CONSISTENT TESTING FOR LAG LENGTH IN COINTEGRATED RELATIONSHIPS

by

Limin Liu

Submitted in fulfilment of the requirements for the degree of

Magister Scientiae

in the

Faculty of Science

at the Nelson Mandela Metropolitan University

January 2007

Supervisor: Mr G. D. Sharp

Abstract

In the past few decades the theory of cointegration has been widely used in the empirical analysis of economic data. The reason is that, it captures the economic notion of a long-run economic relation. One of the problems experienced when applying cointegrated techniques to econometric modelling is the determination of lag lengths for the modelled variables. Applied studies have resulted in contradictory choices for lag length selection. This study reviews and compares some of the well-known information criteria using simulation techniques for bivariate models.

Key words

Unit root test

Test of cointegration

Lag length selection criteria

Contents

Acknowledgements

I would like to express my gratitude to the following:

- My father, mother and brother for their support. I love you all very much.

- My cousin and my friends, thanks for your encouragement and friendship when I needed it the most.

- My supervisor, Mr Sharp, thanks very much for all your hard work and guidance. Thanks also for your patience and encouragement. I really appreciated it.

- To Vodacom, the NRF and NMMU whose sponsorship funded this project.

Chapter 1

Introduction

In the past few decades there has been a significant increase in development of econometric models. One of the more recent approaches is based on the theory of cointegrated variables. The theory of cointegration provides econometricians with a practical procedure to analyze non-stationary time series data. The methodology of cointegration has generated substantial interest amongst both econometricians and statisticians and examples of the applications of cointegrated modelling have been published in international journals, see for example Diamandis and Kouretas (1995) and Dwyer and Wallace (1992), and in local South African journals, see for example Ferret and Page (1998) and Wilson, Okunev, du Plessis and Ta (1998).

Soren Johansen (1988) developed a maximum likelihood estimation procedure for multivariate cointegrated time series models. The phenomenal growth in software development has seen the algorithms developed by Johansen being incorporated into econometric software packages such as Shazam, EViews and Cats for Rats. The availability of these packages has resulted in widespread application of cointegrated modelling.

The determination of model lag length is a problem analysts experience when using cointegrated techniques. Criteria such as Akaike's information criterion (AIC), Schwarz's information criterion (SIC), Final Prediction Error (FPE) and Hannan-Quinn's information criterion (HQ) have been used in cointegrated modelling to assist in selecting the lag length of a model from a dataset.

Although the theory of model selection using information criteria was developed some thirty years ago, the benefits from a practical point of view were largely ignored. This changed when the practical benefits of information criteria were acknowledged by Anderson, Burnham and White (1994) in their capture-recapture studies.

Traditionally, likelihood ratio hypothesis tests were used to compare models of different lag lengths. Despite the usefulness of likelihood ratio tests for model comparisons, Sclove (1994) illustrates several examples where an information criterion approach is preferable for model selection. One of Sclove's illustrative examples considers the scenario of comparing two treatment means with unequal population variances. This two treatment means comparison realizes four possible models and using an information criteria approach yields better results than a sequence of likelihood ratio hypothesis tests.

The four information criteria mentioned previously have been used in several research disciplines. In economics, Akinboade and Niedermeier (2002) use AIC in determining the relationship between labour costs and inflation in South Africa. Mainardi (2000) uses AIC in estimating the consumption rate of interest for Trinidad and Tobago. In a biological study Bozdogan, Sclove and Gupta (1994) use AIC to identify the best fitting parametric multivariate model of male Egyptian skulls from Thebes. In psychological research, Takane (1994) reviews the application of AIC and reveals some of the difficulties in modelling psychological phenomena.

AIC is arguably the most well-known of the information criteria used in model selection. The AIC algorithms have been integrated into econometric and statistical software packages such as EViews, Stata, Statistica, Shazam and Cats for Rats. The value of AIC is calculated for the purpose of model comparison and the minimum AIC indicates the most appropriate model to select.

In spite of the popularity of AIC, Lütkepohl (1985) and Gonzalo and Pitarakis (2002) have illustrated that the information criterion has a tendency to over-parameterize when sample size is unrealistically small. Bozdogan (1987) extended the methodology of AIC to be asymptotically consistent to neutralize the problem of over-parameterization and developed a criterion called the Consistent Akaike's information criterion (CAIC). In the Gonzalo and Pitarakis's (2002) study, they show that a criterion such as the Bayesian information criterion (BIC) and Hannan-Quinn's information criterion (HQ) lead to consistent estimates in both stationary and non-stationary systems.

In research published during 2004, Khim and Liew conduct a simulation study to determine the optimal lag length for an autoregressive process. In their study, they used several information criteria to analyze models with different lag lengths. In addition, they used various sample sizes ranging from 30 to 960 observations for their models.

Initially, Khim and Liew used different criteria to calculate the probability of correctly estimating the true lag length of an autoregressive process for their different sample sizes. Thereafter, they utilised the information criteria to determine the probability of under-estimating the true lag length of an autoregressive process as well as the probability of over-estimating the true lag length of an autoregressive process for the various sample sizes. This research follows the methodology of Khim and Liew but unlike Khim and Liew this study conditions the autoregressive process to be non-stationary and cointegrated.

Anderson, Burnham and White (1994) provide an illustration of lag length selection criteria used for making inferences from ringing data (capture-recapture data). They explain, in their paper, that the theory of information criteria was developed along two ideas. The first theory assumes that a true model exists and that the information criteria will determine the best fitting true model for the data. This theory consists of criteria that are "*dimension consistent*" and requires sufficiently large sample sizes. The second theory assumes that a true model does not exist but by using the information criteria the best fitting model for the data will be determined. This theory tries to minimize the loss of information and Kullback-Leibler (1951) addressed this issue when developing the measure, Kullback-Leibler information coefficients, which represent the information lost when approximating reality.

The information criteria based on Kullback-Leibler information coefficients include Akaike's information criterion, corrected Akaike's information criterion, Final Prediction Error and Takeuchi's information criterion. The information criteria that are "*dimension consistent*" include Bayesian information criterion, Minimum Description Length and Hannan-Quinn's information criterion.

The determination of the lag length for a model with the information criterion methodology has played an important role in many studies. Several studies found that certain criteria are better than others (see Lütkepohl (1985) and Khim and Liew (2004)) in determining the order of autoregressive processes. It is the objective of this study to review four of the well-known criteria and compare these criteria using simulation analysis for cointegrated time series bivariate data.

Chapter 2

Literature Review

Over the last two decades several studies have been conducted to determine the order of an autoregressive (AR) process. The studies of both Lütkepohl (1985) and Ng and Perron (2005) use information criteria as aids for choosing the order of the AR process. The use of AIC, the better known information criteria, for determining the order of the AR process, has shown how easy the information criteria are to apply whilst emphasizing the benefits of the techniques. To date, a review of the available literature has yielded little research on information criteria decision making for determining the number of lag terms in a cointegrated time series[1]. The application of the cointegration theory to econometric studies has provided a platform for numerous research publications, both locally and internationally. A brief review of some of the studies that have applied the methodology successfully is provided in the forthcoming sections. The one noticeable inconsistency in the published studies is the lack of a clear approach for determining the lag lengths of the analyzed cointegrated models.

2.1 Review of selected publications using Akaike's information criterion

In the equity market study by Botha and Apostolellis (2003) an analysis is performed of the financial integration between equity markets using cointegration techniques. Their analysis seeks to provide a better understanding of the long-term relationships and short-term dynamics that exist between emerging and developed markets. They opt to use Akaike's information criterion for the determination of the number of lag terms in their multi-equation model. A lag of one was selected, given that the stock prices

[1] A recent publication by Ivanov and Kilian (2005) attempts to address lag length selection in cointegrated models. This paper was only recently discovered and not reviewed in this study.

tend upward over time and a linear deterministic trend is included in the equation model.

The role of future markets in the South African financial markets was examined by Fedderke and Joao (2001) during a period of considerable volatility on world financial markets. The motivation for their study was that this issue has implications for the fundamental concepts in financial theory, particularly in market efficiency and arbitrage. In their empirical analysis, two cointegration techniques were used to determine the linear relationships between the variables and these techniques are the ARDL cointegration approach and Johansen's method of testing for cointegration. The ARDL cointegration technique was developed by Pesaran and Shin (1995a, 1995b) and Pesaran, Shin and Smith (1996) and was used to determine the nature of the patterns of association between the variables. Both methods provide a cointegrating relationship between the variables and the coefficients of the cointegrating equation are similar. Akaike's information criterion was used to determine the number of lag terms in the estimated multi-equation model. As a result, a lag of eight was used for each of the equations in the Johansen's test and for the ARDL cointegration approach a different lag term was used for different variables in each equation.

In research published during 2001, Fedderke, de Kadt and Luiz used several new measures (such as political and property rights) to explore the link between institutions and economic activities for South Africa. They used Johansen's test of cointegration to establish the importance of property rights and political instability as determinants of the level of desired per capita output, capital-labour ratio and investment expenditure. Akaike's information criterion was used to determine the number of lag terms in the VAR models. Akaike's information criterion was also used to select the number of lag terms in the ARDL models for ARDL cointegration technique.

2.2 Review of selected publications using General-to-specific model selection procedure

The general-to-specific (Gets) model selection procedure is a commonly used approach for the selection of the number of lag terms in an autoregressive model. This procedure uses a likelihood ratio test statistic to test the lag terms in the model. An example illustrating the likelihood ratio test on the different lag terms was provided in Barr and Kantor (1990). The Gets procedure requires that one start with a relatively large number of lag terms and pares down the model using a likelihood ratio test statistic. As an example, one could estimate a regression equation using a lag of (p). If the likelihood ratio test statistic is insignificant at some specific critical value, re-estimate the regression using a lag of (p - 1). Repeat the process until the lag is significantly different from zero. It is worth noting that for the general-to-specific approach one can also perform this procedure by keeping the initial lag (m) the same in each test or setting $p = m - i$ where $i = 0, 1, \ldots, m - 1$.

This method is the equivalent to the backwards stepwise regression procedure advocated in many undergraduate text books (see Mendenhall and Sincich (2003, pg. 327)) with the exception that this is specific to autoregressive models. Once a tentative lag term has been determined, diagnostic checking is then used to test for the presence of heteroscedasticity in the lagged model. Krolzig and Hendry (2000) developed a software package called PcGets to analyze the Gets procedure of the Vector Autoregressive models. PcGets is a computer-automated approach to econometric modelling focusing on general-to-specific reduction approaches for linear, dynamic and regression models. Several studies used the Gets procedure to determine the number of lag terms for their time series models and some of these studies are reviewed below.

The use of an information criterion to determine the number of lag terms has been used extensively in distributed lag models such as autoregressive models, vector autoregressive models and moving average models. Barr and Kantor (1990) used a Vector Autoregressive (VAR) process in their time series models. This process was used to analyze the interrelationship between the GDP at constant 1985 prices, the rand value of notes in circulation, the private consumption expenditure deflator and the ratio of the trade balance to nominal GDP in the South African economy. The general-to-specific model selection procedure was used to determine the number of lag terms in the models. The highest lag selected in the model was eight. This model was tested against the alternative ones with a progressively smaller number of lags. The optimal lag chosen for each of the time series functions was four.

Barr and Kahn (1995) tested the behaviour of the Purchasing Power Parity (PPP) relationship in a South African study using cointegration techniques. Their study was sub-divided into three periods to take into account various shocks to the economy and the changes in the policy regime. In the empirical analysis of their study, they used the general-to-specific model selection procedure to select the number of lag terms in the multi-equation model. The initial lag selected for the variables in all equations was eight. This lag was "*tested down*" to a lag that is statistically significant at the 5% significance level. The result concluded that the explanatory variable lag terms were different, a finding which contradicted the results of other studies one could consider as similar. In the studies by Diamandis and Kouretas (1995), Barr and Kantor (1990) and Gumede (2000), all explanatory variable equations selected had the same number of lag terms.

It is widely acknowledged that the lack of foreign exchange is the main obstacle to economic growth in South Africa. Many studies have focused on exports and export expansions as a means to eradicate this economic dilemma (Gumede (2000)). The Gumede (2000) study provided a contribution to the understanding of South Africa's foreign trade outlook by examining the import demand elasticities for South Africa with respect to the real income and relative prices using econometric analysis. The estimation of the long-run relationship between these variables was implemented using the Engle-Granger two-step procedure. The general-to-specific model selection criterion was used to estimate the number of lag terms in the model. The result of the number of lag terms in the model was not provided.

Madsen (1997) evaluates the macroeconomic implications of fiscal and monetary policies in South Africa using a four equation model which allows for demand and supply side interaction. The model identifies several channels along which monetary and fiscal policies feed into prices, wages, employments and outputs, in the short and long-run. Error correction models are estimated for that purpose. Parameter estimates of the cointegration equations provide the information for long-run relationships, whereas the error correction term gives insight into short-run adjustment towards the long-run equilibrium. Since the ordinary least square estimates of the price and the wage equation model give inefficient estimates and the distribution of the estimators are non-normal, the Engle and Yoo (1989) three step procedure, which gives efficient and normal distributed parameter estimates, was used to estimate the cointegrating model. The lag structure for the estimated models were determined by the general-to-specific model selection procedure. The initial lag selected was three and the insignificant lags sequentially deleted with a 5% significance level. The result of the Gets procedure for lag selection was not indicated and therefore one is unable to identify the number of lag terms incorporated in the model.

Kouassi (1997) examines the impact of terms of trade shocks on Ivorian macroeconomic variables, in the context of an open economy, using the theory of cointegration. Johansen’s method of testing for cointegration was used to analyze possible linear relationship between the Ivorian macroeconomic variables. The choice of the lag structure in the model was selected by the Sims (1980) likelihood ratio test. As a result a lag of two for the macro-model was selected. Diagnostic test statistics such as the

Box-Ljung Q statistic and the coefficient of determination was used to test for the presence of serial correlation in the model. In the lag two model, no presence of serial correlation was detected, therefore supporting the choice of two lag terms in the model.

To summarise, the Gets procedure has been used quite extensively as evidenced by the review of the papers in this section. A distinct drawback to the procedure is the lack of explanation of significant levels used for lag selection and how the author(s) selected the maximum lag to start the procedure.

2.3 Review of selected publications using Final Prediction Error

In recent years, South Africa's Rand to US dollar exchange rate has been volatile which has created some concern for the South African economy. Damoense (2003) used the cointegration method to investigate some of the key determinants of the South African exchange rate. Variables of interest to the study were money supply, differential inflation rate, differential interest rate, and relative national income. Johansen's test of cointegration was used to test the linear relationships between the variables. They opted to use Akaike's Final Prediction Error to determine the number of lag terms in the multi-equation model. The result indicates a two-year lag for inflation rate differential, a one-year lag for money supply and interest differential and no lag for relative national income.

2.4 Review of selected publications using Akaike's information criterion and Schwarz's information criterion

The livestock sector has played a prominent role in South Africa in the past several decades and factors affecting livestock supply in South Africa were examined by Townsend (1999), using econometric techniques. A test of cointegration was used to examine whether factors such as producer price index, technology shifts, weather and others provided a linear relationship to the livestock supply in South Africa. Having established a cointegrating relationship and estimated the long-run elasticities, an error correction model was estimated to provide a valid representation of the data.

In order to reduce the number of estimated variables and so increase the degrees of freedom, a constrained form of the error correction model was estimated. A lag of three was estimated for the real livestock price index and a zero lag for the other variables. The author failed to explain how the number of lag terms were determined using an inferential test. It was assumed that a *t*-test was used to determine the significance of the lagged terms. The estimation of the short and long run elasticities revealed little information on the structure of the lagged relationship between the two factors, price index and research and development expenditure.

After due consideration it was decided to investigate the structure of the lagged relationship between these factors using an Almon polynomial lag approach. This approach allowed the lagged effects to be captured on output and to avoid the collinearity problems of the unrestricted model (Evenson (1967), Knutson and Tweeten (1979)). However, the Almon lag model required that the number of lag terms in the model be specified. Akaike's information criterion and Schwarz's information criterion were used to determine the number of lag terms in this case. The result of both criteria indicated a lag of thirteen for the research and development expenditure factor and a lag of five for the real livestock price index factor.

2.5 Review of selected publications using Schwarz's information criterion and Hannan-Quinn's information criterion

In research published during 2002, Leng performed an investigation on the efficiency of the All Share Index (ALSI) 40 futures contract which was traded on the South African Futures Exchange (SAFEX). This investigation was accomplished by studying the temporal causal dynamics between the futures price and its underlying spot index price before, during and after the Asian economic crisis. Leng used Johansen's method to test for cointegration between the ALSI futures contract and the spot market index over the period January 1996 to June 2001.

The result of the study indicated that a long-run equilibrium relationship existed between these two markets. As a result of this relationship, the error correction model for these two markets was estimated. In the error-correction model fitted, Leng opted to use a combination of Schwarz's information criterion and Hannan-Quinn's information criterion to determine the number of lag terms for the model. The result of these two criteria indicated a lag of one was to be incorporated in the error-correction model. Surprisingly, rather than accepting this, Leng selected a lag of three. He motivated this decision by indicating that he wanted to minimize the presence of heteroscedasticity that was present in the model.

It is these inconsistencies in deciding lag length of cointegrated models that provided the motivation for this research study. In Leng's paper, he omits to discuss how he determined the presence of heteroscedasticity and how the three lag model corrected it.

2.6 Review of selected publications which omitted to disclose the method(s) used to determine the lag length in the model

In recent years, several economists have asserted that asset prices determined in the efficient asset markets are not cointegrated and the first person who introduced this idea was Granger (1986, pg. 218). The implication that asset prices cannot be cointegrated relies on the definition of '*efficient markets*' as markets in which changes in asset prices are unpredictable. A more useful alternative definition of an efficient market is provided in Dwyer and Wallace (1992). Their definition indicated that there are no risk-free returns above opportunity cost available to agents given that there are transaction costs and agents' information. In their study, they examined some of the implications of this definition of efficient markets for the cointegration of asset prices and demonstrated that market efficiency does not preclude cointegration. The Engle-Granger two-step procedure was used to test for cointegration, however, no information was provided by the authors on how they determined the number of lag terms in the model.

Mainardi (1995) investigated the possible links between short and medium term variations in gold price, the corresponding performance of the nominal exchange rate with US dollar and real effective exchange rate in South Africa. In order to analyze the linear relationship between these variables, the model was assessed using the Pesaran, Shin and Smith ARDL cointegrating technique. The estimated equation for this method was a substantial improvement on the fitted model using the Engle-Granger procedure. The number of lag terms incorporated in the model was four. Unfortunately the method of lag length selection was omitted from the study, an omission which in this studies context is important.

The Bureau for Economic Research (BER) at the University of Stellenbosch has been using macroeconomic models for the purpose of short and medium-term economic forecasting of the South African economy. A recent version of the macroeconomic models was presented by Smit and Pellissier (1997) and their model was estimated using cointegration analysis. They followed the Engle-Granger two-step procedure whereby the long-run cointegrating equation was first estimated, followed by a short-run equation which included an error-correction term derived from the long-run equation. In their paper, Smit and Pellissier stipulate that standard statistical and economic criteria were used to determine the number of lag terms in the model but do not reveal which criteria were used.

The theory of Purchasing Power Parity (PPP) equilibrium has been widely used in sectors such as the academic, public and business sectors, however, the use of PPP equilibrium to analyze the exchange rate between countries is less extensive. De Wet (2000) used the theory of PPP to study the relationship between South Africa's exchange rate and its major trading partners' exchange rates. In his study, the time series variables were subjected to tests for unit roots and cointegration. The published findings make no reference to how the number of lag terms in the model were selected.

The importance of manufacturing to the South African economy was discussed in the study by Gumede (2003). The study informs us that about 20% of the economically active in South Africa are employed in the manufacturing sector and more than half of the South African exports are manufactured. In his study, where he estimated the export elasticities of the total economy and the manufacturing sector in South Africa, he used a time series model. Although the models are subjected to tests for cointegration, no disclosure on how the number of lag terms were selected was provided.

Neubrech and Pienaar (2001) examined the possible implications of the price, cross-price and income elasticities of the demand for public road transport in the Cape Metropole, using cointegration modelling. The results showed that there was a linear relationship between the price, cross-price and income elasticities of the demand for public road transport. Although a lag of three was selected in the multi-equation model, no disclosure on lag length selection was provided.

In research published during 1998, Sinha re-examines the export expansion hypothesis of Alfred Maizels using data from Asian countries. Following the earlier studies, Sinha used the Engle-Granger two-step procedure to determine the cointegrating relationships between the variables for each of the Asian countries. The method used to determine the number of lag terms in the multi-equation model was omitted from the paper.

All these published findings have one glaring omission, they did not include information on lag length selection of their cointegrated study. This omission provides an opportunity to question the validity of the studies' statistical procedures.

2.7 Econometric software package used

New software and upgraded versions of existing software regularly become available, thus choosing a suitable software package is essential. The choice of the optimal software package for this study is based on the cost of the software, the capability features that are incorporated in the software and the availability of the software package.

Wesso (1999) reviews twenty commonly used econometric software packages for personal computers. These software packages include SAS/ETS, Econometric Software Package (ESP), Shazam 8.0, SPSS 8.0 and EViews 3.0. The twenty software packages are compared based on the techniques the software is designed to perform.

Wesso's 1999 study followed his 1997 survey which was previously based on the econometric forecasting methods taught and practiced in South Africa. Vendors were asked to complete and return a questionnaire. Telephone and personal interviews were also used for follow-up in a few cases of incomplete or un-returned questionnaires.

Wesso discussed some of the capability features that were incorporated into these computer software packages. The capabilities include graph capabilities, read/write capabilities, hardware/software capabilities and miscellaneous procedures. The capabilities associated with each of the twenty software packages assisted us in selecting the ideal software for this study. Five out of the twenty software packages were suitable for this study based on the capabilities associated with them. The five software packages are EViews, Microfit, RATS, SAS and SHAZAM.

The SAS software package is most suitable for this study as it has all of the features required by this study. Although SAS is an excellent software package to use for this study, it is expensive so SAS was not chosen for this study.

The software package EViews would be the alternative package to use for this study. EViews is a powerful and easy-to-use econometric software package with satisfactory time series routines. Wesso (1999) provided information on the earlier version of EViews, EViews 3.0, and identified features incorporated into this software as well as the features that were lacking in this software package. The latest version of EViews is EViews 5.1 and several improvements have been introduced since EViews 3.0. The modifications include the improvement of the number of tests for unit roots, the increase of the number of model selection criteria tests and the ability to compute standard errors for impulse response functions and variance decomposition.

EViews has become an acceptable package for econometric analysis of windows-based software. This software is well explained and referenced in undergraduate/post-graduate textbooks (see Gujarati (1995)) and used in several research publications (see Botha and Apostolellis (2003) and de Wet (2000)). Given the features, cost and availability, it was decided to use EViews 5.1 in this study.

Chapter 3

Econometric Theory

3.1 Unit root tests

When analyzing stationary and non-stationary time series, it is necessary to test for the presence of a unit root to avoid the problem of spurious regression (Harris (1995, Pg. 6)). If a series contains a unit root then this series is non-stationary otherwise the series is stationary. The inferential tests of the unit root hypothesis are of interest to economists because they help to evaluate the nature of the non-stationarity that many macroeconomic series exhibit. Examples of macroeconomic variables whose data are possibly non-stationary include variables such as real exchange rate (Parikh and Kahn (1997)), spot market indices (Ferret and Page (1998)), business cycles (Moolman (2002)) and relative income (Gumede (2000)).

Currently, several unit root tests are used for testing for the presence of unit roots in the data. Examples of unit root tests, appropriately named after the researcher(s) who developed the underlying models, include the Augmented Dickey-Fuller (ADF) test, Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test and Phillips-Perrron (PP) test. Arguably the most commonly used test for detecting the presence of unit roots in applied econometrics is the ADF test (see Barr and Kahn (1995), Madsen (1997), Parikh and Kahn (1997) and Mainardi (2000)). This method allows for the analysis of slightly more complicated time series processes than the simple random walk models on which the original theory was developed and for which many believe is a sufficient representation of the underlining data generated process. The null hypothesis of this test statistic states that the series is non-stationary, whilst the alternative hypothesis test claims that the series is stationary. The popularity and ease of the ADF test for analyzing the stationarity of time series data has meant that the routine has been successfully integrated into several software econometric packages (for example EViews 5.1, RATS 4.02 and SAS). Given that this is the case, it was decided to use the ADF test as an appropriate method for analyzing the simulated data for this study.

To utilise the ADF test, the selection of the correct lag order of the autoregressive model is necessary. If one were to use a model which does not provide a reasonable representation of the data, the ADF test will not be well-estimated. Including too many lags in the model reduces the power of the test to reject the null hypothesis as the increased number of lags necessitates the estimation of additional parameters and a subsequent loss of degrees of freedom. As such, the presence of unnecessary lags will reduce the power of the ADF test to detect a unit root. The loss of power may be so severe that the test may indicate a unit root for some lag terms but not for the others. In practice, the true order of the autoregressive process is unknown, so the order of this process must be chosen by the researcher.

Prior to the analysis, there are two strategies for making this choice. The first strategy is to arbitrarily choose the order of the autoregressive process to be relatively large. The reason for this approach is that if the order is chosen to be too small, then the ensuring inference about the unit root is biased. This bias is discussed further in Schwert (1989) and readers who require more information are referred to his paper. If the order of the autoregressive process is chosen to be too large, it may cause deterioration in the finite-sample properties of the ADF test. This point was clearly illustrated by the simulation results reported in Phillips and Perron (1988).

The second strategy addresses this problem by using the data to estimate the order of the autoregressive process. This strategy uses either a general-to-specific approach or one or more of the information criterion methods proposed by Hannan and Quinn (1979), Schwarz (1978) and Akaike

(1969, 1973). Typically, when this approach is followed, it is assumed that the ADF test has the distribution tabulated by Dickey and Fuller (1979, 1981).

Dickey and Fuller (1979, 1981) developed several tests to determine whether a p^{th} order autoregressive (AR) process was stationary. Stationarity implies that the roots of the lag polynomial lies inside the unit circle and the series has a finite variance and constant mean. The null hypotheses of these tests state that the AR process contains one unit root, so the sum of the autoregressive coefficients equals one. In research published during 1989, Schwert describes the recent extensions of the Dickey-Fuller test procedure in an attempt to account for mixed ARIMA processes as well as pure AR processes in performing unit-root tests. It has been shown in Schwert (1989) that the tests for unit roots 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 an MA component, the distribution of the unit-root test statistics can be different from the distributions reported by Dickey and Fuller. Therefore, if economic time series models contain MA components, then tests for unit root that use the Dickey-Fuller (1979, 1981) test are inappropriate.

Dickey and Fuller (1979) considered three different single-equation regression models that can be used to test for the presence of a unit root:

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$$
$$\Delta y_t = a_0 + \gamma y_{t-1} + \varepsilon_t$$
$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \varepsilon_t$$

The difference between the three single-equation models is the presence of the deterministic elements a_0 and $a_2 t$. The first equation is a pure random walk model, the second equation adds an intercept or a drift variable, and the third equation includes both a drift term and a linear time trend variable. This study only considers the pure random walk models such that no intercept or trend variables are included. The parameter of interest in the equation is γ, if γ = 0, then the equations are non-stationary and have a unit root. This test involves estimating one (or more) of the equations above using ordinary least squares in order to obtain the estimated value of γ and the associated standard error. The *t*-test statistic value of the regression equation is calculated by dividing γ by the corresponding standard error. Comparing the resulting *t*-statistic with the appropriate values reported in the Dickey-Fuller's tables allows one to determine whether or not to reject the null hypothesis, γ = 0 (non-stationary series). The methodology is the same regardless of which of the three forms of the equations is estimated. However, the critical values of the *t*-statistics do depend on whether an intercept and/or time trend is included in the regression equation. Dickey and Fuller (1979) found that the critical values for γ = 0 were dependent on both the form of the regression equation and the sample size. As the sample size increases, the critical value of the *t*-statistic becomes more negative.

Dickey and Pantula (1987) proposed a sequential test procedure, one that allows testing for series containing more than one unit root. As an illustration, consider a series that has two unit roots. This series is analyzed by testing the null hypothesis of two unit roots against the alternative of one unit root. If the null hypothesis is rejected, we may then test the hypothesis of exactly one unit root against the alternative of no unit root. An example of this testing procedure is provided in Mills (1999, pg. 90).

A procedure for testing for the presence of a unit root in a general time series setting has also been proposed by Phillips (1987). This approach is a non-parametric technique with respect to nuisance parameters and therefore allows for a wide class of time series models in which there is a unit root. This procedure can be used for ARIMA models with heterogeneously as well as identically distributed innovations. The method seems to have significant advantages when there are moving average components in the time series and, offers an alternative to the Dickey-Fuller procedure. Moreover, Phillips and Perron (1988) extend this method to the cases where a drift and/or a linear trend are included in the specification. These extensions are important for practical applications, where the presence of a non-zero drift is common. In addition, in many cases and, particularly, with economic time series, the main competing alternative to the presence of a unit root is a deterministic linear time trend. It is therefore important that regression tests for unit roots allow for these possibilities.

In conclusion, the inclusion of unit root testing is provided as we use the ADF tests to determine the order of integration of the simulated data for our models.

3.2 Test of cointegration

In the last two decades the theory of cointegrated variables has provided an exciting and new approach to econometric modelling. The theory of cointegration provides econometricians with a practical procedure to analyze non-stationary time series data. These non-stationary time series are cointegrated if the linear combination of these series is stationary. These stationary linear combinations are referred to as the cointegrating equations and may be interpreted as the long-run equilibrium relationships among the variables. The purpose of the test of cointegration is to determine whether a group of non-stationary series are cointegrated.

There are several methods to test for cointegration in a multi-equation model. The initial procedure, developed by Engle and Granger, was referred to as the Engle-Granger two-step procedure. This was followed up by Johansen who determined the maximum likelihood approach using the trace of the error correction vector. Engle and Granger (1987) were the first to propose the theory of cointegration. Although the Engle-Granger two-step procedure is easy to implement, it have several important defects. The estimation of the long-run equilibrium regression using the Engle-Granger two-step procedure requires that one place one variable on the left-hand-side of the equation and uses the other variables as regressors. In practice, it is possible to find that one regression indicates that the variables are cointegrated, whereas reversing the order of the variables indicates no cointegration. This is an undesirable feature of the procedure because the test for cointegration should not change when the order of the variables are reversed. The problem is obviously compounded when using three or more variables since any of the variables can be selected as the dependent variable. Moreover, in tests using three or more variables, we know that there may be more than one cointegrating vector. The method has no systematic procedure for the separate estimation of the multiple cointegrating vectors. Another drawback of the Engle-Granger two-step procedure is that it relies on a two-step estimator and any error introduced in the first step is carried into the second step. Fortunately, several methods have been developed that avoid these problems.

The Johansen (1988) maximum likelihood estimator can estimate and test for the presence of multiple cointegrating vectors as an alternative to using the problematic two-step estimators. The Johansen's maximum likelihood estimator allows one to test the restricted versions of the cointegrating vector(s) and the speed of adjustment parameters. The limitations associated with the Engle and Granger two-step procedure and the fact that most econometric packages are now capable of running the matrix algorithms of the Johansen method means that this is the procedure we have followed in our analysis.

Johansen's method of testing for cointegration has become an essential tool for applied economists who model time series data (see Kouassi (1997), Parikh and Kahn (1997), Sinha (1998) and Leng (2002)). The phenomenal growth in software development has seen the method of Johansen being incorporated into econometric software packages such as Shazam, EViews and Cats for Rats. The availability of these packages has resulted in widespread application of cointegrated modelling.

The Johansen methodology uses the mathematics of the rank of a matrix to test for cointegration. The theory shows that cointegration can be tested as the hypothesis of a reduced rank of a regression coefficient matrix in a vector error correction model. There are two tests used by Johansen for testing the number of cointegrating relations. They are the trace statistic (λ_{trace}) and the maximal eigenvalue statistic (λ_{max}).

The two statistical tests proposed by Soren Johansen are shown below.

Trace statistic:

$$\lambda_{trace} = -T\sum_{i=r+1}^{n}\ln(1-\hat{\lambda}_i)$$

where $\hat{\lambda}_i$ = estimated values of the characteristic roots

T= the actual number of observations used

r = 0, 1, 2, …, n -2, n -1

n = number of characteristic roots

Maximal eigenvalue statistic:

$$\lambda_{max} = -T\ln(1-\hat{\lambda}_{r+1})$$

where $\hat{\lambda}_i$ = estimated values of the characteristic roots

T= the actual number of observations used

r = 0, 1, 2, …, n -2, n -1

n = number of characteristic roots

The λ_{trace} tests the null hypothesis that the number of distinct cointegrating vectors is less than or equal to *r* against a general alternative. In the λ_{trace}, the further the estimated characteristic roots ($\hat{\lambda}_i$) are from zero, the more negative the $\ln(1-\hat{\lambda}_i)$ and the larger the λ_{trace}. If the λ_{trace} is large, then the chances of rejecting the null hypothesis, that the number of distinct cointegrating vectors is less than or equal to *r*, increases. The λ_{max} tests the null hypothesis that the number of cointegrating vectors is *r* against the alternative of *r* + 1 cointegrating vectors. Again, if the estimated value of the characteristic roots ($\hat{\lambda}_i$) is close to zero in the λ_{max}, then the λ_{max} will be small. If the λ_{max} is small, then the chances of rejecting the null hypothesis of *r* cointegrating vectors will decrease.

To implement Johansen's test of cointegration, it is necessary to determine the number of lag terms in the vector autoregression model. In practice the lag order is unknown and needs to be chosen prior to conducting any statistical tests since these tests may suffer from serious size distortions if the lag order is not chosen appropriately (see for example Cheung and Lai (1993)). To this end, it has been proposed that information criteria such as Akaike's (1973) AIC, the Bayesian information criterion proposed by Schwarz's (1978) and the other information criteria be used to determine the number of lag terms in the cointegrated model. However, the optimum information criterion used for determining the number of lag terms in the cointegrated model is still debatable and as such is the focus of this study.

When performing Johansen's method of testing for cointegration, one needs to decide on the type of deterministic components to incorporate into the model. A model with a linear deterministic trend includes terms such as an intercept and/or a linear time trend, while a model without a deterministic trend may include an intercept or no intercept term and a model with a quadratic deterministic trend includes both an intercept and a linear time trend term.

Therefore, it is important for one to be aware of the different types of deterministic components so that the correct components are included into the model.

Johansen's method of testing for cointegration is based on several assumptions. Arguably the most important assumption is that all series are non-stationary and each of the series is integrated of a same order. In this study we only consider series which are integrated of order one, denoted I (1). The pre-dominant applied cointegration literature considers the case where the series is I (1), examples of applications can be found in Barr and Kahn (1995), Akinboade and Niedermeier (2002) and Leng (2002). The reason provided is that most economic variables are integrated of an order one. Several authors used the term cointegration to refer to the case in which variables are I (1). To test whether the series satisfies the assumption of I (1), the Augmented Dickey-Fuller test is often used.

Cheung and Lai (1993) were concerned with the performance of Johansen's likelihood ratio tests for cointegration in finite sample sizes. The likelihood ratio tests proposed by Johansen were derived from asymptotic results and statistical inferences in finite samples could be inappropriate. In particular, the critical values based on asymptotic distributions could be misleading. In the Cheung and Lai (1993) study, the finite sample critical values of the Johansen's likelihood ratio tests were assessed and the sensitivity of the likelihood ratio tests for the different lag specifications were examined. It was found that Johansen's tests of cointegration were biased toward finding cointegration more often than asymptotic theory suggests. Furthermore, the finite-sample bias increases as the dimension of the lag length increases.

In summary, Johansen's method of testing for cointegration was used to test the number of cointegrating relationships between the simulated time series equations and those that satisfied pre-determined conditions were analyzed further.

Chapter 4

Model Selection Criteria

A standard problem in time series analysis is the choice of an appropriate model to represent the data. This is a common problem when a statistical model contains many variables. According to Parzen (1982), statistical data modelling is a field of statistical reasoning that seeks to fit models to data without knowing what the "*true*" model is or might be.

Consequently, one seeks to learn the model and study the quality of the model by a process which is called statistical model identification or evaluation. In recent years, in the literature, the necessity of introducing the concept of model selection or model evaluation has been recognized. Sclove (1994) describes model selection as the choice of selecting the best model(s) from a set of models and the different type of models that one compares and selects can be characterized according to the number of lags, the different number of explanatory variables and other factors. Also, there is presently a great deal of interest in simple criteria represented by parsimony of parameters for choosing one of a set of competing models to describe a given data set. As discussed in Stone (1981), parsimony can take in to account a variety of attributes of the selected model. One such attribute is the cost of measuring the models that required implementing the model and a second attribute is the complexity of the selected model. The general principle is that for a given level of accuracy, a simpler or a more parsimonious model is preferable to a more complex one.

This study focuses on four well-known model selection criteria to determine the order of the model and each of these criteria is discussed in the literature that follows. The four criteria are Akaike's information criterion, Schwarz's information criterion, Hannan-Quinn's information criterion and Final Prediction Error. In this study, these criteria are used to analyze simulated data from a theoretical cointegrated model. The criterion which identifies the correct model most often is identified as the most appropriate criterion.

The four well-known information criteria that are used in this research follow a similar format to the general information criterion (GIC) and the formula of the GIC is illustrated below. The first term of the GIC measures the lack of fit of the model and the second term is a penalty function for the number of parameters in the model. The lack of fit of the model involves a measure of the lack of parsimony or complexity of the model. One of the issues that lead to model complexity is the number of parameters incorporated in the model.

$$\text{GIC} = -2\log(L_k) + P_k$$

where L_k is the likelihood value of the *k*-th model

P_k is the penalty for the *k*-th model

4.1 Akaike's information criterion

During the last three decades, Akaike's information criterion (AIC) has had an important impact in statistical model evaluation problems. He developed the information-theoretic, or the entropic AIC criterion for the identification of an optimal and parsimonious model in data analysis from a class of competing models which take model complexity into account. The literature review in Chapter 2 presented a few of the published references which use AIC for model selection. There are many other publications but some of the more recent include Akinboade and Niedermeier (2002), Botha and Apostolellis (2003) and Nwokoma and Olofin (2003).

The introduction of AIC furthered the recognition of the importance of good modelling statistics. As a result, many important statistical modelling techniques have been developed in various field of statistics, control theory, econometrics, engineering, psychometrics, and in many other fields (Bozdogan (1987)). Despite the accumulation of many successful results using AIC, and despite its extreme popularity and growing school of adherents, AIC has been almost universally accepted in some areas of statistics, whilst in other areas it is still unknown or misunderstood (Bozdogan (1987)).

The model selection strategy of AIC has the objective of selecting a model based on simply minimizing the Kullback-Leibler discrepancy between the unknown (true) and the approximating data based models. Finding of the true model can be very complex and may require a great amount of time, since the model may incorporate an infinite number of parameters. Therefore, obtaining a true model is not an ideal manner to represent the recorded data but rather allow for the best approximating model and that is what AIC does.

In this study, we use the IC formula

$$\text{AIC}(p) = \ln|\hat{\Sigma}| + \frac{2k^2 p}{T}$$

where	$k =$ the number of variables in the model
	$p =$ the number of lag terms in the model
	$T =$ the number of observations used
	$\ln|\hat{\Sigma}|=$ the estimated covariance matrix of the fitted multivariate model

taken from Lütkepohl (1985) and Gonzalo and Pitarakis (1998) and it consists of two measurement terms. The first term (i.e. $\ln|\hat{\Sigma}|$) measures the inaccuracy or poorness of fit of the model. The second term (i.e. $\frac{2k^2 p}{T}$)

measures the complexity or the penalty due to the increase of unreliability in the first term which depends upon the number of parameters used to fit the data.

Consequently, when there are several competing models the parameters within the models are estimated by the method of maximum likelihood and the values of the AIC are computed and compared to find a model with the minimum value of AIC. This approach is called the minimum AIC procedure and the model with the minimum AIC value is called the minimum AIC estimator and is chosen to be the best model. For us the best model is the one with least complexity, or equivalent, the highest information gain. In applying AIC, the emphasis is on comparing the goodness of fit of various models with an allowance made for parsimony.

4.2 Schwarz's information criterion

This model selection criterion is used when a true model exists and has a finite and small dimension that does not increase with sample size. This criterion does not receive any benefit from the theory of Kullback-Leibler discrepancy, but is derived based on a Bayesian viewpoint. The best fitting true model is chosen from the list of candidate models as the one that has the lowest Schwarz's information criterion (SIC) value.

Lütkepohl (1985) performed a comparison of several information criteria used for determining the order of a vector autoregressive process for different sample sizes. The result indicated that the Schwarz's information criterion estimated the order of an autoregressive process correctly most often and estimated correctly more often when the sample size increased. Lütkepohl suggested that the Schwarz's information criterion and the Hannan-Quinn's criterion were the most parsimonious criteria as these two criteria produced the smallest average squared forecasting error and estimated the order of an autoregressive process correctly most often.

Several studies have applied the criterion introduced by Schwarz (1978). The criterion developed by Schwarz is often referred to as SIC, Bayesian information criterion (BIC) or even Schwarz Bayesian criterion (SBC). A review of the literature illustrates that all three notations are in use.

In this study, we use the IC formula

$$\text{SIC}(p) = \ln|\hat{\Sigma}| + \frac{k^2 p \ln(T)}{T}$$

where k, p, T and $\ln|\hat{\Sigma}|$ are as previously defined

taken from Lütkepohl (1985) and Gonzalo and Pitarakis (1998).

4.3 Hannan-Quinn's information criterion

Hannan and Quinn (1979) provide a brief discussion on methods used for the determining the order of an autoregressive model. They realized that a method such as Shibata's information criterion was inconsistent in the estimation of the order of the autoregressive model.

Hannan and Quinn (1979) claimed that the best-known rule for estimating the true order of an autoregression was to make use of the method developed by Akaike (1969). They followed a similar estimation procedure where the method was strongly consistent for estimating the order of the autoregression. This method they called the Hannan-Quinn's information criterion (HQ) and it has been used in analysis by Lütkepohl (1985), Quinn (1980) and Gonzalo and Pitarakis (1998).

Lütkepohl (1985) illustrated in his analysis that the method developed by Hannan and Quinn was consistent in the estimation of the true order of an autoregressive process. This was established when performing comparison with other consistent criteria of various sample sizes. Lütkepohl suggested that the Schwarz's information criterion and Hannan-Quinn's information criterion were the best criteria when one was interested in forecasting (minimizing the mean square forecasting error) or estimating the order of a finite order vector autoregressive model.

Quinn (1980) extended the procedure developed by HQ to the larger dimension case. This larger dimension case was referred to as the multivariate autoregressive process. This procedure was developed in such a way that it was strongly consistent just as in the situation of a univariate autoregression. During the same period, Hannan (1980) extended the original work of HQ by determining the order of an autoregressive moving average process.

In this study, we use the IC formula

$$\text{HQ}\,(p) = \ln|\hat{\Sigma}| + \frac{2k^2 p \ln\ln T}{T}$$

where k, p, T and $\ln|\hat{\Sigma}|$ are as previously defined

taken from Lütkepohl (1985) and Gonzalo and Pitarakis (1998).

4.4 Final Prediction Error

Akaike (1969) provided a brief discussion on the practical use of the Final Prediction Error (FPE) in determining the order of an autoregressive model. The practical application of the FPE is to estimate the FPE of each autoregressive model within a prescribed sufficiently wide range of possible orders and to select the one that gives the minimum of the estimates. Akaike (1969) claimed that by seeking the minimum of FPE, we would be able to arrive at an autoregressive model of an order that did not have a significant bias and simultaneously did not have a large mean square prediction error.

In research published during 1969, Akaike performed a comparison of three types of predictors that were used for model selection. These predictors were the original minimum FPE, the modified version denoted by the minimizing $(FPE)^{1/4}$ and the FPE proposed by Anderson (1963) for the decision of the order of a Gaussian autoregressive process. These three predictors were compared based on various simulated time series models, the predictor that indicated the true model most often was the one selected. The results showed that for practical applications, the original procedure, minimum FPE, was the best procedure to use for model comparison.

Lütkepohl (1985) also compared several types of information criteria and found that the predictor FPE had a tendency to over-estimate the order of an autoregressive process. In addition, the criteria FPE, AIC and Shibata all had a tendency to obtain the same number of lag terms for large sample sizes.

In this study, we use the IC formula

$$\text{FPE}(p) = |\hat{\Sigma}|\left(\frac{T + pk^2 + 1}{T - pk^2 - 1}\right)^{k^2}$$

taken from Lütkepohl (1985) which has an equivalent minimum for

$$\ln \text{FPE}(p) = \ln|\hat{\Sigma}| + k^2 \ln\left(\frac{T + pk^2 + 1}{T - pk^2 - 1}\right)$$

where k, p, T and $\ln|\hat{\Sigma}|$ are as previously defined.

Chapter 5

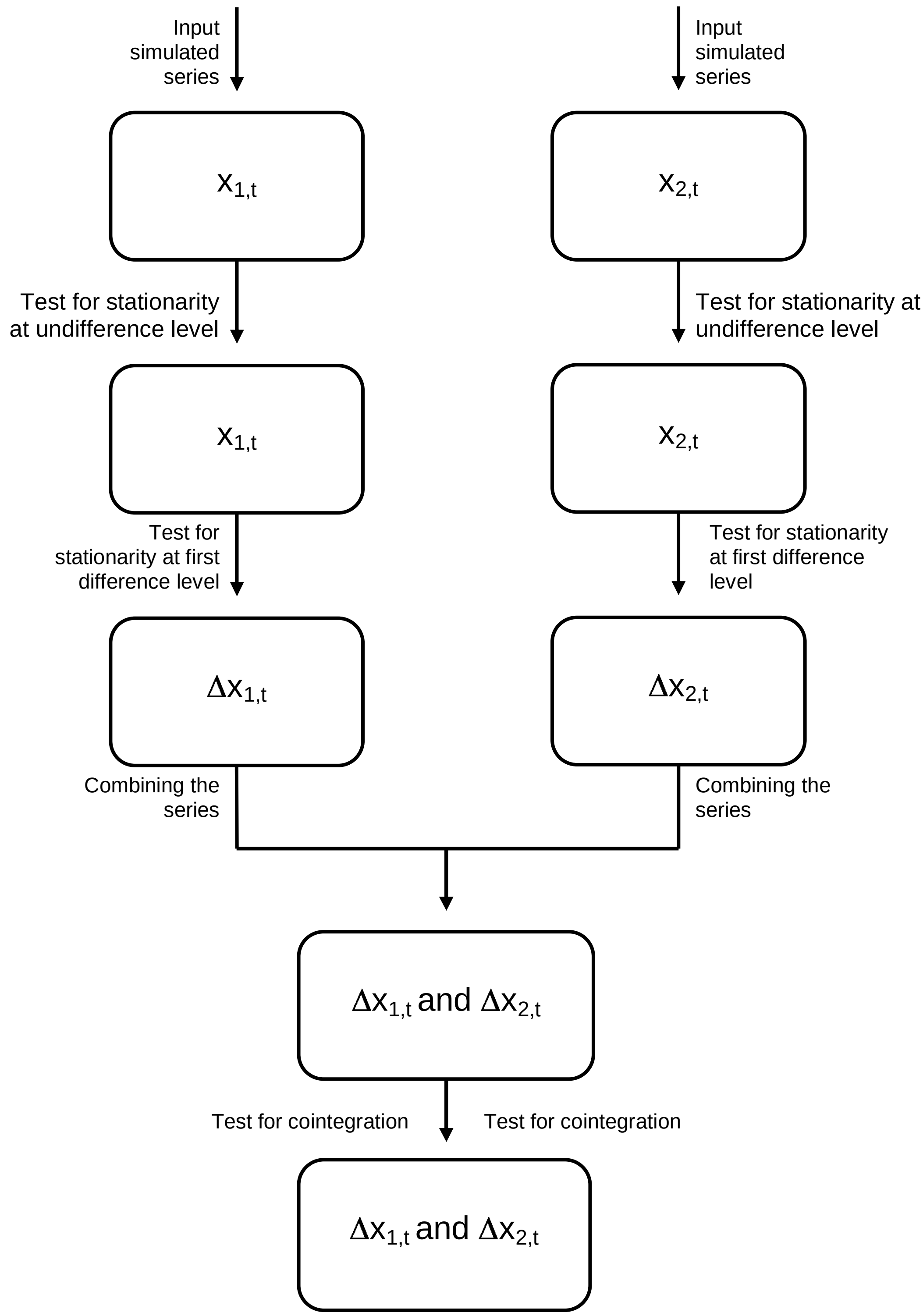
Methodology

In this study, we simulate several lag length models for our empirical analysis. We restricted our highest lag length term to four. For each of the lag length models, a sample size of 30, 60, 120 and 240 observations was simulated. The sample size selection followed the study of Khim and Liew (2004) where they use the same number of observations for ARMA models. To illustrate the methodology, we use a dataset of size 30 and a model with one lag term.

In this study, we simulated 15300 series for each of the models. These series were simulated using the random number generator in Excel and are recorded on the attached DVD. The simulated series were then exported into EViews 5.1 for analysis purposes.

An illustration of the procedure used to test whether the simulated bivariate time series equation is cointegrated is illustrated in Figure 5.1.

Figure 5.1: Analysis flow chart



5.1 Unit root tests

In this study, the ADF test is used to determine the order of integration of each series in the bivariate model. The order of integration is established by determining whether the series is stationary or non-stationary. If the series is non-stationary, the series is differenced and the differenced series is then tested to determine whether it is stationary or non-stationary. This sequence is repeated until all series are stationary. A series (for example $x_{1,t}$) that does not need to be differenced to achieve stationarity is called a series of integrated order zero, denoted I (0) and a series that is differenced once is called a series of integrated order one, denoted I (1).

This study only considers series of order one i.e. I (1) such that they are non-stationary series but difference stationary. Many of the methods used here can be extended to higher order cases, as an example a series that is I (2) or even higher. In our analysis we only retain the simulated data that meets the criteria I (1). If the criteria is not met, the dataset for both variables is discarded.

5.2 Test of cointegration

The Johansen method of testing for cointegration was used to analyze the linear relationship between these simulated series. For ease of exposition, this study applies the Trace statistic for the testing of cointegration as the results are similar to those of the Maximum Eigenvalue statistic. As an example, consider the series $x_{1,t}$ and $x_{2,t}$ analyzed in the next section. Both series $x_{1,t}$ and $x_{2,t}$ were I (1) and thus were tested for cointegration. The Johansen test of cointegration for the series $x_{1,t}$ and $x_{2,t}$ is shown in Table 5.2.

Table 5.2: Test of cointegration for the series $x_{1,t}$ and $x_{2,t}$			
Sample (adjusted): 2 30			
Included observations: 29 after adjustments			
Trend assumption: No deterministic trend			
Series: x_{1t} x_{2t}			
Lags interval (in first differences): No lags			
Unrestricted Cointegration Rank Test (Trace):			
Hypothesized	**Trace**	**Critical**	
Number of Cointegrating Equation (s)	**Statistic**	**Value**	**Prob.****
None *	31.57393	12.3209	0
At most 1	3.075338	4.129906	0.0941
Trace test indicates 1 cointegrating equation(s) at the 0.05 level			
* denotes rejection of the hypothesis at the 0.05 level			
**MacKinnon-Haug-Michelis (1999) *p*-values			

The result of the Trace statistic indicates that there is at most one cointegrating relationship (at the 5% significance level) between the series $x_{1,t}$ and $x_{2,t}$ i.e. for a 5% significance level test we would reject the null hypothesis lending support to the claim that the bivariate series has at most one cointegrating relations. In Table 5.2 the Johansen test of cointegration, the null hypothesis of this test indicates the series has no cointegrating relations against the alternative the series has one cointegrating relations between them.

5.3 Model selection criteria

In this study, we assessed the information criteria given in Chapter 4. To assess these criteria we need to calculate the value of each of these information criteria for each of the lag length models. Then evaluate the percentage that correctly estimated the true lag for the model.

For each of the information criteria, we estimate the determinant of the covariance matrix ($|\hat{\Sigma}|$) for the number of lag terms (p), the number of variables in the model (k) and the number of observations (T) used for estimating the fitted model. The function of the Akaike's information criterion is provided below:

$$\text{AIC}(p) = \ln|\hat{\Sigma}| + \frac{2k^2 p}{T}$$

In summary, the simulated time series are tested to ensure they are I (1) using the ADF test. The linear relationship between the bivariate time series equations are analyzed using the Johansen test of cointegration and only bivariate equations that have one cointegrated relationship between them were analyzed further in this study. The information criteria such as AIC, SIC, HQ and FPE estimate the lag length for each of the cointegrated models.

Chapter 6

The Theoretical and Simulation Models

The theoretical bivariate cointegrated models are presented in the forthcoming sections of this study. We restrict our study in the following way: we only use a two variable model, we exclude intercept and trend terms and limit the number of lag terms to four. The lag restriction is based on the methodology of Khim and Liew (2004). The variable restriction is by choice, the trivariate model is a study currently under investigation by another researcher. The exclusion of the intercept and trend terms is to simplify the analysis and should have little impact on the outcome of the study.

Initially, the cointegrated relationship for a lag one bivariate autoregressive model is illustrated. This is then followed by models of lag two, three and four. As discussed in Chapter 3, the theory of cointegration requires these lagged models to be non-stationary and integrated of order one, such that the first difference process $\Delta x_t = x_t - x_{t-1}$ must be stationary. These non-stationary time series models are cointegrated if the linear combination of these series is stationary. The linear combination of these series is called the cointegrating equation and is illustrated in the forthcoming section.

6.1 Autoregressive model of lag length one

The example of a two dimensional autoregressive model for a lag length of one is represented below. These autoregressive models depend on the previous changes of both variables $x_{1,t}$ and $x_{2,t}$ and the white-noise disturbances. Similar models can be found in the study of Cheung and Lai (1993) and are used illustratively in the text of Enders (2004). We define the bivariate model with one lag term as

$$x_{1,t} = a_{11}x_{1,t-1} + a_{12}x_{2,t-1} + \varepsilon_{1,t}$$

$$x_{2,t} = a_{21}x_{1,t-1} + a_{22}x_{2,t-1} + \varepsilon_{2,t}$$

where

a_{ij} denotes the coefficient of the ith equation for the jth variable (where i = 1, 2 and j = 1, 2)

$x_{j,t}$ denotes the jth variable at time period *t*

$\varepsilon_{i,t}$ denotes the error term of the ith equation at time period *t*.

Assuming the bivariate model is cointegrated and that each single equation is first order stationary, the model can be re-written as a cointegrated model by subtracting $x_{j,t-1}$ from both sides of each equation i.e.

$x_{1,t} = a_{11}x_{1,t-1} + a_{12}x_{2,t-1} + \varepsilon_{1,t}$ is re-written as

$$x_{1,t} - x_{1,t-1} = (a_{11}x_{1,t-1} - 1x_{1,t-1}) + (a_{12}x_{2,t-1} - 0x_{2,t-1}) + \varepsilon_{1,t}$$

and

$x_{2,t} = a_{21}x_{1,t-1} + a_{22}x_{2,t-1} + \varepsilon_{2,t}$ is re-written as

$$x_{2,t} - x_{2,t-1} = (a_{21}x_{1,t-1} - 0x_{1,t-1}) + (a_{22}x_{2,t-1} - 1x_{2,t-1}) + \varepsilon_{2,t} .$$

And thus for the cointegrated model we have

$$\Delta x_{1,t} = (a_{11}x_{1,t-1} - 1x_{1,t-1}) + (a_{12}x_{2,t-1} - 0x_{2,t-1}) + \varepsilon_{1,t}$$

$$\Delta x_{2,t} = (a_{21}x_{1,t-1} - 0x_{1,t-1}) + (a_{22}x_{2,t-1} - 1x_{2,t-1}) + \varepsilon_{2,t}$$

Re-written in matrix form

$$\begin{bmatrix} \Delta x_{1,t} \\ \Delta x_{2,t} \end{bmatrix} = \left[\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}.$$

This can be represented in matrix notation and simplified to

$$\Delta X_t = (A_1 - I)X_{t-1} + E_t \,.$$

Simplifying, the notation further we have

$$\begin{aligned} \Delta X_t &= (A_1 - I)X_{t-1} + E_t \\ &= -(I - A_1)X_{t-1} + E_t \\ &= \Pi X_{t-1} + E_t \end{aligned}$$

where $\Pi = -(I - A_1)$ and $A_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Our assumption that the bivariate model is cointegrated means that the coefficient matrix Π has rank *r*, where $0 < r < k$, with *k* denoting the number of variables in the multivariate model (Enders (2004) and Harris (1995)). The rank *r* is the number of cointegrating relations (or the cointegrating rank) between the variables.

An example of the two simulated series of lag length one is illustrated below. The simulated series consists of an autoregressive irregular component, $\sigma_{j,t}$. The idea for the inclusion of this component was taken from Enders (2004) with the objective of ensuring a "*pure*" irregular component.

The series $x_{1,t}$ and $x_{2,t}$ is constructed as

$$x_{1,t} = 1x_{1,t-1} + 0x_{2,t-1} + \sigma_{1,t} + \varepsilon_{1,t}$$
$$x_{2,t} = 0.5x_{1,t-1} + 0.7x_{2,t-1} + \sigma_{2,t} + \varepsilon_{2,t}$$

where $\sigma_{1,t} = 0.5\sigma_{1,t-1} + \mu_{1,t}$

$\sigma_{2,t} = 0.5\sigma_{2,t-1} + \mu_{2,t}$

where $x_{1,0} = 0$ and $x_{2,0} = 0$

$\sigma_{1,0} = 0$ and $\sigma_{2,0} = 0$

$\varepsilon_{1,t} \sim N(0,1)$ and $\varepsilon_{2,t} \sim N(0,1)$

$\mu_{1,t} \sim N(0,1)$ and $\mu_{2,t} \sim N(0,1)$.

Using the coefficients given, the equations are restricted to provide a cointegrated series with one cointegrating equation. This is seen by substituting the values into the error correction model and observing that the error correction parameter matrix, Π, has rank of one.

$$x_{1,t} - x_{1,t-1} = (1x_{1,t-1} - x_{1,t-1}) + 0x_{2,t-1} + \sigma_{1,t} + \varepsilon_{1,t}$$

$$\Delta x_{1,t} = (1-1)x_{1,t-1} + 0x_{2,t-1} + \sigma_{1,t} + \varepsilon_{1,t}$$

and

$$x_{2,t} - x_{2,t-1} = 0.5x_{1,t-1} + (0.7x_{2,t-1} - x_{2,t-1}) + \sigma_{2,t} + \varepsilon_{2,t}$$

$$\Delta x_{2,t} = 0.5x_{1,t-1} + (0.7 - 1.0)x_{2,t-1} + \sigma_{2,t} + \varepsilon_{2,t}$$

$$\begin{bmatrix} \Delta x_{1,t} \\ \Delta x_{2,t} \end{bmatrix} = \begin{bmatrix} (1-1) & 0 \\ 0.5 & (0.7-1.0) \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} \sigma_{1,t} \\ \sigma_{2,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}.$$

This can be written in matrix notation and simplified to

$$\Delta X_t = \begin{bmatrix} (1-1) & 0 \\ 0.5 & (0.7-1.0) \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} \sigma_{1,t} \\ \sigma_{2,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 \\ 0.5 & -0.3 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} \sigma_{1,t} \\ \sigma_{2,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

where $\Pi = \begin{bmatrix} 0 & 0 \\ 0.5 & -0.3 \end{bmatrix}.$

An example of the first 20 observations of the above bivariate model is shown in Table 6.1.

Table 6.1: Simulated data for the lag length one autoregressive time series equations				
	Observation	**x1,t**	**x2,t**	
	1	-21.69270671	-35.97325524	
	2	-21.320594	-35.33114314	
	3	-21.35507191	-34.32587148	
	4	-21.36299082	-34.85752817	
	5	-21.52136315	-35.4419631	
	6	-21.22716764	-34.69287283	
	7	-21.38725171	-33.59283254	
	8	-19.08886052	-35.00819907	
	9	-20.24699954	-34.05227178	
	10	-21.84365777	-35.19364626	
	11	-22.60585607	-36.67884141	
	12	-23.7244929	-38.68570684	
	13	-23.44721829	-36.77601806	
	14	-24.10946042	-37.00822383	
	15	-22.92577941	-38.4908229	
	16	-22.19188038	-38.77070955	
	17	-20.36361162	-36.18285526	
	18	-21.99904782	-35.63420184	
	19	-24.39075223	-37.58770944	
	20	-25.57779484	-39.97814559	

6.2 Autoregressive model of lag length two

The example of a two dimensional autoregressive model for a lag length of two is represented below. This model is similar to the lag one model except that an additional lag term is added. We define the bivariate model with two lag terms as

$$x_{1,t} = a_{11}x_{1,t-1} + a_{12}x_{2,t-1} + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + \varepsilon_{1,t}$$

$$x_{2,t} = a_{21}x_{1,t-1} + a_{22}x_{2,t-1} + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + \varepsilon_{2,t}$$

where a_{ij}, $x_{j,t}$ and $\varepsilon_{i,t}$ are as defined for autoregressive model of lag length one and b_{ij} denotes the coefficient of the ith equation for the jth variable (where i = 1, 2 and j = 1, 2 at time period t – 2).

As for the lag one autoregressive model, assuming the bivariate model is cointegrated and that each single equation is first order stationary, the model can be re-written as a cointegrated model by subtracting $x_{j,t-1}$ from both sides of each equation i.e.

$x_{1,t} = a_{11}x_{1,t-1} + a_{12}x_{2,t-1} + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + \varepsilon_{1,t}$ is re-written as

$$x_{1,t} - x_{1,t-1} = (a_{11}x_{1,t-1} - 1x_{1,t-1}) + (a_{12}x_{2,t-1} - 0x_{2,t-1}) + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + \varepsilon_{1,t}$$

and

$x_{2,t} = a_{21}x_{1,t-1} + a_{22}x_{2,t-1} + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + \varepsilon_{2,t}$ is re-written as

$$x_{2,t} - x_{2,t-1} = (a_{21}x_{1,t-1} - 0x_{1,t-1}) + (a_{22}x_{2,t-1} - 1x_{2,t-1}) + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + \varepsilon_{2,t}.$$

And thus for the cointegrated model we have

$$\Delta x_{1,t} = (a_{11}x_{1,t-1} - 1x_{1,t-1}) + (a_{12}x_{2,t-1} - 0x_{2,t-1}) + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + \varepsilon_{1,t}$$

$$\Delta x_{2,t} = (a_{21}x_{1,t-1} - 0x_{1,t-1}) + (a_{22}x_{2,t-1} - 1x_{2,t-1}) + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + \varepsilon_{2,t}$$

Re-written in matrix form

$$\begin{bmatrix} \Delta x_{1,t} \\ \Delta x_{2,t} \end{bmatrix} = \left[\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}.$$

This can be represented in matrix notation and simplified to

$$\Delta X_t = (A_1 - I)X_{t-1} + A_2 X_{t-2} + E_t .$$

Now, we add and subtract a term $A_2 X_{t-1}$ from the right-hand-side of the equation. This is then simplified as illustrated below.

$$\begin{aligned} \Delta X_t &= (A_1 - I)X_{t-1} + A_2 X_{t-2} + (A_2 X_{t-1} - A_2 X_{t-1}) + E_t \\ &= (A_1 + A_2 - I)X_{t-1} + A_2 X_{t-2} - A_2 X_{t-1} + E_t \\ &= -(I - A_1 - A_2)X_{t-1} - A_2(X_{t-1} - X_{t-2}) + E_t \\ &= -(I - A_1 - A_2)X_{t-1} - A_2 \Delta X_{t-1} + E_t \end{aligned}$$

where $A_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $A_2 = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$, $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $E_t = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$.

Simplifying, the notation further we have

$$\Delta X_t = \Pi X^{\;t1} + \Pi_1 \Delta X_{t-1} + E_t$$

where $\Pi = -(I - A_1 - A_2)$, $\Pi_1 = -A_2$.

As for the lag one model, the assumption of a cointegrated model implies that the rank of Π , denoted *r*, must have rank, 0 < *r* < *k*, where *k* denotes the number of variables in the model.

An example of the two simulated series of lag length two is illustrated below. The simulated series consists of no autoregressive irregular components and is constructed as

$$x_{1,t} = -0.7x_{1,t-1} + 0.45x_{2,t-1} + 0.3x_{1,t-2} + 0.3x_{2,t-2} + \varepsilon_{1,t}$$
$$x_{2,t} = 0x_{1,t-1} + 0.8x_{2,t-1} + 0x_{1,t-2} + 0.2x_{2,t-2} + \varepsilon_{2,t}$$

where $x_{1,0} = 0$ and $x_{2,0} = 0$

$\varepsilon_{1,t} \sim N(0,1)$ and $\varepsilon_{2,t} \sim N(0,1)$.

Using the coefficients given, the equations are restricted to provide a cointegrated series with one cointegrating equation. This is seen by substituting the values into the error correction model and observing that the error correction parameter matrix, Π , has rank of one.

$$x_{1,t} - x_{1,t-1} = -0.7x_{1,t-1} - x_{1,t-1} + 0.45x_{2,t-1} + 0.3x_{1,t-2} + 0.3x_{2,t-2} + \varepsilon_{1,t}$$

$$\Delta x_{1,t} = (-0.7 - 1.0)x_{1,t-1} + 0.45x_{2,t-1} + 0.3x_{1,t-2} + 0.3x_{2,t-2} + \varepsilon_{1,t}$$

and

$$x_{2,t} - x_{2,t-1} = 0x_{1,t-1} + 0.8x_{2,t-1} - x_{2,t-1} + 0x_{1,t-2} + 0.2x_{2,t-2} + \varepsilon_{2,t}$$

$$\Delta x_{2,t} = 0x_{1,t-1} + (0.8 - 1.0)x_{2,t-1} + 0x_{1,t-2} + 0.2x_{2,t-2} + \varepsilon_{2,t}$$

$$\begin{bmatrix} \Delta x_{1,t} \\ \Delta x_{2,t} \end{bmatrix} = -\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -0.7 & 0.45 \\ 0 & 0.8 \end{pmatrix} - \begin{pmatrix} 0.3 & 0.3 \\ 0 & 0.2 \end{pmatrix}\right]\begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} +$$

$$\begin{bmatrix} 0.3 & 0.3 \\ 0 & 0.2 \end{bmatrix}\begin{bmatrix} x_{1,t-1} - x_{1,t-2} \\ x_{2,t-1} - x_{2,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

$$= -\left[\begin{pmatrix} 1-(-0.7)-(0.3) & 0-(0.45)-(0.3) \\ 0-0-0 & 1-(0.8)-(0.2) \end{pmatrix}\right]\begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} +$$

$$\begin{bmatrix} 0.3 & 0.3 \\ 0 & 0.2 \end{bmatrix}\begin{bmatrix} x_{1,t-1} - x_{1,t-2} \\ x_{2,t-1} - x_{2,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

$$= -\begin{bmatrix} 1.4 & -0.75 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} 0.3 & 0.3 \\ 0 & 0.2 \end{bmatrix}\begin{bmatrix} x_{1,t-1} - x_{1,t-2} \\ x_{2,t-1} - x_{2,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

where $\Pi = \begin{bmatrix} 1.4 & -0.75 \\ 0 & 0 \end{bmatrix}$.

In the sections that follow, the two dimensional autoregressive models are extended to a lag three and a lag four model. This extension alters the structure of the model. However, the numbers of cointegrating relations is determined in the same manner as the lag one and lag two models.

6.3 Autoregressive model of lag length three

The example of a two dimensional autoregressive model for a lag length of three is represented below. This model is similar to the lag two models except that an additional lag term is added. We define the bivariate model with three lag terms as

$$x_{1,t} = a_{11}x_{1,t-1} + a_{12}x_{2,t-1} + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + c_{11}x_{1,t-3} + c_{12}x_{2,t-3} + \varepsilon_{1,t}$$

$$x_{2,t} = a_{21}x_{1,t-1} + a_{22}x_{2,t-1} + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + c_{21}x_{1,t-3} + c_{22}x_{2,t-3} + \varepsilon_{2,t}$$

where a_{ij}, b_{ij}, $x_{j,t}$ and $\varepsilon_{i,t}$ are as defined for autoregressive model of lag length two and c_{ij} denotes the coefficient of the i[th] equation for the j[th] variable (where i = 1, 2 and j = 1, 2 at time period t – 3).

As for the lag one autoregressive models, assuming the bivariate model is cointegrated and that each single equation is first order stationary, the model can be re-written as a cointegrated model by subtracting $x_{j,t-1}$ from both sides of each equation i.e.

$$x_{1,t} = a_{11}x_{1,t-1} + a_{12}x_{2,t-1} + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + c_{11}x_{1,t-3} + c_{12}x_{2,t-3} + \varepsilon_{1,t}$$

is re-written as

$$x_{1,t} - x_{1,} = (a_{11}x_{1,t-1} - 1x_{1,t-1}) + (a_{12}x_{2,t-1} - 0x_{2,t-1}) + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + c_{11}x_{1,t-3} + c_{12}x_{2,t-3} + \varepsilon_{1,t}$$

and

$$x_{2,t} = a_{21}x_{1,t-1} + a_{22}x_{2,t-1} + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + c_{21}x_{1,t-3} + c_{22}x_{2,t-3} + \varepsilon_{2,t}$$

is re-written as

$$x_{2,t} - x_{2,t-1} = (a_{21}x_{1,t-1} - 0x_{1,t-1}) + (a_{22}x_{2,t-1} - 1x_{2,t-1}) + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + c_{21}x_{1,t-3} + c_{22}x_{2,t-3} + \varepsilon_{2,t} .$$

And thus for the cointegrated model we have

$$\Delta x_{1,t} = (a_{11}x_{1,t-1} - 1x_{1,t-1}) + (a_{12}x_{2,t-1} - 0x_{2,t-1}) + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + c_{11}x_{1,t-3} + c_{12}x_{2,t-3} + \varepsilon_{1,t}$$

$$\Delta x_{2,t} = (a_{21}x_{1,t-1} - 0x_{1,t-1}) + (a_{22}x_{2,t-1} - 1x_{2,t-1}) + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + c_{21}x_{1,t-3} + c_{22}x_{2,t-3} + \varepsilon_{2,t}$$

Re-written in matrix form

$$\begin{bmatrix} \Delta x_{1,t} \\ \Delta x_{2,t} \end{bmatrix} = \left[\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} x_{1,t-3} \\ x_{2,t-3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}.$$

This can be represented in matrix notation and simplified to

$$\Delta X_t = (A_1 - I)X_{t-1} + A_2 X_{t-2} + A_3 X_{t-3} + E_t .$$

Now, we add and subtract the term $A_2 X_{t-1}$, $A_3 X_{t-1}$ and $A_3 X_{t-2}$ from the right-hand-side of the equation. This is then simplified as illustrated below.

$$\begin{aligned}
\Delta X_t &= (A_1 - I)X_{t-1} + A_2 X_{t-2} + A_3 X_{t-3} + E_t \\
&= (A_1 - I)X_{t-1} + A_2 X_{t-2} + A_3 X_{t-3} + (A_2 X_{t-1} - A_2 X_{t-1}) + (A_3 X_{t-1} - A_3 X_{t-1}) + E_t \\
&= (A_1 + A_2 + A_3 - I)X_{t-1} + A_2 X_{t-2} + A_3 X_{t-3} - A_2 X_{t-1} - A_3 X_{t-1} + E_t \\
&= -(I - A_1 - A_2 - A_3)X_{t-1} + (A_2 X_{t-2} - A_2 X_{t-1}) + (A_3 X_{t-3} - A_3 X_{t-1}) + E_t \\
&= -(I - A_1 - A_2 - A_3)X_{t-1} - (A_2 X_{t-1} - A_2 X_{t-2}) + (A_3 X_{t-3} - A_3 X_{t-1}) + E_t \\
&= -(I - A_1 - A_2 - A_3)X_{t-1} - A_2(\Delta X_{t-1}) + (A_3 X_{t-3} - A_3 X_{t-1}) + E_t
\end{aligned}$$

$$= -(I - A_1 - A_2 - A_3)X_{t-1} - A_2(\Delta X_{t-1}) - (A_3 X_{t-1}) + (A_3 X_{t-3}) +$$
$$(A_3 X_{t-2} - A_3 X_{t-2}) + E_t$$
$$= -(I - A_1 - A_2 - A_3)X_{t-1} - A_2(\Delta X_{t-1}) - (A_3 X_{t-1} - A_3 X_{t-2}) -$$
$$(A_3 X_{t-2} - A_3 X_{t-3}) + E_t$$
$$= -(I - A_1 - A_2 - A_3)X_{t-1} - A_2(\Delta X_{t-1}) - A_3(\Delta X_{t-1}) - A_3(\Delta X_{t-2}) + E_t$$
$$= -(I - A_1 - A_2 - A_3)X_{t-1} - (A_2 + A_3)\Delta X_{t-1} - A_3(\Delta X_{t-2}) + E_t$$

where $A_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $A_2 = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$, $A_3 = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$,

$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $E_t = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$.

Simplifying, the notation further we have

$$\Delta X_t = \Pi X_{t-1} - \Pi_1(\Delta X_{t-1}) - \Pi_2(\Delta X_{t-1}) - \Pi_2(\Delta X_{t-2}) + E_t$$
$$= \Pi X_{t-1} - (\Pi_1 + \Pi_2)\Delta X_{t-1} - \Pi_2(\Delta X_{t-2}) + E_t$$

where $\Pi = -(I - A_1 - A_2 - A_3)$, $\Pi_1 = A_2$, $\Pi_2 = A_3$.

As for the lag one model, the assumption of a cointegrated model implies that the rank of Π, denoted *r*, must have rank, 0 < *r* < *k*, where *k* denotes the number of variables in the model.

An example of the two simulated series of lag length three is illustrated below. The simulated series consists of no autoregressive irregular components and is constructed as

$$x_{1,t} = -0.7x_{1,t-1} + 0.4x_{2,t-1} + 0.2x_{1,t-2} + 0.2x_{2,t-2} + 0.1x_{1,t-3} + 0.15x_{2,t-3} + \varepsilon_{1,t}$$

$$x_{2,t} = 0x_{1,t-1} + 0.8x_{2,t-1} + 0x_{1,t-2} + 0.2x_{2,t-2} + 0x_{1,t-3} + 0x_{2,t-3} + \varepsilon_{2,t}$$

where $x_{1,0} = 0$ and $x_{2,0} = 0$

$\varepsilon_{1,t} \sim N(0,1)$ and $\varepsilon_{2,t} \sim N(0,1)$.

Using the coefficients given, the equations are restricted to provide a cointegrated series with one cointegrating equation. This is seen by substituting the values into the error correction model and observing that the error correction parameter matrix, Π , has rank of one.

$$x_{1,t} - x_{1,t-1} = -0.7x_{1,t-1} - x_{1,t-1} + 0.4x_{2,t-1} + 0.2x_{1,t-2} + 0.2x_{2,t-2} + 0.1x_{1,t-3} + 0.15x_{2,t-3} + \varepsilon_{1,t}$$

$$\Delta x_{1,t} = (-0.7 - 1.0)x_{1,t-1} + 0.4x_{2,t-1} + 0.2x_{1,t-2} + 0.2x_{2,t-2} + 0.1x_{1,t-3} + 0.15x_{2,t-3} + \varepsilon_{1,t}$$

and

$$x_{2,t} - x_{2,t-1} = 0x_{1,t-1} + 0.8x_{2,t-1} - x_{2,t-1} + 0x_{1,t-2} + 0.2x_{2,t-2} + 0x_{1,t-3} + 0x_{2,t-3} + \varepsilon_{2,t}$$

$$\Delta x_{2,t} = 0x_{1,t-1} + (0.8 - 1.0)x_{2,t-1} + 0x_{1,t-2} + 0.2x_{2,t-2} + 0x_{1,t-3} + 0x_{2,t-3} + \varepsilon_{2,t}$$

$$\begin{bmatrix} \Delta x_{1,t} \\ \Delta x_{2,t} \end{bmatrix} = -\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -0.7 & 0.4 \\ 0 & 0.8 \end{pmatrix} - \begin{pmatrix} 0.2 & 0.2 \\ 0 & 0.2 \end{pmatrix} - \begin{pmatrix} 0.1 & 0.15 \\ 0 & 0 \end{pmatrix} \right] \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} - \left[\begin{pmatrix} -0.7 & 0.4 \\ 0 & 0.8 \end{pmatrix} + \begin{pmatrix} 0.2 & 0.2 \\ 0 & 0.2 \end{pmatrix} \right] \begin{bmatrix} x_{1,t-1} - x_{1,t-2} \\ x_{2,t-1} - x_{2,t-2} \end{bmatrix} - \begin{bmatrix} 0.1 & 0.15 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{1,t-2} - x_{1,t-3} \\ x_{2,t-2} - x_{2,t-3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

$$= -\begin{bmatrix} 1-(-0.7)-(0.2)-(0.1) & -(0.4)-(0.2)-(0.15) \\ -0-0-0 & 1-(0.8)-(0.2)-0 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} -$$

$$\begin{bmatrix} (-0.7)+(0.2) & (0.4)+(0.2) \\ 0+0 & (0.8)+(0.2) \end{bmatrix} \begin{bmatrix} x_{1,t-1} - x_{1,t-2} \\ x_{2,t-1} - x_{2,t-2} \end{bmatrix} -$$

$$\begin{bmatrix} 0.1 & 0.15 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{1,t-2} - x_{1,t-3} \\ x_{2,t-2} - x_{2,t-3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

$$= -\begin{bmatrix} 1.4 & -0.75 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} - \begin{bmatrix} -0.5 & 0.6 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1,t-1} - x_{1,t-2} \\ x_{2,t-1} - x_{2,t-2} \end{bmatrix} -$$

$$\begin{bmatrix} 0.1 & 0.15 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{1,t-2} - x_{1,t-3} \\ x_{2,t-2} - x_{2,t-3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

where $\Pi = \begin{bmatrix} 1.4 & -0.75 \\ 0 & 0 \end{bmatrix}$.

6.4 Autoregressive model of lag length four

The example of a two dimensional autoregressive model for a lag length of four is represented below. This model is similar to the lag three models except that an additional lag term is added. We define the bivariate model with four lag terms as

$$x_{1,t} = a_{11}x_{1,t-1} + a_{12}x_{2,t-1} + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + c_{11}x_{1,t-3} + c_{12}x_{2,t-3} + d_{11}x_{1,t-4} + d_{12}x_{2,t-4} + \varepsilon_{1,t}$$

$$x_{2,t} = a_{21}x_{1,t-1} + a_{22}x_{2,t-1} + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + c_{21}x_{1,t-3} + c_{22}x_{2,t-3} + d_{21}x_{1,t-4} + d_{22}x_{2,t-4} + \varepsilon_{2,t}$$

where a_{ij}, b_{ij}, c_{ij}, $x_{j,t}$ and $\varepsilon_{i,t}$ are as defined for autoregressive model of lag length three and d_{ij} denotes the coefficient of the ith equation for the jth variable (where i = 1, 2 and j = 1, 2 at time period t – 4).

As for the lag one autoregressive models, assuming the bivariate model is cointegrated and that each single equation is first order stationary, the model can be re-written as a cointegrated model by subtracting $x_{j,t-1}$ from both sides of each equation i.e.

$$x_{1,t} = a_{11}x_{1,t-1} + a_{12}x_{2,t-1} + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + c_{11}x_{1,t-3} + c_{12}x_{2,t-3} + d_{11}x_{1,t-4} + d_{12}x_{2,t-4} + \varepsilon_{1,t}$$

is re-written as

$$x_{1,t} - x_{1,t-1} = (a_{11}x_{1,t-1} - 1x_{1,t-1}) + (a_{12}x_{2,t-1} - 0x_{2,t-1}) + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + c_{11}x_{1,t-3} + c_{12}x_{2,t-3} + d_{11}x_{1,t-4} + d_{12}x_{2,t-4} + \varepsilon_{1,t}$$

and

$$x_{2,t} = a_{21}x_{1,t-1} + a_{22}x_{2,t-1} + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + c_{21}x_{1,t-3} + c_{22}x_{2,t-3} + d_{21}x_{1,t-4} + d_{22}x_{2,t-4} + \varepsilon_{2,t}$$

is re-written as

$$x_{2,t} - x_{2,t-1} = (a_{21}x_{1,t-1} - 0x_{1,t-1}) + (a_{22}x_{2,t-1} - 1x_{2,t-1}) + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + c_{21}x_{1,t-3} + c_{22}x_{2,t-3} + d_{21}x_{1,t-4} + d_{22}x_{2,t-4} + \varepsilon_{2,t}\ .$$

And thus for the cointegrated model we have

$$\Delta x_{1,t} = (a_{11}x_{1,t-1} - 1x_{1,t-1}) + (a_{12}x_{2,t-1} - 0x_{2,t-1}) + b_{11}x_{1,t-2} + b_{12}x_{2,t-2} + c_{11}x_{1,t-3} + c_{12}x_{2,t-3} + d_{11}x_{1,t-4} + d_{12}x_{2,t-4} + \varepsilon_{1,t}$$

$$\Delta x_{2,t} = (a_{21}x_{1,t-1} - 0x_{1,t-1}) + (a_{22}x_{2,t-1} - 1x_{2,t-1}) + b_{21}x_{1,t-2} + b_{22}x_{2,t-2} + c_{21}x_{1,t-3} + c_{22}x_{2,t-3} + d_{21}x_{1,t-4} + d_{22}x_{2,t-4} + \varepsilon_{2,t}$$

Re-written in matrix form

$$\begin{bmatrix} \Delta x_{1,t} \\ \Delta x_{2,t} \end{bmatrix} = \left[\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]\begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}\begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}\begin{bmatrix} x_{1,t-3} \\ x_{2,t-3} \end{bmatrix} + \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}\begin{bmatrix} x_{1,t-4} \\ x_{2,t-4} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}.$$

This can be represented in matrix notation and simplified to

$$\Delta X_t = (A_1 - I)X_{t-1} + A_2X_{t-2} + A_3X_{t-3} + A_4X_{t-4} + E_t .$$

Now, we add and subtract the term A_2X_{t-1}, A_3X_{t-1}, A_4X_{t-1}, A_3X_{t-2}, A_4X_{t-2} and A_4X_{t-3} from the right-hand-side of the equation. This is then simplified as illustrated below.

$$\Delta X_t = (A_1 - I)X_{t-1} + A_2X_{t-2} + A_3X_{t-3} + A_4X_{t-4} + E_t$$

$$= (A_1 - I)X_{t-1} + A_2X_{t-2} + A_3X_{t-3} + A_4X_{t-4} + (A_2X_{t-1} - A_2X_{t-1}) + (A_3X_{t-1} - A_3X_{t-1}) + (A_4X_{t-1} - A_4X_{t-1}) + E_t$$

$$= (A_1 + A_2 + A_3 + A_4 - I)X_{t-1} + (A_2X_{t-2} - A_2X_{t-1}) + (A_3X_{t-3} - A_3X_{t-1}) + (A_4X_{t-4} - A_4X_{t-1}) + E_t$$

$$= -(I - A_1 - A_2 - A_3 - A_4)X_{t-1} - (A_2X_{t-1} - A_2X_{t-2}) - (A_3X_{t-1} - A_3X_{t-3}) - (A_4X_{t-1} - A_4X_{t-4}) + E_t$$

$$= -(I - A_1 - A_2 - A_3 - A_4)X_{t-1} - A_2(\Delta X_{t-1}) - A_3X_{t-1} + A_3X_{t-3} - A_4X_{t-1} + A_4X_{t-4} + (A_3X_{t-2} - A_3X_{t-2}) + (A_4X_{t-2} - A_4X_{t-2}) + (A_4X_{t-3} - A_4X_{t-3}) + E_t$$

$$= -(I - A_1 - A_2 - A_3 - A_4)X_{t-1} - A_2(\Delta X_{t-1}) - (A_3X_{t-1} - A_3X_{t-2}) - (A_3X_{t-2} - A_3X_{t-3}) - (A_4X_{t-1} - A_4X_{t-2}) - (A_4X_{t-2} - A_4X_{t-3}) - (A_4X_{t-3} - A_4X_{t-4}) + E_t$$

$$= -(I - A_1 - A_2 - A_3 - A_4)X_{t-1} - A_2(\Delta X_{t-1}) - A_3(\Delta X_{t-1}) - A_3(\Delta X_{t-2}) - A_4(\Delta X_{t-1}) - A_4(\Delta X_{t-2}) - A_4(\Delta X_{t-3}) + E_t$$

$$= -(I - A_1 - A_2 - A_3 - A_4)X_{t-1} - (A_2 + A_3 + A_4)\Delta X_{t-1} - (A_3 + A_4)\Delta X_{t-2} - A_4(\Delta X_{t-3}) + E_t$$

where $A_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $A_2 = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$, $A_3 = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$,

$A_4 = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}$, $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $E_t = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$.

Simplifying, the notation further we have

$$\Delta X_t = \Pi X_{t-1} - (\Pi_1 + \Pi_2 + \Pi_3)\Delta X_{t-1} - (\Pi_2 + \Pi_3)\Delta X_{t-2} - \Pi_3(\Delta X_{t-3}) + E_t$$

where $\Pi = -(I - A_1 - A_2 - A_3 - A_4)$, $\Pi_1 = A_2$, $\Pi_2 = A_3$, $\Pi_3 = A_4$.

As for the lag one model, the assumption of a cointegrated model implies that the rank of Π, denoted *r*, must have rank, $0 < r < k$, where *k* denotes the number of variables in the model.

An example of the two simulated series of lag length four is illustrated below. The simulated series consists of no autoregressive irregular components and we define the series $x_{1,t}$ and $x_{2,t}$ as

$$x_{1,t} = -0.5x_{1,t-1} + 0.35x_{2,t-1} + 0.25x_{1,t-2} + 0.2x_{2,t-2} + 0.15x_{1,t-3} + 0.15x_{2,t-3} +$$
$$0.1x_{1,t-4} + 0.1x_{2,t-4} + \varepsilon_{1,t}$$
$$x_{2,t} = 0x_{1,t-1} + 0.8x_{2,t-1} + 0x_{1,t-2} + 0.2x_{2,t-2} + 0x_{1,t-3} + 0x_{2,t-3} + 0x_{1,t-4} + 0x_{2,t-4} + \varepsilon_{2,t}$$

where $x_{1,0} = 0$ and $x_{2,0} = 0$

$\varepsilon_{1,t} \sim N(0,1)$ and $\varepsilon_{2,t} \sim N(0,1)$.

Using the coefficients given, the equations are restricted to provide a cointegrated series with one cointegrating equation. This is seen by substituting the values into the error correction model and observing that the error correction parameter matrix, Π , has rank of one.

$$x_{1,t} - x_{1,t-1} = -0.5x_{1,t-1} - x_{1,t-1} + 0.35x_{2,t-1} + 0.25x_{1,t-2} + 0.2x_{2,t-2} + 0.15x_{1,t-3} +$$
$$0.15x_{2,t-3} + 0.1x_{1,t-4} + 0.1x_{2,t-4} + \varepsilon_{1,t}$$

$$\Delta x_{1,t} = (-0.5 - 1.0)x_{1,t-1} + 0.35x_{2,t-1} + 0.25x_{1,t-2} + 0.2x_{2,t-2} + 0.15x_{1,t-3} + 0.15x_{2,t-3} +$$
$$0.1x_{1,t-4} + 0.1x_{2,t-4} + \varepsilon_{1,t}$$

and

$$x_{2,t} - x_{2,t-1} = 0x_{1,t-1} + 0.8x_{2,t-1} - x_{2,t-1} + 0x_{1,t-2} + 0.2x_{2,t-2} + 0x_{1,t-3} + 0x_{2,t-3} +$$
$$0x_{1,t-4} + 0x_{2,t-4} + \varepsilon_{2,t}$$

$$\Delta x_{2,t} = 0x_{1,t-1} + (0.8 - 1.0)x_{2,t-1} + 0x_{1,t-2} + 0.2x_{2,t-2} + 0x_{1,t-3} + 0x_{2,t-3} + 0x_{1,t-4} +$$

$0x_{2,t-4}+\varepsilon_{2,t}$

$$\begin{bmatrix}\Delta x_{1,t}\\ \Delta x_{2,t}\end{bmatrix}=-\left[\begin{pmatrix}1&0\\0&1\end{pmatrix}-\begin{pmatrix}-0.5&0.35\\0&0.8\end{pmatrix}-\begin{pmatrix}0.25&0.2\\0&0.2\end{pmatrix}-\begin{pmatrix}0.15&0.15\\0&0\end{pmatrix}-\begin{pmatrix}0.1&0.1\\0&0\end{pmatrix}\right]\begin{bmatrix}x_{1,t-1}\\x_{2,t-1}\end{bmatrix}-$$
$$\left[\begin{pmatrix}0.25&0.2\\0&0.2\end{pmatrix}+\begin{pmatrix}0.15&0.15\\0&0\end{pmatrix}+\begin{pmatrix}0.1&0.1\\0&0\end{pmatrix}\right]\begin{bmatrix}x_{1,t-1}-x_{1,t-2}\\x_{2,t-1}-x_{2,t-2}\end{bmatrix}-$$
$$\left[\begin{pmatrix}0.15&0.15\\0&0\end{pmatrix}+\begin{pmatrix}0.1&0.1\\0&0\end{pmatrix}\right]\begin{bmatrix}x_{1,t-2}-x_{1,t-3}\\x_{2,t-2}-x_{2,t-3}\end{bmatrix}-$$
$$\begin{bmatrix}0.1&0.1\\0&0\end{bmatrix}\begin{bmatrix}x_{1,t-3}-x_{1,t-4}\\x_{2,t-3}-x_{2,t-4}\end{bmatrix}+\begin{bmatrix}\varepsilon_{1,t}\\\varepsilon_{2,t}\end{bmatrix}$$

$$=-\begin{bmatrix}1-(-0.5)-(0.25)-(0.15)-(0.1)&-(0.35)-(0.2)-(0.15)-(0.1)\\0-0-0-0&1-(0.8)-(0.2)-0-0\end{bmatrix}\begin{bmatrix}x_{1,t-1}\\x_{2,t-1}\end{bmatrix}-$$
$$\begin{bmatrix}(0.25)+(0.15)+(0.1)&(0.2)+(0.15)+(0.1)\\0+0+0&(0.2)+0+0\end{bmatrix}\begin{bmatrix}x_{1,t-1}-x_{1,t-2}\\x_{2,t-1}-x_{2,t-2}\end{bmatrix}-$$
$$\begin{bmatrix}(0.15)+(0.1)&(0.15)+(0.1)\\0+0&0+0\end{bmatrix}\begin{bmatrix}x_{1,t-2}-x_{1,t-3}\\x_{2,t-2}-x_{2,t-3}\end{bmatrix}-$$
$$\begin{bmatrix}0.1&0.1\\0&0\end{bmatrix}\begin{bmatrix}x_{1,t-3}-x_{1,t-4}\\x_{2,t-3}-x_{2,t-4}\end{bmatrix}+\begin{bmatrix}\varepsilon_{1,t}\\\varepsilon_{2,t}\end{bmatrix}$$

$$=-\begin{bmatrix}1&-0.8\\0&0\end{bmatrix}\begin{bmatrix}x_{1,t-1}\\x_{2,t-1}\end{bmatrix}-\begin{bmatrix}0.5&0.45\\0&0.2\end{bmatrix}\begin{bmatrix}x_{1,t-1}-x_{1,t-2}\\x_{2,t-1}-x_{2,t-2}\end{bmatrix}-\begin{bmatrix}0.25&0.25\\0&0\end{bmatrix}\begin{bmatrix}x_{1,t-2}-x_{1,t-3}\\x_{2,t-2}-x_{2,t-3}\end{bmatrix}-$$
$$\begin{bmatrix}0.1&0.1\\0&0\end{bmatrix}\begin{bmatrix}x_{1,t-3}-x_{1,t-4}\\x_{2,t-3}-x_{2,t-4}\end{bmatrix}+\begin{bmatrix}\varepsilon_{1,t}\\\varepsilon_{2,t}\end{bmatrix}$$

where $\Pi=\begin{bmatrix}1&-0.8\\0&0\end{bmatrix}$.

In summary, the four models shown here are often referred to as Vector Autoregressive (VAR) processes with lags of one, two, three and four. Re-written as stationary time series models with an error correction term, which is denoted as ΠX_{t-1}, they are referred to as error correction models with zero, one, two and three lag terms, respectively. It is worth noting that all models have a zero intercept term and zero trend term.

In conclusion, the parameterization of the four models ensure that the error correction matrix, Π, has been restricted to ensure cointegration of the models with one cointegrating relationship.

Chapter 7

Results

The objective of this study was to determine which criterion most correctly estimated the true lag for the model. This is determined by computing the percentage of selection for each lag length for each of these criteria and then observing which criteria is correct most often. Originally, a sample size of 30, 60, 120 and 240 was used to test for unit roots and cointegration. For determining which criteria is correct most often, we reduced the number of observations by four in all sample sizes (i.e. 26, 56, 116 and 236) such that the same number of observations is used for each lag length model for each of these criteria.

7.1 Cointegrated model with a lag length of zero

The percentage of various criteria correctly estimating the true lag, lag zero, for the four sample sizes 26, 56, 116 and 236 are tabulated in Tables 7.1.1 to 7.1.4. The values that are highlighted in the tables are the highest selection percentage for each of the IC.

Table 7.1.1: Simulated cointegrated model of lag length 0 for a sample size of 26

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	63.38	90.99	74.85	99.94
1	21.83	8.12	17.96	0.06
2	7.46	0.67	4.14	0
3	7.33	0.22	3.04	0

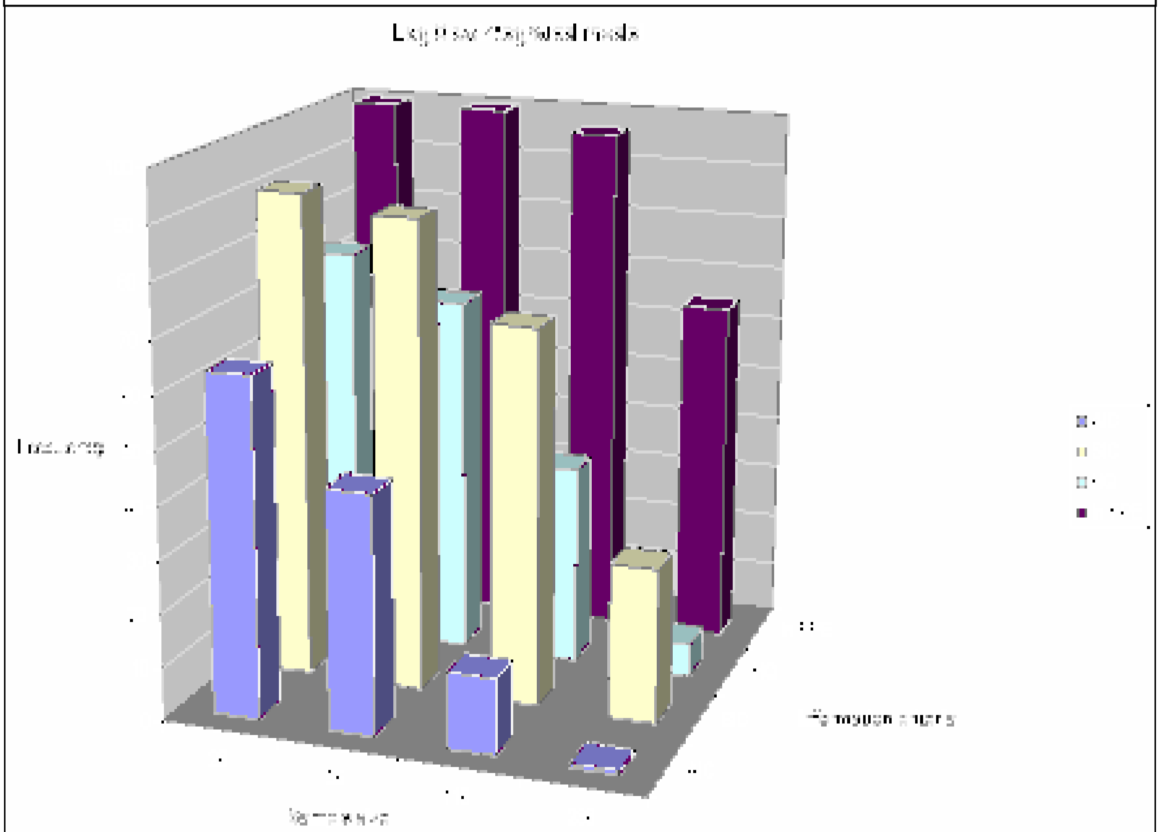
Table 7.1.2: Simulated cointegrated model of lag length 0 for a sample size of 56				
	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	44.58	88.04	67.38	99.85
1	43.35	11.77	30.16	0.15
2	8.36	0.17	2.09	0
3	3.7	0.02	0.37	0

Table 7.1.3: Simulated cointegrated model of lag length 0 for a sample size of 116				
	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	14.08	70.3	37.61	96.51
1	70.85	29.63	60.16	3.49
2	11.29	0.06	2.1	0
3	3.78	0	0.13	0

Table 7.1.4: Simulated cointegrated model of lag length 0 for a sample size of 236				
	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	0.8	28.78	6.33	64.67
1	78.39	71.05	90.57	35.32
2	16.54	0.17	2.98	0.01
3	4.26	0	0.12	0

The results in Table 7.1.1 show that the IC's perform well in estimating the true lag when a sample size is small (for T = 26). In particular, FPE selects the correct lag almost perfectly, whilst SIC selects the correct lag just less than 91% of the time. The HQ and AIC selections for small samples perform indifferently, 75% and 63%, respectively. As the sample size increases, an unusual phenomenon is observed. The correct selection percentages decreases, a finding in contradiction to other studies (see Gonzalo and Pitarakis (1998)). As an example, for a sample size of 56 (in Table 7.1.2), the correct selection percentages for AIC, SIC, HQ and FPE were 45%, 88%, 67% and 100% respectively, whilst for a sample size of 116 (in Table 7.1.3), the selection percentages for the four criteria are 14%, 70%, 38% and 97%. The unusual phenomenon may be caused by the autoregressive irregular variable that was incorporated in the model. Surprisingly, this phenomenon disappears for higher order lag models, a result that was originally expected.

Figure 7.1: Performances of IC's for correctly estimating the lag of the model



In Figure 7.1, the ability of the IC's to correctly select the lag zero cointegrated model is illustrated as the sample size increases. The graph illustrates that FPE performs best, followed by SIC. Of the four IC, AIC performs poorly, in particular when the sample size is large. The correct selection is a miserly 1%, which is exceptionally poor by any standards.

In summary, for the lag zero cointegrated model, i.e. the VAR (1) model, FPE outperforms AIC, SIC and HQ. This is consistent across all sample sizes, despite the somewhat disappointing performance for simulation data with T = 236.

7.2 Cointegrated model with a lag length of one

The percentage of various criteria correctly estimating the true lag, lag one, for the four sample sizes are tabulated in Tables 7.2.1 to 7.2.4. The values that are highlighted in the table are the highest selection percentage for each of the IC.

Table 7.2.1: Simulated cointegrated model of lag length 1 for a sample size of 26

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	16.41	46.86	24.98	98.7
1	60.75	49.89	61.49	1.3
2	13.77	2.61	9.04	0
3	9.07	0.64	4.5	0

Table 7.2.2: Simulated cointegrated model of lag length 1 for a sample size of 56

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	0.78	12.6	3.06	73.58
1	82.39	86.83	92.08	26.42
2	11.92	0.55	4.13	0
3	4.91	0.03	0.72	0

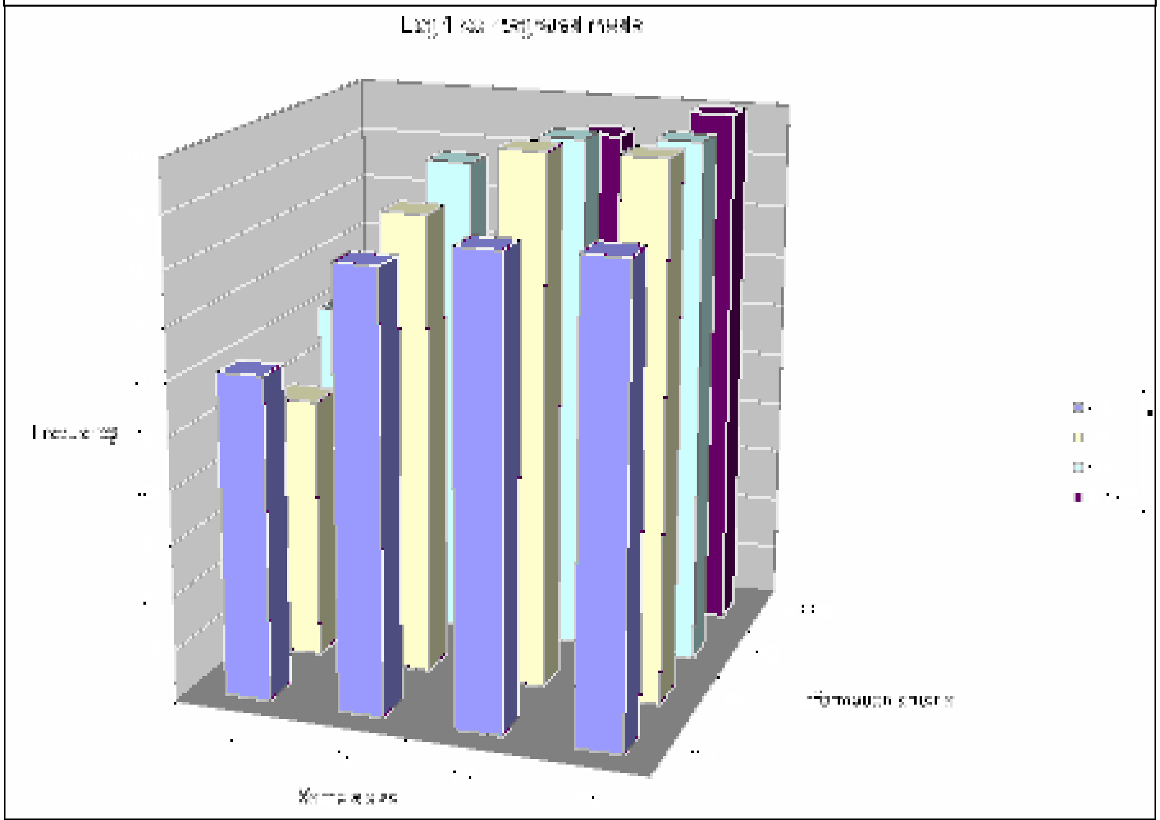
Table 7.2.3: Simulated cointegrated model of lag length 1 for a sample size of 116				
	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**In FPE**
0	0	0.16	0	5.36
1	87.12	99.72	97.96	94.64
2	9.44	0.12	1.85	0
3	3.44	0	0.18	0

Table 7.2.4: Simulated cointegrated model of lag length 1 for a sample size of 236				
	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**In FPE**
0	0	0	0	0
1	87.72	99.97	98.74	100
2	9.18	0.03	1.18	0
3	3.1	0	0.09	0

The results in Tables 7.2.1 to 7.2.4 show that the IC's perform well for large samples ($T \geq 116$), with three of the four IC's scoring above 90% and the fourth, AIC, scoring 87%. When sample size is small (for $T = 26$), the IC's performances are poor. AIC and HQ score approximately 60% with the other two scoring less.

The phenomenon observed, for the lag zero cointegrated model, where selection improves as sample size decreases is not evident for the lag one cointegrated model. The results for this model follow prior studies and lend support to the concern that the lag zero cointegrated model's results are unusual and bear further investigation. This is well illustrated in Figure 7.2, where one sees that selection improves considerably as sample size increases. Of the four IC's compared, HQ is the preferred model selector for smaller samples and SIC is the preferred model selector for larger samples.

Figure 7.2: Performances of IC's for correctly estimating the lag of the model



7.3 Cointegrated model with a lag length of two

The percentage of various criteria correctly estimating the true lag, lag two, for the four sample sizes are tabulated in Tables 7.3.1 to 7.3.4. The values that are highlighted in the table are the highest selection percentage for each of the IC.

Table 7.3.1: Simulated cointegrated model of lag length 2 for a sample size of 26				
	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	27.44	67.55	39.96	99.84
1	37.94	27.21	38.42	0.16
2	23.05	4.56	16.21	0
3	11.57	0.68	5.41	0

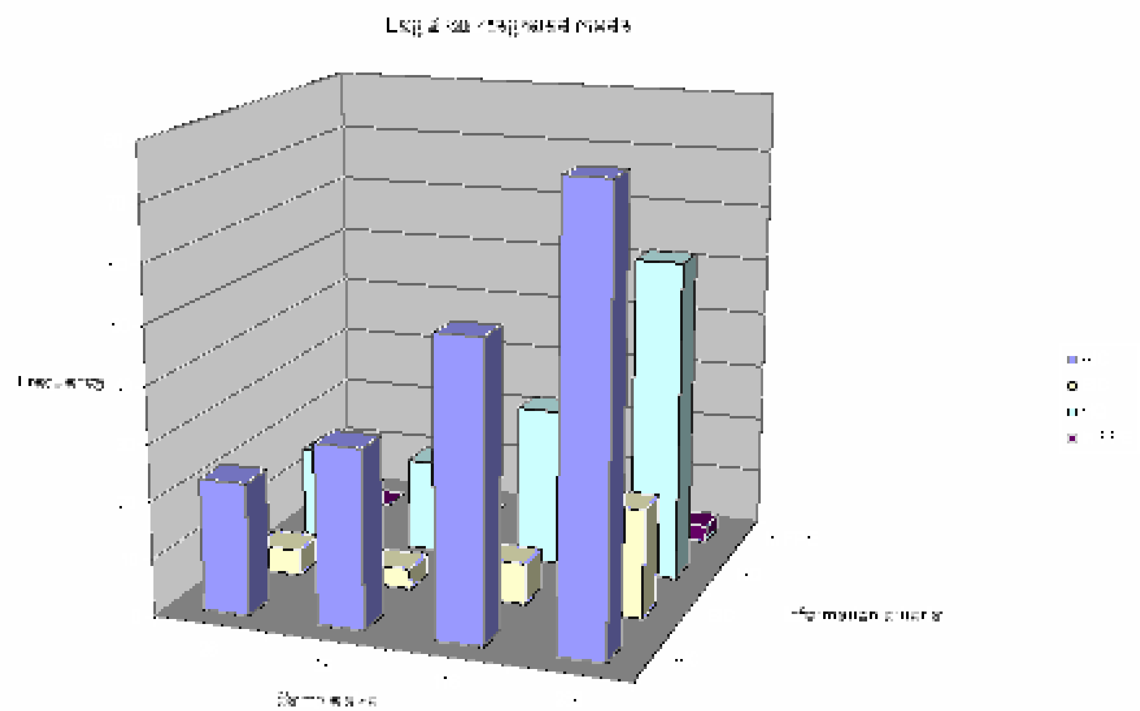
Table 7.3.2: Simulated cointegrated model of lag length 2 for a sample size of 56				
	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	4.35	41.14	15.15	94.62
1	57.01	55.22	67.2	5.38
2	31.27	3.57	16.25	0
3	7.37	0.07	1.39	0

Table 7.3.3: Simulated cointegrated model of lag length 2 for a sample size of 116				
	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	0.01	5	0.54	40.81
1	40.45	87.81	70.62	59.06
2	51.53	7.16	28.05	0.13
3	8.01	0.04	0.79	0

Table 7.3.4: Simulated cointegrated model of lag length 2 for a sample size of 236				
	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	0	0.01	0	0.25
1	13.04	80.88	43.09	96.96
2	77.86	19.11	56.21	2.8
3	9.1	0.01	0.71	0

The results in Tables 7.3.1 to 7.3.4 show that the IC perform poorly in estimating the true lag when a sample size is small (T = 26). AIC and HQ score 23% and 16%, respectively, while the other two criteria score less. When the sample size increases, the correct selection percentage of the various criteria increased. With a sample size of 56 (in Table 7.3.2), the selection percentages for AIC, SIC, HQ and FPE are 31%, 4% 16% and 0%, respectively, whilst for a sample size of 236 (in Table 7.3.4), the percentages for the same criteria reach scores of 78%, 19%, 56% and 2.8%, respectively. This improvement in selection as T increases is illustrated in Figure 7.3.

Figure 7.3: Performances of IC's for correctly estimating the lag of the model



In Figure 7.3, it is observed that AIC and HQ selects the true lag more often than SIC and FPE when the sample size increases. The reason that SIC and FPE select the lag two cointegrated model less often is because they under-estimate the true lag for the model, i.e. they select a lag of zero or a lag of one more often. For example, for T = 236 (in Table 7.3.4), SIC and FPE selects the lag one cointegrated model to be the correct model for the data rather than the lag two cointegrated model. This trend is noticeable for all sample sizes.

In summary, for the lag two cointegrated model, i.e. the VAR (3) model, AIC outperforms SIC, HQ and FPE. This trend is consistent across all sample sizes and is clearly illustrated in Figure 7.3.

7.4 Cointegrated model with a lag length of three

The percentage of various criteria correctly estimating the true lag, lag three, for the four sample sizes are tabulated in Tables 7.4.1 to 7.4.4. The values that are highlighted in the table are the highest selection percentages for each of the IC.

Table 7.4.1: Simulated cointegrated model of lag length 3 for a sample size of 26

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**In FPE**
0	6.56	35.09	13.32	97.27
1	41.27	51.97	49.1	2.73
2	25.98	9.73	21.77	0
3	26.19	3.21	15.8	0

Table 7.4.2: Simulated cointegrated model of lag length 3 for a sample size of 56

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**In FPE**
0	0.06	4.12	0.61	49.12
1	37.76	84.57	63.52	50.83
2	41.03	10.68	29.2	0.05
3	21.15	0.62	6.67	0

Table 7.4.3: Simulated cointegrated model of lag length 3 for a sample size of 116				
	Model selection criteria			
Lag	AIC	SIC	HQ	ln FPE
0	0	0	0	0.49
1	11.76	77.07	39.06	98.34
2	51.28	21.96	49.84	1.17
3	36.96	0.97	11.09	0

Table 7.4.4: Simulated cointegrated model of lag length 3 for a sample size of 236				
	Model selection criteria			
Lag	AIC	SIC	HQ	ln FPE
0	0	0	0	0
1	0.52	40.4	7.03	80.89
2	37.95	56.07	67.13	19.06
3	61.53	3.53	25.84	0.05

The results in Tables 7.4.1 to 7.4.4 show that the IC's perform poorly when the sample size is small but improve as the sample size increases. In a medium sample (for T = 116), the AIC performance is better than the other criteria and as the sample size increases from 26 to 116, AIC improves in ability for selection of the correct lag model whilst little improvement in the other three criteria is observed. In a large sample (for T = 236), AIC again outperforms the other criteria, scoring 62% whilst SIC, HQ and FPE score 4%, 26% and 0%, respectively. SIC, HQ and FPE selections indicate low order lag length models such as 56% and 67% for a lag length of two model and 81% for a lag length of one model. This indicates that SIC, HQ and FPE under-estimate the true lag for the model and are less useful than AIC in selecting the correct lag length of higher order models. This is illustrated in Figure 7.4.

Figure 7.4: Performances of IC’s for correctly estimating the lag of the model

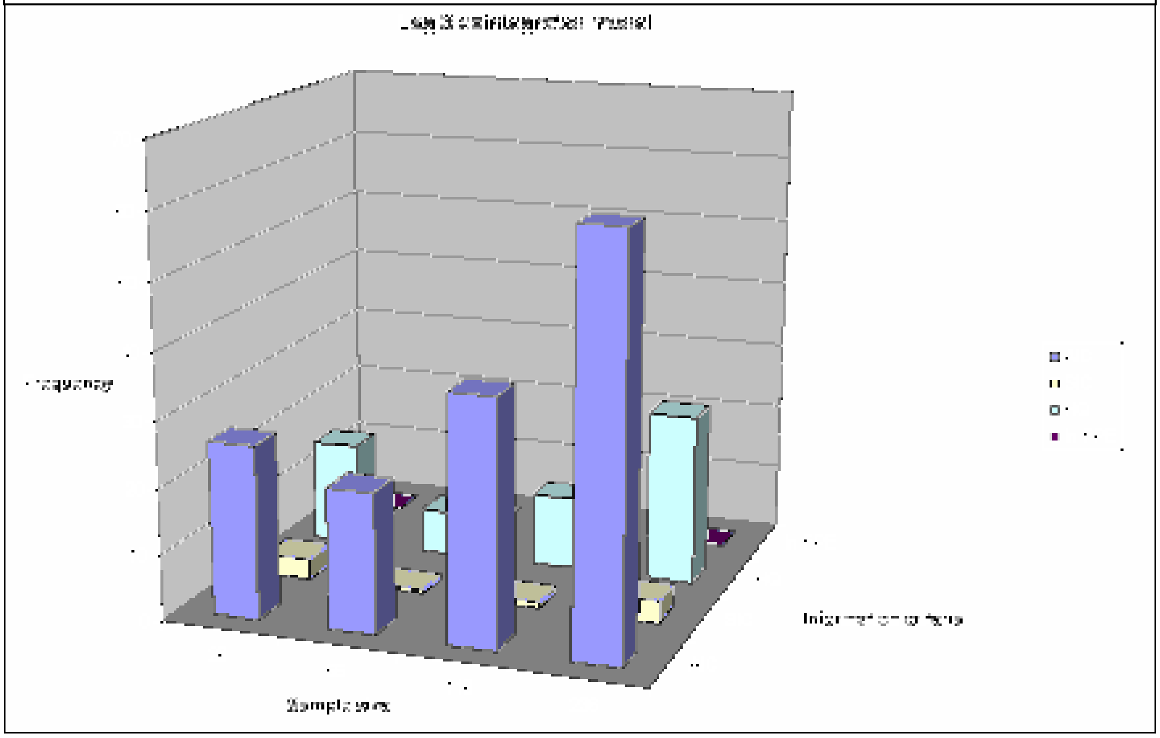


Figure 7.4 illustrates that for the lag three cointegrated model, AIC outperforms SIC, HQ and FPE and this trend is consistent across all sample sizes.

In summary, for the lag length one, two and three cointegrated model, AIC performs the best in selecting the true lag for the model, particularly when the sample size is large. However, in the lag zero cointegrated model, AIC performs poorly in selecting the true lag even when the sample size is large. This may have resulted from the autoregressive irregular variable incorporated in the model. The model selector FPE was recommended for selecting the lag length for the lag zero cointegrated model. As previously discussed, this is an unusual result and requires further investigation. Overall, when the sample size is large AIC performs the best.

The next section attempts to address how well the IC's perform when the lag length of the model is increased keeping the sample size constant. The previous section considers the IC's performances with the sample size increasing and keeping the lag length constant.

7.5 Cointegrated models for a sample size of 26

The percentage of various criteria correctly estimating the true lags, lag zero to lag three for the sample size 26 are tabulated in Tables 7.5.1 to 7.5.4. The values that are highlighted in the table are the highest selection percentage for each of the IC.

Table 7.5.1: Simulated cointegrated model of lag length 0 for a sample size of 26

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	63.38	90.99	74.85	99.94
1	21.83	8.12	17.96	0.06
2	7.46	0.67	4.14	0
3	7.33	0.22	3.04	0

Table 7.5.2: Simulated cointegrated model of lag length 1 for a sample size of 26

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	16.41	46.86	24.98	98.7
1	60.75	49.89	61.49	1.3
2	13.77	2.61	9.04	0
3	9.07	0.64	4.5	0

Table 7.5.3: Simulated cointegrated model of lag length 2 for a sample size of 26

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	27.44	67.55	39.96	99.84
1	37.94	27.21	38.42	0.16
2	23.05	4.56	16.21	0
3	11.57	0.68	5.41	0

Table 7.5.4: Simulated cointegrated model of lag length 3 for a sample size of 26

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	6.56	35.09	13.32	97.27
1	41.27	51.97	49.1	2.73
2	25.98	9.73	21.77	0
3	26.19	3.21	15.8	0

The results in Table 7.5.1 show that for small sample sizes the IC perform well in selecting the lag zero cointegrated model, particularly, for the FPE and SIC where FPE selects the correct lag almost perfectly whilst SIC selects the correct lag just less than 91%. The AIC and HQ selection for lag zero cointegrated model perform indifferently, 64% and 75%, respectively. In the lag one (see Table 7.5.2), two (see Table 7.5.3) and three (see Table 7.5.4) cointegrated models, AIC and HQ perform better than the other two criteria. In the lag one cointegrated model, AIC and HQ score approximately 60% with the other criteria scoring less.

As the number of lag terms increases in the model, the performances of the various criteria in selecting the correct lag diminishes, an observation which one would expect for the small sample sizes. When a model consists of two or more lag terms, all the IC's under-estimate the correct lag for the model. In Table 7.5.3, the model selected by AIC was a lag one model whilst for SIC, HQ and FPE the model selected was a lag zero model despite the fact that the correct model was a lag two model.

In summary, for a small sample size (for T = 26), AIC outperforms SIC, HQ and FPE. In addition the AIC selection percentage for the correct model decreases slightly as the number of lag terms increase.

7.6 Cointegrated models for a sample size of 56

The percentage of various criteria correctly estimating the true lags, lag zero to lag three for the sample size 56 are tabulated in Tables 7.6.1 to 7.6.4. The values that are highlighted in the table are the highest selection percentage for each of the IC.

Table 7.6.1: Simulated cointegrated model of lag length 0 for a sample size of 56

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	44.58	88.04	67.38	99.85
1	43.35	11.77	30.16	0.15
2	8.36	0.17	2.09	0
3	3.7	0.02	0.37	0

Table 7.6.2: Simulated cointegrated model of lag length 1 for a sample size of 56

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	0.78	12.6	3.06	73.58
1	82.39	86.83	92.08	26.42
2	11.92	0.55	4.13	0
3	4.91	0.03	0.72	0

Table 7.6.3: Simulated cointegrated model of lag length 2 for a sample size of 56

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**In FPE**
0	4.35	41.14	15.15	94.62
1	57.01	55.22	67.2	5.38
2	31.27	3.57	16.25	0
3	7.37	0.07	1.39	0

Table 7.6.4: Simulated cointegrated model of lag length 3 for a sample size of 56

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**In FPE**
0	0.06	4.12	0.61	49.12
1	37.76	84.57	63.52	50.83
2	41.03	10.68	29.2	0.05
3	21.15	0.62	6.67	0

The results in Table 7.6.1 show that the IC's perform fairly well in selecting the lag zero cointegrated model, with three of the four IC's scoring above 68% and the fourth, AIC, scoring 45%. When the correct number of lag terms is three, the IC performances are poor, AIC scores 21%, HQ scores 7% and the other two score almost zero. For the higher order model (in Table 7.6.4), the SIC and FPE performances are exceptionally poor compared to the other IC's. In addition, SIC and FPE generally select a lower order model rather than the correct lag length model.

In summary, for a sample size of T = 56, SIC is the preferred model selector for the lower order models and AIC is the preferred model selector for the higher order models i.e. the cointegrated model with a lag length of two or a cointegrated model with a lag length of three.

7.7 Cointegrated models for a sample size of 116

The percentage of various criteria correctly estimating the true lags, lag zero to lag three for the sample size 116 are tabulated in Tables 7.7.1 to 7.7.4. The values that are highlighted in the table are the highest selection percentage for each of the IC.

Table 7.7.1: Simulated cointegrated model of lag length 0 for a sample size of 116

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	14.08	70.3	37.61	96.51
1	70.85	29.63	60.16	3.49
2	11.29	0.06	2.1	0
3	3.78	0	0.13	0

Table 7.7.2: Simulated cointegrated model of lag length 1 for a sample size of 116

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	0	0.16	0	5.36
1	87.12	99.72	97.96	94.64
2	9.44	0.12	1.85	0
3	3.44	0	0.18	0

Table 7.7.3: Simulated cointegrated model of lag length 2 for a sample size of 116

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	0.01	5	0.54	40.81
1	40.45	87.81	70.62	59.06
2	51.53	7.16	28.05	0.13
3	8.01	0.04	0.79	0

Table 7.7.4: Simulated cointegrated model of lag length 3 for a sample size of 116				
	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	0	0	0	0.49
1	11.76	77.07	39.06	98.34
2	51.28	21.96	49.84	1.17
3	36.96	0.97	11.09	0

The results in Table 7.7.1 show that for medium sample sizes, SIC and FPE perform well in selecting the lag zero cointegrated model, scoring above 70%. The other two criteria, AIC and HQ, score less and often over-estimate the lag length of the model. Both indicate lag one as the correct lag with percentages of 71% and 60%, respectively.

In Table 7.7.2, all the IC perform well in estimating the true lag, lag one for the model. In particular, SIC and HQ select the correct lag almost perfectly whilst FPE selects the correct lag just less than 95% of the time. The fourth criterion, AIC, is slightly lower than the others and scores 87%.

For the medium sample size with T = 116, as the number of lag terms increases, the performances of the various criteria decrease and under-estimate the true lag length of the model. As observed for the T = 56 case, AIC is the preferred model selector for higher order terms whilst for lower order models, FPE and SIC perform well.

7.8 Cointegrated models for a sample size of 236

The percentage of various criteria correctly estimating the true lags, lag zero to lag three for the sample size 236 are tabulated in Tables 7.8.1 to 7.8.4. The values that are highlighted in the table are the highest selection percentage for each of the IC.

Table 7.8.1: Simulated cointegrated model of lag length 0 for a sample size of 236

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	0.8	28.78	6.33	64.67
1	78.39	71.05	90.57	35.32
2	16.54	0.17	2.98	0.01
3	4.26	0	0.12	0

Table 7.8.2: Simulated cointegrated model of lag length 1 for a sample size of 236

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	0	0	0	0
1	87.72	99.97	98.74	100
2	9.18	0.03	1.18	0
3	3.1	0	0.09	0

Table 7.8.3: Simulated cointegrated model of lag length 2 for a sample size of 236

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**ln FPE**
0	0	0.01	0	0.25
1	13.04	80.88	43.09	96.96
2	77.86	19.11	56.21	2.8
3	9.1	0.01	0.71	0

Table 7.8.4: Simulated cointegrated model of lag length 3 for a sample size of 236

	Model selection criteria			
Lag	**AIC**	**SIC**	**HQ**	**In FPE**
0	0	0	0	0
1	0.52	40.4	7.03	80.89
2	37.95	56.07	67.13	19.06
3	61.53	3.53	25.84	0.05

The results in Table 7.8.1 shows that all four criteria perform poorly in the selecting the correct lag model with only FPE scoring above 50%.

In Table 7.8.2, all the IC performances are good in particular SIC, HQ and FPE which score above 98% whilst AIC selects the correct lag model just less than 88% of the time. These are good results and not unexpected since the sample size is large and the lag length is small. For the larger sample size (for $T = 236$), AIC outperforms SIC, HQ and FPE. This trend is consistent for the lag one (see Table 7.8.2), two (see Table 7.8.3) and three (see Table 7.8.4) cointegrated models, despite the disappointing performance for the lag zero cointegrated model.

In summary, for all sample sizes, FPE, SIC and HQ select the lag zero and lag one models correctly more often and in particular are correct more often when the sample size is large. For the higher order models, AIC performs the best in selecting the correct lag length for the model and this is consistent for all sample sizes.

Chapter 8

Conclusion and Further Work

8.1 Conclusion

This study compares the selection capabilities of four information criteria used to determine the number of lag terms in a bivariate cointegrated model. In general, Akaike's information criterion dominates the other three criteria considered, in that AIC selects the correct model most often.

In particular when the model is of higher order, AIC consistently outperforms the other IC's. For lower order models, with zero or one lag term, AIC is less successful than one of the other IC's, subject to sample size and order number. As an illustration, comparing AIC and HQ as VEC (1) models for T = 116, we see that HQ performs marginally better then AIC.

Despite the cases where AIC is less successful, for a practioner who has little idea of the order of the model, AIC is the most consistent performer, whilst the other criteria exhibit a tendency to under-fit the model. These findings differ from the results of Gonzalo and Pitarakis (1998) in that this study favours HQ for VAR models with sample sizes $T \geq 100$. In addition, these simulation results concur with the findings of Lütkepohl (1985) in that as sample size increases, ability to select correctly improves for all IC.

8.2 Further work

Model selection strategies for cointegrated models are an active research area and are expected to continue as long as there is a demand from applied researchers doing empirical studies. The results of this study lend support to a strategy that favours the use of AIC rather than SIC, HQ and FPE.

Within this study, we have restricted the model to the bivariate case. Extension to higher dimension models will be a natural route to follow. A trivariate model is currently under investigation by another researcher. In the cointegrated model, the number of lag terms incorporated in each of the time series equations can be extended (for example using a lag length of eight). The increase of lag terms will provide additional information as to how well the IC's perform in selecting the true lag for the model. The number of observations in each of the time series equations can also be extended, as in Khim and Liew (2004) where the highest observations used were 960.

More importantly, further investigation is required for the lag zero cointegrated model, as the number of observation increases, the performances of the IC decreased an unexpected observation. It would be interesting to know whether the irregular component in the lag zero cointegrated model did make such a difference in the performance of the IC's.

Changing the value of the coefficients in the simulated model is another consideration for further examination. There are several opportunities to expand this study. These opportunities should provide researcher(s) with the best criteria used for determining the lag length in cointegrated modelling.

Appendix A

Glossary

1.	Cointegrated model:	A model consists of two or more cointegrated variables.
2.	Cointegrated variables:	The variables that have the same order of integration and the linear combination of these variables are stationary.
3.	Differenced series:	It entails regression a variable on time and saving the residual.
4.	Error correction model:	A model where the movement of the variables in any period is related to the previous period's gap from long-run equilibrium.
5.	Lag:	An event occurring at time $t + k$ ($k > 0$) is said to lag behind an event occurring at time t, the extent of the lag being k.
6.	Lag length:	The number of lag terms used.
7.	Long-run equilibrium relationship:	This relationship exists when the existence of a combination of the non-stationary variables are stationary.

8.	Non-stationary series:	A series of equations that may have a pronounced trend or appear to meander without a constant long-run mean or variance.
9.	Stationary series:	A series of equations that has a constant mean and variance.
10.	Time series:	A sequence of data points, measured typically at successive times.

Appendix B

Programmes

B.1 Programme used to test for unit roots and cointegration

```
wfopen c:\a\lag1\lag1_1\lag1_1
dbcreate c:\a\lag1\lag1_1\doesnot_meetspec
dbcreate c:\a\lag1\lag1_1\meetspec
!count = 0
!gmax =3825
table(!gmax,4) result
for !j=1 to !gmax
        !smax =2
        for !s=1 to !smax
                %sname = "x"+ @str(!s)+@str(!j)+"t"
                fetch(d=c:\a\lag1\lag1_1\lag1_1) %sname
                uroot(adf,none,lag=1,save=matprob) {%sname}
                if matprob(4,1)>0.05 then
                uroot(adf,none,lag=1,dif=1,save=matprob1) {%sname}
                        if matprob1(4,1)<0.05 then
                        !count = !count+1
                        endif
                else
                !count = !count
                endif
        next
                if !count =2 then
                for !s=1 to !smax
                        %sname = "x"+@str(!s)+ @str(!j)+"t"
                        fetch(d=c:\a\lag1\lag1_1\lag1_1) %sname
                        store c:\a\lag1\lag1_1\meetspec::{%sname}
                next
                        %sname1 = "x"+"1"+ @str(!j)+"t"
                        %sname2 = "x"+"2"+ @str(!j)+"t"
                        fetch(d=c:\a\lag1\lag1_1\lag1_1) %sname1
                        fetch(d=c:\a\lag1\lag1_1\lag1_1) %sname2

                group xy.add {%sname1} {%sname2}
                coint(a,0,save=matCoint) {%sname1} {%sname2}
                coint(a,0,save=matLogLike) {%sname1} {%sname2}
                       if matcoint(1,3)>12.32090 and matcoint(2,3)< 4.129906 then
                        result(!j,1) = %sname1+"&"+%sname2
                        result(!j,2) = "1"
                        result(!j,3) = matLogLike(2,4)
                        var a1.ec(a) 0 0 {%sname1} {%sname2}
                       else
                       for !s=1 to !smax
                        %sname = "x"+@str(!s)+@str(!j)+"t"
                        fetch(d=c:\a\lag1\lag1_1\lag1_1) %sname
                        store c:\a\lag1\lag1_1\doesnot_meetspec::{%sname}
                        delete c:\a\lag1\lag1_1\meetspec::{%sname}
                       next
                       endif
                        !count = 0
                else
```

```
            for !s=1 to !smax
                  %sname = "x"+@str(!s)+ @str(!j)+"t"
                  fetch(d=c:\a\lag1\lag1_1\lag1_1) %sname
                  store c:\a\lag1\lag1_1\doesnot_meetspec::{%sname}
            next
                  !count = 0
            endif
next
save c:\a\lag1\lag1_1\testResult
```

B.2 Programme used to obtain the likelihood value

Likelihood value for a bivariate lag 1 autoregressive model

```
wfopen c:\a\lag1\lag1_1\meetspec
dbcreate c:\a\lag1\lag1_1\log_t_1\meetspec
dbcreate c:\a\lag1\lag1_1\log_t_1\doesnot_meetspec

!obs=3
!lag = 1
%lag ="t_1"
!gmax =3825
for !j=1 to !gmax
    !smax =2
    for !s=1 to !smax
        %sname = "x"+@str(!s)+ @str(!j)+"t"
        %newsname = "x"+@str(!s)+ @str(!j)+%lag
        if @isobject(%sname) =1 then
        fetch(d=c:\a\lag1\lag1_1\meetspec) %sname
        series {%newsname} = {%sname}(!obs)
        store c:\a\lag1\lag1_1\log_t_1\meetspec::{%newsname}
        endif
    next
next

    table(!gmax,4) result1
    for !j=1 to !gmax
            %sname1 = "x"+"1"+ @str(!j)+%lag
            %sname2 = "x"+"2"+ @str(!j)+%lag
            if @isobject(%sname1) =1 and @isobject(%sname2) =1 then
                fetch(d=c:\a\lag1\lag1_1\log_t_1\meetspec) %sname1
                fetch(d=c:\a\lag1\lag1_1\log_t_1\meetspec) %sname2
                group xy.add {%sname1} {%sname2}
                coint(a,0,save=matLogLike1) {%sname1} {%sname2}
                result1(!j,1) = %sname1+"&"+%sname2
                result1(!j,2) = "1"
                result1(!j,3) = matLogLike1(2,4)
                var a1.ec(a) 0 0 {%sname1} {%sname2}
            endif
    next
save c:\a\lag1\lag1_1\log_t_1\result1

wfopen c:\a\lag1\lag1_1\doesnot_meetspec
for !j=1 to !gmax
    !smax =2
    for !s=1 to !smax
        %sname = "x"+@str(!s)+ @str(!j)+"t"
        %newsname = "x"+@str(!s)+ @str(!j)+%lag
        if @isobject(%sname) =1 then
        fetch(d=c:\a\lag1\lag1_1\doesnot_meetspec) %sname
        series {%newsname} = {%sname}(!obs)
        store c:\a\lag1\lag1_1\log_t_1\doesnot_meetspec::{%newsname}
        endif
    next
next

    table(!gmax,4) result2
```

```
	for !j=1 to !gmax
		%sname1 = "x"+"1"+ @str(!j)+%lag
		%sname2 = "x"+"2"+ @str(!j)+%lag
		if @isobject(%sname1) =1 and @isobject(%sname2) =1 then
			fetch(d=c:\a\lag1\lag1_1\log_t_1\doesnot_meetspec)
%sname1
			fetch(d=c:\a\lag1\lag1_1\log_t_1\doesnot_meetspec)
%sname2
			group xy.add {%sname1} {%sname2}
			coint(a,0,save=matLogLike2) {%sname1} {%sname2}
			result2(!j,1) = %sname1+"&"+%sname2
			result2(!j,2) = "0"
			result2(!j,3) = matLogLike2(2,4)
			var a1.ec(a) 0 0 {%sname1} {%sname2}
		endif
	next
save c:\a\lag1\lag1_1\log_t_1\result2
```

Likelihood value for a bivariate lag 2 autoregressive model

```
wfopen c:\a\lag1\lag1_1\meetspec
dbcreate c:\a\lag1\lag1_1\log_t_2\meetspec
dbcreate c:\a\lag1\lag1_1\log_t_2\doesnot_meetspec

!obs=2
!lag = 2
%lag ="t_2"
!gmax =3825
for !j=1 to !gmax
        !smax =2
        for !s=1 to !smax
                %sname = "x"+@str(!s)+ @str(!j)+"t"
                %newsname = "x"+@str(!s)+ @str(!j)+%lag
                if @isobject(%sname) =1 then
                fetch(d=c:\a\lag1\lag1_1\meetspec) %sname
                series {%newsname} = {%sname}(!obs)
                store c:\a\lag1\lag1_1\log_t_2\meetspec::{%newsname}
                endif
        next
next

        table(!gmax,4) result1
        for !j=1 to !gmax
                        %sname1 = "x"+"1"+ @str(!j)+%lag
                        %sname2 = "x"+"2"+ @str(!j)+%lag
                        if @isobject(%sname1) =1 and @isobject(%sname2) =1 then
                                fetch(d=c:\a\lag1\lag1_1\log_t_2\meetspec) %sname1
                                fetch(d=c:\a\lag1\lag1_1\log_t_2\meetspec) %sname2
                                group xy.add {%sname1} {%sname2}
                                coint(a,1,save=matLogLike1) {%sname1} {%sname2}
                                result1(!j,1) = %sname1+"&"+%sname2
                                result1(!j,2) = "1"
                                result1(!j,3) = matLogLike1(2,4)
                                var a1.ec(a) 1 1 {%sname1} {%sname2}
                        endif
        next
save c:\a\lag1\lag1_1\log_t_2\result1

wfopen c:\a\lag1\lag1_1\doesnot_meetspec
for !j=1 to !gmax
        !smax =2
        for !s=1 to !smax
                %sname = "x"+@str(!s)+ @str(!j)+"t"
                %newsname = "x"+@str(!s)+ @str(!j)+%lag
                if @isobject(%sname) =1 then
                fetch(d=c:\a\lag1\lag1_1\doesnot_meetspec) %sname
                series {%newsname} = {%sname}(!obs)
                store c:\a\lag1\lag1_1\log_t_2\doesnot_meetspec::{%newsname}
                endif
        next
next

        table(!gmax,4) result2
        for !j=1 to !gmax
                        %sname1 = "x"+"1"+ @str(!j)+%lag
                        %sname2 = "x"+"2"+ @str(!j)+%lag
                        if @isobject(%sname1) =1 and @isobject(%sname2) =1 then
                                fetch(d=c:\a\lag1\lag1_1\log_t_2\doesnot_meetspec)
```

```
%sname1
                        fetch(d=c:\a\lag1\lag1_1\log_t_2\doesnot_meetspec)
%sname2
                        group xy.add {%sname1} {%sname2}
                        coint(a,1,save=matLogLike2) {%sname1} {%sname2}
                        result2(!j,1) = %sname1+"&"+%sname2
                        result2(!j,2) = "0"
                        result2(!j,3) = matLogLike2(2,4)
                        var a1.ec(a) 1 1 {%sname1} {%sname2}
                  endif
      next
save c:\a\lag1\lag1_1\log_t_2\result2
```

Likelihood value for a bivariate lag 3 autoregressive model

```
wfopen c:\a\lag1\lag1_1\meetspec
dbcreate c:\a\lag1\lag1_1\log_t_3\meetspec
dbcreate c:\a\lag1\lag1_1\log_t_3\doesnot_meetspec

!obs= 1
!lag = 3
%lag ="t_3"
!gmax =3825
for !j=1 to !gmax
        !smax =2
        for !s=1 to !smax
                %sname = "x"+@str(!s)+ @str(!j)+"t"
                %newsname = "x"+@str(!s)+ @str(!j)+%lag
                if @isobject(%sname) =1 then
                fetch(d=c:\a\lag1\lag1_1\meetspec) %sname
                series {%newsname} = {%sname}(!obs)
                store c:\a\lag1\lag1_1\log_t_3\meetspec::{%newsname}
                endif
        next
next

        table(!gmax,4) result1
        for !j=1 to !gmax
                        %sname1 = "x"+"1"+ @str(!j)+%lag
                        %sname2 = "x"+"2"+ @str(!j)+%lag
                        if @isobject(%sname1) =1 and @isobject(%sname2) =1 then
                                fetch(d=c:\a\lag1\lag1_1\log_t_3\meetspec) %sname1
                                fetch(d=c:\a\lag1\lag1_1\log_t_3\meetspec) %sname2
                                group xy.add {%sname1} {%sname2}
                                coint(a,2,save=matLogLike1) {%sname1} {%sname2}
                                result1(!j,1) = %sname1+"&"+%sname2
                                result1(!j,2) = "1"
                                result1(!j,3) = matLogLike1(2,4)
                                var a1.ec(a) 1 2 {%sname1} {%sname2}
                        endif
        next
save c:\a\lag1\lag1_1\log_t_3\result1

wfopen c:\a\lag1\lag1_1\doesnot_meetspec
for !j=1 to !gmax
        !smax =2
        for !s=1 to !smax
                %sname = "x"+@str(!s)+ @str(!j)+"t"
                %newsname = "x"+@str(!s)+ @str(!j)+%lag
                if @isobject(%sname) =1 then
                fetch(d=c:\a\lag1\lag1_1\doesnot_meetspec) %sname
                series {%newsname} = {%sname}(!obs)
                store c:\a\lag1\lag1_1\log_t_3\doesnot_meetspec::{%newsname}
                endif
        next
next

        table(!gmax,4) result2
        for !j=1 to !gmax
                        %sname1 = "x"+"1"+ @str(!j)+%lag
                        %sname2 = "x"+"2"+ @str(!j)+%lag
```

```
            if @isobject(%sname1) =1 and @isobject(%sname2) =1 then
                fetch(d=c:\a\lag1\lag1_1\log_t_3\doesnot_meetspec) %sname1
                fetch(d=c:\a\lag1\lag1_1\log_t_3\doesnot_meetspec) %sname2
                group xy.add {%sname1} {%sname2}
                coint(a,2,save=matLogLike2) {%sname1} {%sname2}
                result2(!j,1) = %sname1+"&"+%sname2
                result2(!j,2) = "0"
                result2(!j,3) = matLogLike2(2,4)
                var a1.ec(a) 1 2 {%sname1} {%sname2}
            endif
    next
save c:\a\lag1\lag1_1\log_t_3\result2
```

Likelihood value for a bivariate lag 4 autoregressive model

```
wfopen c:\a\lag1\lag1_1\meetspec
dbcreate c:\a\lag1\lag1_1\log_t_4\meetspec
dbcreate c:\a\lag1\lag1_1\log_t_4\doesnot_meetspec

!lag = 4
%lag ="t_4"
!gmax =3825
for !j=1 to !gmax
        !smax =2
        for !s=1 to !smax
                %sname = "x"+@str(!s)+ @str(!j)+"t"
                %newsname = "x"+@str(!s)+ @str(!j)+%lag
                if @isobject(%sname) =1 then
                fetch(d=c:\a\lag1\lag1_1\meetspec) %sname
                series {%newsname} = {%sname}
                store c:\a\lag1\lag1_1\log_t_4\meetspec::{%newsname}
                endif
        next
next

        table(!gmax,4) result1
        for !j=1 to !gmax
                        %sname1 = "x"+"1"+ @str(!j)+%lag
                        %sname2 = "x"+"2"+ @str(!j)+%lag
                        if @isobject(%sname1) =1 and @isobject(%sname2) =1 then
                                fetch(d=c:\a\lag1\lag1_1\log_t_4\meetspec) %sname1
                                fetch(d=c:\a\lag1\lag1_1\log_t_4\meetspec) %sname2
                                group xy.add {%sname1} {%sname2}
                                coint(a,3,save=matLogLike1) {%sname1} {%sname2}
                                result1(!j,1) = %sname1+"&"+%sname2
                                result1(!j,2) = "1"
                                result1(!j,3) = matLogLike1(2,4)
                                var a1.ec(a) 1 3 {%sname1} {%sname2}
                        endif
        next
save c:\a\lag1\lag1_1\log_t_4\result1

wfopen c:\a\lag1\lag1_1\doesnot_meetspec
for !j=1 to !gmax
        !smax =2
        for !s=1 to !smax
                %sname = "x"+@str(!s)+ @str(!j)+"t"
                %newsname = "x"+@str(!s)+ @str(!j)+%lag
                if @isobject(%sname) =1 then
                fetch(d=c:\a\lag1\lag1_1\doesnot_meetspec) %sname
                series {%newsname} = {%sname}
                store c:\a\lag1\lag1_1\log_t_4\doesnot_meetspec::{%newsname}
                endif
        next
next

        table(!gmax,4) result2
        for !j=1 to !gmax
                        %sname1 = "x"+"1"+ @str(!j)+%lag
                        %sname2 = "x"+"2"+ @str(!j)+%lag
                        if @isobject(%sname1) =1 and @isobject(%sname2) =1 then
```

```
                    fetch(d=c:\a\lag1\lag1_1\log_t_4\doesnot_meetspec)
%sname1
                    fetch(d=c:\a\lag1\lag1_1\log_t_4\doesnot_meetspec)
%sname2
                    group xy.add {%sname1} {%sname2}
                    coint(a,3,save=matLogLike2) {%sname1} {%sname2}
                    result2(!j,1) = %sname1+"&"+%sname2
                    result2(!j,2) = "0"
                    result2(!j,3) = matLogLike2(2,4)
                    var a1.ec(a) 1 3 {%sname1} {%sname2}
               endif
     next
save c:\a\lag1\lag1_1\log_t_4\result2
```

References

Akaike, H. (1969). Fitting autoregressive models for prediction. Annals of the Institute of Statistical Mathematics, no. 20, pp. 243-247.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In 2nd International Symposium on Information Theory (Ed. B.N. Petrov & F. Csaki), Akademia Kiado, Budapest, pp. 267-281.

Akinboade, O. A. & Niedermeier, E. W. (2002). Labour costs and inflation in South Africa: an econometric study. Journal for Studies in Economics and Econometrics, vol. 26, no. 2, pp. 1-18.

Anderson, T. W. (1963). Determination of the order of dependence in normally distributed time series. Time Series Analysis (Ed. Rosenblatt), New York: John Wiley, pp. 425-446.

Anderson, D. R., Burnham, K. P. & White, G. C. (1994). AIC model selection in overdispersed capture-recapture data. Ecology, vol. 76, no. 6, pp. 1780-1793.

Barr, G. D. & Kahn, B. (1995). Testing for Purchasing Power Parity in South Africa in the presence of real shocks. Journal for Studies in Economics and Econometrics, vol. 19, no. 1, pp. 69-86.

Barr, G. D. & Kantor, B. S. (1990). The application of a vector autoregressive model to money, income and price links in the South African economy. Journal for Studies in Economics and Econometrics, vol. 14, no. 1, pp. 39-49.

Botha, I. & Apostolellis, G. (2003). Analysis of financial cointegration between emerging and developed equity markets. In Eighth Annual Conference on Econometric Modelling for Africa. RAU University, Aucklandpark, pp. 1-31.

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. Psychometrika, vol. 52, no. 3, pp. 345-370.

Bozdogan, H., Sclove, S. L. & Gupta, A. K. (1994). AIC-replacements for some multivariate tests of homogeneity with applications in multisample clustering and variable selection. Ed. by H. Bozdogan, proceedings of the first US/Japan conference on the frontiers of statistical modeling: an informational approach. Netherlands: Kluwer Academic Publishers, pp. 199-232.

Cheung, Y. & Lai, K. S. (1993). Finite-sample sizes of Johansen's likelihood ratio tests for cointegration. Oxford Bulletin of Economics and Statistics, vol. 55, no. 3, pp. 313-328.

Damoense, M. Y. (2003). The determinants of South Africa's exchange rate. In Eight Annual Conference on Econometric Modeling for Africa, Department of Economics, Stellenbosch University, Cape Town, South Africa, pp. 1-14.

De Wet, T. (2000). Purchasing Power Parity in the long run: an empirical re-evaluation of the South African evidence. Journal for Studies in Economics and Econometrics, vol. 24, no. 2, pp. 19-34.

Diamandis, P. F. & Kouretas, G. P. (1995). Cointegration and market efficiency: a time series analysis of the Greek drachma. Applied Economics Letters, vol. 2, no. 8, pp. 271-277.

Dickey, D. A. & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, no. 74, pp. 427-431.

Dickey, D. A. & Fuller, W. A. (1981). The likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, no. 49, pp. 1057-1072.

Dickey, D. A. & Pantula, S. (1987). Determining the order of differencing in autoregressive processes. Journal of Business and Economic Statistics, no. 5, pp. 455-461.

Dwyer, G. P. & Wallace, M. S. (1992). Cointegration and market efficiency. Journal of International Money and Finance, no. 11, pp. 318-327.

Enders, W. (2004). Applied econometric time series. 2nd Edition. John Wiley and Sons, Inc.

Engle, R. F. & Granger, C. W. J. (1987). Cointegration and error correction: representation, estimation and testing. Econometrica, no. 55, pp. 251-276.

Engle, R. F. & Yoo, B. S. (1989). Cointegrated economic time series: a survey with new results. Discussion paper 89-138, University of California, San Diegao.

Evenson, R. E. (1967). The contribution of agricultural research to production. Journal of Farm Economics, no. 49, pp. 1415-1425.

Fedderke, J. W. & Joao, M. (2001). Arbitrage, cointegration and efficiency in financial markets in the presence of financial crises. The South Africa Journal of Economics, vol. 69, no. 3, pp. 366-384.

Fedderke, J. W., de Kadt, R. H. J. & Luiz, J. M. (2001). Growth and institutions: a study of the link between political institutions and economic growth in South Africa - a time series study: 1935-97. Journal for Studies in Economics and Econometrics, vol. 25, no. 1, pp. 1-26.

Ferret, A. & Page, M. J. (1998). Cointegration of South African index futures contracts and their underlying spot market indices. Journal for Studies in Economics and Econometrics, vol. 22, no. 1, pp. 69-90.

Gonzalo, J. & Pitarakis, J. (1998). Specification via model selection in vector error correction models. Economics Letters, no. 60, pp. 321-328.

Gonzalo, J. & Pitarakis, J. (2002). Lag length estimation in large dimensional systems. Journal of Time Series Analysis, vol. 23, no. 4, pp. 401-423.

Granger, C. W. (1986). Developments in the study of cointegrated variables. Oxford Bulletin of Economics and Statistics, no. 48, pp. 213-228.

Gujarati, D. N. (1995). Basic Econometrics. 3rd Edition, McGraw-Hill, New York, Chapter 21.

Gumede, V. (2000). Import demand elasticities for South Africa: a cointegration analysis. Journal for Studies in Economics and Econometrics, vol. 24, no. 1, pp. 21-37.

Gumede, V. (2003). Manufacturing export elasticities in South Africa - a time series approach. Journal for Studies in Economics and Econometrics, vol. 27, no. 1, pp. 39-51.

Hannan, E. J. & Quinn, B. G. (1979). The determination of the order of an autoregression. Journal of the Royal Statistical Society, Series B, no. 41, pp. 190-195.

Hannan, E. J. (1980). The estimation of order of an ARMA process. The Annuals of Statistics, vol. 8, no. 5, pp. 1071-1081.

Harris, R. (1995). Cointegration analysis in econometric modelling. 1st Edition, Prentice Hall, Pearson Education Limited.

Ivanov, V. & Kilian, L. (2005). A practitioner's guide to lag order selection for VAR impulse response analysis. Studies in Nonlinear Dynamics & Econometrics, vol. 9, no. 1, article 2.

Johansen, S. (1988). Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control, no. 12, pp. 231-254.

Khim, V. & Liew, S. (2004). Which lag length selection criteria should we employ?. Economics Bulletin, vol. 3, no. 33, pp. 1-9.

Knutson, M. & Tweeten, L. G. (1979). Toward an optimal rate of growth in agricultural production research and extension. American Journal of Agricultural Economics, no. 61, pp. 70-76.

Kouassi, E. (1997). Analyzing and simulating terms of trade shocks on Ivorian macro-economic variables: an error corrected-VAR approach. Journal for Studies in Economics and Econometrics, vol. 21, no. 1, pp. 91-108.

Krolzig, H. M. & Hendry, D. F. (2000). Computer automation of general-to-specific model selection procedures. Forthcoming, Journal of Economic Dynamics and Control.

Kullback, S. & Leibler, R. A. (1951). On information and sufficiency. Annals of Mathematical Statistics, no. 22, pp. 79-86.

Leng, H. M. J. (2002). The South African share index futures and share markets: efficiency and causality revisited. Journal for Studies in Economics and Econometrics, vol. 26, no. 3, pp. 1-18.

Lütkepohl, H. (1985). Comparison of criteria for estimating the order of a vector autoregressive process. Journal of Time Series Analysis, vol. 6, no. 1, pp. 35-52.

MacKinnon, J. G., Haug, A. A. & Michelis, L. (1999). Numerical distribution functions of likelihood ratio tests for cointegration. Journal of Applied Econometrics, no. 14, pp. 563-577.

Madsen, J. B. (1997). Macroeconomic adjustment and policy in South Africa. Journal for Studies in Economics and Econometrics, vol. 21, no. 2, pp. 23-43.

Mainardi, S. (1995). The influence of the gold price on exchange rates in South Africa. Journal for Studies in Economics and Econometrics, vol. 19, no. 1, pp. 51-67.

Mainardi, S. (2000). Social rates of discount revisited, with derivations for Trinidad and South Africa. Journal for Studies in Economics and Econometrics, vol. 24, no. 2, pp. 67-85.

Mendenhall, W. & Sincich, T. (2003). A second course in statistics regression analysis. 6th Edition. United State of America: Pearson Education, Inc.

Mills, T. C. (1999). The econometric modelling of financial time series. 2nd Edition. United Kingdom: Cambridge University Press.

Moolman, E. (2002). The term structure as a predictor of recessions. Journal for Studies in Economics and Econometrics, vol. 26, no. 3, pp. 43-52.

Neubrech, S. E. & Pienaar, W. J. (2001). Possible implications of the price, cross-price and income elasticities of the demand for public road transport in the Cape Metropolitan Area - a cointegration analysis. Journal for Studies in Economics and Econometrics, vol. 25, no. 3, pp. 57-73.

Ng, S. & Perron, P. (2005). Practitioners' corner: a note on the selection of time series models. Oxford Bulletin of Economics and Statistics, vol. 67, no. 1, pp. 115-134.

Nwokoma, N. I. & Olofin, S. O. (2003). An empirical evaluation of globalisation and the Nigerian stock market. In Annual Conference of the African Econometrics Society (AES), University of Stellenbosch, South Africa, pp. 1-20.

Parikh, A. & Kahn, B. (1997). Determinants of real exchange rates in South Africa: a short-run and long-run analysis. Journal for Studies in Economics and Econometrics, vol. 21, no. 2, pp. 1-22.

Parzen, E. (1982). Data modelling using quantile and density-quantile function. London: Academic Press, pp. 23-52.

Pesaran, M. H. & Shin, Y. (1995a). Long run structural modeling. Unpublished manuscript, University of Cambridge.

Pesaran, M. H. & Shin, Y. (1995b). An autoregressive distributed lag modelling approach to cointegration analysis. DAE Working Paper no. 9514, Department of Applied Economics, University of Cambridge.

Pesaran, M. H., Shin, Y. & Smith, R. J. (1996). Testing for the existence of a long run relationship. Unpublished manuscript, University of Cambridge.

Phillips, P. C. B. (1987). Time series regression with a unit root. Econometrica, no. 55, pp. 227-301.

Phillips, P. C. B. & Perron, P. (1988). Testing for unit roots in time series regression. Biometrika, no. 75, pp. 335-346.

Quinn, B. G. (1980). Order determination for a multivariate autoregression. Journal of the Royal Statistical Society, Series B 42, no. 2, pp. 182-185.

Schwert, G. W. (1989). Why does stock market volatility change over time?. Journal of Finance, no. 44, pp. 1115-1153.

Schwarz, G. (1978). Estimating the dimension of a model. The Annals of Statistics, vol. 6, no. 2, pp. 461-464.

Sclove, S. L. (1994). Some aspects of model-selection criteria. H. Bozdogan (Ed), proceedings of the first US/Japan conference on the frontiers of statistical modeling: an informational approach, Netherlands: Kluwer Academic Publishers, pp. 37-67.

Sims, C. A. (1980). Macroeconomics and reality. Econometrica, vol. 48, no. 1, pp. 1-48.

Sinha, D. (1998). Exports and savings in Asia: a re-examination. Journal for Studies in Economics and Econometrics, vol. 22, no. 2, pp. 77-86.

Smit, B. W. & Pellissier, G. M. (1997). The BER annual econometric model of the South African economy: a cointegration version. Journal for Studies in Economics and Econometrics, vol. 21, no. 1, pp. 1-35.

Stone, C. J. (1981). Admissible selection of an accurate and parsimonious normal linear regression model. Annals of Statistics, no. 9, pp. 475-485.

Takane, Y. (1994). A review of applications of AIC in psychometrics. H. Bozdogan (Ed), proceedings of the first US/Japan conference on the frontiers of statistical modeling: an informational approach, Netherlands: Kluwer Academic Publishers, pp. 379-403.

Townsend, R. F. (1999). Livestock supply response in South Africa: an investigation of producer price and technology dynamics. Journal for Studies in Economics and Econometrics, vol. 23, no. 3, pp. 49-61.

Wesso, G. R. (1999). A comparative review of twenty econometric packages. Journal for Studies in Economics and Econometrics, vol. 23, no. 2, pp. 89-118.

Wilson, P. J., Okunev, J., du Plessis, P. G. & Ta, G. (1998). The impact of structural breaks on the integration of property and stock markets in South Africa and Australia. Journal for Studies in Economics and Econometrics, vol. 22, no. 3, pp. 43-70.