon arises.

We consider the general framework (5.1.1) in which a k -integrated series y_t is wrongly regressed on another p -integrated series $X_{t,p}$ and its lags, i.e., $\nabla^k y_t = u_t$, $\nabla^p X_{t,p} = x_t$, where ∇ is the difference operator such that $\nabla X_{t,j} = X_{t,j} - X_{t-1,j}$. For simplicity of analysis and notation we assume that $u_t = 0$ and $x_t = 0$ for $t \leq 0$. We also assume that both $\{x_t\}$ and $\{u_t\}$ satisfy the above condition CC1 or CC2 for some $\sigma_x^2 > 0$ and $\sigma_u^2 > 0$. Then application of Theorems 5.1 and 5.2 gives the following (Shin and Sarkar, 1996, p. 11).

Corollary 5.3. Under the assumptions of Section 5.3.2, $\hat{\beta}_j$ diverges for $j \leq k$, while $\hat{\beta}_j$ converges for $j > k$; $\hat{\sigma}^2$ diverges; all the t-statistics diverges at the rate $n^{1/2}$; $DW \xrightarrow{P} 0$; R^2 has a nondegenerate limit which is less than one.

Proof. We first give a proof for the case when $\{y_t\}$ and $\{x_t\}$ are independent. The true y_t satisfy model (5.1.1) with $U_{t,k} = y_t = \nabla^k u_t$ and $\beta_0 = \beta_1 = \dots = \beta_p = 0$, and Assumption 5.1 is then satisfied with

$$a_n = n^{1/2}, \quad d_n = n^{1/2}, \quad g(r) = \sigma_x W^x(r), \quad f(r) = \sigma_u W^u(r),$$

where $W^x(r)$ and $W^u(r)$ are two independent standard Brownian motions on $[0,1]$. Therefore, the limiting distribution of the normalized LS estimator $n^{-k+1-1/2}(\hat{\beta}_0, \sqrt{n}\hat{\beta}_1, \sqrt{n^3}\hat{\beta}_2, \dots, n^{-1/2+p}\hat{\beta}_p)$ in the spurious regression is given by Theorem 5.1-(1) with

$$\mathbf{G}(r) = [1, \sigma_x W_1^x(r), \dots, \sigma_x W_p^x(r)]' \text{ and } f_k(r) = \sigma_u W_k^u(r),$$

where $W_j^x(r)$ is the j -fold integral of $W^x(r)$, and $W_k^u(r)$ is the k -fold integral of $W^u(r)$.

Therefore, $\hat{\beta}_j$ diverges for $j \leq k$, while $\hat{\beta}_j$ converges for $j > k$; also $\hat{\sigma}^2$ diverges. Since $X_{t,j}$'s are independent of $y_t = U_{t,k}$, $W_k^u(r)$ is not a linear combination of $W_1^x(r), \dots, W_p^x(r)$, and since $\beta_0 = \beta_1 = \dots = \beta_p = 0$, τ of Theorem 5.2 is zero. In this case the regressors do not explain nonstationarity of the regressand at all and Theorem 5.1-(1) holds and results follow from Theorem 5.2.

Suppose next $X_{t,p}$ and y_t are not independent. If $k > p$ or y_t and $X_{t,p}$ are not cointegrated, then $W_k^u(r)$ is not a linear combination of $W_1^x(r), \dots, W_p^x(r)$ and the asymptotics of $\hat{\beta}$, $\hat{\sigma}^2$, t_{β_j} 's and DW are the same as those when $X_{t,p}$ and its lags and y_t are independent. Q.E.D.

Example 5.1. Engle and Granger (1987, Table IV) analyzed the US quarterly real per capita consumption on nondurables (C_t) and real per capita disposable income (Y_t) from 1947: I to 1981: II and they obtained $\hat{C}_t = 0.52 + 0.23Y_t$, $\hat{\sigma} = 0.016$, $DW = 0.46$, $R^2 = 0.99$. The moderate value of DW and high value of R^2 indicate cointegration between C_t and Y_t , supporting the conclusion of Engle and Granger who proposed to use DW as a test statistic for integration when $p = k = 1$.

5.4. Misspecification of nonstationary AR and polynomial regression models

In analyzing a time series showing some kind of a trend as in macroeconomic variables, the deterministic trend (or polynomial regression) model and stochastic trend y_t , defined by $\nabla^s y_t = u_t$ for some positive integer s , are two main alternatives for representing the trend. We consider the effect of misspecification on the asymptotic behavior of the LS estimate in which a polynomial regression model

$$y_t = \delta_0 + \delta_1 t + \delta_2 (t^2/2!) + \delta_p (t^p/p!) + \varepsilon_t \quad (5.4.1)$$

is used to estimate a stochastic trend defined by the nonstationary AR model

$$y_t = \phi_0 + \phi_1 \nabla^{q-1} y_{t-1} + \dots + \phi_{q-1} \nabla y_{t-q+1} + \phi_q y_{t-q} + \eta_t \quad (5.4.2)$$

or vice versa. The ε_t and η_t are the error components and are assumed to satisfy condition CC1 or CC2 given in Section 5.1. For model (5.4.2) we also assume that the s roots of the polynomial $A(L) = [1 - \phi_1(1 - L)^{q-1}L - \dots - \phi_q m^q]$ lie on the unit circle and the remaining $(q-s)$ roots lie outside the unit circle.

5.4.1. Previous works

Nelson and Plosser (1982) argued that a large number of macroeconomic aggregates are better modeled as random walks or integrated processes of order 1 (I(1) processes) rather than stationary processes with a trend. Furthermore, many variables such as real GNP and employment were found to be reasonably characterized as random walks with drift.

Nelson and Kang (1984) investigated the properties of standard regression statistics including R^2 , t and sample autocorrelations of the residuals when time is the only explanatory variable, namely polynomial model of order one in time when in fact the time series we are interested in explaining belongs to the difference stationary process (DSP) class such as a driftless random walk. They explored the fact that this regression results in the inappropriate inference that the trend is significant.

Durlauf and Phillips (1988) analyzed the effect of misspecification of the generating mechanism of a nonstationary time series in terms of deterministic trends and derived the limiting properties of the LS estimator in the linear time trend regression model when the true model is a random walk. In particular, the DW statistic is shown to possess promising asymptotic properties as a regression diagnostic tool in the sense that it converges to zero when an integrated process is erroneously treated as stationary.

5.4.2. Misspecification of a nonstationary AR model as a polynomial regression model

Suppose the true y_t satisfies $\nabla^s y_t = \eta^*$ with η^* satisfying condition CC1 or CC2 of Section 5.1 for some $\sigma^2 > 0$, suppose model (5.4.1) is fit to the $\{y_t\}_{t=1}^n$ series. Then applying Theorem 5.1 and 5.2 with $x_t = 1$, $k = s$, and $U_{t,s} = y_t$ we obtain the next result (Shin and Sarkar, 1996, p. 14).

Corollary 5.4.1. Under the above assumptions of Section 5.4.2, for $j < s$, $\hat{\delta}_j$ diverges; all t-statistics diverges at the rate $n^{1/2}$; DW converges to zero; R^2 converges to a degenerate limit which is less than one.

Proof. Regression (5.4.2) has the form of model (5.1.1) with true parameters $\delta_0 = \delta_1 = \delta_p = 0$, $x_t = 1$, $k = s$ and regression error $U_{t,s} = y_t = \nabla^s \eta_t^*$. Let $a_n = n$, $d_n = (\sigma^2 n)^{1/2}$, $g(r) = r$, $f(r) = W(r)$,

$$\mathbf{A}_n = \text{diag}(1, n, \dots, n^p), \quad \mathbf{G} = (1, r/1!, \dots, r^p/p!)', \quad f_s(r) = W_s(r),$$

where $W_s(r)$ is the s -fold integral of $W(r)$. Then by Theorem 5.1-(1), upon evaluating the integrals, we obtain

$$n^{1-s} (\sigma^2 n)^{-1/2} (\hat{\delta}_0, n\hat{\delta}_1, \dots, n^p \hat{\beta}_p)' \xrightarrow{L} \mathbf{K}^{-1} \mathbf{M}^{-1} \int_0^1 \mathbf{G}(r) W_s(r) dr,$$

where $\mathbf{K} = \text{diag}(1, 1/1!, 1/2!, \dots, 1/p!)$ and \mathbf{M} is a $(p+1) \times (p+1)$ matrix with (i, j) th element $= (i+j+1)^{-1}$. Note that for $j < s$, $\hat{\delta}_j$ diverges. We have positive $(z - \mathbf{V}' \mathbf{H}^{-1} \mathbf{V})$ in Theorem

5.1-(2) because $W_s(r)$ is not a linear combination of $(1, r/1!, \dots, r^p/p!)$. Also τ in Theorem 5.2 is zero because $\delta_0 = \dots = \delta_p = 0$. Therefore, the asymptotics of statistics follow from Theorem 5.2. Q.E.D.

Note that this section is a case in which the regressors do not express non-stationarity of y_t at all and hence the model is totally misspecified.

Example 5.2. Consider the log of the US consumer prices analyzed by Nelson and Plosser (1982, Table 4). Their reported DW statistic value for this series in the regression $\hat{y}_t = \hat{\delta}_0 + \hat{\delta}_1 t$ is approximately 0.06, indicating a misspecification. However, the DW statistic value in the autoregressive fitting in Table 5 of Nelson and Plosser (1982) is 2.12, which indicates a good fit. This leads to the same conclusion of Nelson and Plosser that log of consumer price is well represented by a unit root AR model instead of the trend regression model.

5.4.3. Misspecification of a polynomial regression model as a nonstationary AR model

We assume that the true y_t is generate by model (5.1.1) with errors $\{\varepsilon_t\}$ satisfying condition CC1 or CC2 with some $\sigma^2 > 0$, model (5.4.2) is fit to the $\{y_t\}_{t=1}^n$ series. Assume $\delta_p \neq 0$ and $p \geq q \geq 1$. Let $X_{t,j} = \delta_p (t^{p-q+j}/(p-q+j)!)$, $j = 1, 2, \dots, q$. Then $\nabla^{q-j} y_{t-j} = X_{t,j} + O_p(n^{p-q+j-1})$, uniformly in t . Hence, the asymptotics in the estimated model (5.4.2) in the same as those in

$$\hat{y}_t = \hat{\phi}_0 + \hat{\phi}_1 X_{t,1} + \dots + \hat{\phi}_q X_{t,q}$$

and we have the following result (Shin and Sarkar, 1996, p. 15).

Corollary 5.4.2. Assume that the above assumptions of Section 5.4.2 hold

(1) If $\delta_0 = \delta_1 = \dots = \delta_{p-q} = 0$ then $\hat{\phi}_j$ converges; $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$; t_{ϕ_j} converges; DW $\xrightarrow{P} 2(\sigma^2 - \varphi)/\sigma^2$, where $\sum_{t=1}^n \varepsilon_t \varepsilon_{t-1} \xrightarrow{P} \varphi$; $R^2 \xrightarrow{P} 1$.

(2) If some of $\delta_0, \delta_1, \dots, \delta_{p-q}$ are nonzero, then $\hat{\phi}_j$ converges; $\hat{\sigma}^2$ diverges; DW $\xrightarrow{P} 0$; $R^2 \xrightarrow{P} 1$; t_{γ_j} diverges for j such that $(\mathbf{H}^{-1}\mathbf{V})_{j+1,1} \neq 0$ where \mathbf{H} and \mathbf{V} are as defined in (5.2.1) with

$$\mathbf{G}(r) = [1, r^{p-q+1}/(p-q+1)!, \dots, r^p/p!]', \quad f_k(r) = \delta_k r^k/k!$$

where $k = \kappa = \max \{j: \delta_j \neq 0, 0 \leq j \leq p-q\}$.

Proof. Observe that the true y_t in (5.4.1) can be written as

$$y_t = (\delta_{p-q+1}X_{t,1} + \dots + \delta_q X_{t,q})/\delta_p + U_t$$

which is of the form (5.1.1) with $x_t = \delta_p t^{p-q}/(p-q)!$ and $U_t = \delta_{p-q} t^{p-q}/(p-q)! + \dots + \delta_0 + \varepsilon_t$.

Note that

$$\sum_{t=1}^{[nr]} x_t / \delta_p n^{p-q+1} \rightarrow g(r) = r^{p-q+1}/(p-q+1)!$$

Therefore, with $a_n = \delta_p n^{p-q+1}$, $\mathbf{G}(r) = [1, r^{p-q+1}/(p-q+1)!, \dots, r^p/p!]$.

The asymptotics depend on whether $U_t = \varepsilon_t$ or not.

First proof of part (1) is given. Assume $\delta_0 = \delta_1 = \dots = \delta_{p-q} = 0$. Then $U_t = \varepsilon_t$ and conditions of Theorem 5.3 are satisfied with $\zeta = \int_0^1 \mathbf{G}(r) d\mathbf{W}(r)$, $\psi = \sigma^2$. Hence, by

Theorem 5.3

$$\delta_p (n^{p-q+1} n^j / \sigma n^{1/2}) (\hat{\phi}_j - \delta_{p-q+j} \delta_p^{-1}) \xrightarrow{L} (\mathbf{H}^{-1} \zeta)_{j+1,1},$$

which is a normal distribution, and the results follow. Q.E.D.

Now proof of part (2) is given. Assume some of the δ_j 's are not equal to zero. Let $\kappa = \max\{j: \delta_j \neq 0, 0 \leq j \leq p-q\}$. Then $\delta_\kappa \neq 0$ and $U_t \cong \delta_\kappa t^\kappa / \kappa!$. Hence, with $d_n = n^{\kappa+1}$,

$$d_n^{-1} \mathbf{A}_n^{-1} \mathbf{X}' \mathbf{U} \xrightarrow{L} \delta_\kappa \int_0^1 \mathbf{G}(r) \frac{r^\kappa}{\kappa!} dr = \mathbf{V}.$$

Since $r^\kappa / \kappa!$ is not a linear combination of $r^{p-q+1} / (p-q+1)!, \dots, r^p / p!$, condition of Theorem 5.2 is satisfied. Therefore, by Theorem 5.1 and 5.2

$$(\delta_p (n^{p-q+1}) / n^{\kappa+1}) n^j (\hat{\phi}_j - \delta_{p-q+j} \delta_p^{-1}) \xrightarrow{L} (\mathbf{H}^{-1} \mathbf{V})_{j+1,1},$$

which is nonrandom; $\hat{\phi}_j$ converges; $\hat{\sigma}^2$ diverges; $DW \xrightarrow{P} 0$; $t_{\hat{\phi}_j}$ diverges for j such that $(\mathbf{H}^{-1} \mathbf{V})_{j+1,1} \neq 0$. Since $X_{t,p} \cong t^p / p!$ increases faster than $U_t \cong \delta_\kappa t^\kappa / \kappa!$, τ in Theorem 5.2 is ∞ . Hence, $R^2 \xrightarrow{P} 1$. Q.E.D.

In the case of part (1) of Corollary 5.4.2 the estimated model (5.4.2) is well specified model in the sense that all the nonstationarity of y_t is explained by the regressors. Whereas in the case of part (2) the model is partially misspecified in the sense that the trends of y_t of the order greater than $(p-q)$ are explained by the regressors, the lags of y_t , while trends of the order less than or equal to κ is still unexplained.

5.5. Underspecification of orders in nonstationary autoregressive and polynomial regression models

The order values p and q in the polynomial regression and nonstationary autoregressive models (5.4.1) and (5.4.2) are usually not known and they might be overspecified or underspecified.

5.5.1. Previous works

Dickey and Pantula (1987) developed a sequential testing procedure for testing three unit roots against two unit roots and then two unit roots against one unit root, which could be used to determine the number of unit roots. In addition Hasza and Fuller (1979), Sen and Dickey (1987) and Haldrup (1994) investigated tests for unit roots in the $I(2)$ time series.

5.5.2. Underspecification of the order in a nonstationary autoregressive model

If the data are underdifferenced, by which we mean that the AR model does in fact have some unit roots and/or cointegrating relationships but is nevertheless estimated completely in levels, so that no unit roots or cointegrating relationships are imposed on the data, then the unit roots and cointegrating relationships (if present) will nevertheless

be satisfied asymptotically (although some efficiency is lost in finite samples), and moreover, convergence is typically at rates faster than $O(n^{1/2})$. The fact that the distribution theory for some of the estimated coefficients is nonstandard is of no consequence for construction of point forecasts. Thus, the costs of underdifferencing are likely to be low. Overdifferencing, on the other hand, discards low-frequency information and destroys cointegrating relationships, and may cause difficulties for numerical estimation algorithm, due to the induced unit moving-average roots.

Suppose the true y_t satisfies $\nabla^s y_t = \eta^*$ with η^* satisfying condition CC1 or CC2 for some $\sigma^2 > 0$ and is estimated by model (5.4.2) with $q < s$. In this case the model is partially misspecified. Using Theorem 5.1 and 5.2 we get the following (Shin and Sarkar, 1996, p. 17).

Corollary 5.5.1. Under the above assumptions of section 5.2, $\hat{\phi}_j$ converges; $\hat{\sigma}^2$ diverges at the rate $n^{2(s-q)}$; $R^2 \xrightarrow{P} 1$; $DW \xrightarrow{P} 0$; t_{ϕ_j} diverges at the rate $n^{1/2}$.

Proof. Let $X_{tj}^* = \nabla^{q-j} y_{t-j}$, $X_{tj} = \nabla^{q-j} y_{t-1}$ and $x_t = \nabla^q y_t$. Then

$$X_{tj}^* = X_{tj} + O_p(n^{(j-1)/2}),$$

$$y_t = (1 - L + L)^q y_t = \sum_{j=1}^q \binom{q}{j} \nabla^{q-j} y_{t-j} + x_t = \sum_{j=1}^q \binom{q}{j} X_{t-j}^* + x_t$$

which is of the form (5.1.1) with $U_{t,k} = x_t$. Also observe that $k = s - q$ because $x_t = \nabla^{q-j} \eta^*$.

Since

$$n^{-(s-q)}(\sigma^2 n)^{-1/2} \sum_{t=1}^{\lfloor nr \rfloor} X_t \xrightarrow{L} W_{s-q+1}(r), \quad a_n = n^{s-q}(\sigma^2 n)^{1/2}$$

and

$$\mathbf{G}(r) = (1, W_{s-q+1}(r), \dots, W_s(r))'.$$

Also, since $U_{t,k} = \nabla^{q-s} \eta_t^*$, by Lemma 5.1-(2), $d_n = a_n$ and $\mathbf{V} = \int_0^1 \mathbf{G}(r) W_{s-q}(r) dr$, $z =$

$\int_0^1 W_{s-q}^2(r) dr$. By Theorem 5.1

$$n^j (\hat{\phi}_j - \begin{pmatrix} q \\ j \end{pmatrix}) \xrightarrow{L} (\mathbf{H}^{-1} \mathbf{V})_{j+1,1},$$

$\hat{\phi}_j$ converges; $\hat{\sigma}^2$ diverges at the rate $n^{2(s-q)}$. Observe that $W_{s-q}(r)$ is not a linear combination of $W_{s-q+1}(r), \dots, W_s(r)$, and since the probabilistic order of $X_{t,q} = \nabla^s \eta_t^*$ is greater than that of $U_{t,k} = \nabla^{s+q} \eta_t^*$, τ in Theorem 5.2 is ∞ . Therefore, $R^2 \xrightarrow{P} 1$; also $DW \xrightarrow{P} 0$ and t_{ϕ_j} diverges at the rate $n^{1/2}$. Q.E.D.

We now consider an application of the above. If one is fitting an AR(1) model to

an I(2) process, then $p = 1$, $s = 2$, and $\mathbf{G} = (1, W_2(r))'$ and $\mathbf{V} = \int_0^1 \mathbf{G}(r) W(r) dr = (\int_0^1 W(r) dr,$

$\int_0^1 W(r) W_2(r) dr)'$. Therefore,

$$n(\hat{\phi}_1 - 1) \xrightarrow{d} \int_0^1 W(r) W_2(r) dr / \int_0^1 W_2^2(r) dr.$$

Hence $n(\hat{\phi}_1 - 1)$ converges in distribution. However, $t_{\hat{\phi}_1} = (\hat{\phi}_1 - 1) / s_{\hat{\phi}_1}$ diverges at the rate $n^{1/2}$ and $t_{\hat{\phi}_1}$ is expected to reject the single unit root hypothesis with higher probability than $n(\hat{\phi}_1 - 1)$.

Example 5.3. Consider the data on commercial bank real estate loans analyzed by Dickey and Pantula (1987). Dickey and Pantula argued that the series was well represented by an AR process with two unit roots. For testing H_0 : one unit root vs. H_1 : not H_0 , after adjusting for the mean and lags of first differences, Dickey and Pantula (1987, Table 1) obtained 1.265 for the t-statistic $\hat{\tau}_\mu$ and $70 \times (0.006197) = 0.434$ for $n(\hat{\rho}_\mu - 1)$. Using Tables 8.5.1 - 8.5.2 of Fuller (1976) the $n(\hat{\rho}_\mu - 1)$ value is less than 97.5-th percentile point and $\hat{\tau}_\mu$ is much higher than the 99-th percentile point. Therefore, the t-statistic $\hat{\tau}_\mu$ has more power for the double unit roots alternative. In this example, $\hat{\tau}_\mu$ provides stronger evidence than $n(\hat{\rho}_\mu - 1)$ against the single unit root null hypothesis.

5.5.3. Underspecification of the order in a polynomial regression model

We next consider the polynomial regression. Let the true model for y_t be a q th order polynomial regression $y_t = \delta_0 + \delta_1 t + \delta_2 (t^2/2!) + \dots + \delta_q (t^q/q!) + \varepsilon_t$, $\delta_q \neq 0$, and is estimated by the p th order polynomial regression $\hat{y}_t = \hat{\delta}_0 + \hat{\delta}_1 t + \hat{\delta}_2 t^2 / 2! + \dots + \hat{\delta}_p (t^p/p!)$,

where $p < q$. This is a case of a totally misspecified model. By Theorem 5.1 and 5.2 the following is obtained:

Corollary 5.5.2. Under the assumptions of Section 5.5.3, all the $\hat{\delta}_j$'s and $\hat{\sigma}^2$ diverges; the t-statistics diverge at the rate $n^{1/2}$; DW converges to 0; R^2 converges to a constant lying strictly between zero and one.

Proof. The p th order polynomial model can be considered as model (5.1.1) with $p = p$, $k = q$, $x_t = 1$, $U_{t,q} = y_t$ and $u_t = \nabla^q (\delta_1 t + \dots + \delta_q (t^q/q!)) + \varepsilon_t \cong \delta_q$, $a_n = n$ and $d_n = \delta_q n$, and zero as the true parameter values. Then,

$$\mathbf{G}(r) = [1, r, \dots, r^p/p!]', \quad f_q(r) = r^q/q!$$

and by Theorem 5.1 the limiting distribution of $(n^{-q}\hat{\delta}_0, n^{-q+1}\hat{\delta}_1, \dots, n^{-q+p}\hat{\delta}_p)'$ is $\delta_q \int_0^1 \mathbf{G}(r)$
 $\mathbf{G}(r)' dr]^{-1} [\int_0^1 \mathbf{G}(r) f_q(r) dr]$. We see that all the $\hat{\delta}_j$'s and $\hat{\sigma}^2$ diverge. Since $U_{t,q} \cong \delta_q t^q/q!$, τ
 in Theorem 5.2 is zero. Also $f_q(r)$ is not a linear combination of $1, r, \dots, r^p/p!$. Hence, by
 Theorem 5.2 the result follows. Q.E.D.

CHAPTER VI

AN EMPIRICAL STUDY

In this chapter we consider a Monte Carlo study based on Chapter 5 regarding the model specification problems. We generate 10,000 samples of size n for $n = 25, 50, 100$ and 250. For a fixed sample size n , the corresponding empirical means of parameter estimates, $\hat{\sigma}^2$ and other conventional regression statistics such as t -statistics, R^2 and DW are considered. The subroutine DRNNOA is used to generate the normal random numbers. Section 6.1 discusses algorithms based on misspecified models in Sections 5.3-5.5, and Section 6.2 contains Monte Carlo results and a discussion.

6.1. Algorithms

In this section several algorithms derived on the basis of corollaries from Chapter 5 are introduced.

6.1.1. Spurious regression

We consider the following special cases of the spurious regression based on $\nabla^k y_t = u_t$, $\nabla^p X_{t,p} = x_t$:

(A) $k = 1$ and $p = 2$

(B) $k = 2$ and $p = 1$, and

(C) $k = 2$ and $p = 2$.

The following different algebraic equations will be used in data generation.

$$\nabla y_t = u_t \Rightarrow y_t = y_{t-1} + u_t \Rightarrow y_t = \sum_{i=1}^t u_i \quad (6.1.1)$$

$$\nabla^2 y_t = u_t \Rightarrow y_t = 2y_{t-1} - y_{t-2} + u_t \Rightarrow y_t = 2 \sum_{i=1}^{t-1} u_i - \sum_{i=1}^{t-2} u_i + u_t \quad (6.1.2)$$

$$\nabla X_{t,1} = x_t \Rightarrow X_{t,1} = X_{t-1,1} + x_t \Rightarrow X_{t,1} = \sum_{i=1}^t x_i \quad (6.1.3)$$

$$\nabla X_{t,2} = X_{t,1} \Rightarrow X_{t,2} = X_{t-1,2} + X_{t,1} \Rightarrow X_{t,2} = \sum_{i=1}^t X_{i,1} \quad (6.1.4)$$

$$\begin{aligned} \nabla^2 X_{t,2} = x_t &\Rightarrow X_{t,2} = 2X_{t-1,2} - X_{t-2,2} + x_t \\ &\Rightarrow X_{t,2} = 2 \sum_{i=1}^{t-1} X_{i,1} - \sum_{i=1}^{t-2} X_{i,1} + x_t \end{aligned} \quad (6.1.5)$$

where $x_t = u_t = 0$, if $t \leq 0$.

A. Case 1: $k = 1$ and $p = 2$

Algorithm:

- (1) Generate random numbers u_t 's and x_t 's iid $N(0, 1)$ or $N(0, 5)$, respectively, where $t = 1, \dots, n$ with $n = 25, 50, 100$ and 250 .
- (2) Compute the y_t 's, $X_{t,1}$'s and $X_{t,2}$'s based on equations (6.1.1) and (6.1.4), respectively.
- (3) Regress y_t on $X_{t,1}$ and $X_{t,2}$ where $y_t = \beta_0 + \beta_1 X_{t,1} + \beta_2 X_{t,2} + \varepsilon_t$.
- (4) Calculate regression coefficients and the usual OLS statistics, $\hat{\sigma}^2$, t , DW and R^2 based on the fitted model.
- (5) Repeat steps (1) - (4) ten thousand times.
- (6) Compute the corresponding empirical means of the statistics $\hat{\sigma}^2$, DW and R^2 respectively through step (5).
- (7) Compare the obtained numerical results from steps (1) - (6) with the theoretical results of Corollary 5.3.

B. Case 2: $k = 2$ and $p = 1$

Algorithm:

- (1) Generate random numbers u_t 's and x_t 's iid $N(0, 1)$ or $N(0, 5)$, respectively, where $t = 1, \dots, n$ with $n = 25, 50, 100$ and 250 .
- (2) Compute the y_t 's and $X_{t,1}$'s based on equations (6.1.2) and (6.1.3), respectively.
- (3) Regress y_t on $X_{t,1}$ where $y_t = \beta_0 + \beta_1 X_{t,1} + \varepsilon_t$.
- (4) Calculate regression coefficients and the usual OLS statistics, $\hat{\sigma}^2$, t , DW and R^2 based on the fitted model.
- (5) Repeat steps (1) - (4) ten thousand times.

- (6) Compute the corresponding empirical means of the statistics $\hat{\sigma}^2$, DW and R^2 respectively through step (5).
- (7) Compare the obtained numerical results from steps (1) - (6) with the theoretical results of Corollary 5.3.

C. Case 3: $k = 2$ and $p = 2$

Algorithm:

- (1) Generate random numbers u_t 's and x_t 's iid $N(0, 1)$ or $N(0, 5)$, respectively, where $t = 1, \dots, n$ with $n = 25, 50, 100$ and 250 .
- (2) Compute the y_t 's, $X_{t,1}$'s and $X_{t,2}$'s based on equations (6.1.2) and (6.1.4), respectively.
- (3) Regress y_t on $X_{t,1}$ and $X_{t,2}$ where $y_t = \beta_0 + \beta_1 X_{t,1} + \beta_2 X_{t,2} + \varepsilon_t$.
- (4) Calculate regression coefficients and the usual OLS statistics, $\hat{\sigma}^2$, t, DW and R^2 based on the fitted model.
- (5) Repeat steps (1) - (4) ten thousand times.
- (6) Compute the corresponding empirical means of the statistics $\hat{\sigma}^2$, DW and R^2 respectively through step (5).
- (7) Compare the obtained numerical results from steps (1) - (6) with the theoretical results of Corollary 5.3.

6.1.2. Misspecification of nonstationary AR and polynomial regression models

We consider the case of $p = q = 2$ in models (5.4.1) and (5.4.2) to observe the empirical behavior of the regression statistics in misspecified models.

A. Misspecification of a nonstationary AR(2) model as a polynomial regression model of order 2

True model: $\nabla^2 y_t = u_t$, i.e., a nonstationary AR(2) model

Fitted model: $y_t = \delta_0 + \delta_1 t + \delta_2 (t^2/2) + \varepsilon_t$, i.e., a polynomial regression model of order 2.

Algorithm:

- (1) Generate random numbers $u_t \sim \text{NID}(0, 1)$ or $N(0, 5)$, where $t = 1, \dots, n$ with $n = 25, 50, 100$ and 250 .
- (2) Compute the y_t 's based on the following: $y_t = 2y_{t-1} - y_{t-2} + u_t$ where $u_t = y_t = 0$ if $t \leq 0$.
- (3) Fit the data y_t generated by step (2) to the polynomial regression model of order 2.
- (4) Calculate regression coefficients and the usual OLS statistics, $\hat{\sigma}^2$, t , DW and R^2 based on the fitted model.
- (5) Repeat steps (1) - (4) ten thousand times.
- (6) Compute the empirical means of the statistics $\hat{\sigma}^2$, DW, t and R^2 .
- (7) Compare the obtained numerical results from steps (1) - (6) with the theoretical results of Corollary 5.4.1.

B. Misspecification of a polynomial regression model of order 2 as a nonstationary AR(2) model

True model: $y_t = \delta_0 + \delta_1 t + \delta_2(t^2/2) + \varepsilon_t$ with $(\delta_0, \delta_1, \delta_2) = (0, 1, 1)$ or $(0, 2, 2)$ or $(0, 2, 1)$

Fitted model: $y_t = \phi_0 + \phi_1 \nabla y_{t-1} + \phi_2 y_{t-2} + \eta_t$

Algorithm:

- (1) Generate random numbers $\varepsilon_t \sim \text{NID}(0, 1)$ or $\text{N}(0, 5)$, where $t = 1, \dots, n$ with $n = 25, 50, 100$ and 250 .
- (2) Compute the y_t 's based on the following: $y_t = t + t^2/2 + \varepsilon_t$ where $\varepsilon_t = y_t = 0$ if $t \leq 0$.
- (3) Regress y_t on $\{(y_{t-1} - y_{t-2}), y_{t-2}\}$.
- (4) Calculate regression coefficients and the usual OLS statistics, $\hat{\sigma}^2$, t , DW and R^2 based on the fitted model.
- (5) Repeat steps (1) - (4) ten thousand times.
- (6) Compute the empirical means of the statistics $\hat{\sigma}^2$, DW, t and R^2 .
- (7) Compare the obtained numerical results from steps (1) - (6) with the theoretical results of Corollary 5.4.2.

6.1.3. Underspecification of the orders in nonstationary AR and polynomial regression model

A. Underspecification of the order in a nonstationary AR model.

Let the true generating mechanism for y_t be $\nabla^2 y_t = \eta_t^*$, with η_t^* satisfying condition CC1 or CC2 for some $\sigma^2 > 0$, and is estimated by model $\nabla y_t = \eta_t$:

True model: $\nabla^2 y_t = u_t$, i.e., nonstationary AR(2) model

Fitted model: $y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$.

Algorithm:

- (1) Generate random number $u_t \sim \text{NID}(0, 1)$ or $\text{NID}(0, 5)$, where $t = 1, \dots, n$ with $n = 25, 50, 100$ and 250 .
- (2) Compute the y_t 's based on the following: $y_t = 2y_{t-1} - y_{t-2} + u_t$ where $u_t = y_t = 0$ if $t \leq 0$.
- (3) Regress y_t on y_{t-1} .
- (4) Calculate regression coefficients and the usual OLS statistics, $\hat{\sigma}^2$, t , DW and R^2 based on the fitted model.
- (5) Repeat steps (1) - (4) ten thousand times.
- (6) Compute the empirical means of the statistics $\hat{\sigma}^2$, DW, t and R^2 .
- (7) Compare the obtained numerical results from steps (1) - (6) with the theoretical results of Corollary 5.5.1.

B. Underspecification of the order in a polynomial regression model

Let the true model for y_t be a second order polynomial regression and is estimated by the first order polynomial regression:

True model: $y_t = \delta_0^* + \delta_1^* t + \delta_2^* (t^2/2) + \varepsilon_t^*$ with $(\delta_0^*, \delta_1^*, \delta_2^*) = (0, 1, 1)$ or $(0, 2, 2)$

Fitted model: $y_t = \delta_0 + \delta_1 t + \varepsilon_t$.

Algorithm:

- (1) Generate random number $\varepsilon_t \sim \text{NID}(0, 1)$ or $\text{NID}(0, 5)$, where $t = 1, \dots, n$ with $n = 25, 50, 100$ and 250 .

- (2) Compute the y_t 's based on the following: $y_t = t + t^2/2 + \varepsilon_t$ where $\varepsilon_t = y_t = 0$ if $t \leq 0$.
- (3) Regress y_t on t .
- (4) Calculate regression coefficients and the usual OLS statistics, $\hat{\sigma}^2$, t , DW and R^2 based on the fitted model.
- (5) Repeat steps (1) - (4) ten thousand times.
- (6) Compute the empirical means of the statistics $\hat{\sigma}^2$, DW, t and R^2 .
- (7) Compare the obtained numerical results from steps (1) - (6) with the theoretical results of Corollary 5.5.2.

6.2. Results and Discussion

In Chapter 5 a general regression model with integrated errors and one system of integrated regressors has been considered and asymptotic properties of the conventional regression statistics under this model have been discussed. Empirical results on model specification problems for nonstationary time series are presented in Table 6.1 through Table 6.7.

Tables 6.1-6.3 show the empirical results in spurious regression against the theoretical background given by Corollary 5.3. The following common empirical results from Tables 6.1-6.3 can be seen: all the DW's appear to converge in probability to 0; all $\hat{\sigma}^2$'s diverge; all the R^2 's seem to have a nondegenerate limit which is less than one (0.49, 0.37 and 0.78 in Tables 6.1-6.3, respectively); all the t-statistics seem to diverge slowly; all the parameter estimates $\hat{\beta}_j$'s may diverge very slowly for $j \leq k$ and converge very slowly for $j > k$. From these results we can observe the fact that LS statistics such as R^2 and DW can be used as diagnostic tools to check the spurious regression in the sense that they can be used as remarkable symptoms of spurious regression when we obtain DW close to 0 and R^2 much less than one. Furthermore, as the orders of k and/or p are getting larger, the rate of divergence for $\hat{\beta}_j$'s and t_{β_j} 's are getting faster.

Table 6.4-Table 6.5 display the empirical results under misspecification of polynomial regression models and nonstationary AR against the theoretical backdrop provided by Corollary 5.4.1 and Corollary 5.4.2 respectively. From Table 6.4 the following can be observed: $\hat{\delta}_j$'s, for $j = 0, 1, 2$, diverge; all three t-statistics diverge; DW

seems to converge to 0; R^2 appears to converge to 0.96; $\hat{\sigma}^2$ diverges. In addition, from Table 6.5 the following can be seen: $\hat{\phi}_1$ appears to converge to 1.9 and $\hat{\phi}_2$ to 1.00; DW seems to converge in probability to 3; $\hat{\sigma}^2$ appears to converge in probability to 4.01; R^2 seems to converge in probability to 1; t_{ϕ_1} and t_{ϕ_2} converge to numbers near 9.5 and 5.78 respectively. All these empirical values in Table 6.4-6.5 fully support the theoretical results of Corollary 5.4.1 and Corollary 5.4.2 respectively. Similarly, LS statistics such as R^2 and DW may be exploited to check for misspecification of polynomial regression models and nonstationary AR as follows:

(1) If DW is close to 0 and R^2 is close to 0.96, we can conclude that a nonstationary AR model has been misspecified as a polynomial regression model.

(2) If DW is close to 3 and R^2 is close to 1.0, we can conclude that a polynomial regression model has been misspecified as a nonstationary AR model.

Table 6.6-Table 6.7 give the simulation results under underspecification of the order in a nonstationary AR and underspecification of the order in polynomial regression. In these cases Corollary 5.5.1 and Corollary 5.5.2 respectively provide the theoretical background. From Table 6.6 the following can be observed: $\hat{\phi}_0$ seems to converge to 0.01 and $\hat{\phi}_1$ to 1.00; $\hat{\sigma}^2$ appears to diverge; R^2 seems to converge in probability to 1; DW seems to converge in probability to 0; two t-statistics appear to diverge. The following facts can be seen from Table 6.7: all the parameter estimates, $\hat{\sigma}^2$ and two t-statistics seem to diverge; DW seems to converge in probability to 0; R^2 seems to converge in probability to 0.94. All these empirical values in Tables 6.6-6.7 fully support the results of Corollary 5.4.1 and Corollary 5.4.2 respectively.

In summary, it is observed from the simulation results that DW and R^2 can be in general used as diagnostic tools to detect spurious regression, misspecification of non-stationary AR and polynomial regression models.

Table 6.1

SPURIOUS REGRESSION 1 ($k = 1, p = 2$)

statistics	n	25	50	100	250
$\hat{\beta}_0$		-0.0000	0.0075	0.0292	0.0673
$\hat{\beta}_1$		0.0041	0.0001	-0.0118	-0.1915
$\hat{\beta}_2$		-0.0004	-0.0002	-0.0001	-0.0001
t_{β_0}		-0.008	0.006	-0.022	0.092
t_{β_1}		0.002	-0.003	-0.118	-0.249
t_{β_2}		0.082	0.063	0.017	-0.271
R^2		0.497	0.496	0.490	0.489
DW		0.907	0.486	0.253	0.104
$\hat{\sigma}^2$		1.768	3.339	6.572	16.218

Table 6.2

SPURIOUS REGRESSION 2 ($k = 2, p = 1$)

statistics	n	25	50	100	250
$\hat{\beta}_0$		0.185	-0.281	-1.569	-9.702
$\hat{\beta}_1$		-0.063	-0.135	0.058	-0.090
t_{β_0}		0.039	0.026	-0.044	-0.142
t_{β_1}		-0.057	-0.024	-0.013	-0.001
R^2		0.376	0.377	0.372	0.374
DW		0.447	0.177	0.093	0.015
$\hat{\sigma}^2$		360.4	2630.9	20221.2	307268.1

Table 6.3

SPURIOUS REGRESSION 3 ($k = 2, p = 2$)

statistics	n	25	50	100	250
$\hat{\beta}_0$		0.049	0.093	0.111	6.951
$\hat{\beta}_1$		-0.014	-0.002	0.001	0.149
$\hat{\beta}_2$		-0.011	-0.042	-0.090	-0.230
t_{β_0}		0.013	-0.003	-0.171	-0.246
t_{β_1}		-0.018	0.005	0.067	0.165
t_{β_2}		0.038	-0.152	-0.536	-0.740
R^2		0.774	0.786	0.781	0.780
DW		1.001	0.318	0.103	0.033
$\hat{\sigma}^2$		90.3	644.3	5020.1	74303.8

Table 6.4

MISSPECIFICATION OF A NONSTATIONARY AR(2) AS
A POLYNOMIAL REGRESSION MODEL
OF ORDER 2

statistics	n	25	50	100	250
$\hat{\delta}_0$		-0.028	-0.109	-0.147	-1.498
$\hat{\delta}_1$		0.026	0.010	0.025	0.121
$\hat{\delta}_2$		-0.001	0.001	0.000	-0.001
t_{δ_0}		-0.017	-0.022	-0.053	-0.189
t_{δ_1}		-0.123	0.005	0.147	0.857
t_{δ_2}		-0.015	0.027	0.182	0.308
R^2		0.965	0.965	0.965	0.964
DW		0.414	0.113	0.028	0.004
$\hat{\sigma}^2$		3.304	24.774	200.348	3070.343

Table 6.5

MISSPECIFICATION OF A POLYNOMIAL REGRESSION MODEL
OF ORDER 2 AS A NONSTATIONARY AR(2) MODEL

statistics	n	25	50	100	250
$\hat{\phi}_0$		2.988	2.354	1.748	1.285
$\hat{\phi}_1$		1.604	1.860	1.961	1.994
$\hat{\phi}_2$		1.030	1.005	1.000	1.000
t_{ϕ_0}		3.092	2.354	1.748	1.701
t_{ϕ_1}		10.240	9.658	9.585	9.526
t_{ϕ_2}		5.264	5.658	5.783	5.789
R^2		0.998	0.998	0.998	0.999
DW		2.759	3.222	3.411	3.415
$\hat{\sigma}^2$		3.487	3.845	4.009	4.019

Table 6.6

UNDERSPECIFICATION OF A NONSTATIONARY AR(2) MODEL

statistics	n	25	50	100	250
$\hat{\phi}_0$		0.019	0.014	0.017	0.013
$\hat{\phi}_1$		1.019	1.012	1.007	1.003
$t_{\hat{\phi}_0}$		0.058	0.042	0.042	0.167
$t_{\hat{\phi}_1}$		93.065	257.153	704.187	2722.667
R^2		0.982	0.994	0.997	0.998
DW		0.706	0.369	0.187	0.075
$\hat{\sigma}^2$		1.186	2.273	4.618	11.459

Table 6.7

UNDERSPECIFICATION OF A POLYNOMIAL REGRESSION MODEL OF ORDER 2

statistics	n	25	50	100	250
$\hat{\delta}_0$		-58.497	-220.999	-858.501	-5271.000
$\hat{\delta}_1$		14.000	26.500	51.500	126.500
$t_{\hat{\delta}_0}$		-5.861	-8.094	-11.315	-17.756
$t_{\hat{\delta}_1}$		20.838	28.441	39.470	61.691
R^2		0.949	0.943	0.939	0.938
DW		0.087	0.023	0.006	0.001
$\hat{\sigma}^2$		585.196	9018.000	141570.278	5468059.648

REFERENCES

- Billingsley, P. (1968) *Convergence of probability measures*, Wiley, New York.
- Binder, D. A. and Hidioglou, M. A. (1988) Sampling in time, *Handbook of Statistics* 6, 187-211, North Holland, New York.
- Box, G. E. P. and Jenkins, G. M. (1976) *Time Series Analysis: Forecasting and Control*, revised edition, Holden Day, San Francisco.
- Dickey, D. A. (1976) *Estimation and testing of nonstationary time series*, Ph. D. thesis, Department of Statistics, Iowa State University, Ames, IA.
- Dickey, D. A. and Fuller, W. A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc.* 74, 427-431.
- Dickey, D. A. and Fuller, W. A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 49, 1057-1072.
- Dickey, D. A. and Pantula, S. G. (1987) Determining the order of differencing in autoregressive process. *Journal of Business and Economic Statistics* 5, 455-461.
- Dunsmuir, W. (1979) A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise, *Ann. Statist* 7, 490-506.
- Dunsmuir, W. and Hannan, E. J. (1976) Vector linear time series model. *Adv. Statist.* 7, 339-364.
- Durlauf, S. N. and Phillips, S. G. (1988) Trends versus random walks in time series analysis. *Econometrica* 56, 1333-1354.
- Engle, R. F. and Granger, C. W. J. (1987) Cointegration and error correction: representation, estimation, and testing. *Econometrica* 55, 251-276.
- Fountis, N. G. and Dickey, D. A. (1989) Testing for a unit root nonstationarity in multivariate autoregressive time series, *Ann. Statist.* 17, 419-428.

- Fuller, W. A. (1976) *Introduction to statistical time series*. New York: John Wiley.
- Fuller, W. A. (1987) *Measurement error models*. New York: John Wiley.
- Granger, C. W. J. and Newbold, P. (1974) Spurious regression in econometrics. *Journal of Econometrics* 2, 111-120.
- Haldrup, N. (1994) Semiparametric tests for double unit roots. *Journal of Business and Economic Statistics* 12, 109-122.
- Hall, R. E. (1978) Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. *J. Polit. Econ.* 86, 971-987.
- Hall, A. (1989) Testing for a unit root in the presence of moving average errors. *Biometrika* 76, 49-56.
- Hannan, E. J. and Kavalieris, L. (1984a) A method for autoregressive-moving average estimation. *Biometrika* 71, 273-280.
- Hannan, E. J. and Kavalieris, L. (1984b) Multivariate linear time series model, *Adv. Appl. prob.* 16, 975-981.
- Hannan, E. J. and Rissanen, J. (1982) Recursive estimation of mixed autoregressive moving average order. *Biometrika* 69, 81-94.
- Hasza, D. P. and Fuller, W. A. (1979) Estimation for autoregressive with unit roots. *Ann. Statist.* 7, 1106-1120.
- Herrndorf, N. (1984) A functional central limit theorem for weakly dependent sequences of random variables. *Ann. Probab.* 12, 141-153.
- Hillmer, S. and Tiao, G. (1979) likelihood function of stationary multiple autoregressive moving average models. *J. Amer. Statist. Assoc.* 74, 654-660.
- Jenkins, G. M. and Alavi, A. S. (1981) Some aspects of modeling and forecasting multivariate time series. *J. Time Ser. Anal.* 2, 1-47.
- Judge, G. G., Griffiths, W. E., Hill, R. C., Lutkepohl, H. and Lee, T. C. (1985) *The theory and practice of econometrics*. New York: John Wiley.
- Kashyap, R. L. (1970) Maximum likelihood identification of stochastic linear systems, *I.E.E.E. Trans. autom. Control* 15, 25-34.
- Kleidon, A. W. (1986) Variances bounds tests and stock price valuation models. *J. Polit. Econ.* 94, 959-1001.
- Kohn, R. (1979) Asymptotic estimation and hypothesis testing results for vector

- linear time series models. *Econometrica* 47, 1005-1030.
- Koreisha, S. G. and Pukkila, T. M. (1989) Fast linear estimation methods for vector ARMA models. *J. Time Ser. Anal.* 10, 325-339.
- Meese, R. A. and Singleton, K. J. (1983) On unit roots and the empirical modeling of exchange rates. *Int. Econ. Rev.* 24, 1029-1035.
- Miazaki, E. S. and Dorea, C. C. Y. (1993) Estimation of the parameters of a time series subject to the error of rotational sampling, *Commun. Stat., Theory Methods* 22, 805-825.
- Nelson, C. R. and Kang, H. J. (1984) Pitfalls in the use of time as an explanatory variable in regression, *Journal of Business and Economic Statistics* 2, 73-82.
- Nelson, C. R. and Plosser, C. I. (1982) Trends and random walks in macroeconomic time series. *Journal of Monetary Economics* 10, 139-162.
- Newbold, P. and Davies, N. (1978) Error misspecification and spurious regressions. *International Economic Review* 19, 513-519.
- Osborn, D. R. (1977) Exact and approximate maximum likelihood estimators for vector moving average processes. *J. Roy. Statist. Soc., Series B* 39, 114-118.
- Pagano, N. (1974) Estimation of models of autoregressive signal plus white noise. *Ann. Statist.* 2, 99-108.
- Pankratz, A. (1983) *Forecasting with univariate Box-Jenkins models*. New York: John Wiley.
- Pantula, S. G. (1991) Asymptotic distributions of the unit root tests when the process is nearly stationary, *Journal of Business and Economic Statistics* 9, 63-71.
- Pantula, S. G. and Hall, A. (1991) Testing for unit roots in autoregressive moving average models. *Journal of Econometrics* 48, 325-353.
- Parzen, E. (1967) *Advanced seminar on spectral analysis of time series*, John Wiley.
- Phillips, P. C. B. (1986) Understanding spurious regression. *Journal of Econometrics* 33, 311-340
- Phillips, P. C. B. (1987) Time series regression with a unit root. *Econometrica* 55, 277-301.
- Phillips, P. C. B. and Perron, P. (1988) Testing for a unit root in time series regression. *Biometrika* 75, 335-346.

- Phillips, P. C. B. and Solo, V. (1992) Asymptotics for linear processes. *Ann. Statist.* 20, 971-1001.
- Reinsel, G. C., Basu, S. B. and Yap, S. F. (1992) Maximum likelihood estimators in the multivariate autoregressive moving average model from a generalized least squares viewpoint. *J. Time Ser. Anal.* 13, 133-145.
- Sakai, H. and Arase, M. (1979) Recursive parameter estimation of an autoregressive process distributed by white noise. *Int. J. Control* 30, 949-966.
- Said, S. E. and Dickey, D. A. (1984) Testing for unit roots in autoregressive-moving average of unknown order. *Biometrika* 71, 599-607.
- Said, S. E. and Dickey, D. A. (1985) Hypothesis testing in ARIMA(p,1,q) models. *J. Amer. Statist. Assoc.* 80, 369-374.
- Samuelson, P. A. (1965) Proof that properly anticipated prices fluctuate randomly. *Indust. Manag. Rev.* 6, 41-50.
- Samuelson, P. A. (1973) Proof that properly discounted present values of assets vibrate randomly Bell. *J. Econ. Manag. Sci.* 4, 69-74.
- Schwert, G. W. (1989) Tests for unit roots: A Monte Carlo investigation. *J. Amer. Statist. Assoc.* 83, 147-159.
- Sen, D. L. and Dickey, D. A. (1987) Symmetric test for second differencing in univariate time series. *Journal of Business and Economic Statistics* 5, 463-473.
- Shin, D. W. and Fuller, W. A. (1992) *Estimation for the autoregressive moving average with an autoregressive unit root*. Unpublished manuscript, Iowa State University, Ames, IA.
- Shin, D. W. (1993) Maximum likelihood estimation for autoregressive processes distributed by a moving average. *J. Time Ser. Anal.* 14, 629-643.
- Shin, D. W. and Pantula, S. G. (1993) Testing for a unit root in autoregressive processes with systematic but incomplete sampling. *Statistics and Probability Letters* 18, 183-190.
- Shin, D. W. and Sarkar, S. (1994) Testing for a unit root in an ARIMA(p,1,0) signal observed with MA(q) noise. *Commun. Stat., Theory Methods* 29, 2643-2670.
- Shin, D. W. and Sarkar, S. (1995) Estimation of the multivariate autoregressive moving average having parameter restrictions and an application to rotational sampling.

J. Time Ser. Anal. 16, 431-444.

Shin, D. W. and Sarkar, S. (1996) Model misspecification problem for nonstationary time series. *Technical report #130*, Dept. of Statistics, Oklahoma State University, Stillwater, OK.

Spliid, H. (1983) A fast estimation method for the vector autoregressive moving average model with exogenous variables. *J. Amer. Statist. Assoc.* 78, 843-849.

Walker, A. M. (1960) Some consequences of superimposed error in time series analysis, *Biometrika* 47, 33-43.

APPENDICES

APPENDIX A

FORTRAN PROGRAM FOR HANNAN AND RISSANEN'S METHOD
BASED ON KOHN'S APPROACH

```
*****
* Note that  $(N, S, \sigma_e^2) = (250, 4, 5.0)$  is considered as an example. *
*****
```

```
*****
* MAIN PROGRAM *
*****
```

```
PROGRAM HR-KOHN
PARAMETER(N=250, NUM=10000, NN=254, NNN=252, N2=4, N3=3, N5=6)
IMPLICIT REAL*8 (A-H, O-Z)
REAL*8 E(N), W(N), Z(N), U(N), Y(N)
REAL*8 A1(6), C1, D1, D2
REAL*8 R(NN), ETILDA(N), X(N,N2), T(NNN)
REAL*8 RR(NNN), ABTILDA(N3), HAP, SIGMAET
INTEGER ISEED1, ISEED2
EXTERNAL RNSET, DRNNOA, DLINRG
```

```
DATA A1/1.0,0.99,0.95,0.90,0.80,0.70/
ISEED1 = 13579
CALL RNISD(ISEED1, ISEED2)
```

```
DO 260 I1 = 1, 6
  DO 255 C1 = -0.5, 0.5, 0.5
```

```
  ICOUNT = 0
  COUNT1 = 0.0
  COUNT2 = 0.0
  COUNT3 = 0.0
  COUNT4 = 0.0
  COUNT5 = 0.0
  COUNT6 = 0.0
```

```
  D1 = A1(I1) + C1
  D2 = A1(I1) * C1
```

```
DO 100 I = 1, 18000
```

```
  CALL RNSET(ISEED1)
  CALL DRNNOA(N,E)
  CALL RNGET(ISEED1)
  CALL RNSET(ISEED2)
  CALL DRNNOA(N,W)
  CALL RNGET(ISEED2)
  Z(1) = E(1) * SQRT(5.0)
  U(1) = W(1)
```



```

Y(1) = Z(1) + U(1)

DO 10 J = 2, N
  Z(J) = -A1(I1)*Z(J-1) + E(J)*SQRT(5.0)
  U(J) = W(J) + C1*W(J-1)
  Y(J) = Z(J) + U(J)
10  CONTINUE

DO 1 K = 1, N
  R(K+4) = Y(K)
1  CONTINUE

DO 3 K = 1, 4
  R(K) = 0.0
3  CONTINUE

CALL ETINIT(Y, R, X, N, NN, ETILDA)

DO 5 K = 1, N
  RR(K+2) = ETILDA(K)
5  CONTINUE

DO 7 K = 1, 2
  RR(K) = 0.0
7  CONTINUE

CALL ABINIT(N, ABTILDA, R, RR, NN, NNN, N3)

CALL SIGMAINIT(Y,R,T,N,NN,NNN,ABTILDA,N3,HAP,SIGMAET)

CALL NRMETHOD(N,NN,NNN,N3,N5,R,T,ABTILDA,HAP,SIGMAET,
/      C1,D1,D2,ICOUNT,COUNT1,COUNT2,COUNT3,COUNT4,
/      COUNT5,COUNT6)

IF (ICOUNT.GE.10000) GO TO 115

100 CONTINUE
115 CONTINUE

PWR1 = (COUNT1/NUM)*100.0
PWR2 = (COUNT2/NUM)*100.0
PWR3 = (COUNT3/NUM)*100.0
PWR4 = (COUNT4/NUM)*100.0
PWR5 = (COUNT5/NUM)*100.0
PWR6 = (COUNT6/NUM)*100.0
WRITE(*,*) 'A1 = ', A1(I1), 'C1 = ', C1

```

```

WRITE(*,*) COUNT1, PWR1
WRITE(*,*) COUNT2, PWR2
WRITE(*,*) COUNT3, PWR3
WRITE(*,*) COUNT4, PWR4
WRITE(*,*) COUNT5, PWR5
WRITE(*,*) COUNT6, PWR6

255  CONTINUE
260  CONTINUE
      END

*****
*   SUBROUTINE 1   *
*****
*** ESTIMATE INITIAL VALUE OF ET.

SUBROUTINE ETINIT(Y, R, X, N, NN, ETILDA)
  PARAMETER(LDA=4, LDAINV=4, N2=4)
  REAL*8 R(NN), A(LDA,LDA), B(N2), AINV(LDAINV,LDAINV)
  REAL*8 X(N,N2), Y(N), ETILDA(N), SUM1(N2)

      DO 210 I = 1, 4
          DO 200 J = 1, 4
              A(I,J) = 0.0
200          CONTINUE
              B(I) = 0.0
210          CONTINUE

      DO 250 L = 1, N
          A(1,1) = A(1,1) + R(L+3)*R(L+3)
          A(2,1) = A(2,1) + R(L+2)*R(L+3)
          A(3,1) = A(3,1) + R(L+1)*R(L+3)
          A(4,1) = A(4,1) + R(L)*R(L+3)
          A(2,2) = A(2,2) + R(L+2)*R(L+2)
          A(2,3) = A(2,3) + R(L+1)*R(L+2)
          A(2,4) = A(2,4) + R(L)*R(L+2)
          A(3,3) = A(3,3) + R(L+1)*R(L+1)
          A(3,4) = A(3,4) + R(L)*R(L+1)
          A(4,4) = A(4,4) + R(L)*R(L)
          B(1) = B(1) + R(L+3)*R(L+4)
          B(2) = B(2) + R(L+2)*R(L+4)
          B(3) = B(3) + R(L+1)*R(L+4)
          B(4) = B(4) + R(L)*R(L+4)
250          CONTINUE
          A(1,2) = A(2,1)

```

```

      A(1,3) = A(3,1)
      A(1,4) = A(4,1)
      A(3,2) = A(2,3)
      A(4,2) = A(2,4)
      A(4,3) = A(3,4)

      CALL DLINRG(N2, A, LDA, AINV, LDAINV)

      DO 265 K1 = 1, 4
        DO 260 K2 = 1, N
          IF (K2.LE.K1) THEN
            X(K2,K1) = 0.0
          ELSE
            X(K2,K1) = Y(K2 - K1)
          ENDIF
        CONTINUE
      260 CONTINUE
      265 CONTINUE

      DO 280 LDAINV1 = 1, 4
        SUM = 0.0
        DO 270 LDAINV2 = 1, 4
          SUM = SUM + AINV(LDAINV1,LDAINV2)*B(LDAINV2)
        270 CONTINUE
        SUM1(LDAINV1) = SUM
      280 CONTINUE

      DO 290 K3 = 1, N
        SUM2 = 0.0
        DO 285 K4 = 1, 4
          SUM2 = SUM2 + X(K3,K4)*SUM1(K4)
        285 CONTINUE
        ETILDA(K3) = Y(K3) - SUM2
      290 CONTINUE

```

```

      RETURN
      END

```

```

*****
*   SUBROUTINE 2   *
*****

```

```

*** ESTIMATE INITIAL VALUE OF A1, B1 AND B2.

```

```

      SUBROUTINE ABINIT(N,ABTILDA,R,RR,NN,NNN,N3)
      PARAMETER(MDA=3, MDAINV=3)
      REAL*8 R(NN), RR(NNN), ABTILDA(N3)
      REAL*8 E(MDA,MDA), EINV(MDAINV,MDAINV), F(MDA)

```

```

DO 320 I = 1, 3
  DO 310 J = 1, 3
    E(I,J) = 0.0
310  CONTINUE
    F(I) = 0.0
320  CONTINUE

DO 350 K = 1, N
  E(1,1) = E(1,1) + R(K+3)*R(K+3)
  E(2,1) = E(2,1) - R(K+3)*RR(K+1)
  E(3,1) = E(3,1) - R(K+3)*RR(K)
  E(2,2) = E(2,2) + RR(K+1)*RR(K+1)
  E(2,3) = E(2,3) + RR(K)*RR(K+1)
  E(3,3) = E(3,3) + RR(K)*RR(K)
  F(1) = F(1) - R(K+3)*R(K+4)
  F(2) = F(2) + R(K+4)*RR(K+1)
  F(3) = F(3) + R(K+4)*RR(K)
350  CONTINUE
  E(1,2) = E(2,1)
  E(1,3) = E(3,1)
  E(3,2) = E(2,3)

CALL DLINRG(3, E, MDA, EINV, MDAINV)

DO 390 MDAINV1 = 1, 3
  SUM3 = 0.0
  DO 370 MDAINV2 = 1, 3
    SUM3 = SUM3 + EINV(MDAINV1,MDAINV2)*F(MDAINV2)
370  CONTINUE
  ABTILDA(MDAINV1) = SUM3
390  CONTINUE

RETURN
END

*****
*   SUBROUTINE 3   *
*****
*** ESTIMATE INITIAL VALUE OF  $\sigma_e^2$ .
    SUBROUTINE SIGMAINIT(Y, R, T, N, NN, NNN, ABTILDA,
/
    N3, HAP, SIGMAET)
    REAL*8 Y(N), R(NN), T(NNN)
    REAL*8 ABTILDA(N3), HAP, SIGMAET
    DO 400 L = 1, NNN
      T(L) = 0.0
400  CONTINUE

```

```

HAP = 0.0
DO 430 L = 1, N
    T(L+2) = R(L+4) + ABTILDA(1)*R(L+3) - ABTILDA(2)*T(L+1)
    /
    - ABTILDA(3)*T(L)
    HAP = HAP + T(L+2)*T(L+2)
430 CONTINUE

SIGMAET = HAP/N

RETURN
END

*****
* SUBROUTINE 4 *
*****
*** NEWTON-RAPHSON METHOD BASED ON KOHN'S APPROACH

SUBROUTINE NRMETHOD(N,NN,NNN,N3,N5,R,T,ABTILDA,HAP,
/ SIGMAET,C1,D1,D2, ICOUNT,COUNT1,COUNT2,
/ COUNT3,COUNT4,COUNT5,COUNT6)
PARAMETER(LDA=6, LDAINV=6, N6=6)
REAL*8 A(LDA,LDA), AINV(LDAINV,LDAINV)
REAL*8 R(NN), T(NNN), ABTILDA(N3)
REAL*8 A11, B11, B22, SIGMAET
REAL*8 SS1, SS2

A11 = ABTILDA(1)
B11 = ABTILDA(2)
B22 = ABTILDA(3)

WA1 = 0.0
WB1 = 0.0
WB2 = 0.0
WA1WB1 = 0.0
WA1WB2 = 0.0
WB1WB1 = 0.0
WB1WB2 = 0.0
WB2WB2 = 0.0

DO 510 I = 1, 6
    DO 500 J = 1, 6
        A(I,J) = 0.0
500 CONTINUE
510 CONTINUE

```

W1TEMP = 0.0
 W2TEMP = 0.0
 W3TEMP = 0.0
 WAB1TMP = 0.0
 WAB2TMP = 0.0
 WWB1TMP = 0.0
 WB12TMP = 0.0
 WWB2TMP = 0.0
 WA1TEMP1 = 0.0
 WA1TEMP2 = 0.0
 WB1TEMP1 = 0.0
 WB1TEMP2 = 0.0
 WB2TEMP1 = 0.0
 WB2TEMP2 = 0.0
 WAB1TMP1 = 0.0
 WAB1TMP2 = 0.0
 WAB2TMP1 = 0.0
 WAB2TMP2 = 0.0
 WWB1TMP1 = 0.0
 WWB1TMP2 = 0.0
 WB12TMP1 = 0.0
 WB12TMP2 = 0.0
 WWB2TMP1 = 0.0
 WWB2TMP2 = 0.0

DO 550 L = 1, N

$WA1 = R(L+3) - B11*WA1TEMP1 - B22*WA1TEMP2$
 $WB1 = -T(L+1) - B11*WB1TEMP1 - B22*WB1TEMP2$
 $WB2 = -T(L) - B11*WB2TEMP1 - B22*WB2TEMP2$
 $WA1WB1 = -WA1TEMP1 - B11*WAB1TMP1 - B22*WAB1TMP2$
 $WA1WB2 = -WA1TEMP2 - B11*WAB2TMP1 - B22*WAB2TMP2$
 $WB1WB1 = -2*WB1TEMP1 - B11*WWB1TMP1 - B22*WWB1TMP2$
 $WB1WB2 = -WB2TEMP1 - B11*WB12TMP1 - WB1TEMP2$
 $\quad - B22*WB12TMP2$
 $WB2WB2 = -2*WB2TEMP2 - B11*WWB2TMP1 - B22*WWB2TMP2$

$A(1,1) = A(1,1) + (WA1*WA1)$
 $A(2,2) = A(2,2) + (WB1*WB1 + T(L+2)*WB1WB1)$
 $A(3,3) = A(3,3) + (WB2*WB2 + T(L+2)*WB2WB2)$
 $A(4,4) = A(4,4) + (T(L+2)*T(L+2))$
 $A(1,2) = A(1,2) + (WA1*WB1 + T(L+2)*WA1WB1)$
 $A(1,3) = A(1,3) + (WA1*WB2 + T(L+2)*WA1WB2)$
 $A(1,4) = A(1,4) - (T(L+2)*WA1)$
 $A(2,3) = A(2,3) + (WB1*WB2 + T(L+2)*WB1WB2)$
 $A(2,4) = A(2,4) - (T(L+2)*WB1)$
 $A(3,4) = A(3,4) - (T(L+2)*WB2)$

W1TEMP = WA1TEMP1
 WA1TEMP1 = WA1
 WA1TEMP2 = W1TEMP
 W2TEMP = WB1TEMP1
 WB1TEMP1 = WB1
 WB1TEMP2 = W2TEMP
 W3TEMP = WB2TEMP1
 WB2TEMP1 = WB2
 WB2TEMP2 = W3TEMP
 WAB1TMP = WAB1TMP1
 WAB1TMP1 = WA1WB1
 WAB1TMP2 = WAB1TMP
 WAB2TMP = WAB2TMP1
 WAB2TMP1 = WA1WB2
 WAB2TMP2 = WAB2TMP
 WWB1TMP = WWB1TMP1
 WWB1TMP1 = WB1WB1
 WWB1TMP2 = WWB1TMP
 WB12TMP = WB12TMP1
 WB12TMP1 = WB1WB2
 WB12TMP2 = WB12TMP
 WWB2TMP = WWB2TMP1
 WWB2TMP1 = WB2WB2
 WWB2TMP2 = WWB2TMP

550 CONTINUE

A(1,1) = A(1,1)/SIGMAET
 A(1,2) = A(1,2)/SIGMAET
 A(1,3) = A(1,3)/SIGMAET
 A(1,4) = A(1,4)/(SIGMAET**2)
 A(2,2) = A(2,2)/SIGMAET
 A(2,3) = A(2,3)/SIGMAET
 A(2,4) = A(2,4)/(SIGMAET**2)
 A(3,3) = A(3,3)/SIGMAET
 A(3,4) = A(3,4)/(SIGMAET**2)
 A(4,4) = A(4,4)*SIGMAET**3 - (N/(2*SIGMAET**2))
 A(1,5) = A(1,5) - (1+2*D2+C1**2)
 A(1,6) = A(1,6) - C1
 A(2,5) = A(2,5) + (1+B22)*SIGMAET
 A(3,5) = A(3,5) + B11*SIGMAET
 A(3,6) = A(3,6) + SIGMAET
 A(4,5) = A(4,5) + B11*(1+B22)
 A(4,6) = A(4,6) + B22
 A(2,1) = A(1,2)
 A(3,1) = A(1,3)
 A(4,1) = A(1,4)
 A(5,1) = A(1,5)

A(6,1) = A(1,6)
A(3,2) = A(2,3)
A(4,2) = A(2,4)
A(5,2) = A(2,5)
A(4,3) = A(3,4)
A(5,3) = A(3,5)
A(6,3) = A(3,6)
A(5,4) = A(4,5)
A(6,4) = A(4,6)

CALL DLINRG(N6, A, LDA, AINV, LDAINV)

IF (AINV(1,1).GT.0.0) THEN
 SS1 = N*(A11 - 1.0)
 SS2 = (A11 - 1.0)/SQRT(AINV(1,1))
 IF (SS1.LT.-5.7) THEN
 COUNT1 = COUNT1 + 1.0
 ENDIF
 IF (SS1.LT.-8.0) THEN
 COUNT2 = COUNT2 + 1.0
 ENDIF
 IF (SS1.LT.-13.6) THEN
 COUNT3 = COUNT3 + 1.0
 ENDIF
 IF (SS2.LT.-1.62) THEN
 COUNT4 = COUNT4 + 1.0
 ENDIF
 IF (SS2.LT.-1.95) THEN
 COUNT5 = COUNT5 + 1.0
 ENDIF
 IF (SS2.LT.-2.58) THEN
 COUNT6 = COUNT6 + 1.0
 ENDIF
 ICOUNT = ICOUNT + 1
ENDIF

RETURN
END

* END OF PROGRAM

APPENDIX B

FORTRAN PROGRAM FOR HANNAN AND RISSANEN'S METHOD
BASED ON SHIN AND SARKAR'S APPROACH

```

*****
* Note that  $(N, S, \sigma_e^2) = (250, 4, 5.0)$  case is considered as example.      *
* In addition, only Subroutine 4 is described because the rest of all subroutines *
* and main program are exactly the same as those of HR-KOHN.                    *
*****

```

```

*****
* SUBROUTINE 4 *
*****

```

```

SUBROUTINE NRMETHOD(N,NN,NNN,N3,N5,R,T,ABTILDA,HAP,
/      SIGMAET,C1,D1,D2,ICOUNT,COUNT1,COUNT2,
/      COUNT3,COUNT4,COUNT5,COUNT6)
PARAMETER(LDA=6, LDAINV=6, N6=6)
REAL*8 A(LDA,LDA), AINV(LDAINV,LDAINV)
REAL*8 R(NN), T(NNN), ABTILDA(N3)
REAL*8 A11, B11, B22, SIGMAET, SS1, SS2

A11 = ABTILDA(1)
B11 = ABTILDA(2)
B22 = ABTILDA(3)

WA1 = 0.0
WB1 = 0.0
WB2 = 0.0
W1TEMP = 0.0
W2TEMP = 0.0
W3TEMP = 0.0
WA1TEMP1 = 0.0
WA1TEMP2 = 0.0
WB1TEMP1 = 0.0
WB1TEMP2 = 0.0
WB2TEMP1 = 0.0
WB2TEMP2 = 0.0

DO 510 I = 1, 6
  DO 500 J = 1, 6
    A(I,J) = 0.0
500 CONTINUE
510 CONTINUE

DO 550 L = 1, N
  WA1 = R(L+3) - B11*WA1TEMP1 - B22*WA1TEMP2
  WB1 = -T(L+1) - B11*WB1TEMP1 - B22*WB1TEMP2
  WB2 = -T(L) - B11*WB2TEMP1 - B22*WB2TEMP2

```

```

A(1,1) = A(1,1) + WA1*WA1
A(2,2) = A(2,2) + WB1*WB1
A(3,3) = A(3,3) + WB2*WB2
W1TEMP = WA1TEMP1
WA1TEMP1 = WA1
WA1TEMP2 = W1TEMP
W2TEMP = WB1TEMP1
WB1TEMP1 = WB1
WB1TEMP2 = W2TEMP
W3TEMP = WB2TEMP1
WB2TEMP1 = WB2
WB2TEMP2 = W3TEMP
550 CONTINUE
A(1,1) = A(1,1)/SIGMAET
A(2,2) = A(2,2)/SIGMAET
A(3,3) = A(3,3)/SIGMAET
A(5,1) = A(5,1) - (1 + 2*D2 + C1**2)
A(2,5) = A(2,5) + (1 + B22)*SIGMAET
A(3,5) = A(3,5) + B11*SIGMAET
A(6,1) = A(6,1) - C1
A(3,6) = A(3,6) + SIGMAET
A(4,4) = A(4,4) - N/(2*(SIGMAET**2))
A(4,5) = A(4,5) + B11*(1 + B22)
A(4,6) = A(4,6) + B22
A(1,5) = A(5,1)
A(1,6) = A(6,1)
A(5,2) = A(2,5)
A(5,3) = A(3,5)
A(6,3) = A(3,6)
A(5,4) = A(4,5)
A(6,4) = A(4,6)

CALL DLINRG(N6, A, LDA, AINV, LDAINV)

IF (AINV(1,1).GT.0.0) THEN
  SS1 = N*(A11 - 1.0)
  SS2 = (A11 - 1.0)/SQRT(AINV(1,1))
  IF (SS1.LT.-5.7) THEN
    COUNT1 = COUNT1 + 1.0
  ENDIF
  IF (SS1.LT.-8.0) THEN
    COUNT2 = COUNT2 + 1.0
  ENDIF
  IF (SS1.LT.-13.6) THEN
    COUNT3 = COUNT3 + 1.0
  ENDIF

```

```
IF (SS2.LT.-1.62) THEN
  COUNT4 = COUNT4 + 1.0
ENDIF
IF (SS2.LT.-1.95) THEN
  COUNT5 = COUNT5 + 1.0
ENDIF
IF (SS2.LT.-2.58) THEN
  COUNT6 = COUNT6 + 1.0
ENDIF
ICOUNT = ICOUNT + 1
ENDIF
```

```
RETURN
END
```

```
* END OF PROGRAM
```

APPENDIX C

FORTRAN PROGRAM FOR KOHN'S METHOD

```
*****
* (N, S,  $\sigma_e^2$ ) = (250, 4, 5.0) is considered as an example. *
*****
```

```
*****
*   MAIN PROGRAM   *
*****
```

```
PROGRAM KOHN
PARAMETER(NUM=10000, N= 250, NN=254, NNN=252, N2=4, N3=3, N5=6)
IMPLICIT REAL*8 (A-H, O-Z)
REAL*8 E(N), W(N), Z(N), U(N), Y(N)
REAL*8 A1(6), C1, D1, D2
REAL*8 T(NNN), TT(NNN), T1(NNN)
REAL*8 R(NN), ETILDA(N), X(N,N2), FL_EST(10)
REAL*8 RR(NNN), ABTILDA(N3), C(N5), HAP, ZIGMAET, SIGMAET
REAL*8 CMIN1, CMIN2, CMIN3, CMIN4, CMIN5, CMIN6
INTEGER ISEED1, ISEED2
EXTERNAL RNSET, DRNNOA, DLINRG
```

```
DATA A1/1.0,0.99,0.95,0.90,0.80,0.70/
```

```
ISEED1 = 13579
CALL RNISD(ISEED1, ISEED2)
```

```
DO 260 I1 = 1, 6
  DO 255 C1 = -0.5, 0.5, 0.5
```

```
    ICOUNT = 0
    COUNT1 = 0.0
    COUNT2 = 0.0
    COUNT3 = 0.0
    COUNT4 = 0.0
    COUNT5 = 0.0
    COUNT6 = 0.0
```

```
    D1 = A1(I1) + C1
    D2 = A1(I1) * C1
```

```
DO 100 I = 1, 18000
  CALL RNSET(ISEED1)
  CALL DRNNOA(N,E)
  CALL RNGET(ISEED1)
  CALL RNSET(ISEED2)
  CALL DRNNOA(N,W)
  CALL RNGET(ISEED2)
```

```

Z(1) = E(1) * SQRT(5.0)
U(1) = W(1)
Y(1) = Z(1) + U(1)

DO 10 J = 2, N
  Z(J) = -A1(I1)*Z(J-1) + E(j)*SQRT(5.0)
  U(J) = W(J) + C1*W(J-1)
  Y(J) = Z(J) + U(J)
10 CONTINUE

DO 1 K = 1, N
  R(K+4) = Y(K)
1 CONTINUE

DO 3 K = 1, 4
  R(K) = 0.0
3 CONTINUE

CALL ETINIT(Y, R, X, N, NN, ETILDA)

DO 5 K = 1, N
  RR(K+2) = ETILDA(K)
5 CONTINUE

DO 7 K = 1, 2
  RR(K) = 0.0
7 CONTINUE

CALL ABINIT(N, NN, NNN, N3, R, RR, ABTILDA)

CALL SIGMAINIT(Y, R, T, N, NN, NNN, N3, ABTILDA,
/      HAP, ZIGMAET)

CALL NR1(N, NN, NNN, N3, N5, R, T, TT, C1, D1, D2, C,
/      ABTILDA, HAP, ZIGMAET, FL-EST, CMIN1, CMIN2,
/      CMIN3, CMIN4, CMIN5, CMIN6)

CALL NR2(N, NN, NNN, N3, N5, R, T1, C1, D1, D2, SIGMAET,
/      CMIN1, CMIN2, CMIN3, CMIN4, CMIN5, CMIN6,
/      ICOUNT, COUNT1, COUNT2,
/      COUNT3, COUNT4, COUNT5, COUNT6)

IF (ICOUNT.GE.10000) GO TO 115

100 CONTINUE
115 CONTINUE

```

```

PWR1 = (COUNT1/NUM)*100.0
PWR2 = (COUNT2/NUM)*100.0
PWR3 = (COUNT3/NUM)*100.0
PWR4 = (COUNT4/NUM)*100.0
PWR5 = (COUNT5/NUM)*100.0
PWR6 = (COUNT6/NUM)*100.0

```

```

WRITE(*,*) 'A1 =', A1(I1), 'C1 =', C1
WRITE(*,*) COUNT1, PWR1
WRITE(*,*) COUNT2, PWR2
WRITE(*,*) COUNT3, PWR3
WRITE(*,*) COUNT4, PWR4
WRITE(*,*) COUNT5, PWR5
WRITE(*,*) COUNT6, PWR6

```

```

255     CONTINUE
260     CONTINUE
      END

```

```

*****

```

```

*   SUBROUTINE 1   *

```

```

*****

```

```

*** ESTIMATE THE INITIAL VALUE OF  $e_t$ .

```

```

SUBROUTINE ETINIT(Y, R, X, N, NN, ETILDA)
PARAMETER(LDA = 4, LDAINV = 4, N2 = 4)
REAL*8 R(NN), A(LDA,LDA), B(N2), AINV(1LDAINV,LDAINV)
REAL*8 X(N,N2), Y(N), ETILDA(N), SUM1(N2), SUM, SUM2

```

```

      DO 200 I = 1, 4
        DO 210 J = 1, 4
          A(I,J) = 0.0
210     CONTINUE
          B(I) = 0.0
200     CONTINUE

```

```

      DO 250 L = 1, N
        A(1,1) = A(1,1) + R(L+3)*R(L+3)
        A(2,1) = A(2,1) + R(L+2)*R(L+3)
        A(3,1) = A(3,1) + R(L+1)*R(L+3)
        A(4,1) = A(4,1) + R(L)*R(L+3)
        A(2,2) = A(2,2) + R(L+2)*R(L+2)
        A(2,3) = A(2,3) + R(L+1)*R(L+2)
        A(2,4) = A(2,4) + R(L)*R(L+2)
        A(3,3) = A(3,3) + R(L+1)*R(L+1)
        A(3,4) = A(3,4) + R(L)*R(L+1)

```



```

      A(4,4) = A(4,4) + R(L)*R(L)
      B(1) = B(1) + R(L+3)*R(L+4)
      B(2) = B(2) + R(L+2)*R(L+4)
      B(3) = B(3) + R(L+1)*R(L+4)
      B(4) = B(4) + R(L)*R(L+4)
250  CONTINUE
      A(1,2) = A(2,1)
      A(1,3) = A(3,1)
      A(1,4) = A(4,1)
      A(3,2) = A(2,3)
      A(4,2) = A(2,4)
      A(4,3) = A(3,4)

      CALL DLINRG(N2, A, LDA, AINV, LDAINV)

      DO 260 K1 = 1, 4
        DO 265 K2 = 1, N
          IF (K2.LE.K1) THEN
            X(K2,K1) = 0.0
          ELSE
            X(K2,K1) = Y(K2-K1)
          ENDIF
265  CONTINUE
260  CONTINUE

      DO 270 LDAINV1 = 1, 4
        SUM = 0.0
        DO 280 LDAINV2 = 1, 4
          SUM = SUM + AINV(LDAINV1,LDAINV2)*B(LDAINV2)
280  CONTINUE
        SUM1(LDAINV1) = SUM
270  CONTINUE

      DO 285 K3 = 1, N
        SUM2 = 0.0
        DO 290 K4 = 1, 4
          SUM2 = SUM2 + X(K3,K4)*SUM1(K4)
290  CONTINUE
        ETILDA(K3) = Y(K3) - SUM2
285  CONTINUE

      RETURN
      END

```

* SUBROUTINE 2 *

*** ESTIMATE THE INITIAL VALUE OF A1, B1 AND B2.

```

SUBROUTINE ABINIT(N, NN, NNN, N3, R, RR, ABTILDA)
PARAMETER(MDA=3, MDAINV=3)
REAL*8 R(NN), RR(NNN), ABTILDA(N3)
REAL*8 E(MDA,MDA), EINV(MDAINV,MDAINV), F(MDA)
REAL*8 SUM3

DO 310 I = 1, 3
  DO 320 J = 1, 3
    E(I,J) = 0.0
320  CONTINUE
    F(I) = 0.0
310  CONTINUE

DO 350 K = 1, N
  E(1,1) = E(1,1) + R(K+3)*R(K+3)
  E(2,1) = E(2,1) - R(K+3)*RR(K+1)
  E(3,1) = E(3,1) - R(K+3)*RR(K)
  E(2,2) = E(2,2) + RR(K+1)*RR(K+1)
  E(2,3) = E(2,3) + RR(K)*RR(K+1)
  E(3,3) = E(3,3) + RR(K)*RR(K)
  F(1) = F(1) - R(K+3)*R(K+4)
  F(2) = F(2) + R(K+4)*RR(K+1)
  F(3) = F(3) + R(K+4)*RR(K)
350  CONTINUE
  E(1,2) = E(2,1)
  E(1,3) = E(3,1)
  E(3,2) = E(2,3)

  CALL DLINRG(3, E, MDA, EINV, MDAINV)

DO 390 MDAINV1 = 1, 3
  SUM3 = 0.0
  DO 370 MDAINV2 = 1, 3
    SUM3 = SUM3 + EINV(MDAINV1,MDAINV2)*F(MDAINV2)
370  CONTINUE
    ABTILDA(MDAINV1) = SUM3
390  CONTINUE

RETURN
END

```

* SUBROUTINE 3 *

*** ESTIMATE THE INITIAL VALUE OF σ_e^2 .

```

SUBROUTINE SIGMAINIT(Y, R, T, N, NN, NNN, N3, ABTILDA,
/
                      HAP, ZIGMAET)
  REAL*8 Y(N), R(NN), T(NNN)
  REAL*8 ABTILDA(N3), HAP, ZIGMAET

  DO 400 L = 1, NNN
    T(L) = 0.0
400  CONTINUE

  HAP = 0.0
  DO 430 L = 1, N
    T(L+2) = R(L+4) + ABTILDA(1)*R(L+3) - ABTILDA(2)*T(L+1)
/
    - ABTILDA(3)*T(L)
    HAP = HAP + T(L+2)*T(L+2)
430  CONTINUE
  ZIGMAET = HAP/N

  RETURN
  END

```

* SUBROUTINE 4 *

*** NEWTON-RAPHSON METHOD BASED ON KOHN'S.

```

SUBROUTINE NR1(N, NN, NNN, N3, N5, R, T, TT, C1, D1, D2, C,
/
                      ABTILDA, HAP, ZIGMAET, FL_EST,
/
                      CMIN1, CMIN2, CMIN3, CMIN4, CMIN5, CMIN6)
  PARAMETER(LDA=6, LDAINV=6, N6=6)
  REAL*8 A(LDA), AINV(LDAINV,LDAINV), B(N6), C(N5)
  REAL*8 R(NN), T(NNN), TT(NNN), ABTILDA(N3)
  REAL*8 D(N6), HAP1
  REAL*8 CC1(10), CC2(10), CC3(10), CC4(10), CC5(10), CC6(10)
  REAL*8 FL_EST(10), FMIN
  REAL*8 CMIN1, CMIN2, CMIN3, CMIN4, CMIN5, CMIN6

  A11 = ABTILDA(1)
  B11 = ABTILDA(2)
  B22 = ABTILDA(3)
  C(5) = 0.0

```

```

C(6) = 0.0
WA1 = 0.0
WB1 = 0.0
WB2 = 0.0
WA1WB1 = 0.0
WA1WB2 = 0.0
WB1WB1 = 0.0
WB1WB2 = 0.0
WB2WB2 = 0.0

DO 500 I = 1, 6
    DO 510 J = 1, 6
        A(I,J) = 0.0
510    CONTINUE
        B(I) = 0.0
500    CONTINUE

W1TEMP = 0.0
W2TEMP = 0.0
W3TEMP = 0.0
WAB1TMP = 0.0
WAB2TMP = 0.0
WWB1TMP = 0.0
WB12TMP = 0.0
WWB2TMP = 0.0
WA1TEMP1 = 0.0
WA1TEMP2 = 0.0
WB1TEMP1 = 0.0
WB1TEMP2 = 0.0
WB2TEMP1 = 0.0
WB2TEMP2 = 0.0
WAB1TMP1 = 0.0
WAB1TMP2 = 0.0
WAB2TMP1 = 0.0
WAB2TMP2 = 0.0
WWB1TMP1 = 0.0
WWB1TMP2 = 0.0
WB12TMP1 = 0.0
WB12TMP2 = 0.0
WWB2TMP1 = 0.0
WWB2TMP2 = 0.0

DO 550 L = 1, N
    WA1 = R(L+3) - B11*WA1TEMP1 - B22*WA1TEMP2
    WB1 = -T(L+1) - B11*WB1TEMP1 - B22*WB1TEMP2
    WB2 = -T(L) - B11*WB2TEMP1 - B22*WB2TEMP2

```

$$\begin{aligned}
WA1WB1 &= -WA1TEMP1 - B11*WAB1TMP1 - B22*WAB1TMP2 \\
WA1WB2 &= -WA1TEMP2 - B11*WAB2TMP1 - B22*WAB2TMP2 \\
WB1WB1 &= -2*WB1TEMP1 - B11*WWB1TMP1 - B22*WWB1TMP2 \\
WB1WB2 &= -WB2TEMP1 - B11*WB12TMP1 - WB1TEMP2 \\
&\quad - B22*WB12TMP2 \\
WB2WB2 &= -2*WB2TEMP2 - B11*WWB2TMP1 - B22*WWB2TMP2
\end{aligned}$$

$$\begin{aligned}
A(1,1) &= A(1,1) + (WA1*WA1) \\
A(2,2) &= A(2,2) + (WB1*WB1 + T(L+2)*WB1WB1) \\
A(3,3) &= A(3,3) + (WB2*WB2 + T(L+2)*WB2WB2) \\
A(4,4) &= A(4,4) + (T(L+2)*T(L+2)) \\
A(1,2) &= A(1,2) + (WA1*WB1 + T(L+2)*WA1WB1) \\
A(1,3) &= A(1,3) + (WA1*WB2 + T(L+2)*WA1WB2) \\
A(1,4) &= A(1,4) - (T(L+2)*WA1) \\
A(2,3) &= A(2,3) + (WB1*WB2 + T(L+2)*WB1WB2) \\
A(2,4) &= A(2,4) - (T(L+2)*WB1) \\
A(3,4) &= A(3,4) - (T(L+2)*WB2) \\
B(1) &= B(1) + WA1*T(L+2) \\
B(2) &= B(2) + WB1*T(L+2) \\
B(3) &= B(3) + WB2*T(L+2) \\
B(4) &= B(4) + T(L+2)*T(L+2)
\end{aligned}$$

$$\begin{aligned}
W1TEMP &= WA1TEMP1 \\
WA1TEMP1 &= WA1 \\
WA1TEMP2 &= W1TEMP \\
W2TEMP &= WB1TEMP1 \\
WB1TEMP1 &= WB1 \\
WB1TEMP2 &= W2TEMP \\
W3TEMP &= WB2TEMP1 \\
WB2TEMP1 &= WB2 \\
WB2TEMP2 &= W3TEMP \\
WAB1TMP &= WAB1TEMP \\
WAB1TMP1 &= WA1WB1 \\
WAB1TMP2 &= WAB1TMP \\
WAB2TMP &= WAB2TMP1 \\
WAB2TMP1 &= WA1WB2 \\
WAB2TMP2 &= WAB2TMP \\
WWB1TMP &= WWB1TMP1 \\
WWB1TMP1 &= WB1WB1 \\
WWB1TMP2 &= WWB1TMP \\
WB12TMP &= WB12TMP1 \\
WB12TMP1 &= WB1WB2 \\
WB12TMP2 &= WB12TMP \\
WWB2TMP &= WWB2TMP1 \\
WWB2TMP1 &= WB2WB2 \\
WWB2TMP2 &= WWB2TMP
\end{aligned}$$

```

550 CONTINUE
      A(1,1) = A(1,1)/ZIGMAET
      A(1,2) = A(1,2)/ZIGMAET
      A(1,3) = A(1,3)/ZIGMAET
      A(1,4) = A(1,4)/(ZIGMAET**2)
      A(2,2) = A(2,2)/ZIGMAET
      A(2,3) = A(2,3)/ZIGMAET
      A(2,4) = A(2,4)/(ZIGMAET**2)
      A(3,3) = A(3,3)/ZIGMAET
      A(3,4) = A(3,4)/(ZIGMAET**2)
      A(4,4) = A(4,4)*ZIGMAET**3 - (N/(2*ZIGMAET**2))
      A(1,5) = A(1,5) - (1+2*D2+C1**2)
      A(1,6) = A(1,6) - C1
      A(2,5) = A(2,5) + (1+B22)*ZIGMAET
      A(3,5) = A(3,5) + B11*ZIGMAET
      A(3,6) = A(3,6) + ZIGMAET
      A(4,5) = A(4,5) + B11*(1+B22)
      A(4,6) = A(4,6) + B22
      B(1) = B(1)/ZIGMAET
      B(2) = B(2)/ZIGMAET
      B(3) = B(3)/ZIGMAET
      B(4) = (-B(4)/(2*ZIGMAET*ZIGMAET)) + (N/(2*ZIGMAET))
      B(5) = ZIGMAET*(1+B22)*B11 - D1*(1+D2)
      B(6) = ZIGMAET*B22 - D2
      A(2,1) = A(1,2)
      A(3,1) = A(1,3)
      A(4,1) = A(1,4)
      A(5,1) = A(1,5)
      A(6,1) = A(1,6)
      A(3,2) = A(2,3)
      A(4,2) = A(2,4)
      A(5,2) = A(2,5)
      A(4,3) = A(3,4)
      A(5,3) = A(3,5)
      A(6,3) = A(3,6)
      A(5,4) = A(4,5)
      A(6,4) = A(4,6)

      CALL DLINRG(N6, A, LDA, AINV, LDAINV)

      DO 580 LDAINV1 = 1, 6
          SUM = 0.0
          DO 570 LDAINV2 = 1, 6
              SUM = SUM + AINV(LDAINV1,LDAINV2)*B(LDAINV2)
570          CONTINUE
          D(LDAINV1) = SUM

```

```

580 CONTINUE

DO 700 K = 1, 10
  C(1) = A11 - (1/K)*D(1)
  C(2) = B11 - (1/K)*D(2)
  C(3) = B22 - (1/K)*D(3)
  C(4) = ZIGMAET - (1/K)*D(4)
  C(5) = C(5) - (1/K)*D(5)
  C(6) = C(6) - (1/K)*D(6)
  HAP1 = 0.0

      DO 600 I1 = 1, NNN
        TT(I1) = 0.0
600 CONTINUE

      DO 650 I2 = 1, N
        TT(I2+2) = R(I2+4) + C(1)*R(I2+3) - C(2)*TT(I2+1) - C(3)*TT(I2)
        HAP1 = HAP1 + TT(I2+2)*TT(I2+2)
650 CONTINUE
      FL_EST(K) = (1/(2*C(4)))*HAP1 + (N/2)*LOG(C(4))
      CC1(K) = C(1)
      CC2(K) = C(2)
      CC3(K) = C(3)
      CC4(K) = C(4)
      CC5(K) = C(5)
      CC6(K) = C(6)
700 CONTINUE

      M = 1
      DO 750 I3 = 2, 10
        IF(I3.EQ.0.OR.FL_EST(I3).GE.FL_EST(M)) GO TO 750
        M = I3
750 CONTINUE
      FMIN = FL_EST(M)
      CMIN1 = CC1(M)
      CMIN2 = CC2(M)
      CMIN3 = CC3(M)
      CMIN4 = CC4(M)
      CMIN5 = CC5(M)
      CMIN6 = CC6(M)

RETURN
END

```

* SUBROUTINE 5 *

```

SUBROUTINE NR2(N, NN, NNN, N3, N5, R, T1, C1, D1, D2, SIGMAET,
/          CMIN1, CMIN2, CMIN3, CMIN4, CMIN5, CMIN6, ICOUNT,
/          COUNT1, COUNT2, COUNT3, COUNT4, COUNT5, COUNT6)
PARAMETER(LDA = 6, LDAINV = 6, N6 = 6)
REAL*8 A(LDA,LDA), AINV(LDAINV,LDAINV)
REAL*8 R(NN), T1(NNN)
REAL*8 CMIN1, CMIN2, CMIN3, CMIN4, CMIN5, CMIN6
REAL*8 AA1, B1, B2, SIGMAET
REAL*8 SS1, SS2

AA1 = CMIN1
B1 = CMIN2
B2 = CMIN3
SIGMAET = CMIN4

DO 800 L = 1, NNN
    T1(L) = 0.0
800 CONTINUE

DO 830 L = 1, N
    T1(L+2) = R(L+4) + AA1*R(L+3) - B1*T1(L+1) - B2*T1(L)
830 CONTINUE

WA1 = 0.0
WB1 = 0.0
WB2 = 0.0
WA1WB1 = 0.0
WA1WB2 = 0.0
WB1WB1 = 0.0
WB1WB2 = 0.0
WB2WB2 = 0.0

DO 870 I = 1, 6
    DO 850 J = 1, 6
        A(I,J) = 0.0
850 CONTINUE
870 CONTINUE

W1TEMP = 0.0
W2TEMP = 0.0
W3TEMP = 0.0
WAB1TMP = 0.0

```


WAB2TMP = 0.0
 WWB1TMP = 0.0
 WB12TMP = 0.0
 WWB2TMP = 0.0
 WA1TEMP1 = 0.0
 WA1TEMP2 = 0.0
 WB1TEMP1 = 0.0
 WB1TEMP2 = 0.0
 WB2TEMP1 = 0.0
 WB2TEMP2 = 0.0
 WAB1TMP1 = 0.0
 WAB1TMP2 = 0.0
 WAB2TMP1 = 0.0
 WAB2TMP2 = 0.0
 WWB1TMP1 = 0.0
 WWB1TMP2 = 0.0
 WB12TMP1 = 0.0
 WB12TMP2 = 0.0
 WWB2TMP1 = 0.0
 WWB2TMP2 = 0.0

DO 900 L = 1, N

WA1 = R(L+3) - B1*WA1TEMP1 - B2*WA1TEMP2
 WB1 = -T1(L+1) - B1*WB1TEMP1 - B2*WB1TEMP2
 WB2 = -T1(L) - B1*WB2TEMP1 - B2*WB2TEMP2
 WA1WB1 = -WA1TEMP1 - B1*WAB1TMP1 - B2*WAB1TMP2
 WA1WB2 = -WA1TEMP2 - B1*WAB2TMP1 - B2*WAB2TMP2
 WB1WB1 = -2*WB1TEMP1 - B1*WWB1TMP1 - B2*WWB1TMP2
 WB1WB2 = -WB2TEMP1 - B1*WB12TMP1 - WB1TEMP2
 - B2*WB12TMP2
 WB2WB2 = -2*WB2TEMP2 - B1*WWB2TMP1 - B2*WWB2TMP2

A(1,1) = A(1,1) + (WA1*WA1)
 A(2,2) = A(2,2) + (WB1*WB1 + T1(L+2)*WB1WB1)
 A(3,3) = A(3,3) + (WB2*WB2 + T1(L+2)*WB2WB2)
 A(4,4) = A(4,4) + (T1(L+2)*T1(L+2))
 A(1,2) = A(1,2) + (WA1*WB1 + T1(L+2)*WA1WB1)
 A(1,3) = A(1,3) + (WA1*WB2 + T1(L+2)*WA1WB2)
 A(1,4) = A(1,4) - (T1(L+2)*WA1)
 A(2,3) = A(2,3) + (WB1*WB2 + T1(L+2)*WB1WB2)
 A(2,4) = A(2,4) - (T1(L+2)*WB1)
 A(3,4) = A(3,4) - (T1(L+2)*WB2)

W1TEMP = WA1TEMP1
 WA1TEMP1 = WA1
 WA1TEMP2 = W1TEMP

```

W2TEMP = WB1TEMP1
WB1TEMP1 = WB1
WB1TEMP2 = W2TEMP
W3TEMP = WB2TEMP1
WB2TEMP1 = WB2
WB2TEMP2 = W3TEMP
WAB1TMP = WAB1TMP1
WAB1TMP1 = WA1WB1
WAB1TMP2 = WAB1TMP
WAB2TMP = WAB2TMP1
WAB2TMP1 = WA1WB2
WAB2TMP2 = WAB2TMP
WWB1TMP = WWB1TMP1
WWB1TMP1 = WB1WB1
WWB1TMP2 = WWB1TMP
WB12TMP = WB12TMP1
WB12TMP1 = WB1WB2
WB12TMP2 = WB12TMP
WWB2TMP = WWB2TMP1
WWB2TMP1 = WB2WB2
WWB2TMP2 = WWB2TMP
900 CONTINUE
A(1,1) = A(1,1)/SIGMAET
A(1,2) = A(1,2)/SIGMAET
A(1,3) = A(1,3)/SIGMAET
A(1,4) = A(1,4)/(SIGMAET**2)
A(2,2) = A(2,2)/SIGMAET
A(2,3) = A(2,3)/SIGMAET
A(2,4) = A(2,4)/(SIGMAET**2)
A(3,3) = A(3,3)/SIGMAET
A(3,4) = A(3,4)/(SIGMAET**2)
A(4,4) = A(4,4)*SIGMAET**3 - (N/(2*SIGMAET**2))
A(1,5) = A(1,5) - (1+2*D2+C1**2)
A(1,6) = A(1,6) - C1
A(2,5) = A(2,5) + (1+B2)*SIGMAET
A(3,5) = A(3,5) + B1*SIGMAET
A(3,6) = A(3,6) + SIGMAET
A(4,5) = A(4,5) + B1*(1+B2)
A(4,6) = A(4,6) + B2
A(2,1) = A(1,2)
A(3,1) = A(1,3)
A(4,1) = A(1,4)
A(5,1) = A(1,5)
A(6,1) = A(1,6)
A(3,2) = A(2,3)
A(4,2) = A(2,4)

```

```
A(5,2) = A(2,5)
A(4,3) = A(3,4)
A(5,3) = A(3,5)
A(6,3) = A(3,6)
A(5,4) = A(4,5)
A(6,4) = A(4,6)
```

```
CALL DLINRG(N6, A, LDA, AINV, LDAINV)
```

```
IF (AINV(1,1).GT.0.0) THEN
  SS1 = N*(AA1 - 1.0)
  SS2 = (AA1 - 1.0)/SQTR(AINV(1,1))
  IF (SS1.LT.-5.7) THEN
    COUNT1 = COUNT1 + 1.0
  ENDIF
  IF (SS1.LT.-8.0) THEN
    COUNT2 = COUNT2 + 1.0
  ENDIF
  IF (SS1.LT.-13.6) THEN
    COUNT3 = COUNT3 + 1.0
  ENDIF
  IF (SS2.LT.-1.62) THEN
    COUNT4 = COUNT4 + 1.0
  ENDIF
  IF (SS2.LT.-1.95) THEN
    COUNT5 = COUNT5 + 1.0
  ENDIF
  IF (SS2.LT.-2.58) THEN
    COUNT6 = COUNT6 + 1.0
  ENDIF
  ICOUNT = ICOUNT + 1
ENDIF
```

```
RETURN
END
```

```
* END OF PROGRAM
```

APPENDIX D

FORTRAN PROGRAM FOR SHIN AND SARKAR'S METHOD

```

*****
*   Note that  $(n, s, \sigma_e^2) = (250, 4, 5.0)$  is considered as an example.           *
*   This program is the same as that of Kohn's except Subroutine 4 and 5 and thus   *
*   we have only to replace Subroutine 4 and 5 in Kohn's with these two subroutines.*
*****

```

```
*****
```

```
*   SUBROUTINE 4   *
```

```
*****
```

```
*** NEWTON-RAPHSON METHOD BASED ON SHIN AND SARKAR'S
```

```

SUBROUTINE NR1(N,NN,NNN,N3,N5,R,T,ABTILDA,HAP,ZIGMAET,TT,
/   C1,D1,D2,C,FL_EST,CMIN1,CMIN2,CMIN3,CMIN4,CMIN5,CMIN6)
PARAMETER(LDA=6, LDAINV= 6, N6=6)
REAL*8 A(LDA,LDA), B(N6), C(N5), AINV(LDAINV,LDAINV)
REAL*8 R(NN), T(NNN), ABTILDA(N3), TT(NNN)
REAL*8 D(N6), HAP1
REAL*8 CC1(10), CC2(10), CC3(10), CC4(10), CC5(10), CC6(10)
REAL*8 FL_EST(10), FMIN
REAL*8 CMIN1, CMIN2, CMIN3, CMIN4, CMIN5, CMIN6

A11 = ABTILDA(1)
B11 = ABTILDA(2)
B22 = ABTILDA(3)
C(5) = 0.0
C(6) = 0.0

WA1 = 0.0
WB1 = 0.0
WB2 = 0.0
W1TEMP = 0.0
W2TEMP = 0.0
W3TEMP = 0.0
WA1TEMP1 = 0.0
WA1TEMP2 = 0.0
WB1TEMP1 = 0.0
WB1TEMP2 = 0.0
WB2TEMP1 = 0.0
WB2TEMP2 = 0.0

DO 510 I = 1, 6
  DO 500 J = 1, 6
    A(I, J) = 0.0
500  CONTINUE
    B(I) = 0.0
510  CONTINUE

```

```

DO 550 L = 1, N
  WA1 = R(L+3) - B11*WA1TEMP1 - B22*WA1TEMP2
  WB1 = -T(L+1) - B11*WB1TEMP1 - B22*WB1TEMP2
  WB2 = -T(L) - B11*WB2TEMP1 - B22*WB2TEMP2
  AA(1,1) = AA(1,1) + WA1*WA1
  AA(2,2) = AA(2,2) + WB1*WB1
  AA(3,3) = AA(3,3) + WB2*WB2
  BB(1) = BB(1) + WA1*T(L+2)
  BB(2) = BB(2) + WB1*T(L+2)
  BB(3) = BB(3) + WB2*T(L+2)
  BB(4) = BB(4) + T(L+2)*T(L+2)
  W1TEMP = WA1TEMP1
  WA1TEMP1 = WA1
  WA1TEMP2 = W1TEMP
  W2TEMP = WB1TEMP1
  WB1TEMP1 = WB1
  WB1TEMP2 = W2TEMP
  W3TEMP = WB2TEMP1
  WB2TEMP1 = WB2
  WB2TEMP2 = W3TEMP
550 CONTINUE
  A(1,1) = A(1,1)/ZIGMAET
  A(2,2) = A(2,2)/ZIGMAET
  A(3,3) = A(3,3)/ZIGMAET
  A(5,1) = A(5,1) - (1+2*D2 + C1**2)
  A(2,5) = A(2,5) + (1+B22)*ZIGMAET
  A(3,5) = A(3,5) + B11*ZIGMAET
  A(6,1) = A(6,1) - C1
  A(3,6) = A(3,6) + ZIGMAET
  A(4,4) = A(4,4) - N/(2*(ZIGMAET**2))
  A(4,5) = A(4,5) + B11*(1+B22)
  A(4,6) = A(4,6) + B22
  B(1) = B(1)/ZIGMAET
  B(2) = B(2)/ZIGMAET
  B(3) = B(3)/ZIGMAET
  B(4) = (-B(4)/(2*ZIGMAET*ZIGMAET))+(N/(2*ZIGMAET))
  B(5) = ZIGMAET*(1+B22)*B11 - D1*(1+D2)
  B(6) = ZIGMAET*B22 - D2
  A(1,5) = A(5,1)
  A(1,6) = A(6,1)
  A(5,2) = A(2,5)
  A(5,3) = A(3,5)
  A(6,3) = A(3,6)
  A(5,4) = A(4,5)
  A(6,4) = A(4,6)

```

```

CALL DLINRG(N6, A, LDA, AINV, LDAINV)

DO 580 LDAINV1 = 1, 6
    SUM = 0.0
    DO 570 LDAINV2 = 1, 6
        SUM = SUM + AINV(LDAINV1,LDAINV2)*B(LDAINV2)
570    CONTINUE
    D(LDAINV1) = SUM
580    CONTINUE

DO 800 K = 1, 10
    C(1) = A11 - (1/K)*D(1)
    C(2) = B11 - (1/K)*D(2)
    C(3) = B22 - (1/K)*D(3)
    C(4) = ZIGMAET - (1/K)*D(4)
    C(5) = C(5) - (1/K)*D(5)
    C(6) = C(6) - (1/K)*D(6)

    HAP1 = 0.0
    DO 650 I1 = 1, NNN
        TT(I1) = 0.0
650    CONTINUE

    DO 700 I2 = 1, N
        TT(I2+2) = R(I2+4) + C(1)*R(I2+3) - C(2)*TT(I2+1) - C(3)*TT(I2)
        HAP1 = HAP1 + TT(I2+2)*TT(I2+2)
700    CONTINUE
    FL_EST(K) = (1/(2*C(4)))*HAP1 + (N/2)*LOG(C(4))
    CC1(K) = C(1)
    CC2(K) = C(2)
    CC3(K) = C(3)
    CC4(K) = C(4)
    CC5(K) = C(5)
    CC6(K) = C(6)
800    CONTINUE

    M = 1
    DO 850 I3 = 2, 10
        IF(I3.EQ.0.OR.FL_EST(I3).GE.FL_EST(M)) GO TO 850
        M = I3
850    CONTINUE
    FMIN = FL_EST(M)
    CMIN1 = CC1(M)
    CMIN2 = CC2(M)
    CMIN3 = CC3(M)
    CMIN4 = CC4(M)

```

CMIN5 = CC5(M)
 CMIN6 = CC6(M)

RETURN
 END

* SUBROUTINE 5 *

```

SUBROUTINE NR2(N,NN,NNN,N3,N5,R,T1,C1,D1,D2,SIGMAET,
/          CMIN1,CMIN2,CMIN3,CMIN4,CMIN5,CMIN6,ICOUNT,
/          COUNT1,COUNT2,COUNT3,COUNT4,COUNT5,COUNT6)
  PARAMETER(LDA=6, LDAINV=6, N6=6)
  REAL*8 A(LDA,LDA), AINV(LDAINV,LDAINV)
  REAL*8 R(NN), T1(NNN)
  REAL*8 AA1, B1, B2, SIGMAET
  REAL*8 CMIN1, CMIN2, CMIN3, CMIN4
  REAL*8 SS1, SS2

  AA1 = CMIN1
  B1 = CMIN2
  B2 = CMIN3
  SIGMAET = CMIN4

  DO 900 L = 1, NNN
    T1(L) = 0.0
900  CONTINUE

  DO 930 L = 1, N
    T1(L+2) = R(L+4) + AA1*R(L+3) - B1*T1(L+1) - B2*T1(L)
930  CONTINUE

  DO 950 I = 1, 6
    DO 940 J = 1, 6
      A(I,J) = 0.0
940  CONTINUE
950  CONTINUE

  WA1 = 0.0
  WB1 = 0.0
  WB2 = 0.0
  W1TEMP = 0.0
  W2TEMP = 0.0
  W3TEMP = 0.0

```



```

WA1TEMP1 = 0.0
WA1TEMP2 = 0.0
WB1TEMP1 = 0.0
WB1TEMP2 = 0.0
WB2TEMP1 = 0.0
WB2TEMP2 = 0.0

```

```
DO 980 L = 1, N
```

```

    WA1 = R(L+3) - B1*WA1TEMP1 - B2*WA1TEMP2
    WB1 = -T1(L+1) - B1*WB1TEMP1 - B2*WB1TEMP2
    WB2 = -T1(L) - B1*WB2TEMP1 - B2*WB2TEMP2
    A(1,1) = A(1,1) + WA1*WA1
    A(2,2) = A(2,2) + WB1*WB1
    A(3,3) = A(3,3) + WB2*WB2
    W1TEMP = WA1TEMP1
    WA1TEMP1 = WA1
    WA1TEMP2 = W1TEMP
    W2TEMP = WB1TEMP1
    WB1TEMP1 = WB1
    WB1TEMP2 = W2TEMP
    W3TEMP = WB2TEMP1
    WB2TEMP1 = WB2
    WB2TEMP2 = W3TEMP

```

```
980 CONTINUE
```

```

    A(1,1) = A(1,1)/SIGMAET
    A(2,2) = A(2,2)/SIGMAET
    A(3,3) = A(3,3)/SIGMAET
    A(5,1) = A(5,1) - (1+2*D2 + C1**2)
    A(2,5) = A(2,5) + (1+B2)*SIGMAET
    A(3,5) = A(3,5) + B1*SIGMAET
    A(6,1) = A(6,1) - C1
    A(3,6) = A(3,6) + SIGMAET
    A(4,4) = A(4,4) - N/(2*(SIGMAET**2))
    A(4,5) = A(4,5) + B1*(1+B2)
    A(4,6) = A(4,6) + B2
    A(1,5) = A(5,1)
    A(1,6) = A(6,1)
    A(5,2) = A(2,5)
    A(5,3) = A(3,5)
    A(6,3) = A(3,6)
    A(5,4) = A(4,5)
    A(6,4) = A(4,6)

```

```
CALL DLINRG(N6, A, LDA, AINV, LDAINV)
```

```
IF (AINV(1,1).GT.0.0) THEN
```

```
SS1 = N*(AA1 - 1.0)
SS2 = (AA1 - 1.0)/SQRT(AINV(1,1))
```

```
IF (SS1.LT.-5.7) THEN
  COUNT1 = COUNT1 + 1.0
ENDIF
IF (SS1.LT.-8.0) THEN
  COUNT2 = COUNT2 + 1.0
ENDIF
IF (SS1.LT.-13.6) THEN
  COUNT3 = COUNT3 + 1.0
ENDIF
IF (SS2.LT.-1.62) THEN
  COUNT4 = COUNT4 + 1.0
ENDIF
IF (SS2.LT.-1.95) THEN
  COUNT5 = COUNT5 + 1.0
ENDIF
IF (SS2.LT.-2.58) THEN
  COUNT6 = COUNT6 + 1.0
ENDIF
ICOUNT = ICOUNT + 1
ENDIF
```

```
RETURN
END
```

```
* END OF PROGRAM
```

APPENDIX E

FORTRAN PROGRAM FOR
EMPIRICAL CUMULATIVE DISTRIBUTION OF
UNIT ROOT TEST STATISTICS
BASED ON
SHIN AND SARKAR'S METHOD

```

*****
* (N, S,  $\sigma_e^2$ ) = (250, 4, 5.0) is considered as an example. *
*****
*****
*   MAIN PROGRAM   *
*****

```

```
PROGRAM MAIN
```

```
PARAMETER(N=250,NUM=10000,NN=254,NNN=252,N2=4,N3=3,N5=6)
```

```
PARAMETER(I1=100,I2=250,I3=500,I4=1000)
```

```
PARAMETER(I5=9000,I6=9500,I7=9750,I8=9900)
```

```
IMPLICIT REAL*8 (A-H, O-Z)
```

```
REAL*8 E(N), W(N), Z(N), U(N), Y(N)
```

```
REAL*8 STAT1(NUM), STAT2(NUM), SS1, SS2
```

```
REAL*8 A1, C1, D1, D2
```

```
REAL*8 FL_INIT, FL_EST(10)
```

```
REAL*8 T(NNN), TT(NNN), T1(NNN)
```

```
REAL*8 R(NN), ETILDA(N), X(N,N2)
```

```
REAL*8 RR(NNN), ABTILDA(N3), C(N5), HAP, SIGMAET, ZIGMAET
```

```
REAL*8 CMIN1, CMIN2, CMIN3, CMIN4, CMIN5, CMIN6
```

```
INTEGER ISEED1, ISEED2
```

```
EXTERNAL RNSET, DRNNOA, DLINRG
```

```
A1 = 1.0
```

```
ISEED1 = 13579
```

```
CALL RNISD(ISEED1, ISEED2)
```

```
DO 260 C1 = -0.5, 0.5, 0.5
```

```
ICOUNT = 0
```

```
D1 = A1 + C1
```

```
D2 = A1 * C1
```

```
DO 100 I = 1, 18000
```

```
CALL RNSET(ISEED1)
```

```
CALL DRNNOA(N, E)
```

```
CALL RNGET(ISEED1)
```

```
CALL RNSET(ISEED2)
```

```
CALL DRNNOA(N, W)
```

```
CALL RNGET(ISEED2)
```

```
Z(1) = E(1) * SQRT(5.0)
```

```
U(1) = W(1)
```

```
Y(1) = Z(1) + U(1)
```

```
DO 10 J = 2, N
```

```

      Z(J) = -A1*Z(J-1) + E(J)*SQRT(5.0)
      U(J) = W(J) + C1*W(J-1)
      Y(J) = Z(J) + U(J)
10   CONTINUE

      DO 1 K = 1, N
          R(K+4) = Y(K)
1   CONTINUE

      DO 3 K = 1, 4
          R(K) = 0.0
3   CONTINUE

      CALL ETINIT(Y,R,X,N,NN,ETILDA)

      DO 5 K = 1, N
          RR(K+2) = ETILDA(K)
5   CONTINUE

      DO 7 K = 1, 2
          RR(K) = 0.0
7   CONTINUE

      CALL ABINIT(N,ABTILDA,R,RR,NN,NNN,N3)

      CALL SIGMAINIT(Y,R,T,N,NN,NNN,ABTILDA,N3,HAP,SIGMAET,FL_INIT)

      CALL NR1(N,NN,NNN,N3,N5,R,T,ABTILDA,HAP,SIGMAET,
/          TT,C1,D1,D2,C,FL_INIT,FL_EST,
/          CMIN1,CMIN2,CMIN3,CMIN4,CMIN5,CMIN6)

      CALL NR2(N,NN,NNN,N3,N5,R,T1,C1,D1,D2,ZIGMAET,
/          CMIN1,CMIN2,CMIN3,CMIN4,CMIN5,CMIN6,
/          ICOUNT,SS1,SS2)

      STAT1(ICOUNT) = SS1
      STAT2(ICOUNT) = SS2

      IF (ICOUNT.GE.10000) GO TO 112

100  CONTINUE
112  CONTINUE

      CALL SORT(NUM,STAT1)
      CALL SORT(NUM,STAT2)

```

```

PRINT*,I1,'-th SMALLEST = ', STAT1(I1)
PRINT*,I2,'-th SMALLEST = ', STAT1(I2)
PRINT*,I3,'-th SMALLEST = ', STAT1(I3)
PRINT*,I4,'-th SMALLEST = ', STAT1(I4)
PRINT*,I5,'-th SMALLEST = ', STAT1(I5)
PRINT*,I6,'-th SMALLEST = ', STAT1(I6)
PRINT*,I7,'-th SMALLEST = ', STAT1(I7)
PRINT*,I8,'-th SMALLEST = ', STAT1(I8)
PRINT*,I1,'-th SMALLEST = ', STAT2(I1)
PRINT*,I2,'-th SMALLEST = ', STAT2(I2)
PRINT*,I3,'-th SMALLEST = ', STAT2(I3)
PRINT*,I4,'-th SMALLEST = ', STAT2(I4)
PRINT*,I5,'-th SMALLEST = ', STAT2(I5)
PRINT*,I6,'-th SMALLEST = ', STAT2(I6)
PRINT*,I7,'-th SMALLEST = ', STAT2(I7)
PRINT*,I8,'-th SMALLEST = ', STAT2(I8)

```

260 CONTINUE

END

* SUBROUTINE 1 *

*** ESTIMATE THE INITIAL VALUE OF ET.

```

SUBROUTINE ETINIT(Y,R,X,N,NN,ETILDA)
PARAMETER(LDA=4, LDAINV=4, N2=4)
REAL*8 R(NN), A(LDA,LDA), B(N2), AINV(LDAINV,LDAINV)
REAL*8 X(N,N2), Y(N), ETILDA(N), SUM1(N2)

```

```

DO 200 I = 1, 4
  DO 210 J = 1, 4
    A(I,J) = 0.0
210 CONTINUE
  B(I) = 0.0
200 CONTINUE

```

```

DO 250 L = 1, N
  A(1,1) = A(1,1) + R(L+3)*R(L+3)
  A(2,1) = A(2,1) + R(L+2)*R(L+3)
  A(3,1) = A(3,1) + R(L+1)*R(L+3)
  A(4,1) = A(4,1) + R(L)*R(L+3)
  A(2,2) = A(2,2) + R(L+2)*R(L+2)

```

```

A(2,3) = A(2,3) + R(L+1)*R(L+2)
A(2,4) = A(2,4) + R(L)*R(L+2)
A(3,3) = A(3,3) + R(L+1)*R(L+1)
A(3,4) = A(3,4) + R(L)*R(L+1)
A(4,4) = A(4,4) + R(L)*R(L)
B(1) = B(1) + R(L+3)*R(L+4)
B(2) = B(2) + R(L+2)*R(L+4)
B(3) = B(3) + R(L+1)*R(L+4)
B(4) = B(4) + R(L)*R(L+4)
250 CONTINUE
A(1,2) = A(2,1)
A(1,3) = A(3,1)
A(1,4) = A(4,1)
A(3,2) = A(2,3)
A(4,2) = A(2,4)
A(4,3) = A(3,4)

CALL INVMAT(AINV,A,4,4)

DO 260 K1 = 1, 4
  DO 265 K2 = 1, N
    IF (K2.LE.K1) THEN
      X(K2,K1) = 0.0
    ELSE
      X(K2,K1) = Y(K2-K1)
    ENDIF
  CONTINUE
265 CONTINUE
260 CONTINUE

  DO 270 LDAINV1 = 1, 4
    SUM = 0.0
    DO 280 LDAINV2 = 1, 4
      SUM = SUM + AINV(LDAINV1,LDAINV2)*B(LDAINV2)
280 CONTINUE
    SUM1(LDAINV1) = SUM
270 CONTINUE

  DO 285 K3 = 1, N
    SUM2 = 0.0
    DO 290 K4 = 1, 4
      SUM2 = SUM2 + X(K3,K4)*SUM1(K4)
290 CONTINUE
    ETILDA(K3) = Y(K3) - SUM2
285 CONTINUE
RETURN
END

```

* SUBROUTINE 2 *

*** ESTIMATE THE INITIAL VALUE OF A1, B1 AND B2.

SUBROUTINE ABINIT(N,ABTILDA,R,RR,NN,NNN,N3)

PARAMETER(MDA=3, MDAINV=3)

REAL*8 R(NN), RR(NNN), ABTILDA(N3)

REAL*8 A(MDA,MDA), AINV(MDAINV,MDAINV), F(MDA)

DO 310 I = 1, 3

DO 320 J = 1, 3

A(I,J) = 0.0

320 CONTINUE

F(I) = 0.0

310 CONTINUE

DO 350 K = 1, N

A(1,1) = A(1,1) + R(K+3)*R(K+3)

A(2,1) = A(2,1) - R(K+3)*RR(K+1)

A(3,1) = A(3,1) - R(K+3)*RR(K)

A(2,2) = A(2,2) + RR(K+1)*RR(K+1)

A(2,3) = A(2,3) + RR(K)*RR(K+1)

A(3,3) = A(3,3) + RR(K)*RR(K)

F(1) = F(1) - R(K+3)*R(K+4)

F(2) = F(2) + R(K+4)*RR(K+1)

F(3) = F(3) + R(K+4)*RR(K)

350 CONTINUE

A(1,2) = A(2,1)

A(1,3) = A(3,1)

A(3,2) = A(2,3)

CALL INVMAT(AINV,A,3,3)

DO 370 MDAINV1 = 1, 3

SUM3 = 0.0

DO 390 MDAINV2 = 1, 3

SUM3 = SUM3 + AINV(MDAINV1,MDAINV2)*F(MDAINV2)

390 CONTINUE

ABTILDA(MDAINV1) = SUM3

IF (ABTILDA(1).LT.-2.0) THEN

ABTILDA(1) = -2.0

ELSE IF (ABTILDA(1).GT.2.0) THEN

ABTILDA(1) = 2.0

ENDIF

370 CONTINUE

RETURN
END

* SUBROUTINE 3 *

*** ESTIMATE THE INITIAL VALUE OF SIGMAET.

SUBROUTINE SIGMAINIT(Y,R,T,N,NN,NNN,ABTILDA,N3,HAP,
/ SIGMAET,FL_INIT)

REAL*8 Y(N), R(NN), T(NNN)

REAL*8 ABTILDA(N3), HAP, SIGMAET

DO 400 L = 1, NNN

T(L) = 0.0

400 CONTINUE

HAP = 0.0

DO 430 L = 1, N

T(L+2) = R(L+4) + ABTILDA(1)*R(L+3) - ABTILDA(2)*T(L+1)
/ - ABTILDA(3)*T(L)

HAP = HAP + T(L+2)*T(L+2)

430 CONTINUE

SIGMAET = HAP/N

IF (SIGMAET.LT.EXP(-100.0)) THEN

SIGMAET = EXP(-100.0)

ELSE IF (SIGMAET.GT.EXP(10.0)) THEN

SIGMAET = EXP(10.0)

ENDIF

FL_INIT = (1/(2*SIGMAET))*HAP + (N/2)*LOG(SIGMAET)

RETURN

END

* SUBROUTINE 4 *

*** NEWTON-RAPHSON METHOD BASED ON SS'S APPROACH.

SUBROUTINE NR1(N,NN,NNN,N3,N5,R,T,ABTILDA,HAP,SIGMAET,TT,
/ C1,D1,D2,C,FL_INIT,FL_EST,

```

/      CMIN1,CMIN2,CMIN3,CMIN4,CMIN5,CMIN6)
PARAMETER(LDA=6, LDAINV=6, N6=6)
REAL*8 A(LDA,LDA), B(N6), C(N5), AINV(LDAINV,LDAINV)
REAL*8 R(NN), T(NNN), ABTILDA(N3), TT(NNN)
REAL*8 D(N6), HAP1
REAL*8 CC1(10), CC2(10), CC3(10), CC4(10), CC5(10), CC6(10)
REAL*8 FL_EST(10), FMIN, FL_INIT
REAL*8 CMIN1, CMIN2, CMIN3, CMIN4, CMIN5, CMIN6

      A11 = ABTILDA(1)
      B11 = ABTILDA(2)
      B22 = ABTILDA(3)
      C(5) = 0.0
      C(6) = 0.0
      WA1 = 0.0
      WB1 = 0.0
      WB2 = 0.0

      DO 500 I = 1, 6
        DO 510 J = 1, 6
          A(I,J) = 0.0
510    CONTINUE
        B(I) = 0.0
500    CONTINUE

      W1TEMP = 0.0
      W2TEMP = 0.0
      W3TEMP = 0.0
      WA1TEMP1 = 0.0
      WA1TEMP2 = 0.0
      WB1TEMP1 = 0.0
      WB1TEMP2 = 0.0
      WB2TEMP1 = 0.0
      WB2TEMP2 = 0.0

      DO 550 L = 1, N
        WA1 = R(L+3) - B11*WA1TEMP1 - B22*WA1TEMP2
        WB1 = -T(L+1) - B11*WB1TEMP1 - B22*WB1TEMP2
        WB2 = -T(L) - B11*WB2TEMP1 - B22*WB2TEMP2
        A(1,1) = A(1,1) + WA1*WA1
        A(2,2) = A(2,2) + WB1*WB1
        A(3,3) = A(3,3) + WB2*WB2
        B(1) = B(1) + WA1*T(L+2)
        B(2) = B(2) + WB1*T(L+2)
        B(3) = B(3) + WB2*T(L+2)
        B(4) = B(4) + T(L+2)*T(L+2)

```

```

W1TEMP = WA1TEMP1
WA1TEMP1 = WA1
WA1TEMP2 = W1TEMP
W2TEMP = WB1TEMP1
WB1TEMP1 = WB1
WB1TEMP2 = W2TEMP
W3TEMP = WB2TEMP1
WB2TEMP1 = WB2
WB2TEMP2 = W3TEMP
550  CONTINUE
A(1,1) = A(1,1)/SIGMAET
A(2,2) = A(2,2)/SIGMAET
A(3,3) = A(3,3)/SIGMAET
A(5,1) = A(5,1) - (1+2*D2+C1**2)
A(2,5) = A(2,5) + (1+B22)*SIGMAET
A(3,5) = A(3,5) + B11*SIGMAET
A(6,1) = A(6,1) - C1
A(3,6) = A(3,6) + SIGMAET
A(4,4) = A(4,4) - N/(2*(SIGMAET**2))
A(4,5) = A(4,5) + B11*(1+B22)
A(4,6) = A(4,6) + B22
B(1) = B(1)/SIGMAET
B(2) = B(2)/SIGMAET
B(3) = B(3)/SIGMAET
B(4) = (-B(4)/(2*SIGMAET*SIGMAET)) + (N/(2*SIGMAET))
B(5) = SIGMAET*(1+B22)*B11 - D1*(1+D2)
B(6) = SIGMAET*B22 - D2
A(1,5) = A(5,1)
A(1,6) = A(6,1)
A(5,2) = A(2,5)
A(5,3) = A(3,5)
A(6,3) = A(3,6)
A(5,4) = A(4,5)
A(6,4) = A(4,6)

CALL INVMAT(AINV,A,6,6)

DO 570 LDAINV1 = 1, 6
  SUM = 0.0
  DO 580 LDAINV2 = 1, 6
    SUM = SUM + AINV(LDAINV1,LDAINV2)*B(LDAINV2)
580  CONTINUE
  D(LDAINV1) = SUM
570  CONTINUE

DO 600 K = 1, 10

```

```

C(1) = A11 - (1/K)*D(1)
C(2) = B11 - (1/K)*D(2)
C(3) = B22 - (1/K)*D(3)
C(4) = SIGMAET - (1/K)*D(4)
C(5) = C(5) - (1/K)*D(5)
C(6) = C(6) - (1/K)*D(6)
HAP1 = 0.0

DO 650 I1 = 1, NNN
  TT(I1) = 0.0
650 CONTINUE

DO 700 I2 = 1, N
  TT(I2+2) = R(I2+4) + C(1)*R(I2+3) - C(2)*TT(I2+1)
  /      - C(3)*TT(I2)
  HAP1 = HAP1 + TT(I2+2)*TT(I2+2)
700 CONTINUE

IF (C(4).LT.EXP(-100.0)) THEN
  C(4) = EXP(-100.0)
ELSE IF (C(4).GT.EXP(10.0)) THEN
  C(4) = EXP(10.0)
ENDIF

FL_EST(K) = (1/(2*C(4)))*HAP1 + (N/2)*LOG(C(4))
CC1(K) = C(1)
CC2(K) = C(2)
CC3(K) = C(3)
CC4(K) = C(4)
CC5(K) = C(5)
CC6(K) = C(6)
600 CONTINUE

M = 1
DO 800 I3 = 2, 10
  IF(I3.EQ.0.OR.FL_EST(I3).GE.FL_EST(M)) GO TO 800
  M = I3
800 CONTINUE
FMIN = FL_EST(M)
CMIN1 = CC1(M)
CMIN2 = CC2(M)
CMIN3 = CC3(M)
CMIN4 = CC4(M)
CMIN5 = CC5(M)
CMIN6 = CC6(M)

```

RETURN
END

* SUBROUTINE 5 *

SUBROUTINE NR2(N,NN,NNN,N3,N5,R,T1,C1,D1,D2,ZIGMAET,
/ CMIN1,CMIN2,CMIN3,CMIN4,CMIN5,CMIN6,ICOUNT,SS1,SS2)

PARAMETER(LDA=6, LDAINV=6, N6=6)
REAL*8 A(LDA,LDA), AINV(LDAINV,LDAINV)
REAL*8 R(NN), T1(NNN)
REAL*8 AA1, B1, B2, ZIGMAET
REAL*8 CMIN1, CMIN2, CMIN3, CMIN4
REAL*8 SS1, SS2
AA1 = CMIN1
B1 = CMIN2
B2 = CMIN3
ZIGMAET = CMIN4

DO 400 L = 1, NNN
T1(L) = 0.0
400 CONTINUE

DO 430 L = 1, N
T1(L+2) = R(L+4) + AA1*R(L+3) - B1*T1(L+1) - B2*T1(L)
430 CONTINUE

DO 500 I = 1, 6
DO 510 J = 1, 6
A(I,J) = 0.0
510 CONTINUE
500 CONTINUE

WA1 = 0.0
WB1 = 0.0
WB2 = 0.0
W1TEMP = 0.0
W2TEMP = 0.0
W3TEMP = 0.0
WA1TEMP1 = 0.0
WA1TEMP2 = 0.0
WB1TEMP1 = 0.0
WB1TEMP2 = 0.0
WB2TEMP1 = 0.0

WB2TEMP2 = 0.0

DO 550 L = 1, N

WA1 = R(L+3) - B1*WA1TEMP1 - B2*WA1TEMP2
 WB1 = -T1(L+1) - B1*WB1TEMP1 - B2*WB1TEMP2
 WB2 = -T1(L) - B1*WB2TEMP1 - B2*WB2TEMP2

A(1,1) = A(1,1) + WA1*WA1
 A(2,2) = A(2,2) + WB1*WB1
 A(3,3) = A(3,3) + WB2*WB2
 W1TEMP = WA1TEMP1
 WA1TEMP1 = WA1
 WA1TEMP2 = W1TEMP
 W2TEMP = WB1TEMP1
 WB1TEMP1 = WB1
 WB1TEMP2 = W2TEMP
 W3TEMP = WB2TEMP1
 WB2TEMP1 = WB2
 WB2TEMP2 = W3TEMP

550 CONTINUE

A(1,1) = A(1,1)/ZIGMAET
 A(2,2) = A(2,2)/ZIGMAET
 A(3,3) = A(3,3)/ZIGMAET
 A(5,1) = A(5,1) - (1+2*D2+C1**2)
 A(2,5) = A(2,5) + (1+B2)*ZIGMAET
 A(3,5) = A(3,5) + B1*ZIGMAET
 A(6,1) = A(6,1) - C1
 A(3,6) = A(3,6) + ZIGMAET
 A(4,4) = A(4,4) - N/(2*(ZIGMAET**2))
 A(4,5) = A(4,5) + B1*(1+B2)
 A(4,6) = A(4,6) + B2
 A(1,5) = A(5,1)
 A(1,6) = A(6,1)
 A(5,2) = A(2,5)
 A(5,3) = A(3,5)
 A(6,3) = A(3,6)
 A(5,4) = A(4,5)
 A(6,4) = A(4,6)

CALL INVMAT(AINV,A,6,6)

IF (AINV(1,1).GT.0.0) THEN
 SS1 = N*(AA1-1.0)
 SS2 = (AA1-1.0)/SQRT(AINV(1,1))
 ICOUNT = ICOUNT + 1
ENDIF

```

RETURN
END

```

```

*****
*   SUBROUTINE 6   *
*****
      SUBROUTINE SORT(N,DATA)

* LOG-SPACE QUICK-SORT
  PARAMETER (MAXSTK=10)
* MAX. DATA (ARRAY) SIZE: (MAXSTK+1)-TH POWER OF 2
  INTEGER N
*****
* DATA TYPE
  REAL*8 DATA(N), PIVOT, TEMP
*****
  INTEGER I, J, LEFT, RIGHT, L, R
* STACK FRAMES AND POINTER
  INTEGER LSTACK(MAXSTK), RSTACK(MAXSTK), SP
*
  SP = 0
*
  LEFT = 1
  RIGHT = N
*
* RECURSIVE CALL ENTRY (CALL RETURNS HERE IF STACK NOT EMPTY)
1000  CONTINUE
*
  I = LEFT
  J = RIGHT
  PIVOT = DATA(INT((I+J)/2))
*
2000  CONTINUE
*
100  IF (DATA(I).LT.PIVOT) THEN
      I = I+1
      GOTO 100
    ENDIF
*
200  IF (DATA(J).GT.PIVOT) THEN
      J = J-1
      GOTO 200
    ENDIF
*
  IF (I.LE.J) THEN

```

```
* SWAP
  TEMP = DATA(I)
  DATA(I) = DATA(J)
  DATA(J) = TEMP
  I = I+1
  J = J-1
  ENDIF
  IF (I.LE.J) GOTO 2000
* RECURSIVE CALL
  R = RIGHT - I
  L = J - LEFT
  IF (R.GE.L) THEN
  IF (L.GT.0) THEN
* PUSH STACK
  SP = SP + 1
  LSTACK(SP) = I
  RSTACK(SP) = RIGHT
  RIGHT = J
  GOTO 1000
  ELSE IF (R.GT.0) THEN
  LEFT = I
  GOTO 1000
  ENDIF
  ELSE
  IF (R.GT.0) THEN
* PUSH STACK
  SP = SP + 1
  LSTACK(SP) = LEFT
  RSTACK(SP) = J
  LEFT = I
  GOTO 1000
  ELSE IF (L.GT.0) THEN
  RIGHT = J
  GOTO 1000
  ENDIF
  ENDIF
*
* RECURSIVE RETURN
  IF (SP.GT.0) THEN
* POP STACK
  RIGHT = RSTACK(SP)
  LEFT = LSTACK(SP)
  SP = SP - 1
  GOTO 1000
  ENDIF
```


RETURN
END

* SUBROUTINE 7 *

*** COMPUTE INVERSE OF A MATRIX.

```

SUBROUTINE INVMAT(AINV,A,N,NMAX)
REAL*8 A(NMAX,NMAX), L(NMAX,NMAX), U(NMAX,NMAX)
REAL*8 AINV(NMAX,NMAX), UINV(NMAX,NMAX), LINV(NMAX,NMAX)
CALL LUDCOM(L,U,A,N,NMAX)
DO 100 I = 1, N
100  LINV(I,I) = 1.0
    DO 110 I = 1, N-1
        DO 110 J = I+1, N
110  LINV(I,J) = 0.0
        DO 130 J = 1, N-1
            DO 130 I = J+1, N
                SUM = 0.0
                DO 120 K = J+1, I-1
120  SUM = SUM + L(I,K)*LINV(K,J)
                LINV(I,J) = -L(I,J) - SUM
130 CONTINUE
        DO 140 I = 1, N
140  UINV(I,I) = 1/U(I,I)
        DO 150 J = 1, N-1
            DO 150 I = J+1, N
150  UINV(I,J) = 0.0
            DO 170 I = N-1, 1, -1
                DO 170 J = N, I+1, -1
                    SUM = 0.0
                    DO 160 K = I+1, J
160  SUM = SUM + U(I,K)*UINV(K,J)
                    UINV(I,J) = UINV(I,I)*(-SUM)
170 CONTINUE
        CALL PROMAT(AINV,UINV,LINV,N,NMAX)
        RETURN
    END

```

*

* LU-DECOMPOSITION ROUTINE

```

SUBROUTINE LUDCOM(L,U,A,N,NMAX)
REAL*8 L(NMAX,NMAX), U(NMAX,NMAX)
REAL*8 A(NMAX,NMAX), WM(NMAX,NMAX)
DO 100 K = 1, N

```

```

      L(K,K) = 1.0
      U(1,K) = A(1,K)
100  CONTINUE
      DO 150 I = 2, N
        DO 120 J = 1, I-1
          WM(I,J) = 0.0
          DO 110 M = 1, J-1
110    WM(I,J) = WM(I,J) + L(I,M)*U(M,J)
          L(I,J) = (A(I,J) - WM(I,J))/U(J,J)
120  CONTINUE
        DO 140 J = I, N
          WM(I,J) = 0.0
          DO 130 M = 1, I-1
130    WM(I,J) = WM(I,J) + L(I,M)*U(M,J)
          U(I,J) = A(I,J) - WM(I,J)
140  CONTINUE
150  CONTINUE
      RETURN
      END
*
*  MATRIX MULTIPLICATION (C = A * B)
SUBROUTINE PROMAT(C,A,B,N,NMAX)
  REAL*8 C(NMAX,NMAX), A(NMAX,NMAX), B(NMAX,NMAX)
  DO 100 I = 1, N
    DO 100 J = 1, N
      C(I,J) = 0.0
      DO 100 K = 1, N
        C(I,J) = C(I,J) + A(I,K)*B(K,J)
100  CONTINUE
  RETURN
  END
* END OF PROGRAM

```

APPENDIX F

FORTRAN PROGRAM FOR SPURIOUS REGRESSION

```

*****
*   SPURIOUS REGRESSION: K = 1, P = 2   *
*****
*****
*   MAIN PROGRAM   *
*****

PROGRAM MAIN
PARAMETER(N=250, NUM=10000, N2=3)
IMPLICIT REAL*8 (A-H,O-Z)
REAL*8 U(N), X(N), XT1(N), XT2(N), Y(N)
REAL*8 XX(N,3), RESID(N), SUMGAMMA(N2)
REAL*8 EPGAMMA1, EPGAMMA2, EPGAMMA3
REAL*8 EPRSQ, EPDW, EPSGSQ
REAL*8 EPT1, EPT2, EPT3
EXTERNAL RNSET,RNNOA
ISEED1= 13579

CALL RNISD(ISEED1,ISEED2)

DO 300 I= 1, NUM

    CALL RNSET(ISEED1)
    CALL DRNNOA(N,U)
    CALL RNGET(ISEED1)
    CALL RNSET(ISEED2)
    CALL DRNNOA(N,X)
    CALL RNGET(ISEED2)

    DO 100 J= 1, N
        SUM1 = 0.0
        SUM2 = 0.0
        DO 50 K= 1, J
            SUM1 = SUM1 + U(K)
            SUM2 = SUM2 + X(K)
50        CONTINUE
            Y(J) = SUM1
            XT1(J) = SUM2
100       CONTINUE
            XT2(1) = X(1)
            XT2(2) = 2*X(1) + X(2)

    DO 250 J= 3, N
        SUM3 = 0.0
        DO 200 K= 1, J-2
            SUM3 = SUM3 + XT1(K)

```

```

200  CONTINUE
      SUM4 = SUM3 + XT1(K+1)
      XT2(J) = 2*SUM4 - SUM3 + X(J)
250  CONTINUE

      CALL GAMMA(N,N2,XT1,XT2,Y,XX,RESID,SUMGAMMA,SUMRSQ,
/      SUMDW,SUMSGSQ,SUMT1,SUMT2,SUMT3)

300  CONTINUE
      EPGAMMA1 = SUMGAMMA(1)/NUM
      EPGAMMA2 = SUMGAMMA(2)/NUM
      EPGAMMA3 = SUMGAMMA(3)/NUM
      EPRSQ = SUMRSQ/NUM
      EPDW = SUMDW/NUM
      EPSGSQ = SUMSGSQ/NUM
      EPT1 = SUMT1/NUM
      EPT2 = SUMT2/NUM
      EpT3 = SUMT3/NUM

      WRITE(*,*) EPGAMMA1,EPGAMMA2,EPGAMMA3
      WRITE(*,*) EPT1,EPT2,EPT3
      WRITE(*,*) EPRSQ,EPDW
      WRITE(*,*) EPSGSQ

      END

*****
*   SUBROUTINE 1   *
*****

      SUBROUTINE GAMMA(N,N2,XT1,XT2,Y,XX,RESID,SUMGAMMA,SUMRSQ,
/      SUMDW,SUMSGSQ,SUMT1,SUMT2,SUMT3)
      PARAMETER(MDA=3, MDAINV=3, P=2)
      REAL*8 A(MDA,MDA), AINV(MDAINV,MDAINV), B(MDA)
      REAL*8 XT1(N), XT2(N), XX(N,3), Y(N), GAMMAHAT(MDA), RESID(N)
      REAL*8 SUMGAMMA(N2), SQY, YBAR, RSQDENO, RSQNUM, RSQUARE
      REAL*8 DENO, DNUM, DW, SUMDW
      REAL*8 SUM, SUM2, SUM3, SUMRSQ
      REAL*8 T_STAT1, T_STAT2, T_STAT3
      REAL*8 SUMT1, SUMT2, SUMT3
      REAL*8 SIGMASQ

      SQY = 0.0
      DO 150 K1 = 1, 3
        DO 120 K2 = 1, 3

```

```

      A(K1,K2) = 0.0
120  CONTINUE
      B(K1) = 0.0
150  CONTINUE

***** OLS, R-SQUARE(DENOMINATOR), T-STATISTICS
      DO 200 L = 1, N
        A(1,2) = A(1,2) + XT1(L)
        A(1,3) = A(1,3) + XT2(L)
        A(2,2) = A(2,2) + XT1(L)**2
        A(2,3) = A(2,3) + XT1(L)*XT2(L)
        A(3,3) = A(3,3) + XT2(L)**2
        B(1) = B(1) + Y(L)
        B(2) = B(2) + XT1(L)*Y(L)
        B(3) = B(3) + XT2(L)*Y(L)
        SQY = SQY + Y(L)*Y(L)
200  CONTINUE
        A(1,1) = N
        A(2,1) = A(1,2)
        A(3,1) = A(1,3)
        A(3,2) = A(2,3)
        YBAR = B(1)*B(1)/N
        RSQDENO = SQY - YBAR

      CALL INVMAT(AINV,A,3)

      DO 240 L2 = 1, N
        XX(L2,1) = 1.0
        XX(L2,2) = XT1(L2)
        XX(L2,3) = XT2(L2)
240  CONTINUE

      DO 280 MDAINV1 = 1, 3
        SUM = 0.0
        DO 250 MDAINV2 = 1, 3
          SUM = SUM + AINV(MDAINV1,MDAINV2)*B(MDAINV2)
250  CONTINUE
        GAMMAHAT(MDAINV1) = SUM
280  CONTINUE
        SUMGAMMA(1) = SUMGAMMA(1) + GAMMAHAT(1)
        SUMGAMMA(2) = SUMGAMMA(2) + GAMMAHAT(2)
        SUMGAMMA(3) = SUMGAMMA(3) + GAMMAHAT(3)

***** RESIDUAL, R-SQUARE(NUMERATOR)
      SUM3 = 0.0
      DO 320 M1 = 1, N

```

```

        SUM2 = 0.0
        DO 300 M2 = 1, 3
        SUM2 = SUM2 + XX(M1,M2)*GAMMAHAT(M2)
300    CONTINUE
        SUM3 = SUM3 + Y(M1)*SUM2
        RESID(M1) = Y(M1) - SUM2
320    CONTINUE
        RSQNUM = SUM3 - YBAR
        RSQUARE = RSQNUM/RSQDENO
        SUMRSQ = SUMRSQ + RSQUARE

*** DW, SIGMASUARE HAT
        DENO = 0.0
        DNUM = 0.0
        DO 350 M3 = 1, N-1
        DNUM = DNUM + (RESID(M3+1) - RESID(M3))**2
        DENO = DENO + RESID(M3+1)**2
350    CONTINUE
        DENO = DENO + RESID(1)**2
        DW = DNUM/DENO
        SIGMASQ = DENO/(N-P-1)
        T_STAT1 = GAMMAHAT(1)/SQRT(AINV(1,1)*SIGMASQ)
        T_STAT2 = GAMMAHAT(2)/SQRT(AINV(2,2)*SIGMASQ)
        T_STAT3 = GAMMAHAT(3)/SQRT(AINV(3,3)*SIGMASQ)
        SUMT1 = SUMT1 + T_STAT1
        SUMT2 = SUMT2 + T_STAT2
        SUMT3 = SUMT3 + T_STAT3
        SUMDW = SUMDW + DW
        SUMSGSQ = SUMSGSQ + SIGMASQ
        RETURN
        END

```

```
*****
```

```
* SUBROUTINE 2 *
```

```
*****
```

```
*
```

```

SUBROUTINE INVMAT(AINV,A,N)
REAL*8 A(3,3), L(3,3), U(3,3)
REAL*8 AINV(3,3), UINV(3,3), LINV(3,3)
CALL LUDCOM(L,U,A,N)
        DO 100 I = 1, N
100    LINV(I,I) = 1.0
        DO 110 I = 1, N-1
        DO 110 J = I+1, N
110    LINV(I,J) = 0.0
        DO 130 J = 1, N-1

```

```

      DO 130 I = J+1, N
        SUM = 0.0
        DO 120 K = J+1, I-1
120     SUM = SUM + L(I,K)*LINV(K,J)
        LINV(I,J) = -L(I,J) - SUM
130    CONTINUE
      DO 140 I = 1, N
140     UINV(I,I) = 1/U(I,I)
      DO 150 J = 1, N-1
        DO 150 I = J+1, N
150     UINV(I,J) = 0.0
      DO 170 I = N-1, 1, -1
        DO 170 J = N, I+1, -1
          SUM = 0.0
          DO 160 K = I+1, J
160     SUM = SUM + U(I,K)*UINV(K,J)
          UINV(I,J) = UINV(I,I)*(-SUM)
170    CONTINUE
      CALL PROMAT(AINV,UINV,LINV,N,N,N)
      RETURN
      END

```

*

```

SUBROUTINE LUDCOM(L,U,A,N)
REAL*8 L(3,3), U(3,3), A(3,3), WM(3,3)
DO 100 K = 1, N
  L(K,K) = 1.0
  U(1,K) = A(1,K)
100  CONTINUE
  DO 150 I = 2, N
    DO 120 J = 1, I-1
      WM(I,J) = 0.0
      DO 110 M = 1, J-1
110     WM(I,J) = WM(I,J) + L(I,M)*U(M,J)
      L(I,J) = (A(I,J) - WM(I,J))/U(J,J)
120    CONTINUE
    DO 140 J = I, N
      WM(I,J) = 0.0
      DO 130 M = 1, I-1
130     WM(I,J) = WM(I,J) + L(I,M)*U(M,J)
      U(I,J) = A(I,J) - WM(I,J)
140    CONTINUE
150    CONTINUE
      RETURN
      END

```

*


```
SUBROUTINE PROMAT(C,A,B,N,L,M)
  REAL*8 C(3,3), A(3,3), B(3,3)
  DO 100 I = 1, N
    DO 100 J = 1, M
      C(I,J) = 0.0
      DO 100 K = 1, L
        C(I,J) = C(I,J) + A(I,K)*B(K,J)
100    CONTINUE
  RETURN
END
```

* END OF PROGRAM

```

*****
*   SPURIOUS REGRESSION: K = 2, P = 1   *
*****
*****
*   MAIN PROGRAM   *
*****

PROGRAM MAIN
PARAMETER(N=250, NUM=10000, N2=2)
IMPLICIT REAL*8 (A-H, O-Z)
REAL*8 U(N), X(N), XT(N), YT(N), Y(N)
REAL*8 XX(N,2), RESID(N), SUMGAMMA(N2)
REAL*8 SUM1, SUM2, SUM3, SUM4
REAL*8 EPGAMMA1, EPGAMMA2
REAL*8 EPRSQ, EPDW, EPSGSQ1, EPSGSQ2
REAL*8 EPT1, EPT2, EPT3
REAL*8 SUMSGSQ, SUMRSQ, SUMDW
REAL*8 SUMT1, SUMT2, SUMT3
EXTERNAL RNSET, RNNOA
ISEED1 = 13579

CALL RNISD(ISEED1, ISEED2)

DO 2000 I = 1, NUM

    CALL RNSET(ISEED1)
    CALL DRNNOA(N, U)
    CALL RNGET(ISEED1)
    CALL RNSET(ISEED2)
    CALL DRNNOA(N, X)
    CALL RNGET(ISEED2)

    DO 100 J = 1, N
        SUM1 = 0.0
        SUM2 = 0.0
        DO 10 K = 1, J
            SUM1 = SUM1 + X(K)
            SUM2 = SUM2 + U(K)
10        CONTINUE
        XT(J) = SUM1
        YT(J) = SUM2
100    CONTINUE

    Y(1) = U(1)
    Y(2) = U(1)*2 + U(2)
    DO 250 J = 3, N

```

```

        SUM3 = 0.0
        DO 200 K = 1, J-2
            SUM3 = SUM3 + YT(K)
200    CONTINUE
        SUM4 = SUM3 + YT(K+1)
        Y(J) = 2*SUM4 - SUM3 + U(J)
250    CONTINUE

        CALL GAMMA(N,N2,XT,Y,XX,RESID,SUMGAMMA,SUMRSQ,
/          SUMDW,SUMSGSQ,SUMT1,SUMT2)

2000 CONTINUE
        EPGAMMA1 = SUMGAMMA(1)/NUM
        EPGAMMA2 = SUMGAMMA(2)/NUM
        EPRSQ = SUMRSQ/NUM
        EPDW = SUMDW/NUM
        EPSGSQ = SUMSGSQ/NUM
        EPT1 = SUMT1/NUM
        EPT2 = SUMT2/NUM
        WRITE(*,*) EPGAMMA1,EPGAMMA2
        WRITE(*,*) EPT1,EPT2
        WRITE(*,*) EPRSQ,EPDW
        WRITE(*,*) EPSGSQ
        END

*****
*   SUBROUTINE 1   *
*****

SUBROUTINE GAMMA(N,N2,XT,Y,XX,RESID,SUMGAMMA,SUMRSQ,
/          SUMDW,SUMSGSQ,SUMT1,SUMT2)
    PARAMETER(MDA=2, MDAINV=2, P=1)
    REAL*8 A(MDA,MDA), AINV(MDAINV,MDAINV), B(MDA)
    REAL*8 XT(N), XX(N,2), Y(N), GAMMAHAT(MDA), RESID(N)
    REAL*8 SQY, YBAR, RSQDENO, RSQNUM, RSQUARE
    REAL*8 SUMGAMMA(N2), SQY, YBAR, RSQDENO, RSQNUM, RSQUARE
    REAL*8 SUM, SUM2, SUM3
    REAL*8 SUMRSQ, DENO, DNUM, DW, SUMDW
    REAL*8 T_STAT1, T_STAT2
    REAL*8 SUMT1, SUMT2
    REAL*8 SIGMASQ, SUMSGSQ

    SQY = 0.0
    DO 150 K1 = 1, 2
        DO 120 K2 = 1, 2
            A(K1,K2) = 0.0

```

```

120    CONTINUE
      B(K1) = 0.0
150    CONTINUE

*** OLS, R-SQUARE(DENOMINATOR), T-STATISTICS
      DO 200 L = 1, N
        A(1,2) = A(1,2) + XT(L)
        A(2,2) = A(2,2) + XT(L)**2
        B(1) = B(1) + Y(L)
        B(2) = B(2) + XT(L)*Y(L)
        SQY = SQY + Y(L)*Y(L)
200    CONTINUE
      A(1,1) = N
      A(2,1) = A(1,2)
      YBAR = B(1)*B(1)/N
      RSQDENO = SQY - YBAR

      CALL INVMAT(AINV,A,2)

      DO 240 L2 = 1, N
        XX(L2,1) = 1.0
        XX(L2,2) = XT(L2)
240    CONTINUE
      DO 280 MDAINV1 = 1, 2
        SUM = 0.0
        DO 250 MDAINV2 = 1, 2
          SUM = SUM + AINV(MDAINV1,MDAINV2)*B(MDAINV2)
250    CONTINUE
        GAMMAHAT(MDAINV1) = SUM
280    CONTINUE
      SUMGAMMA(1) = SUMGAMMA(1) + GAMMAHAT(1)
      SUMGAMMA(2) = SUMGAMMA(2) + GAMMAHAT(2)

*** RESIDUAL, R-SQUARE(NUMERATOR)
      SUM3 = 0.0
      DO 320 M1 = 1, N
        SUM2 = 0.0
        DO 300 M2 = 1, 2
          SUM2 = SUM2 + XX(M1,M2)*GAMMAHAT(M2)
300    CONTINUE
        SUM3 = SUM3 + Y(M1)*SUM2
        RESID(M1) = Y(M1) - SUM2
320    CONTINUE
      RSQNUM = SUM3 - YBAR
      RSQUARE = RSQNUM/RSQDENO
      SUMRSQ = SUMRSQ + RSQUARE

```

```

*** DW, SIGMASUARE HAT
    DENO = 0.0
    DNUM = 0.0
    DO 350 M3 = 1, N-1
        DNUM = DNUM + (RESID(M3+1) - RESID(M3))**2
        DENO = DENO + RESID(M3+1)**2
350  CONTINUE
    DENO = DENO + RESID(N)**2
    DW = DNUM/DENO
    SIGMASQ = DENO/(N-P-1)
    T_STAT1 = GAMMAHAT(1)/SQRT(AINV(1,1)*SIGMASQ)
    T_STAT2 = GAMMAHAT(2)/SQRT(AINV(2,2)*SIGMASQ)
    SUMT1 = SUMT1 + T_STAT1
    SUMT2 = SUMT2 + T_STAT2
    SUMDW = SUMDW + DW
    SUMSGSQ = SUMSGSQ + SIGMASQ
    RETURN
    END

```

```

*****
*   SUBROUTINE 2   *
*****

```

```

SUBROUTINE INVMAT(AINV,A,N)
    REAL*8 A(2,2), L(2,2), U(2,2)
    REAL*8 AINV(2,2), UINV(2,2), LINV(2,2)
    CALL LUDCOM(L,U,A,N)
    DO 100 I = 1, N
100  LINV(I,I) = 1.0
        DO 110 I = 1, N-1
            DO 110 J = I+1, N
110  LINV(I,J) = 0.0
        DO 130 J = 1, N-1
            DO 130 I = J+1, N
                SUM = 0.0
                DO 120 K = J+1, I-1
120  SUM = SUM + L(I,K)*LINV(K,J)
                LINV(I,J) = -L(I,J) - SUM
130  CONTINUE
            DO 140 I = 1, N
140  UINV(I,I) = 1/U(I,I)
                DO 150 J = 1, N-1
                    DO 150 I = J+1, N
150  UINV(I,J) = 0.0
                DO 170 I = N-1, 1, -1

```

```

DO 170 J = N, I+1, -1
  SUM = 0.0
  DO 160 K = I+1, J
160   SUM = SUM + U(I,K)*UINV(K,J)
      UINV(I,J) = UINV(I,I)*(-SUM)
170  CONTINUE
      CALL PROMAT(AINV,UINV,LINV,N,N,N)
      RETURN
      END

```

*

```

SUBROUTINE LUDCOM(L,U,A,N)
REAL*8 L(2,2), U(2,2), A(2,2), WM(2,2)
DO 100 K = 1, N
  L(K,K) = 1.0
  U(1,K) = A(1,K)
100  CONTINUE
      DO 150 I = 2, N
        DO 120 J = 1, I-1
          WM(I,J) = 0.0
          DO 110 M = 1, J-1
110   WM(I,J) = WM(I,J) + L(I,M)*U(M,J)
          L(I,J) = (A(I,J) - WM(I,J))/U(J,J)
120  CONTINUE
          DO 140 J = I, N
            WM(I,J) = 0.0
            DO 130 M = 1, I-1
130   WM(I,J) = WM(I,J) + L(I,M)*U(M,J)
            U(I,J) = A(I,J) - WM(I,J)
140  CONTINUE
150  CONTINUE
      RETURN
      END

```

*

```

SUBROUTINE PROMAT(C,A,B,N,L,M)
REAL*8 C(2,2), A(2,2), B(2,2)
DO 100 I = 1, N
  DO 100 J = 1, M
    C(I,J) = 0.0
    DO 100 K = 1, L
      C(I,J) = C(I,J) + A(I,K)*B(K,J)
100  CONTINUE
      RETURN
      END

```

* END OF PROGRAM

```
*****
* SPURIOUS REGRESSION: K = 2, P = 2 *
*****
```

```
*****
* MAIN PROGRAM *
*****
```

```
PROGRAM MAIN
PARAMETER(N=500, NUM=10000, N2=3)
IMPLICIT REAL*8 (A-H, O-Z)
REAL*8 U(N), X(N), XT1(N), XT2(N), YT(N), Y(N)
REAL*8 XX(N,3), RESID(N), SUMGAMMA(N2)
EXTERNAL RNSET, RNNOA
ISEED1= 13579
```

```
*** GENERATE 2 DIFFERENT RANDOM NUMBERS.
```

```
CALL RNISD(ISEED1, ISEED2)
DO 300 I = 1, NUM
  CALL RNSET(ISEED1)
  CALL DRNNOA(N, U)
  CALL RNGET(ISEED1)
  CALL RNSET(ISEED2)
  CALL DRNNOA(N, X)
  CALL RNGET(ISEED2)
```

```
DO 100 J = 1, N
  SUM1 = 0.0
  SUM2 = 0.0
  DO 50 K = 1, J
    SUM1 = SUM1 + X(K)
    SUM2 = SUM2 + U(K)
50  CONTINUE
  XT1(J) = SUM1
  YT(J) = SUM2
100 CONTINUE
  XT2(1) = X(1)
  XT2(2) = X(1)*2 + X(2)
  Y(1) = U(1)
  Y(2) = U(1)*2 + U(2)
DO 250 J = 3, N
  SUM3 = 0.0
  SUM5 = 0.0
  DO 200 K = 1, J-2
    SUM3 = SUM3 + XT1(K)
    SUM5 = SUM5 + YT(K)
```

```

200  CONTINUE
      SUM4 = SUM3 + XT1(K+1)
      SUM6 = SUM5 + YT(K+1)
      XT2(J) = 2*SUM4 - SUM3 + X(J)
      Y(J) = 2*SUM6 - SUM5 + U(J)
250  CONTINUE

      CALL GAMMA(N,N2,XT1,XT2,YT,Y,XX,RESID,SUMGAMMA,SUMRSQ,
/          SUMDW,SUMSGSQ,SUMT1,SUMT2,SUMT3)
300  CONTINUE
      EPGAMMA1 = SUMGAMMA(1)/NUM
      EPGAMMA2 = SUMGAMMA(2)/NUM
      EPGAMMA3 = SUMGAMMA(3)/NUM
      EPRSQ = SUMRSQ/NUM
      EPDW = SUMDW/NUM
      EPSGSQ = SUMSGSQ/NUM
      EPT1 = SUMT1/NUM
      EPT2 = SUMT2/NUM
      EPT3 = SUMT3/NUM
      WRITE(*,*) EPGAMMA1, EPGAMMA2, EPGAMMA3
      WRITE(*,*) EPT1, EPT2, EPT3
      WRITE(*,*) EPRSQ, EPDW
      WRITE(*,*) EPSGSQ
      END

*****
*   SUBROUTINE 1   *
*****

      SUBROUTINE GAMMA(N,N2,XT1,XT2,YT,Y,XX,RESID,SUMGAMMA,
/          SUMRSQ,SUMDW,SUMSGSQ,SUMT1,SUMT2,SUMT3)
      PARAMETER(MDA=3, MDAINV=3, P=2)
      REAL*8 A(MDA,MDA), AINV(MDAINV,MDAINV), B(MDA), YT(N)
      REAL*8 XT1(N), XT2(N), XX(N,3), Y(N), GAMMAHAT(MDA), RESID(N)
      REAL*8 SUMGAMMA(N2), SQY, YBAR, RSQDENO, RSQNUM, RSQUARE
      REAL*8 DENO, DNUM, DW, SUMDW
      REAL*8 SUM, SUM2, SUM3, SUMRSQ
      REAL*8 T_STAT1, T_STAT2, T_STAT3
      REAL*8 SUMT1, SUMT2, SUMT3
      REAL*8 SIGMASQ

      SQY = 0.0
      DO 150 K1 = 1, 3
        DO 120 K2 = 1, 3
          A(K1,K2) = 0.0
120    CONTINUE

```



```

      B(K1) = 0.0
150  CONTINUE

*** OLS, R-SQUARE(DENOMINATOR), T-STATISTICS
      DO 200 L = 1, N
        A(1,2) = A(1,2) + XT1(L)
        A(1,3) = A(1,3) + XT2(L)
        A(2,2) = A(2,2) + XT1(L)**2
        A(2,3) = A(2,3) + XT1(L)*XT2(L)
        A(3,3) = A(3,3) + XT2(L)**2
        B(1) = B(1) + Y(L)
        B(2) = B(2) + XT1(L)*Y(L)
        B(3) = B(3) + XT2(L)*Y(L)
        SQY = SQY + Y(L)*Y(L)
200  CONTINUE
      A(1,1) = N
      A(2,1) = A(1,2)
      A(3,1) = A(1,3)
      A(3,2) = A(2,3)
      YBAR = B(1)*B(1)/N
      RSQDENO = SQY - YBAR

      CALL INVMAT(AINV,A,3)

      DO 240 L2 = 1, N
        XX(L2,1) = 1.0
        XX(L2,2) = XT1(L2)
        XX(L2,3) = XT2(L2)
240  CONTINUE

      DO 280 MDAINV1 = 1, 3
        SUM = 0.0
        DO 250 MDAINV2 = 1, 3
          SUM = SUM + AINV(MDAINV1,MDAINV2)*B(MDAINV2)
250  CONTINUE
        GAMMAHAT(MDAINV1) = SUM
280  CONTINUE
      SUMGAMMA(1) = SUMGAMMA(1) + GAMMAHAT(1)
      SUMGAMMA(2) = SUMGAMMA(2) + GAMMAHAT(2)
      SUMGAMMA(3) = SUMGAMMA(3) + GAMMAHAT(3)

*** RESIDUAL, R-SQUARE(NUMERATOR)
      SUM3 = 0.0
      DO 320 M1 = 1, N
        SUM2 = 0.0
        DO 300 M2 = 1, 3

```

```

SUM2 = SUM2 + XX(M1,M2)*GAMMAHAT(M2)
300  CONTINUE
SUM3 = SUM3 + Y(M1)*SUM2
RESID(M1) = Y(M1) - SUM2
320  CONTINUE
RSQNUM = SUM3 - YBAR
RSQUARE = RSQNUM/RSQDENO
SUMRSQ = SUMRAQ + RSQUARE

*** DW, SIGMASUARE HAT
DENO = 0.0
DNUM = 0.0
DO 350 M3 = 1, N-1
  DNUM = DNUM + (RESID(M3+1) - RESID(M3))**2
  DENO = DENO + RESID(M3+1)**2
350  CONTINUE
DENO = DENO + RESID(N)**2
DW = DNUM/DENO
SIGMASQ = DENO/(N-P-1)
T_STAT1 = GAMMAHAT(1)/SQRT(AINV(1,1)*SIGMASQ)
T_STAT2 = GAMMAHAT(2)/SQRT(AINV(2,2)*SIGMASQ)
T_STAT3 = GAMMAHAT(3)/SQRT(AINV(3,3)*SIGMASQ)
SUMT1 = SUMT1 + T_STAT1
SUMT2 = SUMT2 + T_STAT2
SUMT3 = SUMT3 + T_STAT3
SUMDW = SUMDW + DW
SUMSGSQ = SUMSGSQ + SIGMASQ
RETURN
END

```

```

*****
* SUBROUTINE 2 *
*****

```

```

SUBROUTINE INVMAT(AINV,A,N)
REAL*8 A(3,3), L(3,3), U(3,3)
REAL*8 AINV(3,3), UINV(3,3), LINV(3,3)
CALL LUDCOM(L,U,A,N)
DO 100 I = 1, N
100 LINV(I,I) = 1.0
DO 110 I = 1, N-1
  DO 110 J = I+1, N
110 LINV(I,J) = 0.0
DO 130 J = 1, N-1
  DO 130 I = J+1, N
SUM = 0.0

```

```

      DO 120 K = J+1, I-1
120   SUM = SUM + L(I,K)*LINV(K,J)
      LINV(I,J) = -L(I,J)-SUM
130   CONTINUE
      DO 140 I = 1, N
140   UINV(I,I)=1/U(I,I)
      DO 150 J = 1, N-1
      DO 150 I = J+1, N
150   UINV(I,J) = 0.0
      DO 170 I = N-1, 1, -1
      DO 170 J = N, I+1, -1
      SUM = 0.0
      DO 160 K = I+1, J
160   SUM = SUM + U(I,K)*UINV(K,J)
      UINV(I,J) = UINV(I,I)*(-SUM)
170   CONTINUE
      CALL PROMAT(AINV,UINV,LINV,N,N,N)
      RETURN
      END

```

*

```

SUBROUTINE LUDCOM(L,U,A,N)
REAL*8 L(3,3), U(3,3), A(3,3), WM(3,3)
DO 100 K = 1, N
  L(K,K) = 1.0
  U(1,K) = A(1,K)
100  CONTINUE
  DO 150 I = 2, N
  DO 120 J = 1, I-1
    WM(I,J) = 0.0
    DO 110 M = 1, J-1
110   WM(I,J) = WM(I,J) + L(I,M)*U(M,J)
    L(I,J) = (A(I,J)-WM(I,J))/U(J,J)
120  CONTINUE
  DO 140 J = I, N
    WM(I,J) = 0.0
    DO 130 M = 1, I-1
130   WM(I,J) = WM(I,J) + L(I,M)*U(M,J)
    U(I,J) = A(I,J) - WM(I,J)
140  CONTINUE
150  CONTINUE
      RETURN
      END

```

*

```

SUBROUTINE PROMAT(C,A,B,N,L,M)
REAL*8 C(3,3), A(3,3), B(3,3)
DO 100 I = 1, N

```

```
DO 100 J = 1, M
  C(I,J) = 0.0
  DO 100 K = 1, L
    C(I,J) = C(I,J) + A(I,K)*B(K,J)
100  CONTINUE
RETURN
END
```

* END OF PROGRAM

APPENDIX G

FORTRAN PROGRAM
FOR
MISSPECIFICATION
OF A NONSTATIONARY AR(2) AND
A POLYNOMIAL REGRESSION MODEL OF ORDER 2

```
*****
*  MISSPECIFICATION AR(2) AS POLYNOMIAL REGRESSION MODEL  *
*  OF ORDER 2.                                           *
*****
```

```
*****
*  MAIN PROGRAM  *
*****
```

```
PROGRAM MAIN
PARAMETER(N=500, NUM=10000, N2=3)
IMPLICIT REAL*8 (A-H,O-Z)
REAL*8 U(N), Y(N), X(N,3), RESID(N), UMBETA(N2)
EXTERNAL RNSET, RNNOA
ISEED = 13579
```

```
CALL RNSET(ISEED)
```

```
DO 50 I=1, NUM
  CALL RNNOA(N,U)
  Y(1) = U(1)
  Y(2) = 2*Y(1) + U(2)
  DO 10 J=3, N
    Y(J) = 2*Y(J-1) - Y(J-2) + U(J)
10  CONTINUE
  CALL BETA(N,N2,X,Y,RESID,SUMBETA,SUMRSQ,SUMDW,
 /      SIGMASQ,SUMT1,SUMT2,SUMT3)
50  CONTINUE
  EPBETA1 = SUMBETA(1)/NUM
  EPBETA2 = SUMBETA(2)/NUM
  EPBETA3 = SUMBETA(3)/NUM
  EMPRSQ = SUMRSQ/NUM
  EMPDW = SUMDW/NUM
  EMPSIGSQ = SUMSIGSQ/NUM
  EMPT1 = SUMT1/NUM
  EMPT2 = SUMT2/NUM
  EMPT3 = SUMT3/NUM
  WRITE(*,*) EPBETA1, EPBETA2, EPBETA3
  WRITE(*,*) EMPT1, EMPT2, EMPT3
  WRITE(*,*) EMPRSQ, EMPDW, EMPSIGSQ
END
```

```
*****
*  SUBROUTINE 1  *
*****
```

```
SUBROUTINE BETA(N,N2,X,Y,RESID,SUMBETA,SUMRSQ,SUMDW,
```

```

/          SUMSIGSQ,SUMT1,SUMT2,SUMT3)
PARAMETER(MDA=3, MDAINV=3, P=2)
REAL*8 A(MDA,MDA), AINV(MDAINV,MDAINV), B(MDA)
REAL*8 X(N,3), Y(N), BETAHAT(MDA), RESID(N), SQY
REAL*8 RSQDENO, YBAR, SUM, SUM2, SUM3, RSQNUM, RSQUARE
REAL*8 SUMBETA(N2), HAP, DENO, DNUM, DW, SIGMASQ
REAL*8 T_STAT1, T_STAT2, T_STAT3, SUMT1, SUMT2, SUMT3
REAL*8 SUMRSQ, SUMDW, SUMSIGSQ
  SQY = 0.0
  DO 240 L2 = 1, N
    X(L2,1) = 1.0
    X(L2,2) = L2
    X(L2,3) = (L2*L2)/2
240  CONTINUE
  DO 150 K1 = 1, 3
    DO 120 K2 = 1, 3
      A(K1,K2) = 0.0
120  CONTINUE
    B(K1) = 0.0
150  CONTINUE

*** OLS, R-SQUARE(DENOMINATOR), T-STATISTICS
  HAP = 0
  DO 200 L = 1, N
    A(1,1) = A(1,1) + X(L,1)*X(L,1)
    A(1,2) = A(1,2) + X(L,1)*X(L,2)
    A(1,3) = A(1,3) + X(L,1)*X(L,3)
    A(2,2) = A(2,2) + X(L,2)*X(L,2)
    A(2,3) = A(2,3) + X(L,2)*X(L,3)
    A(3,3) = A(3,3) + X(L,3)*X(L,3)
    B(1) = B(1) + X(L,1)*Y(L)
    B(2) = B(2) + X(L,2)*Y(L)
    B(3) = B(3) + X(L,3)*Y(L)
    SQY = SQY + Y(L)*Y(L)
    HAP = HAP + Y(L)
200  CONTINUE
    A(2,1) = A(1,2)
    A(3,1) = A(1,3)
    A(3,2) = A(2,3)
    YBAR = HAP**2/N
    RSQDENO = SQY - YBAR

  CALL DLINRG(3,A,MDA,AINV,MDAINV)

  DO 280 MDAINV1 = 1, 3
    SUM = 0.0

```

```

        DO 250 MDAINV2 = 1, 3
            SUM = SUM + AINV(MDAINV1,MDAINV2)*B(MDAINV2)
250    CONTINUE
        BETAHAT(MDAINV1) = SUM
280    CONTINUE
        SUMBETA(1) = SUMBETA(1) + BETAHAT(1)
        SUMBETA(2) = SUMBETA(2) + BETAHAT(2)
        SUMBETA(3) = SUMBETA(3) + BETAHAT(3)

*** RESIDUAL, R-SQUARE(NUMERATOR)
        SUM3 = 0.0
        DO 320 M1 = 1, N
            SUM2 = 0.0
            DO 300 M2 = 1, 3
                SUM2 = SUM2 + X(M1,M2)*BETAHAT(M2)
300    CONTINUE
            SUM3 = SUM3 + Y(M1)*SUM2
            RESID(M1) = Y(M1) - SUM2
320    CONTINUE
        RSQNUM = SUM3 - YBAR
        RSQUARE = RSQNUM/RSQDENO
        SUMRSQ = SUMRSQ + RSQUARE

*** DW, SIGMASUARE HAT
        DENO = 0.0
        DNUM = 0.0
        DO 350 M3 = 1, N-1
            DNUM = DNUM + (RESID(M3+1) - RESID(M3))**2
            DENO = DENO + RESID(M3+1)**2
350    CONTINUE
        DENO = DENO + RESID(N)**2
        DW = DNUM/DENO
        SIGMASQ = DENO/(N-P-1)
        T_STAT1 = BETAHAT(1)/SQRT(AINV(1,1)*SIGMASQ)
        T_STAT2 = BETAHAT(2)/SQRT(AINV(2,2)*SIGMASQ)
        T_STAT3 = BETAHAT(3)/SQRT(AINV(3,3)*SIGMASQ)
        SUMT1 = SUMT1 + T_STAT1
        SUMT2 = SUMT2 + T_STAT2
        SUMT3 = SUMT3 + T_STAT3
        SUMDW = SUMDW + DW
        SUMSIGSQ = SUMSIGSQ + SIGMASQ
        RETURN
        END

```

* END OF PROGRAM


```
*****
* MISSPECIFICATION OF POLYNOMIAL REGRESSION MODEL OF ORDER 2 *
* AS AN AR(2) MODEL. *
*****
```

```
*****
* MAIN PROGRAM *
*****
```

```
PROGRAM MAIN
PARAMETER(N=250, NUM=10000, N2=3)
IMPLICIT REAL*8 (A-H, O-Z)
REAL*8 E(N), Y(N), X(N,3), RESID(N), SUMGAMMA(N2)
EXTERNAL RNSET, RNNOA
ISEED = 13579
CALL RNSET(ISEED)
```

```
DO 50 I= 1, NUM
```

```
CALL RNNOA(N,E)
DO 10 J= 1, N
Y(J) = J + J*J/2 + E(J)
10 CONTINUE
```

```
CALL GAMMA(N,N2,Y,X,RESID,SUMGAMMA,SUMRSQ,SUMDW,
SUMSIGSQ,SUMT1,SUMT2,SUMT3)
```

```
50 CONTINUE
EPGAMMA1 = SUMGAMMA(1)/NUM
EPGAMMA2 = SUMGAMMA(2)/NUM
EPGAMMA3 = SUMGAMMA(3)/NUM
EPRSQ = SUMRSQ/NUM
EPDW = SUMDW/NUM
EPSIGSQ = SUMSIGSQ/NUM
EPT1 = SUMT1/NUM
EPT2 = SUMT2/NUM
EPT3 = SUMT3/NUM
WRITE(*,*) EPGAMMA1, EPGAMMA2, EPGAMMA3
WRITE(*,*) EPT1, EPT2, EPT3
WRITE(*,*) EPRSQ, EPDW, EPSIGSQ
END
```

```
*****
* SUBROUTINE 1 *
*****
```

```

SUBROUTINE GAMMA(N,N2,Y,X,RESID,SUMGAMMA,SUMRSQ,
/ SUMDW,SUMSIGSQ,SUMT1,SUMT2,SUMT3)
```

```

PARAMETER(MDA =3, MDAINV=3, P=2)
REAL*8 A(MDA,MDA), AINV(MDAINV,MDAINV), B(MDA)
REAL*8 X(N,3), Y(N), GAMMAHAT(MDA), RESID(N), SQY, YBAR
REAL*8 SUMGAMMA(N2)
REAL*8 RSQDENO, RSQNUM, RSQUARE
REAL*8 SUM, SUM2, SUM3, DNUM, DENO, DW, SIGMASQ
REAL*8 T_STAT1, T_STAT2, T_STAT3
REAL*8 SUMT1, SUMT2, SUMT3
REAL*8 SUMDW, SUMSIGSQ

SQY = 0.0
DO 150 K1 = 1, 3
  DO 120 K2 = 1, 3
    A(K1,K2) = 0.0
120   CONTINUE
    B(K1) = 0.0
150   CONTINUE

X(1,1) = 1.0
X(2,1) = 1.0
X(1,2) = 0.0
X(2,2) = Y(1)
X(1,3) = 0.0
X(2,3) = 0.0
DO 180 L2 = 3, N
  X(L2,1) = 1.0
  X(L2,2) = Y(L2-1) - Y(L2-2)
  X(L2,3) = Y(L2-2)
180   CONTINUE

*** OLS, R-SQUARE(DENOMINATOR), T-STATISTICS
HAP = 0.0
DO 200 L = 1, N
  A(1,1) = A(1,1) + X(L,1)**2
  A(1,2) = A(1,2) + X(L,1)*X(L,2)
  A(1,3) = A(1,3) + X(L,1)*X(L,3)
  A(2,2) = A(2,2) + X(L,2)**2
  A(2,3) = A(2,3) + X(L,2)*X(L,3)
  A(3,3) = A(3,3) + X(L,3)**2
  B(1) = B(1) + X(L,1)*Y(L)
  B(2) = B(2) + X(L,2)*Y(L)
  B(3) = B(3) + X(L,3)*Y(L)
  SQY = SQY + Y(L)*Y(L)
  HAP = HAP + Y(L)
200   CONTINUE
  A(2,1) = A(1,2)

```

```

A(3,1) = A(1,3)
A(3,2) = A(2,3)
YBAR = HAP**2/N
RSQDENO = SQY - YBAR

CALL INVMAT(AINV,A,3)

DO 280 MDAINV1 = 1, 3
  SUM = 0.0
  DO 250 MDAINV2 = 1, 3
    SUM = SUM + AINV(MDAINV1,MDAINV2)*B(MDAINV2)
250  CONTINUE
  GAMMAHAT(MDAINV1) = SUM
280  CONTINUE
  SUMGAMMA(1) = SUMGAMMA(1) + GAMMAHAT(1)
  SUMGAMMA(2) = SUMGAMMA(2) + GAMMAHAT(2)
  SUMGAMMA(3) = SUMGAMMA(3) + GAMMAHAT(3)

*** RESIDUAL, R-SQUARE(NUMERATOR)
SUM3 = 0.0
DO 320 M1 = 1, N
  SUM2 = 0.0
  DO 300 M2 = 1, 3
    SUM2 = SUM2 + X(M1,M2)*GAMMAHAT(M2)
300  CONTINUE
  SUM3 = SUM3 + Y(M1)*SUM2
  RESID(M1) = Y(M1) - SUM2
320  CONTINUE
RSQNUM = SUM3 - YBAR
RSQUARE = RSQNUM/RSQDENO
SUMRSQ = SUMRSQ + RSQUARE

*** DW, SIGMASUARE HAT, T-STATISTICS
DENO = 0.0
DNUM = 0.0
DO 350 M3 = 1, N-1
  DNUM = DNUM + (RESID(M3+1)-RESID(M3))**2
  DENO = DENO + RESID(M3)**2
350  CONTINUE
DENO = DENO + RESID(N)**2
DW = DNUM/DENO
SIGMASQ = DENO/(N-P-1)
T_STAT1 = GAMMAHAT(1)/SQRT(AINV(1,1)*SIGMASQ)
T_STAT2 = GAMMAHAT(2)/SQRT(AINV(2,2)*SIGMASQ)
T_STAT3 = GAMMAHAT(3)/SQRT(AINV(3,3)*SIGMASQ)

```

```

SUMT1 = SUMT1 + T_STAT1
SUMT2 = SUMT2 + T_STAT2
SUMT3 = SUMT3 + T_STAT3
SUMDW = SUMDW + DW
SUMSIGSQ = SUMSIGSQ + SIGMASQ
RETURN
END

```

```
*****
```

```
* SUBROUTINE 2 *
```

```
*****
```

```
*** COMPUTE INVERSE OF A MATRIX.
```

```

SUBROUTINE INVMAT(AINV,A,N)
REAL*8 A(3,3), L(3,3), U(3,3)
REAL*8 AINV(3,3), UINV(3,3), LINV(3,3)
CALL LUDCOM(L,U,A,N)
DO 100 I = 1, N
100 LINV(I,I) = 1.0
DO 110 I = 1, N-1
DO 110 J = I+1, N
110 LINV(I,J) = 0.0
DO 130 J = 1, N-1
DO 130 I = J+1, N
SUM = 0.0
DO 120 K = J+1, I-1
120 SUM = SUM + L(I,K)*LINV(K,J)
LINV(I,J) = -L(I,J)-SUM
130 CONTINUE
DO 140 I = 1, N
140 UINV(I,I) = 1/U(I,I)
DO 150 J = 1, N-1
DO 150 I = J+1, N
150 UINV(I,J) = 0.0
DO 170 I = N-1, 1, -1
DO 170 J = N, I+1, -1
SUM = 0.0
DO 160 K = I+1, J
160 SUM = SUM + U(I,K)*UINV(K,J)
UINV(I,J) = UINV(I,I)*(-SUM)
170 CONTINUE
CALL PROMAT(AINV,UINV,LINV,N,N,N)
RETURN
END

```

```
*
```

```
SUBROUTINE LUDCOM(L,U,A,N)
REAL*8 L(3,3), U(3,3), A(3,3), WM(3,3)
DO 100 K = 1, N
  L(K,K) = 1.0
  U(1,K) = A(1,K)
100 CONTINUE
  DO 150 I = 2, N
    DO 120 J = 1, I-1
      WM(I,J) = 0.0
      DO 110 M = 1, J-1
110      WM(I,J) = WM(I,J) + L(I,M)*U(M,J)
        L(I,J) = (A(I,J)-WM(I,J))/U(J,J)
120 CONTINUE
      DO 140 J = I, N
        WM(I,J) = 0.0
        DO 130 M = 1, I-1
130      WM(I,J) = WM(I,J) + L(I,M)*U(M,J)
          U(I,J) = A(I,J) - WM(I,J)
140 CONTINUE
150 CONTINUE
  RETURN
END
```

*

```
SUBROUTINE PROMAT(C,A,B,N,L,M)
REAL*8 C(3,3), A(3,3), B(3,3)
DO 100 I = 1, N
  DO 100 J = 1, M
    C(I,J) = 0.0
    DO 100 K = 1, L
      C(I,J) = C(I,J) + A(I,K)*B(K,J)
100 CONTINUE
  RETURN
END
```

* END OF PROGRAM

APPENDIX H

FORTRAN PROGRAM
FOR
UNDERSPECIFICATION
OF
THE ORDER IN A NONSTATIONARY AR(2) MODEL AND
POLYNOMIAL REGRESSION MODEL OF ORDER 2

```
*****
* UNDERSPECIFICATION OF AR(2) AS AR(1) *
*****
```

```
*****
* MAIN PROGRAM *
*****
```

```
PROGRAM MAIN
PARAMETER(N = 250, NUM = 10000, N2 = 2)
IMPLICIT REAL*8 (A-H, O-Z)
REAL*8 U(N), X(N), Y(N), XX(N,2), RESID(N), SUMGAMMA(N2)
EXTERNAL RNSET, RNNOA
ISEED = 13579
```

```
CALL RNSET(ISEED)
DO 50 I = 1, NUM
    CALL RNNOA(N,U)
    Y(1) = U(1)
    Y(2) = 2*Y(1) + U(2)
    DO 10 J = 3, N
        Y(J) = 2*Y(J-1) - Y(J-2) + U(J)
10 CONTINUE

    CALL GAMMA(N,N2,X,Y,XX,RESID,SUMGAMMA,SUMRSQ,
/          SUMDW,SUMSIGSQ,SUM1,SUMT2)
50 CONTINUE
EPGAMMA1 = SUMGAMMA(1)/NUM
EPGAMMA2 = SUMGAMMA(2)/NUM
EMPRSQ = SUMRSQ/NUM
EMPDW = SUMDW/NUM
EMPSIGSQ = SUMSIGSQ/NUM
EMPT1 = SUMT1/NUM
EMPT2 = SUMT2/NUM
WRITE(*,*) EPGAMMA1, EPGAMMA2
WRITE(*,*) EMPT1, EMPT2
WRITE(*,*) EMPRSQ, EMPDW, EMPSIGSQ
END
```

```
*****
* SUBROUTINE *
*****
```

```
SUBROUTINE GAMMA(N,N2,X,Y,XX,RESID,SUMGAMMA,SUMRSQ,
/          SUMDW,SUMSIGSQ,SUMT1,SUMT2)
PARAMETER(MDA=2, MDAINV=2, P=1)
```

```

REAL*8 A(MDA,MDA), AINV(MDAINV,MDAINV), B(MDA)
REAL*8 X(N), XX(N,2), Y(N), GAMMAHAT(MDA), RESID(N)
REAL*8 SUMGAMMA(N2), SQY, YBAR, RSQDENO, RSQNUM, RSQUARE
REAL*8 SUM, SUM2, SUM3, SUMRSQ, DENO, DNUM, DW, SIGMASQ
REAL*8 T_STAT1, T_STAT2, SUMT1, SUMT2, SUMDW, SUMSIGSQ

SQY = 0.0
XX(1,1) = 1.0
XX(1,2) = 0.0
DO 240 L2 = 2, N
    XX(L2,1) = 1.0
    XX(L2,2) = Y(L2-1)
240 CONTINUE

DO 150 K1 = 1, 2
    DO 120 K2 = 1, 2
        A(K1,K2) = 0.0
120 CONTINUE
    B(K1) = 0.0
150 CONTINUE

*** OLS, R-SQUARE(DENOMINAOR), T-STATISTICS
HAP = 0.0
DO 200 L = 1, N
    A(1,1) = A(1,1) + XX(L,1)**2
    A(1,2) = A(1,2) + XX(L,1)*X(L,2)
    A(2,2) = A(2,2) + XX(L,2)**2
    B(1) = B(1) + XX(L,1)*Y(L)
    B(2) = B(2) + XX(L,2)*Y(L)
    SQY = SQY + Y(L)*Y(L)
    HAP = HAP + Y(L)
200 CONTINUE
A(2,1) = A(1,2)
YBAR = HAP**2/N
RSQDENO = SQY - YBAR

CALL DLINRG(2,A,MDA,AINV,MDAINV)

DO 280 MDAINV1 = 1, 2
    SUM = 0.0
    DO 250 MDAINV2 = 1, 2
        SUM = SUM + AINV(MDAINV1,MDAINV2)*B(MDAINV2)
250 CONTINUE
    GAMMAHAT(MDAINV1) = SUM
280 CONTINUE

```



```
SUMGAMMA(1) = SUMGAMMA(1) + GAMMAHAT(1)
SUMGAMMA(2) = SUMGAMMA(2) + GAMMAHAT(2)
```

```
*** RESIDUAL, R-SQUARE(NUMERATOR)
```

```
  SUM3 = 0.0
  DO 320 M1 = 1, N
    SUM2 = 0.0
    DO 300 M2 = 1, 2
      SUM2 = SUM2 + XX(M1,M2)*GAMMAHAT(M2)
300    CONTINUE
      SUM3 = SUM3 + Y(M1)*SUM2
      RESID(M1) = Y(M1) - SUM2
320    CONTINUE
  RSQNUM = SUM3 - YBAR
  RSQUARE = RSQNUM/RSQDENO
  SUMRSQ = SUMRSQ + RSQUARE
```

```
*** DW, SIGMASUARE HAT
```

```
  DENO = 0.0
  DNUM = 0.0
  DO 350 M3 = 1, N-1
    DNUM = DNUM + (RESID(M3+1) - RESID(M3))**2
    DENO = DENO + RESID(M3+1)**2
350  CONTINUE
  DENO = DENO + RESID(N)**2
  DW = DNUM/DENO
  SIGMASQ = DENO/(N-P-1)
  T_STAT1 = GAMMAHAT(1)/SQRT(AINV(1,1)*SIGMASQ)
  T_STAT2 = GAMMAHAT(2)/SQRT(AINV(2,2)*SIGMASQ)
  SUMT1 = SUMT1 + T_STAT1
  SUMT2 = SUMT2 + T_STAT2
  SUMDW = SUMDW + DW
  SUMSIGSQ = SUMSIGSQ + SIGMASQ
  RETURN
  END
```

```
* END OF PROGRAM
```

```
*****
*   UNDERSPECIFICATION OF POLYNOMIAL REGRESSION OF ORDER 2   *
*   ORDER 2 AS POLYNOMIAL REGRESSION MODEL OF ORDER 1.       *
*****
```

```
*****
*   MAIN PROGRAM   *
*****
```

```
PROGRAM MAIN
PARAMETER(N=250, NUM=10000, N2=2)
IMPLICIT REAL*8 (A-H,O-Z)
REAL*8 E(N), X(N), Y(N), XX(N,2), RESID(N), SUMBETA(N2)
EXTERNAL RNSET, RNNOA
ISEED = 13579
```

```
CALL RNSET(ISEED)
```

```
DO 50 I = 1, NUM
  CALL RNNOA(N,E)
  DO 10 J = 1, N
    X(J) = J
    Y(J) = X(J) + X(J)*X(J)/2 + E(J)
10  CONTINUE
```

```
CALL BETA(N,N2,X,Y,XX,RESID,SUMBETA,SUMRSQ,SUMDW,
/      SUMSIGSQ,SUMT1,SUMT2)
```

```
50 CONTINUE
  EMPBETA1 = SUMBETA(1)/NUM
  EMPBETA2 = SUMBETA(2)/NUM
  EMPRSQ = SUMRSQ/NUM
  EMPDW = SUMDW/NUM
  EMPSIGSQ = SUMSIGSQ/NUM
  EMPT1 = SUMT1/NUM
  EMPT2 = SUMT2/NUM
  WRITE(*,*) EMPBETA1, EMPBETA2
  WRITE(*,*) EMPT1, EMPT2
  WRITE(*,*) EMPRSQ, EMPDW, EMPSIGSQ
END
```

```
*****
*   SUBROUTINE   *
*****
```

```
SUBROUTINE BETA(N,N2,X,Y,XX,RESID,SUMBETA,SUMRSQ,
```

```

/          SUMDW,SUMSIGSQ,SUMT1,SUMT2)
PARAMETER(MDA=2, MDAINV=2, P=1)
REAL*8 A(MDA,MDA), AINV(MDAINV,MDAINV), B(MDA)
REAL*8 X(N), XX(N,2), Y(N), BETAHAT(MDA), RESID(N)
REAL*8 SUMBETA(N2)
REAL*8 SQY, YAR, RSQDENO, SUM, SUM2, SUM3
REAL*8 RSQNUM, RSQUARE, SUMRSQ, DENO, DNUM
REAL*8 SIGMASQ, DW, SUMDW, SUMSIGSQ
REAL*8 T_STAT1, T_STAT2, SUMT1, SUMT2
SQY = 0.0
DO 150 K1 = 1, 2
    DO 120 K2 = 1, 2
        A(K1,K2) = 0.0
120    CONTINUE
        B(K1) = 0.0
150    CONTINUE

*** OLS, R-SQUARE(DENOMINATOR), T-STATISTICS
DO 200 L = 1, N
    A(1,2) = A(1,2) + X(L)
    A(2,2) = A(2,2) + X(L)**2
    B(1) = B(1) + Y(L)
    B(2) = B(2) + X(L)*Y(L)
    SQY = SQY + Y(L)*Y(L)
200 CONTINUE
    A(1,1) = N
    A(2,1) = A(1,2)
    YBAR = B(1)*B(1)/N
    RSQDENO = SQY - YBAR

CALL DLINRG(2,A,MDA,AINV,MDAINV)

DO 240 L2 = 1, N
    XX(L2,1) = 1.0
    XX(L2,2) = L2
240 CONTINUE

DO 280 MDAINV1 = 1, 2
    SUM = 0.0
    DO 250 MDAINV2 = 1, 2
        SUM = SUM + AINV(MDAINV1,MDAINV2)*B(MDAINV2)
250    CONTINUE
    BETAHAT(MDAINV1) = SUM
280 CONTINUE
    SUMBETA(1) = SUMBETA(1) + BETAHAT(1)
    SUMBETA(2) = SUMBETA(2) + BETAHAT(2)

```

```

*** RESIDUAL, R-SQUARE(NUMERATOR)
    SUM3 = 0.0
    DO 320 M1 = 1, N
        SUM2 = 0.0
        DO 300 M2 = 1, 2
            SUM2 = SUM2 + XX(M1,M2)*BETAHAT(M2)
300    CONTINUE
        SUM3 = SUM3 + Y(M1)*SUM2
        RESID(M1) = Y(M1) - SUM2
320    CONTINUE
        RSQNUM = SUM3 - YBAR
        RSQUARE = RSQNUM/RSQDENO
        SUMRSQ = SUMRSQ + RSQUARE

*** DW, SIGMASUARE HAT
    DENO = 0.0
    DNUM = 0.0
    DO 350 M3 = 1, N-1
        DNUM = DNUM + (RESID(M3+1) - RESID(M3))**2
        DENO = DENO + RESID(M3)**2
350    CONTINUE
    DENO = DENO + RESID(N)**2
    DW = DNUM/DENO
    SIGMASQ = DENO/(N-P-1)
    T_STAT1 = BETAHAT(1)/SQRT(AINV(1,1)*SIGMASQ)
    T_STAT2 = BETAHAT(2)/SQRT(AINV(2,2)*SIGMASQ)
    SUMT1 = SUMT1 + T_STAT1
    SUMT2 = SUMT2 + T_STAT2
    SUMDW = SUMDW + DW
    SUMSIGSQ = SUMSIGSQ + SIGMASQ
    RETURN
    END

```

* END OF PROGRAM

VITA

DONG-BIN JEONG

Candidate for the Degree of

Doctor of Philosophy

Thesis: TESTING FOR A UNIT ROOT IN AN AR(p) SIGNAL OBSERVED
WITH MA(q) NOISE AND MODEL MISSPECIFICATION

Major Field: Statistics

Biographical:

Personal Data: Born in Iksan, South Korea, August 11, 1967, the son of
Dr. Jeong, Gab-Won and Kim, Won-Kyong.

Education: Graduated from Nam-Sung High school, Iksan, South Korea, in
February 1985; received Bachelor of Science degree in Statistics from
Chung-Ang University, Seoul, South Korea, in February 1989;
received Master of Science degree at Chung-Ang University in August,
1991; completed requirements of the Doctor of Philosophy Degree at
Oklahoma State University in December 1996.

Professional Experience: Teaching Assistant, Language Institute in Chung-
Ang University, 1990 - 1991; Research Assistant, Statistics, Oklahoma
State University, 1993; Teaching Assistant, Statistics, Oklahoma State
University, 1993 - 1996.