**INTRODUCTION TO TIME SERIES ECONOMETRICS**

In the time series econometrics framework, the starting point is to exploit the information that we can get from a variable that is available through the variable itself. An analysis of a single time series is called a univariate time series, and in time series econometrics we can also have multivariate time series models. Traditional econometricians have emphasized the use of economic theory and the study of contemporaneous relationships in order to explain relationships among dependent and explanatory variables.

There are various aspects to time series analysis but one common theme to them all is to fully exploit the dynamic structure in the data, by this we mean that we extract as much information as possible from the past history of the series. The two principle types of time series analysis are time series forecasting and dynamic modelling. Time series forecasting is unlike most other econometrics in that it is not concerned with building structural models, understanding the economy or testing hypothesis. All that it is concerned with is building efficient models, which forecast well. This is usually done by exploiting the dynamic inter-relationship, which exists over time for any single variable. Dynamic modelling on the other hand is still concerned with understanding the structure of the economy and testing hypothesis however it starts from the view that most economic series are slow to adjust to any shock and so to understand the process we must fully capture the adjustment process which may be long and complex. Over the past couple of decades the techniques developed in the time series forecasting literature have become increasingly useful in econometrics generally. Hence we begin this chapter with an account of the basic 'work horse' of time series forecasting, the ARIMA model.

**ARIMA models**

Box and Jenkins (1976) first introduced ARIMA models, the term deriving from:

AR = Autoregressive

I = Integrated

MA = moving average.

**Stationarity**

A key concept underlying time series processes is that of stationary. A time series is covariance stationary when it has the following three characteristics:

(a) Exhibits mean reversion in that it fluctuates around a constant long-run mean;

(b) Has a finite variance that is time-invariant; and

(c) Has a theoretical correlogram that diminishes as the lag length increases.

In its simplest terms a time series *Yt* is said to be stationary if:

(a) Constant for all *t;*

(b) Constant for all *t and*

(c)  Constant for all *t* and all *k* or if, its mean, its variance and its covariances remain constant over time.

**AUTOREGRESSIVE PROCESS**

**AUTOREGRESSIVE PROCESS OF ORDER ONE **

**** Where**, =** white noise process**,**  and are arbitrary constants.

In the analysis of the first order difference equation, if  the time path of is said to unstable. In this case, the impact of a change in on accumulates rather than die out over time.

In time series when there is no stationary covariance stationary process for.

**MOMENTS**

**(i) Mean:**

****

 ****

 ****

 **=**

For stationary ****

****

****

****

**(ii) Variance**

 

 

 

Therefore 

 

 

 

Squaring on both sides we have:





Taking the expectation on both sides we have:



The left hand side and the first term on the right hand side will equal and the third term will equal to zero (0) while the last term will equal to .





 

 

**MOVING AVERAGE PROCESESS**

**FIRST ORDER MOVING AVERAGE, MA (1) PROCESS**

The first order moving average, **MA (1)** process can be represented by the following equation.

 ****

Where, **=** white noise process,and  are arbitrary constants. The model process is said to be a moving average because it is weighted sum of the adjacent values (in time) of.

(i) **The Mean:**

 ****

 ****

****

 ****

(ii) 

 ****

 ****

 ****

****

****

****

 

Where, 

**(iii) Autocovariance**

The first autocovariance,is given by

****

 

 

 

 =

 

 

**SECOND AUTOCOVARIANCE, **

****

**=**

**=******

****

****

**=**

 = 0, this is because of the nature of the moving average process (our moving average is of order one (1). Hence the highest order of autocorrelation we can get. Generally, higher order autocovariances are zero I.e.

  For 

**Conclusion:**

First order moving average is stationary or weakly stationary since the mean  and autocovariances (and) are time independent, and MA (1) process is covariance stationary regardless of the.

**TIME SERIES DATA**

One of the important types of the data used in empirical analysis is time series data. Time series data is a type of data collected over time e.g weekly, monthly, quarterly, yearly etc. Such data poses several challenges to econometricians practitioners due to the following reasons:

1.Empirical based on time series data assumed that underlying time serves in stationary.

2. **C**oncerning autocorrelation arises because the underlying times series is non-stationary.

3.In regression a time series variable on another time series variable(s) one obtains a very high $ R^{2}$(In excess of 0.9) even though there is no meaningful relationship between the two variables. Sometimes we expect no relationship between two variables yet a regression of one on the other variable often shows a significant relationship. This situation exemplifies the problem of spurious or nonsense regression.

4. Some financial time series, such as stock prices, exhibit what is known as the random walk phenomenon. This means the best prediction of the price of a stock tomorrow is equal to its price today plus a purely random stock (or error term). If this were in fact the case forecasting asset prices would be futile exercise.

5. Regression models involving time series data are often used for forecasting. Thus we would like to know if such forecasting is valid if the underlying time series is not stationary.

**CHARECTARISTICS OF STATIONERY AND NON-STATIONERY SERIES**

To illustrate stationary and non stationary series we use AR (I) process and we define:

 $y\_{t}= ∝y\_{t-1}+ μ\_{t}$ . In this case we say that if the absolute value of $∝$ is less than one (1) the series is stationery and if $∝,\geq \left(1\right)$ the series ($y\_{t})$ is non stationary.

|  |  |  |
| --- | --- | --- |
| **CHARECTARISTIC** | **STATIONARY** | **NON-STATIONARY** |
| Value of $∝$ | 1$∝1<1$  | 1$∝1\geq 1$ |
| Series mean | Constant | Time-dependent varies with time |
| Variance | Finite  | Infinite.$t.δ^{2}$ |
| Innovations | Transitory, Derivation from the mean are not permanent | Permanent |
| Plot | Returns to the mean value | Varies from the mean value |
| Order of integration | I(0) | I (1) or higher order. |

When $∝ =1 $then, we have a special case of non-stationary series know as random walk, where random walk refers to a time series process where next period’s value is obtained as this period’s value, plus an independent (or at least an uncorrelated ) error term.

**THE ORDER OF INTERGRATION**

The structure of the series can be referred to in terms of their order of integration which provides the direct link between stationary and non-stationary series. A series is said to be integrated of order  if it has a stationary invertible  representation after differencing the series  times but which is not stationary after differencing only  times. In this case th series is said to have  unit roots. Such a series is denoted as ~where  is the order of integration. A series integration of order can be described as having an autoregressive integrated moving coverage representation know as. Therefore a stationery series is denoted as series. A random walk Series. If we difference the random walk once we get stationary series. Differentiating i.e . If we difference and I (0) series we get another stationary series.

**WHY NON-STATIONARY A PROBLEM**

There are two cases:

1. Spurious Regression results
2. Inconsistent Regression.

**SPURIOUS PROBLEM:-**

This arises in the case where uncorrelated series are seen to be related because of the fact that they share the common time frame for example suppose that.

Two series *X* and *Y* are known to be uncorrelated random walk.

 , 

Also,$R^{2}$ will tend to zero i.e 

This is because of the assumption no correlations between X and Y. In a non stationary series, this may not be the case. This is because of the existence of one unit root in $y and X\_{t}$ will make $β \ne 0$ and$R^{2}\ne 0$. Increasing the sample size will tend to worsen the problem. One way of detecting spurious correlation is by use of autocorrelation statistics in particular the Durbin Watson.

If there is no correlation between the series then the Durbin Watson will converge to a value 2. If $R^{2} $is greater than Durbin Watson statistics then there is high probability of spurious correlation. In the presence of that we cannot rely on $R^{2}$ as a measure of goodness of fit because $R^{2}$ is telling us two series which are uncorrelated explain the relations. Also we cannot rely on the t and F statistics in evaluating regression model one way of solving this problem is by use of co integration. This provides an important way of dealing with modeling of non-stationary series.

**2. INCONSISTENCY REGRESSION**

Here we consider a regression of a stationary series on a non- stationary series because the non-stationary series will have the time dependent mean. The value of the coefficient of the regression will not itself be constant.

Consider: **** Where ****

**

The inconsistency here is caused by the fact that the regressors and the regressand are of different orders of integration. Therefore if inference is to be valid and not time dependent then all the variables in the model must be in the same order of integration.

**TESTING THE ORDER OF INTEGRATION OF A SERIES**

Consider the following model:

 $y\_{r}= ∝ y\_{t-1}+ μ\_{r }$ ;

Testing for the order of integration is same as testing the significance of $∝$ in the equation.If $∝ <1$ $⇒ $the series is stationary. If $∝ \geq 1$ which implies the serve is non-stationary. We do this by carrying out a t-test by constructing a t-test against null hypothesis $H\_{0}: ∝ =1$

If you do not reject $H\_{0}: ∝ =1 $then the series is non stationary but we do not know the order of integration.

**Dickey-Fuller Test for Unit Roots**

Dickey and Fuller (1979, 1981) devised a procedure to formally test for non stationarity (hereafter refer to DF test). The key insight of their test is that testing for non stationarity is equivalent to testing for the existence of a unit root. Thus the obvious test is the following which is based on the following simple *AR(1)* model:

 *Yt = Yt-1 + ut 1*

What we need to examine here is = 1 (unity and hence ‘unit root’). Obviously, the null hypothesis is H0: = 1, and the alternative hypothesis is H1: < 1.

We obtain a different (more convenient) version of the test by subtracting *Yt-1* from both sides of

 ` *Yt – Yt-1 = Yt-1 – Yt-1 + ut*

*ΔYt = ( - 1)Yt-1 + ut*

*ΔYt = Yt-1 + ut* ***Random walk***  *2*

Where δ= (- 1). Then, now the null hypothesis is *H0: = 0,* and the alternative hypothesis is

*H1: < 0* (why?). In this case, if = 0, then *Yt* follows a pure random walk (and, of course, *Yt* is non stationary).

Dickey and Fuller (1979) also proposed two alternative regression equations that can be used for testing for the presence of a unit root. The first contains a constant in the random walk process as in the following equation:

 *Yt = + δYt-1 + ut* ***Random walk with a drift***3

According to Asteriou (2007), this is an extremely important case, because such processes exhibit a definite trend in the series when = 0, which is often the case for macroeconomic variables.

The second case is also allow, a non-stochastic time trend in the model, so as to have:

 *Yt = + δYt-1 + ut* ***Random walk with a drift and deterministic trend* 4**

The Dickey-Fuller test for Stationarity is the simply the normal‘t’ test on the coefficient of the lagged dependent variable *Yt-1* from one of the three models (2, 3, and 4). This test does not, however, have a conventional ‘t’ distribution and so we must use special critical values which were originally calculated by Dickey and Fuller. This is also known as the Dickey-Fuller *tau* statistic (Gujarati, 2003; 2011). However, most modern statistical packages such as Stata and Eviews routinely produce the critical values for Dickey-Fuller tests at 1%, 5%, and 10% significant levels.

MacKinnon (1991,1996) tabulated appropriate critical values for each of the three above models and these are presented in Table 1.

 **Table 1: Critical values for DF test**

|  |  |  |  |
| --- | --- | --- | --- |
| Model | 1% | 5% | 10% |
| *ΔYt = Yt-1 + ut* | -2.56 | -1.94 | -1.62 |
| *Yt = + δYt-1 + ut* | -3.43 | -2.86 | -2.57 |
| *Yt = + δYt-1 + ut* | -3.96 | -3.41 | -3.13 |
| Standard critical values | -2.33 | -1.65 | -1.28 |

 Source: Asteriou (2007)

In all cases, the test concerns whether = 0. The DF test statistic is the *t* statistic for the lagged dependent variable. If the DF statistical value is smaller in absolute terms than the critical value then we reject the null hypothesis of a unit root and conclude that *Yt* is a stationary process.

**Augmented Dickey-Fuller Test for Unit Roots**

As the error term is unlikely to be white noise process, Dickey and Fuller extended their test procedure suggesting an augmented version of the test (hereafter refer to ADF test) which includes extra lagged terms of the dependent variable in order to eliminate autocorrelation in the test equation.

The lag length of 14 on these extra terms is either determined by Akaike Information Criterion (AIC) or Schwarz Bayesian/Information Criterion (SBC), or more usefully by the lag length necessary to whiten the residuals (i.e. after each case, we check whether the residuals of the ADF regression are autocorrelated or not through *LM* tests and not the *DW* test.

The three possible forms of the ADF test are given by the following equations:

  5

 6  7

The difference between the three regressions concerns the presence of the deterministic elements α and . The critical values for the ADF test are the same as those given in Table 1 above for the DF test.

According to Asteriou (2007), unless the econometrician knows the actual data-generating process, there is a question concerning whether it is most appropriate to estimate (5), (6), or (7). Daldado, Jenkinson and Sosvilla-Rivero (1990) suggest a procedure which starts from estimation of the most general model given by (7) and then answering a set of questions regarding the appropriateness of each model and moving to the next model. It needs to be stressed here that, although useful, this procedure is not designed to be applied in a mechanical fashion. Plotting the data and observing the graph is sometimes very useful because it can clearly indicate the presence or number of deterministic regressors. However, this procedure is the most sensible way to test for unit roots when the form of the data-generating process is typically unknown

In practical studies, researchers mostly use both the ADF and the Phillips-Perron (PP) tests. This is because the distribution theory that supporting the Dickey-Fuller tests is based on the assumption of random error terms [*iid*(0,)], when using the ADF methodology we have to make sure that the error terms are uncorrelated and they really have a constant variance. Phillips and Perron (1988) developed a generalization of the ADF test procedure that allows for fairly mild assumptions concerning the distribution of errors. The regression for the PP test is similar to equation (4).

 *ΔYt = Yt-1 + et 8*

While the ADF test corrects for higher order serial correlation by adding lagged differenced terms on the right-hand side of the test equation, the PP test makes a correction to the *t* statistic of the coefficient from the AR(1) regression (semi-difference method) to account for the serial correlation in *et*.

So, the PP statistics are just modifications of the ADF *t* statistics that take into account the less restrictive nature of the error process. The expressions are extremely complex to derive and are beyond the scope of this lecture. Luckily, since most statistical packages have routines available to calculate these statistics, it is good for researcher to test the order of integration of a series performing the PP test as well. The asymptotic distribution of the PP *t* statistic is the same as the ADF *t* statistic and therefore the MacKinnon (1991, 1996) critical values are still applicable.

**CO INTEGRATION**

Co integration refers to linear combination of two series, each of which is integrated of order one,, is integrated of order zero. Co-integration represents a statistical characteristic of equilibrium relationship. Co-integration allows us to capture equilibrium relationship between non stationery series within a stationary model. So it is a method of avoiding both spurious and inconsistent regression problem which otherwise occur with regression of non-stationary data series.

Economic theory says much about equilibrium but it’s quiet in explaining short term variation from the long run relationship. In this case co integration provides a means of splitting (Portioning) evolution of time series data into two components.

1. The long-run disequilibrium characteristics.
2. Short – run disequilibrium dynamic and it does these through a direct which between co-integration and the so called or using error correction model (Error equilibrium correction model).

We can say that co-integration permits combination of long-run and short run information in the same model and therefore overcomes some of the drawbacks associated with the loss of the information which occurs from simple attempts to address non stationary through differencing.

According to Asteriou (2007), the concept of co integration was first introduced by Granger (1981) and elaborated further Engle and Granger (1987), Engle and Yoo (1987), Phillips and Ouliaris (1990), Stock and Watson (1988), Phillips (1986 and 1987), and Johansen (1988, 1991, and 1995).

It is known that trended time series can potentially create major problems in empirical econometrics due to spurious regressions. One way of resolving this is to difference the series successively until stationary is achieved and then use the stationary series for regression analysis. According to Asteriou (2007), this solution, however, is not ideal because it not only differences the error process in the regression, but also no longer gives a unique long-run solution. If two variables are non stationary, then we can represent the error as a combination of two cumulated error processes. These cumulated error processes are often called stochastic trends and normally we could expect that they would combine to produce another non-stationary process. However, in the special case that two variables, say Xt and Yt, are really related, then we would expect them to move together and so the two stochastic trends would be very similar to each other and when we combine them together it should be possible to find a combination of them which eliminates the non Stationarity. In this special case, we say that the variables are co integrated (Asteriou, 2007). Co integration becomes an overriding requirement for any economic model using non stationary time series data. If the variables do not co-integrate, we usually face the problems of spurious regression and econometric work becomes almost meaningless. On the other hand, if the stochastic trends do cancel to each other, then we have co integration.

Suppose that, if there really is a genuine long-run relationship between Yt and Xt, the although the variables will rise overtime (because they are trended), there will be a common trend that links them together. For an equilibrium, or long-run relationship to exist, what we require, then, is a linear combination of Yt and Xt that is a stationary variable [an *I*(0) variable]. A linear combination of Yt and Xt can be directly taken from estimating the following regression:

*Yt = β1 + β2Xt + ut***9**

And taking the residuals:

  **10**

If ~*I*(0), then the variables *Yt* and *Xt* are said to be co-integrated and then we estimate the error correction model as follows:

  11

Where lagged error correction term. The coefficient of the error correction term  is theoretically suppose to negative and significant to illustrate that the deviation from the long-run is corrected over time.

**Important of Error correction model (ECM):-**

1. It captures equilibrium relationship between non-stationary serve within a stationary model hence avoided spurious and inconsistent regression results.
2. Allows combinations of long-run and short-run information in the same model.
3. It minimizes the problem of multicolinearity in the model.
4. It retains information about the levels of variables and hence any long-run relationship between such variables within the model.

**Test for Co integration**

Test for co integration we need to test for the presence of the unit root in the residuals from a regression of two variables. This can be done by using the following steps:

**Step 1.** Regress  on and obtain the residuals.

**Step 2.** Create the differenced residuals, and lagged residuals if need to.

**Step 3.** Test for unit root on the residuals. If we can reject unit root in the residuals, then the two series are co integrated, otherwise no**.**

**EXAMPLES OF CO INTEGRATION IN FINANCE**

1. Spot and futures prices
2. Ratio of relative prices and an exchange rate
3. Equity prices and dividends

Market forces arising from no arbitrage conditions should ensure an equilibrium relationship.

No co integration implies that series could wander apart without bound in the long run.

**Long-run Solution:** When the concept of non-stationarity was first considered, a usual response was to independently take the first differences of a series of I (1) variables. The problem with this approach is that pure first difference models have no long run solution.

Consider two nonstationary **** variable  and , we can estimate the model:



One definition of the long-run implies that the variables have converged upon some steady-state value: . Hence, the above equation has no long-run solution.

One way to get around this problem is to use both first difference and levels terms:



This model is known as an Error Correction Model or an Equilibrium Correction Model. Provided that  and are cointegrated then the linear combination  will be I (0). This term is known as the error correction term. The model can be estimated using OLS.

**RESIDUAL-BASED TEST**

The model for the equilibrium correction term can be generalized to include more than two variables:o



 should be if the variables in this equation are co integrated

Then testing for co integration is to see whether the residuals are stationary. An *ADF* test can be used to . Note that the critical value for such test is different from a standard *ADF* test on a series of raw data.

**ENGLE-GRANGER TWO STEP METHOD**

A two-step procedure testing for Cointegration is proposed by Engle and Granger (1987), which is a single equation technique;

**Step 1:**

* Make sure that all the individual variables are I (1);
* Then estimate the cointegrating regression using OLS;
* Save the residuals of the cointegrating regression;
* Test these residuals to ensure that they are I (0).

**Step 2:** Use the step 1 residuals as one variable in the error correction model.

**Limitations of Engle-Granger two Step Method**

* Unit root and co integration tests have low power in finite Samples;
* There Can be a simultaneous equations bias if the causality between  and runs in both directions;
* Can only find one co integrating vector
* For the other problems, we may have to refer to the Johansen Approach.

All these problems are resolved with the use of the Johansen approach that will be examined later.

**The Steps of the Johansen Approach in Practice**

**Step 1:** *Testing the order of Integration of the Variables*

As with the EG approach, the first step in the Johansen approach is to test for the order of integration of the variables under examination. It was noted earlier that most economic time series are non-stationary and therefore integrated. Indeed, the issue here is to have non-stationary variables in order to detect among them stationary cointegrating relationship(s) and avoid the problem of spurious regressions. It is clear that the most desirable case is when all the variables are integrated of the same order, and then to proceed with the cointegration test. However, it is important to stress that this is not always the case, and that even in cases where a mix of *I(*0*)*, *I(*1*)* and *I(*2*)* variables are present in the model, cointegrating relationships might well exist. The inclusion of these variables, though, will massively affect researchers’ results and more consideration should be applied in such cases.

Consider, for example, the inclusion of an *I(*0*)* variable. In a multivariate framework, for every *I(*0*)* variable included in the model the number of cointegrating relationships will increase correspondingly.We stated earlier that the Johansen approach amounts to testing for the rank of (that is finding the number of linearly independent columns in ), and since each *I(*0*)* variable is stationary by itself, it forms a cointegrating relationship by itself and therefore forms a linearly independent vector in .

Matters become more complicated when we include *I(*2*)* variables. Consider, for example, a model with the inclusion of two *I(*1*)* and two *I(*2*)* variables. There is a possibility that the two *I(*2*)* variables cointegrate down to an *I(*1*)* relationship, and then this relationship may further cointegrate with one of the two *I(*1*)* variables to form another cointegrating vector. In general, situations with variables in differing orders of integration are quite complicated, though the positive thing is that it is quite common in macroeconomics to have *I(*1*)* variables. Those who are interested in further details regarding the inclusion of *I(*2*)* variables can refer to Johansen’s (1995b) paper, which develops an approach to treat *I(*2*)* models.

**Step 2:** *Setting the Appropriate Lag Length of the Model*

The issue of finding the appropriate (optimal) lag length is very important because we want to have Gaussian error terms (that is standard normal error terms that do not suffer from non-normality, autocorrelation, heteroskedasticity and so on). Setting the value of the lag length is affected by the omission of variables that might affect only the shortrun behaviour of the model. This is because omitted variables instantly become part of the error term. Therefore very careful inspection of the data and the functional relationship is necessary before proceeding with estimation, to decide whether to include additional variables. It is quite common to use dummy variables to take into account short-run ‘shocks’ to the system, such as political events that had important effects on macroeconomic conditions.

The most common procedure in choosing the optimal lag length is to estimate a VAR model including all our variables in levels (non-differenced data). This VAR model should be estimated for a large number of lags, then reducing down by re-estimating the model for one lag less until zero lags are reached (that is we estimate the model for 12 lags, then 11, then 10 and so on until we reach 0 lags).

In each of these models we inspect the values of the AIC and the SBC criteria, as well as the diagnostics concerning autocorrelation, heteroskedasticity, possible ARCH effects and normality of the residuals. In general the model that minimizes AIC and SBC is selected as the one with the optimal lag length. This model should also pass all the diagnostic checks.

**Step 3:** *choosing the appropriate model regarding the deterministic*

Components in the multivariate system Another important aspect in the formulation of the dynamic model is whether an intercept and/or a trend should enter either the short-run or the long-run model, or both models.

**Step 4:** *determining the rank of*  *******or the number of cointegrating vectors*

According to Johansen (1988) and Johansen and Juselius (1990), there are two methods (and corresponding test statistics) for determining the number of cointegrating relations, and both involve estimation of the matrix. This is a *k*×*k* matrix with rank *r*. The procedures are based on propositions about eigenvalues.

**Step 5:** *Testing for Weak Exogeneity*

After determining the number of cointegrating vectors we proceed with tests of weak exogeneity.

**Step 6:** *Testing for Linear Restrictions in the Cointegrating Vectors*

An important feature of the Johansen approach is that it allows us to obtain estimates of the coefficients of the matrices ***α*** and ***β***, and then test for possible linear restrictions regarding those matrices. Especially for matrix ***β***, the matrix that contains the longrun parameters, this is very important because it allows us to test specific hypotheses regarding various theoretical predictions from an economic theory point of view. So, for example, if we examine a money–demand relationship, we might be interested in testing restrictions regarding the long-run proportionality between money and prices, or the relative size of income and interest-rate elasticities of demand for money and so on. For more details regarding testing linear restrictions in the Johansen framework, see Enders (1995) and Harris (1997).

**Box-Jenkins Model** **selection**

A fundamental idea in the Box-jenkins approach is the principle of *parsimony.* Parsimony (meaning sparseness or stinginess) should come as second nature toeconomists and financial analysts. Incorporating additional coefficients will necessarilyincrease the fit of the regression equation (i.e. the value of the *R2* will increase),but the cost will be a reduction of the degrees of freedom. Box and jenkins arguethat parsimonious models produce better forecasts than overparametrized models.In general Box and Jenkins popularized a three-stage method aimed at selecting anappropriate (parsimonious) ARIMA model for the purpose of estimating and forecastinga univariate time series. The three stages are: (a) Identification, (b) Estimation, and

(c) Diagnostic checking and are presented below.

**Identification**

In the identification stage (this identification should not be confused with the identification procedure explained in the simultaneous equations chapter), the researcher visually examines the time plot of the series autocorrelation function, and partial correlation function. Plotting each observation of the *Yt* sequence against *t* provides useful information concerning outliers, missing values, and structural breaks in the data. Most economic and financial time series are trended and therefore non-stationary. Typically, non-stationary variables have a pronounced trend (increasing or declining) or appear to meander without a constant long-run mean or variance. Missing values and outliers can be corrected at this point. At one time, the standard practice was to first-difference any series deemed to be non-stationary.

**Estimation**

In the estimation stage, each of the tentative models is estimated and the various coefficients are examined. In this second stage, the estimated models are compared using the Akaike information criterion (AI C) and the Schwartz Bayesian criterion (SBC). We want a parsimonious model, so we will choose the model with the smallest AIC and SBC values. Of the two criteria, the SBC is preferable. Also at this stage we have to be aware of the common factor problem. The Box-Jenkins approach necessitates that the series is stationary and the model invertible.

**Diagnostic checking**

In the diagnostic checking stage we examine the goodness of fit of the model. The standard practice at this stage is to plot the residuals and look for outliers and evidence of periods in which the model does not fit the *data* well. We must be careful here to avoid overfitting (the procedure of adding another coefficient in an appropriate model). The special statistics that we use here are the Box-Pierce statistic (BP) and the Ljung Box (LB) Q-statistic which serve to test for autocorrelations- of-the-residuals.