CHAPTER - VI

Research Methodology on Time-Series Econometrics with an Application to Macroeconomic Instability

METHODOLOGICAL LINKAGE

This section describes the underlying rationale for constructing a comparative perspective of macroeconomic instability and outlines the methodology. Macroeconomic instability analysis of the effects of policies and exogenous shocks on economic growth is generally conducted within one of the three alternative frameworks: (a) a macroeconomic growth model, (b) business cycle approach; (c) a multifactor technological production function. In (a) the emphasis is on how the macroeconomic policies affect the growth parameters such as the savings – income ratio and the capital – output ratio. In (c), the main emphasis is on the scale, productivity and efficiency parameters of the production function. Thus framework (c) falls potentially in the ambit of endogenous growth theory provided foreign investment can be treated as an endogenously generated (or market determined) input factor to be used in the production function. However in (b), time-series econometrics like unit root tests, co-integration within error correction framework, and causality tests are essentially useful. The unit root methodology was developed as a means to provide a summary of economic series that exhibits a potential long-run relationship with the fluctuations of economic activity. Methodologically, this chapter explains the concepts of stationary, non-stationary and co-integration, and their relation to theoretical and methodological perspectives of instability linkages between macroeconomic variables and economic outcomes. It also discusses for unit root tests the Augmented Dickey Fuller (ADF) test and the LJung – Box (LB) test. It also has a brief discussion of co-integration, Co-integrating Regression Durbin – Watson Test (CRDW), and error correction mechanism (ECM), bi-variate Granger causality tests and their extension to multivariate co-integration framework.

UNIT ROOT TESTS

An econometric model illustrating some of the instability linkages is discussed in this chapter. Suppose, the theory suggests that there is a linear relationship between the values of some univariate time-series at time ‘t’ and (t-1) for which time-series data are available for a period of sufficient length. Then the obvious useful way to test the theory is to estimate by considering a linear equation of the form: $Y_t = \alpha + \rho Y_{t-1} + e_t$; where, e_t is the stochastic (random) error term with the classical assumptions of the Ordinary Least Square (OLS) method of zero expected value (mean), constant finite variance (homoscedastic), zero auto (serial) correlation (un or non correlated) due to zero auto covariance as are independent of time (time-invariant), absence of correlation between the error term and the regressors, and correct specification of the conditional mean function, i.e., no omitted variables and appropriate function form. Such an error term is also known as white noise error term. This stochastic process is also called stationary stochastic process when the estimation is conventionally done by the OLS method, the overall fit is tested by R^2, and the statistically significance of ‘ ρ’ is tested by using the standard t-test. The coefficient ‘ ρ’ measures the degree of persistence of deviations of Y_t from mean. The classical approach of estimation to statistical inference routinely used in the applied macroeconomic research work is based on the assumption that the data sources are stationary. That is, data-series over time grow in a fairly steady, constant manner, reflecting smoothly evolving economic forces, and in this sense, fluctuations in the time-series are taken to imply the influence of cyclical or temporary factors and eventually the series bounce back to their trend growth values, as the mean, variance and covariance are assumed tend to remain constant over time.

When $\rho = 1$, these deviations are permanent what is known as the unit root problem, i.e., a non-stationary situation. In this case, the time-series (Y_t) that has a unit root is said to follow a random walk – it can wander arbitrarily far from any given constant if enough time passes. It is important to note that the applications of the unit root tests have shown that the assumptions of classical estimation method are not satisfied by most number of macroeconomic time-series as large of them are not stationary around a deterministic trend. Instead, they have stochastic trends; i.e., some economic shocks that affect a series would have a permanent effect on the level of the series, making the series wander without a tendency to return to mean value. Thus, the mean, variance and covariance of the series will be time-variant or time-dependent, i.e., tending to vary over time. If variables are non-stationary, using classical estimation methods, such as OLS, to estimate relationships with unit root variables gives misleading inferences. In other words, in this case, the OLS estimates tend to have sample distributions with properties very different from those assumed under the conventional procedure, and regression coefficients tend to appear spuriously significant, which is known as the spurious regression problem.

An intuitive explanation of its significance is as follows. As the means and variances of the unit root variables change over time, all the computed statistics using OLS estimation in a regression model fail to converge to their true values as the sample size increases. Furthermore, conventional tests of hypotheses will then tend to be seriously biased towards rejecting the null hypothesis of no relationship between the dependent and independent variables. This is a serious problem if the null hypothesis is true. In other words, when the OLS estimation method works with non-stationary variables, there is always the possibility of 'Type-I' error; accepting the relationship as significant when, in fact, they are uncorrelated. However, measures of the overall fit of the regression, such the coefficient of determination R^2 tend to provide an artificially high explanatory power of the regression, as is called the problem of 'nonsense correlation' or 'spurious regression'[Yule 1926]. The spurious regression problem has also other implication that low D-W statistics (converges towards zero) often indicates that the variables in a regression model are non-stationary [Phillips 1986].

The existing business cycle theories having alternative macroeconomic paradigms have so far treated economic fluctuations as temporary deviations from a stable trend rate of growth of output and offered different explanations for these fluctuations in which a distinction is made between the determinants of the trend rate of growth of output and cycles. However, the assumption about the stability of the long-run trend rate of growth of output is untenable because aggregate output in many countries found non-stationary. Thus this finding casts a doubt about the usefulness of the existing theories on the merits of short-run stabilisation policies to promote the trend growth rate of output as were less noticeable such disagreement. Together unit roots and cointegration have important implications for the specification and estimations of dynamic economic approach. It is convenient to view co-integration as a technique to estimate the equilibrium or long-run parameters in a relationship among unit root variables. It is worth noting that cointegration techniques are useful for estimation not only to neo and new classical theories that use the equilibrium framework but also for Keynesian economics though based on disequilibrium framework as it can be interpreted as equilibrium economics in which resources are underemployed. In this way, the usefulness of cointegration techniques to Keynesian approach should not be underestimated.

The methodological conflict between the equilibrium framework of the theory and the disequilibrium environment from which data gathered is resolved by extending the equilibrium specifications to include disequilibrium adjustment mechanisms. The extended equation is then estimated from which estimates of the long-run or equilibrium parameters are derived by imposing equilibrium conditions, which are then used to test the underlying theory. The short-run or the dynamic disequilibrium relationships are estimated utilising the estimates of the long-run parameters of cointegration techniques within the error correction framework. In other words, cointegration facilitates utilisation of the estimated long-run parameters into the estimation of the short-run disequilibrium relationships, as there is a trade-off. Finally, the robustness of the estimated dynamic disequilibrium relationships is determined by subjecting them to the post-regression standard diagnostic test statistics. And thus the traditional approach can be criticised as it neglects the problems caused by the presence of the unit root variables while the main advantage of cointegration is that it being capable directly to test or falsify the underlying theory. In this way, it is advisable for the theoretical framework of this present research to apply essentially estimation methods of cointegrating regressions and joint estimation of both the long and short-run specifications that are computationally demanding. It has thus necessarily paid relatively more attention to the time-series underlying theoretical insights to test for unit roots, estimate cointegrating regressions and error correction specifications.

It is important to note that until the publication of Granger and Newbold (1974), the problem of ‘nonsense correlation’ or ‘spurious regression’ got paid far too little attention by applied econometricians. They by simulation showed that how misleading standard regression procedure can be if applied to non-stationary variables. According to them, high R^2 values can occur together with apparently significant t-values in regression entailing two completely unrelated series if both are integrated at order 1(a series is said to be integrated of order 1 when the differenced series is stationary with positive spectrum at zero frequency).

A stochastic process means a random process. A stochastic process may be stationary or non-stationary. It is stationary if its mean, variance and serial auto-covariance and thereby serial auto correlation remain constant over time; otherwise the time-series is said to be non-stationary or having presence of unit root or random walk. If a time-series is differenced once and the differenced series is stationary then the original series is said to be integrated of order 1, denoted by I(1). This I(1) series is called non-stationary or random walk as there is unit root. Similarly, an original series has to be differenced once or more times, say 'd'times before the differenced series becomes stationary, then the original series is said to be integrated of order 'd' or I(d) and is non-stationary and having presence of unit-root or random walk. If a series is originally stationary and does not have to be difference any time i.e., if d=0, then the original series is called stationary time series or stationary process and having presence of no unit root and is denoted by I(0). The terms a 'stationary process' and an I(0) process are as synonymous.

Moreover, they recommended that if more than two independent I (1) series are used in a regression, the possibility of a spurious relationship rises further. In upshot, Granger-Newbold results emphasised the need to rework the basic statistical techniques used in applied econometrics.

A simple alternative to cure the problem of spurious correlation dealing with non-stationary variables can be to estimate relationships using differenced series instead using them at level values [Nelson and Plosser 1982; Mills 1990]. The effect of taking the rate of change in a variable typically removes any trend. This indicates that the non-stationary time-series would become stationary if they are first differenced. However, doing that would have real danger as important components of the potential relationship between variables may remain undiscovered (Granger 1987). Thus, estimation based on differences are likely to fail yielding a long-run relationship, if exists. Differencing can thus in way be a satisfactory method tackling a spurious correlation problem.

Recent advancements in time-series econometrics revolutions came with two alternative approaches to redress the spurious regression problem without losing any information about the data at level value. These are the general to specific modelling procedure formulated by Hendry (1996) and the other one is the cointegration approach pioneered by Engle and Granger (1987). The general to specific methodology on error correction approach to time-series analysis was first developed by Sargan (1964) but was started from the early 1960s in the works of London School of Economics. The merit of this methodology is to accommodate the relationship being investigated within a reasonable complex dynamic specification, including lagged dependent and independent variables so that a parsimonious specification of the model can be captured [Hendry, Pagan and Sargan 1985]. An important advantage of this method is that it yields equations with first-differenced and stationary dependent variables unlike simple first-differenced equations and makes use of long-run information in the data appropriately [Wickens and Breucsh 1988].

This methodology starts with an estimation procedure of over-parameterized autoregressive distributive lag (ADL) order:

$$Y_t = \alpha + \sum_{i=1}^{m} A_i Y_{t-i} + \sum_{i=0}^{m} B_i X_{t-i} + \mu_t \quad (6.1)$$

Where α is a vector of constants, Y_i is a (nx1) vector of endogenous variables, X_i is a (kx1) vector of explanatory variables, and A_i and B_i are (nxn) and (nxk) matrices of parameters.

The above equation (7.1) can be re-parameterised in terms of differences and lagged levels in order to separate the short-run and long-run multipliers of the system as follows:

$$\Delta Y_t = \alpha + \sum_{i=1}^{m-1} A_i^2 \Delta Y_{t-i} + \sum_{i=0}^{m-1} B_i^2 \Delta X_{t-i} + C_0 Y_{t-m} + C_1 X_{t-m} + \mu_t \quad (6.2)$$

Where

$C_0 = -I - [\sum_{i=1}^{m} A_i]$ and $C_1 = -I - [\sum_{i=1}^{m} B_i]$

Where the long-run multipliers of the system are given by $C_0^{-1} C_1$ and I is the identity matrix. Equation 6.2 is known as the error correction mechanism (ECM). This is the particular formulation that is generally used as the 'maintained hypotheses' of the specification search. However, there are many different ways by which the dynamics of the general model could be presented as all of them would result in the same estimates of the unknown parameters, but each of them carries information differently that consequently make easier to interpret and understand (Pagan 1987). But there are number of reasons as is why the ECM is generally much more preferred over the other formulations: first, the ECM has a first-differenced dependent variable and not a typical trending variable; second, it provides the model is correctly specified and the disequilibrium error of the specification would also be stationary, which can therefore be estimated by standard classical estimation methods such as OLS; third, an ECM involves parameterisation that clearly distinguishes between long- and short-run effects and this separation between long-and short-run parameters in an ECM makes possible to assess for either the validity of the long-run implications of theory or of incorporating them into the estimation process.

The estimation procedure having first estimated the unrestricted equation (6.2) usually goes progressively to simplifying it by restricting statistically insignificant coefficients to zero and reformulating the lag patterns where appropriate in terms of levels and differences to achieve orthogonality. As part of the specification search, it is essential to check rigorously at every stage as to whether the more general of models for possible misspecifications. Such checks would provide guide for both to have visual inspection of the residual from the fitted version of the model and tests for serial correlation, heteroskedasticity and normality in the residual, and the appropriateness of the particular functional form used. More importantly, any suggestion of the autocorrelation in the residual should guide to rethink about the form of the general model. Above all theoretical consistency must be born in this methodology used throughout the testing down procedure.

COINTEGRATION TECNIQUES

Stationary of a time-series is very essential as the use of a non-stationary time-series in regression analysis can yield dubious and spurious relationships. Data-series (Y) can be either : (i) trend stationary: if $Y^* = a + bt$, then $e = Y - Y^*$ is stationary; (ii) differenced stationary: ΔY_t is stationary and (iii) the series may have to be de-trended and differenced in order to obtain a stationary series [Enders, 1995].

A time-series y_t is said to be integrated of order 1 or I(1) if Δy_t is a stationary time-series. A stationary time-series is said to be I(0). The concept of cointegration on the spurious regression problem was developed by Engle and Granger (1987) as an extension to the earlier work by Granger and Newbold (1974). They observed that despite the individual economic series may be non-stationary, a vector of variables, taken together, may be cointegrated as generating a stationary residual subject to these variables are individually integrated of the same order. If the particular vector of the variables is placed together on the basis of sound economic theory, then a cointegration relationship among them can be explored as the long-run (equilibrium) relationship explained by the theory. A prerequisite for testing the sets of variables for cointegration is to establish the properties of the individual series and the order of integration of each variable must be determined because series with different orders of integration cannot be cointegrated.

Based on the conceptual analogy between cointegration and the economic concept of equilibrium, Engle and Granger propose a consecutive two-stage approach to modeling economic relationships entailing non-stationary variables, subject to the variables are integrated of the same order. The first stage captures modeling the long-run or cointegrating the relationship. The short-run dynamic disequilibrium relationship among the variables is estimated at the second stage. The second stage disequilibrium relationship can best be represented with an error correction mechanism. Engle and Granger (1987) prove that if the variables of same order are cointegrated i.e., if there exists an equilibrium relationship, then the short-run disequilibrium relationship between them can be captured by the ECM. This mechanism is referred to as the Granger representation theorem.

The order of integration of the series is determined by using the Dickey – Fuller (DF) and the Augmented Dickey-Fuller (ADF) tests for a Unit Root. The test statistic is the 't-ratio' for the lagged level of the variable with the null being the series having unit root. A data-series is said to be integrated of order'd', if it requires to be differenced d times to become a stationary series. Suppose, there is a vector (a set of variables) X_t containing n variables, all of which is I (1). This set of variables are said to be cointegrated if there exists a linear combination of these variables $Z_t = \alpha X_t$ such that Z_t is I (0), where α is known as the cointegrating vector. Therefore, in a long-run relationship between two variables both of them must be cointegrated of the same order if the error term is to be I(0) [Deadman and Charemza, 1992]. The following tests such as Cointegration Regression Durbin-Watson Test (CRDW), Dickey-Fuller Test, and Ljung-Box Test are used to test whether a set of variables is cointegrated. The null hypothesis is both the variables are not cointegrated. Prior to using the three tests for cointegration, cointegration regression is run, if the errors are found stationary i.e., I(0), then the two series are said to be cointegrated.

The econometric investigation of this research draws upon recent advances in time-series econometrics in regard to minimising the risk of uncovering spurious relations. It aims to redress the spurious regression problem through continuous interplay of theory and data in the process of specification search. This analysis has helped to suggest which variables should enter into a relationship, and the data are left to determine whether this relationship is static or dynamic. Using multiple cointegration and causal relationship techniques, the study has examined the short-run and long-run relationships between real and nominal macroeconomic variables. Prior to causality testing, the study employed the cointegration technique to determine whether the set of variables are cointegrated or not. The Dickey-Fuller (DF) and the Augmented Dickey-Fuller (ADF) tests for the presence of unit-root properties used wherever appropriate. Application of econometric TSA techniques helps in the identification of stable long-run and disequilibrium short-run relationships between the key macroeconomic variables under study. The interdependence of series behaviour on the policy framework establishes the degree of endogeneity or exogeneity of macroeconomic behaviour. Before analysing the relevant economic relationships, this study has carried out exogeneity tests to spotting complex simultaneous links, otherwise, estimates would have been unreliable. Make use of such tests on a priori basis rules out the possibility of the bias and gives more consistent results and helps in detecting the direction of causalities and ensures the smooth running of Recursive Systems whereby findings would remain robust.

CONCLUSION

The methodological conflict between the equilibrium framework of the theory and the disequilibrium environment from which data gathered is resolved by extending the equilibrium specifications including disequilibrium adjustment mechanisms. The extended equations are then estimated from which estimates of the long-run or equilibrium parameters are derived by restricting equilibrium conditions. These equilibrium parameters are then used to verify the underlying theory. Before analysing short-run macro dynamics and long-run equilibrium relationships devising error correction models and cointegration, this research has on a priory basis examined the nature of data particularly when variables are non-stationary for the presence of unit root with a view to minimising the risk of uncovering spurious relations and carried out endogeniety tests to rule out the possibility of estimates to be biased, unreliable and simultaneity; detecting the direction of causality, causality tests were run to quantify more consistent instability results . Together unit roots, co-integration and causality tests have important implications for the specification and estimations of dynamic macroeconomic study.

By extending the ideas of unit root tests, ECM, co-integration and causality the next two chapters would establish short-run disequilibrium instability and long-run stability channels for the Indian economy.